Report

The Use of Predictive Analytics in the Canadian Property and Casualty Insurance Industry

Authors:

Denise Cheung, FCIA, PricewaterhouseCoopers LLP

Megan Kang, ACIA, PricewaterhouseCoopers LLP

Adam Goldfarb, PricewaterhouseCoopers LLP

Document 222067

Use of this report

The Canadian Institute of Actuaries (CIA) funded this report with the intention of advancing the development of predictive analytics in the property and casualty insurance (P&C) industry. It seeks to spur discussion of novel applications of predictive analytics in the insurance space and draw the attention of actuaries in other practice areas to get inspired by some of the use cases in P&C.

PricewaterhouseCoopers LLP was engaged by the CIA to perform the survey and prepare the report. PricewaterhouseCoopers LLP does not assume any duty, obligation, responsibility, or liability to users of this report.

No part of this report may be used or presented without reference to it.

Table of contents

Use of this report 2

Executive summary 6

Suggested strategic focus areas by company size 8

Suggested strategic focus areas by type of insurer 8

Introduction 10

Report methodology 12

Current maturity of advanced analytics applications 13

Management and internal support 13

Data 25

Technology 32

Model and business usage 37

Implementation 45

Considerations and opportunities in the age of analytics 49

Management and internal support 51

COVID-19 changes 51

Rate environment and market flexibility 53

Changing talent profile 56

InsurTech and mergers 57

Data 59

Utilizing unstructured data 60

Size of the data and the Internet of Things 61

Data access, governance, and privacy threats 62

Technology 63

Hardware changes 63

Embracing open-source 64

Blockchain 65

Model and business usage 67

Modeling techniques 68

Responsible AI 71

Explainable AI 73

Segmentation and knowing your customer 75

Price optimization and customer lifetime value 76

Reserving 77

Geospatial analysis and climate change 79

Fraud 81

Implementation 83

Application Programming Interfaces 83

Rating algorithm 84

Automation 85

Customer-centric 86

On-demand microinsurance 87

Summary remarks 88

Actionable approach for advanced analytics application 90

Management and internal support 91

Data 93

Technology 95

Modeling 97

Implementation 99

Conclusion 102

Appendices 103

Appendix 1 – Survey result summary 103

Appendix 2 – List of survey questions 151

Appendix 3 – Survey glossary 180

Appendix 4 – Works cited 183

Executive summary

The increasing availability of big data and the use of advanced analytics is changing how insurers have traditionally operated. Canadian property and casualty (P&C) insurers are exploring and implementing new ways of utilizing these resources in everyday practice to optimize results, and there are many new approaches and opportunities that can be adopted from other industries. The purpose of this research study is to benchmark the current utilization of predictive analytics, machine learning, and artificial intelligence (AI) in the Canadian P&C insurance industry, to examine potential areas for improvement and to accelerate the industry's use and development of advanced analytics.

This research was prepared by surveying a total of 61 Canadian P&C insurance companies. Responses were received from 24 companies, including eight out of the top 10 companies by market share based on 2020 gross written premium (GWP). Of the 24 respondents, 10 companies were interviewed and asked to expand upon their responses.

The survey evaluated the current sophistication of advanced analytics applications for each insurer. Modeling use cases were selected based on experience, research, and input from an oversight committee. In order to have a comprehensive understanding of each company's strategy and preparedness for advanced analytics, the support, resources, and limitations were also evaluated, leading to the following evaluation facets: management support, data practices, modeling process and use cases, implementation, and general questions. Key highlights of our survey results from these evaluation facets are as follows:

- Management was aware of and encouraged the use of advanced analytics; however, there were significant challenges in converting that enthusiasm into hiring and retaining talent.
- Significant investment is being made into modernizing data storage and upgrading hardware to better enable the use of advanced analytics and future-proof data infrastructure, with many larger insurers having either transitioned or begun the process of transitioning their systems.
- Respondents overwhelmingly identified pricing as the business area that has yielded the most significant gain from predictive analytics, followed by claims.
- There were few techniques used outside of generalized linear models (GLMs) and tree-based machine learning (ML) methods, although academic literature is increasingly testing neural networks in both the traditional areas of pricing and reserving, as well as in less traditional uses such as image classification.
- Fraud had the highest number of models used outside of the usual traditional and GLM methods, among the more essential analytics use cases.
- Smaller insurers took longer to implement changes to their most business-critical models compared to their larger counterparts, with IT infrastructure identified as a source of difficulty across the industry.

Research was conducted on the application of advanced analytics applications in other financial industries and within the insurance industry from other countries to better assess the maturity of Canadian P&C

insurer practices, organized utilizing the familiar survey evaluation facets: management support, data practices, modeling process and use cases, and implementation.

On the company management level, innovative product design and enhancing customer centricity have assisted insurers in acquiring and retaining market share. Microinsurance, from insuring Uber drivers during their trips to insuring personal possessions such as a camera, is part of a changing shared economy paradigm. Customers are also demanding better experiences, whether that be in the form of chatbots or automated claims triaging to speed up payment and minimize human intervention. In order to deal with these new developments, companies are looking to InsurTechs for help to jumpstart these processes.

Data form the foundation of advanced analytics initiatives. In addition to structured data commonly collected by insurers, utilizing unstructured data could reveal more insights into the insurer's risks through techniques such as text mining and image recognition. Big data availability and the Internet of Things (IoT) enable insurers to apply complex model algorithms; however, increasing data diversity and size will pose an additional burden to the insurer's hardware, software and IT infrastructures, which are the engines of advanced analytics projects. In the meantime, maintaining better data governance policies and procedures is more important than ever. Additional investment may be required to upgrade to big data solutions and cloud-based modeling platforms with open-source software to utilize the most up-to-date packages for modeling and data wrangling.

With the increasing computing power and data availability, insurers are exploring more complex modeling algorithms such as ML and neural networks to capture non-linear effects which may not be reflected in GLM models. Insurers also need to balance predictivity and explainability when utilizing these "black box" types of algorithms, especially for applications in pricing for regulated lines of business. Ethical considerations are another important aspect when testing the model. The main body of the report provides additional discussion of further considerations and opportunities for common business use cases of advanced analytics: segmentation and knowing your customer, price optimization and customer lifetime value, reserving, geospatial analysis and climate change, and fraud.

Finally, implementation is the last piece of the puzzle, and it is a pain point for most insurers due to IT capacity, infrastructure, and model complexity. Leveraging new techniques such as Application Programming Interfaces (APIs) and real-time data processing could impact the implementation process, but it is important to review the insurer's infrastructure from data collection to final model implementation as a whole. Other possible improvements are discussed in further detail in this report, including pricing, claims, products, and customer interaction perspectives.

Based on our survey findings and research, areas of focus for insurers organized by size and type are summarized in the table below.

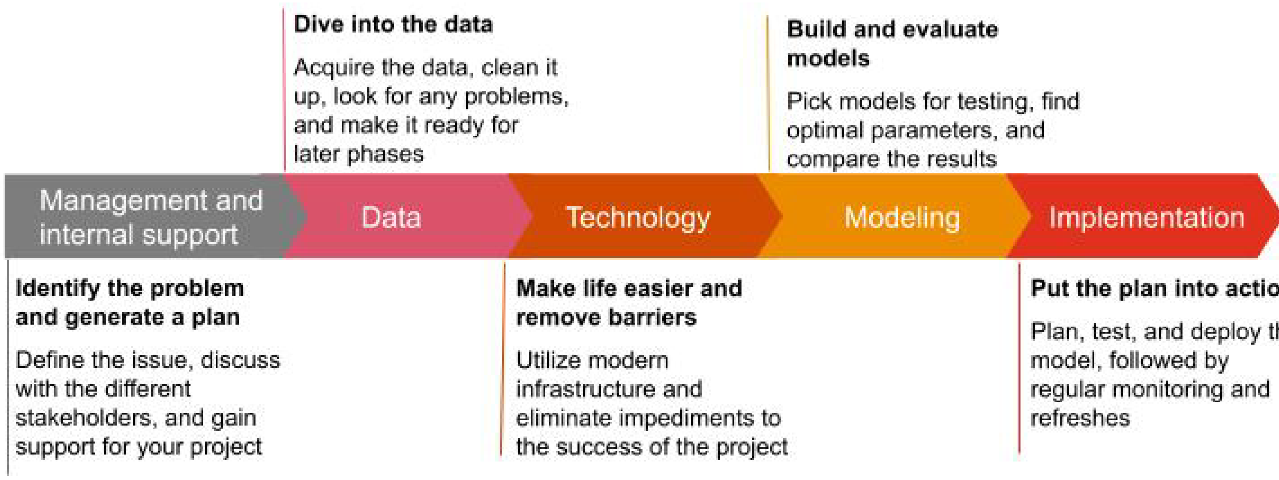
Suggested strategic focus areas by company size

Small	• Expand your analytics efforts by slowly growing a team. Automate simple tasks, and identify and address quick wins to build confidence in your teams before tackling longer-term targets. • Update the infrastructure to better manage new sources of data. Instead of a collection of linked files, create a pipeline connected to your source of truth to facilitate the diagnosis of errors. • Look to external data sources in order to supplement the existing information. Having more information is a key enabler for better decision-making.
Large	• Upskill current employees and leverage your existing talent to meet the analytics demands and to address labour force talent shortages. • Ensure the quality of the data and documentation. Being nimble is important; however, mistakes are amplified with larger books of business. • Build an API ecosystem. Doing so will help to achieve increased business velocity by making it easier to manage the many services and models.

Suggested strategic focus areas by type of insurer

Personal	• Broaden your use of advanced analytics and use responsible AI to manage black-box risks. Adopting explainable AI can help make models less opaque and help build trust. • Make your experiences more customer-centric by providing omni-channel experiences and by developing your automation tools to create personalized and seamless user journeys.
Commercial	• Look for strategic partnerships such as finding new data sources to supplement existing data. • Place additional focus on using analytics to drive fraud detection efforts and to help strengthen segmentation capabilities in better understanding your risks through objective metrics.
Reinsurer	• Look to new technologies, such as blockchain, to help decrease expense ratios, streamline data processing, and increase transparency in the process. • Keep abreast of the recent changes to catastrophe-modeling software and continue to monitor new participants that are seeking to unseat the current market leaders.

Together with our survey results and research findings, an actionable approach to implementing advanced analytics applications is illustrated within the report, where a generic method is laid out to go from an idea to post-implementation. Examples are provided to give context to the different facets of building management support, data practices, acquiring technology, modeling, and implementation.



Overall, the industry continues to invest heavily in advanced analytics in order to keep pace with other sectors. Canadian P&C insurers are generally on par with their peers abroad who are focusing their efforts on different parts of the business, with Asia concentrating on marketing, the United States on customer experience, Australia on catastrophe modeling, and Europe on optimization. Canadian P&C insurers are still utilizing GLMs for most use cases, although significant investment is being made to enable the application of new technologies and to staff analytics teams with people having the knowledge to make use of modern infrastructure. In more business-critical parts of the insurance process, such as pricing, reserving, and fraud detection, significant experimentation is being done to get a better understanding of customers and events, with a host of new methods finding use, from tree-based methods to language processing. Despite their gaining prominence, these newer models remain a challenge to explain, and require additional subject knowledge to understand their results.

A combination of new methods, an increased focus on using technology, and an additional spotlight on customer experience will be key growth drivers in the evolving Canadian P&C insurance market.

Introduction

With the rise of big data and the rapid development of computing power, industries are increasingly looking to advanced analytics to enable greater use of their data and to support business decision-making. Advanced analytics is finding greater prominence in the property and casualty (P&C) insurance industry as well. New approaches to modeling data are constantly being developed, and the skills necessary to deal with these techniques and technologies are changing. Companies have also adjusted their infrastructure, training, distribution, and talent acquisition to deal with stay-at-home orders. To tackle these developments, the Canadian P&C insurance industry is undergoing significant changes in modernizing the way it functions. Significant investment is being placed into upgrading infrastructure to ensure insurers are able to take advantage of the advances of predictive analytics, machine learning (ML), and artificial intelligence (AI). To this end, the Canadian Institute of Actuaries (CIA) identified the use and maturity of advanced analytics technology as an important object of study.

The goal of the study was to investigate and benchmark the current practice of predictive analytics, ML, and AI within the Canadian P&C insurance industry. This research was also performed to examine potential areas of improvement with regard to the application of these new methods based on P&C insurance practices abroad and academic literature, as well as on the investigation of other sectors. In order to gauge the Canadian P&C insurance industry's preparedness for advanced analytics, an assessment of management's support as well as the data practices of the various P&C insurers was necessary.

In this study, "the Canadian P&C insurance industry" is meant to include all products sold by the Canadian P&C insurance industry, whether they be providing personal, commercial, or reinsurance products. The term "insurers" will include personal insurers, commercial insurers, and reinsurance companies. Life and health products are excluded from this study.

In this study, the term "predictive analytics" refers to the practice of using statistical methods for predicting outcomes, whereas "machine learning" (ML) refers to a host of techniques, wherein the machine "learns" based on some iterative method. The term "advanced analytics" spans all of predictive analytics, ML, and AI.

This research was prepared by surveying a total of 61 Canadian P&C insurance companies, of which 24 companies responded, including eight out of the top 10 companies by market share based on 2020 GWP. The breakdown of the distribution by respondent type, origin, and GWP size is displayed below, respectively. Insurance companies with over $1 billion GWP were considered as large insurers.

Figure 1 – Distribution of Respondents by Type

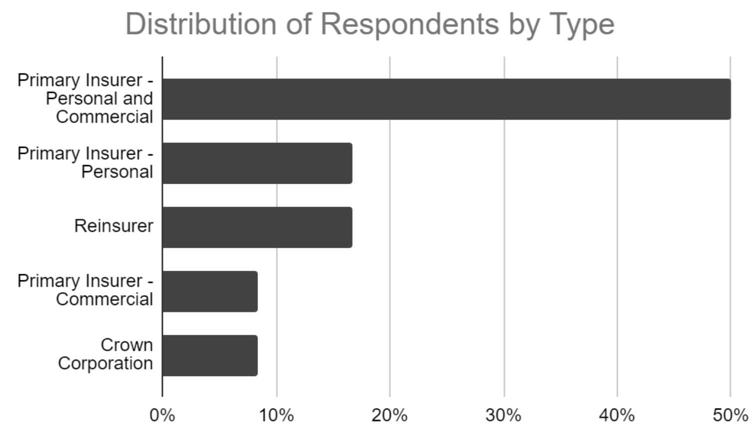


Figure 2 – Distribution of Respondents by Origin

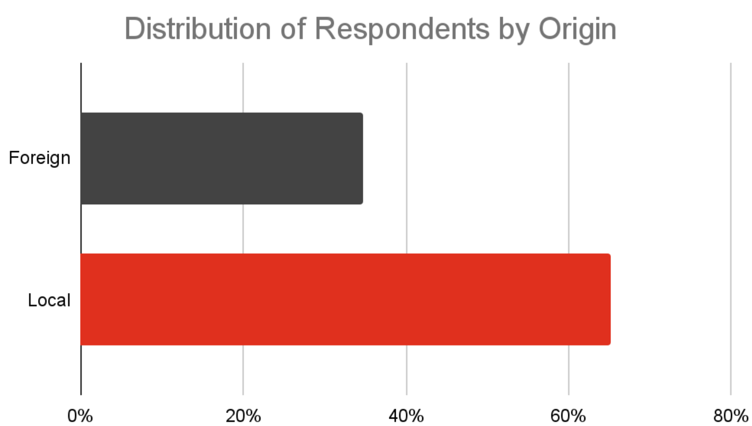
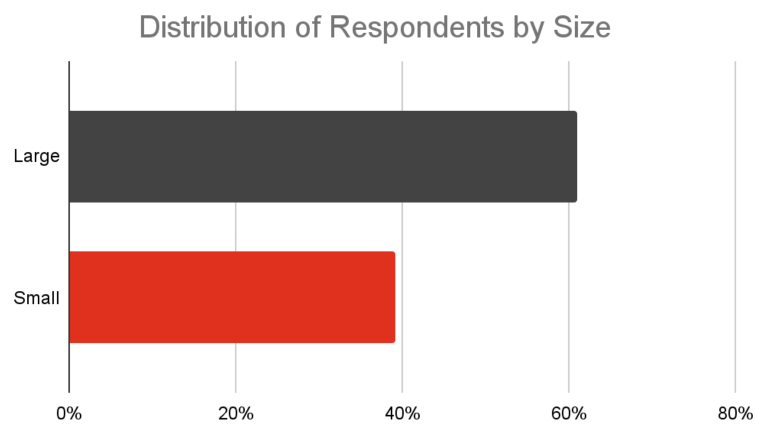


Figure 3 – Distribution of Respondents by Size



Of the 24 companies that responded, a further 10 companies were interviewed. Within these interviews, companies were asked to expand upon their responses, and preliminary discussions were had regarding their strengths and weaknesses relative to the industry. In the event of companies modifying their answer during interviews, the survey responses were altered accordingly, and the final benchmarking analysis was performed following these modifications.

Report methodology

The form of the survey was largely structured based on the modeling lifecycle and the practical data science framework from the Institute and Faculty of Actuaries (IFoA). A full outline of the method we anticipate insurers to use for the development of advanced analytics business cases is outlined in the *Actionable approach for advanced analytics application* section. Questions were constructed in accordance with Qualtrics' suggestions as well as in accordance with input and consultation from the CIA-appointed project oversight group.

In order to have consistency and to increase the response rate, the majority of questions had fixed levels, although free-form responses were allowed in many questions. Furthermore, in order to segment further for potential trends, insurers were split by their origin and size.

Research into the practices of P&C insurance companies abroad was performed. An investigation into the academic literature and actuarial journals was carried out, as was a review of industry papers from the various actuarial societies and independent advisory bodies. Finally, several discussions were had with actuaries abroad to hear the personal experiences of those across the United States, Europe, and Asia.

The report is structured into the three sections listed below:

- Survey findings outlining the present state of the industry
- Considerations and opportunities for the industry at large
- Actionable approach to implementing advanced analytics applications

Appendices at the end contain the summary graphs of the survey results, original survey, survey glossary, and works cited.

Current maturity of advanced analytics applications

One of the main purposes of this research is to survey the current state of advanced analytics usage in the Canadian P&C insurance industry. To this end, the discussion of the industry is broken into five categories, which are outlined below:

- Management and internal support – establishing the support from management and a general analytics culture
- Data – creating well-documented, accurate data, with strong oversight and privacy considerations
- Technology – utilizing newer infrastructure and increased processing power to remove impediments to growth and development
- Model and business usage – using novel techniques and applying them throughout the business
- Implementation – speeding up deployment of models to drive results

The full results of the survey have been summarized and are included in Appendix 1. Below, we discuss specific questions that were most critical to a company's analytics operations and also had a credible response rate. In addition, we further discuss survey results when obvious differences by company country of origin or company size arise. These segmentation analyses are not performed for every question due to data availability and the need to protect a company's confidentiality.

Management and internal support

Active support from management regarding the need for advanced analytics as well as its understanding of the latter's importance were necessary components of the survey in order to understand the enthusiasm for analytics throughout organizations. To gauge the support from management and senior leadership on their commitment to the use of advanced analytics, we asked about the senior leaderships' awareness, support of advanced analytics initiatives, and the presence of advanced analytics on the executive agenda. Roughly 70% of respondents stated a score of four or above, out of five, for management support (Figure 4). This indicates a large focus from management to take strong action and invest in resources to further P&C insurers' analytics capabilities. As shown in Figures 5 and 6, Canadian-based companies and larger insurers have more senior leadership support for developing advanced analytics, compared to foreign branches and smaller insurers, respectively.

Figure 4 – Overall

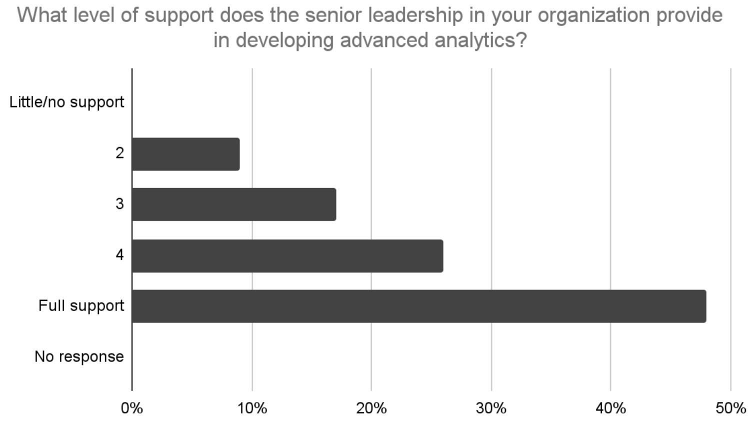


Figure 5 – By Origin

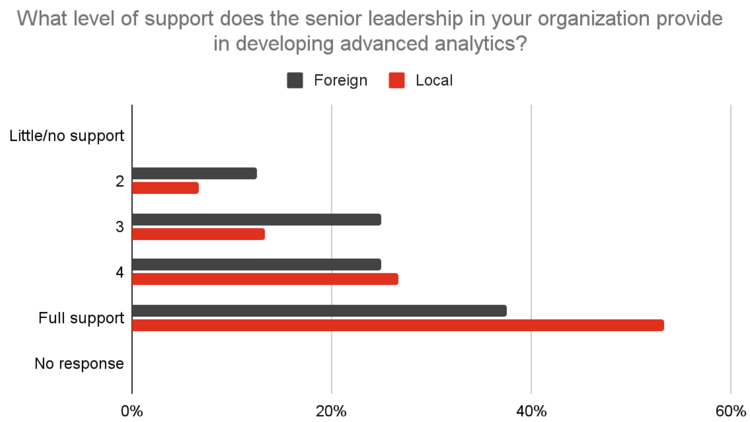
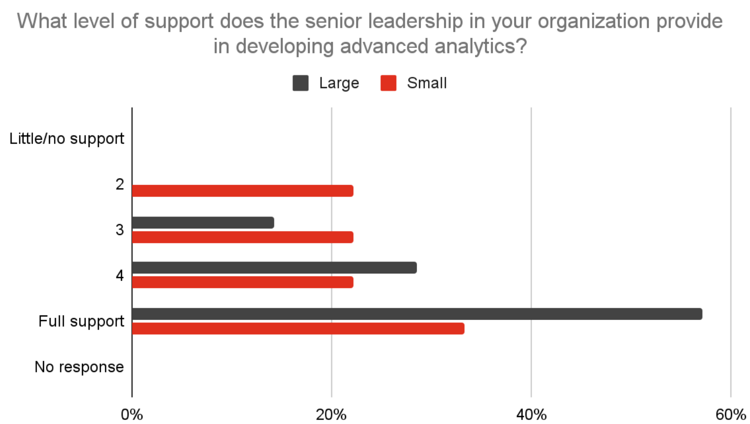


Figure 6 – By Size



One challenge insurers are having is translating the enthusiasm for analytics into action. Part of the reason for this is due to difficulties in hiring, as 60% of respondents stated that acquiring and retaining analytics talent is either extremely or somewhat difficult (Figure 7). This is especially true for smaller insurers (Figure 8). This challenge in hiring might account for the majority of insurers' centralization of analytics talent, with few insurers having advanced analytics talent organized at the business level let alone the product level (Figure 9); however, it might alternatively signal a management shift towards centres of excellence. When compared to smaller insurers, large insurers are organizing their analytics capabilities into more centralized structures (Figure 11) and local insurers are more centralized than foreign branches (Figure 10).

Figure 7 – Overall

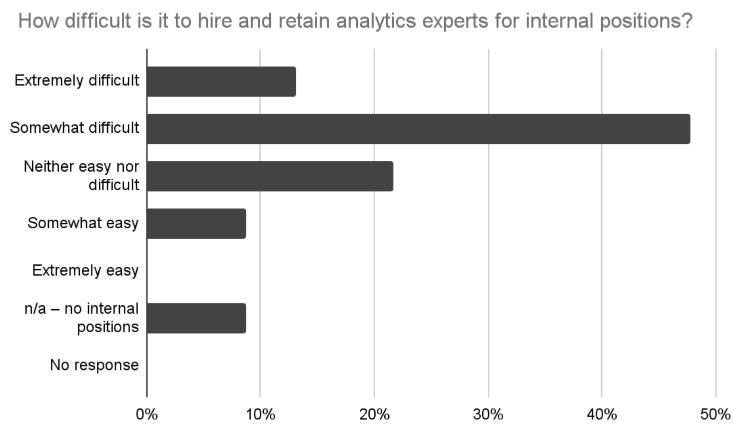
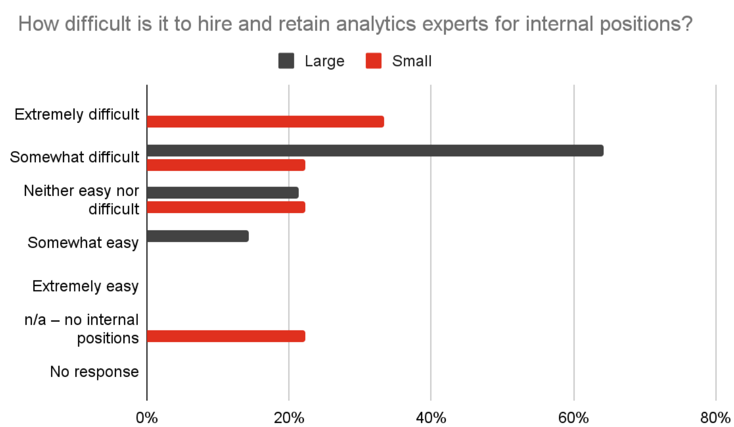


Figure 8 – By Size



Despite these hiring challenges, when analyzing the competencies of current analytics talent, respondents overwhelmingly indicated a strong understanding of technical, statistical, and business knowledge in rating such knowledge as either Medium-High or High between 70% and 90% of the time (Figure 12). Of these, statistical knowledge was rated the highest, whereas business skills tended to lag. This again might be attributed to hiring practices, as insurers hire to attract analytics talent with the intention to educate them about insurance domain knowledge. In addition, we noticed that larger insurers rated their analytics talent's business knowledge lower than smaller insurers did. This may be a result of the more centralized analytics structure for large insurers (Figure 13), which decreases the importance of business expertise for analytics talent.

Figure 9 – Overall

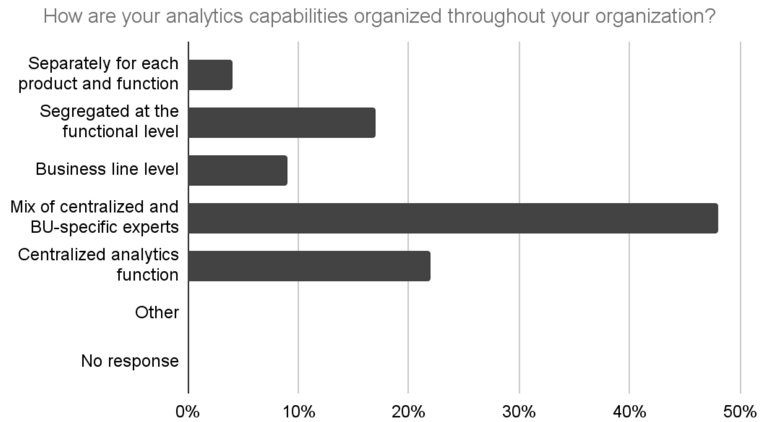


Figure 10 – By Origin

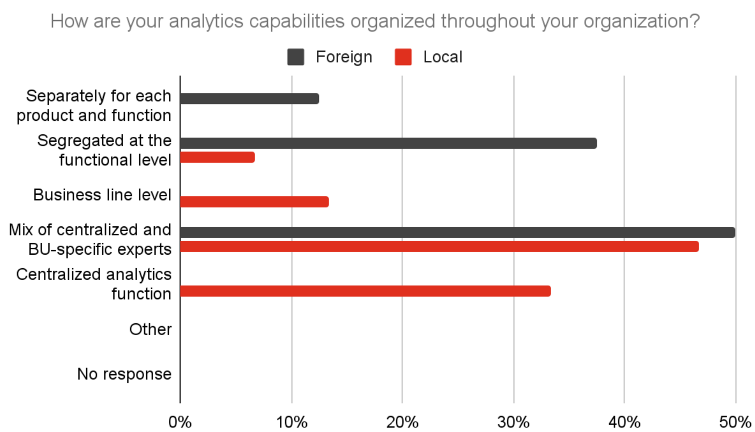


Figure 11 – By Size

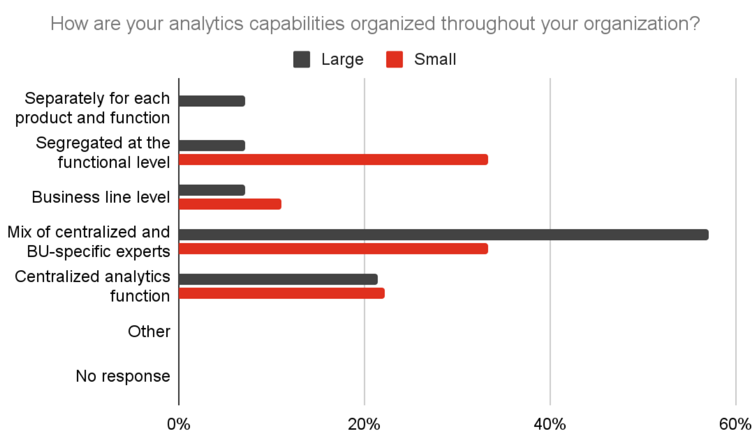


Figure 12 – Overall

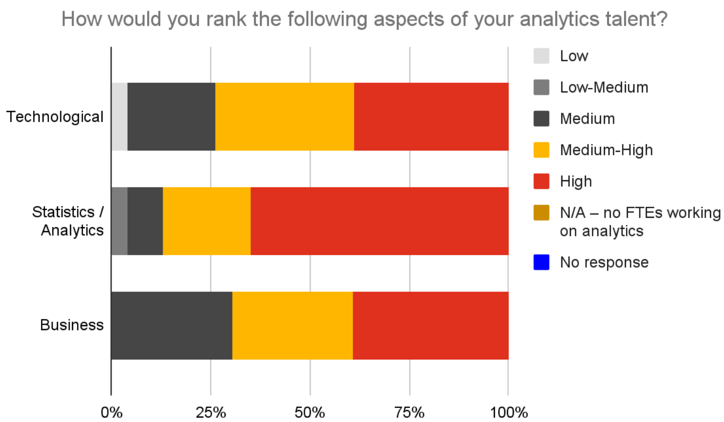
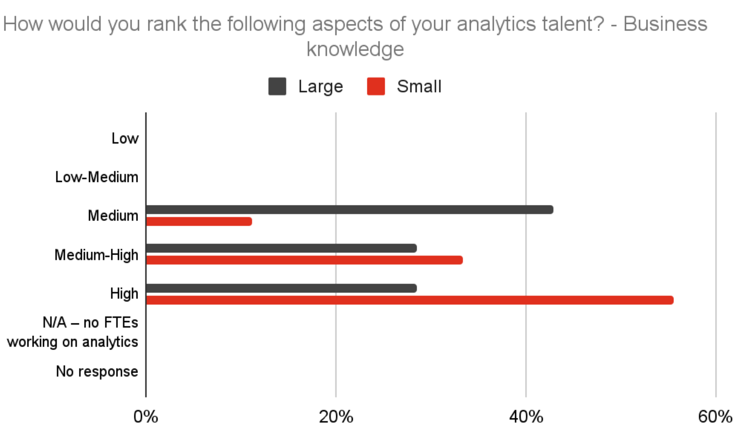


Figure 13 – By Size



The number of Full Time Equivalents (FTEs) allocated to advanced analytics falls into a wide range across the respondents. About 40% of the respondents indicated that they employed over 50 FTEs, while just under 50% indicated fewer than 24 FTEs (Figure 14). As expected, larger insurers are capable of allocating more FTEs for advanced analytics (Figure 15). Foreign insurers allocated fewer FTEs to advanced analytics compared to local insurers (Figure 16); however, during interviews, foreign insurers stated that analytics projects may involve global teams, with fewer FTEs dedicated to conducting advanced analytics work at the Canadian branch.

Figure 14 – Overall

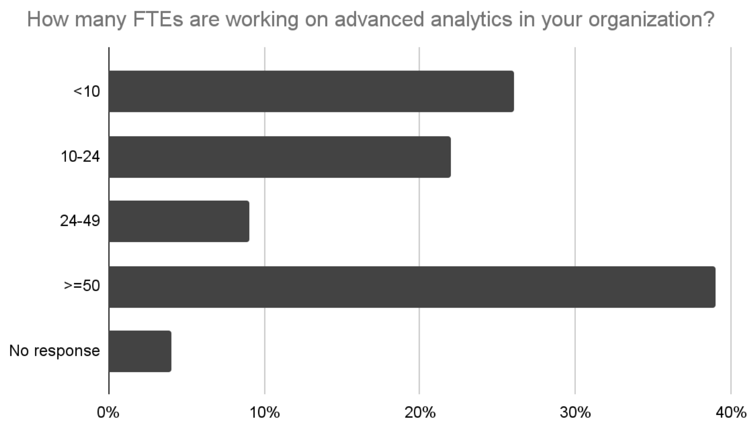


Figure 15 – By Origin

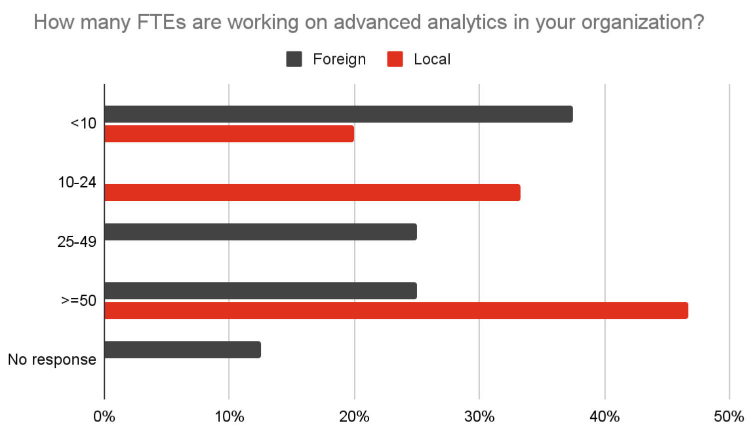
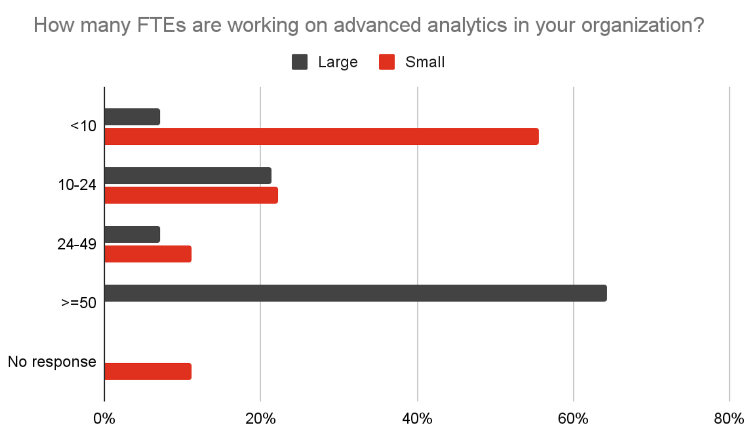


Figure 16 – By Size



The composition of the talent conducting advanced analytics varied based on the survey respondents' interpretation and organizational structure. Some respondents defined the whole data operations and analytics team as contained within the advanced analytics umbrella, whereas others were far narrower in scope, limiting their definition to only include those employees who performed advanced modeling techniques (being primarily data scientists and data engineers). In our interviews, we also asked how a data scientist differs from a statistician, and received a range of answers. Some insurers indicated employees had to be in an internal program to qualify to be a data scientist, whereas others required differing amounts of experience. Statisticians were considered distinct and more focused on statistical inference than on predictive analytics. We have surveyed the types of FTEs working on advanced analytics (Figure 17), and asked respondents to distinguish how the FTEs are being deployed in maintaining databases, building advanced analytics models, and running the advanced analytics models (Figures 20, 23, and 26).

We have observed that in all three dimensions, larger insurers usually build larger advanced analytics teams with a broader mix of talent (Figures 22, 25, and 28).

Figure 17 – Overall

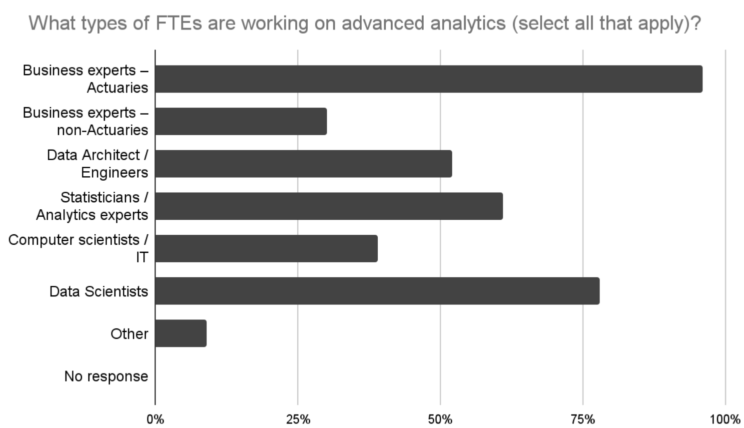


Figure 18 – By Origin

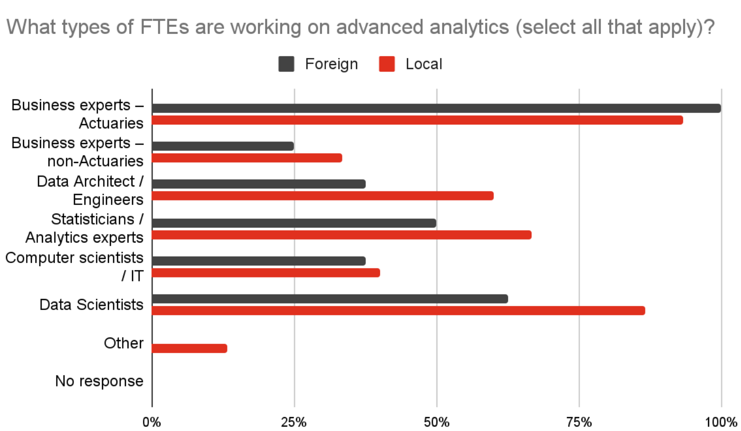


Figure 19 – By Size

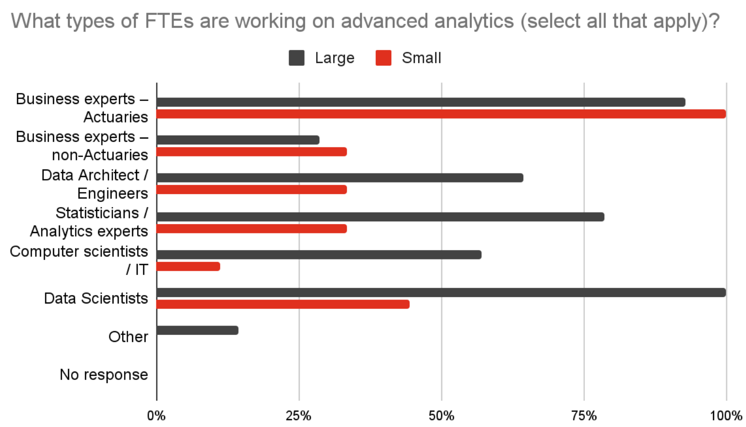


Figure 20 – Overall

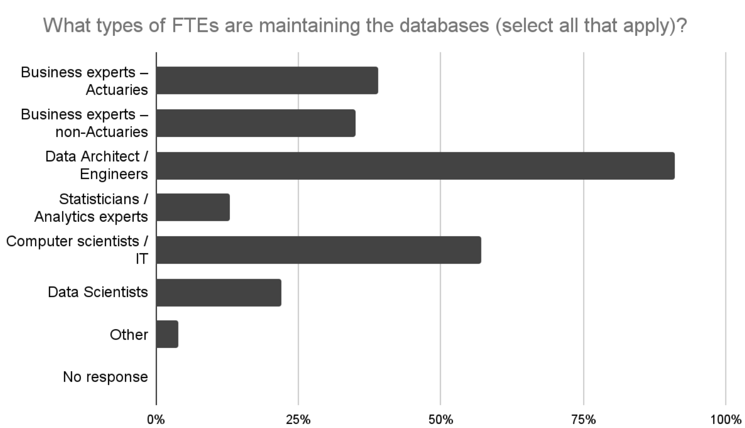


Figure 21 – By Origin

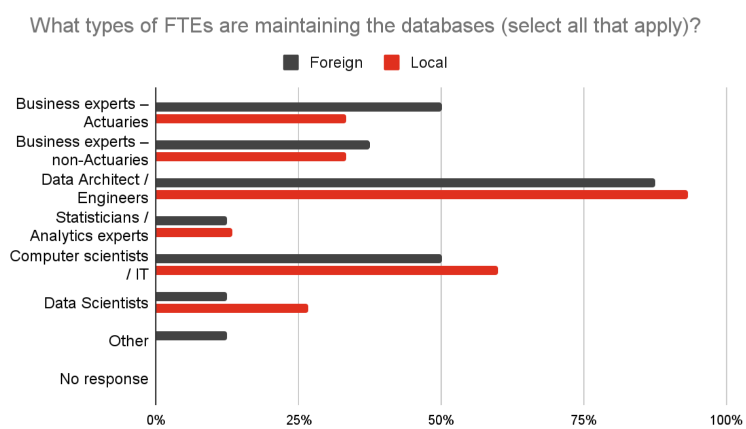


Figure 22 – By Size

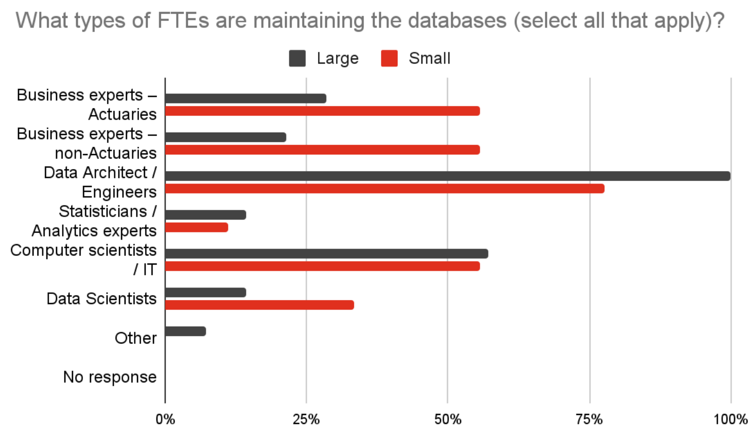


Figure 23 – Overall

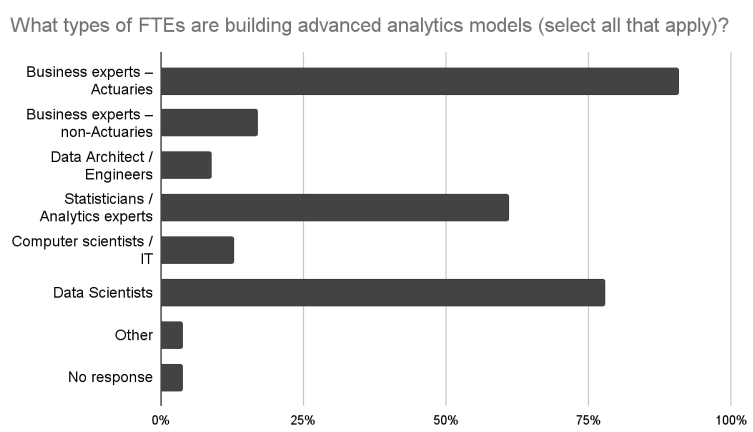


Figure 24 – By Origin

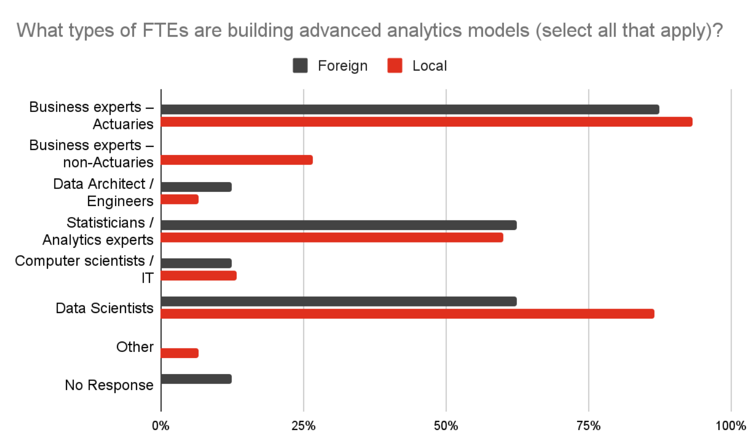


Figure 25 – By Size

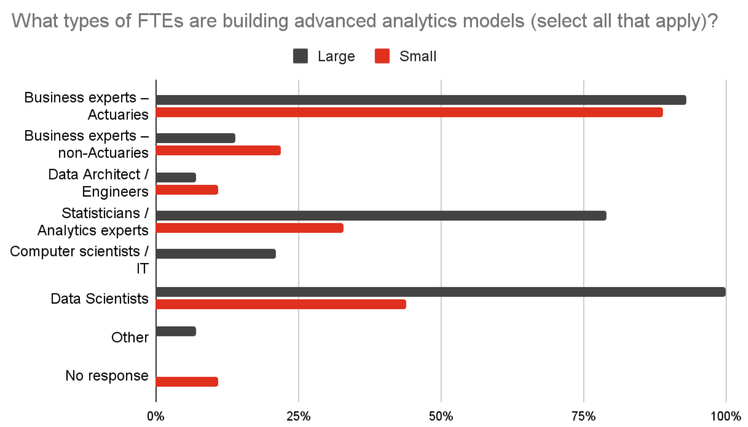


Figure 26 – Overall

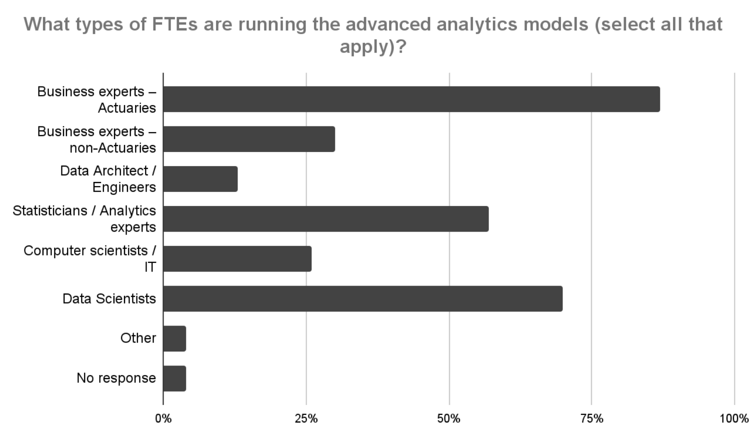


Figure 27 – By Origin

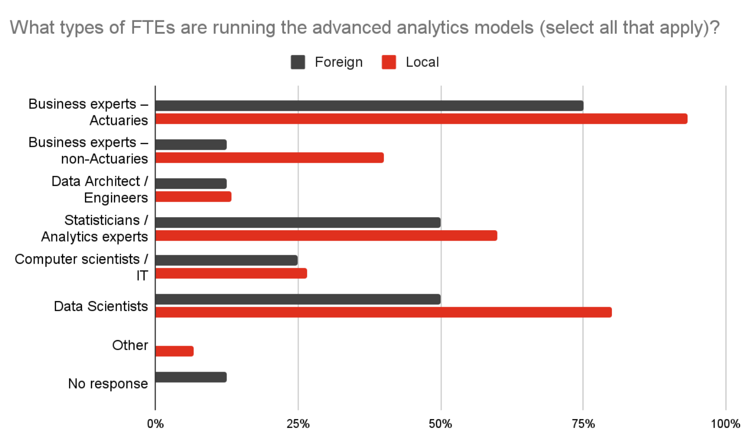
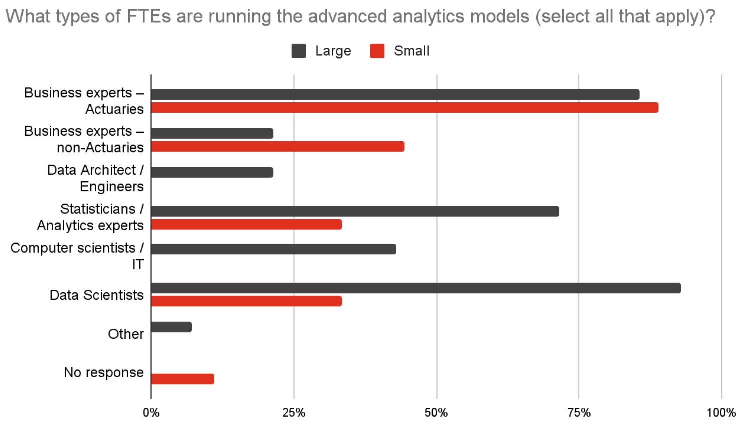


Figure 28 – By Size



In addition, despite the push from leadership to use advanced analytics, the collaboration and use of technical teams does not seem to be utilized to its full potential. While collaboration between technical teams with end-users happens most of the time, only 30% of respondents stated that it always happens (Figure 29), and under 5% stated that advanced analytics results are always used in business decision-making (Figure 30). This might, as some of our interviewees indicated, be a result of a lack of staffing to dedicate towards advanced analytics. That is, there are a lot of tasks where advanced analytics can be applied; however, there is a bottleneck in terms of capable staff, as indicated above.

Figure 29 – Overall

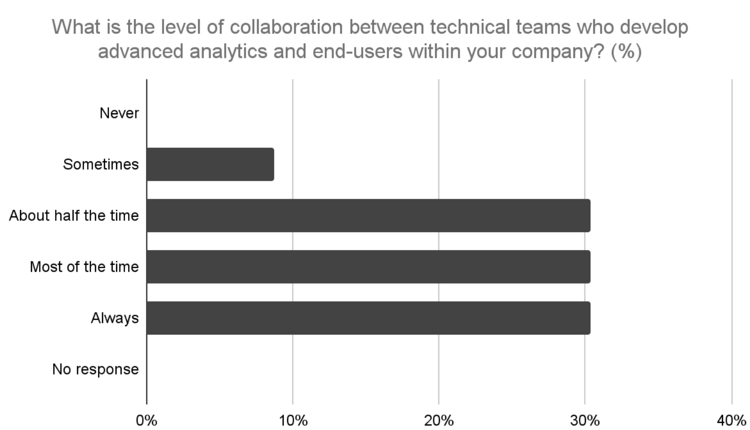
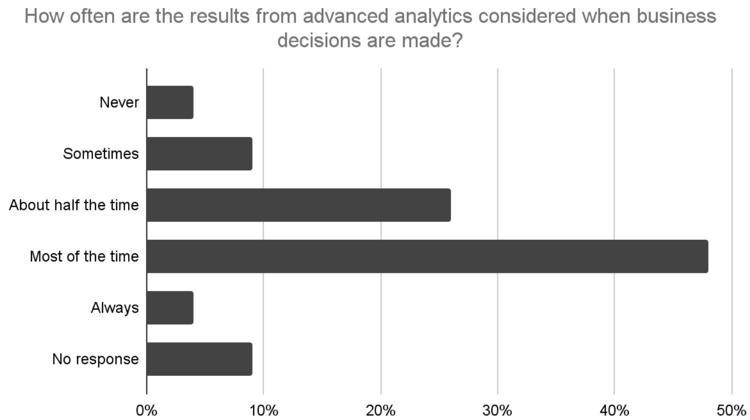


Figure 30 – Overall



Leveraging existing staff to train employees from other departments might be one way of bridging the gap between need and capacity. This, however, is not being done frequently enough, as almost 60% of respondents train their staff either yearly or even less frequently (Figure 31). Drilling down further to distinguish by company size, it seems that larger insurers provide more regular training to their staff than smaller insurers (Figure 32). Having a well-defined training process to upskill current employees might be worth pursuing.

Figure 31 – Overall

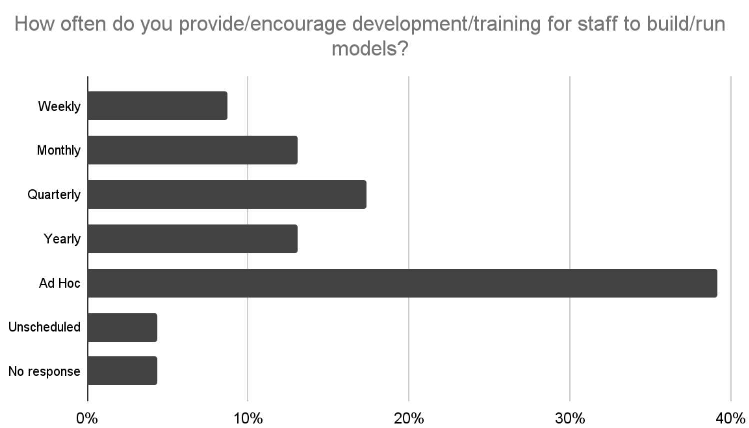
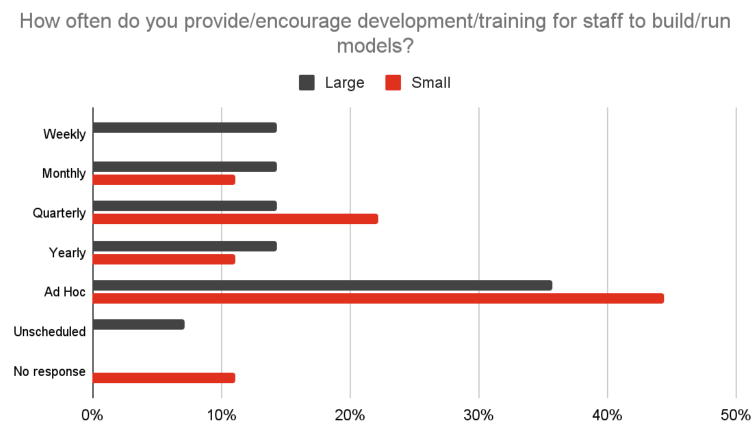


Figure 32 – By Size



Overall, it appears that there is a strong motivation from executives to push for advanced analytics; however, there are currently challenges in translating this enthusiasm into adding quality staffing, both through hiring and through the direct training of current employees.

Data

A model is only as capable as the data on which it is based, and thus data form the foundation of any advanced analytics pipeline. To this end, it was necessary to establish the quality of insurers' data practices. Furthermore, with insurers dealing with personally identifiable information (PII), governance is a key component in managing the associated privacy risks with such data.

Data quality was consistently good, with end-users giving high ratings for both completeness and accuracy of the data, most often scoring 4 or above (Figure 33). The data were also stated as being frequently reconciled or better, with the reconciliation process characterized as very comprehensive (Figure 35). Regarding data accuracy, smaller insurers rated themselves higher than large insurers (Figure 34); however, they also indicated longer time lags between data reconciliation processes when compared to larger insurers (Figure 36).

Figure 33 – Overall

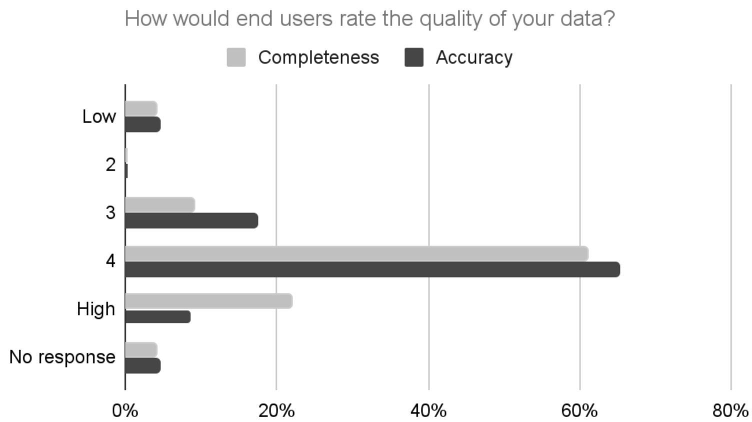


Figure 34 – By Size

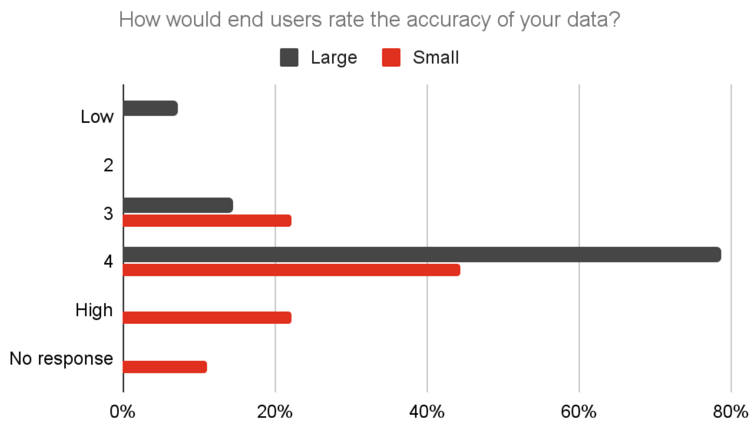


Figure 35 – Overall

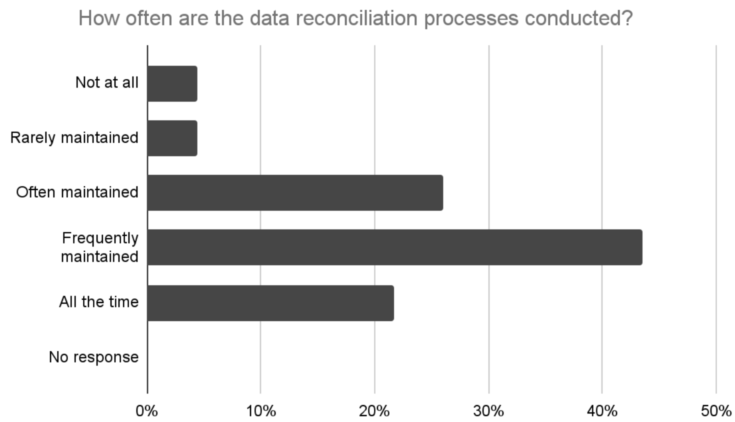
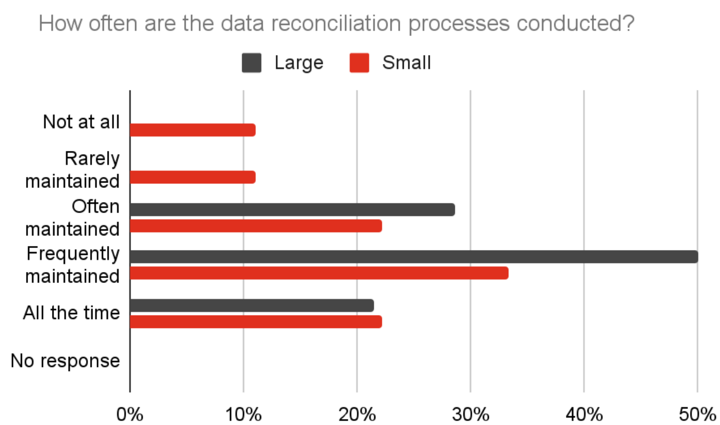


Figure 36 – By Size



Strong documentation greatly decreases confusion when new users have to access a data source, and generally makes working with data less burdensome. In comparison with the data quality, the amount of documentation with respect to data dictionaries was not as high as anticipated. Only about 40% of respondents stated that 75% of their data or better was defined in dictionaries (Figure 37), with the accuracy and the frequency of maintenance averaging 3.3 and 3.1 out of 5, respectively (Figures 40 and 43). Of note is the disparity between local and foreign insurers with respect to the completeness of data dictionaries, where international companies consistently stated higher scores (Figure 41). In addition, larger insurers consistently scored themselves lower than smaller insurers in terms of these two metrics of accuracy and frequency (Figures 42 and 45). This seems reasonable, given that larger insurers indicated utilizing significantly more data sources in business decisions in later questions (Figure 48).

Figure 37 – Overall

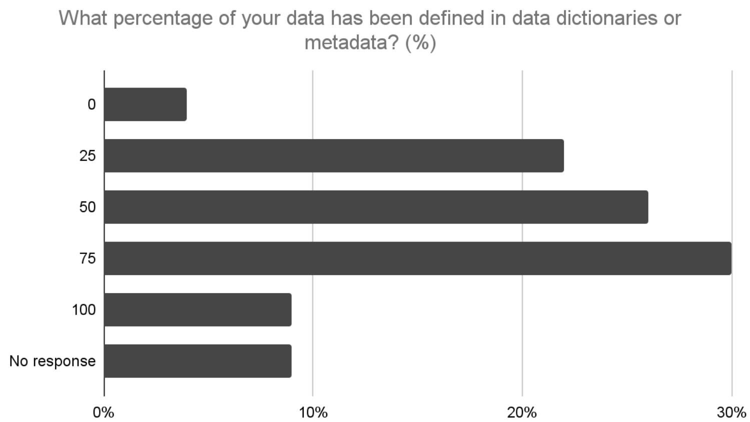


Figure 38 – By Origin

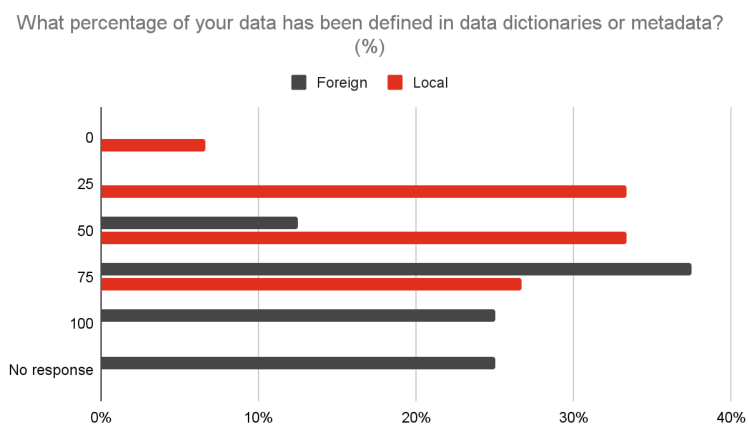


Figure 39 – By Size

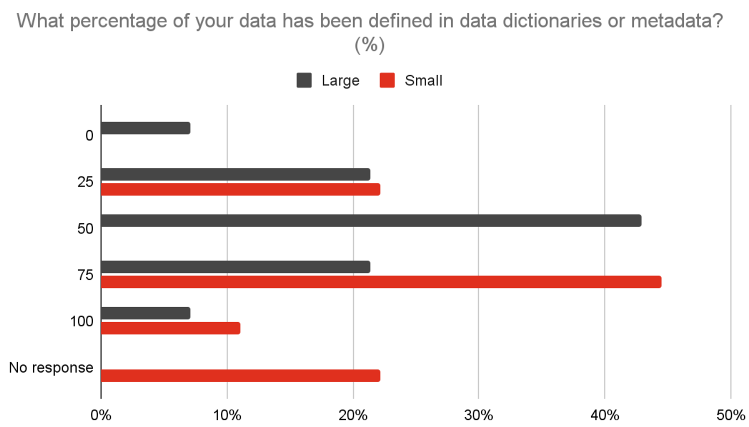


Figure 40 – Overall

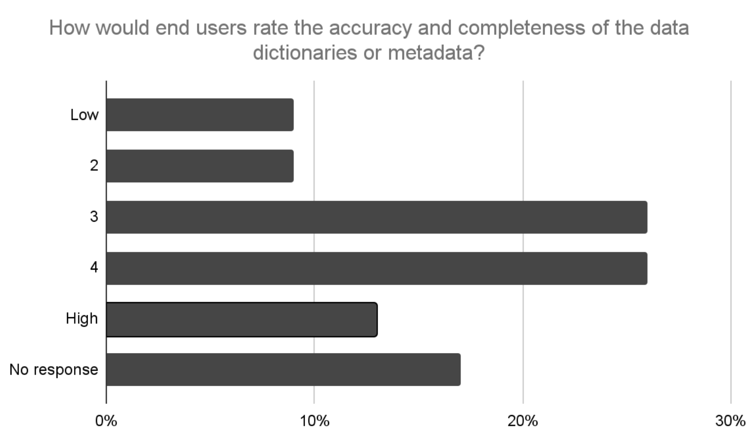


Figure 41 – By Origin

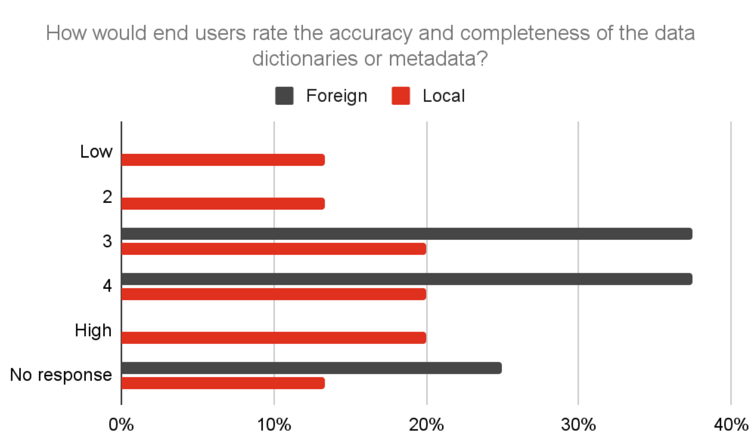


Figure 42 – By Size

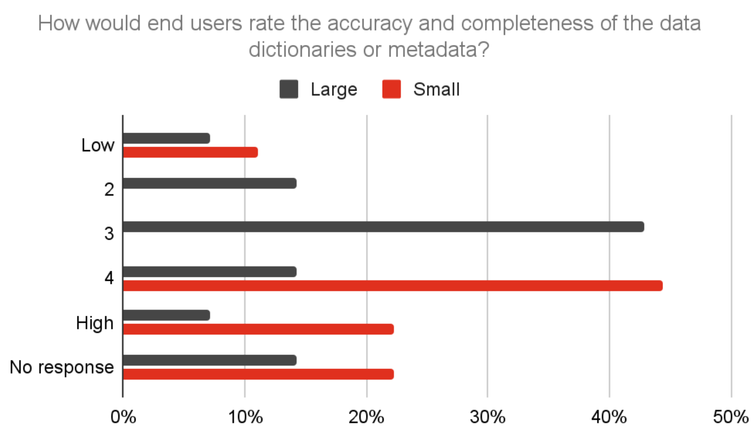


Figure 43 – Overall

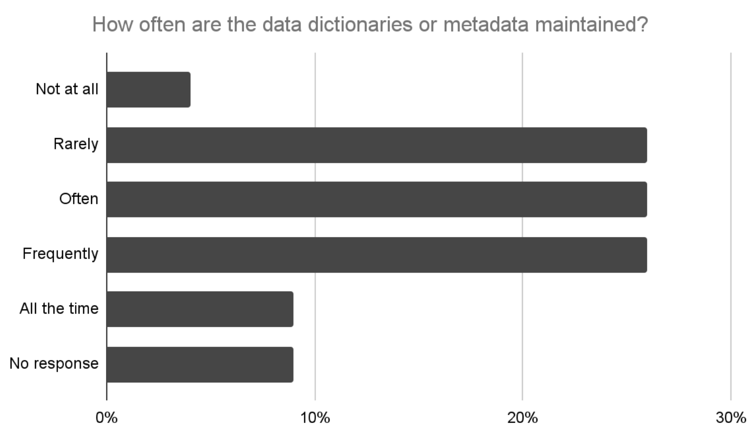


Figure 44 – By Origin

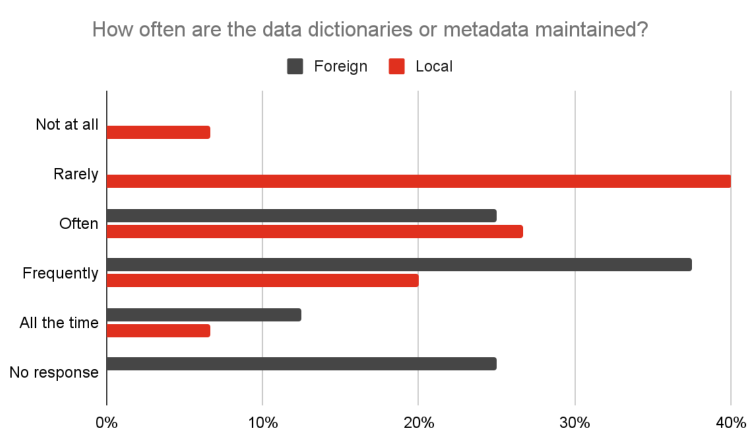
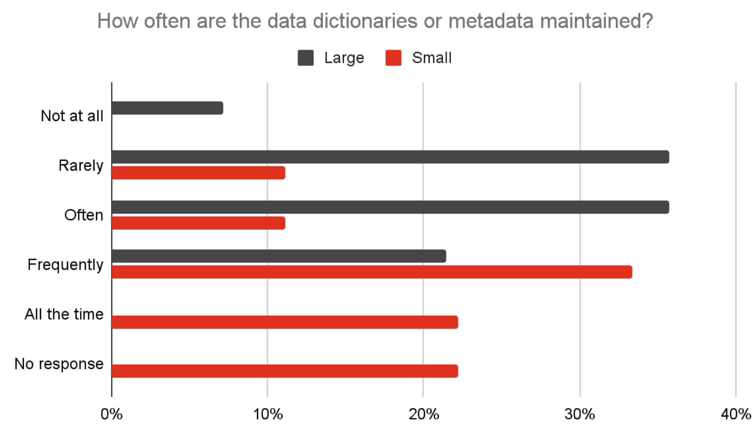
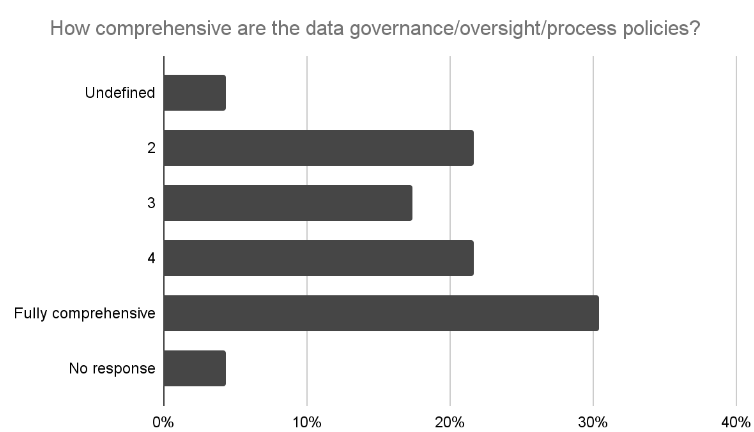


Figure 45 – By Size



Governance and oversight procedures within the P&C insurance industry were fairly strong, with policies rated very comprehensive, especially those concerned with data privacy and retention (Figure 46). As a consequence of stringent data governance, gaining access to data is more challenging, although such caution regarding data protection is warranted in light of the confidential information necessary for the insurance business.

Figure 46 – Overall



The number of data sources regularly used when performing analytics varied widely by insurers, with almost 50% of insurers stating fewer than 10 data sources, but a couple stating as high as 100 data sources (Figure 47). Larger insurers stated they leverage more data sources in analytics to assist in their business decisions (Figure 48). Most data sources that were directly mentioned in the survey related to the policy management system, although some insurers are starting to leverage additional external sources such as weather data and specialized location data about demographics beyond purely census data.

Figure 47 – Overall

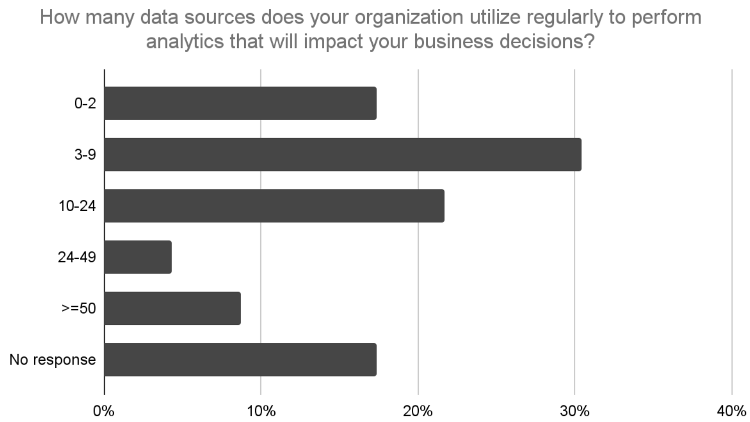
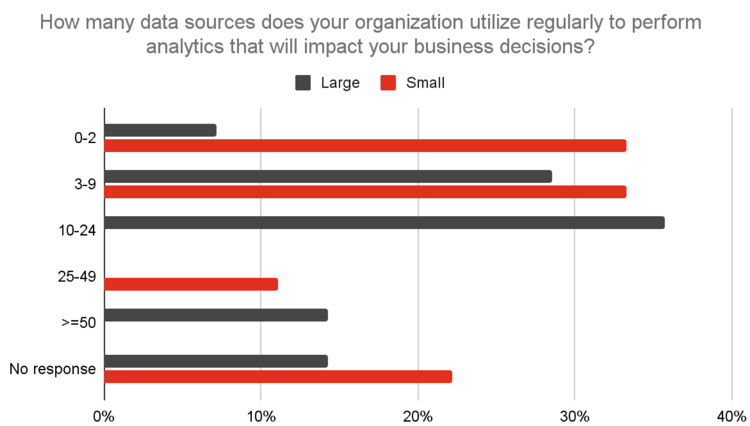


Figure 48 – By Size



In general, respondents stated that the quality of the data was very good, and while there are some challenges, some insurers mentioned in interviews that they gave lower scores to their governance and documentation not because the governance and documentation are lacking but because there is still significant potential to improve.

Technology

Investment in technology is an important enabler, ensuring fast turnaround times for developing analyses, enabling the use of new algorithms and efficient use of resources. Based on our interviews with select insurers, investment in technology seems to be playing a major role in the minds of respondents. That said, in a later question (Figure 78), over 60% of respondents said their ability to implement models was negatively impacted by the age of their IT infrastructure.

To deal with operational constraints, some of the larger insurers are moving to cloud solutions and Hadoop implementations, which can be deployed on a local cluster or in the cloud. Almost 40% of respondents indicated they had a Hadoop implementation. One of the major drivers was to use Spark and its toolset (Figure 49). During the interviews, respondents mentioned that their transitions away from legacy systems often took three to four years, with other respondents still in the midst of transitions. Larger insurers tend to use newer platforms while smaller ones are still heavily dependent on Excel for data storage.

Figure 49 – Overall

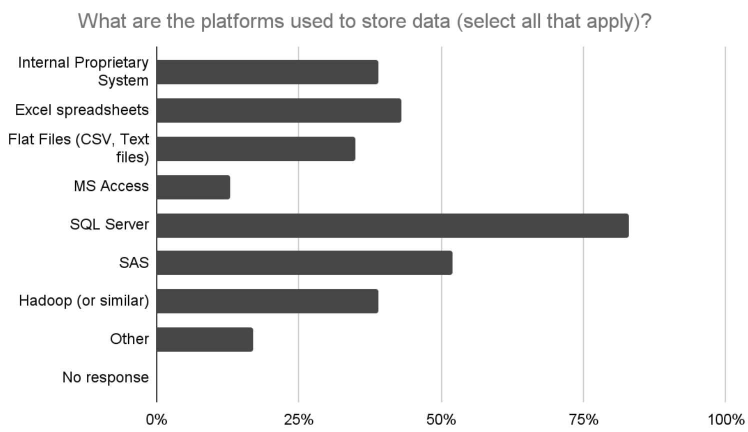
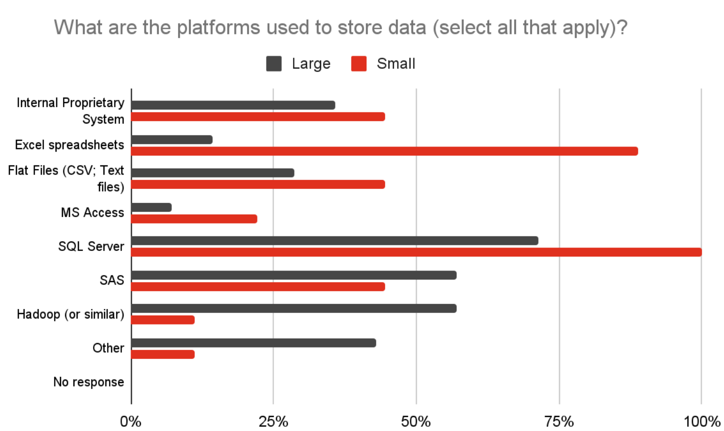
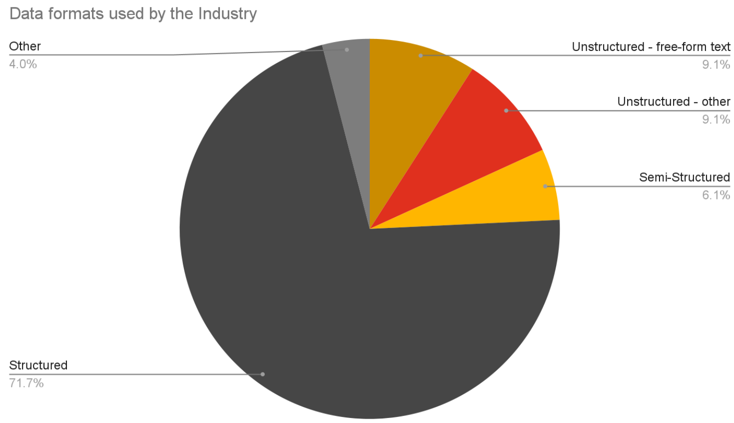


Figure 50 – By Size



The structure of the data is also an important consideration when utilizing advanced analytics. Some insurers specifically highlighted how they are trying to better deal with free-form text; however, over 70% of the data used by respondents was structured data, compared to just under 20% of the data being unstructured (Figure 51).

Figure 51 – Overall



To deal with the data, insurers are increasingly looking to use the open-source programming languages R and Python in data processing. Respondents are often still using SAS for some parts of their data workflow, with just over 50% of respondents using it for processing; however, for conducting analytics and exploratory data analysis, this number drops to just under 40% (Figures 52 and 54). Figure 53 demonstrates that larger insurers are more willing to use coding languages for data processing rather than dashboarding software like Power BI and Tableau. In Figure 55, we see that large insurers tend to be using a greater variety of software.

Figure 52 – Overall

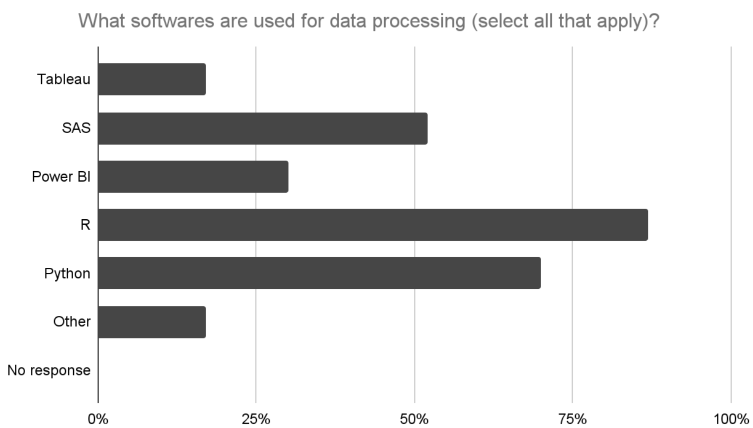


Figure 53 – By Size

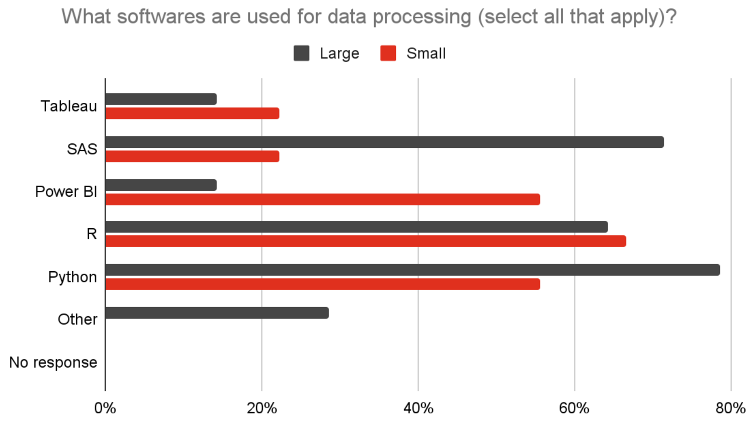


Figure 54 – Overall

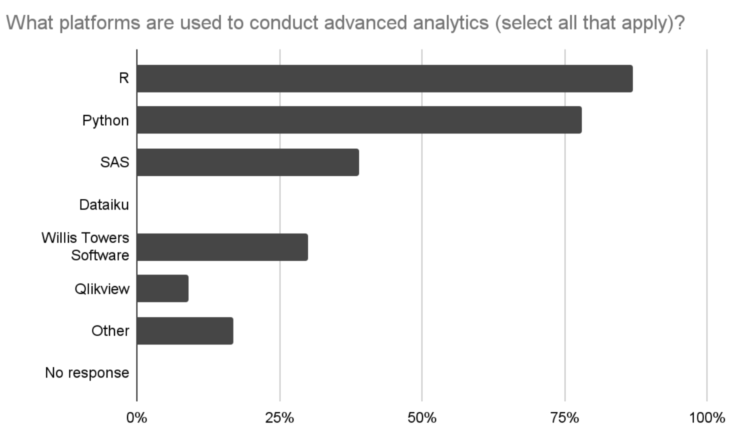
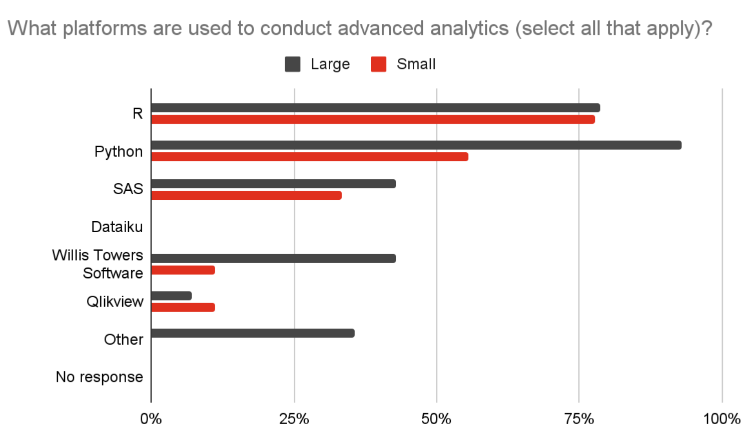


Figure 55 – By Size



For exploratory data analysis (EDA), Excel is still a frequently used tool, being used nearly two-thirds of the time (Figure 56); however, for visualization in reporting, other tools such as Power BI and Tableau are finding increasing use (Figure 58). Larger insurers are increasingly using Python for EDA, with SAS also being used by large insurers for such purposes (Figure 57).

Figure 56 – Overall

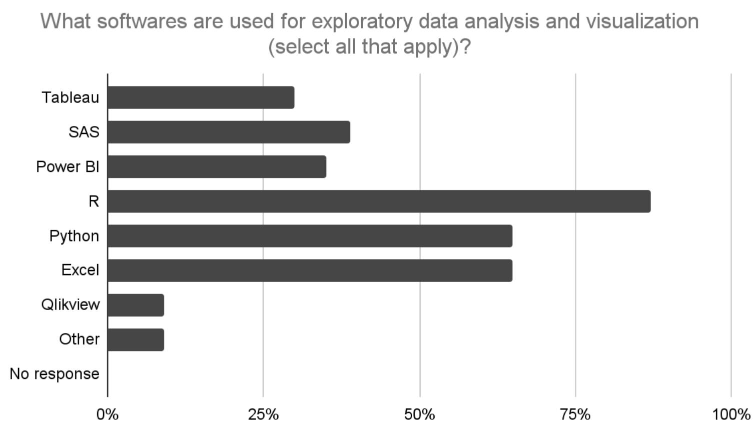


Figure 57 – By Size

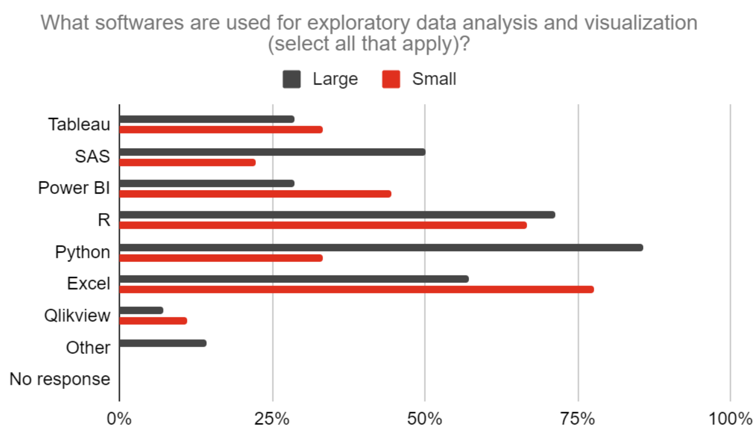
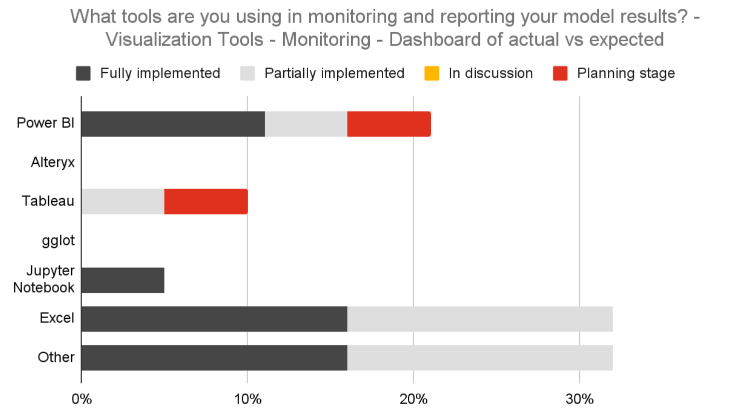
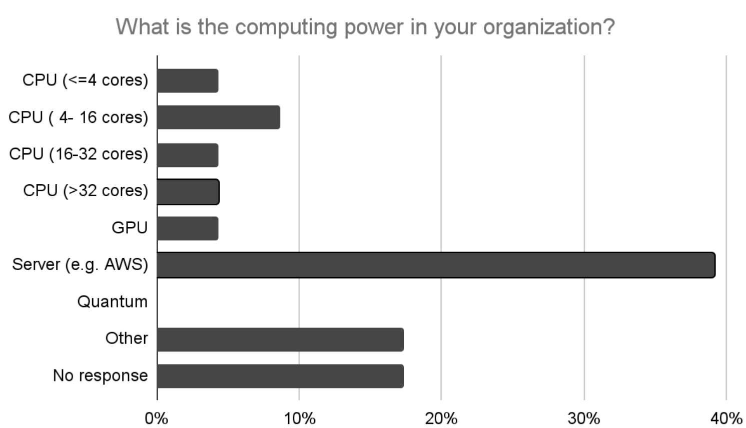


Figure 58 – Overall



Finally, to deal with tuning models, insurers are moving away from running processes on their local machine, with many of them tuning models remotely either on dedicated machines with many cores or on dedicated servers. One team mentioned that it is leveraging graphics processing units (GPUs) to speed up the modeling process, while others are taking advantage of Amazon SageMaker to run their models in the cloud (Figure 59).

Figure 59 – Overall



In general, insurers are investing significant resources on updating their infrastructure, regarding both their data architectures and their computing power, to decrease modeling time. To take advantage of these sophisticated resources, insurers are putting a large emphasis on coding skills, with much of the analyses and EDA being done using R and Python.

Model and business usage

Modeling is a crucial component of the insurance business and was a heavy focus for the survey. Respondents were asked to state both their methodology and progress with regard to many potential analytics use cases within their organization. The use of more advanced techniques is becoming more prevalent across the whole business, and not solely for the tasks of product development, pricing, and underwriting. While these latter tasks remain the ones with the highest number of responses, indicating an understandable heavy focus given that accurate pricing of a policy is the most critical task of the insurer, increased focus is being paid to using advanced models for reserving practices. Some insurers also mentioned the challenge in keeping their increasing number of models up to date, which often takes precedence over developing new models.

It is important to have governance procedures and policies when developing a new model, as model governance plays an important role in risk management and in the model lifecycle. As such, respondents were asked to state the comprehensiveness of model review and the frequency of guideline review (Figures 60 and 62). Over 50% of respondents scored themselves 4 or above on the former, and most companies reviewed their policies either yearly or on an ad hoc basis. Differences arose when comparing larger to smaller insurers, with larger insurers generally performing better (Figures 61 and 63).

Figure 60 – Overall

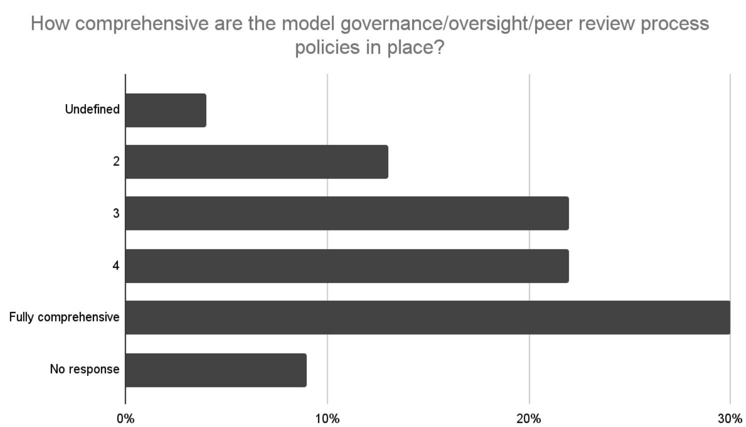


Figure 61 – By Size

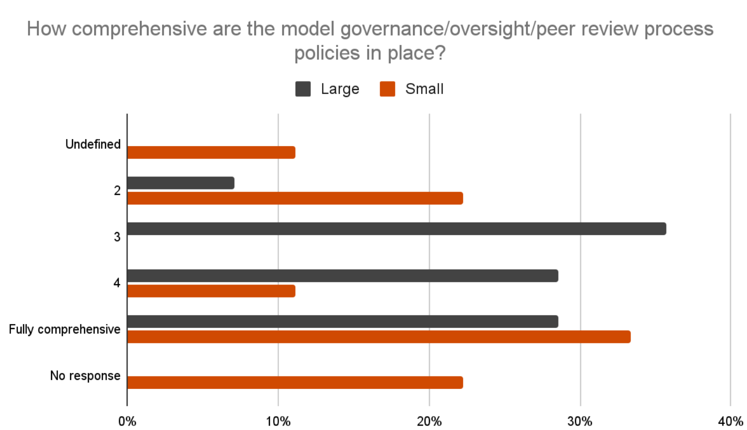


Figure 62 – Overall

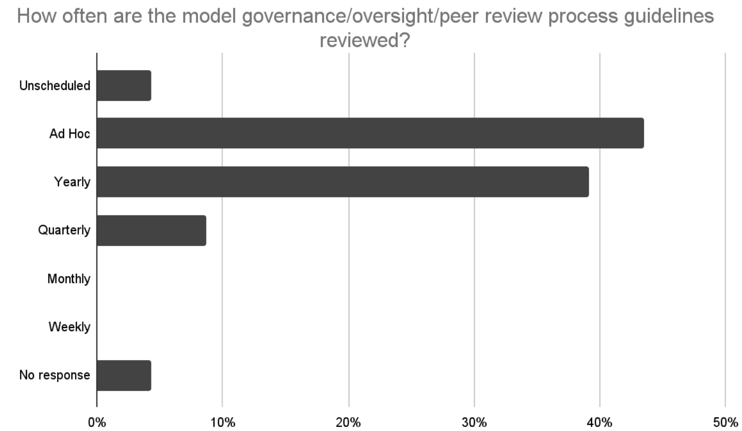
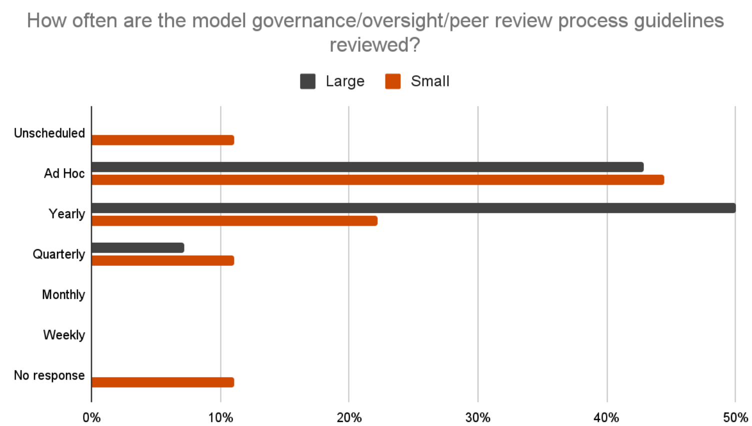


Figure 63 – By Size



Behavioural economics, a mix between economic theory and psychology, is a tool that is being explored for price optimization, but also, as some insurers have mentioned, to help maximize the number of quotes being offered. Under 10% of respondents stated that they use behavioural economics most of the time or always (Figure 64). In interviews, we were told it is less of a concern for reinsurers and commercial underwriters, with their primary concern being purely the risk aspect of whether to insure or not as opposed to maximizing long-term profit. Large insurers tended to use these behaviour-based techniques more frequently (Figure 65).

Figure 64 – Overall

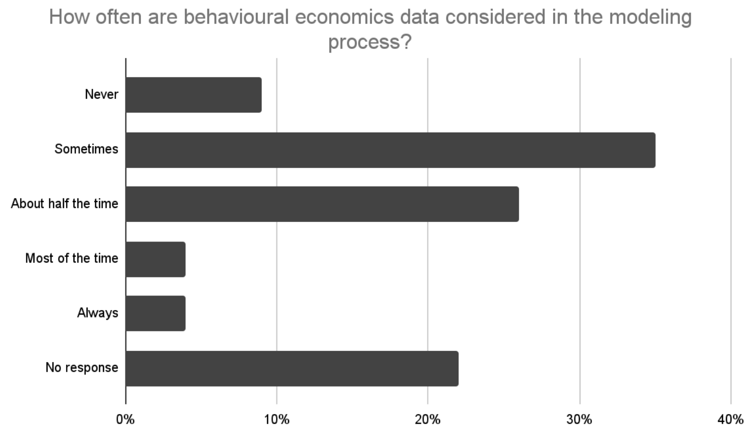
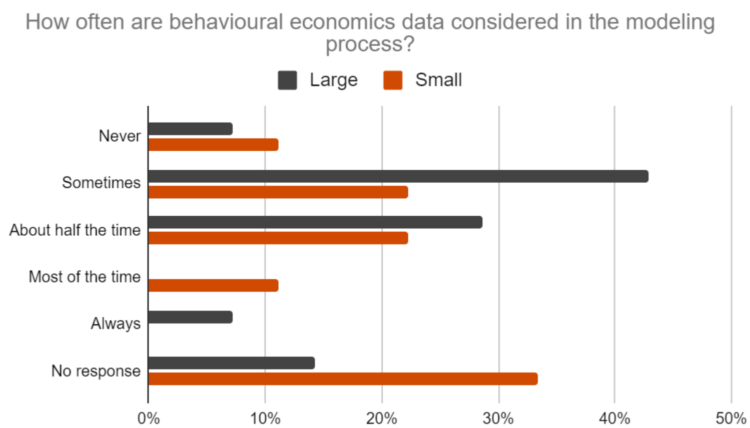


Figure 65 – By Size



Model selection-criteria metrics are integrated into most insurers' processes (Figure 66), although there is a large disparity between large and small insurers, with the former having additional metrics to supplement their modeling decisions (Figure 67). Unsurprisingly, larger companies who tended to use more "black box" techniques needed to use other criteria, as the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are not well defined for such models.

Figure 66 – Overall

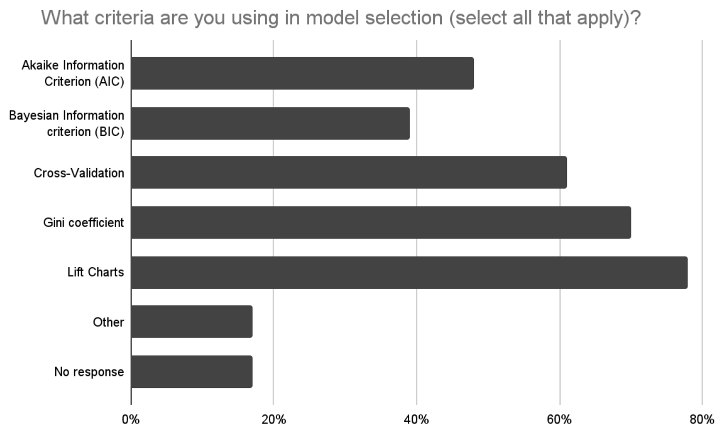
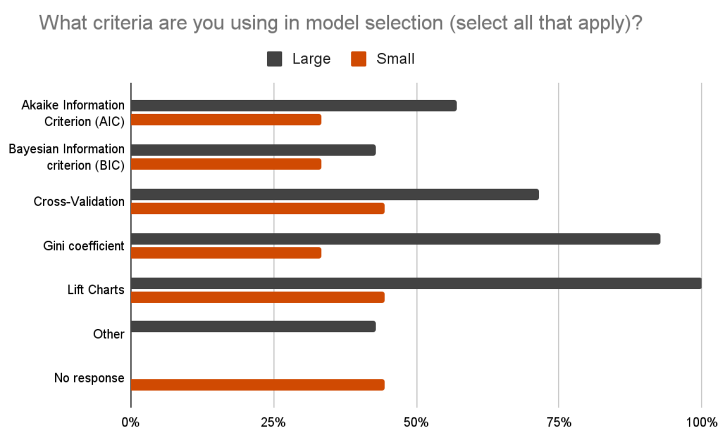


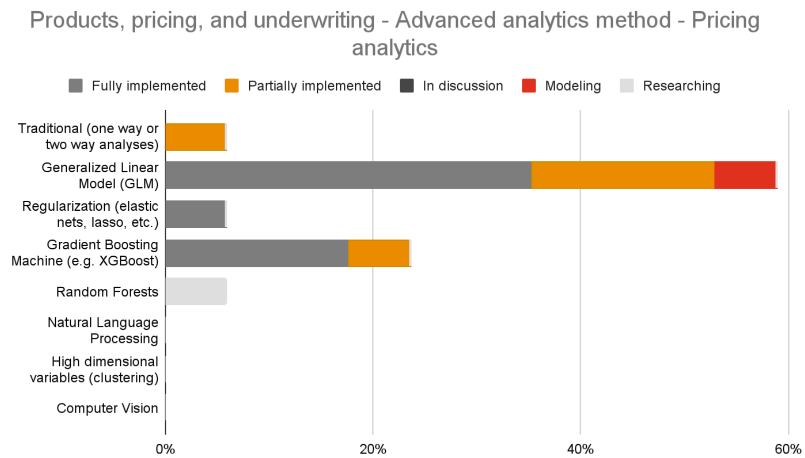
Figure 67 – By Size



Questions concerning model implementation for particular use cases are displayed as two-way graphs, with the method and progress level displayed. Some respondents stated a method but listed the progress status as "Not Started." These "Not Started" responses were not included in the analysis, as they were deemed not to be of consequence. Thus, only survey respondents with a progress level above "Researching" were included and displayed in the calculation of percentages.

Regarding pricing and underwriting, almost 60% of respondents continue to use GLMs, with many auto insurers pointing out that they are constrained by the regulatory burden of well-explained models to justify pricing decisions. Additionally, some insurers mentioned that the techniques used to generate the initial models are sometimes different from updating the model, in that the initial model will be tuned but subsequent iterations will simply be trended due to time constraints.
A couple of respondents also mentioned that they might use a GLM, for example, but that they would generate variables to capture additional signals from a more advanced model and back out the result (Figure 68).

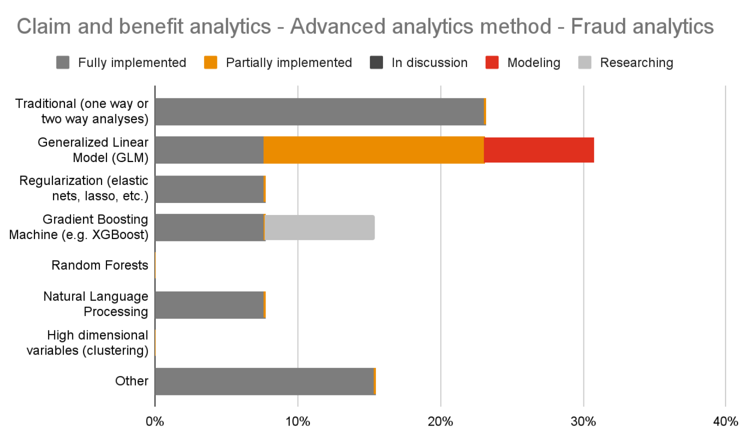
Figure 68 – Overall



These regulatory constraints are less present in other parts of the business and lines outside of personal auto insurance. Without such constraints, as in the analysis of fraudulent claims, insurers have more leeway to focus on predictability as opposed to explainability, since the primary concern is whether a claim is fraudulent or not as opposed to whether the pricing is equitable.

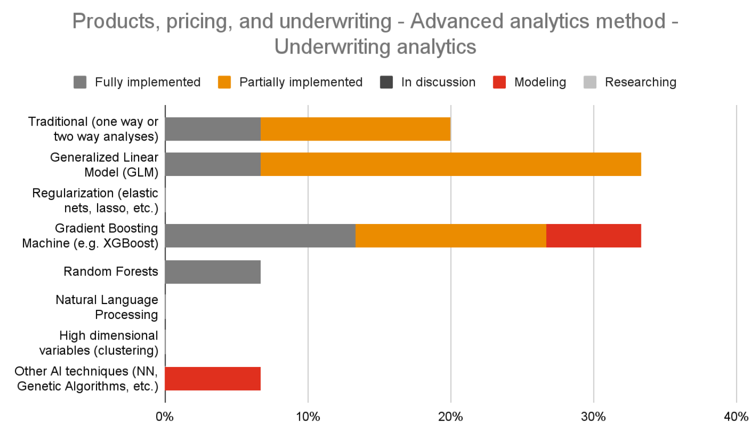
Areas where we see the most developed and largest array of advanced analytics are in fraud and claim loss analytics. The need to model fraud appears more important for primary insurers as opposed to reinsurers. Almost 30% of respondents have turned to regularization, Gradient Boosting Machines (GBMs) and Natural Language Processing (NLP) to identify possible fraudulent claims (Figure 69). One challenge some interviewees pointed to is in mining the data, as phone transcripts are often long, giving the standard NLP algorithms challenges, let alone the lack of consistency in customer stories. Nonetheless, interviewees often highlighted improved identification of fraud as one of the business cases with major success.

Figure 69 – Overall



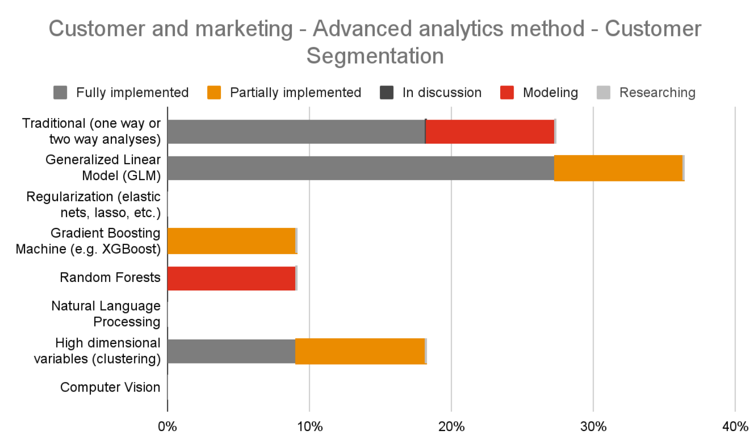
Regarding more traditional areas of analytics including underwriting and customer segmentation, respondents are also trying to increase their use of advanced analytics. Almost 50% of respondents are using neither traditional methods nor GLMs to analyze their underwriting practices (Figure 70).

Figure 70 – Overall



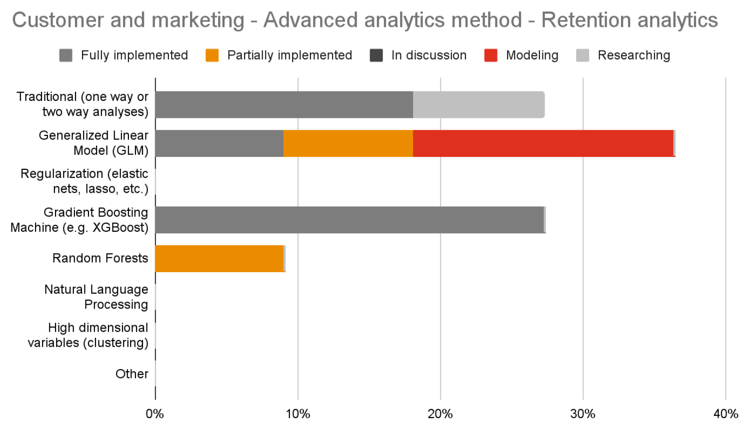
Customer segmentation is another area where insurers are placing a large amount of attention in using advanced analytics to better understand their customers, from both a pricing and an acquisition standpoint (Figure 71). Almost 40% of respondents are using alternatives to traditional techniques and GLMs to generate homogenous groups of policies for better modeling.

Figure 71 – Overall



Retention analytics is another area where there has been a lot of time invested into generating better models. Almost 40% of respondents stated that they were using a non-traditional and non-GLM technique (Figure 72). Some of the related areas are less developed, those being the related Customer Lifetime Value and Price Optimization, although in our interviews a few insurers mentioned that Price Optimization specifically has long been a target. Others, on the other hand, have been using Price Optimization for many years.

Figure 72 – Overall



Overall, insurers are increasingly moving towards black-box solutions for the more traditional actuarial tasks of pricing, reserving, fraud, and churn. There were far fewer responses regarding the use of analytics in marketing, strategic acquisition, and operations, and as such we did not comment due to a lack of credible data. This lack of responses might be a function of the particular respondents being primarily linked to those more traditional actuarial tasks and to a lack of knowledge about other internal teams.

Implementation

The ability to implement a model forms the last step of the process workflow. Two-thirds of respondents are able to deploy changes to their business-critical models within a quarter, although several of the larger insurers take six months (Figure 73). Nonetheless, larger insurers generally implement their models faster (Figure 74).

Figure 73 – Overall

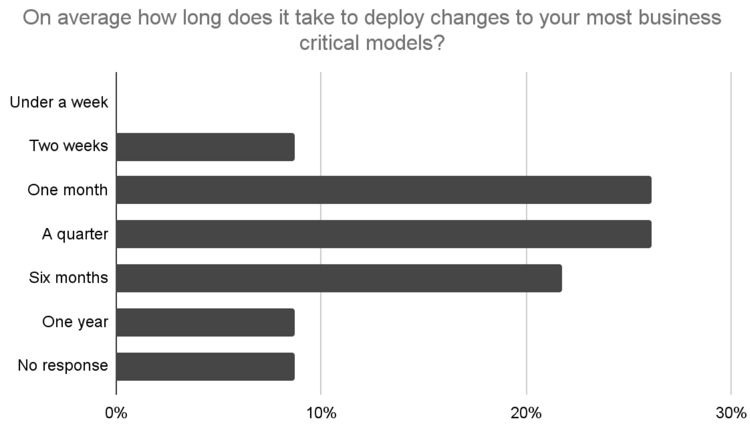
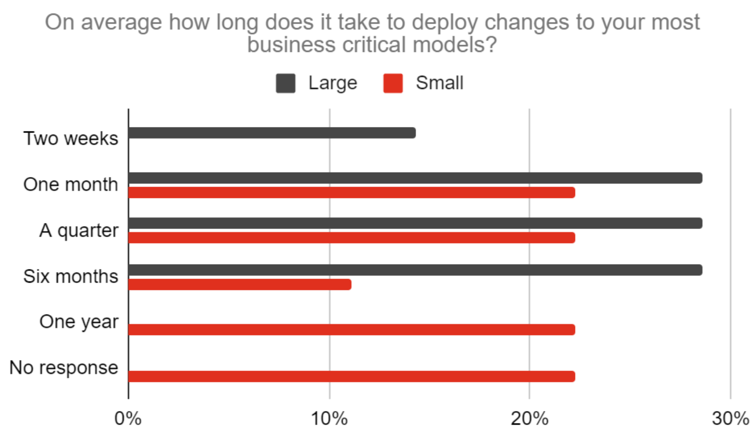


Figure 74 – By Size



Additionally, insurers are increasingly implementing models to deal with real-time data processing instead of merely batch processing, with some of the more advanced insurers having in excess of 50% of their models doing real-time processing (Figure 75). Similarly, insurers are building Application Programming Interfaces (APIs) to help connect to their models, with the best insurers having over 50% of their models accessible via APIs (Figure 76).

Figure 75 – Overall

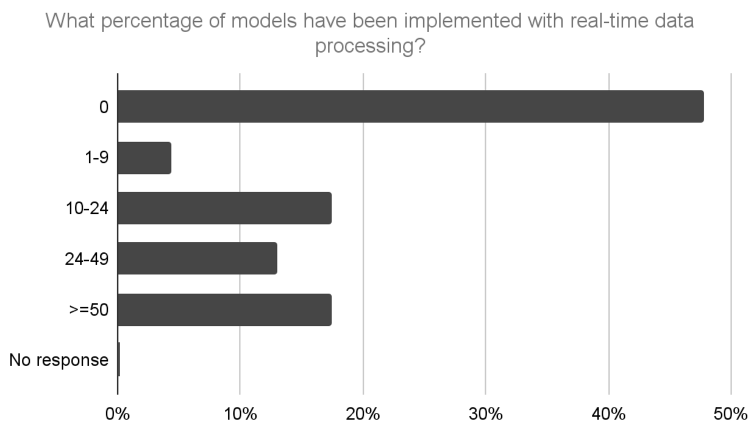
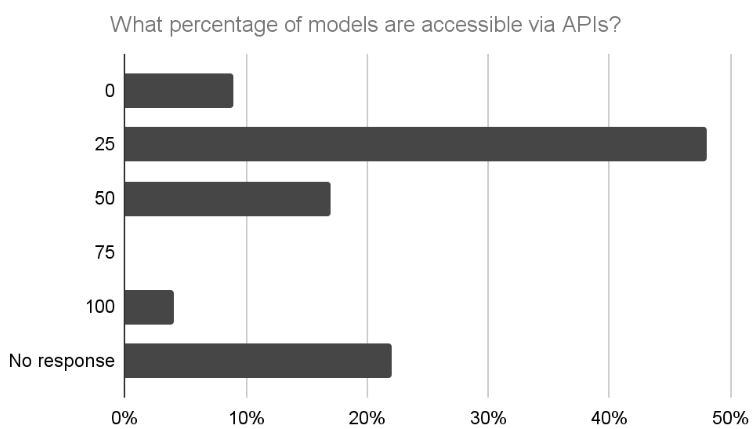


Figure 76 – Overall



There was a major distinction between larger insurers and smaller ones regarding the amount of prototyped work that tends to get scrapped due to implementation difficulties (Figure 78). It remains unclear whether these difficulties are a result of large insurers focusing on experimentation afforded by additional staffing or due to larger insurers having greater difficulty in leveraging their models.

Figure 77 – Overall

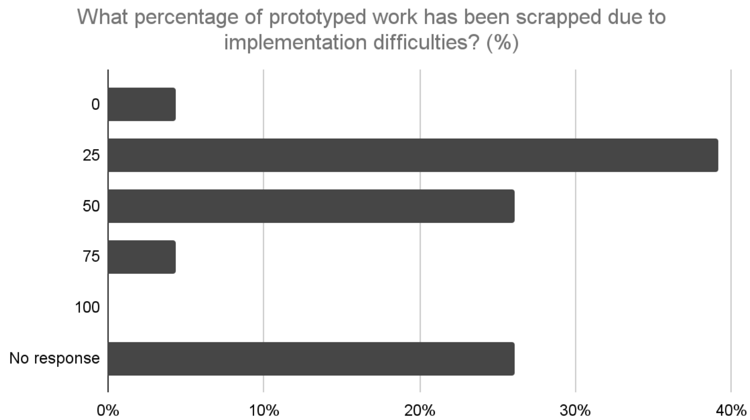
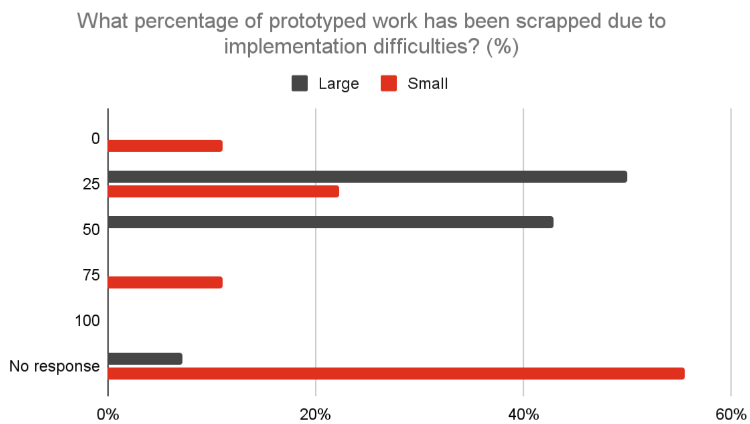


Figure 78 – By Size



Most insurers appear to be affected by aging infrastructure inhibiting their ability to utilize advanced analytics (Figure 79). The age of the IT infrastructure appears to be impacting larger and local insurers more than smaller and foreign insurers, respectively (Figure 80). Those larger insurers that have fully updated to a newer infrastructure stated they are no longer having difficulties in implementation; however, as stated above, this infrastructure transformation is burdensome and time-consuming.

Figure 79 – Overall

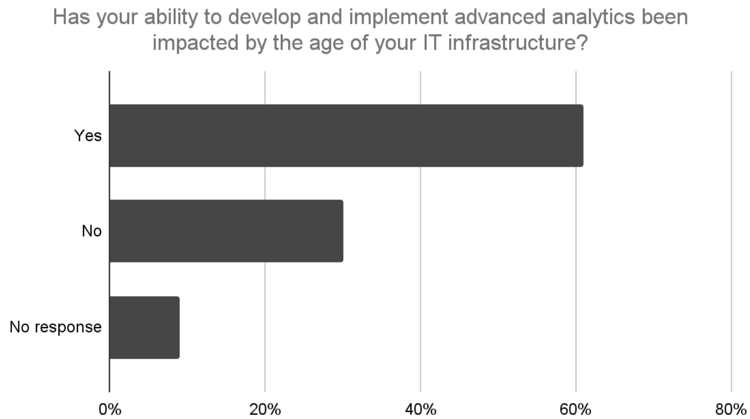
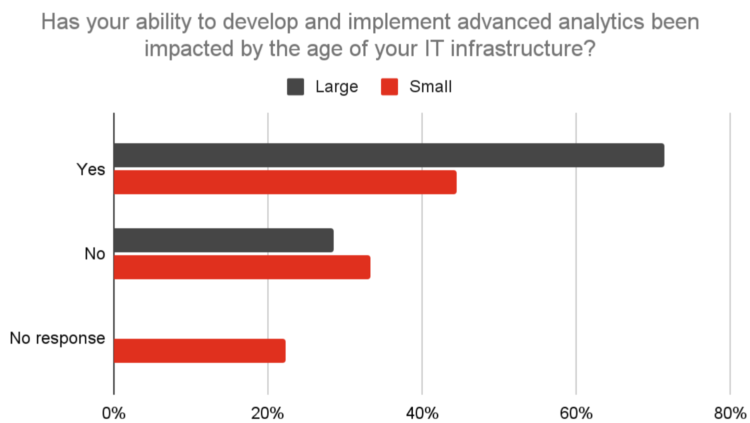


Figure 80 – By Size



Overall, challenges persist either in dealing with aging infrastructure or in transitioning to new ones. One of the challenges consistently highlighted was the availability of IT resources as analytics teams are finding it challenging to deploy advanced analytics solutions. Nevertheless, companies are increasingly able to streamline their implementation process with the goal of speeding up the time lag between model creation and model implementation, with many utilizing APIs to help update these processes.

Considerations and opportunities in the age of analytics

In this section, we seek to describe some of the challenges faced by the P&C insurance market and some of the possible benefits of taking advantage of the opportunities afforded by the advances in advanced analytics. Examples of advanced analytics applications we have observed in other countries and industries are also discussed in this section. Similar to the previous benchmark analysis section, we again organize our discussion as follows:

- Management and internal support – establishing the support from management and a general analytics culture
- Data – creating well-documented, accurate data, with strong oversight and privacy considerations
- Technology – utilizing newer infrastructure and increased processing power to remove impediments to growth and development
- Model and business usage – using novel techniques and applying them throughout the business
- Implementation – speeding up the deployment of models to drive results

Based on our survey responses and interviews, insurers benefited most from predictive analytics in pricing, claim, and underwriting functionalities (Figure 81). Reinsurers indicated that catastrophe modeling has been heavily utilized. Furthermore, respondents indicated that IT infrastructure and business buy-in have been major roadblocks to implementing advanced analytics (Figure 82). Data issues, implementation lag, and the implementation process are inhibiting insurers as well. In addition to guiding insurers' data privacy and retention policies, the regulators' impact has been restricted to those highly regulated lines of business, such as personal auto insurance.

Figure 81

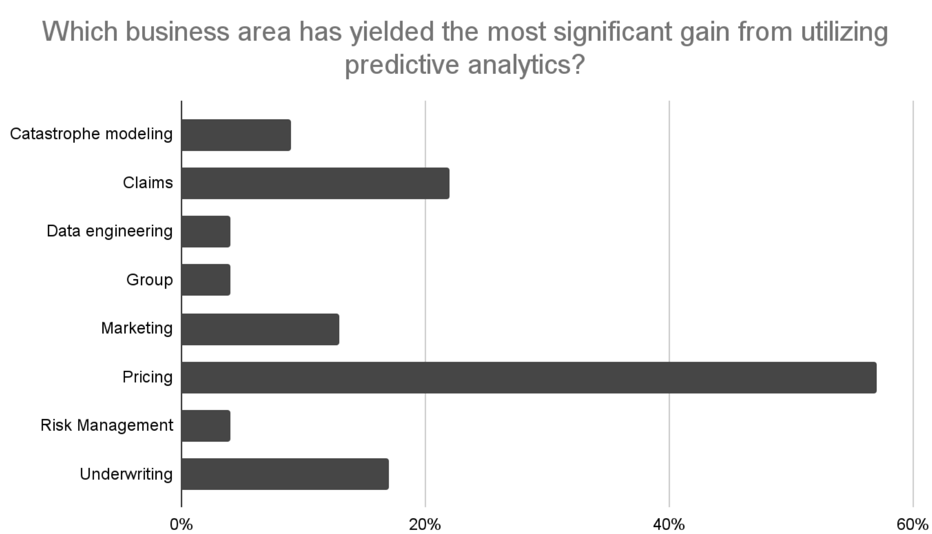
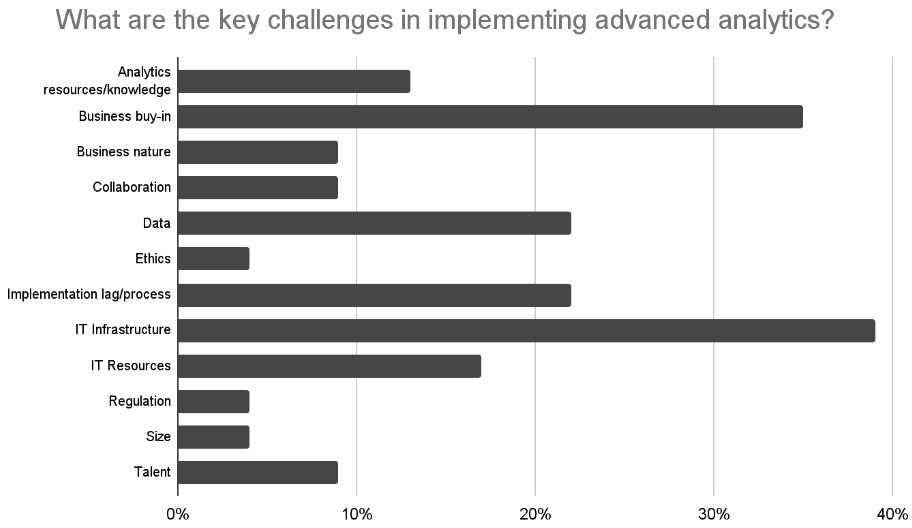


Figure 82



Businesses in every sector are eager to claim their piece of the potential AI windfall, which one report estimates could translate to an infusion of US$15.7 trillion into the global economy by 2030.[1] These financial gains will require changes in how processes are treated and are not without their own risks, as questions about unintentional bias and data concerns loom large.

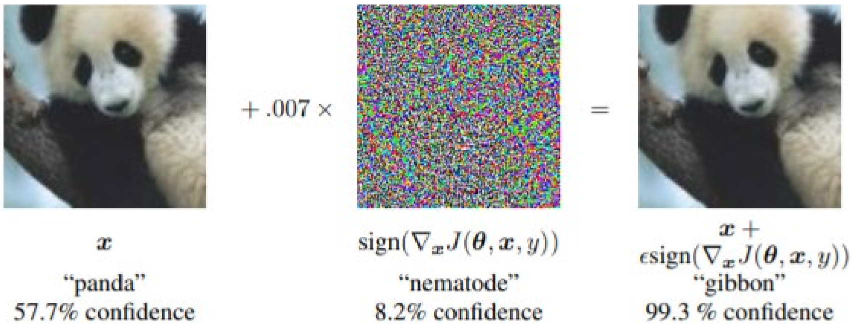
One major issue is that "when artificial intelligence makes a mistake, that mistake looks completely foolish to humans, or almost evil."[2] An example of strange ML algorithm results from adversarial research can be seen below.[3] As a result, it is often difficult to trust these abstract systems. At the same time, questions are also raised about the limits of AI and the capabilities of the current algorithms that we use. In a recent article,[4] one of the points highlighted was the inability of AI to use knowledge-based reasoning as opposed to pure pattern recognition.

[1] Anand S. Rao and Gerard Verweij. Sizing the prize: What's the real value of AI for your business and how can you capitalise? PwC Global, 2017. www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.

[2] Tim Adam. Daniel Kahneman: "Clearly AI is going to win. How people are going to adjust is a fascinating problem." *The Guardian*, 2021. www.theguardian.com/books/2021/may/16/daniel-kahneman-clearly-ai-is-going-to-win-how-people-are-going-to-adjust-is-a-fascinating-problem-thinking-fast-and-slow.

[3] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. *Explaining and Harnessing Adversarial Examples*. Google, 2015. https://arxiv.org/pdf/1412.6572.pdf.

[4] Christopher Mims. Self-driving cars could be decades away, no matter what Elon Musk said. *Wall Street Journal*, 2021. www.wsj.com/articles/self-driving-cars-could-be-decades-away-no-matter-what-elon-musk-said-11622865615.



Goodfellow et al. *Explaining and Harnessing Adversarial Examples*[5]

This in no way diminishes the advances that predictive modeling and AI have had, let alone the way business automation is transforming the way we work. These technological advances are improving user interactions with companies, cutting costs, and making our lives easier.

Management and internal support

The entire P&C insurance industry has recently experienced a significant change in the way companies interact with their employees and customers. Following the pandemic, large swathes of the population have transitioned to working from home, and this has drastically changed the way we interact both with each other and with our work. For employees, there are now more emails and check-ins than before,[6] increased distractions of dealing with the rest of the family at home,[7] and a push towards increased flexibility. These flexible working conditions might have lasting effects in the way people interact with each other in the workplace. For customers, there are also many questions about how their behaviour will change with regard to the way they interact with technology, their buying preferences, and their associated risks. It is important for senior management to be aware of the new technologies and then blend them in business decisions, daily operation, product design, and customer interactions, in order to adapt to the dynamic market and new norms post the pandemic depending on the insurer’s unique circumstances. In the following sections, we provide some considerations and opportunities to the senior management regarding the most recent macro environment changes that may impact insurers.

COVID-19 changes

The pandemic has led to changes in consumer preferences and behaviours, resulting in customers being more open to telematics and other new product offerings. These changes have also manifested themselves as changes in levels of exposure to different risks.

The pandemic has posed a direct challenge to the way insurers have previously behaved. Prior to the pandemic, industry executives employed a more deliberate, careful strategy that allowed tech capabilities

[5] Goodfellow et al., *Explaining and Harnessing Adversarial Examples*.
[6] Danielle Kost. You’re right! You are working longer and attending more meetings. Harvard Business School Working Knowledge, 2020. https://hbswk.hbs.edu/item/you-re-right-you-are-working-longer-and-attending-more-meetings.https://hbswk.hbs.edu/item/you-re-right-you-are-working-longer-and-attending-more-meetings.
[7] Sophie Leroy and Theresa M. Glomb. A plan for managing (constant) interruptions at work. *Harvard Business Review*, 2020. https://hbr.org/2020/06/a-plan-for-managing-constant-interruptions-at-work.

to evolve through incremental innovations rather than sudden breakthroughs. The emergence of COVID-19 forced an industry slow to embrace change to figure out new ways of working in a matter of weeks, implementing major shifts in technology deployment and usage, agile problem-solving, and more.[8]

This shift of people moving to work-from-home and hybrid working situations might have lasting effects on the risks to personal property. Neos, a connected home insurance provider in the United Kingdom, revealed to the Casualty Actuarial Society (CAS) that during the lockdowns in England in 2020 overall loss ratios dropped by over 30%.[9] It reasoned that this drop in claims was a function of people being home to deal with at-home disasters and reduced crime. Market research seems to indicate that homeowners are also increasingly open to insurer-connected property telematics.[10] Privacy concerns remain; however, given the success of Ring and Nest, the benefits of such devices might help people become more comfortable with connected homes. While a question concerning the modeling of connected homes had few respondents in our survey, there might be a large opportunity to gain additional information, alert homeowners to issues sooner, and decrease the frequency of claims. That is, management could further support their insurer's analytics efforts by looking to change their products and services, with home telematics growing in popularity. To further this point, a recent survey showed that both Canadian and US insurers are paying increasing attention to telematics in the home, with almost 9% of carriers using it.[11]

Another knock-on effect is in product design. For instance, as a result of customers staying home more, there has been increased interest in Pay-As-You-Drive (PAYD), policies that use mileage as an exposure basis, as a means to grow one's business. For a Canadian issuer, year-over-year growth of their PAYD policies was 300% in 2020.[12] Existing companies with Pay-How-You-Drive (PHYD) would also have been able to monitor driving behaviour during the pandemic with greater understanding. It is yet unclear whether changes in driving behaviour and consumer preference will persist in the coming years; however, companies that are able to more quickly implement changes to their models should have a distinct advantage.

These examples illustrate the changing insurance landscape due to or accelerated by the pandemic with transitions to new product offerings changing the way insurers are trying to attract new customers and retain existing ones. These changes, however, are not limited to the types of products offered but also directly impact an insurer's ability to use past data to predict future events. If, for example, these are structural changes leading to a "new normal" in property claim frequency, then companies that quickly adjust their rates to better account for this change might have an opportunity to significantly grow their business in a short period of time, with fewer concerns about adverse selection. On the other hand, companies might inadvertently subject themselves to unforeseen risks, with some risky driving behaviours

[8] PwC United States. Moment of truth: Why insurance carriers should rethink their customer service models now. 2020. www.pwc.com/us/en/industries/insurance/library/carriers-customer-service-models-covid-19.html.

[9] Annmarie Geddes Baribeau. Domestic perils: 2020 – a pivotal year for homeowners insurance. *Actuarial Review*, 2021. https://ar.casact.org/domestic-perils-2020-a-pivotal-year-for-homeowners-insurance/.

[10] Annmarie Geddes Baribeau. Getting personal: Can IoT do for homeowners insurance what telematics did for auto coverage? *Actuarial Review*, 2021. https://ar.casact.org/getting-personal-can-iot-do-for-homeowners-insurance-what-telematics-did-for-auto-coverage/.

[11] Willis Towers Watson. 2019/2020 P&C Insurance Advanced Analytics Survey Report (North America): Fields of dreams – three areas dominate the field of insurers' aspirations for advanced analytics. 2020. www.willistowerswatson.com/en-CA/Insights/2020/01/fields-of-dreams-three-areas-dominate-the-field-of-insurers-aspirations-for-advanced-analytics.

[12] Donald Light. *Building a First-to-Market Pay as You Drive Offering*. Guidewire, 2021. https://explore.guidewire.com/c/report-celent-buildi?x=-QNqMK.

having increased above pre-pandemic levels.[13] Such changes would impact both pricing and reserving and will put a larger focus on monitoring. One way to test for changes to reserves is by using reserving validation.[14]

For commercial underwriters and reinsurers, one of the changes would be in the way insurers are obligated to indemnify in the event of pandemics. *Force majeure* clauses might become ever more important in light of recent events, with business interruption insurance being one such example[15] and event cancellation being another. The latter has featured prominently in the news, in light of insurance policies taken against the possibility of the Olympics not going forward in Tokyo.[16]

Other impacts include general liability insurance in the form of reduced foot traffic, and customers being exposed to increased risk of exposure to the coronavirus.[17] For cyber-insurance, there is increased risk with employees working from home instead of from the secure confines of the office. As such, increased focus will be placed on securing internet connections and avoiding cyber risks such as man-in-the-middle attacks. Ransomware attacks, where access is blocked or data are threatened to be released, jumped 148% in March 2020 from the previous month.[18]

For reinsurers, a greater focus might be put on counterparty arrangements as they seek to adjust their risk profile in light of recent tail risks. Considerations of public–private partnerships might also be worth considering in light of the pandemic in order to build a more proactive and agile response mechanism.[19] Such partnerships have occurred in the past, specifically for terrorism, as seen in the US government's Terrorism Risk Insurance Act.[20]

Rate environment and market flexibility

Macroeconomic changes have affected balance sheets, which has led to increased cost-cutting through automation. Meanwhile, insurers abroad are targeting new offerings and using intuitive platform interfaces to better connect with consumers.

[13] Scott Calvert. Rise in car crash deaths prompts new seat-belt push. *Wall Street Journal*, 2021. www.wsj.com/articles/rise-in-car-crash-deaths-prompts-new-seat-belt-push-11627637400.

[14] William Diffey, Malcolm Cleugh, Laura Hobern and Ed Harrison on behalf of members of the TORP Working Party. *Can You Trust Your Reserving? Reserving Validation Under Covid-19*. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/system/files/field/document/Can%20you%20trust%20your%20reserving%20-%20Reserving%20validation%20under%20Covid-19%20-%20v1.3.pdf.

[15] Ponora Ang, Sébastien Richemont, and Xin Jia Wang. Insurance coverage during a pandemic and force majeure. Fasken, 2020. www.fasken.com/en/knowledge/2020/03/27-covid19-assurance-pandemie-et-force-majeure.

[16] Alicja Grzadkowska. Cancellation of Tokyo Olympics could cripple the insurance industry. *Insurance Business Canada*, 2021. www.insurancebusinessmag.com/ca/news/columns/cancellation-of-tokyo-olympics-could-cripple-the-insurance-industry-246029.aspx.

[17] Brian A. Fannin. *COVID-19: The Property-Casualty Perspective*. Casualty Actuarial Society, 2020. www.casact.org/sites/default/files/2021-03/COVID-19_The_PC_Perspective_3-27-2020.pdf.

[18] VMware Security and Compliance Blog. Amid Covid-19, global orgs see a 148% spike in ransomware attacks; finance industry heavily targeted. 2020. https://blogs.vmware.com/security/2020/04/amid-covid-19-global-orgs-see-a-148-spike-in-ransomware-attacks-finance-industry-heavily-targeted.html.

[19] Angat Sandhu, Steven Chen, Ajit Rochlani, Jun Hao Tay, and Bella Thamrin. *Insurance Redefined: A Roadmap for Insurers and Insurtechs*. Oliver Wyman and Singapore Fintech Association, 2020. www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2020/dec/insurance-redefined.pdf.

[20] Carol Lyons. Does Canada need a terrorism risk insurance scheme? McMillan, 2015. https://mcmillan.ca/insights/does-canada-need-a-terrorism-risk-insurance-scheme-2/.

Insurer underwriting profit has been volatile albeit largely profitable over the past 10 years.[21] Based on the Property and Casualty Insurance Compensation Corporation's quarterly report, the Canadian P&C industry's Q1 2021 performance was more profitable compared to the same period in 2020, with a 27.8% drop in net claim incurred and a dramatic net loss ratio drop from 72% to 50%.[22] The broader Canadian economy also shrank by 5.4% in 2020,[23] and it is not clear how this will impact the insurance premiums in the long term. A third of restaurants, for example, are expected to close as a result of the pandemic,[24] although many have remained open due to changes to a takeout-based operating model.

Source: PACICC based on data from MSA Research[25]

Large changes are occurring with regard to consumer habits, with online shopping becoming more important and consumers being increasingly eco-conscious,[26] with both likely to persist post-pandemic. The consumer's changing preferences are being seen in the insurance industry as well, with company websites for the first time beating brokers in a US customer satisfaction study.[27] Canadian consumers also express significant concern about spending,[28] which could lead to customers being more willing to shop for quotes. As such, it is important for management to investigate how customer acquisition is changing in the face of these evolving trends. The gig economy and new revenue streams are also changing the

[21] Property and Casualty Insurance Compensation Corporation. *Insolvency Protection for Home, Automobile and Business Insurance Customers, Annual Report 2019*. 2020. www.pacicc.ca/wp-content/uploads/2020/04/PACICC-2019-Annual-Report-ENG.pdf.

[22] Property and Casualty Insurance Compensation Corporation. *Solvency Matters: A Quarterly Report on Solvency Issues Affecting P&C Insurers in Canada*. 2021. www.pacicc.ca/wp-content/uploads/2021/06/Solvency_Matters_14_June.pdf.

[23] *Financial Post*. Canadian economy suffers biggest contraction since the Great Depression. 2021. https://financialpost.com/news/economy/canadian-press-newsalert-canadian-economy-contracted-5-4-per-cent-in-2020.

[24] Nicole Brockbank. Fewer Toronto restaurants cancelled business licences in first year of Covid-19 than normal. CBCnews, CBC/Radio Canada, 2021. www.cbc.ca/news/canada/toronto/toronto-business-licences-covid-1.5965568.

[25] *Solvency Matters: A Quarterly Report on Solvency Issues affecting P&C Insurers in Canada*. 2020.

[26] PwC Canada. The retail landscape of the future: Canadian Consumer Insights 2021, Pulse 1. 2021. www.pwc.com/ca/en/industries/consumer-markets/consumer-insights-2021.html.

[27] David Gambrill. How insurance company websites just beat out brokers in a U.S. consumer satisfaction study. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/insurance/why-insurance-company-websites-just-beat-out-brokers-in-a-u-s-consumer-satisfaction-study-1004179552/.

[28] PwC Canada. Understanding the Canadian consumer of the moment: Canadian Consumer Insights 2021, Pulse 2. 2021. www.pwc.com/ca/en/industries/consumer-markets/consumer-insights-2021-pulse-2.html.

market, with two larger insurers[29, 30] targeting the ridesharing industry. Additionally, changes in the types of products and services provided are taking place, as highlighted with PAYD and with Asia showing a different model of what to be offering in terms of smaller products that are more specialized.[31] The Asian market seems focused on rapid innovation by bringing digital products to market quickly and being open to risk-taking.[32]

On the other end, there are the balance sheet components, with insurers having to deal with macroeconomic changes affecting investments. Backbook distress caused by low interest rates poses a challenge for P&C insurers.[33] These developments are driving increased focus on cost-cutting measures, partly through automation in the form of robotic process automation (RPA).

Another avenue companies are looking to is digital innovation in the form of omni-channel experience and intuitive platform interfaces. Integrating systems across channels to provide an omni-channel experience, wherein a customer can start a process online and complete the transaction with an agent, gives the customer a cohesive buying and servicing experience. Established insurers are investing heavily in digitizing their customer journeys by increasing availability and easing navigation of online purchasing, as well as making it easier to integrate sales support.[34] Reaching out to customers through familiar means is also another opportunity to innovate, with WeChat being used to market insurance products in Asia[35] and insurance companies in the United States using chatbots through Facebook Messenger to reach small businesses.[36]

Admittedly, the Asian market is very different, both in the way data flows between banks, tech companies, and insurers, and in the way customers interact with technology, with WeChat being "China's app for everything."[37] The Bancassurance model has not been as successful in Canada as it has been in Europe, Asia, and Australia due to regulatory restrictions,[38] although insurance companies are increasingly partnering with technology firms. Regardless, there are many lessons to be learned in the way digitization and data are affecting other countries.

---

[29] Aviva Canada. Ride sharing insurance quotes. Accessed September 9, 2021. www.aviva.ca/en/find-insurance/add-ons/ride-sharing/.

[30] Greg Meckbach. Why Uber Canada dropped Intact as its insurance provider. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/insurance/why-uber-canada-dropped-intact-as-its-insurance-provider-1004196996/.

[31] Erika Krizsan. Insuretech boosts China's online insurance market. *Insurance Factory*, 2020. https://insurance-factory.eu/insuretech-boosts-chinas-online-insurance-market/.

[32] PwC United States. How can US insurers become more innovative? Look to Asia. 2020. www.pwc.com/us/en/industries/insurance/library/insights-from-asia.html.

[33] Maarssen Mei. *Insurance 2020 & Beyond: The Current Agenda of the CEO*. PwC Netherlands, 2017. www.ag-ai.nl/view/34908-JH+Lobregt+Insurance+2020+%26+beyond.pdf.

[34] Stephan Binder, Ulrike Deetjen, Simon Kaesler, Jörg Mußhoff, and Felix Schollmeier. Moving to a user-first, omnichannel approach. McKinsey and Company, 2021. www.mckinsey.com/industries/financial-services/our-insights/moving-to-a-user-first-omnichannel-approach.

[35] Wanwan Huang. *The Development of WeChat Marketing and Distribution of Insurance Products in China*. Society of Actuaries, 2018. www.soa.org/globalassets/assets/files/resources/research-report/2018/wechat-marketing-distribution.pdf.

[36] Eileen Brown. Next Insurance launches Facebook Messenger chatbot to replace the insurance agent. *Social Business*, 2017. www.zdnet.com/article/next-insurance-launches-facebook-messenger-chatbot-to-replace-the-insurance-agent/.

[37] Tjun Tang, Michelle Hu, and Angelo Candreia. Why Chinese insurers lead the way in digital innovation. Boston Consulting Group, 2018. www.bcg.com/en-ca/publications/2018/chinese-insurers-digital-innovation.

[38] Lyle Adriano. DBRS: "Bancassurance" model fails to take off in North America. *Insurance Business Canada*, 2019. www.insurancebusinessmag.com/ca/news/breaking-news/dbrs-bancassurance-model-fails-to-take-off-in-north-america-166980.aspx.

Changing talent profile

Insurers are now competing with a larger market for data scientists and IT specialists. The actuarial organizations are trying to bridge the skills gap by providing additional offerings to teach actuaries data science skills.

Despite the pandemic's impact and the current market environment, acquiring and retaining qualified talent continues to be a challenge for insurers. Furthermore, the insurer's talent profile will need to be shifted and diversified to adapt to a dynamic insurance market and to enable further innovations. In a 2019 survey on the financial services industry's productivity, 80% of global insurance respondents said they were somewhat or very concerned about the availability of key skills, "putting pressure on costs and impairing organizations' ability to innovate and meet customer expectations."[39] As can be seen from this survey's results, insurers are increasingly targeting data scientists, data engineers, and IT specialists. Changes to the workforce will become more important given the changes to data collection and the resulting larger volumes of data. For instance, data visualization techniques for a dataset containing a million data points are distinctly different from those for one containing 20, as a scatterplot would no longer be as insightful with points superimposed on each other. Similarly, big data architects will be needed to help construct, maintain, and optimize the performance of distributed computer architectures.

Admittedly, these changes are not universal in the insurance industry: reinsurers and commercial insurers naturally have different problems compared to personal insurers. Nonetheless, all insurers are trying to get any additional technological advantages they can get. For example, trucking companies are trying to monitor fatigue among their drivers through wearable devices,[40] which in turn would lead to better evaluation of risk. Unfortunately, competition for technical positions is fierce, with companies having trouble attracting and replacing talent. The challenge with digitization is that insurers are no longer merely competing within the insurance sector for talent, but across the full spectrum of companies requiring tech resources.

For data science in particular, the different actuarial societies are trying to bridge the skills gap between actuaries and data scientists. The CAS started offering the Certified Specialist in Predictive Analytics (CSPA) credential, the IFoA has a certificate in data science, and the Australian Actuaries Institute has added a module on data analytics. During interviews, one of the respondents stated that they were encouraging their employees to enrol in the iCAS program. There was no uniform practice among Canadian P&C insurers, with many having their own internal development programs. Upskilling current employees is an option for insurers, as opposed to trying to attract technologically skilled employees who understand the subtleties of the industry.[41] Dealing with these hiring challenges as long-term structural problems will be critical in dealing with these skill shortages.

[39] *Insurance Trends 2019: Digital Transformation Shifts from Threat to Opportunity*. PwC Global, 2020. www.pwc.com/mu/en/pwc-2019-ceo-survey-insurance-report.pdf.

[40] Julie Weed. Wearable tech that tells drowsy truckers it's time to pull over. *New York Times*, 2020. www.nytimes.com/2020/02/06/business/drowsy-driving-truckers.html.

[41] PwC United States. Top insurance industry issues in 2021: Talent strategies for today's insurers. 2021. www.pwc.com/us/en/industries/insurance/library/top-issues/talent-strategies.html.

InsurTech and mergers

Strategic partnerships between insurers and InsurTechs are improving product offerings. Insurers are also considering mergers to help increase modeling capacity and increase investment in technology.

In addition to transforming the internal talent profile, partnerships with InsurTech and Mergers & Acquisitions (M&A) provide cost-effective opportunities for the insurance industry to keep up with the fast-changing developments in technology. Unlike FinTechs, which have had a significant impact on the way banks interact with technology, most InsurTechs are not directly competing with incumbent insurers, but instead partnering with them, or with several of them.[42] Many of the InsurTech companies are trying to fix particular issues in the insurance field, or to help incumbents move into new fields. As described below, there are a number of areas, from claims and customer service to telematics providers and new data sources, that InsurTech is trying to fix, instead of providing the full-stack solution.

In this context, established insurers must determine how the InsurTech extended ecosystem fits with their strategies. To do this, they have to monitor developments and refine products accordingly, partner with companies to address specific challenges, and consider playing an active role in helping these companies grow through incubators, as some already have.[43]

Some recent global examples of strategic partnerships include:

- An InsurTech and a leading reinsurer partnering to provide weather insurance using science-based analytics and AI[44]
- A leading broker partnering with two InsurTechs[45, 46] to produce innovative real-time integrations for small fleet insurance and to drive digital client solutions, respectively
- A market-leading commercial insurer in the United States partnering with an InsurTech[47] to get new sources of data for small commercial underwriting
- A leading insurer in Japan partnering with an InsurTech to help accelerate accident recovery for auto claims[48]

---

[42] Chris Clague. The perfect time for tech in insurance. The Economist Intelligence Unit, 2020. https://eiuperspectives.economist.com/financial-services/perfect-time-tech-insurance.

[43] DMZ. Aviva and DMZ are working to make Toronto the insurtech capital of Canada. 2019. https://dmz.ryerson.ca/partner_profiles/aviva/.

[44] Jinjer Lorenz. Farmers Edge and Munich Re announce strategic partnership to implement large-scale parametric weather insurance solutions. Farmers Edge, 2020. www.farmersedge.ca/farmers-edge-and-munich-re-announce-strategic-partnership-to-implement-large-scale-parametric-weather-insurance-solutions/.

[45] CarrierHQ. Aon: CarrierHQ's Small Fleet Advantage adjustable rate insurance for trucking wins 2021 Celent Model Insurer Award for data, analytics, and AI. Cision PR Newswire, 2021. www.prnewswire.com/news-releases/aon---carrierhqs-small-fleet-advantage-adjustable-rate-insurance-for-trucking-wins-2021-celent-model-insurer-award-for-data-analytics-and-ai-301250532.html.

[46] Aon. Aon completes acquisition of CoverWallet, the leading digital insurance platform for small and medium-sized businesses. 2020. https://aon.mediaroom.com/2019-01-07-Aon-completes-acquisition-of-CoverWallet-the-leading-digital-insurance-platform-for-small-and-medium-sized-businesses.

[47] Carpe Data. The Hartford taps new data sources for small business underwriting. 2020. https://carpe.io/the-hartford-taps-carpe-data-for-small-business-data/.

[48] Tractable. MS&AD to use Tractable's AI across Japan to accelerate recovery from auto accidents. Cision PR Newswire, 2020. www.prnewswire.com/news-releases/msad-to-use-tractables-ai-across-japan-to-accelerate-recovery-from-auto-accidents-301162486.html.

These examples seek to highlight some of the ways that InsurTech is helping to fill gaps in current insurers' strategies, such as on a product development front, digitizing insurers' interactions, finding new sources of data, or strategic ways to reduce expenses. This is not to say that these are the only ways InsurTech is changing the industry. Much attention is being paid to how one particular AI-driven US insurer progresses as it matures and spreads to other segments and markets. Other companies have also gone public, such as one US online homeowners' insurance company through a special-purpose acquisition company with a valuation of $5 billion[49], and a title insurer with an expected $3 billion[50] valuation.

In Canada, there have been a number of partnerships or strategic investments, including:

- A leading local insurer investing in two InsurTechs to automate the claim process[51] and offer photo estimate claim services[52]
- A commercial insurer partnering with a Dutch company to create its own fully digital insurance proposition[53]
- A branch of a world-leading insurer funding an InsurTech that is offering the largest selection of online insurance in Canada[54]

As can be seen, insurers are trying to keep in step with changing consumer preferences, such as increasing digital interactions and a changing economy, with new developments such as ridesharing changing the demands regarding what type of coverage to offer. With so much retail choice, those insurance providers that fail to respond to the changes that the back-to-the-future model are creating will fall away over time.[55] Having seen astronomical investment in Q1 of 2021, the same consultancy group stated that it expects record-breaking investments to continue in Q2.

During a 2020 panel, the CEO of a leading insurer in Québec talked about a recent merger, stating: "We both in our strategic plans came to the conclusion that size matters more and more going forward, as we needed to invest in technology and digitization."[56] When looking at the survey results, we also found differences between small and larger insurers, particularly in the way larger insurers were able to deploy resources towards advanced analytics. While having more resources allocated to a team proportional to your company size indicates investment in analytics, sheer size also seemed to be a larger determining factor. The reason should be clear, as having 100 people dedicated to advanced analytics should still yield

[49] Sohini Podder. Home insurance agency Hippo to go public in $5b SPAC, *Insurance Journal*, 2021. www.insurancejournal.com/news/national/2021/03/04/603732.htm.

[50] Mary Ann Azevedo and Alex Wilhelm. Proptech startup States Title, now Doma, going public via SPAC in $3b deal, *TechCrunch*, 2021. https://techcrunch.com/2021/03/02/proptech-startup-states-title-now-doma-going-public-via-spac-in-3b-deal/.

[51] Paul Sawers. Pay-per-mile car insurance company Metromile raises $90 million to automate the claims process. *VentureBeat*, 2018. https://venturebeat.com/2018/07/24/pay-per-mile-car-insurance-company-metromile-raises-90-million-to-automate-the-claims-process/.

[52] Insurance-Canada.ca. Intact partners with Snapsheet to offer photo claims estimating service. 2019. www.insurance-canada.ca/2019/06/21/intact-snapsheet-photo-claims-estimating/.

[53] Zeist. Achmea launches Canadian online insurance proposition in partnership with Fairfax. Achmea, 2018. https://news.achmea.nl/achmea-launches-canadian-online-insurance-proposition-in-partnership-with-fairfax/.

[54] Apollo Insurance Solutions. Canadian insurtech Apollo closes $13.5 million Series A financing round. Cision Newswire, 2021. www.newswire.ca/news-releases/canadian-insurtech-apollo-closes-13-5-million-series-a-financing-round-855112967.html.

[55] Willis Towers Watson. Quarterly Insurtech Briefing Q2 2021. 2021. https://www.datocms-assets.com/24091/1627554491-wtw-quarterly-insurtech-briefing-q2-20212.pdf.

[56] Greg Meckbach. Who is funding many of these insurance mergers. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/mergers-and-aqcuisitions/who-is-funding-many-of-these-insurance-mergers-1004200880/.

more research and insights into your data than 10 people could. A recent survey in the United States found similar results[57] with the top 20 carriers having ML and AI adoption rates 20% higher than the top 21–50.

A recent article discussed some of the trends regarding advantages of scale in the P&C industry.[58] While not as obvious as in life insurance, it noted that in non-motor retail larger players on average achieved 40% lower cost ratios in claims handling, with better results for both personal and commercial insurers. It pointed to IT costs and digital innovation as points where scale plays a large role in cutting costs.

Data

Historically, there have been few touchpoints between the insurer and the customer, with most of these interactions occurring during acquisition, renewal, or customer claims. Recently, there has been an explosion in data available to insurers in the proliferation of telematics data and other sensors, satellite images, weather data, and others. As seen in the survey results, insurers are upgrading their data assets to deal with these new streams of data, and to help consolidate their sprawling legacy systems. Unfortunately, this transformation is both challenging and time-consuming, as described to us in our interviews, requiring either the parallel running of both the new and old systems or the use of a phased implementation approach. Those insurers who have already dealt with migrating data from their legacy systems, or are in the process of transitioning their accounting system to International Financial Reporting Standard (IFRS) 17, will know full well the challenges present in the transformation process. Transformation to the new financial reporting standard will provide insurers with the perfect opportunity to upgrade their legacy systems.

Challenges also arise out of the changing work environment. When dealing with large amounts of data, transfers over even the speediest of internet connections are long. While insurance companies do not yet have data that are too large compared to other industries, such as banking, there are limitations to internet transfer speeds. Instead of uploading data directly to the cloud, Amazon has helped customers save time by physically moving data centres using trucks.[59] To circumvent having to do data transfers, many companies are setting up virtual machines or looking to take advantage of cloud storage and computing to deal with their data needs. Privacy issues naturally arise, with companies needing to secure their connection and restrict data privileges. Such processes are part of good governance practices and are further discussed below.

Furthermore, dealing with big data requires using data structures with increasing sophistication. As relayed in our interviews, some companies are moving away from more traditional data warehouses and transitioning to data lakes.
Several questions will have to be answered regarding what database technology to use and whether to transition to NoSQL from a traditional relational database. A non-exhaustive list of such questions include:

[57] LexisNexis Risk Solutions. 2019 study results: How U.S. insurance carriers are using artificial intelligence and machine learning. 2021. https://risk.lexisnexis.com/insights-resources/research/state-of-ai-ml-in-the-insurance-industry.

[58] Nagendra Bommadevara, Björn Münstermann, Sanaya Nagpal, and Ulrike Vogelgesang. Scale matters … to an extent: Playing the scale game in insurance. McKinsey & Company, 2021. www.mckinsey.com/industries/financial-services/our-insights/scale-matters-to-an-extent-playing-the-scale-game-in-insurance.

[59] *Wired*. Amazon's Snowmobile is actually a truck hauling a huge hard drive. 2016. www.wired.com/2016/12/amazons-snowmobile-actually-truck-hauling-huge-hard-drive/.

- What type of data is intended to be stored?
- How frequently are the business requirements changing?
- Are there key-value stores, which could include those designed for managing document-oriented information?
- Will the data need to be compatible with SQL applications?
- Will the data be more READ-heavy or will there be more WRITE transactions in the form of inserts or updates?
- What granularity would the database need to enable user-based security (e.g., at the attribute, row, or table level)?
- Do you need the technology to enable integrations with analytics and visualization tools that are "What you see is what you get" in nature, such as Tableau or Qlik?
- Is there a plan to employ emerging architecture techniques such as Microservices?

One's responses to the above questions will help to determine whether there is a need for a NoSQL database and what type to consider. For instance, key-value stores like Apache Cassandra are effective for fast read and write operations when able to query by a key; however, it would be less optimal for data that could be highly compressed, where Apache HBase or BigTable would be preferable. Some offerings, such as Snowflake, Amazon Redshift, and Google BigQuery, are hybrid offerings, having the advantages of being compatible with many SQL applications while providing the scalability of NoSQL. That is, there are numerous considerations when deciding on a big data solution, tailored to one's specific needs.

Migration to the cloud is a decision some large Canadian P&C carriers are making, as relayed in interviews and survey responses. On-premises systems can be scaled to deliver such results, although they would normally require high capital and operational expenditure.[60] It would also necessitate an insurer investing a considerable effort to implement the analytical platform and derive value from it. Should an insurance provider opt to use a cloud solution, cloud training and certification will be important for its business and technical resource team.[61] Most cloud solution providers insist clients become cloud trained and certified before the implementation phase and create a centre of excellence (CoE) within the organization. Care also must be taken not to concentrate too much of the infrastructure, lest one become too reliant on a single provider, resulting in a loss of control in one's cloud spend, and become forced into vendor lock-in.

Utilizing unstructured data

Having databases that can collect unstructured data will be important in making better decisions. Insurers are already making use of this type of data in combination with ML to analyze images.

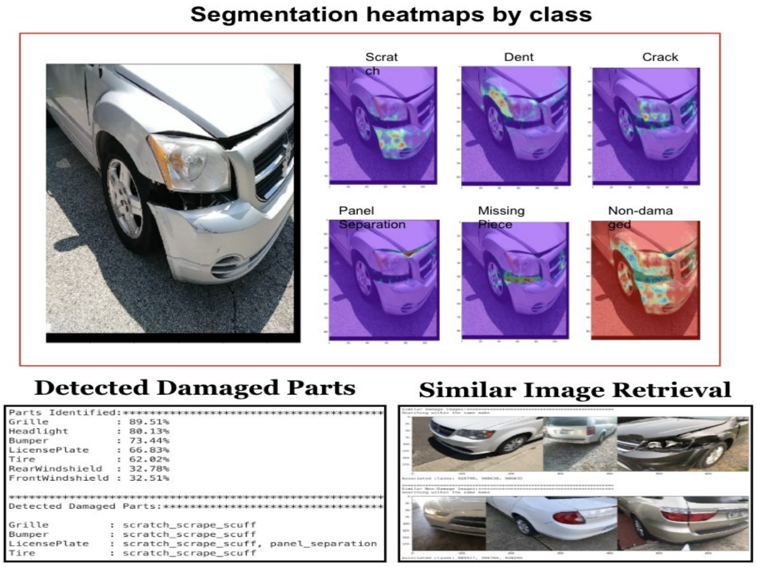
Traditionally, the data that insurance companies collected were easily put into data warehouses and cleaned for use in underwriting and claims. Changes to the nature of the data insurers want to analyze dictate changes to the structure of the data. For instance, developments in computer vision algorithms, wherein a computer tries to pick up information from videos and images, is changing the way insurers are dealing with claims and claims-processing. By having a photo feature in the insurer's app, a customer could take a picture of the damage to a car and immediately have it appraised. That is, an algorithm could

---

[60] Sloan Plumber and Scott Busse. How insurance carriers are modernizing to cloud based analytics. LinkedIn (article), PwC Advisory United States, 2020. www.linkedin.com/pulse/how-insurance-carriers-modernizing-cloud-based-analytics-sloan-plumer/.

[61] Sloan Plumer, Imran Ilyas, Josh Knipp, Yasir Safdar, and Tirath Desai. Eight considerations for insurance carriers migrating to the cloud. LinkedIn (article), PwC Advisory United States, 2020. www.linkedin.com/pulse/eight-considerations-insurance-carriers-migrating-cloud-sloan-plumer/?trackingId=4ER%2BMP1wTuyXxfsyNcIq5Q%3D%3D.

evaluate whether the car is indeed damaged and try to estimate how much is owed based on the parts and labour.



PwC artificial intelligence accelerator case study

The technology has matured considerably as demonstrated by an AI company already having been used in Europe and Asia to settle more than $1 billion in claims.[62] This technology helps generate quick and accurate repair estimates at first notice of loss (FNOL) as well as helping to flag any unnecessary repairs, thus preventing claims leakage.

To deal with image data for ML purposes, insurers might face challenges in upgrading their data infrastructure to be able to store images. To do so, the database has to support images being stored as a binary large object (BLOB).[63] However, backing up the image database would be time-consuming. One alternative would be to store the file location of all the images, although care must be taken as moving images to a different location would cause broken links for the end-users. Another possibility is to use a cloud storage service like Amazon S3 or the Hadoop framework, where the images are stored as objects.[64] Insurers will be exposed to the same data storage challenges with any other form of unstructured data. As such, insurers might have to change their data infrastructure to account for the new streams of data.

Size of the data and the Internet of Things

New streams of information present challenges in storing data that scale at a faster pace than existing sources. Being selective about how to store the data is necessary as the additional size also risks excessively burdening IT with information that will not be used.

[62] Mark Gardiner. A car insurance claim estimate before the tow truck is called. *New York Times*, 2020. www.nytimes.com/2020/09/17/business/car-insurance-claim-estimate-artificial-intelligence.html.

[63] Pearson. Storing and retrieving images in JDBC. InformIT, 2002. www.informit.com/articles/article.aspx?p=25280.

[64] AWS Amazon. Amazon S3: Object storage built to store and retrieve any amount of data from anywhere. Accessed September 10, 2021. https://aws.amazon.com/s3/.

One of the new developments of the technology age has been the use of physical objects that are embedded with sensors and software that connect to the internet: the Internet of Things (IoT). These devices provide new possibilities to prevent losses and detect fraud earlier by identifying whether the home is empty using smart locks, spotting if there are any leaky pipes,[65] distinguishing fatigued drivers through telematics, and using smart security cameras. To deal with the streams of data coming from these devices, insurers have to change the way they deal with data, as records of noteworthy events will have to be stored before being analyzed. This poses a challenge to insurers' IT systems and has serious financial costs.

For instance, if dashcams come to larger prominence as a way for fleets to lower their deductibles by proving fault for accidents,[66] or connected homes continue to gain in popularity to help alert homeowners to leaks sooner, then either the supplier of the product or the insurer will have to store the data.

In addition to challenges managing different data types, unstructured data require much more storage space as well. For instance, to store text on Microsoft SQL Server, one might use varchar to store the text, with each character using a single byte, up to a maximum of 8,000 bytes.[67] Contrast this with a picture from a phone, which might take up 2 megabytes (2,000,000 bytes), which is significantly larger. Admittedly, you will be able to compress the image, but it is still orders of magnitude bigger.

There is a risk of data becoming dark data;[68] that is, data that are collected by insurers but never used. An example given is all the photos that people take on their phones and are not viewed again. The risk is that this will simply be a burden to your IT instead of a benefit to an insurer's data strategy. Making sure that data are used and having a strategy on when to purge data are important for keeping costs at bay.

Data access, governance, and privacy threats

All of these additional data come with their own risks, with cybersecurity risks becoming of greater concern with larger amounts of data. Much in the same way that virtual private networks (VPNs) have to be secured to make sure ill-intentioned actors are not accessing your data, making sure that data are properly siloed is also important in light of high-profile data breaches.

There are also risks associated with the sensors themselves. While some of these exploits have been rather innocuous, like hacking some smart light bulbs,[69] hacking a person's smart lock on their home

[65] Nadine Evans. IoT and the reduction of claims – *Canadian Underwriter*. Eddy Solutions, 2019. https://eddysolutions.com/iot-and-the-reduction-of-claim-canadian-underwriter/.

[66] Hilary Daninhirsch. Some fleets turn to cameras to help mitigate rising insurance costs. *Transport Topics*, 2020. www.ttnews.com/articles/some-fleets-turn-cameras-help-mitigate-rising-insurance-costs.

[67] MikeRayMSFT, julieMSFT, cawrites, icoric, mindlessroman, markingmyname, CarlRabeler, pmasl, PRMerger16, MashaMSFT, WilliamAntonRohm, craigg-msft and edmacauley. Char and varchar (Transact-SQL). SQL Server/Microsoft Docs, 2019. https://docs.microsoft.com/en-us/sql/t-sql/data-types/char-and-varchar-transact-sql?view=sql-server-ver15.

[68] Tom Taulli. What you need to know about dark data. *Forbes*, 2019. www.forbes.com/sites/tomtaulli/2019/10/27/what-you-need-to-know-about-dark-data/?sh=770d4bdf2c79.

[69] Thomas Ricker. Watch a drone hack a room full of smart lightbulbs from outside the window. *The Verge*, 2016. www.theverge.com/2016/11/3/13507126/iot-drone-hack.

would be significantly more worrisome.[70] Weighing the benefits of the IoT must be balanced against the possible risks, both to personal insurers and commercial ones.[71]

Technology

Once the data component and data storage strategy has been ironed out, attention should be placed on how insurers are manipulating their data. This is both in the tools a company uses to model its processes and also in how it conducts its business.

Hardware changes

Hardware is being specifically built to speed up specific algorithms. Pinpointing what hardware to adopt should be done in conjunction with software and algorithm choices when looking for speed improvements.

Companies are experimenting with different hardware to help speed up processes as certain algorithms work more effectively on particular hardware. For instance, Google has made chips for its data centres specifically to run its ML algorithms.[72] In our interviews, we learned that companies are trying to leverage different technologies, from the parallelization of clustered data processing on Hadoop to GPUs being used to help in deep learning. Bandwidth is one of the main reasons why GPUs are faster than CPUs for this purpose, as GPUs have dedicated Video RAM (VRAM) memory, enabling the CPU's memory to be used for other tasks.[73] By switching from a CPU to a GPU, you could achieve a 100x decrease in training time,[74] thereby freeing up resources and speeding up your ability to model.

Of course, to take advantage of these hardware advantages, you will need to use the appropriate software that runs on this hardware. For instance, to take advantage of a GPU, you might use PyTorch to access CUDA for Nvidia graphics cards.

There are many different implementations for big data storage; however, in light of the several companies adopting the Hadoop framework, we briefly describe the different components below. Hadoop is an open-source project managed by the Apache Foundation with companies as large as Twitter using it as their core data platform.[75] It has four main modules, comprising:

- Hadoop Common, which contains the libraries and utilities needed by other Hadoop modules
- Hadoop Distributed File System (HDFS), where data are stored on multiple machines
- Hadoop MapReduce data processing to take advantage of parallelization

[70] Zack Whittaker. Security flaws in a popular smart home hub let hackers unlock front doors. *TechCrunch*, 2019. https://techcrunch.com/2019/07/02/smart-home-hub-flaws-unlock-doors/.

[71] *Insurance Journal*. Internet of Things devices increase risk of cyber attacks on industrial sector: Lloyd's. 2021. www.insurancejournal.com/news/international/2021/02/17/601582.htm.

[72] Google Cloud. Cloud Tensor Processing Units (TPUs). Accessed September 10, 2021. https://cloud.google.com/tpu/docs/tpus.

[73] Jason Dsouza. What is a GPU and do you need one in deep learning? *Towards Data Science*, 2020. https://towardsdatascience.com/what-is-a-gpu-and-do-you-need-one-in-deep-learning-718b9597aa0d.

[74] Dario Radečić. PyTorch: Switching to the GPU. *Towards Data Science*, 2020. https://towardsdatascience.com/pytorch-switching-to-the-gpu-a7c0b21e8a99.

[75] Gerashe Galov. Hadoop filesystem at Twitter. Twitter (blog), 2015. https://blog.twitter.com/engineering/en_us/a/2015/hadoop-filesystem-at-twitter.

- Hadoop YARN resource manager to orchestrate the management of the parallel processes

Beyond HDFS, YARN, and MapReduce, there are a number of other common modules and related projects, such as Oozie and Apache Pig.[76]

End-users are able to query the tables using Apache Hive or Apache Impala before bringing the smaller dataset onto a node for analysis and visualization. The technical tasks of dealing with MapReduce are abstracted away, with users only needing to query using SQL, which would be familiar to many actuaries and analysts. In combination with the HDFS, some companies are turning to Apache Spark to help speed up their analyses. Such implementations could be done on local machines or in the cloud, with many cloud providers giving the option to use Hadoop on their servers.

Embracing open source

Using open-source software will enable experimentation with the most up-to-date packages for modeling and data wrangling. Existing platforms are giving users the ability to supplement pre-built methods with custom Python and R code.

One of the major shifts we have seen from the survey results is the way in which companies are increasingly shifting away from proprietary analytics software to performing tasks using R and Python, both open-source programming languages. These two languages are the most commonly used[77] due to the number of excellent packages available. Other programming languages used for advanced analytics, such as Julia and Scala, were mentioned in interviews and in survey responses, although they were used with far less frequency in the insurance industry.

Proprietary software even allows you to code directly in Python and R, as can be seen in Power BI, SAS, Alteryx, and Amazon SageMaker. Low- or no-code software allows insurers with a shortage of programming talent to perform data manipulation and analysis work.

Part of the reason for the success of these open-source languages are the libraries available that help get analytics projects started easily,[78] as well as access to the newest algorithms, with researchers increasingly posting code with their papers.[79] Each has its advantages, where R is more known for its visualization capabilities, with ggplot and Shiny packages, and Python is better for software development, having been originally created as a fully fledged programming language. Nonetheless, steps have been taken by both languages to reconcile these issues, with R also having the ability to operate using the object-oriented programming paradigm and Python able to interact with D3.js to make aesthetically pleasing visualizations. Popular packages are also being written to be used with both programming languages, such as Plotly for visualization, XGBoost for modeling, and Spark for working on clusters. The fact that these languages are popular outside of the insurance industry, as can be seen in the yearly Stack

[76] Sachin P. Bappalige. An introduction to Apache Hadoop for big data. Opensource.com, 2014. https://opensource.com/life/14/8/intro-apache-hadoop-big-data.

[77] EdX team. 9 top programming languages for data science. EdX Blog, 2021. https://blog.edx.org/9-top-programming-languages-for-data-science.

[78] Burak Karakan. Python vs R for data science. *Towards Data Science*, 2020. https://towardsdatascience.com/python-vs-r-for-data-science-6a83e4541000.

[79] Papers with Code. The latest in machine learning. Accessed September 10, 2021. https://paperswithcode.com/.

Overflow survey,[80] makes it easier to hire talent from other industries with strong modeling ability but with less understanding of the subtleties of the insurance industry.

Many companies still use SAS to store and transform their data using SAS SQL, with analysis done in one of the programming languages mentioned above. Also, as mentioned in our section on data, Hadoop allows one to code using SQL-like queries in Apache Hive or Apache Impala. Thus, regardless of which tool is used for the analytics and visualization processes, SQL skills will likely still remain important for most analysis.

One of the issues when using these software tools is the copyright licence. Depending on whether it is copyleft or permissive, the licence can have massive implications, as it could cause one's software to become open-sourced. In addition, care must be taken with code taken from the internet that could infect your infrastructure. Finally, libraries and programming language versions change, which could cause one's code to not work on coworkers' machines should the functionality change. To deal with the latter issue, choosing a strategy to enable the project to be portable, as well as using version control, will be very useful.[81]

Blockchain

Companies abroad are testing out blockchain to help store a customer's relevant policy information and minimize expenses. The use of smart contracts could alter the way insurers interact with customers and reinsurers.

Blockchain is a subset of distributed ledger technologies (DLTs), using "blocks" of information to keep track of data transactions in a distributed network of multiple nodes or computers.[82] The use of blockchain in the insurance industry is beginning to gain steam. Reinsurers stand to gain significant benefits through the use of smart contracts, where risks are ceded or retroceded using a blockchain application specifically designed to process treaties. One reason reinsurers might consider blockchain is because reinsurance expense ratios are typically 5–10% of premiums.[83] The potential of smart contracts lies in their ability to streamline data processing, allow entry into new markets or products, and allow significant transparency in the process with all documents stored on a reinsurance blockchain. The use of blockchain for such purposes has already begun, with Allianz having piloted the use of blockchain smart contract technology for transacting a natural catastrophe swap.[84]

The use of blockchain is not limited to reinsurers, however, with companies abroad thinking about adding Bitcoin, the most high-profile application of blockchain, to their balance sheets. Recently, several P&C insurers have invested in a Bitcoin investment management company, a provider of technology and

[80] Stack Overflow. Stack Overflow Developer Survey 2021. 2021. https://insights.stackoverflow.com/survey/2021#technology.

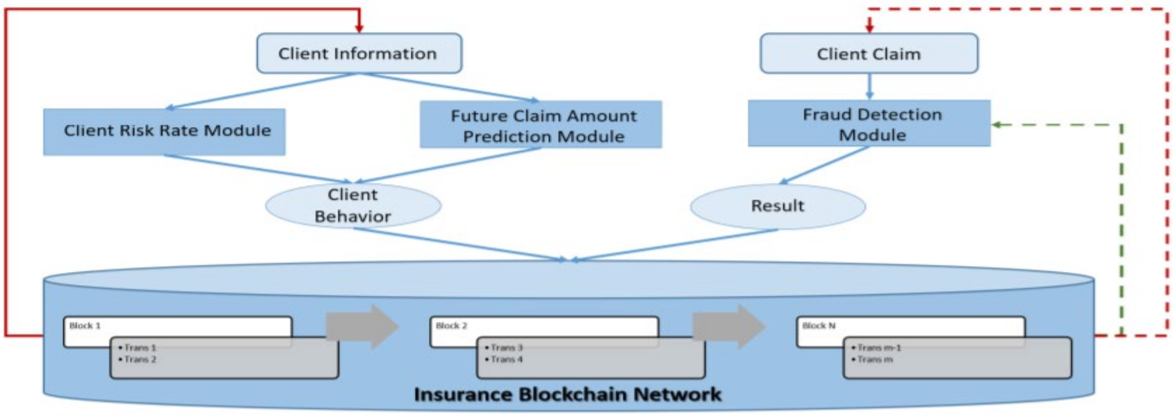
[81] Environments. Reproducible environments. Accessed September 10, 2021. https://environments.rstudio.com/.

[82] Ramona Delcea. Discussion paper on blockchain and smart contracts in insurance: EIOPA invites comments. European Insurance and Occupational Pensions Authority, 2021. www.eiopa.europa.eu/content/discussion-paper-blockchain-and-smart-contracts-insurance-eiopa-invites-comments-0_en.

[83] PwC Legal Estonia. *Blockchain: The $5 Billion Opportunity for Reinsurers*. 2016. www.pwc.com/ee/et/publications/pub/blockchain-for-reinsurers.pdf.

[84] Allianz. Allianz: Blockchain technology successfully piloted by Allianz Risk Transfer and Nephila for catastrophe swap. 2016. www.allianz.com/en/press/news/commitment/sponsorship/160615-blockchain-technology-successfully-piloted.html.

investment solutions for Bitcoin in the United States[85] with ambitions of providing Bitcoin-denominated products to insurers. Blockchain could also find use in storing the ancillary contract documents, shared between the broker and underwriter, with the blockchain also viewable by regulators, tax authorities, and other participants to help simplify reporting and checking processes.[86] Real-world examples of the use of blockchain can already be found, with Chinese media outlets recently reporting that the insurance IT arm of one of the four largest commercial banks in China has launched an insurance blockchain. After their partnership with a world-leading tech company, multiple partners are already onboard in storing electronic policies on the blockchain, and others are currently testing the platform.[87]



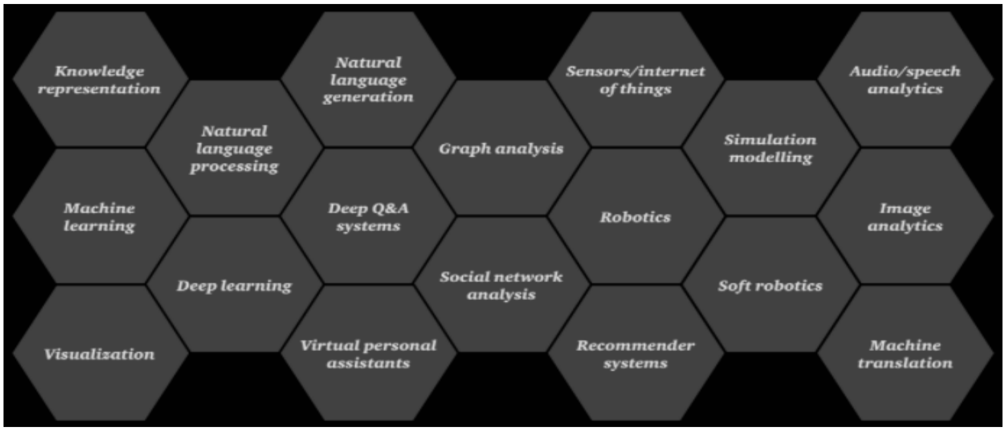
N. Dhieb et al. *A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement*[88]

[85] Gabriella Ben-Hutta. NYDIG raises $100 million. *CoverageR*, 2021. https://coverager.com/nydig-raises-100-million/.

[86] Michael Mainelli and Bernard Manson. *Chain Reaction: How Blockchain Technology Might Transform Wholesale Insurance*. PwC Global, 2016. www.pwc.com/gx/en/financial-services/pdf/how-blockchain-tecnology-might-transform-insurance.pdf.

[87] Miranda Wood. Bank of China officially launches insurance blockchain. *Ledger Insights Enterprise Blockchain News*, 2019. www.ledgerinsights.com/bank-of-china-insurance-blockchain/.

[88] Najmeddine Dhieb et al. A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. *IEEE Access*, 2020. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9046765.



Rao et al. *AI in Insurance: Hype or Reality? The Digital Insurer* [89]

Model and business usage

Companies are increasingly trying to leverage the increasing amount of information available to them and the flurry of new techniques developed to deal with this newfound abundance of data.

These companies are using AI to automate tasks, such as fraud detection, vetting resumés, and loan applications, thereby providing additional resources to perform higher-level work. Doctors are turning to AI to help in reading medical imaging and analyzing electrical impulses from the heart. Chatbots are being used in place of customer service representatives to help customers address simple questions.

As George Box stated: "All models are wrong, but some are useful." Looking to other industries, developments in algorithms have had profound impacts. There have been clear successes in games, with AlphaGo and AlphaZero having significant success against humans in Go, and chess computers significantly changed by the new models.[90]

In "Statistical modeling: The two cultures," Leo Breiman[91] makes the distinction between the two competing goals of explainability and predictiveness, and the practitioners who seek those two different purposes, as follows:

> One assumes that the data are generated by a given stochastic data model. The other uses algorithmic models and treats the data mechanism as unknown. The statistical community has been committed to the almost exclusive use of data models … If our goal as a field is to use data to solve problems, then we need to move away from exclusive dependence on data models and adopt a more diverse set of tools.

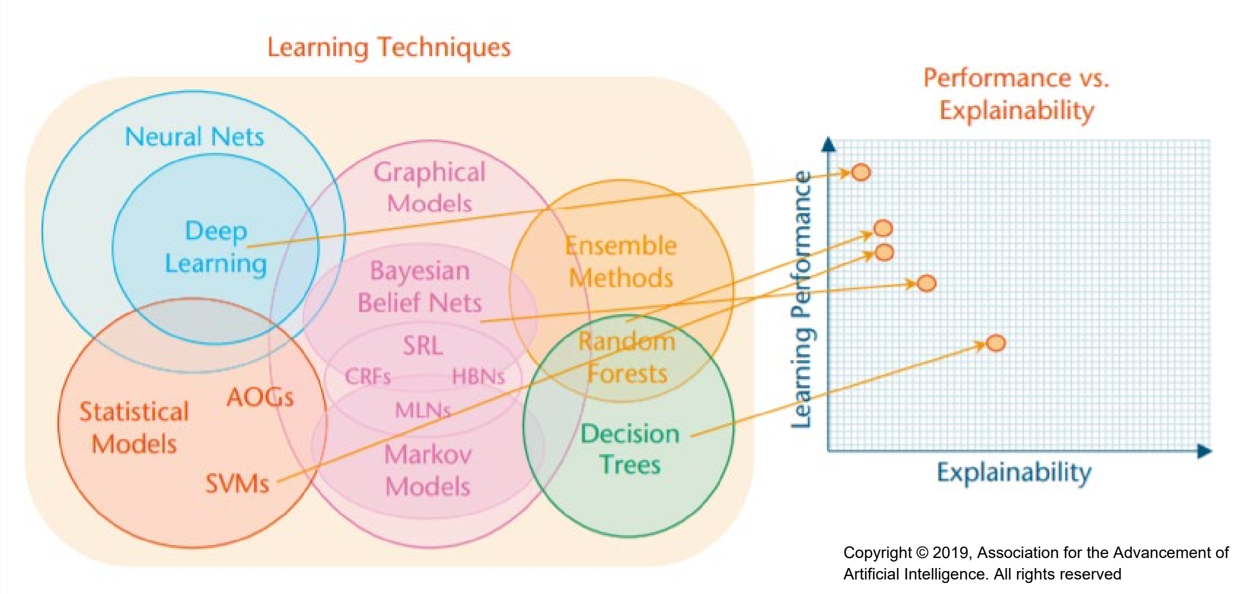
[89] Anand Rao, Jamie Yoder, and Scott Busse. *AI in Insurance: Hype or Reality? The Digital Insurer*. PwC, 2016. www.the-digital-insurer.com/wp-content/uploads/2016/06/716-pwc-top-issues-artificial-intelligence.pdf.

[90] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, and Dharshan Kumaran. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 2018. https://science.sciencemag.org/content/362/6419/1140.

[91] Leo Breiman. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statist. Sci., 2001. https://projecteuclid.org/journals/statistical-science/volume-16/issue-3/Statistical-Modeling--The-Two-Cultures-with-comments-and-a/10.1214/ss/1009213726.full.

These two competing goals still exist 20 years later, with insurers having the need to properly price and segment their data coming in conflict with their ability to explain to regulators and customers the reasons they are being priced as such. We will come back to these ideas later regarding responsible and explainable AI.

There is a large array of techniques currently being applied to model insurance problems, each having different trade-offs between explainability and predictive accuracy.



The Defense Advanced Research Projects Agency's Explainable Artificial Intelligence (XAI) Program[92]

In our interviews, we found that teams were using the full gamut of methods from tree-based techniques to neural networks (NNs) depending on the use case. Depending on the particular use case, the techniques used could vary widely even among the same insurer.

Modeling techniques

Recognize that while GLMs are still common, understand that ML and NNs have a distinct advantage in that they can be used to capture non-linear effects. There is a significant amount of active academic research into these new techniques as applied to insurance.

As demonstrated in the above diagram, there are many different models that can be used to solve problems. The following discussion will in no way be comprehensive, and will instead offer a brief overview of some current techniques. For a more complete survey and technical discussion of techniques, consult

[92] David Gunning and David Aha. DARPA's explainable artificial intelligence (XAI) program. *AI Magazine*, 2019. https://doi.org/10.1609/aimag.v40i2.2850.

The Elements of Statistical Learning.[93] For a survey of AI methods in actuarial science, see the 2020 Hachemeister prize-winning paper.[94]

We begin with a distinction between supervised and unsupervised learning. In supervised learning, each input has an associated response, and the goal is to identify the function that will best model this input/response pair based on some loss function and some restriction on model complexity. Often the loss function is chosen to be the squared error loss, or L^2 norm, and the quality of the predictor will be estimated by comparing it with the observed results, subject to this loss function. Unsupervised learning, on the other hand, has no response, and hence such problems are of the clustering variety. That is, the task of unsupervised learning is to find meaningful patterns from the observations, which can then be used to further understand the data or, in some cases, model it.[95]

In classical ordinary least squares regression, we have n observations, and corresponding predictors. The goal is to find linear relationships between the observations and predictors, subject to some assumptions on the noise. It turns out that this has an elegant solution under the L^2 norm.

For GLMs, the equation is similar except there is a link function. This link function modifies the equation slightly such that a function of the expectation is linearly related to the observations, where this link function is subject to specific criteria and the observations are assumed to come from the exponential family of distributions. The so-called "canonical link" will often be chosen due to some desirable properties, although for Gamma and exponential distributions it is rarely used due to restrictions on the range of expected responses. Often, when dealing with auto insurance, the frequency of accidents will be assumed to follow a Poisson distribution, and the severity will follow a Gamma distribution. When trying to model loss cost without separating the two, a Tweedie distribution will be chosen.[96] The canonical link for the Poisson model is the log-link, and the maximum likelihood estimator is solved using an iterative weighted least squares procedure. An overdispersed Poisson distribution will be used to account for variance different to the mean. Using the log-link is attractive because it leads to the property of having a multiplicative rating structure. For classification problems involving binary variables, logistic regression will be chosen, where observations are Binomial and the link function is the logit function. Interaction terms can be added as the product of two variables, as can higher-order polynomials to try and capture non-linear effects. Further extensions exist in the form of mixed models, generalized additive models (GAMs), and splines. For fat-tailed data, the restriction to exponential families will have to be withdrawn so as to use things like the Pareto distribution.

We now come to more recent developments. Regularization is where another term is added to the regression problem. That is, one adds a term to penalize one's chosen prediction function, with a parameter chosen prior to solving the equation that controls the amount of penalty to levy.[97] Unlike the

[93] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction,* 2nd edition. Stanford Web Spring, 2017. https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12_toc.pdf.

[94] Casualty Actuarial Society. Charles A. Hachemeister Prize. Accessed September 10, 2021. www.casact.org/about/awards-prizes-scholarships/charles-hachemeister-prize.

[95] Ronald Richman. *AI in Actuarial Science*. Presented at the Actuarial Society of South Africa's 2018 Convention, October 24–25, Cape Town. www.actuarialsociety.org.za/wp-content/uploads/2018/10/2018-Richman-FIN.pdf.

[96] Mark Goldburd, Anand Khare, Dan Tevet, and Dmitriy Guller. *Generalized Linear Models for Insurance Rating*, 2nd edition. Casualty Actuarial Society, 2020. www.casact.org/sites/default/files/2021-03/8_GLM.pdf.

[97] Hastie et al. *The Elements of Statistical Learning*, page 168.

previous two methods, however, such methods are not *scale invariant*,[98] meaning there will be different results depending on the units of a variable (e.g., months vs years). As a result, you will likely want to standardize the predictors prior to performing the optimization procedure. For linear regression, the point is to penalize using one of the L^1 norm resulting in LASSO, the L^2 norm resulting in Ridge Regression, or a combination of the two resulting in Elastic Nets. One of the advantages of regularization is that it helps in whittling down the number of predictors for a given problem to a more manageable number.[99] In LASSO specifically, some of the predictors become unused due to the geometry of the optimization problem, and hence that method has variable selection built-in.

Tree-based methods have also grown in prominence, with random forests and boosted trees being especially popular. The idea for decision trees is to create a partition on the feature space. Such trees can be developed by hand, although automating the tree-creation process can be made using greedy algorithms such as by using Classification and Regression Trees (CART), the leading approach. Multiple trees will be created and then combined together in such a way as to improve prediction accuracy. Random forests use bootstrap aggregation, or bagging, to generate predictions, with a modification to decorrelate the trees. The bagging procedure takes bootstrap samples from the training set, generates trees, and averages the predictions. Performance is affected by the number of trees used and the depth of the tree. GBMs work in a similar way, except they work sequentially with each tree grown using information from previously grown trees. The model is updated using a shrunken version of the new tree. The gradient of the loss function is used for the gradient descent algorithm to solve the optimization problem. Other boosting algorithms exist, such as XGBoost and AdaBoost.

In classification problems, Support Vector Machines (SVMs) formalize the idea of drawing a line between the data points in each category. Later variants were created that could handle regression problems. The issue is that it is not often easy to separate things linearly. The magic of SVMs is to use the so-called "kernel trick," to perform all this analysis relatively efficiently, for a range of unfolding functions ("kernels").

Deep learning is a way to use NNs on large datasets using GPUs.[100] The critical idea was to use backpropagation to train NNs. The results have been impressive, such as in the identification of numbers from a collection of handwritten digits. An NN is a multi-stage regression or classification model typically represented by a network diagram. Each neuron takes inputs, which may be outputs from previous layers of neurons, and returns a weighted sum of these that is scaled using an activation function. The backpropagation algorithm is used to calculate the appropriate weights so as to minimize the error. The activation functions will be non-linear – otherwise the model would collapse into a simple linear model – and this allows the model to capture complex non-linearities and interaction effects. Some examples of activation functions include the sigmoid, rectified linear unit (ReLU), and hyperbolic tangent. Many varieties of NNs exist, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

[98] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning with Applications in R*, 2nd edition. Stanford Web, 2021. www.statlearning.com/.

[99] Alex Labram. The machine learning landscape. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/news-and-insights/news/machine-learning-landscape.

[100] Richman, *AI in Actuarial Science*.

These newer algorithms are finding increasing use in the insurance industry. In the survey, GBMs were an especially popular technique, and have been very effective in competitions since their inception.[101]

Of interest to insurers is that better segmentation for pricing might not lead to increased profit or increased market share due to complex real-world factors such as the underwriting cycle which are difficult to model and predict. There have been a couple of iterations of an insurance-pricing game,[102] created by some academic researchers in collaboration with the actuarial societies, to try and test how different algorithms work in a simulated market. That is, they run an iterative game to see how market share and loss ratios change depending on open submissions from participants. In previous iterations of the game, one interesting result was that simpler GLMs might produce worse loss ratios but have higher market share compared to GBMs,[103] and some participants from the most recent iteration remarked that certain models tended to do better on different segments.[104]

Responsible AI

It is critical to ensure that new methods are trustworthy, fair, unbiased, and stable. The use of advanced analytics necessitates a renewed focus on the ethical considerations a modeler must consider when testing a model.

When it comes to implementing AI, the first thoughts that come to mind are usually these: what are the data, what algorithms are we going to use, and are the results predictive? However, there are other things that should be considered in light of high-profile issues. For instance, there have been some high-profile instances of real-world AI going horribly wrong, such as one tech company's experimental chatbot tweeting highly offensive material.[105]

Another possibility, as discussed in Cathy O'Neil's book *Weapons of Math Destruction*, is that some variables can lead to self-fulfilling prophecies in that they create "negative feedback loops."[106] That is, instead of modeling reality, they create that reality. An example of this would be related to the repayment of loans. In the case of payday loans, if poor language skills are a predictor of loan default, then high interest fees are levied against these high-risk people, which itself increases the risk of default. That is, if such a person invariably has trouble paying these loans, then the model is validated even though the difficulty in repaying the loan might instead have been as a result of the high fees. Such processes lead to a vicious feedback loop.

Before continuing further, it would be helpful to consider both bias and fairness, although both will be discussed further in subsequent CIA publications. As an example of bias, one major tech company wanted

[101] *Synced*. Tree boosting with XGBoost: Why does XGBoost win "every" machine learning competition? 2017. https://syncedreview.com/2017/10/22/tree-boosting-with-xgboost-why-does-xgboost-win-every-machine-learning-competition/.

[102] AIcrowd. Insurance pricing game: Challenges. Accessed September 10, 2021. www.aicrowd.com/challenges/insurance-pricing-game.

[103] Arthur Charpentier. *Insurance: Risk Pooling And Price Segmentation*. ESSEC Paris, 2017. http://freakonometrics.free.fr/slides-essec-2017.pdf.

[104] AIcrowd. *Insurance Pricing Game Townhall*. YouTube video, 2:14:41. 2021. www.youtube.com/watch?v=GkU2IqZu1gA.

[105] Oscar Schwartz. In 2016, Microsoft's racist chatbot revealed the dangers of online conversation: The bot learned language from people on Twitter – but it also learned values. *IEEE Spectrum*, 2019. https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation.

[106] Evelyn Lamb. Review: *Weapons of Math Destruction*. *Scientific American*, 2016. https://blogs.scientificamerican.com/roots-of-unity/review-weapons-of-math-destruction/.

to automate its ability to review resumés by assigning a score in order to help filter through all the applications.[107] It built an ML algorithm to help with these tasks based on the people they had hired in the past, but it later realized that the algorithm was not rating in a gender-neutral way, as it was basing its algorithm on prior hiring experience where women were underrepresented.

Fairness, meanwhile, is a social construct with multiple mathematical constructions having been made to evaluate fairness; however, they result in conflicting outcomes, with no decision being fair to all parties. As such, it is important that fairness be well defined by the business leaders so that the data scientists can best try to effectively assess the models they use. For a larger discussion, consult Rao and Golbin's article on fairness in AI.[108]

Insurance companies are not immune to the issues of responsible AI. Questions over whether credit score, for instance, should be used as a rating variable form an important example. While often used because of its predictiveness, questions arise out of the possibility that credit score calculation disproportionately affects people of colour and the poor.[109] For instance, "in Florida, adults with clean driving records and poor credit scores paid an average of $1552 more than the same drivers with excellent credit and a drunk driving conviction."[110] It is possible that the model justifies this pricing, but it is hard to convey such results are fair when drunk-driving convictions should in theory be more causally related to driving behaviour than credit score.

Other recent examples include the European Union banning the use of gender in algorithms and the UK creating legislation on price walking, which is discussed below. **In light of these examples, some major points of consideration are the following:**[111]

1. Fairness: Are you minimizing bias in your data and AI models? Are you addressing bias when you use AI?
2. Interpretability: Can you explain how an AI model makes decisions? Can you ensure those decisions are accurate?
3. Robustness and security: Can you rely on an AI system's performance? Are your AI systems vulnerable to attack?
4. Governance: Who is accountable for AI systems? Do you have the proper controls in place?
5. System ethics: Do your AI systems comply with regulations? How will they impact your employees and customers?

[107] Jeffrey Dastin. Amazon scraps Secret AI recruiting tool that showed bias against women. Reuters, 2018. www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

[108] Anand Rao and Ilana Golbin. What is fair when it comes to AI bias? *Strategy + Business*, 2019. www.strategy-business.com/article/What-is-fair-when-it-comes-to-AI-bias.

[109] Sarah Ludwig. Credit scores in America perpetuate racial injustice. *The Guardian*, 2015. www.theguardian.com/commentisfree/2015/oct/13/your-credit-score-is-racist-heres-why.

[110] Lamb, Review: *Weapons of Math Destruction*.

[111] PwC United States. 2019 AI predictions: Six AI priorities you can't afford to ignore. 2019. https://web.archive.org/web/20211109223336/www.pwc.com/us/en/services/consulting/library/artificial-intelligence-predictions-2019.html.

Trying to deal with all these facets of responsible AI is important in building trust with your customers and mitigate against potential risks as highlighted below:

Source: PwC Responsible AI[112]

The British government has published an ethical framework as a way to deal with the changing technical landscape.[113] The framework was made for government data scientists, although the checklist and assessment tools contained therein would equally apply to a person working at an insurer as it would a person working in the public sector. As such, the quick checklist could be a useful tool when preparing work on a project.

Similar initiatives are underway with the different actuarial governing bodies investigating the challenges further. The CIA is in the process of building a study to analyze the ethical challenges faced by Canadian insurers while also publishing articles on the position of actuaries in AI ethics.[114] Meanwhile, the IFoA has published a guide for ethical data science, outlining a thorough discussion of the types of questions that need to be asked in addition to whether decisions are actuarially justified.[115] The Society of Actuaries (SOA) has also produced a certificate program on the ethical use of predictive models.[116] As such, it is clear that ethical standards for actuaries will be further outlined in the years to come.

Explainable AI

In conjunction with responsible AI, attention needs to be placed on seeking to understand these new black-box algorithms by taking advantage of model-agnostic methods. Using explainable AI can help facilitate conversations with regulators and lead to faster model development time over traditional GLMs.

---

[112] PwC United States.

[113] Matt Hancock. *Data Science Ethical Framework*. Cabinet Office United Kingdom, 2016. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524298/Data_science_ethics_framework_v1.0_for_publication__1_.pdf.

[114] Joel Li, Rolly Molisho, and Harrison Jones. An introduction to AI ethics and regulation. *Seeing Beyond Risk*, 2021. www.seeingbeyondrisk.ca/2021/09/ai-ethics-and-regulation-in-insurance-actuaries-uniquely-positioned-for-success/.

[115] Institute and Faculty of Actuaries and Royal Statistical Society. *A Guide for Ethical Data Science*. 2019. www.actuaries.org.uk/system/files/field/document/An%20Ethical%20Charter%20for%20Date%20Science%20WEB%20FINAL.PDF.

[116] Society of Actuaries. Ethical & Responsible Use of Data & Predictive Models Certificate Program. Accessed October 5, 2021. www.soa.org/programs/ethical-responsible-data-certificate/?utm_medium=Email&utm_source=SNWArticle&utm_campaign=ERUCert_2021&utm_content=2021-08-25.

As briefly touched upon above, companies are facing pressure to explain their models to their customers, employees, and regulators. One of the major developments in recent years is explainable AI (XAI), with the goal of trying to give insight into how black-box algorithms work, as ML algorithms often suffer from opacity regarding their internal mechanisms.

FAT* academics (meaning “fairness, accountability, and transparency” in multiple AI, ML, computer science, legal, social science, and policy applications) are primarily focused on generating principles for best practice in algorithmic implementations by helping modelers to make informed decisions and have publicly accountable algorithms that avoid hazardous social impact.[117] The Defense Advanced Research Projects Agency (DARPA)-funded researchers seem primarily interested in creating more human-understandable AI systems through effective explanations and by drawing on human psychology to help the XAI evaluator define a suitable evaluation framework. Fair Isaac Corporation (FICO) ran a competition to try and push XAI further by partnering with academic institutions to help with algorithmic explainability.[118] A large survey on the different methods can be found in an article by the IEEE.[119]

As well as helping to address pressures such as regulation and adopting good practices around accountability and ethics, there are significant benefits to be gained from being at the forefront of investing in explainability.[120] Greater confidence in AI means you can deploy it faster and more widely, without the fear that it will break or function unexpectedly.

Optimise	Retain	Maintain	Comply
Model performance	Control	Trust	Accountability
Decision making	Safety	Ethics	Regulation

PwC. *Explainable AI*[121]

As a concrete example for insurers, regulatory constraints are a major hurdle for any rate changes for auto insurance products. Issues with protected classes, and making sure pricing is equitable, become easier if one is able to have a better understanding of the relationship between variables, even when such relationships are non-linear. One example of a technique one can use is the partial dependence (PD) plot to give a graphical representation of a single feature importance. For a more in-depth discussion of the different tools available, refer to a PwC article containing plain-language descriptions[122] or an ebook by Christoph Molnar for a more technical perspective.[123] By integrating XAI into the process, additional trust

[117] FAT/ML. Fairness, Accountability, and Transparency in Machine Learning. Accessed October 12, 2021. www.fatml.org/.

[118] FICO Community. Explainable Machine Learning Challenge. Accessed September 10, 2021. https://community.fico.com/s/explainable-machine-learning-challenge.

[119] Filip Karlo Došilović, Mario Brčić, and Nikica Hlupić. *Explainable Artificial Intelligence: A Survey*. IEEE, 2018. https://ieeexplore.ieee.org/document/8400040.

[120] PwC United Kingdom. *Explainable AI: Driving Business Value Through Greater Understanding*. Accessed September 10, 2021. www.pwc.co.uk/audit-assurance/assets/explainable-ai.pdf.

[121] PwC. *Explainable AI*.

[122] PwC. *Explainable AI*.

[123] Christoph Molnar. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. GitHub, 2021. https://christophm.github.io/interpretable-ml-book/.

can be formed, and once-opaque processes become more understandable.[124] Situations where there is no regulatory burden can also be improved by XAI. Being able to justify rating decisions to customers and executives, control for vulnerabilities, and discover prior unforeseen patterns are all important parts of the modeling process.

In addition to using PDs to help understand the underlying behaviour of the model, additional constraints can be added to some of these black-box models to help make them behave more similarly to conventional GLMs. For instance, adding monotonicity constraints in your GBM can give them much of the same interpretability as GLMs, with little impact on the accuracy of the model.[125] For a worked example looking into responsible AI with interpretable models, we advise you to refer to some of the recent academic literature. Several discriminatory measures are included to help gauge fairness measures.[126]

One major advantage of using these automated models in collaboration with markers for interpretability is that one can accelerate the model lifecycle. Once the initial process is in place, one should be able to simply modify the initial set of assumptions and re-run the model. The pre-designated markers can then be analyzed to gauge whether there has been any difference in predictiveness or relative importance of the variables. By removing manual tuning, one should expect to speed up the model lifecycle.

Segmentation and knowing your customer

Companies are taking advantage of the additional sources of information to better understand their customers' needs and risk profiles. More customer information should translate into more equitable pricing, although privacy concerns will increase.

Segmentation was one of the areas where we found insurers heavily focused on analytics, whether it be for improving the companies' marketing capabilities, creating more risk-based pricing, or better assessing their underwriting risk. To this end, better segmentation is useful for generating a better understanding of the customer, in terms of their risk profile and their preferences. As noted by an AI-driven US insurer, the promise of increased segmentation could lead to pricing that is not biased and more equitable,[127] although, as noted earlier concerning responsible AI, others have their doubts. The argument is along the lines that a protected class might be charged more, but it is not because it is a protected class, but merely because that class on average has particular characteristics. Certainly, with better segmentation, questions of privacy and the intrusion of insurers into the private lives of their customers pose a problem both for data protection and for how such information will shape society.

Historically, one of the issues affecting insurers is that of information asymmetry. That is, customers know far more about themselves than the insurer, with insurance companies working to gain more information on customers, so as not to be adversely selected. One reason for this lack of information is that insurers have few touchpoints with customers, as highlighted above. With the development of technology and digitalization, insurers are able to collect customer data through non-traditional channels, such as IoT and

[124] PwC United States. Insurance claims estimator uses AI for efficiency, 2017. www.pwc.com/us/en/library/case-studies/auto-insurance-ai-analytics.html.

[125] Molnar. *Interpretable Machine Learning*.

[126] Navdeep Gill, Patrick Hall, Kim Montgomery, and Nicholas Schmidt. A responsible machine learning workflow with focus on interpretable models, post-hoc explanation, and discrimination testing. *Information*, 2020. https://doi.org/10.3390/info11030137.

[127] Daniel Schreiber. AI can vanquish bias: Algorithms we can't understand can make insurance fairer. Lemonade (blog). Accessed September 10, 2021. www.lemonade.com/blog/ai-can-vanquish-bias/.

social media, to supplement the few traditional touchpoints. However, the big data paradigm, which promises to achieve personalization of risk, comes in direct conflict with some of the main tenets of insurance concerning risk pooling and homogeneity of risks within a class.[128] That is, big data promises to lift the opacity of the individual and in doing so deconstruct the pooling process.

In discussions with consulting partners abroad, GLMs continue to be the dominant tool for modeling in Asia, the United States, the United Kingdom, and Europe. The United Kingdom specifically has stated this in an official publication from 2016.[129]

Outside of insurance, a US bank is using ML to determine whether the discounts given to customers by bankers were in fact targeting high-value customers.[130] An insurance company could perform a similar analysis on the discounts provided to customers by brokers and agents. In addition, it might be worth analyzing claim adjusters' case reserving or claim settlement patterns.

Price optimization and customer lifetime value

Treating price using demand curves instead of point estimates, and discounting future cash flows based on retention modeling, are pervasive abroad. By using models in conjunction, one can maximize the return on acquisition.

Of course, customers do not shop for insurance every year, and price is not the sole consideration when choosing an insurance provider. Companies are able to use price optimization for their portfolios, especially for personal auto and home, to take advantage of a customer's reluctance to change insurance providers. The idea is that calculations are purely related to risk characteristics, but the price charged is a point estimate of a distribution of possible prices a consumer is willing to accept.[131] By considering the price elasticity of demand and competitor prices, one could estimate the gain or loss in the number of policies in response to varying the profit loading for a given segment. Some researchers have compared the performance of different algorithms for price optimization, with the authors noting that the practice is already easily applied in actuarial pricing software.[132] For a proposed methodology on how to implement an optimization model, one can refer to a recent ASTIN/AFIR-ERM (Actuarial Studies in Non-Life Insurance/Actuarial Approach for Financial Risks-Enterprise Risk Management) colloquium paper that used a combination of GLMs and GAMs to calculate the optimal increase in premiums for different groups.[133]

---

[128] Laurence Barry and Arthur Charpentier. *Personalization as a Promise: Can Big Data Change the Practice of Insurance?* PARI, 2019. www.chaire-pari.fr/wp-content/uploads/2019/12/WP-17-Telematics.pdf.

[129] Financial Conduct Authority. *Call for Inputs on Big Data in Retail General Insurance*. 2016. www.fca.org.uk/publication/feedback/fs16-05.pdf.

[130] Amit Garg, Davide Grande, Gloria Macías-Lizaso, and Christoph Sporleder. Analytics in banking: Time to realize the value. McKinsey and Company, 2017. www.mckinsey.com/industries/financial-services/our-insights/analytics-in-banking-time-to-realize-the-value.

[131] Arthur J. Schwartz. *Price Optimization and Insurance Regulation with Examples and Calculations*. Spring Meeting of the Casualty Actuarial Society in Colorado Springs, CO, May 2015. www.casact.org/sites/default/files/presentation/spring_2015_handouts_c-21.pdf.

[132] Giorgio Spedicato, Christophe Dutang, and Leonardo Petrini. Machine learning methods to perform pricing optimization: A comparison with standard generalized linear models. Casualty Actuarial Society, 2018. www.casact.org/abstract/machine-learning-methods-perform-pricing-optimization-comparison-standard-generalized.

[133] Wilson Mayorga and Diego Torres. *A Practical Model for Pricing Optimization in Car Insurance*. ASTIN/AFIR-ERM Colloquium, Panama, 2017. www.actuaries.org/panama2017/docs/papers/3b_ASTIN_Paper_Mayorga.pdf.

This practice of modifying prices based on other factors is already common in other industries, with the airline and hospitality industries using so-called "dynamic pricing."[134] Admittedly, the comparison is not completely the same as in those situations, as dynamic pricing is optimizing the price to fill vacancies based on the time remaining until the flight takes off and possibly on browser history. By contrast, the optimization problem for an insurance company would be on a much longer time horizon.

A related concept is that of customer lifetime value (CLV), wherein you try and model the expected profit of a potential customer over the expected time horizon where they are your customer. To do this, insurers model their ability to convert quotes into new business, the year-on-year retention of customers, and the customer's expected loss cost.

In discussion with a UK consulting actuarial team, price optimization and CLV were widely used, having led to recent regulation of the use of margin optimization. For auto and home insurance, the UK insurance market is less regulated compared to Canada's. Prior to regulatory changes, firms were adopting different prices for new and renewal business, with some companies pricing those policies below cost, with the goal of making back that cost by significantly increasing prices at renewal and by selling ancillary products.[135] The result was referred to as "price walking," a term which refers to the continued increase in margins on renewal, even after making back the initial discount.

For commercial large accounts insurance and reinsurance, these considerations are of lesser importance due to data volume and business nature. While making sure there is no substantial dropping in gross premiums, interviewees relayed that classification of risk profiles was the more important problem as opposed to optimizing pricing.

The use of behavioural economics, however, is not limited to the pricing of contracts. It can also be used for claims management, to increase sales,[136] and to nudge tired drivers to take a break from driving.[137]

Reserving

Insurers can look to the academic literature for ideas on how to test and supplement current reserving practices. Significant research is being done on the use of advanced automated methods.

Advanced analytics has found some applications in P&C claims reserving, especially for data-rich lines of business. Interest in using advanced analytics for reserving has grown, as demonstrated by some of the new research methods being used for reserving. Specifically, NNs have been applied in a few different ways, with one recent example[138] starting with the Mack Chain Ladder method, but replacing the simplified regression assumption with a (non-linear) NN regression model that accounts for the individual

---

[134] Marco Alderighi, Alberto A. Gaggero, and Claudio A. Piga. *The Hidden Side of Dynamic Pricing in Airline Markets*. Munich Personal RePEc Archive, 2016. https://mpra.ub.uni-muenchen.de/71674/1/MPRA_paper_71674.pdf.

[135] Financial Conduct Authority. *General Insurance Pricing Practices*, 2020. www.fca.org.uk/publication/market-studies/ms18-1-3.pdf.

[136] Gregor Becker, Anne Dreller, Anna Güntner and Johannes-Tobias Lorenz. Behavioral science in insurance: Nudges improve decision making. McKinsey and Company, 2020. www.mckinsey.com/industries/financial-services/our-insights/insurance-blog/behavioral-science-in-insurance-nudges-improve-decision-making.

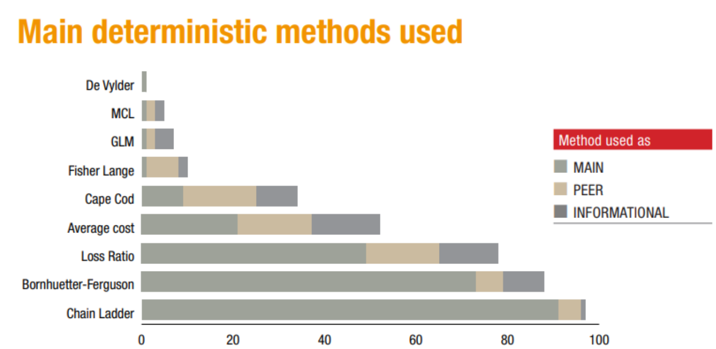
[137] Neil Huzinga. Nudge theory and Insurtech: Happy bedfellows? Insurance-Canada.ca (blog), 2018. www.insurance-canada.ca/2018/01/11/nudge-theory-insurtech-happy-bedfellows/.

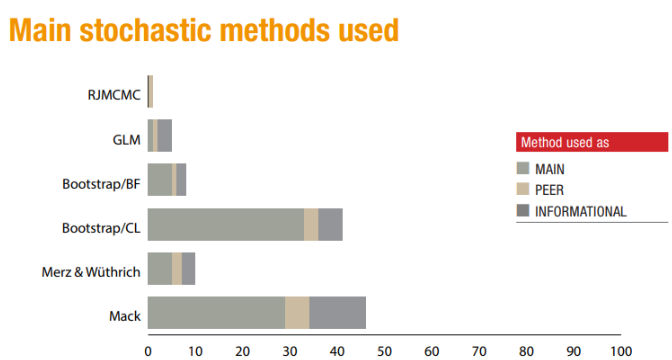
[138] Mario V. Wuthrich. *Neural Networks Applied to Chain-Ladder Reserving*. SSRN, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2966126.

differences between claims. Another approach using NNs[139] allowed for joint modeling of paid losses and claims outstanding, and the incorporation of heterogeneous inputs. Individual claims reserving has also had renewed interest in trying to predict the ultimate claim settlement value directly.[140] Traditionally, individual claim modeling implies fitting a statistical model to claim payments and then additionally modeling the number of claims. As Mario Wüthrich[141] points out, such approaches are not flexible and fail to take into account claim data available on file.

One caveat to the application of ML techniques to claims data is that it is only possible on claims that have already been reported. As a result, additional models need to be built to predict the Incurred But Not Yet Reported (IBNYR) data. In addition, getting buy-in from key stakeholders within the company, auditors, and regulators is difficult due to challenges in explaining the results from “black-box” models as compared to prior period estimates. These challenges were echoed in a recent survey of Canadian P&C insurers on behalf of the IFoA.[142] Challenges also exist in articulating the cost/benefit of implementing these changes, let alone the technical difficulties in familiarizing staff with the methods.

In 2016, ASTIN,[143] the non-life section of the International Actuarial Association, surveyed reserving practices internationally and found that the Chain Ladder and Bornhuetter–Ferguson were the most commonly used methods. Participants also mentioned increased interest in stochastic methods (bootstrap, Mack), and a need to move towards individual claims reserving and to better link the reserving process with the pricing process.





ASTIN. *Non-Life Reserving Practices*[144]

Canadian results were also stated, broken down by method as follows:

[139] Kevin Kuo. *DeepTriangle: A Deep Learning Approach to Loss Reserving*. Cornell University (arXiv.org), 2019. https://arxiv.org/pdf/1804.09253.pdf.

[140] Mario V. Wuthrich. *Machine Learning in Individual Claims Reserving*. Swiss Finance Institute Research Paper No. 16-67, SSRN, 2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2867897.

[141] Wuthrich. *Neural Networks Applied to Chain-Ladder Reserving*.

[142] Jacqueline Friedland. *Survey of Canadian Actuaries on ML in Reserving*. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/system/files/field/document/MLR_CanadaSurvey.pdf.

[143] Pierre Miehe, Judith Lutz, Accroche-com and Fabrice Taillieu. *Non-Life Reserving Practices*. ASTIN Working Party on Non-life Reserving Practices, 2016. www.actuaries.org/ASTIN/Documents/ASTIN_WP_NL_Reserving_Report1.0_2016-06-15.pdf.

[144] Miehe et al. *Non-Life Reserving Practices*.

1. Standard claims: triangle-based technologies

		Main method	Peer method	Informational	Unused
DETERMINISTIC	Percentage	4%	4%	21%	71%
	Loss ratio	58%	8%	17%	17%
	Chain ladder	79%	17%	0%	4%
	Bornhuetter-Ferguson	88%	8%	0%	4%
	Cape Cod	8%	4%	13%	75%
	Average cost	21%	8%	29%	42%
	De Vylder	0%	0%	0%	100%
	Fisher-Lange	4%	0%	0%	96%
	GLM	0%	0%	4%	96%
	Munich Chain Ladder	0%	0%	0%	100%
STOCHASTIC	Market-based std dev	4%	0%	4%	91%
	Internal calibration	5%	0%	5%	91%
	Mack	0%	9%	13%	78%
	Merz & Wüthrich	0%	0%	0%	100%
	GLM	0%	0%	4%	96%
	Bootstrap / CL	17%	0%	9%	74%
	Bootstrap / BF	13%	0%	4%	83%
	RJMCMC	0%	0%	0%	100%

ASTIN. *Non-Life Reserving Practices*[145]

Abroad, the IFoA has been organizing several working parties to research reserving methodologies, including one on ML methods,[146] another on governance as highlighted above,[147] and a disbanded one on pragmatic stochastic reserving.[148] A worked example of ML as applied to reserving by the above reserving methodologies working group can be found on its website.[149] ASTIN also did an analysis of different types of ML and traditional methods,[150] with GBMs performing best in its tests. It does note as well that, while traditional methods were not the most accurate, they also were rarely the "worst" models.

Based on our discussions with some insurers, we observed the number of insurers testing ML techniques for reserving practice and stochastic reserving methods is increasing. Furthermore, those Canadian P&C insurers that have started to explore using these models have reported success with longer-tailed lines. Highlighted in these discussions was the need to have frequently updated dashboards and discussions with other stakeholders to better understand the model output. Some companies are calculating reserves in parallel by performing traditional approaches such as the Chain Ladder method, and performing additional analysis with stochastic and ML models. That is, instead of using ML models to replace existing models, these new methods might be used to supplement current models by providing additional information.

Geospatial analysis and climate change

Increased focus on spatial variables, the interdependence among risks, and an insurer's role in climate change are at the forefront of management and modelers' minds in their efforts to curb present and future

[145] Miehe et al. *Non-Life Reserving Practices*.

[146] Institute and Faculty of Actuaries. General insurance machine learning in reserving. 2020. www.actuaries.org.uk/practice-areas/general-insurance/research-working-parties/general-insurance-machine-learning-reserving.

[147] Diffey et al. *Can You Trust Your Reserving?*

[148] Institute and Faculty of Actuaries. Pragmatic stochastic reserving. Accessed January 26, 2022. www.actuaries.org.uk/practice-areas/general-insurance/disbanded-research-working-parties/pragmatic-stochastic-reserving.

[149] Grainne McGuire and Jacky Poon. ML modelling on triangles: A worked example. Institute and Faculty of Actuaries, 2021. https://institute-and-faculty-of-actuaries.github.io/mlr-blog/post/f-mlr3example/.

[150] Salma Jamal, Stefano Canto, Ross Fernwood, Claudio Giancaterino, Munir Hiabu, Lorenzo Invernizzi, Tetiana Korzhynska, Zachary Martin, and Hong Shen. *Machine Learning and Traditional Methods Synergy in Non-Life Reserving*. ASTIN, 2018. www.actuaries.org/IAA/Documents/ASTIN/ASTIN_MLTMS%20Report_SJAMAL.pdf.

risks. Meanwhile, advances in modeling techniques are forcing established catastrophe modelers to update their models to compete with InsurTechs.

Climate change is having a significant impact on the way insurers are dealing with risk. This can be seen in the way they are modeling catastrophes, from the risk of wildfires in Alberta to hail in the Prairies, and the need to update floodplain risks. These ignore the other possible catastrophe risks not associated with climate change, or, for instance, a large earthquake on the Pacific coast.[151] Repercussions of more granular risk analysis have been highlighted in the New Zealand property market, illustrating the need for better territorial classification as it relates to climate change.[152] Borrowing some of the models used in gerrymandering could prove helpful in better defining territories beyond the forward sortation area (FSA) codes that designate a person's postal code. Once the territories are set, additional consultation of the literature on spatial smoothing might prove useful. A recent CAS E-Forum presented a survey of results on convex optimization as applied for insurance purposes, with Whittaker graduation techniques reformulated as linear programs.[153] Reviewing the literature of spatial analysis as performed in other industries, such as kriging in the geostatistics literature,
might also prove useful. For a worked example outlining the steps to perform spatial smoothing, refer to a 2012 ratemaking seminar presentation on territorial ratemaking.[154]

To deal with climate risks, insurers were already using catastrophe-modeling software, although some of the companies surveyed are turning to InsurTechs[155] to help in their underwriting process to predict wildfire risk. Another insurance company will be using a property data firm to help better understand risk through geospatial analytics and to make recommendations to customers, such as clearing surrounding brush.[156] Some risks are becoming uninsurable as a result of the changing climate, with extreme events such as bushfires in Australia,[157] drought and wildfires in California wine country,[158] and flooding on the coasts.[159] For Canadian insurers, accounting for such risks will require more frequent updates of floodplain maps and testing of wildfire risk when studying these variables in individual-risk rating algorithms.

As relayed in our interviews, companies are looking to tidal data and weather data to help in evaluating these risks. In this regard, the aims of reinsurers and primary insurers are no different in managing accumulation and tail risk to their business, as cat losses could pierce through their treaty limits.

[151] Katie Dangerfield. "Inevitable" 9.0 earthquake, tsunami will hit Canada's West Coast: Expert. *Global News*, 2020. https://globalnews.ca/news/3981536/tsunami-earthquake-canada-the-big-one/.

[152] Mamiko Yokoi-Ara. *The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector*. Organisation for Economic Co-operation and Development, 2020. www.oecd.org/finance/The-Impact-Big-Data-AI-Insurance-Sector.pdf.

[153] Dimitri Semenovich. *Applications of Convex Optimization in Premium Rating*. Casualty Actuarial Society, 2013. www.casact.org/sites/default/files/database/forum_13spforum_semenovich.pdf.

[154] Eliade Micu. *Territorial Ratemaking*. Casualty Actuarial Society and Eagle Eye Analytics, 2012. https://cas.confex.com/cas/rpms12/webprogram/Presentation/Session4723/Terr%20Ratemaking%20EEA%20v2.pdf.

[155] Insurance-Canada.ca. Aon and Zesty.ai revolutionize underwriting with property data solution powered by artificial intelligence. 2019. www.insurance-canada.ca/2019/03/13/aon-zesty-ai-property-data-solution/.

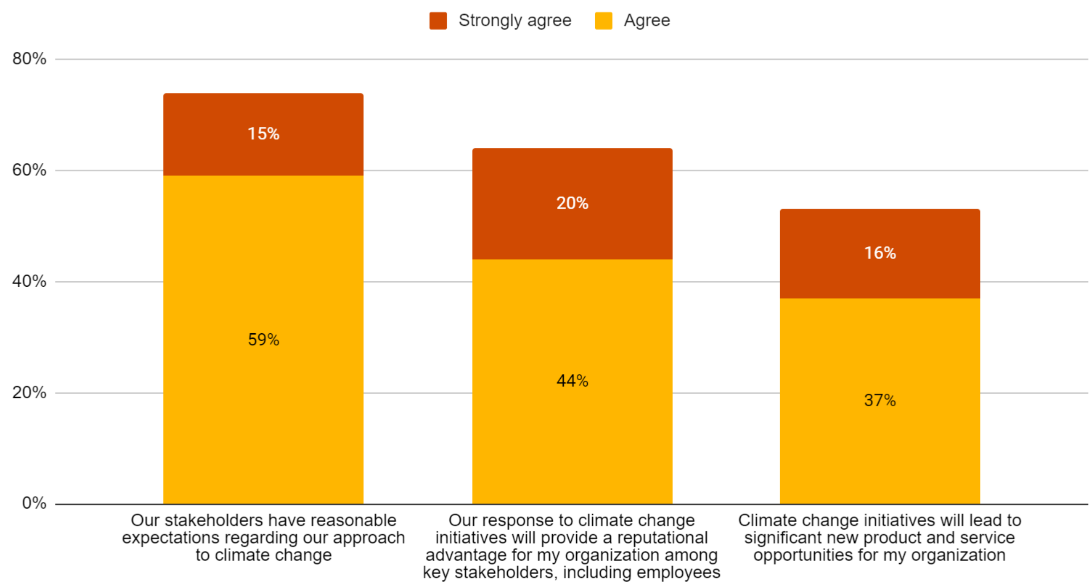
[156] Leslie Scism. Some California homeowners can get coverage again after wildfires. *Wall Street Journal*, 2021. www.wsj.com/articles/some-california-homeowners-can-get-coverage-again-after-wildfires-11623589200.

[157] Royce Kurmelovs. Climate change could put insurance out of reach for many Australians. *The Guardian*, 2021. www.theguardian.com/australia-news/2021/mar/02/climate-change-could-put-insurance-out-of-reach-for-many-australians.

[158] Christopher Flavelle. Scorched, parched and now uninsurable: Climate change hits wine country. *New York Times*, 2021. www.nytimes.com/2021/07/18/climate/napa-wine-heat-hot-weather.html.

[159] Bradley Hope and Nicole Friedman. Climate change is forcing the insurance industry to recalculate. *Wall Street Journal*, 2018. https://web.archive.org/web/20181207130059/www.wsj.com/graphics/climate-change-forcing-insurance-industry-recalculate/.

Finally, we would be remiss not to include details concerning environmental, social, and governance (ESG) factors in the insurance industry and financial analysis. One of the considerations when investing is not to have risks that are too concentrated. During the Fort McMurray incident in 2016, oilsands production dropped, while insurers were liable for the cost of repairing the residents' homes. Insurers are trying to use ESG to help in predictions for their models, although such progress is still in its infancy.[160] In addition, insurers have an incentive to try and avoid the worst climate impacts with some companies, such as one larger commercial writer saying it will no longer invest in companies that derive more than 30% of their revenue from mining coal.[161]



PwC. 23rd Annual Global CEO Survey[162]

A recent survey of global CEOs shows a growing awareness of ESG, with it being increasingly important in an insurer's strategies.

Fraud

Fraud has been a successful area for advanced analytics, with significant gains being found in prediction accuracy and a wide range of techniques being tested. Non-standard techniques such as social network analysis and behavioural economics are also being tested to mitigate fraud.

Fraud is a problem afflicting all parts of the financial services industry, with several high-profile examples including Enron's accounting practices and Bernie Madoff's Ponzi scheme. The insurance industry is no different. One estimate put the amount of P&C fraud in the United States at $38 billion for 2020.[163] Based on 2018 Canadian insurance industry estimates, the Insurance Institute of Canada (IIC) disclosed that

---

[160] Allianz Global Corporate & Specialty and The Value Group. The predictive power of ESG for insurance. 2018. www.agcs.allianz.com/news-and-insights/expert-risk-articles/the-predictive-power-of-esg-for-insurance.html.

[161] Greg Meckbach. The tough question for insurers withdrawing coverage from coal. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/climate-change/the-tough-question-for-insurers-withdrawing-coverage-from-coal-1004198475/.

[162] PwC Global. Insurance trends 2020: Moving from resilience to reinvention will help insurers succeed in uncertain times. 2020. www.pwc.com/gx/en/ceo-agenda/ceosurvey/2020/trends/insurance.html.

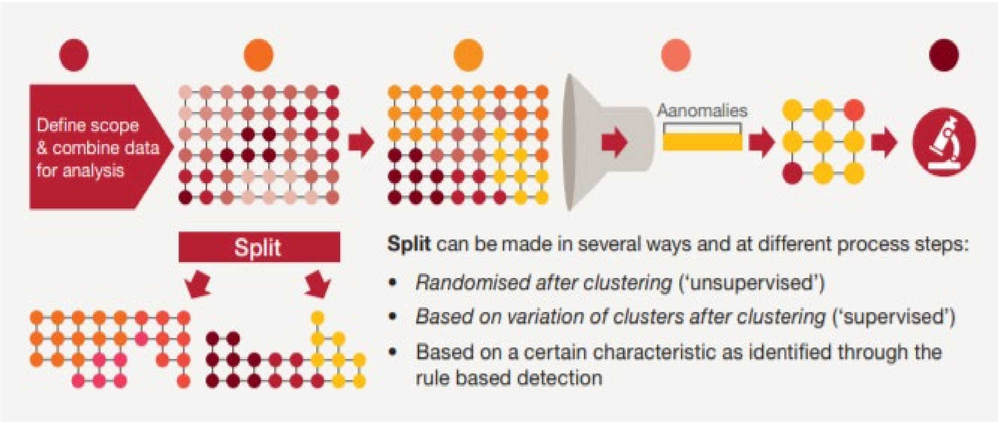
[163] Insurance Information Institute. Background on: Insurance fraud. 2021. www.iii.org/article/background-on-insurance-fraud.

around 5 to 15% of personal automobile premiums were allocated to cover undetected fraudulent claims.[164] In a 2018 US survey of P&C insurers, nearly three-quarters stated that fraud had increased significantly or slightly in the previous three years.[165] Also according to US data, personal property homeowners policies saw a 39% increase from 2017 through 2019 in slips and falls.[166]

To deal with fraud, Équité Association was recently established as the new Canadian non-profit organization integrating the cross-insurer data analytics company CANATICS and the Investigative Services Division (ISD) of the Insurance Bureau of Canada (IBC).[167] Équité will work similarly to the Insurance Fraud Bureau (IFB) in the United Kingdom and the National Insurance Crime Bureau (NICB) in the United States. The pooling of data and use of analytics will help Équité identify potential frauds perpetrated across carriers.

Insurers are using advanced analytics to maximize fraud detection, with high-performing carriers having expedited the identification of fraud.[168] One well-trained special investigation unit (SIU) analyst utilizing the power of analytics can filter and route hundreds (or even thousands) of claims that would ordinarily be reviewed manually. As we note below, insurers are trying to accelerate the triaging process through increased automation in triaging. When many US insurers started making their first major investments in anti-fraud technology 10 years ago, it was believed that automation would grow into a leading source of case referrals.[169] However, the number of accepted referrals has not changed too significantly, although investigators might be dealing with fewer, larger, and more-complex cases. Regardless, insurers will still need field investigators as part of their SIUs to find material evidence that can prove fraud.

An approach for how one might perform fraud analytics is shown in the file of the linked diagram below:



[164] Indrani Nadarajah. Auto insurance fraud. Insurance Institute of Canada, 2018. www.insuranceinstitute.ca/en/cipsociety/information-services/advantage-monthly/0718-insurance-fraud.

[165] Insurance Information Institute. Background on: Insurance fraud.

[166] National Insurance Crime Bureau. Slip & fall incidents rise according to the National Insurance Crime Bureau. 2021. www.nicb.org/news/news-releases/slip-fall-incidents-rise-according-national-insurance-crime-bureau.

[167] Greg Meckbach. New industry anti-fraud group gets its first CEO. *Canadian Underwriter*, 2021. www.canadianunderwriter.ca/associations/new-industry-anti-fraud-group-gets-its-first-ceo-1004210351/.

[168] Michael Skiba, Jeffrey G. Rapattoni, and Chris McKibbin. *Secrets to Combating Insurance Fraud with Data Analytics: Three Insurance Executives Offer a Global Perspective*. Casualty Actuarial Society, 2018. www.casact.org/sites/default/files/presentation/spring_2019_presentations_g-2_rapattoni_1.pdf.

[169] Coalition Against Insurance Fraud. *2020 Insurer SIU Benchmarking Study: Insurers Finding Stability in Their Anti-Fraud Units*. 2020. https://insurancefraud.org/wp-content/uploads/Benchmarking-Study-Summary.pdf.

PwC. *Insurance Fraud Analytics*[170]

As mentioned when discussing the benchmark analysis, fraud often came up as an area with significant results for the business. Also highlighted was the use of numerous different types of techniques. Respondents mentioned the use of NLP to read claims notes, while others used regularized logistic models. One of the respondents stated that their SIU team, in combination with a fraud model, yielded the greatest benefit to their company.

There are also several software companies that partner with insurers to help in their fraud detection processes. Many companies we spoke to, however, seemed to be doing the analysis in-house. In Singapore, one insurer was able to build a minimum viable product in seven months, and achieved more than 92% prediction accuracy for its travel and personal accident lines.[171] However, the company mentioned that motor fraud is not as straightforward.

Regardless of the method used for analyzing claims data, a threshold will have to be crossed regarding whether to send the claim to the SIU. To this end, there has also been some research done on optimal auditing,[172] where the optimal number of audits are made so as to deter fraudulent activities.

Social network analysis is also another avenue of investigation, with one Chinese tech company having success using such a method.[173] In keeping with earlier discussions on nudging, behavioural nudge tactics can also be used to try and decrease fraud.[174]

Implementation

It is necessary to remove sources of friction that impinge upon one's ability to implement new models and make changes to existing models. Updating processes and allocating additional resources for automating one's existing procedures promises to decrease the number of mistakes due to manual entry, accelerate turnaround times for model updates, improve customer experiences, and enable the faster deployment of products to capitalize on new trends and gain a first-mover advantage.

Application Programming Interfaces

APIs provide an effective way to make systems communicate with one another. Insurers are beginning to build API ecosystems to accelerate business processes and facilitate the management of connections between different business partners.

[170] PwC Hong Kong. *Insurance Fraud Analytics*. 2017. www.pwccn.com/en/risk-assurance/publications/insurance-fraud-analytics.pdf.

[171] Basil Han. Improving fraudulent claims detection with AI. AI Singapore, 2021. https://aisingapore.org/2021/07/improving-fraudulent-claims-detection-with-ai/.

[172] Katja Müller, Hato Schmeiser, and Joël Wagner. Insurance claims fraud: Optimal auditing strategies in insurance companies. *Variance Journal*, 2016. www.casact.org/abstract/insurance-claims-fraud-optimal-auditing-strategies-insurance-companies&sa=D&source=editors&ust=1631824051468000&usg=AOvVaw3NaVFYQ7g6G3q2wiD0HoH9.

[173] Chen Liang, Ziqi Liu, Bin Liu, Jun Zhou, Xiaolong Li, Shuang Yang, and Yuan Qi. *Uncovering Insurance Fraud Conspiracy with Network Learning*. Cornell University (arXiv.org), 2020. https://arxiv.org/pdf/2002.12789v1.pdf.

[174] Jim Guszcza. The last-mile problem: How data science and behavioral science can work together. Deloitte Insights, 2015. www2.deloitte.com/us/en/insights/deloitte-review/issue-16/behavioral-economics-predictive-analytics.html.

Modern business ecosystems need to rethink their approach to innovation and integration.[175] APIs are software intermediaries that allow two applications to talk to each other.[176] The application will connect to a server to do some task, and then relay the information back to the application in a readable way. For insurers, this involves service interfaces to aggregators, consumers, and third parties, with APIs as tactical wrappers.

To take a concrete example, a model could be built using a script and an API for other parts of the business to call. Each time you want to set up a quote, information is sent to the API concerning the potential policyholder information, and a JavaScript Object Notation (JSON) is spit out with the correct quote. There will be a verification process, where the person interacting with the API has to be authorized to use the API, as well as an authentication process. InsurTechs have increasingly been given access to insurers' open APIs, scrubbed clean of policyholder data, to develop innovative apps that can help differentiate the incumbents in the market.[177] Taking advantage of open APIs is also something insurers will want to use going forward, such as using Google Maps to help in pricing home insurance,[178] or Twitter to gauge consumer sentiment in one's product. Analyzing Twitter data to understand the level of distrust has already been tested in the banking industry to challenge the predictions of a bank's retail funding model, and to capture possible threats to financial stability deriving from an increase of public distrust in the banking system.[179] Analyzing customer calls through word clouds and sentiment analysis could be used to try and identify customer satisfaction and pain points.[180]

RESTful APIs are APIs with specified API verbs, such as GET, POST, and DELETE.[181] They can be thought of analogously to how one would use SQL to interact with a database.

As insurers move to the cloud, they will want to create an API ecosystem to help them achieve greater business velocity.[182] This allows insurers to develop a services architecture where they can manage connections with partners, regulators, and different parts of their own business through APIs and enable them to deliver innovative services and business models.

Rating algorithm

Critical to one's business is becoming more nimble by identifying pain points in the implementation of changes to the rating algorithm.

[175] Dennis Ashby and Claus T. Jensen. *APIs for Dummies*. John Wiley & Sons and IBM, 2018. www.ibm.com/downloads/cas/GJ5QVQ7X.

[176] MuleSoft. What is an API? (Application Programming Interface). Accessed September 14, 2021. www.mulesoft.com/resources/api/what-is-an-api.

[177] Jeff Picozzi. What APIs mean for an open and connected insurance industry. Red Hat Blog, 2020. www.redhat.com/en/blog/what-apis-mean-open-and-connected-insurance-industry.

[178] Google Cloud. Allstate: Helping agents build better relationships with customers. Accessed September 14, 2021. https://cloud.google.com/customers/allstate.

[179] Financial Stability Board. *Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications*. 2017. www.fsb.org/wp-content/uploads/P011117.pdf.

[180] Xiyue Liao, Guoqiang Chen, Ben Ku, Rahul Narula, and Janet Duncan. Text mining methods applied to insurance company customer calls: A case study. *North American Actuarial Journal*, 2020. www.semanticscholar.org/paper/Text-Mining-Methods-Applied-to-Insurance-Company-A-Liao-Chen/9bc53ae539fa6a96ea974f77768544b545aabb9a.

[181] REST API Tutorial. HTTP methods. Accessed September 14, 2021. https://restfulapi.net/http-methods/.

[182] Capgemini. Capgemini perspectives: Cloud native comes of age in insurance. 2018. www.capgemini.com/article/cloud-native-comes-of-age-in-insurance/.

Critical to an insurer's adaptability is the ability to quickly update its rating algorithm and bring products to market quickly in order to tailor its offerings to changing consumer tastes. As noted in recent findings by a research firm, speed-to-market is a key concern for P&C insurers, with the average time to market being 7–9 months depending on the line and size of the company.[183] Unlike our findings, smaller insurers were more nimble than their large counterparts. Another interesting finding was that newer technology was correlated with faster product modifications. Technology is not the only factor, but updating ageing infrastructure and having a well-thought-out strategy and infrastructure for one's policy administration system (PAS) and associated rating engine, whether external or built into the PAS,[184] should help. Insurers must also focus on speeding up analysis through streamlined analysis, and integrating IT at an earlier stage to speed up change management.[185] Regarding the analysis, making sure there is a comprehensive pricing model, incorporating multiple lines into one tool, instead of tens of individual Excel-based pricing models, should help refine and speed up analysis.[186] Other challenges concern integrating business units, especially after M&A. M&A can lead to a fragmenting of systems, making it difficult to manage[187] and to perform analytics processes.

Automation

Insurers abroad are finding ways to upgrade efficiency by automating processes such as effective claims-handling through improvements to the triaging process.

As noted earlier, the triaging of claims is an area where companies are having significant success in using predictive analytics. As discussed above, the idea is to use analytics to help with the simpler claims and let adjusters deal with the more complex ones. One US personal auto and home insurer has set a goal of automating up to 75% of claims.[188]

One particular example where automation has sped up the process is in claims-handling. In the United States, a leading insurer used ML where the model predicts with a level of confidence if a vehicle is a total loss or repairable. In some cases, the model allows the company to bypass the physical inspection process yet ensure a thorough investigation of the claim.[189] The process reduced the total loss process from as high as 15 days to as little as 30 minutes. Besides speeding up the claims process and decreasing the amount of manual work, ML allows customers to have a significantly better experience in having a faster workup.

---

[183] Matthew Josefowicz and Harry Huberty. Speed to market for property/casualty insurers. AiteNovarica, 2021. https://novarica.com/speed-to-market-for-property-casualty-insurers/.

[184] PwC Hong Kong. *Adding It All Up: Modern Rating Systems for P&C Carriers*. 2015. www.pwchk.com/en/migration/pdf/modern-rating-system-apr2015.pdf.

[185] Novarica. Insurer speed to market depends on process as much as technology. Insurance-Canada.ca, 2019. www.insurance-canada.ca/2019/03/26/novarica-speed-to-market-process.

[186] Deloitte. *Speed to Market: Part of the insurance series – Benefits of a New Policy Administration System: Why Going Live Is Not Enough*. 2015. www2.deloitte.com/content/dam/Deloitte/us/Documents/financial-services/us-cons-policy-admin-systems-speed-to-market-042415.pdf.

[187] Clint Boulton. Insurance firm banks on change management in digital overhaul. CIO United States, 2018. www.cio.com/article/3267641/insurance-firm-banks-on-change-management-in-digital-overhaul.html.

[188] Annmarie Geddes Baribeau. Insurers enjoy benefits from data modeling the claims process. *Actuarial Review*, *2020*. https://ar.casact.org/insurers-enjoy-benefits-from-data-modeling-the-claims-process/.

[189] Microsoft Docs. Identify guiding principles for responsible AI: State Farm case study. Accessed September 14, 2021. https://docs.microsoft.com/en-us/learn/modules/responsible-ai-principles7-responsible-ai-case-study.

In Canada, a leading insurer is similarly trying to accelerate the process of claim notification by getting images from collision-reporting centres.[190] The customer details and accident details are sent directly to the insurer without any need for the customer to call. An Indian insurer reduced claim response time by 30%, and customers were able to be reimbursed in as little as 15 minutes.[191]

Customer-centric

Market leaders in the United States and Asia are improving the customer experience by implementing an omni-channel approach, and by creating chatbots that provide on-demand service.

The customer experience is one of the areas where insurers are paying significant attention to try and improve their current value proposition, helping in acquisition, claims, and marketing. As noted above, insurers are working on having unified omni-channel approaches. Additionally, insurers are trying to improve customers' experience by trying to mimic human conversations through chatbots. Chatbots have been around since the 1960s, although they are becoming more widespread with the adoption of Amazon's Alexa, Google's Home, Apple's Siri, and Microsoft's Cortana.[192] They have been deployed across marketing, sales, and servicing. While a disclosure of chatbot identity can lead to a decrease in purchases,[193] consumers are increasingly willing to engage with them. A large US insurer has already launched an insurance service to connect to Alexa,[194] and another major US insurer expected to save $5 million in 2019 due to fewer conversations with agents. An AI-driven US insurer uses bots for quotes, claims, and customer questions through its AIs Maya, Jim, and CX.AI,[195] and, as noted earlier, even commercial insurers are using them to connect to small businesses using Facebook Messenger.

In Europe, 12% of insurance firms used a chatbot and 42% expected to use one within three years.[196] Of those companies, 43% of them built the chatbot in-house, and one company stated it expected to handle half of its consumer queries through chatbots. Typically, chatbots are produced using NLP and other ML algorithms.

For the moment, companies are using these chatbots to work in tandem with humans, due to their current limitations. Other challenges are clunky interfaces and difficulties in dealing with the nuances of human

[190] Greg Meckbach. Aviva rolls out automated claim notification at collision reporting centres. *Canadian Underwriter*, 2019. www.canadianunderwriter.ca/insurance/this-insurer-rolling-out-automated-claim-notification-at-collision-reporting-centres-1004159316/.

[191] IBM Services. IFFCO Tokio General Insurance Company Limited: Improving customer experience with smarter solutions. 2020. www.ibm.com/case-studies/iffco-tokio-ibm-services-ai.

[192] National Association of Insurance Commissioners. Chatbots. 2020. https://content.naic.org/cipr_topics/topic_chatbots.htm.

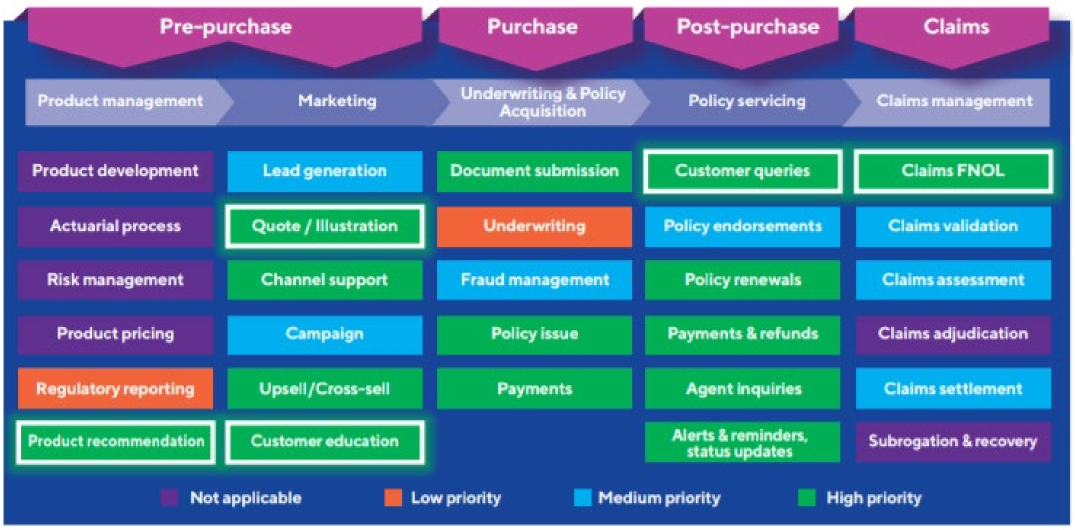
[193] Gil Press. AI stats news: Chatbots lead to 80% sales decline, satisfied customers and fewer employees. *Forbes*, 2019. www.forbes.com/sites/gilpress/2019/09/25/ai-stats-news-chatbots-lead-to-80-sales-decline-satisfied-customers-and-fewer-employees/?sh=5335ba948e05.

[194] *Insurance Journal*. Liberty Mutual giving consumers a voice in insurance via Amazon's Alexa. 2016. www.insurancejournal.com/news/national/2016/09/13/426162.htm.

[195] Tom Taulli. Lemonade IPO shows the power of AI (artificial intelligence). *Forbes*, 2020. www.forbes.com/sites/tomtaulli/2020/07/03/lemonade-ipo-shows-the-power-of-ai-artificial-intelligence/?sh=8263f053aebb.

[196] European Insurance and Occupational Pensions Authority. *Big Data Analytics in Motor and Health Insurance: A Thematic Review*. 2019. https://register.eiopa.europa.eu/Publications/EIOPA_BigDataAnalytics_ThematicReview_April2019.pdf.

language.[197] An example of one of these bots is IntelliBot,[198] designed by an Australian PhD student, with a large workup of various techniques included in the thesis paper.[199]



Cognizant. *The Future of Chatbots in Insurance*[200]

Designing your app and making the quoting process shorter is another way companies are trying to improve their customer experience. For example, a US insurer's AI-driven chatbot only asks 13 questions, allowing you to get a quote in 90 seconds, while producing 1,600 data points to analyze and derive individualized premiums at the back-end.[201] Another US InsurTech company boasts it can get you a quote in 60 seconds. One commercial insurer for small business in Canada helps its customers develop a quote in minutes, and also teamed up with a broker InsurTech to redesign its website after having improved its search engine optimization (SEO).[202] As discussed in interviews, one insurer noted that it is using behavioural economics to try and figure out how best to optimize the quoting process and follow up with customers. One way to test this would be to use A/B testing forms to maximize conversion rate.[203]

On-demand microinsurance

Implementing up-to-date infrastructure will enable insurers to handle new insurance developments such as on-demand forms of insurance, and pursue customers with changing needs.

---

[197] Srinivasan Somasundaram, Akshat Kant, and Prakhar Maheshwari. *The Future of Chatbots in Insurance*. Cognizant, 2019. www.cognizant.com/whitepapers/the-future-of-chatbots-in-insurance-codex4122.pdf.

[198] Mohammad Nuruzzaman and Omar Khadeer Hussain. IntelliBot: A dialogue-based chatbot for the insurance industry. *Science Direct*, 2020. www.sciencedirect.com/science/article/abs/pii/S0950705120301933.

[199] Mohammad Nuruzzaman and Omar Khadeer Hussain. *IntelliBot: A Domain-Specific Chatbot for the Insurance Industry* (awarded by University of New South Wales, Business, 2020). www.unsworks.unsw.edu.au/primo-explore/fulldisplay/unsworks_72771/UNSWORKS.

[200] Somasundaram et al. *The Future of Chatbots in Insurance*.

[201] Sara Morrison. A disturbing, viral Twitter thread reveals how AI-powered insurance can go wrong. *Vox*, 2021. www.vox.com/recode/22455140/lemonade-insurance-ai-twitter.

[202] Trufla Technology. Trufla Technology celebrates Bullfrog Insurance's innovative new website. *Canadian Underwriter*, 2018. www.canadianunderwriter.ca/inspress/trufla-technology-celebrates-bullfrog-insurances-innovative-new-website/.

[203] Formstack. How to A/B test your forms for maximum conversion. Accessed September 14, 2021. www.formstack.com/resources/guide-ab-test-web-forms-maximum-conversion.

On-demand insurance allows consumers to purchase insurance coverage on their smartphone whenever and wherever they want, usually when the asset requiring coverage is in use and at risk.[204] The term encompasses microinsurance, continuous underwriting, and the gig economy. *Microinsurance* refers to rapid underwriting of small risk, *continuous underwriting* to possibly real-time policyholder data determining consumer risk, and *the gig economy* to freelance activities such as those facilitated by Uber and AirBnB.[205]

The idea of on-demand microinsurance is that you swipe right on your phone to insure a possession, and swipe left to turn it off. That is, customers would only insure themselves when the object is at risk. The objective of this type of insurance is to capture the emerging paradigm of the shared economy and changes in consumer behaviour,[206] with some companies particularly targeting millennial consumers.[207]

A consultancy services company noted in its report that convenience and experience are fast replacing price as the key buying criteria for insurance today.[208] This is readily seen from the types of InsurTech companies, with the proliferation of chatbots highlighted above On-demand microinsurance seeks to fill the gap of consumers who feel "over-insured."[209] Implementation of such tasks is challenging, due to the number of products that would have to be priced and the need to make sure the whole process is frictionless.

Multiple InsurTechs have already started providing these services by developing platforms to help incumbent insurers develop their own products. Several examples of partnerships already exist in the UK,[210, 211] and one of these InsurTechs is partnering with an insurer here in Canada.[212]

Summary remarks

Based on the above discussion, the Canadian P&C insurance industry appears to be doing well compared to its international peers. That is, based on discussions with actuaries abroad, and from research highlighted above, Canadian P&C insurers appear to be using a significant amount of analytics despite limitations in their ability to collect as much data as other industries or insurers overseas. This disparity in

[204] National Association of Insurance Commissioners. On-demand insurance. 2021. https://content.naic.org/cipr_topics/topic_ondemand_insurance.htm.

[205] Jeff Goldberg. The 3 pillars of on-demand insurance. Insurance Thought Leadership, 2018. www.insurancethoughtleadership.com/the-3-pillars-of-on-demand-insurance/.

[206] Reni Parameshwaran, Himadri Sikhar Pramanik, Sayantan Datta, and Ujjwal Bunkar. *On-Demand Insurance: Challenges and Opportunities for Large Insurance Carriers*. TATA Consulting Services, 2019. www.tcs.com/content/dam/tcs/pdf/Industries/insurance/rise-of-on-demand-insurance.pdf.

[207] Ageas UK. App-based insurance cover. 2016. www.ageas.co.uk/press-releases/2016-press-releases/app-based-insurance-cover-back-me-up---powered-by-ageas/.

[208] Parameshwaran et al. *On-Demand Insurance*.

[209] Bon. What is on-demand insurance? *Insurance Marketer*, 2021. www.theinsurancem.com/what-is-on-demand-insurance/.

[210] *Finextra*. AXA and Trov bring "on demand" insurance to UK. 2016. www.finextra.com/newsarticle/29804/axa-and-trov-bring-on-demand-insurance-to-uk.

[211] Insurance-Canada.ca. Trov technology enables a new wave of consumer brands to offer digital renters insurance. 2021. www.insurance-canada.ca/2021/04/08/trov-enables-new-brands-digital-renters-insurance/.

[212] Co-operators Group. New insurtech partnership to provide on-demand insurance in Canada. 2018. https://newsreleases.cooperators.ca/2018-07-18-New-Insurtech-Partnership-to-Provide-On-Demand-Insurance-in-Canada.

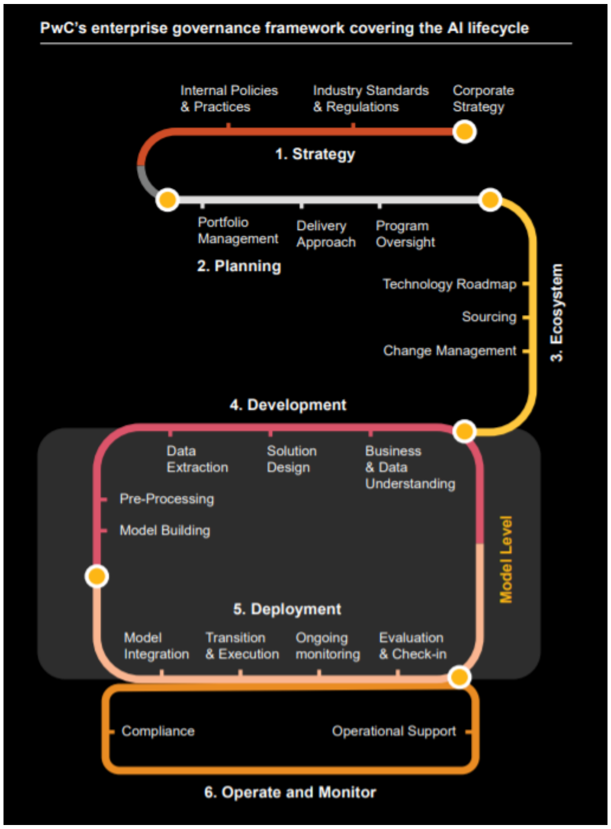
data collection is especially apparent when compared to Asia, as Canada has a significantly different rating environment and Canadian customers have different privacy expectations.

GLMs are still the preferred option for pricing and underwriting abroad, with some Canadian P&C insurers opting for newer techniques; however, Canadian companies are further behind in other parts of the business, such as customer engagement and marketing. When compared to the United States, Canadians have fewer full-stack InsurTech companies threatening their business model, although the threat persists.

Actionable approach for advanced analytics application

Based on our survey results and observations of advanced analytics applications in other industries and countries, the Canadian P&C insurance industry has improvement opportunities. In this section, we will provide an actionable approach for insurers to consider when they plan for advanced analytics initiatives.

Many frameworks have been made to deal with the data science lifecycle, from traditional data-mining methods, such as CRoss-Industry Standard Process for Data Mining (CRISP-DM),[213] to Agile frameworks, such as Scrum. These frameworks are not mutually exclusive, as CRISP-DM, for example, can be done in an Agile way. Leveraging the available data science frameworks[214] and our industry experience, we present an approach following a structure which is consistent with the previous sections in this report: *Management and internal support*, *Data*, *Technology*, *Model and business usage*, and *Implementation*.



The approach demonstrates the key steps throughout the entire project lifecycle. Based on our survey results, we provide suggestions and consideration from a project management perspective regarding planning, communication, and collaboration. These steps are addressed to ensure a seamless process based on common challenges we have observed in similar studies and projects.

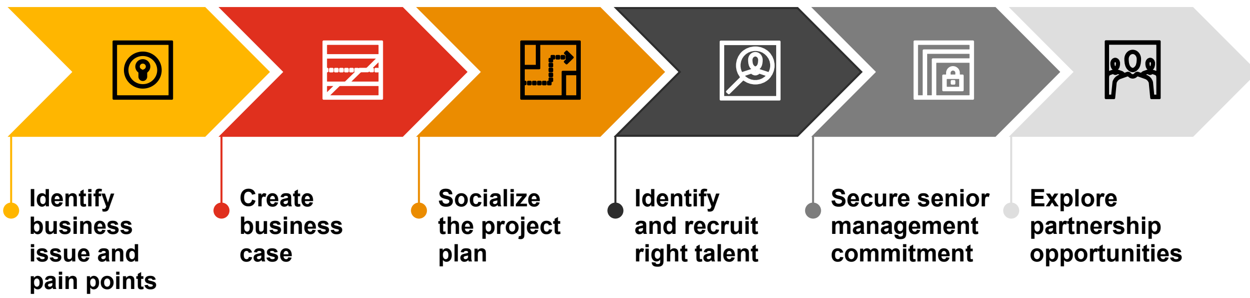
The proposed approach can be used as a reference if starting a new initiative or for ongoing projects. Throughout an advanced analytics application initiative, insurers will need to keep in mind what goal is to be achieved, the capability of existing infrastructure, how to source the data, and what additional investment is needed, and have an understanding of the short- and long-term return of these investments.

[213] Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rüdiger Wirth. *Step-by-Step Data Mining Guide*. SPSS, 2020. www.the-modeling-agency.com/crisp-dm.pdf.

[214] Steven Perkins, Hazel Davis, and Valerie du Preez. *Practical Data Science for Actuarial Tasks*. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/system/files/field/document/Practical%20Data%20Science%20for%20Actuarial%20Tasks%20v1.8.pdf.

PwC. *A Practical Guide to Responsible Artificial Intelligence (AI)*[215]

Management and internal support



To start an advanced analytics project initiative, management and internal support is critical. Most of the respondents in our survey indicated that their senior management are very supportive in developing advanced analytics capabilities. On the other hand, among respondents' challenges business buy-in is ranked second to implementing advanced analytics. Stakeholders may be hesitant about advanced analytics projects for reasons such as:

- Non-traditional modeling techniques can be complex and "black box" in fashion
- Lack of a clear benefit from projects vs competing projects requiring similar financial resources and time commitments
- Concerns about legacy systems unable to handle advanced analytics pipelines

To facilitate the discussions with stakeholders, these considerations need to be addressed in the business plan of each of the advanced analytics projects. Communication throughout the entire project process should be tailored based on the audience. It is important to translate technical topics to a non-technical audience and be transparent about a model's limitations and constraints. Pilot implementations and parallel runs are suggested to increase stakeholders' confidence for advanced analytics projects. Delivering a "lighthouse win" by starting with projects that offer demonstrable benefits with manageable risks can help propel advanced analytics projects into the limelight.[216]

In creating the business plan, it will be necessary to communicate what the problem is, give a reason why it is worth addressing, provide a vision of what the business is trying to achieve, and identify what would qualify as a success. Once these objectives have been identified, the situation must be assessed in terms of personnel, data, timelines, and other constraints, including regulatory ones. Socializing the business plan with all the necessary stakeholders is essential in making sure the solution is politically acceptable to the business. Early discussions will also help in identifying potential problems and constraints in advance, such as inadequate staffing, or challenges with IT implementation. As mentioned earlier, communication of the ideas must be tailored to the audience. For instance, describing how a fraud solution would impact

[215] PwC Global. *A Practical Guide to Responsible Artificial Intelligence (AI)*. 2019. www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.pdf.
[216] George Demarest. Four phases of operating big data. *CIO Review*. Accessed September 14, 2021. https://bigdata.cioreview.com/cxoinsight/four-phases-of-operationalizing-big-data-nid-15251-cid-15.html.

claims teams would be different to actuaries or the head of customer service, despite discussing the same model.

A technical goal is then set regarding prediction accuracy or model lift. Finally, a plan of action is produced, outlining the stages that need to be executed, as well as outlining ongoing support requirements for the developed solution.

We note that, before any task is undertaken, management will have already established the organizational structure in order to develop a clear analytics strategy; that is, whether to use CoEs or have their analytics teams decentralized.[217] In our survey, respondents from larger insurers tended to have CoEs. We expect this model to become most prevalent in the coming years.[218] However, care must be taken not to have teams siloed, with data scientists simply handing off to an IT specialist to code an API. Ideally, teams should be working together from the start. Depending on the needs of the project, additional staffing might be required for initial and continued development of the business case. As noted by one of our survey respondents, limitations on staffing will dictate whether to produce new models or maintain existing ones. Talent might need to be recruited due to technical limitations of the team, based on the goals and size of the project. Larger insurance companies have teams with broad skills, consisting of data architects, administrators, data scientists, statisticians, actuaries, and non-actuarial business experts. A reinsurer, meanwhile, might have a central global team dedicated to analytics with branches dealing less with the modeling aspects. Finally, insurers of all sizes might opt to use external partners, such as InsurTechs and consultants, as a way to help get the project deployed quickly without having to employ significantly more staff.

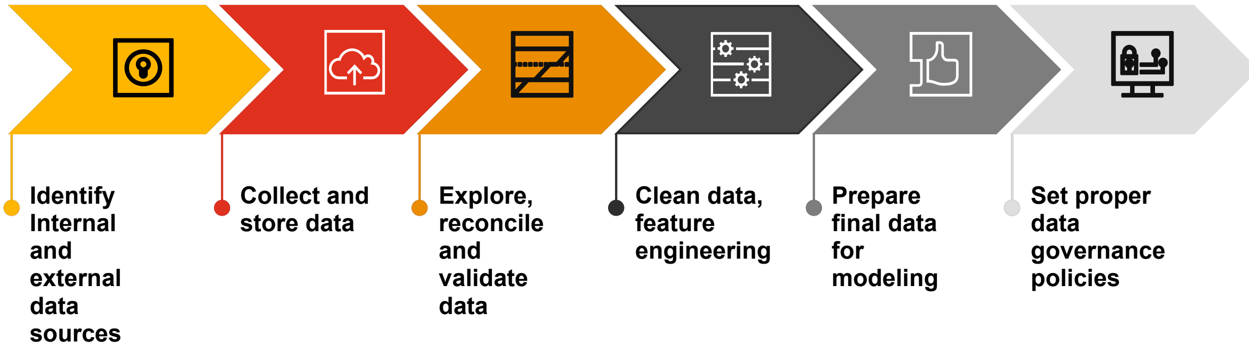
Partnership opportunities can likewise help, in terms of working with InsurTechs or through acquisitions, to augment the available data or jumpstart a particular component of the project. For smaller insurers, doing certain processes in-house might not be feasible, whereas larger insurers might find the capital outlay necessary to develop such a product is more than simply partnering. As a result, whether it be partnering with external data providers and harmonizing their data into a structure more amenable for your analysis, or accessing an external model for feature selection, partnerships can provide information that otherwise would not be feasible.

Once the plan, budgets, and talent are established, it is necessary to secure funding and get the commitment from management to pursue the next steps. We next move to turning the plan into action.

[217] Gloria Macías-Lizaso Miranda. Building an effective analytics organization. McKinsey and Company, 2018. www.mckinsey.com/industries/financial-services/our-insights/building-an-effective-analytics-organization.

[218] PwC United States. 2019 AI predictions: Six AI priorities you can’t afford to ignore. 2019. https://web.archive.org/web/20211109223336/www.pwc.com/us/en/services/consulting/library/artificial-intelligence-predictions-2019.html.

Data



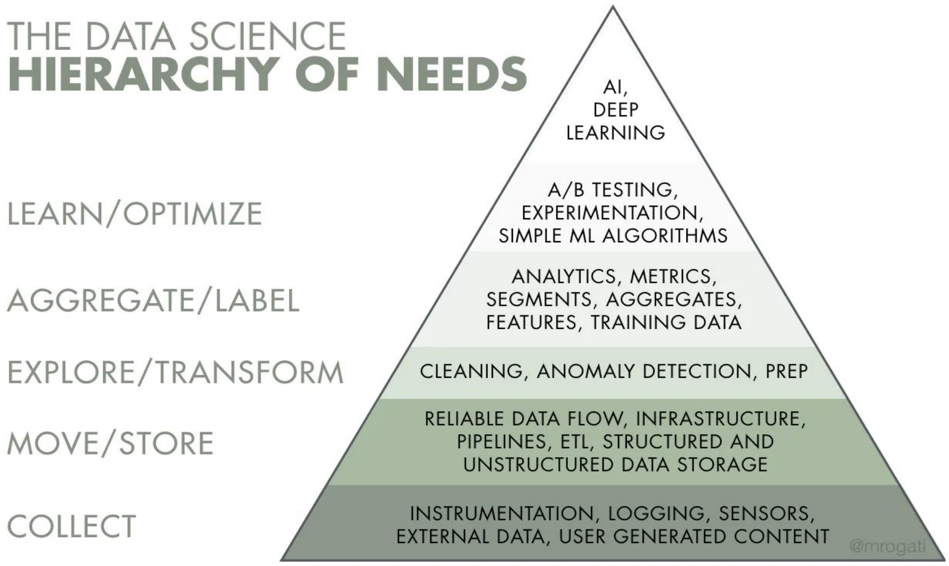
After specifying the problem and how best to tackle it, and socializing the plan, a deeper dive into the data available will need to be considered. Identifying what internal data sources are of interest, among the available data sources for the project, will be the first step. It may be worth holding a conversation with different related stakeholders about the data sources available for the current analysis to save time on the data-collection effort. It is also important to note that, due to the insurance business' nature, not all lines of business share similar volumes of data, especially commercial lines of business and reinsurance. We suggest supplementing your data by using external sources to serve the purpose of your advanced analytics projects, such as climate data and non-insurance-related incident information. The data coming in from external sources might be unstructured, such as in the form of images, JavaScript objects (JSONs), geospatial vector data, or other things. Having people that know how to access these different types of data and knowing how to manipulate them will be important in being able to use them in the subsequent analysis. In other circumstances, such as when providing a new product, data might not exist. For instance, if the insurer is planning on using telematics in its pricing models, then sensors, telematics providers, and data storage capabilities must be identified and priced. Cross-border insurers will want to consider whether foreign data are of relevance in different locales.

Data form the bedrock of your project, and will dictate how effective you can model. One analogy is to liken data science to Maslow's hierarchy of needs,[219] with data and cleaning as the base, and modeling and analytics as the apex. You cannot leapfrog best practices for basic data analytics and go directly to adopting AI and other advanced technologies.[220] The first step is making sure the data are up-to-date. Understanding what data will need to be accessed on an ongoing basis, potentially in real time, is part of understanding what is implementable.

---

[219] Monica Rogati. The AI hierarchy of needs. *Hackernoon*, 2017. https://hackernoon.com/the-ai-hierarchy-of-needs-18f111fcc007.

[220] Nick Harrison and Deborah O'Neill. If your company isn't good at analytics, it's not ready for AI. *Harvard Business Review*, 2017. https://hbr.org/2017/06/if-your-company-isnt-good-at-analytics-its-not-ready-for-ai.



Hackernoon. *The AI hierarchy of needs* [221]

Once the project has been given permission to begin, and the data have been collected, the length of history used for the project should be determined based on the goals of the project. To help modelers understand the data structure and variable distributions, EDA will be undertaken using basic statistics, looking at interesting sub-populations and visualization. As noted in the survey, larger insurers are increasingly looking to use Python and R for exploratory analysis, although this can range across business segments. For smaller insurers, or for less technical teams, the use of business intelligence tools such as Power BI and Tableau are more user-friendly and are easier for creating dashboards and identifying trends. Moving away from Excel towards tools specifically built for these purposes should help save time in getting results.

Reviewing the data structure is important, such as identifying whether tables from different sources have different formatting for the same fields, as well as having consistency in the collection process. While such concerns might be more important for reinsurers than primary personal insurance writers, for a reinsurer receiving data from many different sources, data received from insurance companies might not be uniformly formatted. For commercial insurers, data from the different industry segments may be very different, requiring significant wrangling to have a uniform shape. Errors and missing data will need to be identified, and investigations made, to see if there are systematic reasons for such occurrences. Both errors and missing data fall under the larger umbrella of data validation and reconciliation. Other components of this process involve checking the plausibility of results (such as a sensor indicating a top speed of 500 km/h), checking spelling to see if some values are in different cases or pluralized, making sure keys are unique, or checking if there are redundancies.

Once a better understanding of the data has been established, data selection criteria will be reconsidered in light of data quality and exploration. From there, data pre-processing will have to be performed as well. That is, data must be cleaned, such as reformatting dates, and special and missing values must be

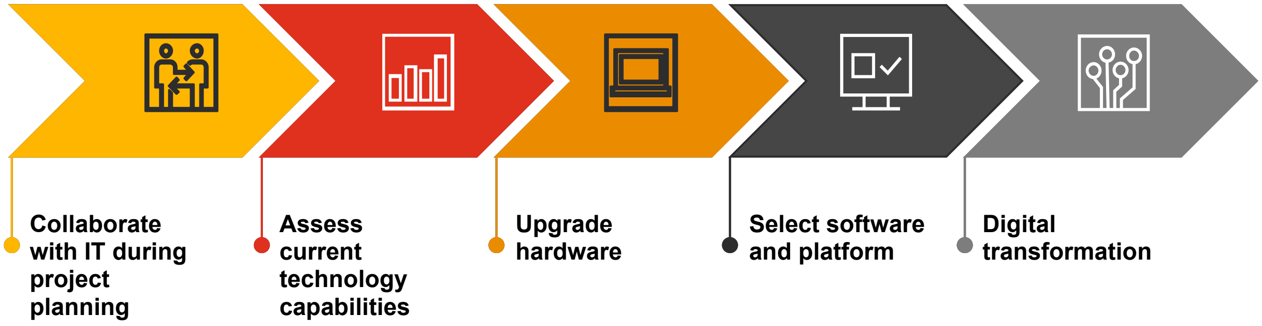
[221] Rogati. The AI hierarchy of needs.

addressed. There are many different ways to deal with missing data, such as imputation, or filtering the missing values out of the data, although such determinations depend on the amount of missing data and the quality of the other fields. From there, new features will be constructed with existing data, such as by normalizing, censoring, one-hot encoding, and performing any other transformations per both modelers' domain knowledge and modeling needs. Once the different data sources are properly formatted, joining different data sources and adjusting historical experience to reflect expected future systematic changes may be required to finalize the joined dataset for modeling use.

After the data pre-processing has been accomplished, data will be partitioned for the modeling process. One possible way of doing this is through random sampling with an 80/20 procedure, wherein models are trained on 80% of the data and models compared on the out-of-sample 20%. Of growing use is k-fold cross-validation, with models compared using the mean square error loss function.

A brief conversation on data accessibility and data protection is necessary in terms of the wider attention that must be paid to data governance. Insurers work with PII, and hence care must be paid to its maintenance, oversight, and access. Policies regarding data access and data retention are necessary requirements when dealing with sensitive information, with regular policy reviews needing to be performed and security assessed. Privacy protections and strong security measures have become more important given the escalation in phishing and other types of hacking. Ensuring that access privileges are set and that hackers are given as little information as possible when trying to infiltrate systems are important parts of minimizing data vulnerabilities. Data systems are also not immune to error, and fault-tolerant designs should be used to allow the continued use of the system should a part fail. A comprehensive data reconciliation and validation process should be set up to guide the initial check on the data completeness and accuracy. Furthermore, data dictionaries should be in place detailing the different fields in the company's databases, and should be maintained regularly to keep documentation up to date. Companies that maintain comprehensive data dictionaries make it easier for users to readily understand their data variables and perform analyses in an effective way.

Technology



Although technically part of the business plan, technology warrants its own considerations. As discussed in the *Data* section, data might be coming from external data sources or additional technology procured for a project's task.

Early discussions concerning implementation are important. Understanding IT's needs, capabilities, and timelines will help to identify early on whether IT's infrastructure can cope with the models you want to test. Understanding how to integrate models within the wider business systems is critical to the success of your project. As noted in a survey, only 45% of companies using ML have successfully deployed a model.[222]

In cases where new data will have to be sourced, solutions will have to be drawn as to the new data's storage and maintenance. Considerations regarding storage will have to be made as to how scalable the data solution is, as well as to how flexible the data solution is, should you want to add unstructured features. Determining whether to host data locally or in the cloud would often be beyond the scope of a single project. How best to store data forms a necessary part of a company's wider data and analytics strategy, as outlined elsewhere.[223, 224] Nonetheless, a larger determination of whether to host data in the cloud might be precipitated by the project's needs.

Storage requirements, technology, and data procurement could all possibly necessitate re-evaluating the initial business proposal. In all, when considering acquisition it is helpful to consider the five Vs: Volume, Velocity, Variety, Veracity, and Value.[225]

Taking the case of telematics mentioned above as an example, IT will want to be consulted regarding the data's storage, building connections with existing systems, and privacy concerns regarding the sensors collecting and transferring data. The accuracy and value of the data will also need to be assessed to establish the necessity of the project vs its costs.

As mentioned in the *Considerations and opportunities* section, discussions might be made regarding investing in upgrading existing hardware, or converting to cloud services, in order to speed up the development of your models. Dealing with big data requires paying attention to run time and scaling. As a result, code profiling, an understanding of how the algorithms work, and testing on small samples become more important. Other advantages can be found both in speeding up the data-wrangling process through specialized algorithms that take advantage of parallelization, or in specialized hardware to speed up specific algorithms. For instance, when considering the use of TensorFlow for ML, consideration should be made regarding the use of Google's cloud Tensor Processing Units (TPUs) or GPUs to speed up compute time. Regardless, the current computing power available should be assessed in order to evaluate how it will affect the model algorithm selection, modeling process, and timelines.

Use of hardware must be properly married with effective software, as otherwise the hardware will not be fully optimized. Open-source programming languages such as Python and R are fully featured in that they have extensive libraries for use in visualization, ML, and software development. The use of an integrated development environment (IDE) is recommended to help with managing your code and helping in debugging. Platforms also exist, with a number of proprietary tools created for the express purpose of modeling and visualization. Some of these platforms are low- or no-code solutions, requiring little training and having drag-and-drop environments, although they permit less flexibility and leave one at the mercy

[222] Algorithmia. *2020 State of Enterprise Machine Learning*. 2019. https://info.algorithmia.com/hubfs/2019/Whitepapers/The-State-of-Enterprise-ML-2020/Algorithmia_2020_State_of_Enterprise_ML.pdf.

[223] Paul Livak. 4 components to a modern insurance data strategy. *Digital Insurance*, 2020. www.dig-in.com/opinion/pwc-insurance-4-key-components-to-a-data-strategy.

[224] Sloan Plumber and Scott Busse. How insurance carriers are modernizing to cloud based analytics. LinkedIn (article), PwC Advisory United States, 2020. www.linkedin.com/pulse/how-insurance-carriers-modernizing-cloud-based-analytics-sloan-plumer/.

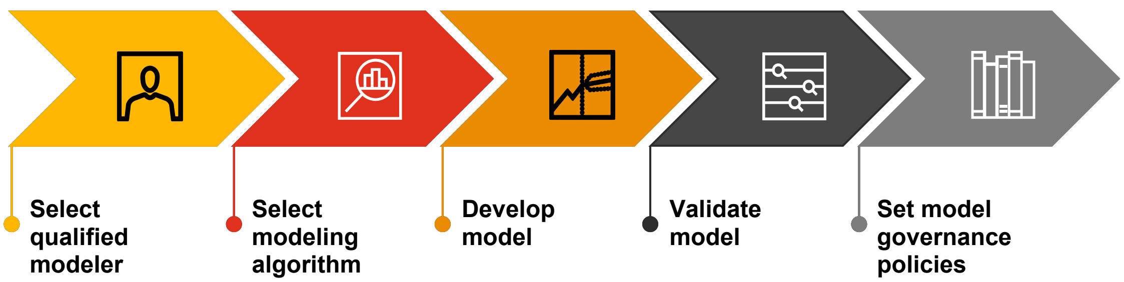
[225] Teradata. What are the 5 V's of Big Data? Accessed March 1, 2022. https://www.teradata.com/Glossary/What-are-the-5-V-s-of-Big-Data.

of developer updates to enable newer techniques. Such platforms will be most useful for teams with less coding experience.

In addition to supporting core business models, having a digital transformation process will further enable implementation of your advanced analytics application. During the pandemic, many insurers were forced to accelerate their digital plan, and this has led to additional touchpoints and data opportunities. For instance, in claims solutions the ability to capture images could be created, whereas when dealing with churn an exit survey could be constructed.

Taking advantage of these new sources and platforms enables enterprise-wide social, mobile, and web solutions. That is, by having this additional infrastructure, it will be easier to gain new droves of data for many common business cases.

Modelling



With the data collected and cleaned, and with additional features made, the initial modeling techniques will be identified. The type of modeling problem will already have been outlined in earlier planning stages, as well as model implementation restrictions assessed related to both IT and regulatory requirements. For instance, should you be working with data that evolve over time, time series techniques might be used to try and identify changes over time. Understanding the problem using a priori beliefs will also inform the modeling techniques to be used in, for instance, choosing between linear and non-linear models. In order to have some expectations of how the model should perform, staff will need to understand how the variables might interact. In addition to examining how a model might work, investigating embedded assumptions is also important. For instance, when dealing with frequency data, use of the default Poisson model might not be appropriate, and a zero-inflated model might make more sense.

The business nature and data volume are important considerations when selecting the best model algorithm. For instance, when there are few data, such as in commercial large accounts, one is restricted in the techniques available due to a lack of training data. In such situations, the use of traditional methods and expert judgment are of greater importance than modeling acumen. Understanding when data can be pooled together to form credible datasets is another important consideration. For large personal auto insurers, such problems of data volume should be of less concern, giving greater freedom to the techniques at their disposal.

Usually, more complex models require more time to develop, whether due to the need for greater volumes of data or due to the longer time it takes to train the model. Therefore, time and data volume constraints must be considered. Part of this process would involve parameter tuning, which can be arduous. One reason an ML model might be chosen over a more traditional one is that parameter tuning can be fully automated. A recent development is the use of automated machine learning (AutoML), which can automatically generate features and tune your hyperparameters.

An insurer might consider how fancier models would perform, even if not implementable. This could give insight into how much lift is being unused with the current data due to structural limitations. Taking the example of the time-dependent data, checking the effectiveness of Long Short-Term Memory (LSTM) networks[226] compared to your current model could indicate that there is additional signal that could be captured, although it would require being able to deploy a recurrent neural network.

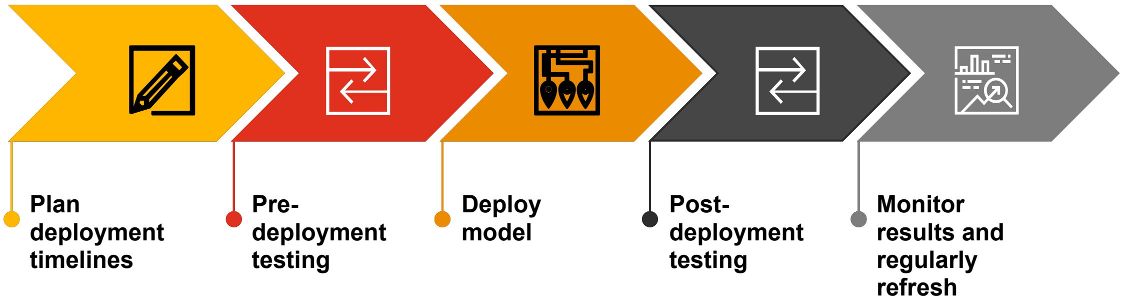
Once the testing stage has finished, the predictive power of the viable models will have to be compared with the expectations as set out by the business plan. To do so, the models chosen must be validated by comparing results using the test dataset across different metrics. What metrics are chosen is dependent on the type of problem being solved, with classification and regression problems being substantially different. For instance, double lift charts and Gini coefficients are very common metrics among survey respondents, with root mean squared error (RMSE) and cross-validation error also discussed. For classification problems, the receiver operating characteristic (ROC) curve is popular. The necessary model performance should already have been defined in the planning stage as a means to gauge the financial impact of the model.

Considerations regarding responsible AI and explainability must also be considered after the number of viable models has been whittled down, balancing the level of transparency with predictive power. Model-agnostic methods have been developed to deal with black-box methods, in the hope of making them more understandable. Understanding which variables are most important in explaining the model, and seeing whether they are necessarily correlated with the results, would prompt further investigation. Questions on how to explain your model and why a prediction is made are of the utmost importance in convincing stakeholders that the model is worth implementing.

Finally, once the model is chosen that fits within the IT infrastructure, the quality of the testing regime will have to be disclosed and documentation made. Setting model governance policies is the last, but not least, part in developing an advanced analytics model. The policies outline the roles and responsibilities of each stakeholder involved, control model access and versions, formalize the modeling process, and manage model risk. Version control is crucial in model governance in order to track changes to the codebase, facilitate team collaboration, back-up files, and document assumptions and parameters. This also allows the branching of the code for different users while the production code remains unchanged, and it also makes it easier to identify possible errors through the log files. Git is an especially popular implementation, with repositories set up either online or on a local drive.

[226] Christopher Olah. Understanding LSTM networks. Colah's blog (GitHub), 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

Implementation



The final part of the journey involves integrating the model with the existing IT infrastructure. Part of this is understanding the lifecycle of the model, versioning, scalability, and monitoring.

In a traditional software world, versioning code is enough, because all behaviour is defined by it.[227] In ML, model versions also have to be tracked, along with the data used to train the model, as well as meta-information such as training hyperparameters.

Depending on the use case and line of business, regulatory approval may be required. Should a novel technique be used, consultation with the regulator becomes more important and extensive conversations should be done at an earlier stage to increase the likelihood of approval. For instance, when dealing with a private-passenger auto-pricing model, extensive discussion will be necessary to justify whether the pricing is fair and the dislocation reasonable. The need for model explainability highly depends on the use case. For targeted advertising, or internally predicting churn, the repercussions of unexpected errors are acceptable as long as the model effectively targets the right segment of customers or predicts retention, respectively, on the whole.

After model governance and regulatory considerations, the model will be deployed to a target environment to serve its role. This deployment can be one of the following:[228]

- Microservices with a REST API to serve online predictions
- An embedded model to an edge or mobile device
- Part of a batch prediction system

Extensive pre-deployment testing must be done to minimize issues with the planned roll-out. This involves following the different stages of the software development lifecycle, with each component ("unit") of the application tested first and later user acceptance testing (UAT) performed. A testing environment must be configured, and a set of user acceptance test libraries need to be created. These test libraries contain a set of tests that are executed with pre-specified data. That is, it will be necessary to check whether the model that was deployed to the test environment produces the same results as expected. Acceptance test

[227] Cristiano Breuel. ML Ops: Machine learning as an engineering discipline. *Towards Data Science*, 2020. https://towardsdatascience.com/ml-ops-machine-learning-as-an-engineering-discipline-b86ca4874a3f.

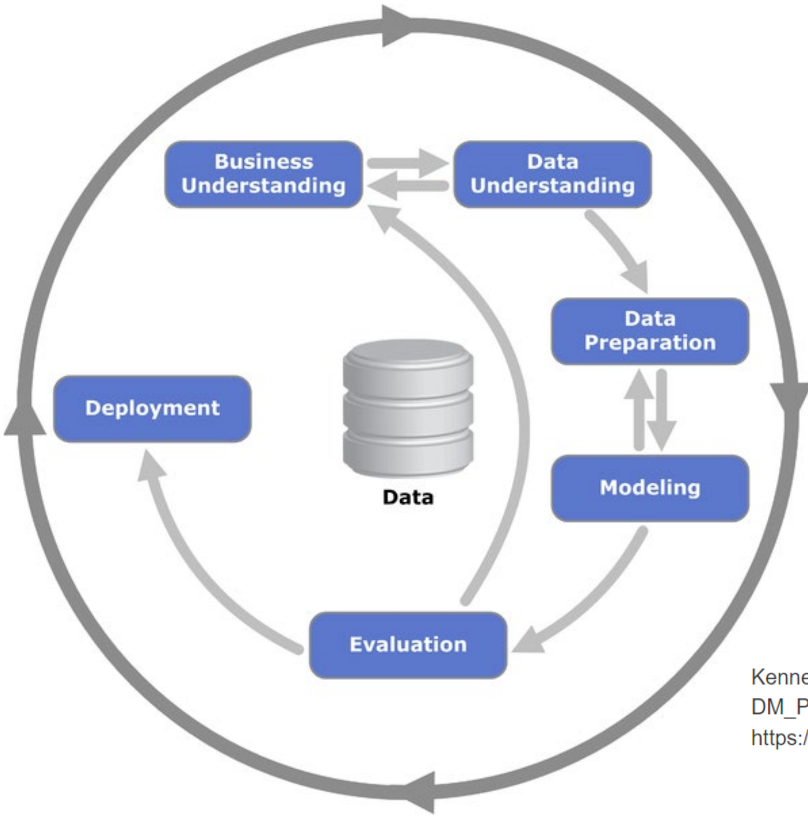
[228] Google Cloud. MLOps: Continuous delivery and automation pipelines in machine learning. 2020. https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning.

libraries should typically contain real end-to-end user journeys through the system.[229] Tests will either be done manually or via automation, depending on the nature of the test. Once there has been adequate testing, the solution can be deployed into production.

After the deployment is completed, post-deployment testing must be performed. The quality assurance team will have to run tests again to make sure there are no issues in the production deployment. Should any issues arise, a back-up should quickly be restored to minimize damage.

Once the model has been deployed, comparing your expected outcomes with actual results will help to validate the model results with actual experience. After an initial monitoring period, the results of your implementation should be shared with stakeholders, and a post-mortem of the process should be undertaken to identify how to better improve processes in the future. Models will also degrade over time, with market competition changing the effectiveness of your model. That is, the model will have to be tested for drift or bias in its predictions. For instance, it will be important to track your model against different segments of your data, such as protected classes, to make sure that the model is not indirectly biased. Moreover, ongoing monitoring should be performed to test for model performance with periodic updates of the model and its parameters done in accordance with results from the monitoring procedure. The ability to quickly update and maintain model performance is important in guaranteeing that the model does not become a liability. Accordingly, having shorter turnaround times for updating critical models is a large priority for any organization. Increased automation by taking advantage of RPA possibilities and optimizing operational efficiency through DevOps should help to accelerate model refreshes and save unnecessary manual labour.

[229] Google Cloud. DevOps tech: Continuous testing. Accessed September 14, 2021. https://cloud.google.com/architecture/devops/devops-tech-test-automation.



Kenneth Jensen (https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png), „CRISP-DM Process Diagram“, https://creativecommons.org/licenses/by-sa/3.0/legalcode

Wikimedia Commons. *Phases of the CRISP-DM reference model*[230]

Lastly, as noted in the diagram, steps do not necessarily proceed in a linear fashion. Lessons learned during one part of the process might precipitate a review of earlier business questions.

Regardless of which framework is chosen, codifying and formalizing a framework into a company-wide approach will help structure your projects and make it easier for them to be replicated.

[230] https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png

Conclusion

Big data and the rapid development of computing power are affecting the way companies are finding new insights and how consumers are interacting with technology. The storage of sensor data, documents, and other forms of unstructured data is opening up new opportunities for companies to mine additional information to support management in making better business decisions. Companies are also having to navigate changes to work environments, and to consumer preferences and behaviours resulting from a global pandemic. For P&C insurers, this means major shifts in technology implementations and placing a greater focus on quick deployments and agile problem-solving.

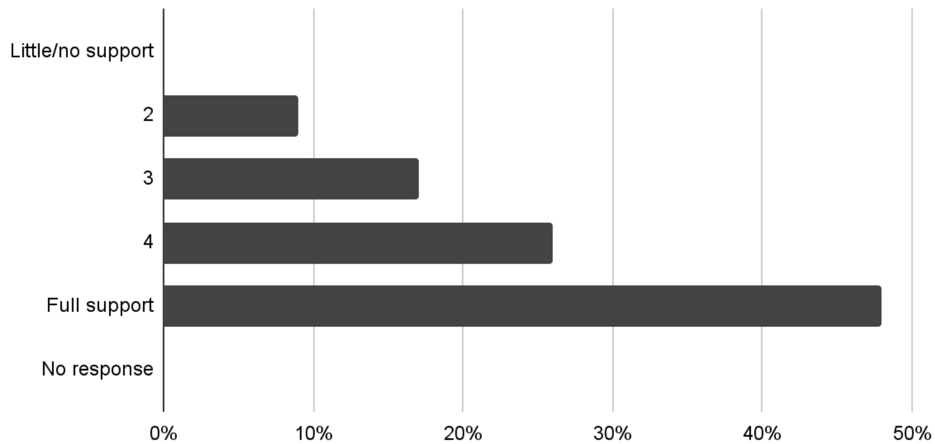
The Canadian P&C insurance industry is making significant investments in upgrading infrastructure, illustrating the foresight of the industry's collective management to ensure insurers are able to take advantage of the advances in predictive analytics, ML, and AI. Transformations are time-consuming, but benefits will be reaped over a long-term horizon, with these changes necessary as part of a P&C insurer's digitization efforts, to enable new data source connections, and to help secure new and existing revenue streams.

The goal of the study was to investigate and benchmark the current practice of predictive analytics, ML, and AI within the Canadian P&C insurance industry. To this end, we found that Canadian P&C insurers are willing to experiment with many methods in the traditional areas of pricing and reserving, in line with their peers abroad; however, with regard to their marketing and customer engagement initiatives, they lag behind. Compared with other financial services, Canadian P&C insurers have historically had far fewer touchpoints and as such have less experience in dealing with large amounts of data. To remedy this problem, Canadian P&C insurers are increasingly looking to hire data scientists and other ML specialists to bridge the knowledge gap, although less is being done to upskill current employees. Nonetheless, much progress is being made with existing analytics applications and in developing new ones as Canadian P&C insurers are eager to experiment with new techniques and optimize existing processes.

Despite the challenges currently facing the Canadian P&C insurance industry, insurers are forging ahead and taking advantage of the new growth opportunities afforded by the changing insurance landscape.

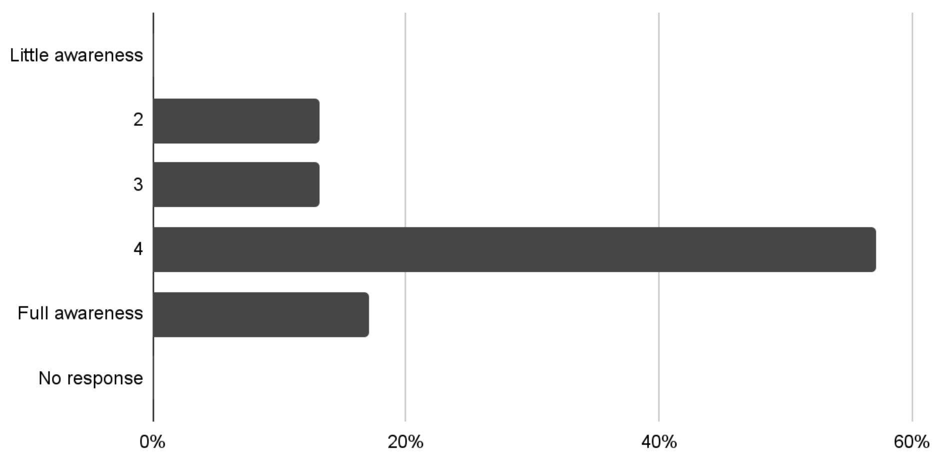
Appendix 1 – Survey result summary

Q2.1. What level of support does the senior leadership in your organization provide in developing advanced analytics?



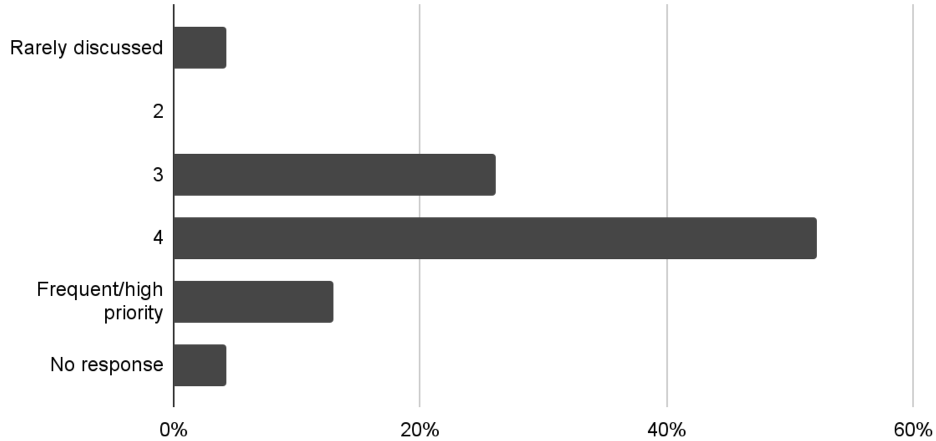
Advanced analytics is fully supported by senior leadership in almost half of the companies surveyed.

Q2.2. What level of awareness does the senior leadership in your organization have about advanced analytics?



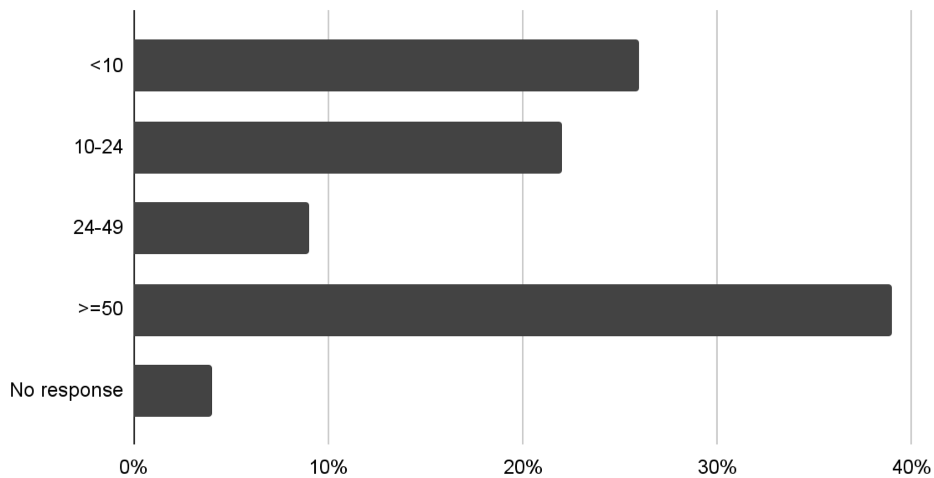
Most responders report that there is good, but not full, awareness of advanced analytics among senior leadership.

Q2.3. What is the presence and priority of data/advanced analytics on the executive meeting agenda?

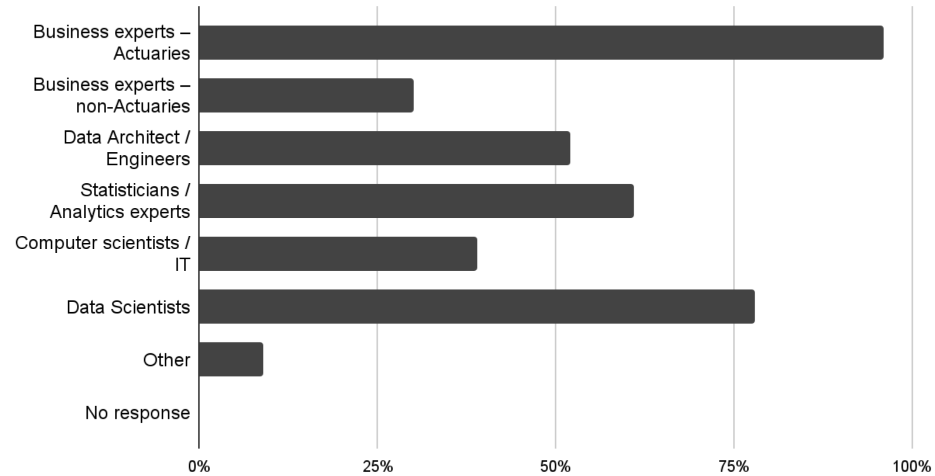


Presence and priority of data/advanced analytics is similar to the level of awareness (Q2.2) of advanced analytics among senior leadership.

Q2.4. How many FTEs are working on advanced analytics in your organization?

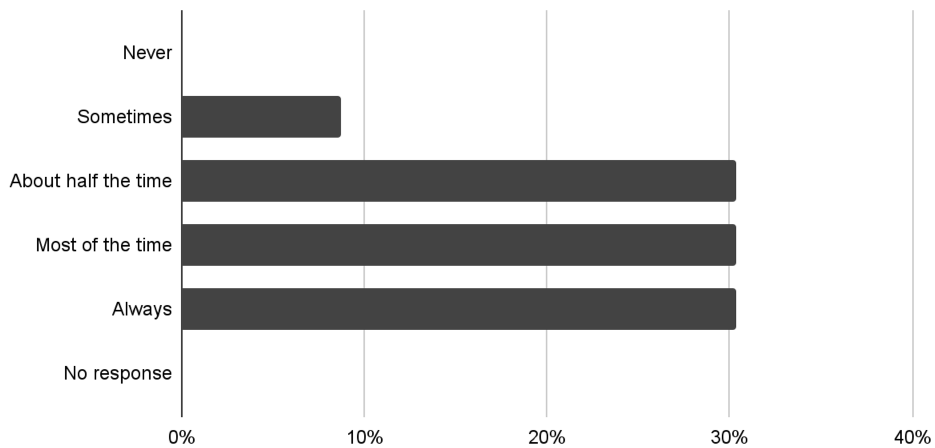


Q2.5. What types of FTEs are working on advanced analytics (select all that apply)?



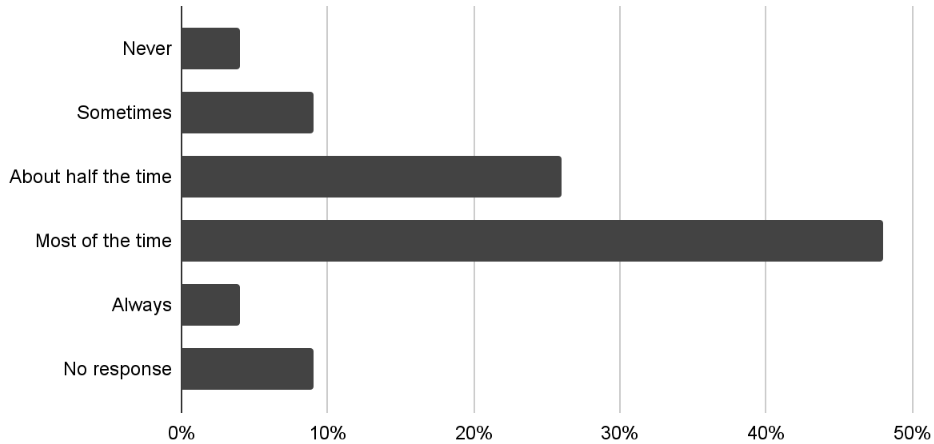
Actuaries and data scientists tend to work in advanced analytics. Larger insurers usually build larger advanced analytics teams with a wider mix of talent.

Q2.6. What is the level of collaboration between technical teams who develop advanced analytics and end-users within your company? (%)



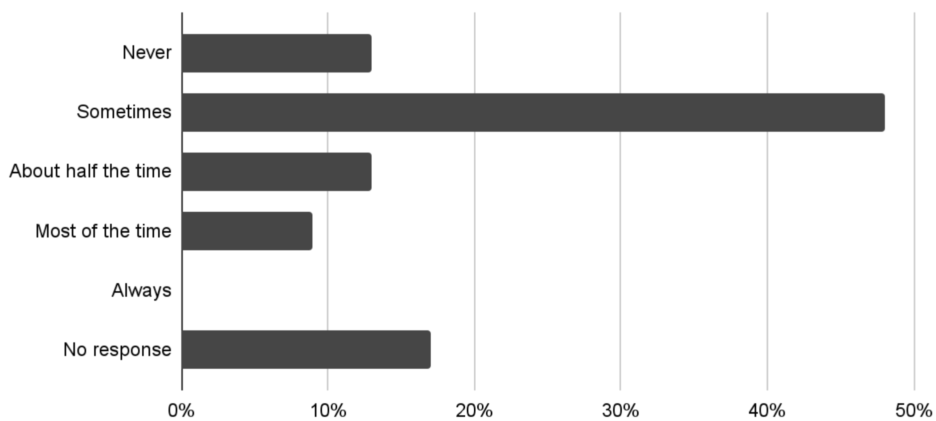
Developers of advanced analytics fully collaborate with end-users in only a third of the survey participants. There is room for increased collaboration/shift to end-user focus.

Q2.7. How often are the results from advanced analytics considered when business decisions are made? (%)



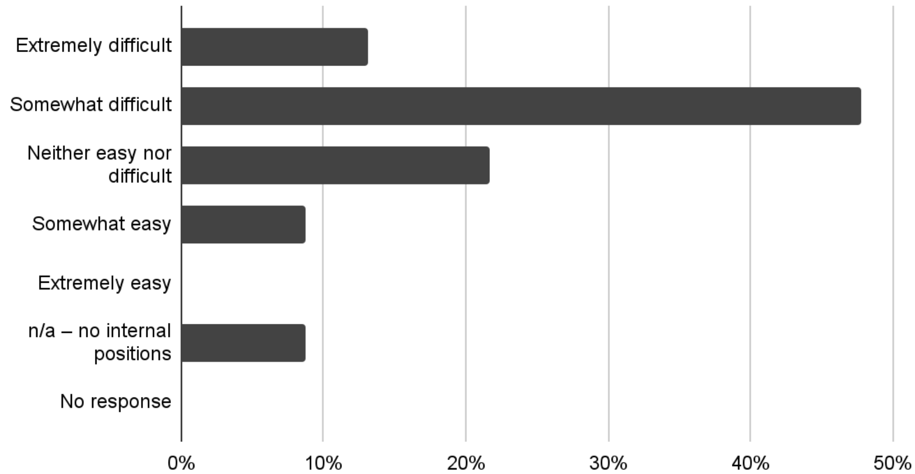
Two-thirds of the survey participants state that advanced analytics is considered in decision-making always or most of the time, but only 5% state that it is always considered.

Q2.8. When deploying a new analytics application that has not been previously executed (new users of GLMs), to what extent would your organization leverage external consultants?



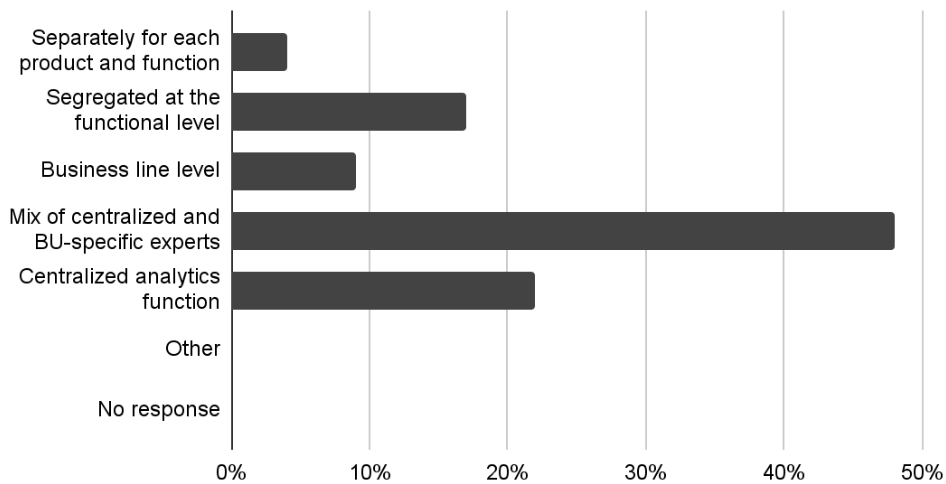
Development of analytics applications is mostly done internally, as only one in four survey participants leverage external consultants half the time or more.

Q2.9. How difficult is it to hire and retain analytics experts for internal positions?



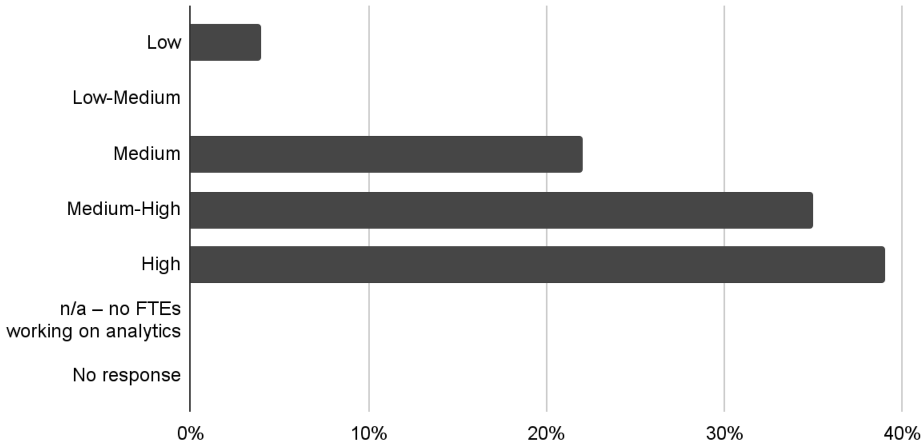
Most companies find it difficult to hire and retain analytics experts for internal positions, but for small insurers the situation is even more difficult (half say that it is "extremely difficult").

Q2.10. How are your analytics capabilities organized throughout your organization?

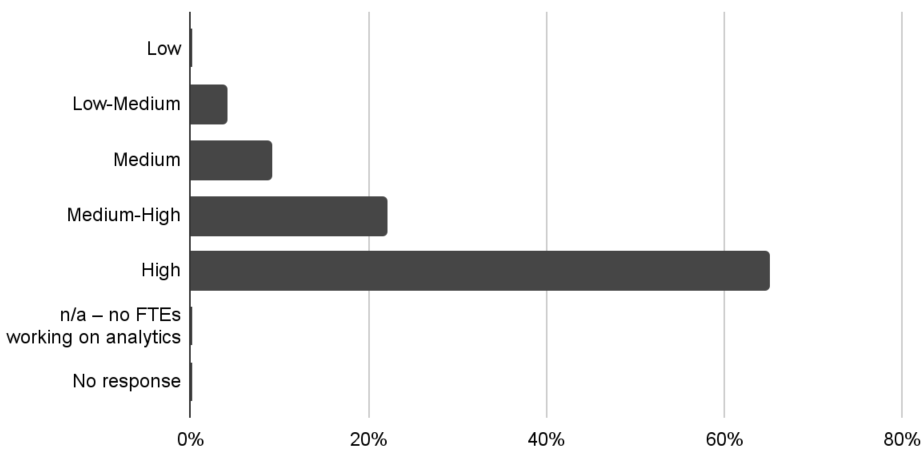


Only one in five of the respondents state that they have a centralized analytics function.

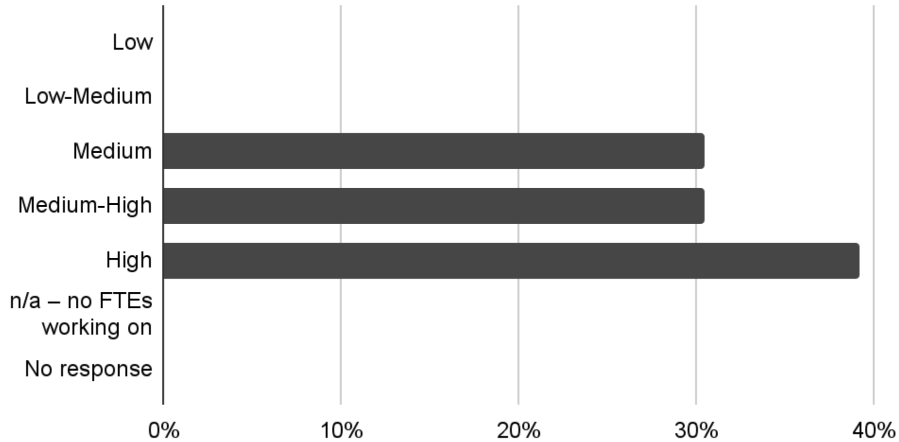
Q2.11a. How would you rank the following aspects of your analytics talent? – Technological capabilities



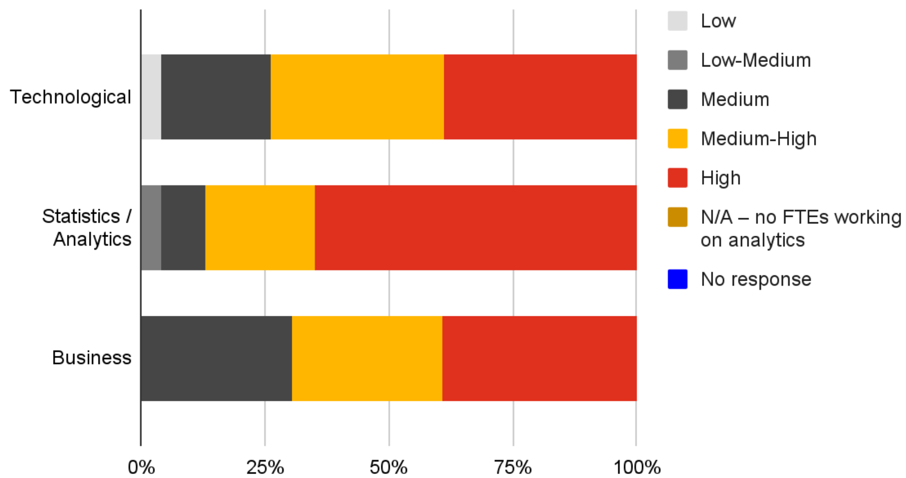
Q2.11b. How would you rank the following aspects of your analytics talent? – Statistics/Analytics knowledge



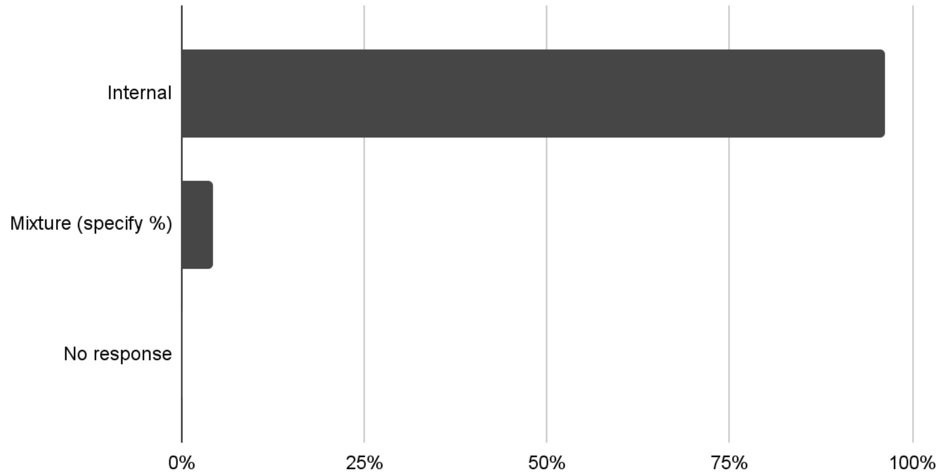
Q2.11c. How would you rank the following aspects of your analytics talent? – Business knowledge



Q2.11. How would you rank the following aspects of your analytics talent?

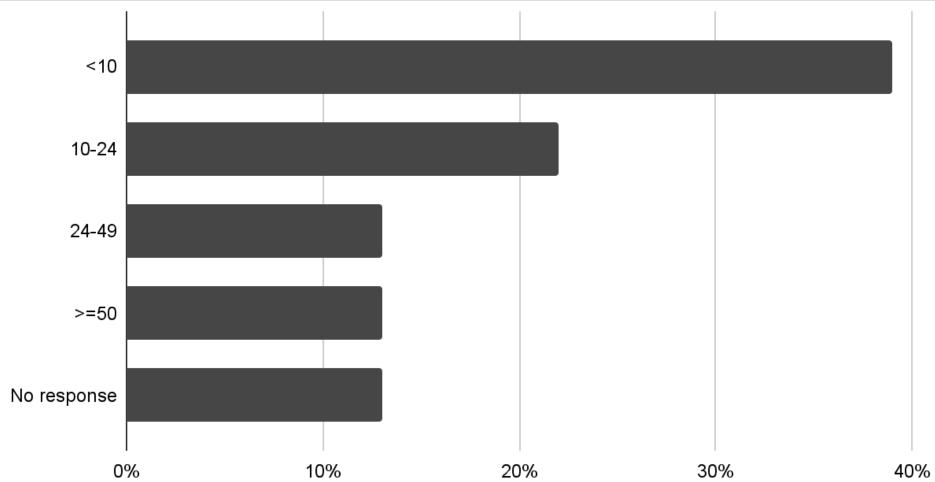


Q3.1. Is the database maintained by internal or external staff?



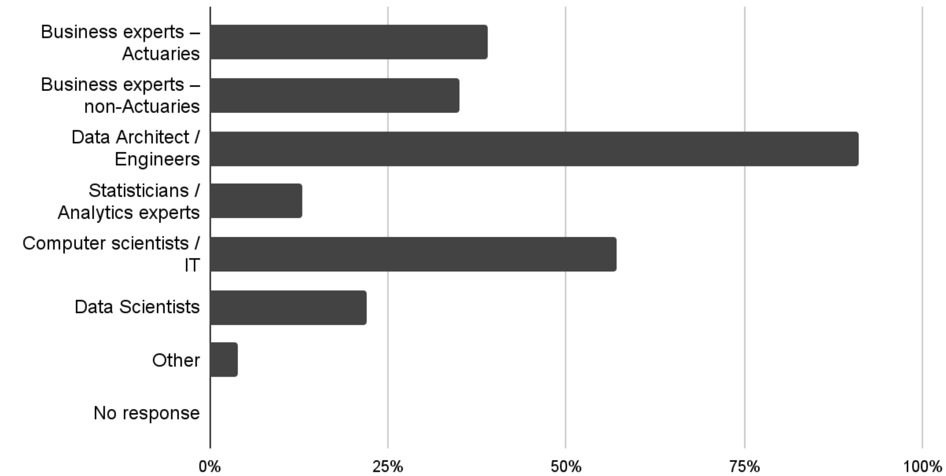
Most of the respondents maintain their databases internally.

Q3.2. How many FTEs are maintaining the databases?



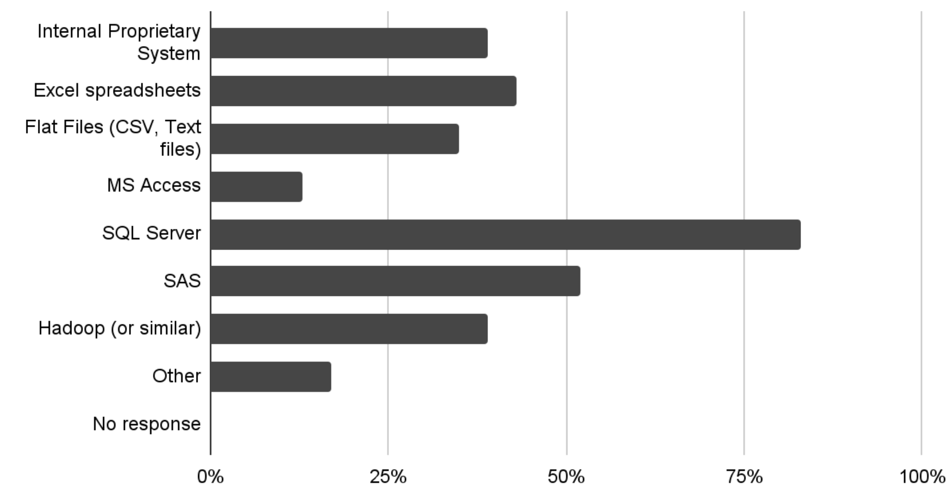
Almost half of the respondents use fewer than 10 FTEs to maintain their databases.

Q3.3. What types of FTEs are maintaining the databases (select all that apply)?

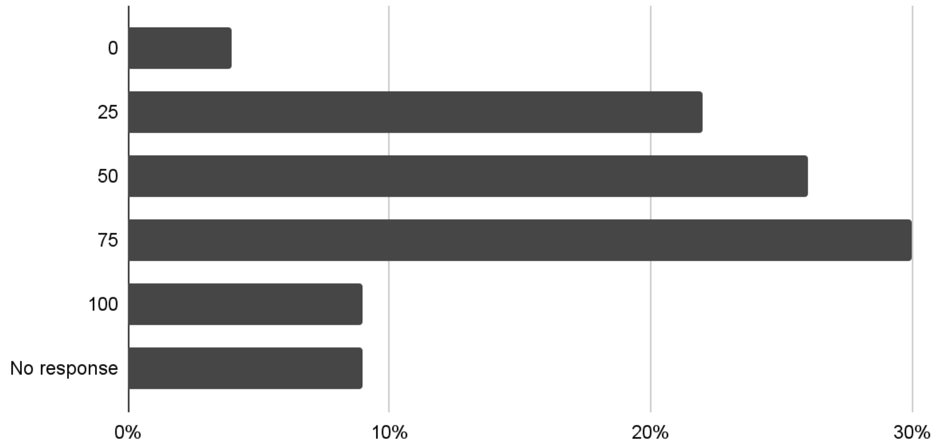


Databases are maintained mostly by data engineers and IT.

Q3.4. What are the platforms used to store data (select all that apply)?

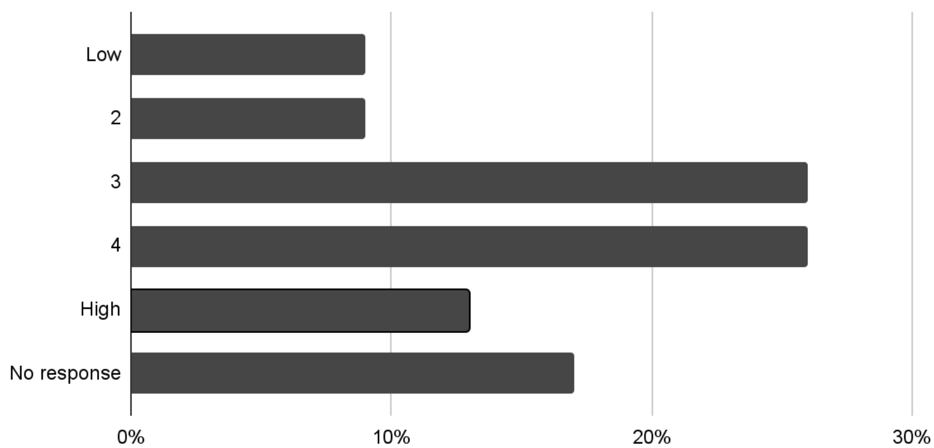


Q3.5. What percentage of your data has been defined in data dictionaries or metadata? (%)

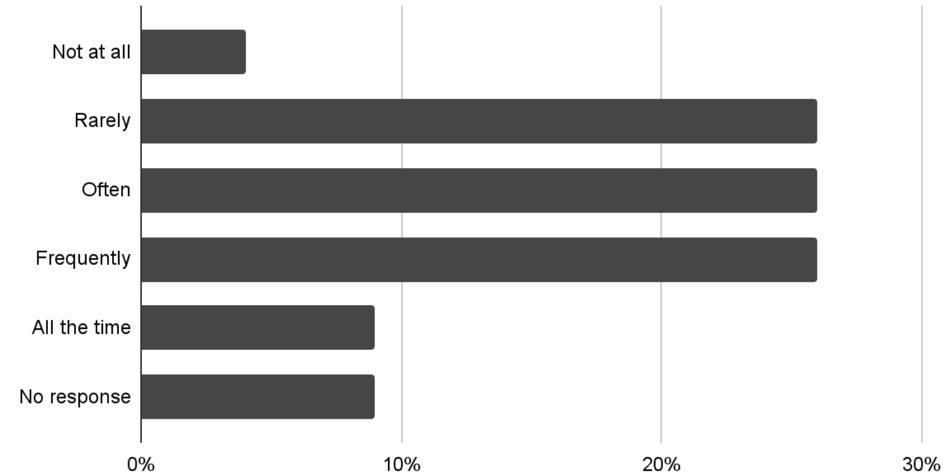


The majority of the respondents state that fewer than half of their data have definitions in data dictionaries or metadata. The situation is worse in Canadian companies compared to branches.

Q3.6. How would end-users rate the accuracy and completeness of the data dictionaries or metadata?

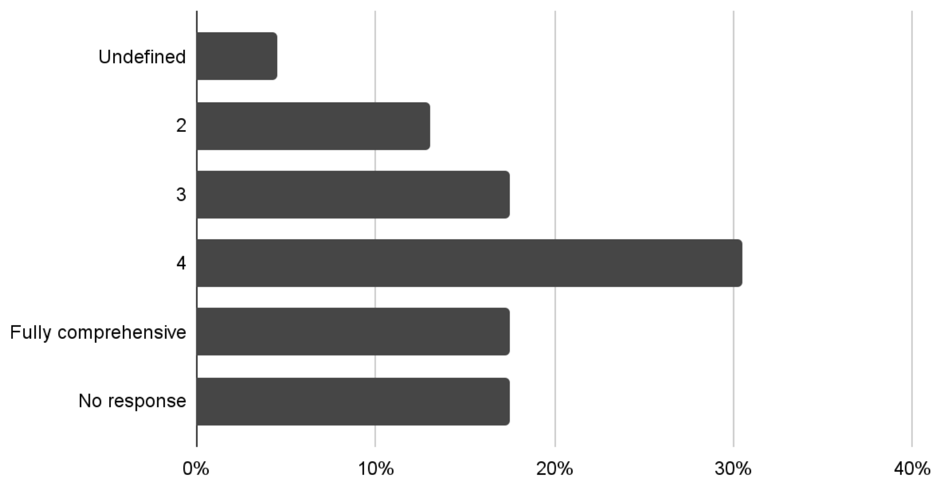


Q3.7. How often are the data dictionaries or metadata maintained?



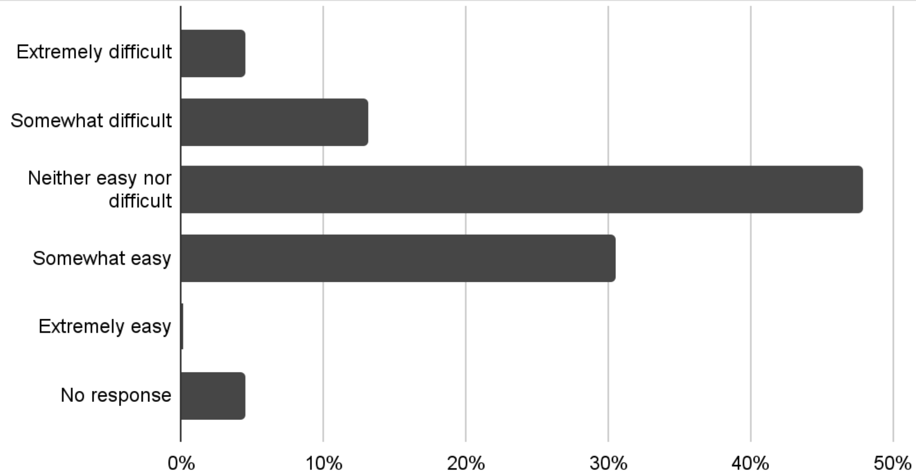
A third of the respondents state that their data dictionaries and metadata are rarely maintained.

Q3.8. How comprehensive are the policies that outline data accessibility?



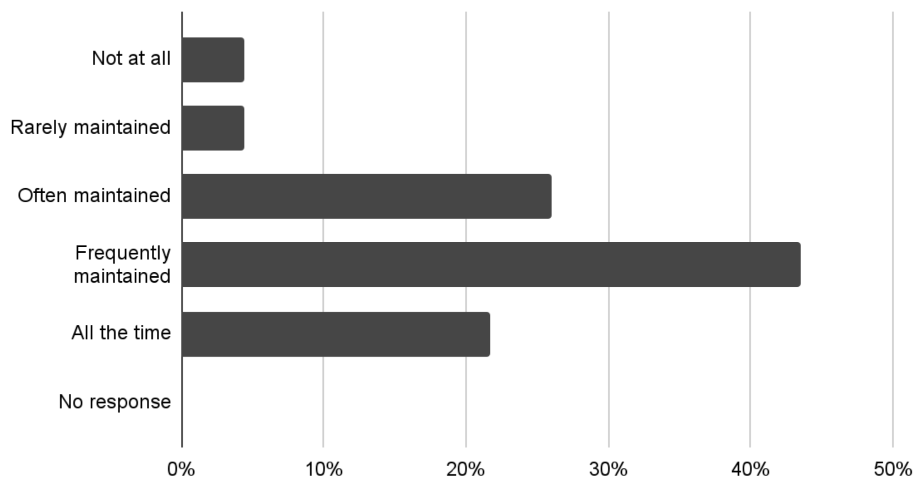
More than half of the respondents state that their data access policies are quite comprehensive. The situation seems to be better at branches and smaller companies, whereas large insurers have less comprehensive data policies.

Q3.9. How time-consuming is the process to gain access to data?



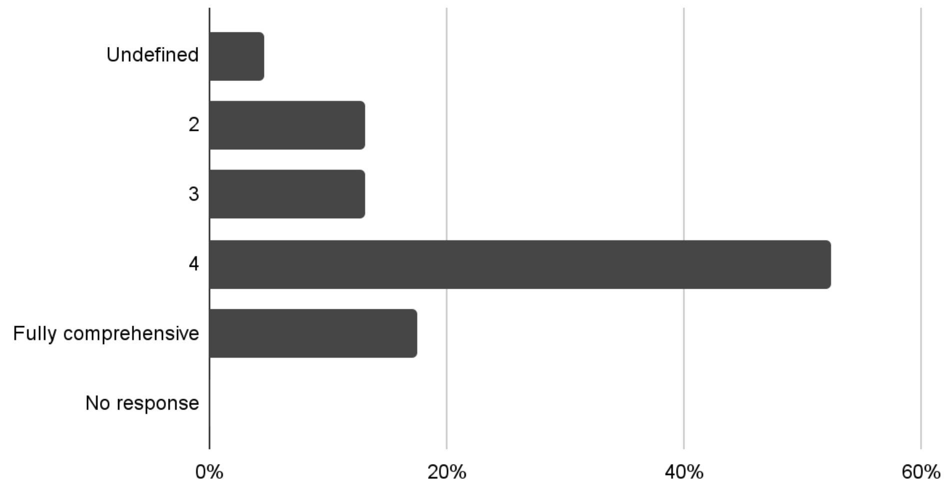
Generally, few respondents find difficulties in gaining access to data.

Q3.10. How often are the data reconciliation processes conducted? (%)



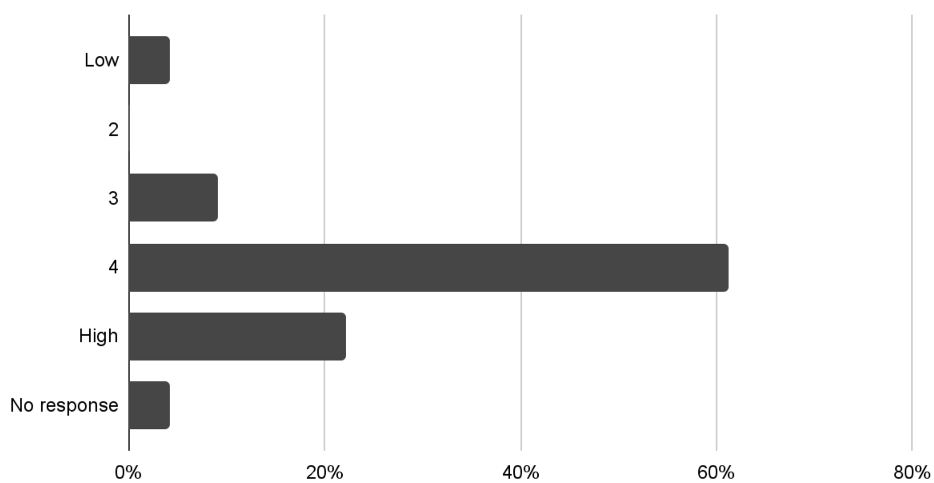
Data reconciliation is performed at least often in more than 90% of the companies surveyed.

Q3.11. How comprehensive are the data reconciliation processes?



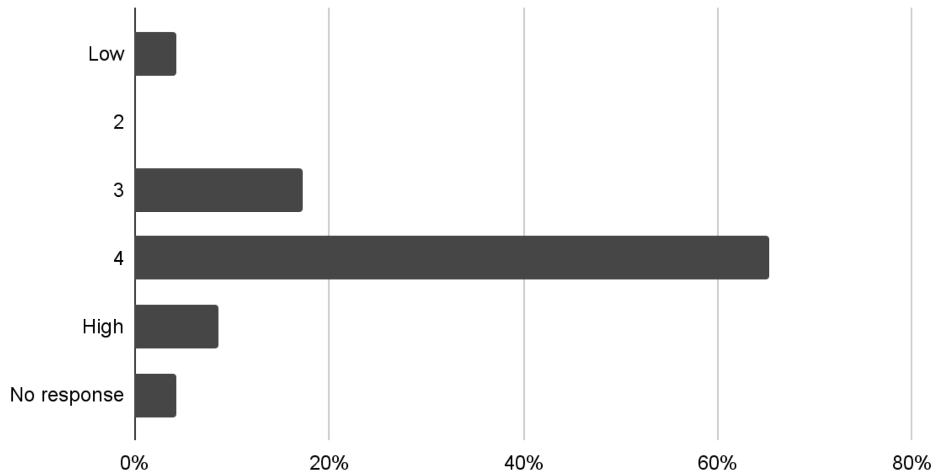
More than two-thirds of the respondents state that their reconciliation process is very to fully comprehensive.

Q3.12. How would end-users rate the completeness of your data?



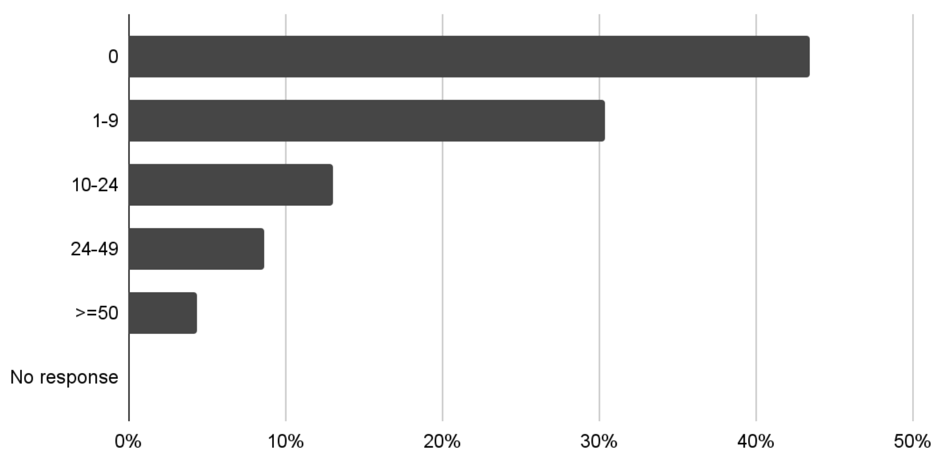
Almost 90% of the companies surveyed have high levels of data completeness.

Q3.13. How would end-users rate the accuracy of your data?

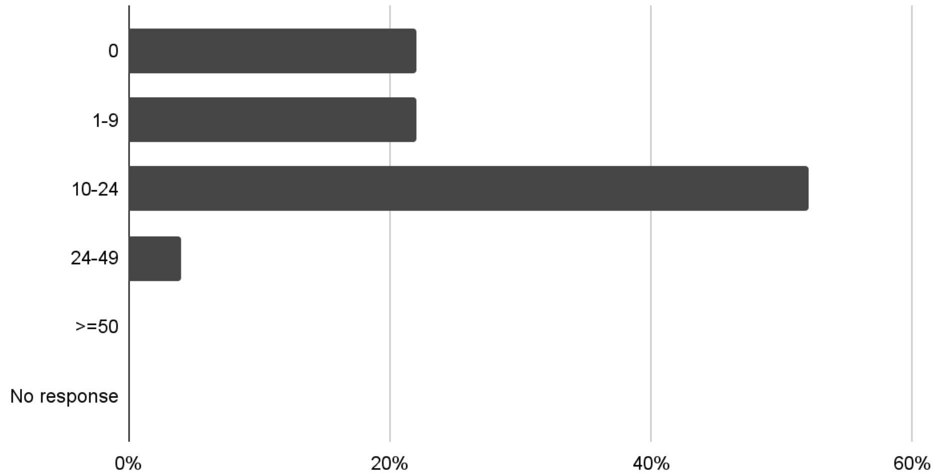


Almost 90% of the companies surveyed have high levels of data accuracy.

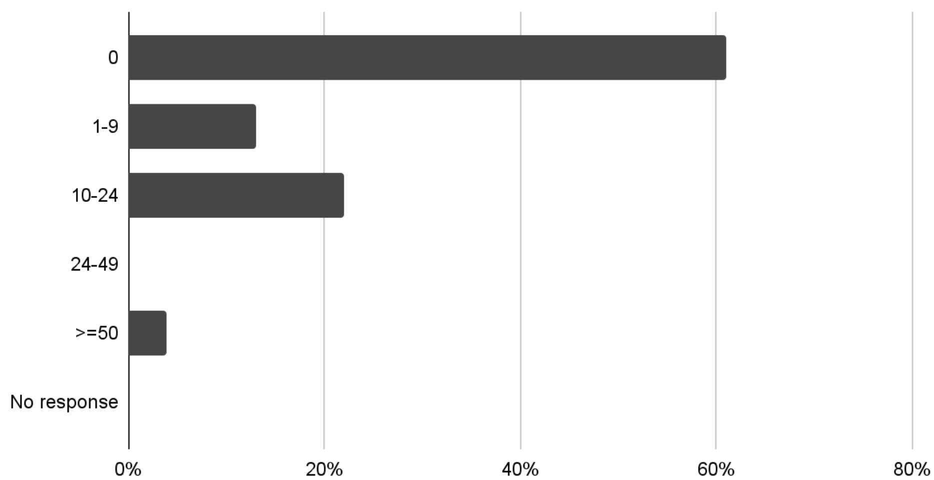
Q3.14a. What proportion of your data is unstructured – voice, images, scanned documents (pre-OCR)? (%)



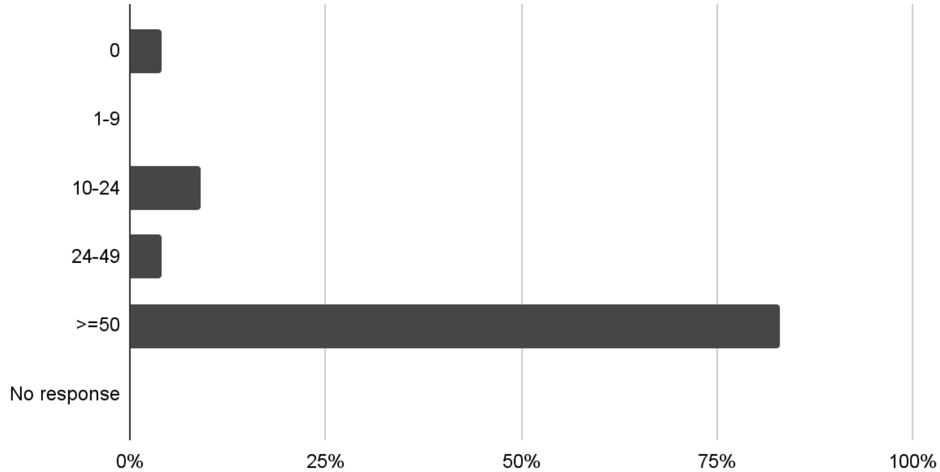
Q3.14b. What proportion of your data is unstructured – Free-form text? (%)



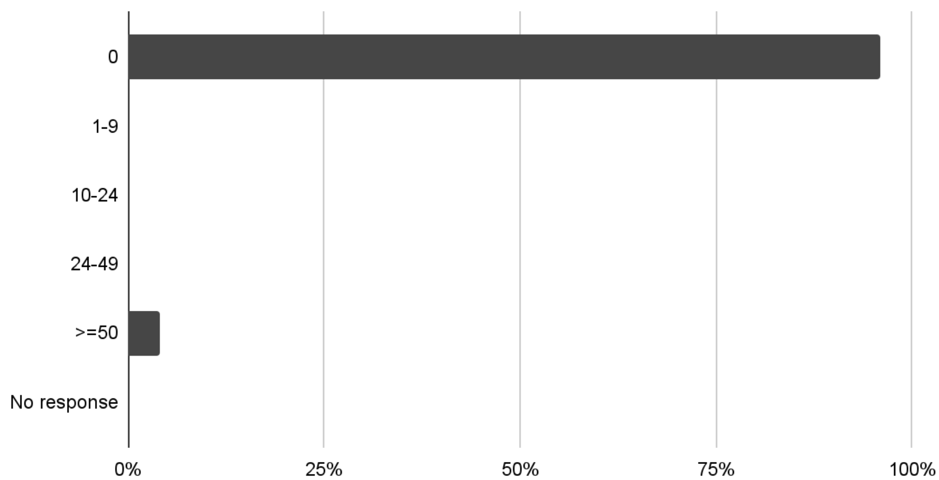
Q3.14c. What proportion of your data is semi-structured – JSON, XML, array? (%)



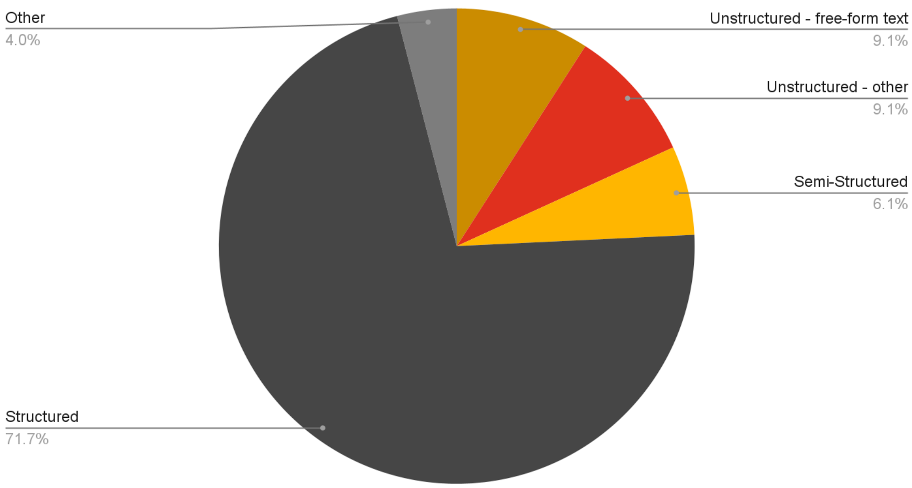
Q3.14d. What proportion of your data is structured data – tables, cubes? (%)



Q3.14e. What proportion of your data is another format? (%)

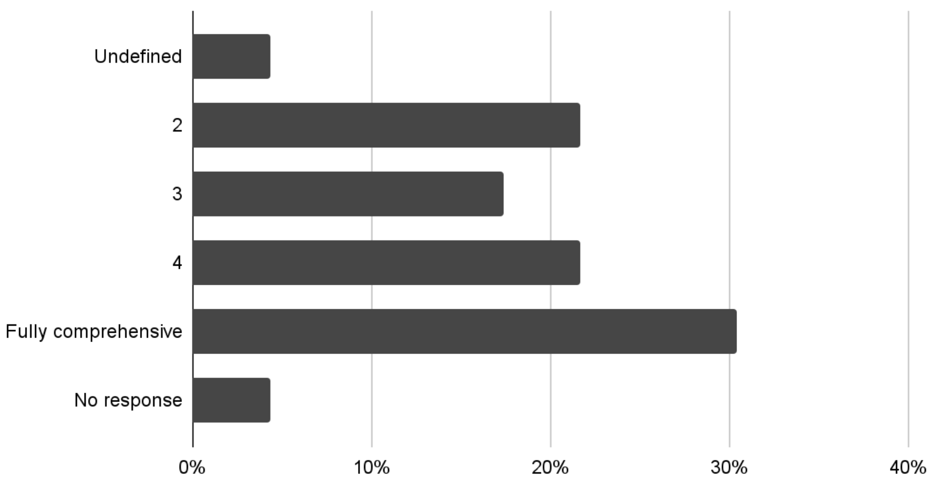


Q3.14. Data formats used by the industry



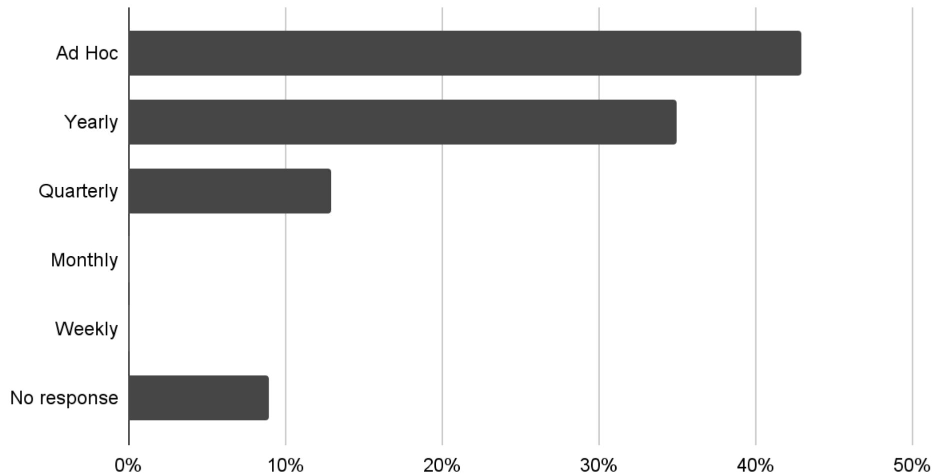
Three-quarters of the data available to insurers are either semi-structured or structured.

Q3.15. How comprehensive are the data governance/oversight/process policies?



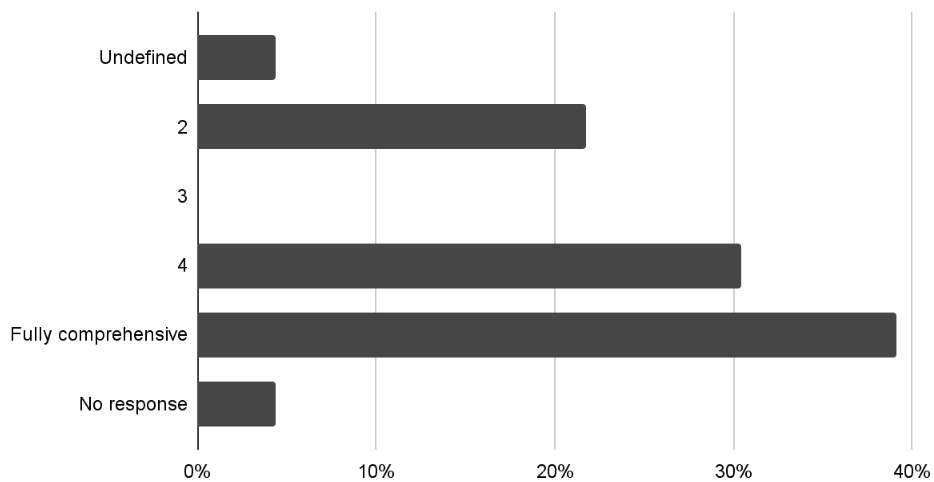
Only half of the respondents state that their data governance/oversight/process policies are very or fully comprehensive.

Q3.16. How often are the data governance/oversight/process guidelines reviewed?



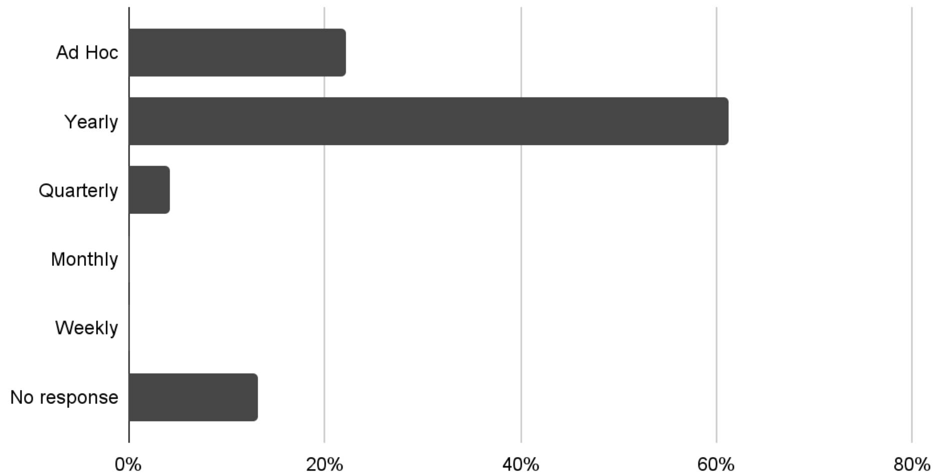
Almost half of the respondents state that they review their data policies on an ad hoc basis.

Q3.17. How comprehensive are the data privacy and retention policies?



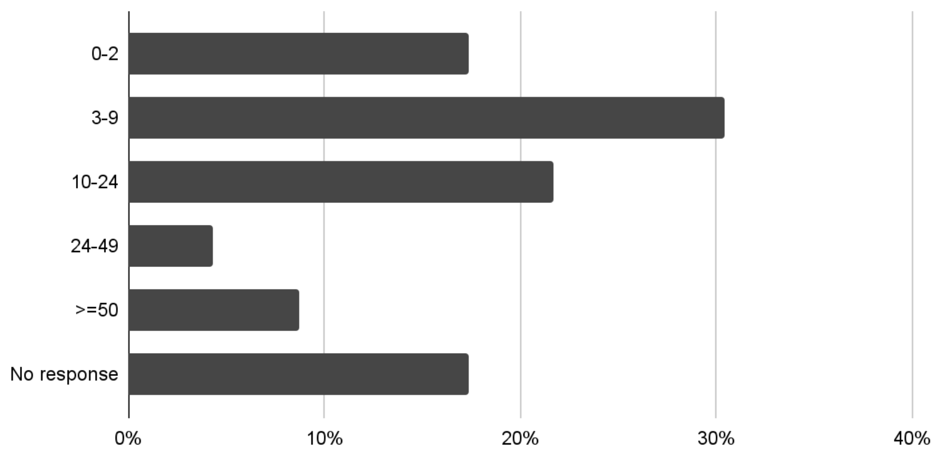
Three-quarters of the respondents have comprehensive data privacy and retention policies.

Q3.18. How often are the data privacy and retention policies reviewed?



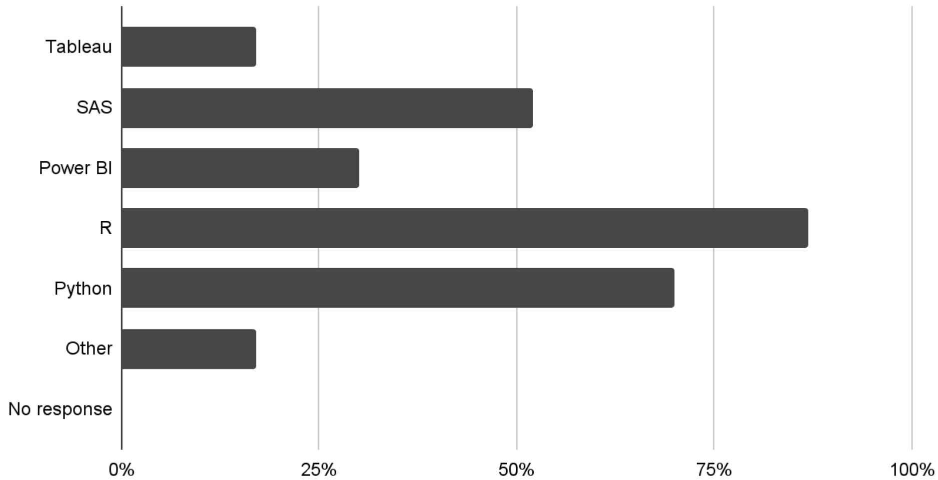
Data privacy and retention policies are reviewed annually or on a more frequent basis in most companies.

Q3.19. How many data sources does your organization utilize regularly to perform analytics that will impact your business decisions?



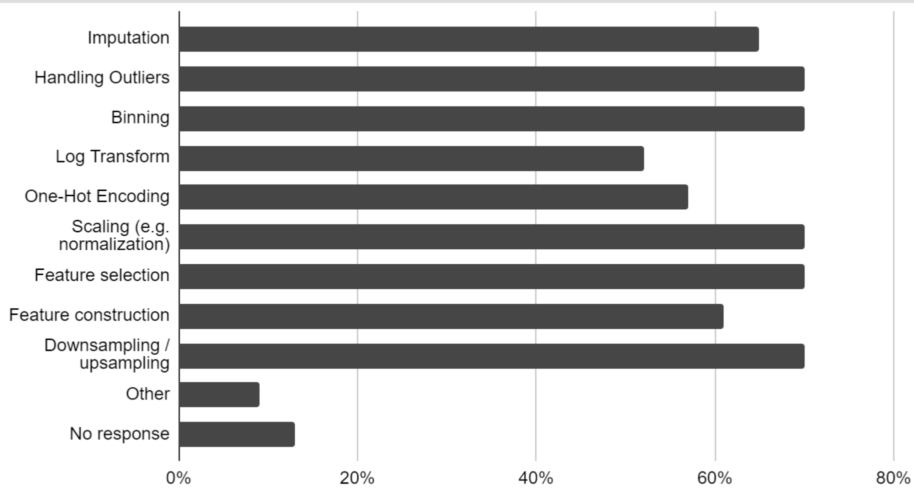
Most respondents utilize regularly fewer than 10 data sources to perform analytics.

Q4.1. What software tools are used for data processing (select all that apply)?



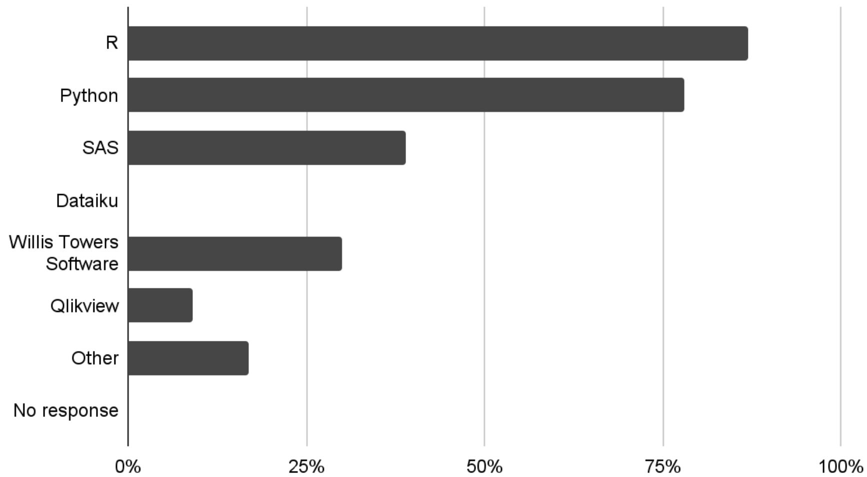
R, Python and SAS are most widely used for data processing.

Q4.2. What feature engineering methods do you currently employ (select all that apply)?



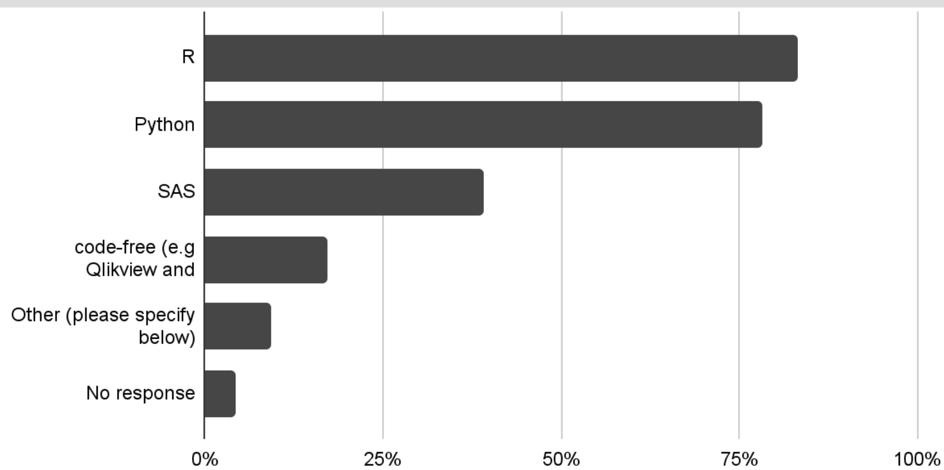
A variety of feature engineering methods are currently employed.

Q4.3. What platforms are used to conduct advanced analytics (select all that apply)?



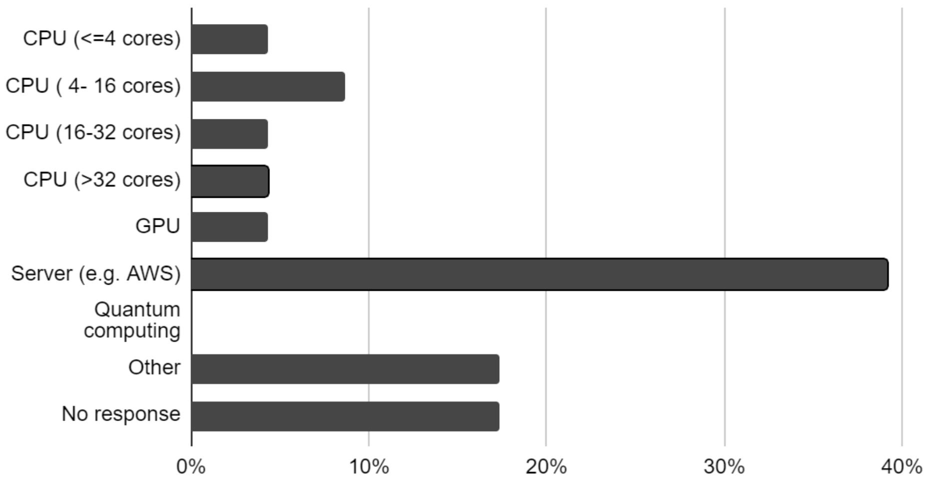
Advanced analytics are mostly performed in R, Python, and SAS.

Q4.4. What programming languages are used to conduct advanced analytics (select all that apply)?



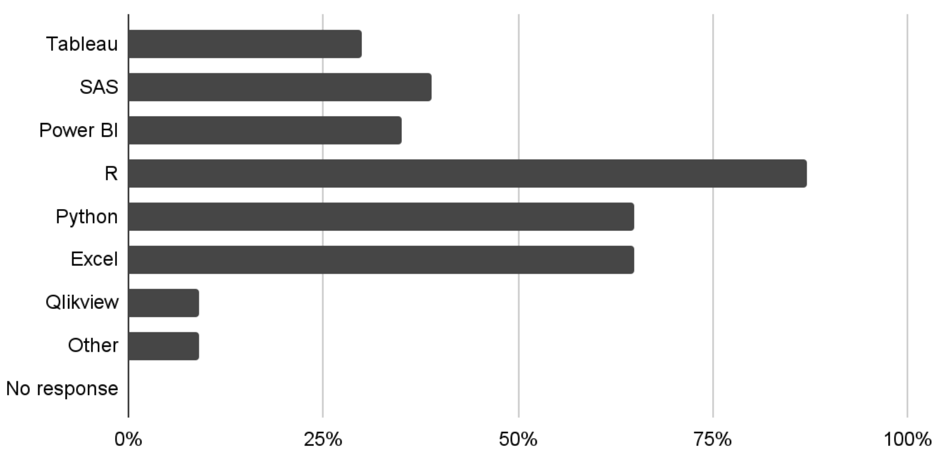
Advanced analytics are mostly performed in R, Python, and SAS.

Q4.5. What is the computing power in your organization?



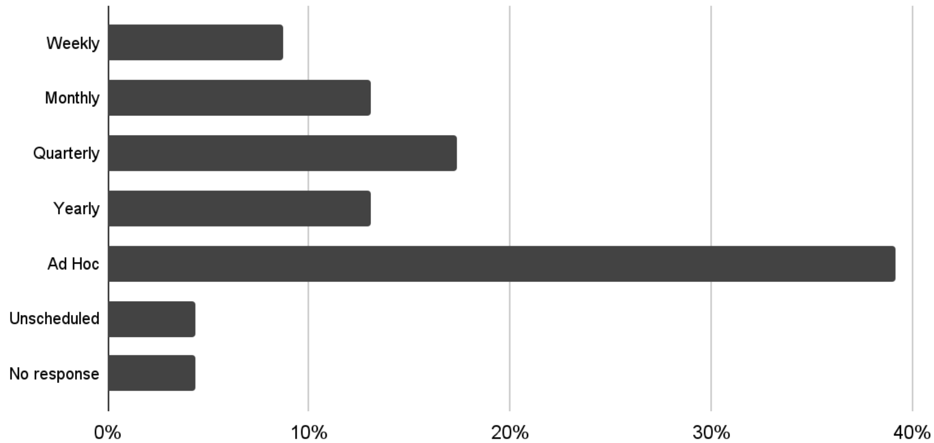
Almost half of the respondents rely on servers (e.g., AWS) to perform advanced analytics computations.

Q4.6. What software tools are used for exploratory data analysis and visualization (select all that apply)?



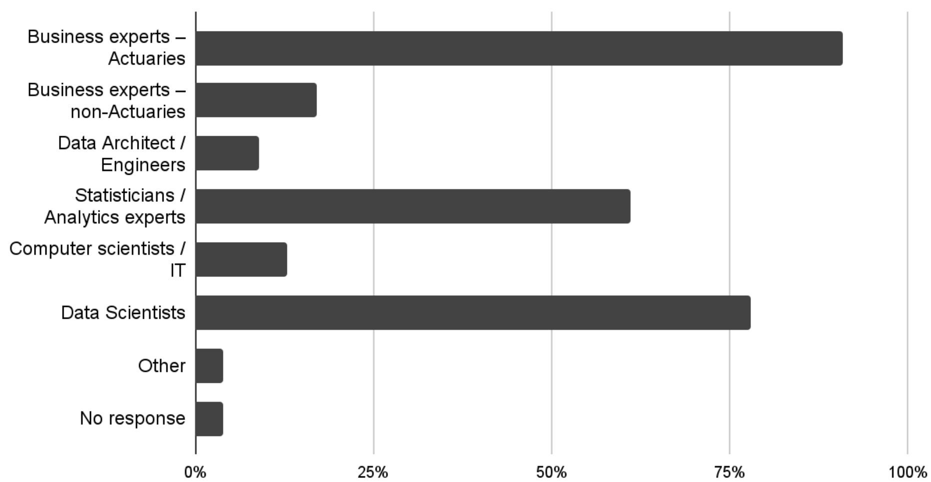
There is a wide mix of software used for exploratory data analysis, with the majority of respondents relying on R, Python, and Excel.

Q4.7. How often do you provide/encourage development/training for staff to build/run models?



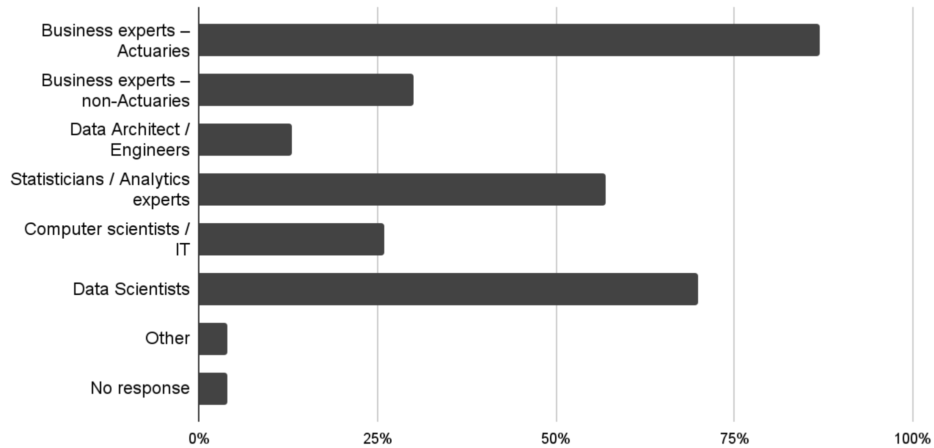
Almost half of the respondents provide training for their employees to build/run models on an ad hoc basis.

Q4.8. What types of FTEs are building advanced analytics models (select all that apply)?



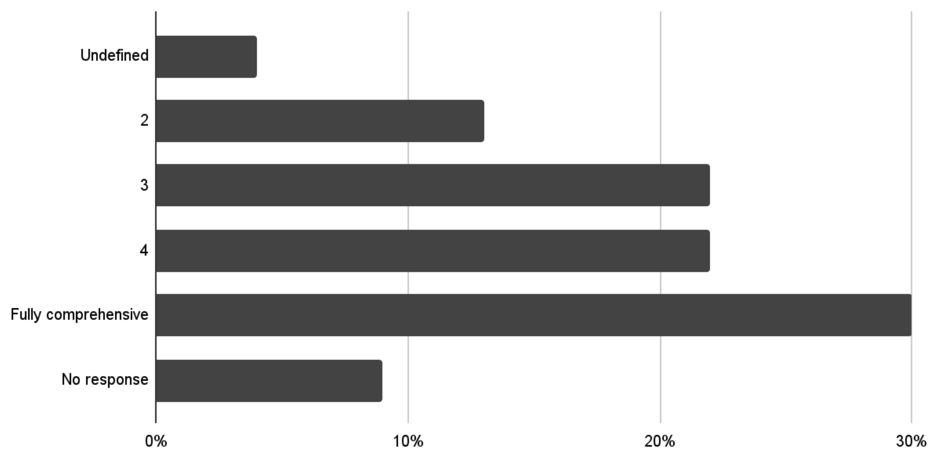
Advanced analytics models are built mostly by actuaries, data scientists and statisticians/analytics experts.

Q4.9. What types of FTEs are running the advanced analytics models (select all that apply)?



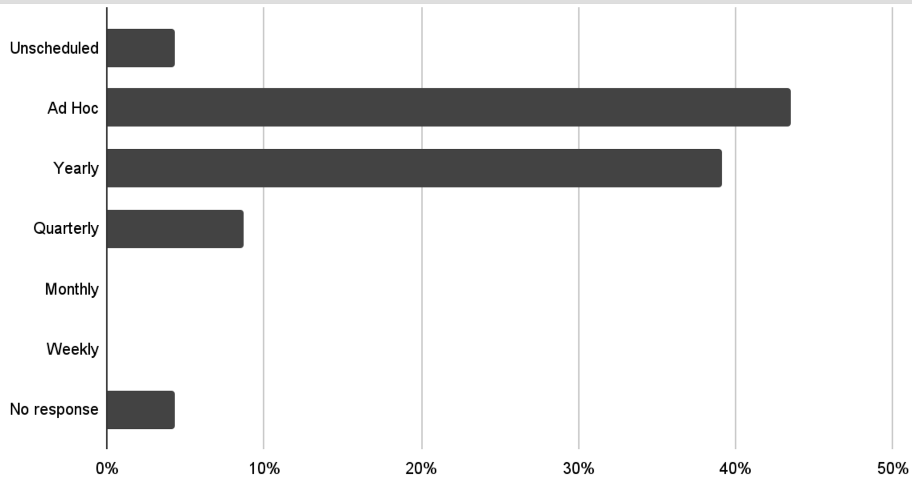
Advanced analytics models are run mostly by actuaries, data scientists and statisticians/analytics experts.

Q4.10. How comprehensive are the model governance/oversight/peer review process policies in place?



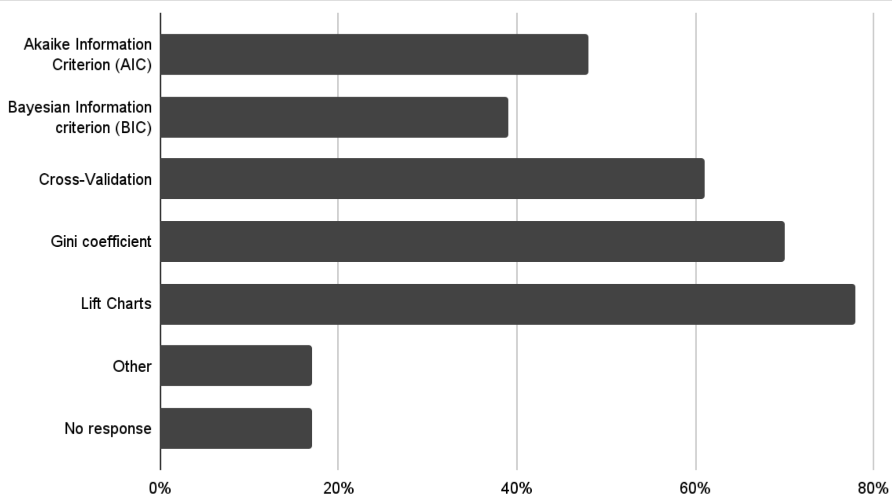
Model governance/oversight/peer review process policies are very comprehensive in at least half of the respondents.

Q4.11. How often are the model governance/oversight/peer review process guidelines reviewed?



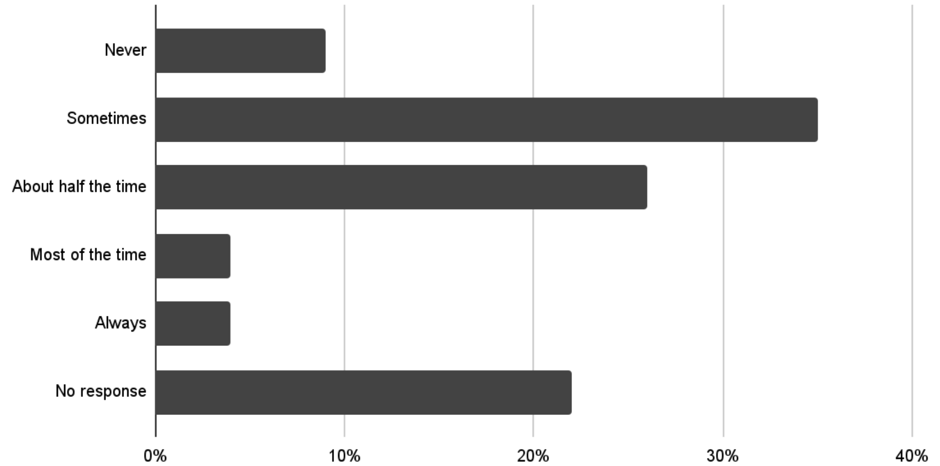
More than half of the respondents state that they review guidelines on an ad hoc or unscheduled basis.

Q4.12. What criteria are you using in model selection (select all that apply)?



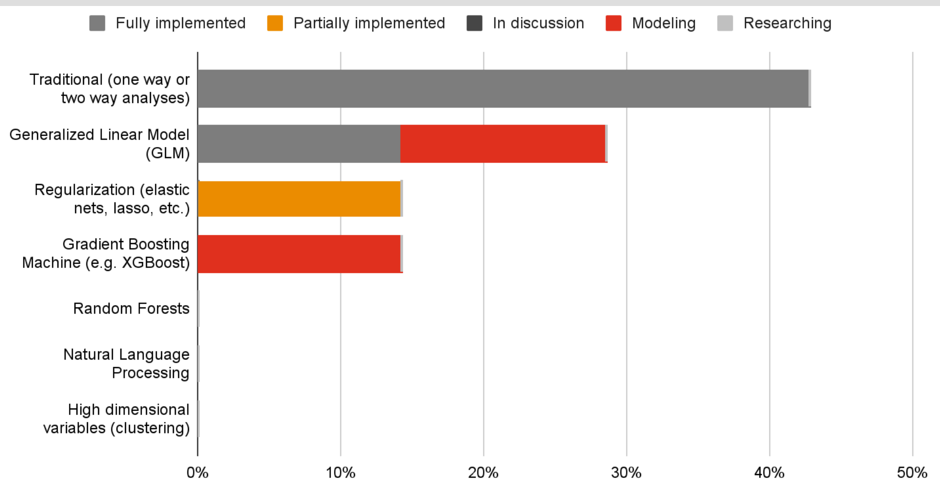
A variety of criteria are used in model selections, with heavy reliance on cross-validation, Gini coefficient and lift charts.

Q4.13. How often are behavioural economics data considered in the modeling process?

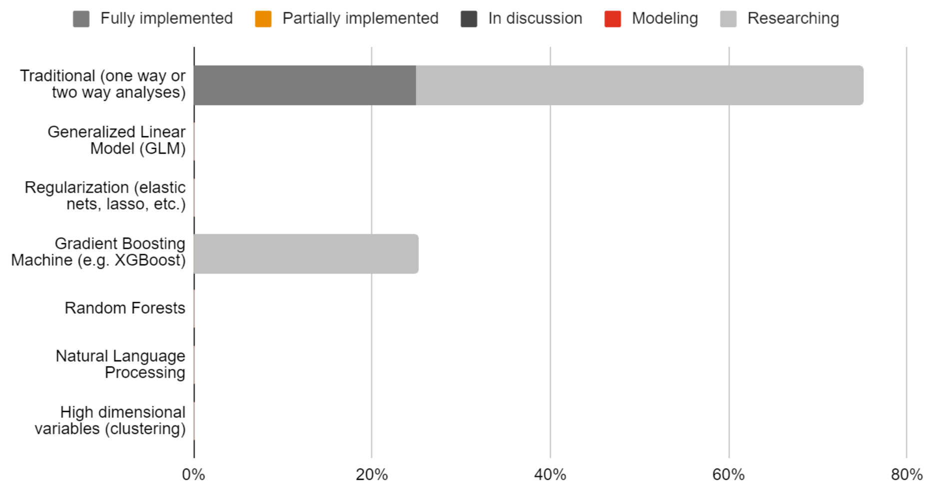


Behavioural economics data are rarely considered in the modeling process.

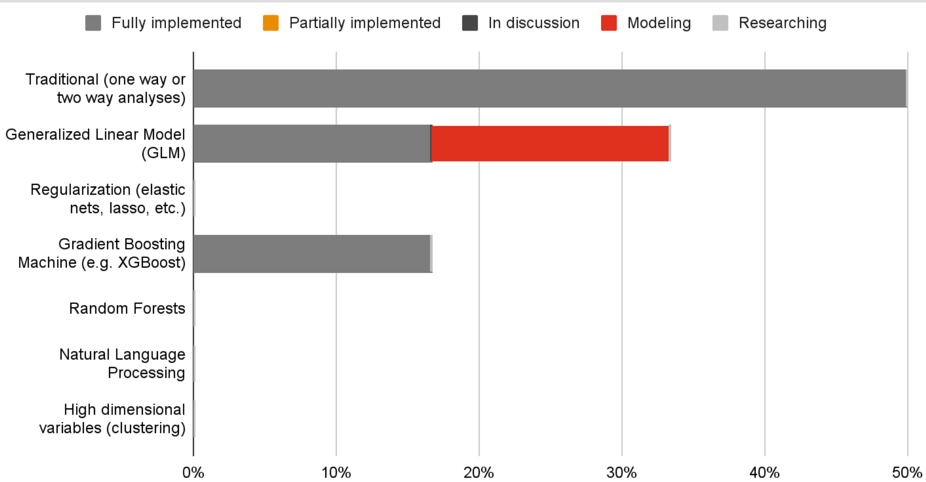
Q4.14.1.1. Strategy and growth – Advanced analytics method – Growth analytics



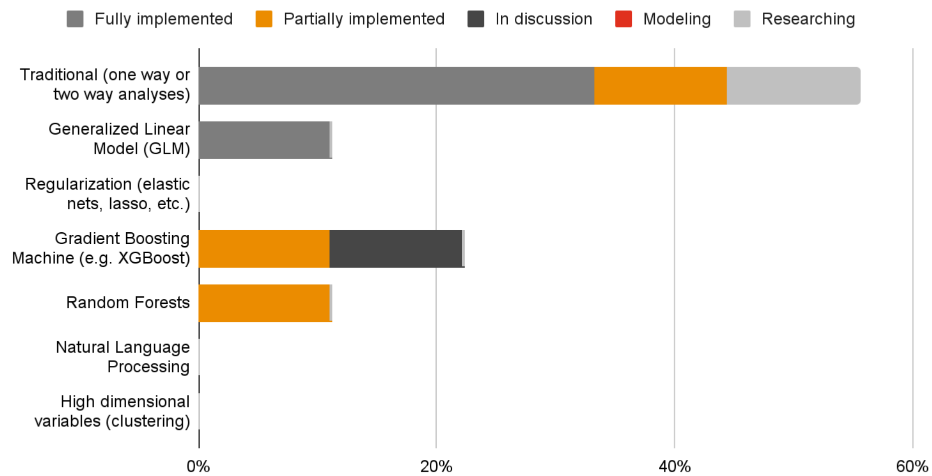
Q4.14.1.2. Strategy and growth – Advanced analytics method – Diffusion analytics



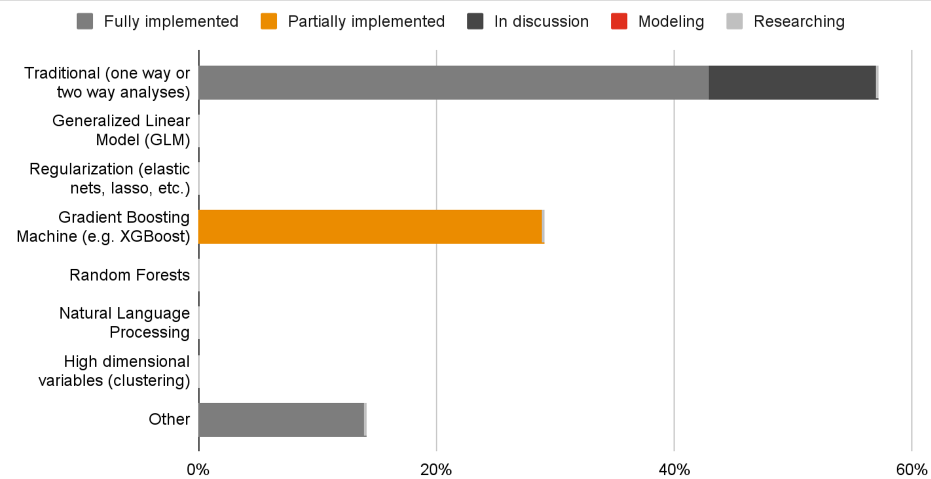
Q4.14.1.3. Strategy and growth – Advanced analytics method – Return analytics



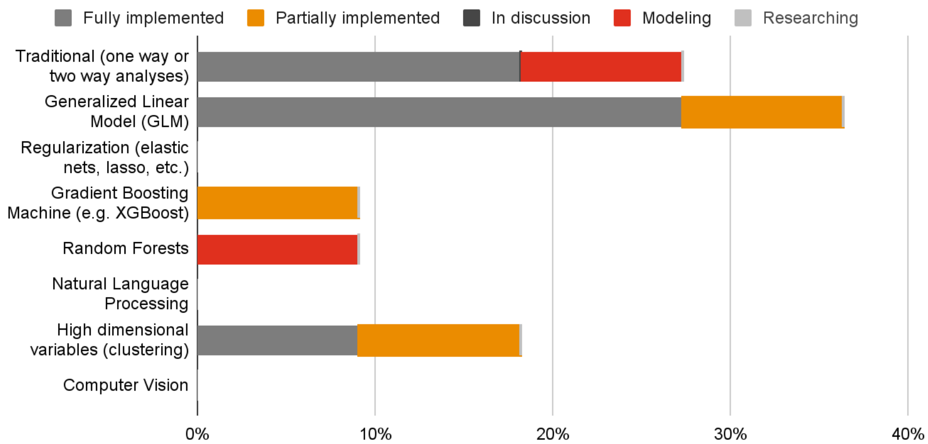
Q4.14.1.4. Strategy and growth – Advanced analytics method – Competitive analytics



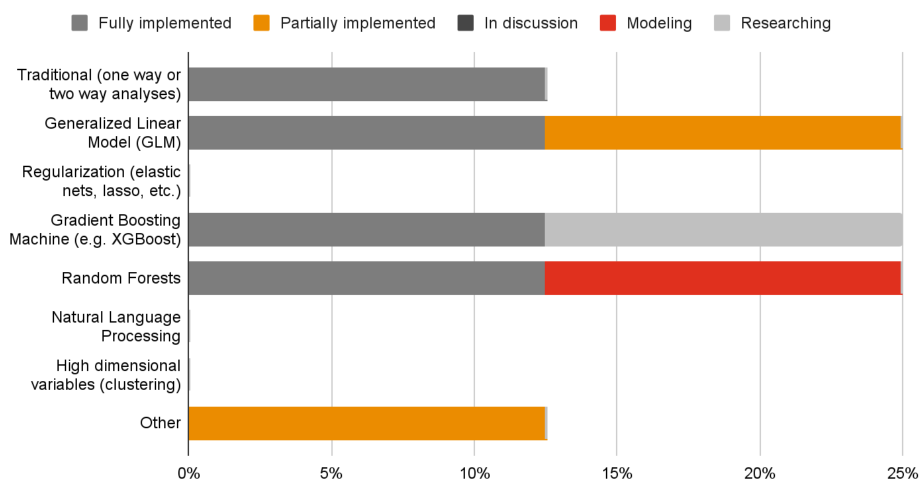
Q4.14.1.5. Strategy and growth – Advanced analytics method – Scenario Analytics



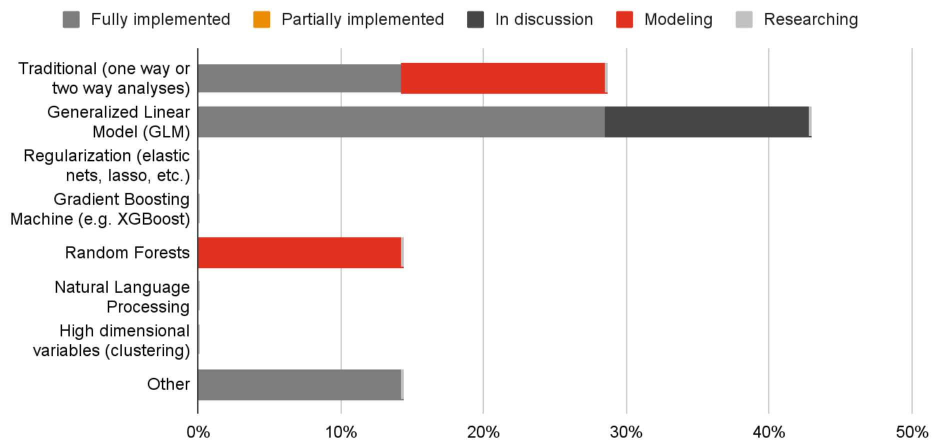
Q4.14.2.1. Customer and marketing – Advanced analytics method – Customer Segmentation



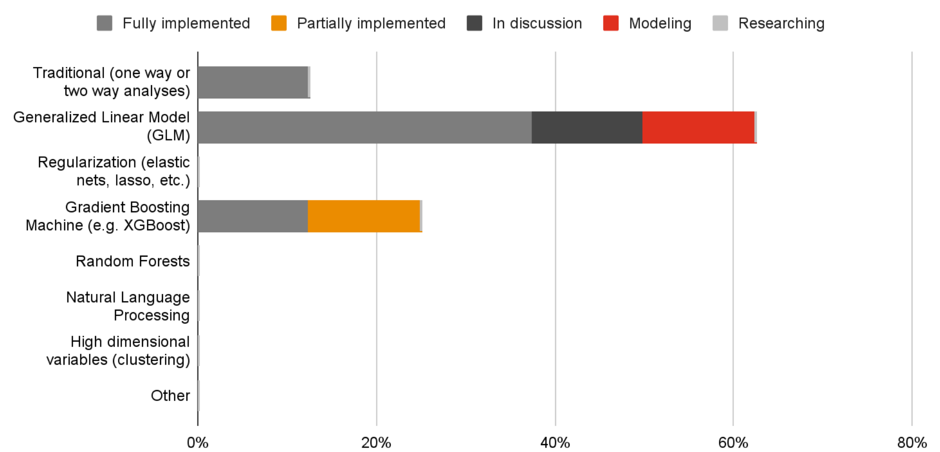
Q4.14.2.2. Customer and marketing – Advanced analytics method – Acquisition analytics



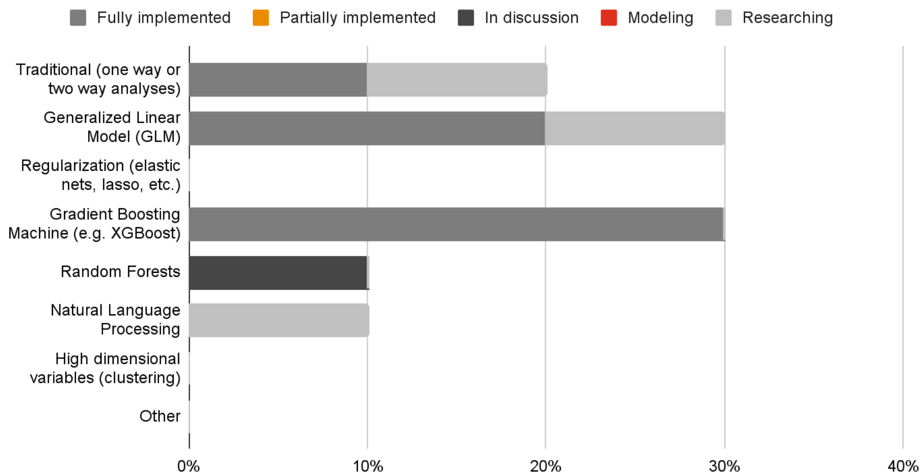
Q4.14.2.3. Customer and marketing – Advanced analytics method – Marketing Spend/Mix analytics



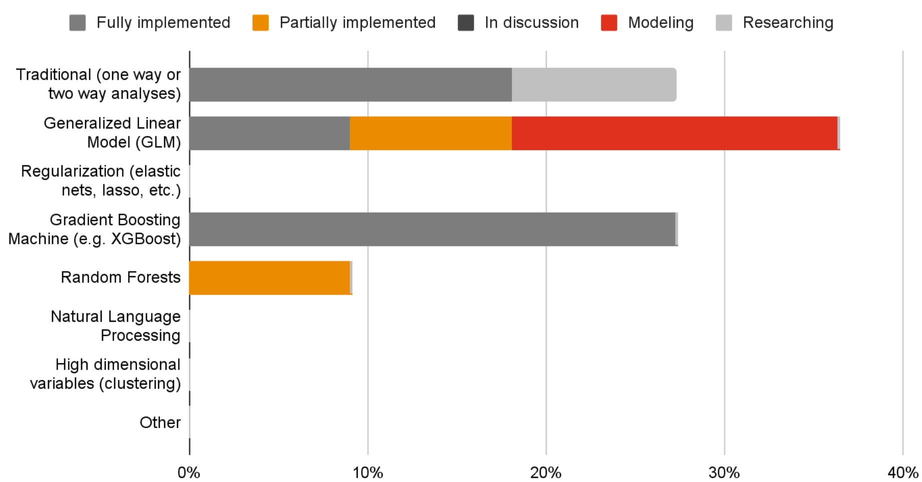
Q4.14.2.4. Customer and marketing – Advanced analytics method – Customer lifetime value (CLV) analytics



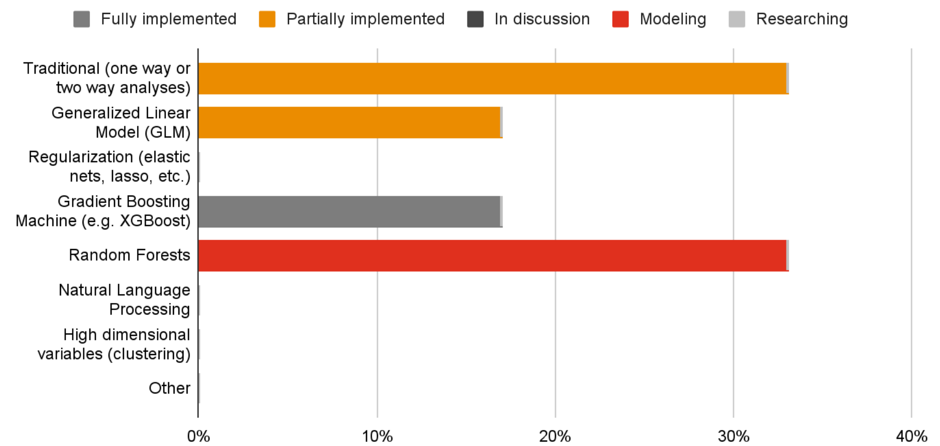
Q4.14.2.5. Customer and marketing – Advanced analytics method – Experience analytics



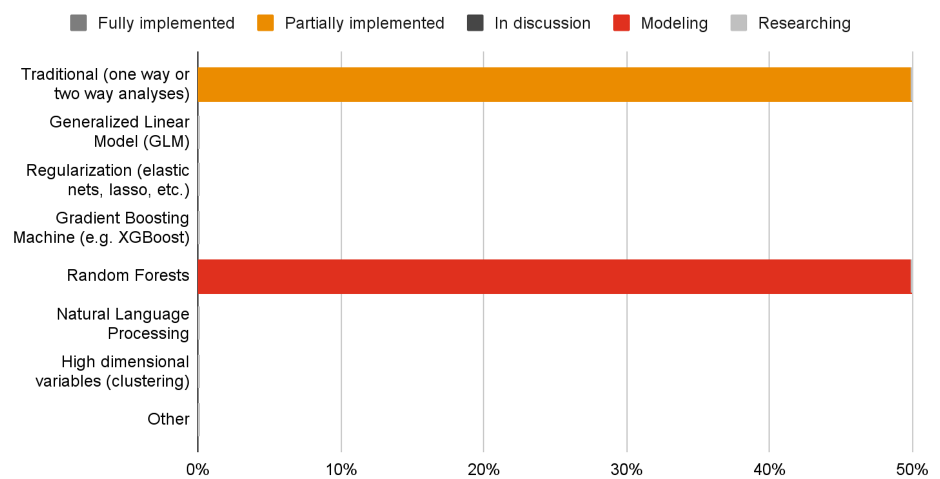
Q4.14.2.6. Customer and marketing – Advanced analytics method – Retention analytics



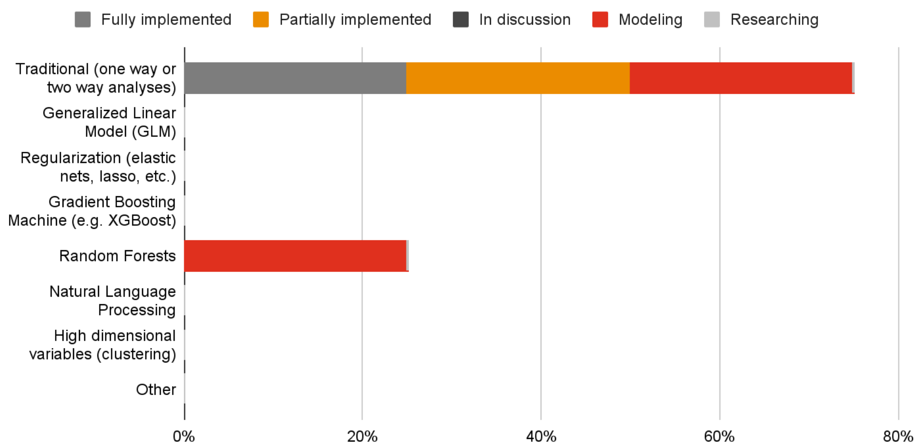
Q4.14.3.1. Sales and distribution – Advanced analytics method – Distribution Segmentation



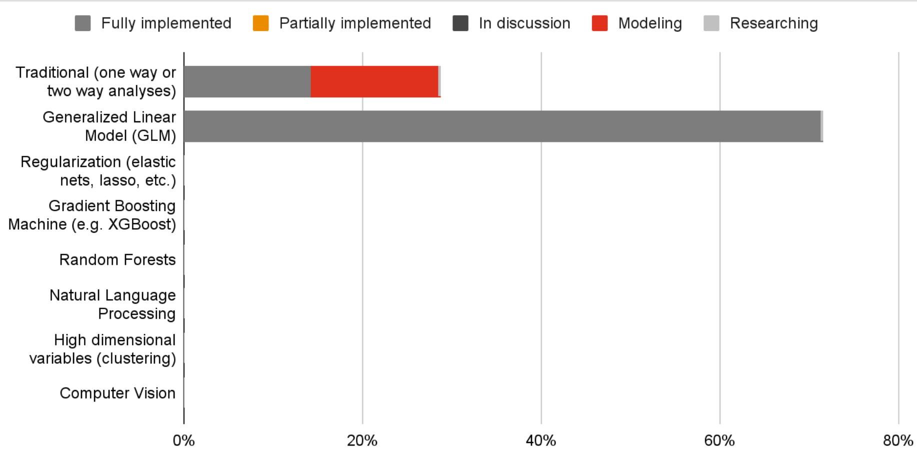
Q4.14.3.2. Sales and distribution – Advanced analytics method – Recruitment analytics



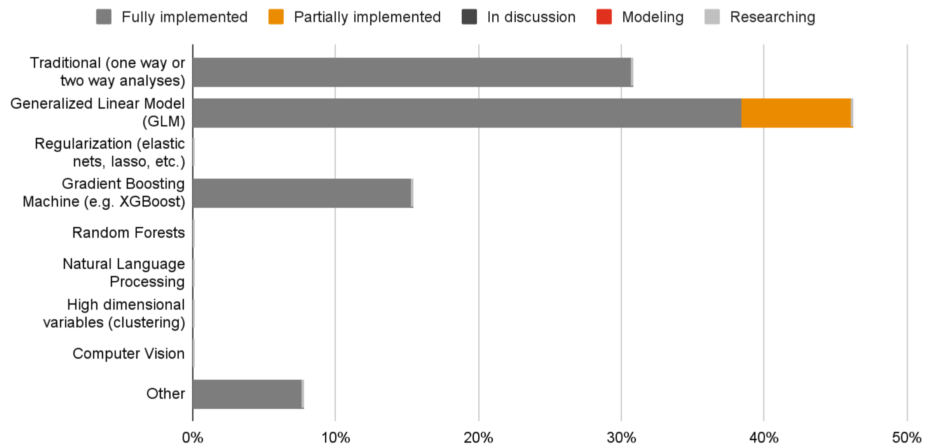
Q4.14.3.3. Sales and distribution – Advanced analytics method – Distribution Value Management



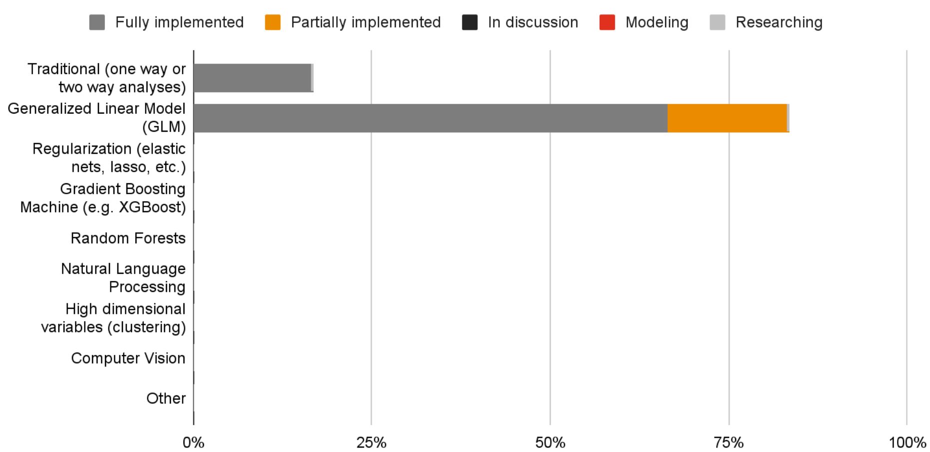
Q4.14.4.1. Products, pricing, and underwriting – Advanced analytics method – Product design analytics



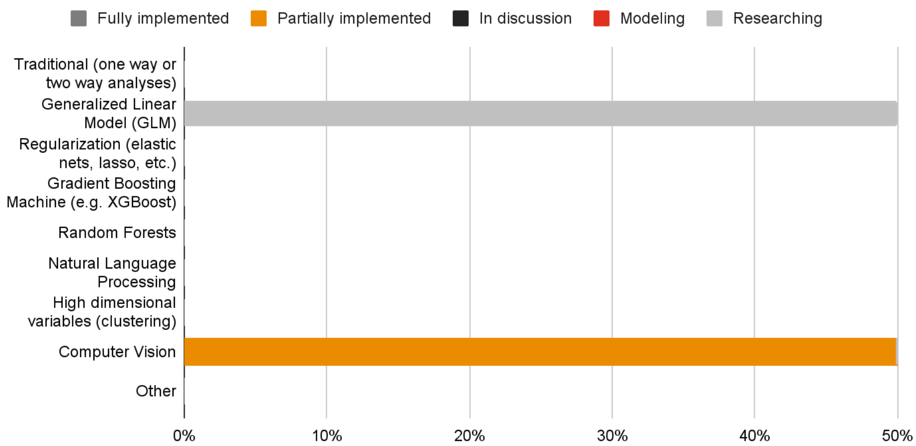
Q4.14.4.2. Products, pricing, and underwriting – Advanced analytics method – Product profitability analytics



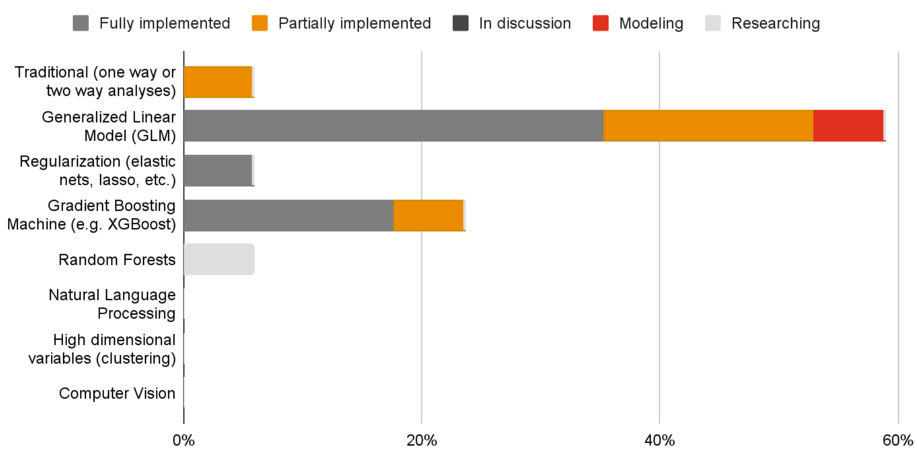
Q4.14.4.3. Products, pricing, and underwriting – Advanced analytics method – Usage Based Insurance



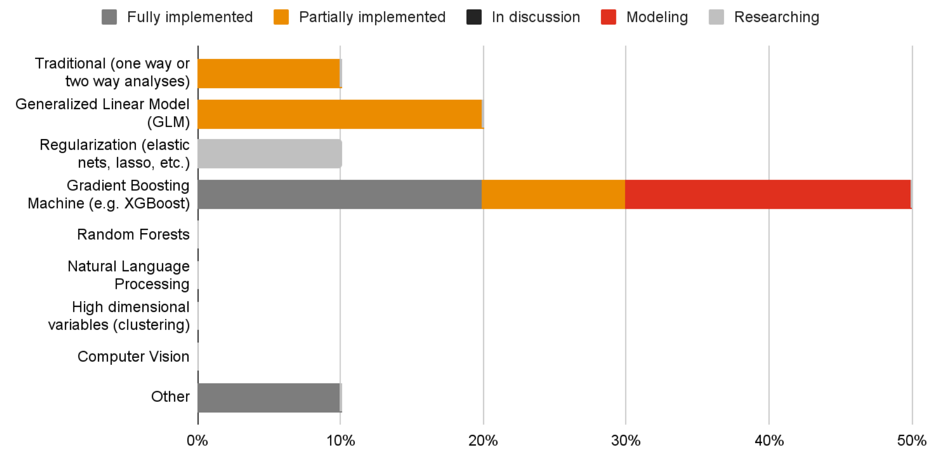
Q4.14.4.4. Products, pricing, and underwriting – Advanced analytics method – Connected homes



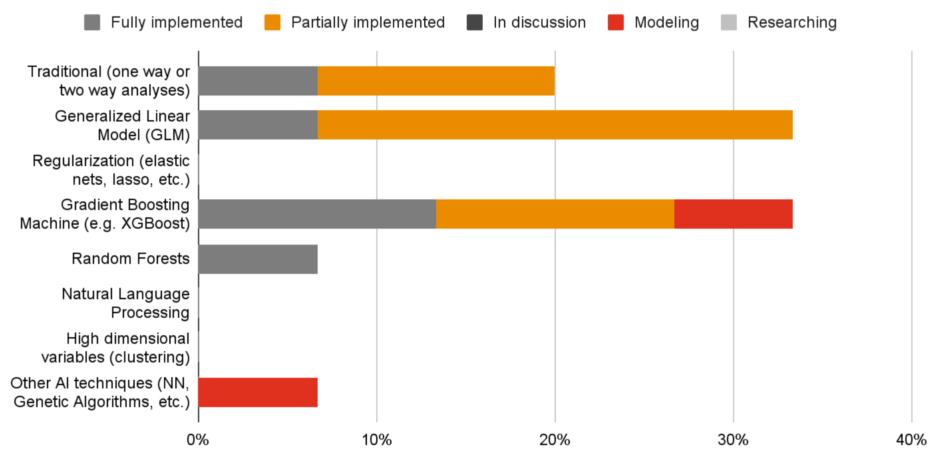
Q4.14.4.5. Products, pricing, and underwriting – Advanced analytics method – Pricing analytics



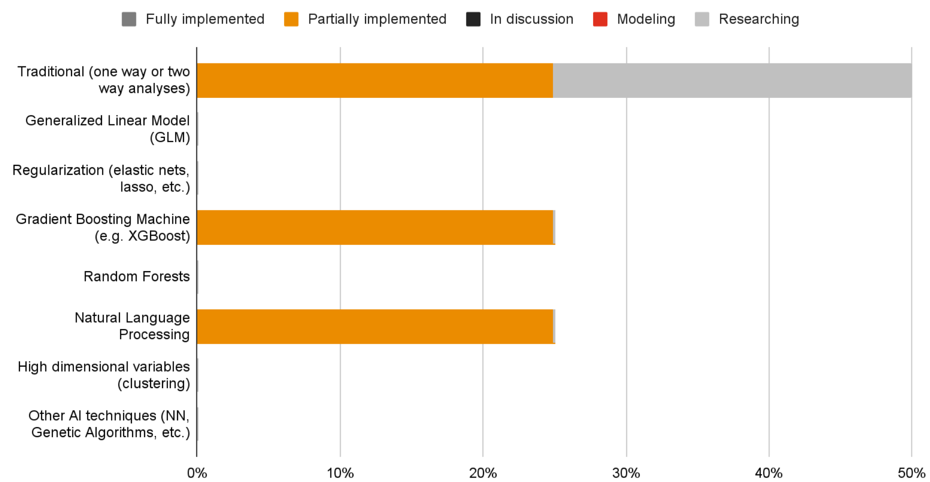
Q4.14.4.6. Products, pricing, and underwriting – Advanced analytics method – Price Optimization



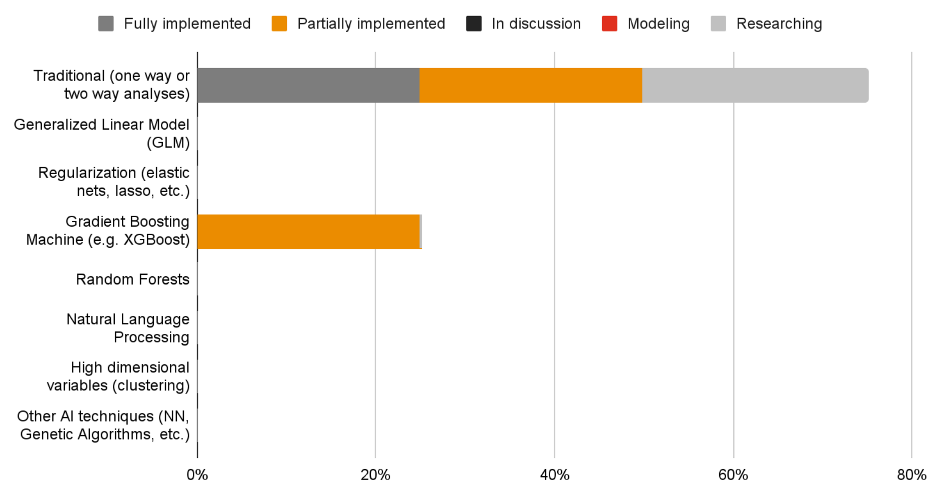
Q4.14.4.7. Products, pricing, and underwriting – Advanced analytics method – Underwriting analytics



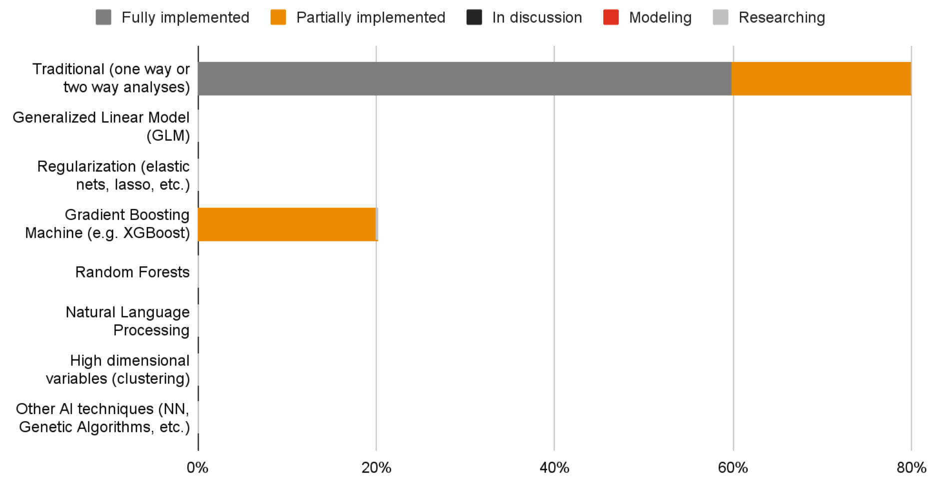
Q4.14.5.1. Process and operations – Advanced analytics method – Multi-channel optimization



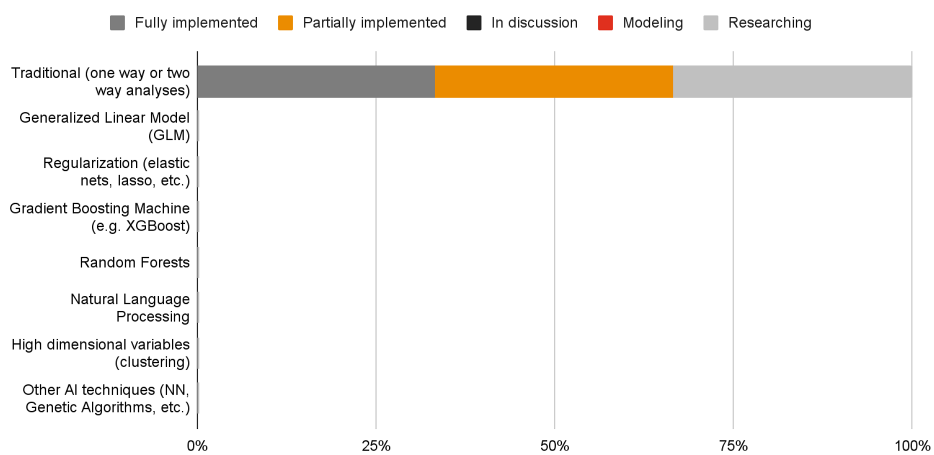
Q4.14.5.2. Process and operations – Advanced analytics method – Policy flow analytics



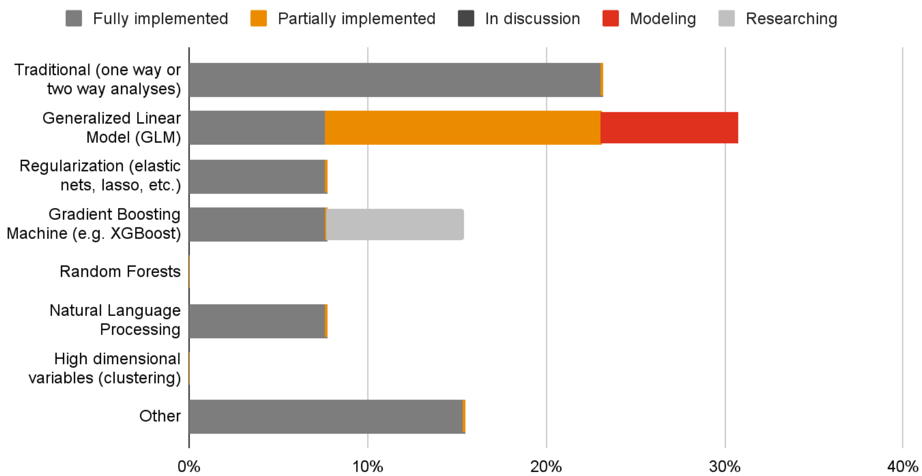
Q4.14.5.3. Process and operations – Advanced analytics method – Portfolio optimization



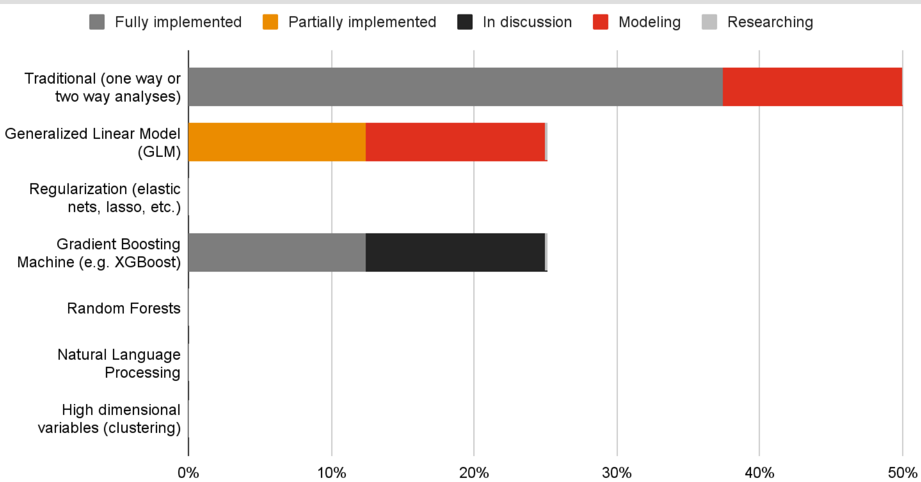
Q4.14.5.4. Process and operations – Advanced analytics method – Program planning & execution analytics



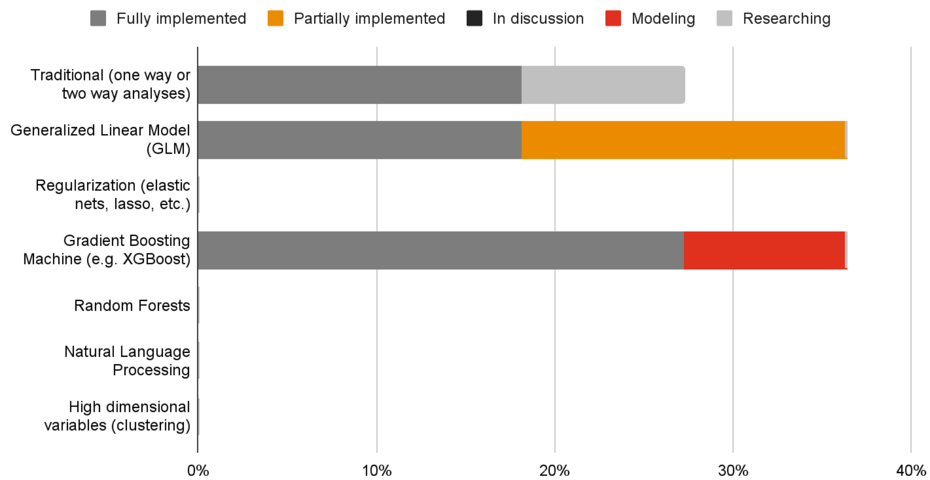
Q4.14.6.1. Claim and benefit analytics – Advanced analytics method – Fraud analytics



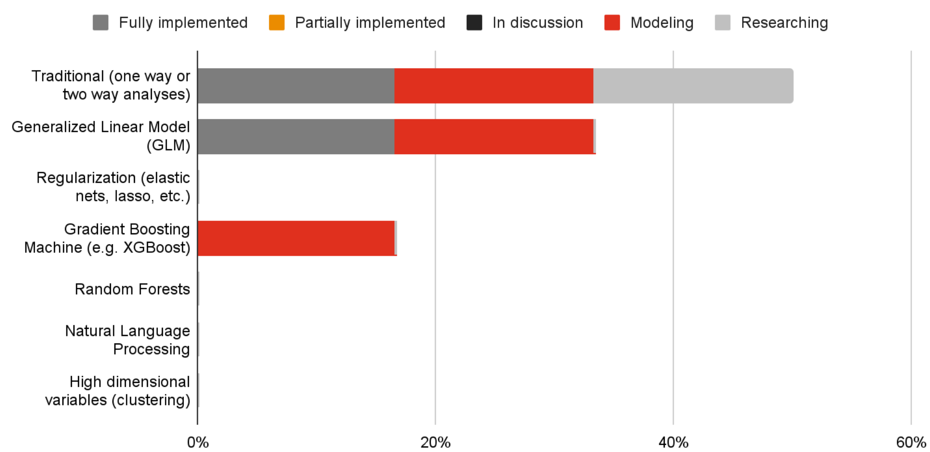
Q4.14.6.2. Claim and benefit analytics – Advanced analytics method – Claim flow analytics



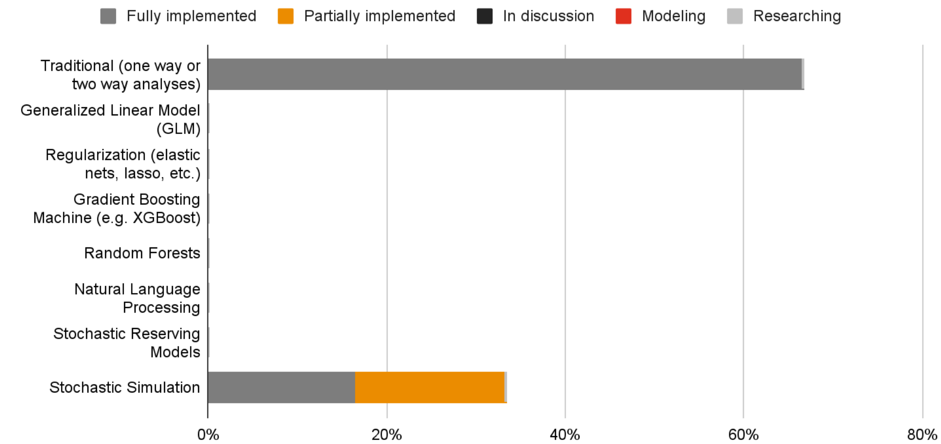
Q4.14.6.3. Claim and benefit analytics – Advanced analytics method – Claim loss analytics



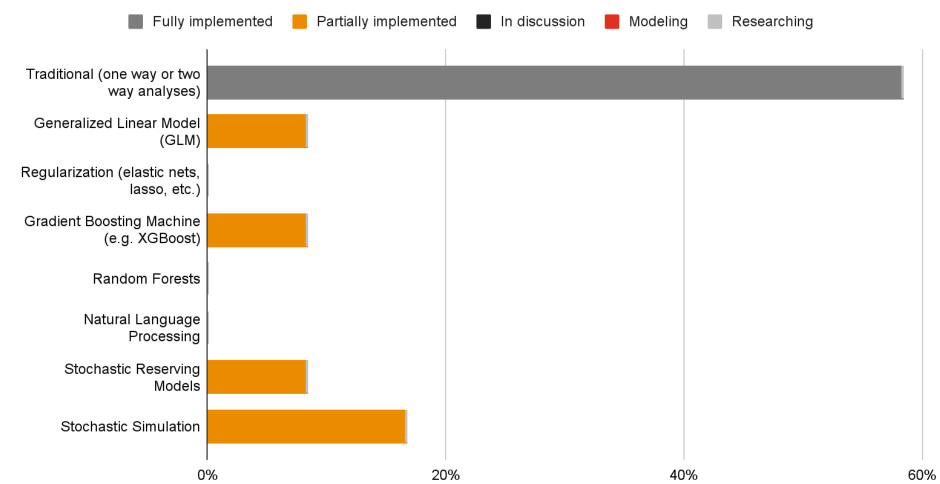
Q4.14.6.4. Claim and benefit analytics – Advanced analytics method – Policyholder behaviour models



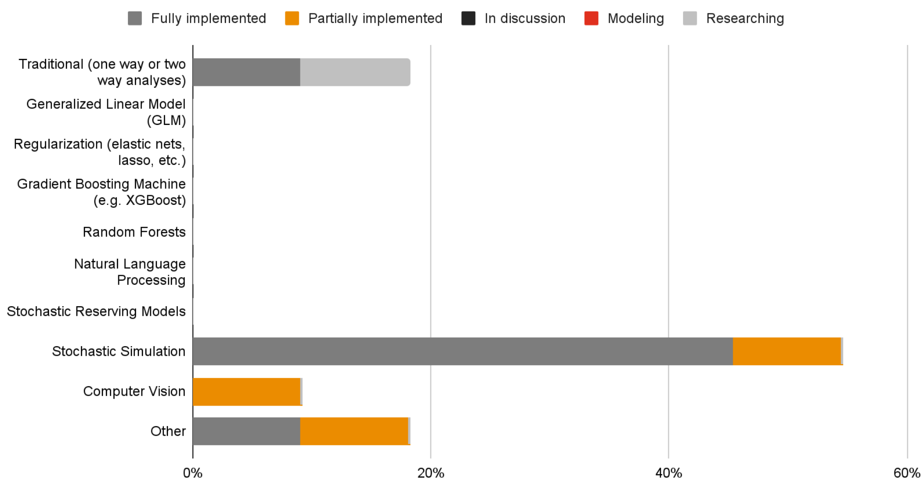
Q4.14.7.1. Capital, risk, and finance – Advanced analytics method – Asset Liability Matching (ALM) Analytics



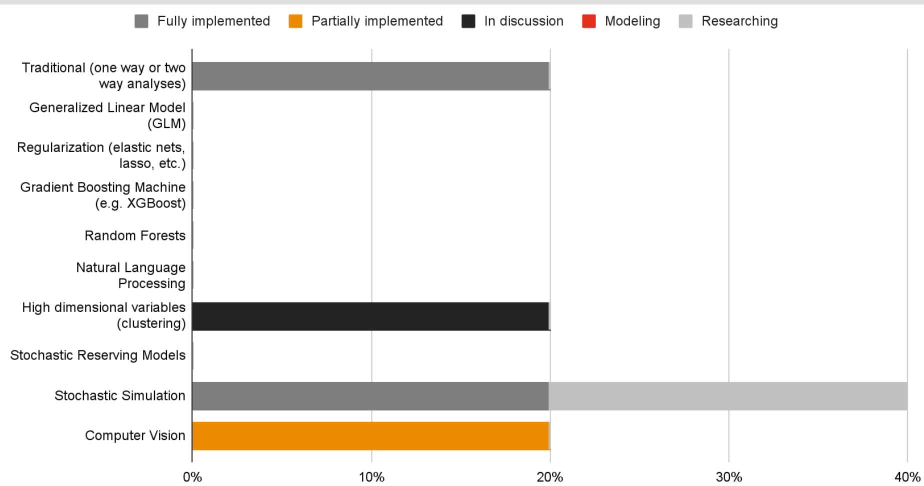
Q4.14.7.2. Capital, risk, and finance – Advanced analytics method – Reserving Analysis



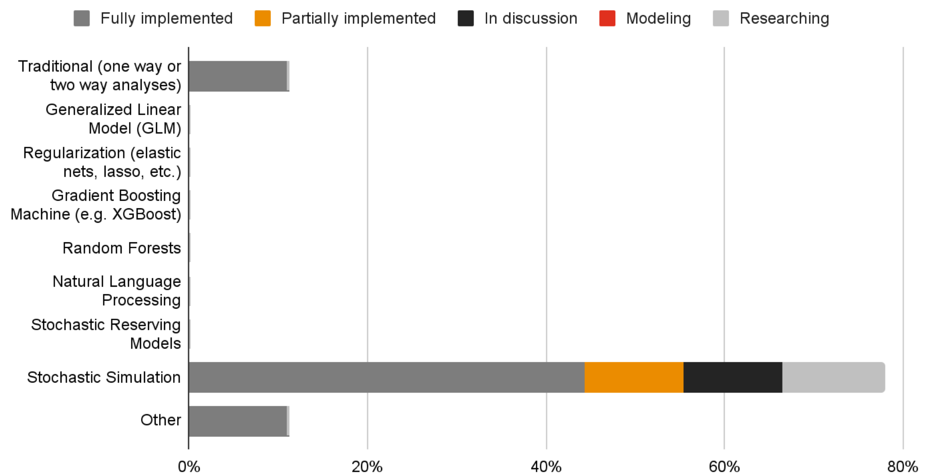
Q4.14.7.3. Capital, risk, and finance – Advanced analytics method – Catastrophe modeling



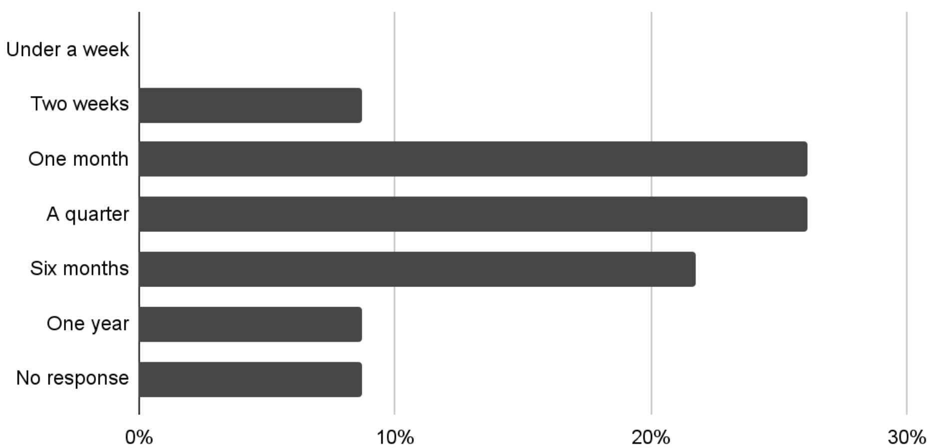
Q4.14.7.4. Capital, risk, and finance – Advanced analytics method – Concentration Analytics



Q4.14.7.5. Capital, risk, and finance – Advanced analytics method – Solvency Models

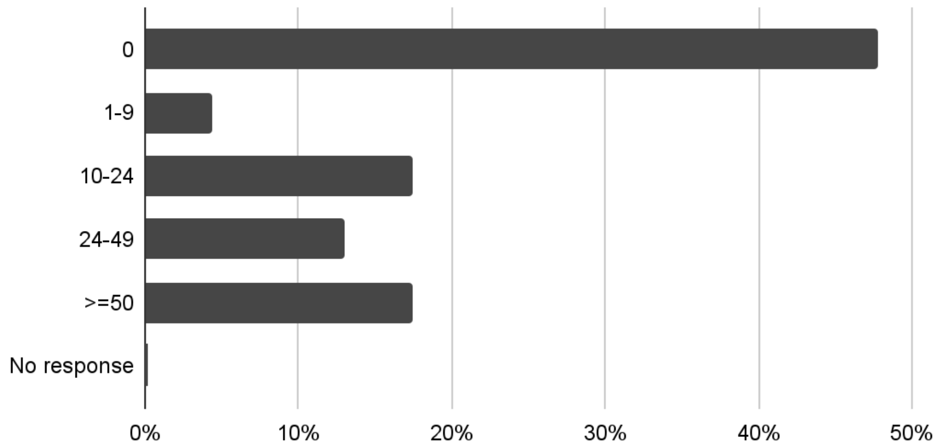


Q4.15. On average how long does it take to deploy changes to your most business-critical models?



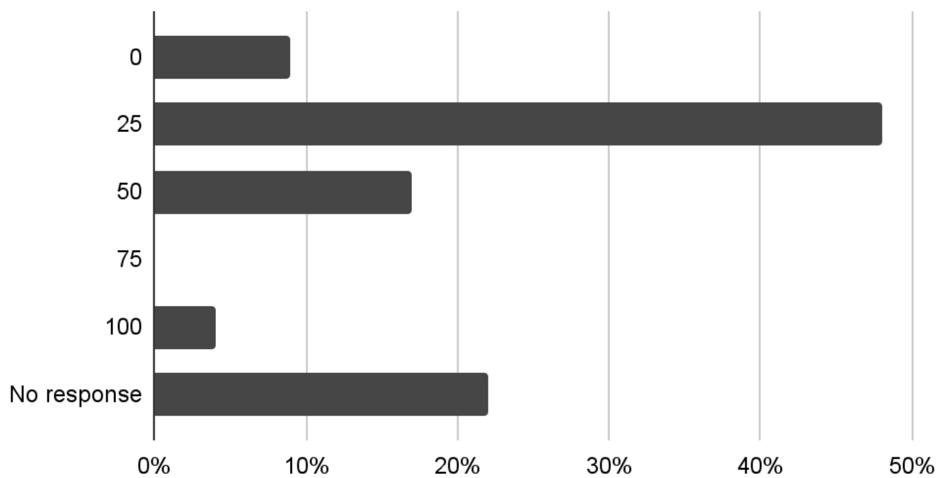
90% of the respondents are able to deploy changes to their most business-critical models within six months, and two-thirds within a quarter.

Q4.16. What percentage of models have been implemented with real-time data processing?



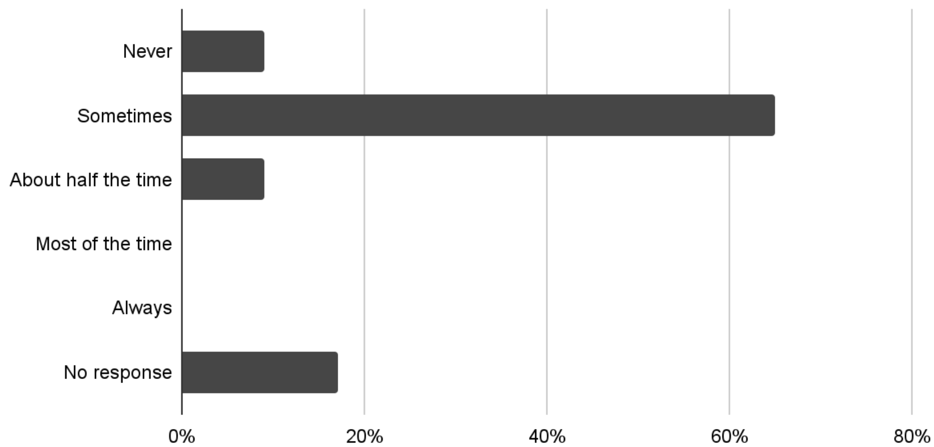
Almost half of the respondents do not have models implemented with real-time data processing.

Q4.17. What percentage of models are accessible via APIs?



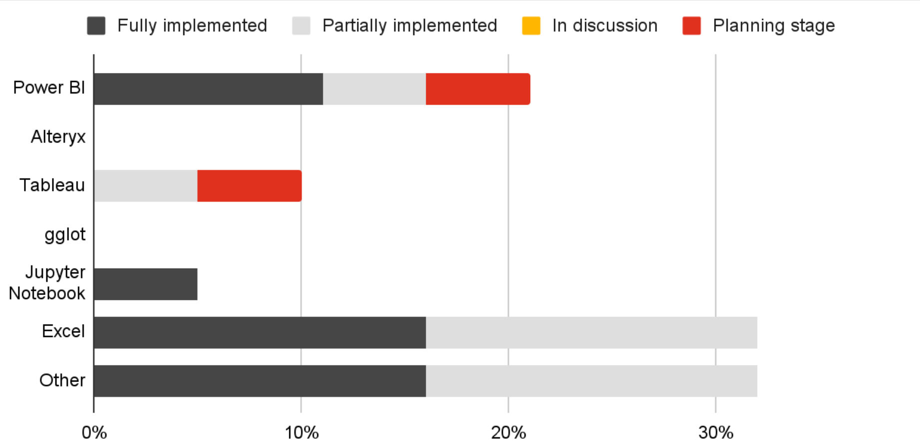
Only one in four respondents have at least half of their models accessible via APIs.

Q4.18. How often do you seek external assistance to implement models (consultants, out-of-the-box applications, etc.)?

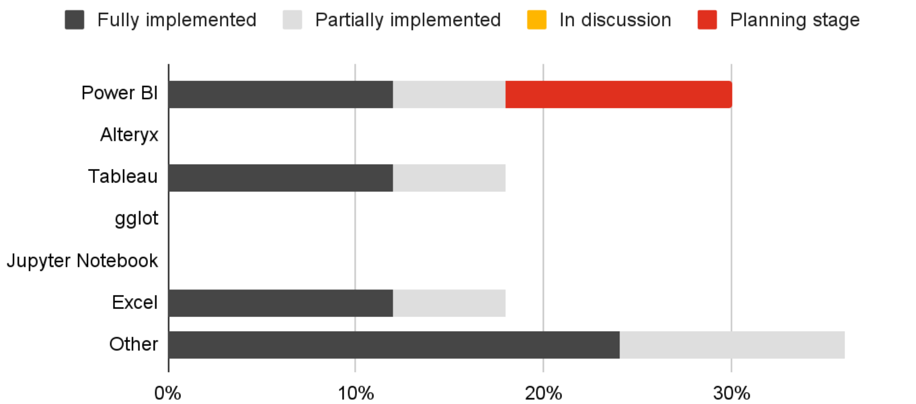


Most of the time model implementation is done internally.

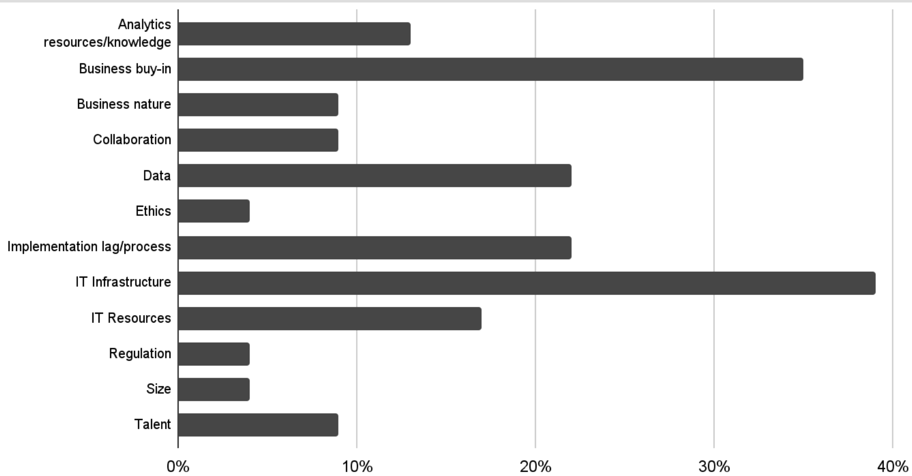
Q4.19.1. What tools are you using in monitoring and reporting your model results? – Visualization Tools – Monitoring – Dashboard of actual vs expected



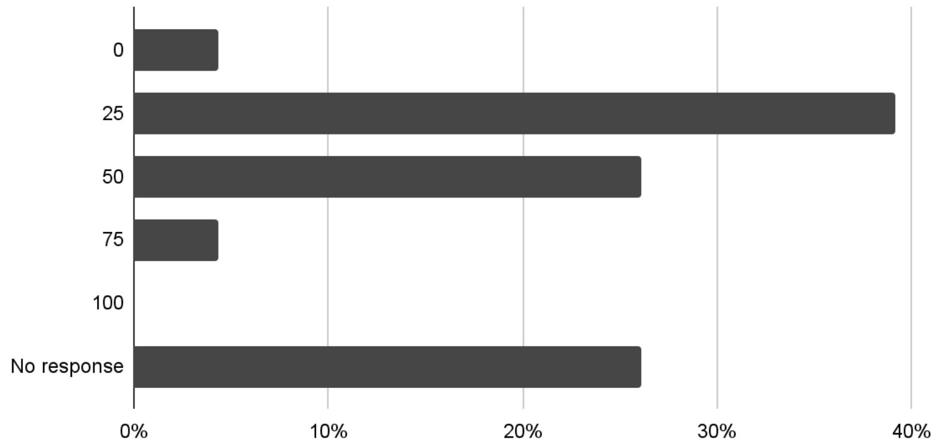
Q4.19.2. What tools are you using in monitoring and reporting your model results? – Business results reporting



Q5.1. What are the key challenges in implementing advanced analytics?

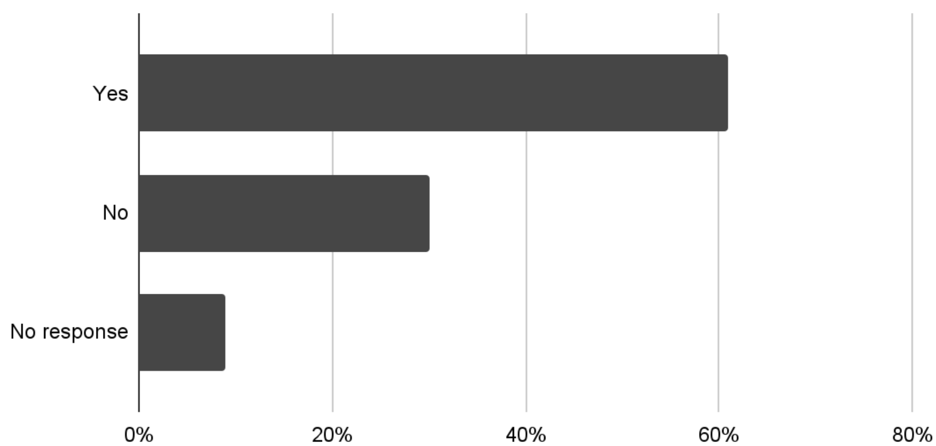


Q5.2. What percentage of prototyped work has been scrapped due to implementation difficulties? (%)

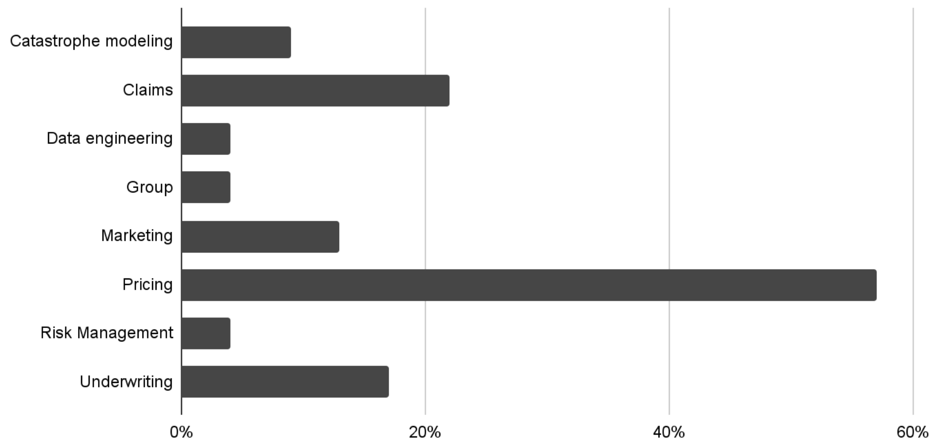


More than one in three companies had to scrap half or more of their prototyped work due to implementation issues.

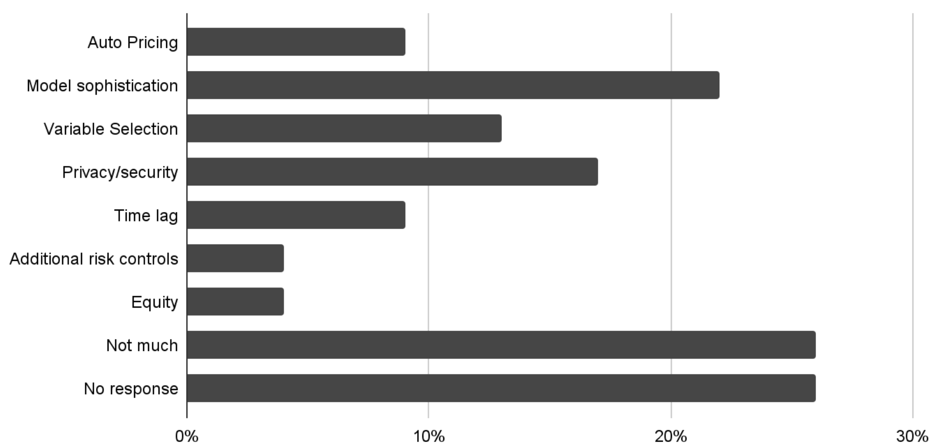
Q5.3 Has your ability to develop and implement advanced analytics been impacted by the age of your IT infrastructure?



Q5.4 Which business area has yielded the most significant gain from utilizing predictive analytics?



Q5.5 How does insurance regulation impact the development and use of advanced analytics in your organization?



Appendix 2 – List of survey questions

#	Question	Response
Q2.1	What level of support does the senior leadership in your organization provide in developing advanced analytics?	• 1 = Little/no support • 5 = Full support
Q2.2	What level of awareness does the senior leadership in your organization have about advanced analytics?	• 1 = Little awareness • 5 = Full awareness
Q2.3	What is the presence and priority of data/advanced analytics on the executive meeting agenda?	• 1 = Rarely discussed • 5 = Frequent/high priority
Q2.4	How many FTEs are working on advanced analytics in your organization?	
Q2.5	What types of FTEs are working on advanced analytics (select all that apply)?	• Business experts – Actuaries • Business experts – non-Actuaries • Data Architect/Engineers • Statisticians/Analytics experts • Computer scientists/IT • Data Scientists • Other (please specify below)
Q2.6	What is the level of collaboration between technical teams who develop advanced analytics and end-users within your company? (%)	• Never (0%) • Sometimes (25%) • About half the time (50%) • Most of the time (75%) • Always (100%)
Q2.7	How often are the results from advanced analytics considered when business decisions are made? (%)	• Never (0%) • Sometimes (25%) • About half the time (50%) • Most of the time (75%) • Always (100%)

#	Question	Response
Q2.8	When deploying a new analytics application that has not been previously executed (e.g., using GLM to predict losses for the first time), to what extent would your organization leverage external consultants? (%)	• Never (0%) • Sometimes (25%) • About half the time (50%) • Most of the time (75%) • Always (100%)
Q2.9	How difficult is it to hire and retain analytics experts for internal positions?	• Extremely difficult • Somewhat difficult • Neither easy nor difficult • Somewhat easy • Extremely easy • n/a – no internal positions
Q2.10	How are your analytics capabilities organized throughout your organization? – Selected choice	a. Separately for each product and function b. Segregated at the functional level (Experience Studies, Pricing, Valuation, etc.) c. Business line level (i.e., across all functions and products) d. Mix of centralized and BU-specific experts e. Centralized analytics function f. Other (Please specify)
Q2.11_a	How would you rank the following aspects of your analytics talent? – Technological capabilities	• Low • Low-Medium • Medium • Medium-High • High • n/a – no FTEs working on analytics
Q2.11_b	How would you rank the following aspects of your analytics talent? – Statistics/Analytics knowledge	• Low • Low-Medium • Medium • Medium-High • High • n/a – no FTEs working on analytics
Q2.11_c	How would you rank the following aspects of your analytics talent? – Business knowledge	• Low • Low-Medium • Medium • Medium-High • High • n/a – no FTEs working on analytics

#	Question	Response
3.1	Is the database maintained by Internal or External Staff?	a. Internal b. External c. Mixture (specify %)
3.2	How many Full Time Equivalents are maintaining the database?	
3.3	What types of FTEs are maintaining the databases (select all that apply)?	a. Business experts – Actuaries b. Business experts – non-Actuaries c. Data Architect/Engineers d. Statisticians/Analytics experts e. Computer scientists/IT f. Data Scientists g. Other (please specify)
3.4	What are the platforms used to store data (select all that apply)?	a. Internal Proprietary System b. Excel spreadsheets c. Flat Files (CSV, Text files) d. MS Access e. SQL Server f. Hadoop (or similar) g. Other (please specify)
3.5	What percentage of your data has been defined in data dictionaries or metadata? (%)	a. None (0%) b. Several (25%) c. Some (50%) d. Most (75%) e. All (100%)
3.6	How would end-users rate the accuracy and completeness of the data dictionaries or metadata?	• 1 = Low • 5 = High
3.7	How often are the data dictionaries or metadata maintained? (%)	a. Not at all (0%) b. Rarely maintained (25%) c. Oftenly maintained (50%) d. Frequently maintained (75%) e. All the time (100%)
3.8	How comprehensive are the policies that outline data accessibility?	• 1 = Undefined • 5 = Fully comprehensive

3.9	How time-consuming is the process to gain access to data?	• Extremely difficult = 1 • Extremely easy = 5

#	Question	Response
3.10	How often are the data reconciliation processes conducted? (%)	a. Not at all (0%) b. Occasionally (25%) c. Often (50%) d. Most of the time (75%) e. All of the time (100%)
3.11	How comprehensive are the data reconciliation processes?	• 1 = Undefined • 5 = Fully comprehensive
3.12	How would end-users rate the completeness of your data?	• 1 = Low • 5 = High
3.13	How would end-users rate the accuracy of your data?	• 1 = Low • 5 = High
3.14	What are the proportions of data in each data format (must total 100)? (%)	a. Unstructured (voice, image, scanned documents) b. Free-form text c. Code that requires a legend to interpret d. Scalars/Values that requires no legend e. Arrays, tables, cubes
3.15	How comprehensive are the data governance/oversight/process policies?	• 1 = Undefined • 5 = Fully comprehensive
3.16	How often are the data governance/oversight/process guidelines reviewed?	a. Ad Hoc b. Yearly c. Quarterly d. Monthly e. Weekly
3.17	How comprehensive are the data privacy and retention policies?	• 1 = Undefined • 5 = Fully comprehensive
3.18	How often are the data privacy and retention policies reviewed?	a. Ad Hoc b. Yearly c. Quarterly d. Monthly e. Weekly
3.19	How many data sources does your organization utilize regularly to perform analytics that will impact your business decisions?	

#	Question	Response	
3.20	Internal data: fill out the key internal data sources and the data format, accuracy and completeness of that data source		
	Data source name	Format	a. Unstructured – voice, images, scanned documents (pre-OCR) b. Unstructured – Free-form text c. Semi-structured – JSON, XML, arrays d. Structured data – tables, cubes e. Other (please specify)
		Accuracy	• 1 = Low • 5 = High
		Completeness	• 1 = Low • 5 = High
		Data Quality Owner	
		Usage	
		How often are the data used to build analytical applications (e.g., predictive models) refreshed?	• Ad Hoc • Yearly • Quarterly • Monthly • Weekly • Daily • Real time

#	Question		Response
3.21	External data: fill out the key external data sources and the data format, accuracy and completeness of that data source		
	Data source name	Format	a. Unstructured – voice, images, scanned documents (pre-OCR) b. Unstructured – Free-form text c. Semi-structured – JSON, XML, arrays d. Structured data – tables, cubes e. Other (please specify)
		Accuracy	• 1 = Low • 5 = High
		Completeness	• 1 = Low • 5 = High
		Data Quality Owner	
		Usage	
		How often are the data used to build analytical applications (e.g., predictive models) refreshed?	• Ad Hoc • Yearly • Quarterly • Monthly • Weekly • Daily • Real time

#	Question	Response
4.1	What softwares are used for data processing (select all that apply)?	a. Tableau b. SAS c. Power BI d. R e. Python f. Other (please specify)
4.2	What feature engineering methods do you currently employ?	a. Imputation b. Handling Outliers c. Binning d. Log Transform e. One-Hot Encoding f. Scaling (e.g., normalization) g. Feature selection h. Feature construction i. Downsampling/upsampling j. Other (Please specify)
4.3	What platforms are used to conduct advanced analytics (select all that apply)?	a. R b. Python c. SAS d. Dataiku e. Willis Towers Software f. Qlikview g. Other (please specify)
4.4	What programming language are used to conduct advanced analytics (select all that apply)?	a. R b. Python c. SAS d. code-free (e.g Qlikview and Tableau) e. Other (please specify)
4.5	What is the computing power in your organization?	a. CPU (<=4 cores) b. CPU (4- 16 cores) c. CPU (16-32 cores) d. CPU (>32 cores) e. GUP f. Server (e.g., AWS) g. Quantum computing h. Other (please specify)

#	Question	Response
4.6	What softwares are used for exploratory data analysis and visualization (select all that apply)?	a. Tableau b. SAS c. Power BI d. R e. Python f. Excel g. Qlikview h. Other (please specify)
4.7	How often do you provide/encourage development/training for staff to build/run models?	a. Unscheduled b. Ad Hoc c. Yearly d. Quarterly e. Monthly f. Weekly
4.8	What types of FTEs are building advanced analytics models (select all that apply)?	a. Business experts – Actuaries b. Business experts – non-Actuaries c. Data Architect/Engineers d. Statisticians/Analytics experts e. Computer scientists/IT f. Data Scientists g. Other (please specify)
4.9	What types of FTEs are running the advanced analytics models (select all that apply)?	a. Business experts – Actuaries b. Business experts – non-Actuaries c. Data Architect/Engineers d. Statisticians/Analytics experts e. Computer scientists/IT f. Data Scientists g. Other (please specify)
4.10	How comprehensive are the model governance/oversight/peer review process policies in place?	• 1 = Undefined • 5 = Fully comprehensive
4.11	How often are the model governance/oversight/peer review process guidelines reviewed?	a. Ad Hoc b. Yearly c. Quarterly d. Monthly e. Weekly

#	Question	Response
4.12	What criteria are used in model selection?	• Akaike Information Criterion (AIC) • Bayesian Information criterion (BIC) • Cross-Validation • Gini coefficient • Lift Charts • Other (please specify)
4.13	How often are behavioural economics data considered in the modeling process? (%)	a. Not at all (0%) b. Rarely (25%) c. Oftenly (50%) d. Frequently (75%) e. All the time (100%)
4.14	For each business unit using predictive analytics, identify the following (select all that apply):	

#	Business unit	Use	Predictive analytics method	Progress
4.14.1	**a. Strategy and growth**	Growth analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Diffusion analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
			Algorithms, etc.) • Other (please specify)	
		Return analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Competitive analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Scenario Analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
			Algorithms, etc.) • Other (please specify)	

#	Business unit	Use	Predictive analytics method	Progress
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
4.14.2	**b. Customer and marketing**	Customer Segmentation	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Acquisition analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Marketing Spend/Mix analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Customer life timevalue (CLV) analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

		Experience analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Retention analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
4.14.3	**c. Sales and distribution**	Distribution Segmentation	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Recruitment analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Distribution Value Management	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
4.14.4	**d. Products, pricing, and underwriting**	Product design analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Product profitability analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Usage Based Insurance	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Connected Homes	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Pricing analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Price Optimization	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Underwriting analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
4.14.5	**e. Process and operations**	Multi-channel optimization	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Policy flow analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Portfolio optimization	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Program planning & execution analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
4.14.6	**f. Claim & benefit analytics**	Fraud analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Claim flow analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Claim loss analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Policyholder behavior models	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

		• Other (please specify)

#	Business unit	Use	Predictive analytics method	Progress
4.14.7	**g. Capital, risk and finance**	Asset Liability Matching (ALM) Analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Stochastic Reserving Models (ODP bootstrapping, GLM) • Stochastic Simulation • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Reserving Analysis	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Stochastic Reserving Models (ODP bootstrapping, GLM) • Stochastic Simulation • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Catastrophe modeling	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Stochastic Reserving Models (ODP bootstrapping, GLM) • Stochastic Simulation • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Concentration Analytics	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Stochastic Reserving Models (ODP bootstrapping, GLM) • Stochastic Simulation • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
		Solvency Models	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Stochastic Reserving Models (ODP bootstrapping, GLM) • Stochastic Simulation • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented
		Other (Please Specify)	• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Stochastic Reserving Models (ODP bootstrapping, GLM) • Stochastic Simulation • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Business unit	Use	Predictive analytics method	Progress
4.14.8	h. Other Use Cases		• Traditional (one- or two-way analyses) • Generalized Linear Model (GLM) • Regularization (elastic nets, lasso, etc.) • Gradient Boosting Machine (e.g., XGBoost) • Random Forests • Natural Language Processing (Text mining, Voice analytics) • High dimensional variables (clustering) • Computer Vision (Satellite data, Image recognition, OCR) • Other AI techniques (NN, HMM, Genetic Algorithms, etc.) • Other (please specify)	• Not Started • Researching • Modeling • In discussion • Partially implemented • Fully implemented

#	Question	Response
4.15	On average how long does it take to deploy changes to your most business-critical models?	
4.16	What percentage of models have been implemented with real-time data processing (must total 100)? (%)	• Real time (%) • Batch (%)
4.17	What percentage of models are accessible via APIs? (%)	• 0% • 25% • 50% • 75% • 100%
4.18	How often do you seek external assistance to implement models (consultants, out-of-the-box applications, etc.)?	• 1 = Never • 5 = Always
4.19.1	What tools are you using in monitoring and reporting your model results? – Visualization Tools – Monitoring – Dashboard of actual vs expected	• Power BI • Alteryx • Tableau • gglot • Jupyter Notebook • Excel • Other (please specify)

#	Question	Response
4.19.2	What tools are you using in monitoring and reporting your model results? – Visualization Tools – Business results reporting	• Power BI • Alteryx • Tableau • gglot • Jupyter Notebook • Excel • Other (please specify)
5.1	What are the key challenges in implementing advanced analytics?	
5.2	What percentage of prototyped work has been scrapped due to implementation difficulties? (%)	a. 0% b. 25% c. 50% d. 75% e. 100%
5.3	Has your ability to develop and implement advanced analytics been impacted by the age of your IT infrastructure?	
5.4	Which business area has yielded the most significant gain from utilizing predictive analytics?	
5.5	How does insurance regulation impact the development and use of advanced analytics in your organization?	

Appendix 3 – Survey glossary

The glossary provided below covers the definitions of the terms as defined in the original survey. It does not seek to provide a list of definitions for all the terms used throughout the report.

Term	Definition
Acquisition analytics	Evaluating how to appeal to already segmented groups with the goal of increasing market share in a more targeted fashion
Asset liability matching (ALM) analytics	Matching short-, medium-, and long-term assets and liabilities to improve liquidity and returns
Binning	A way to group numbers of more or less continuous values into a smaller number of “bins” or intervals
Business results reporting	Reporting of operating and financial data to decision-makers within an organization
Business staff	Staff primarily preoccupied with the task of outlining the problem that needs to be solved, and creating the business case for the project
Catastrophe modeling	Modeling frequency/severity of different catastrophic events and setting risk appetites
Claim flow analytics	Analyzing how best to route and settle claims to reduce loss ratios and litigation expenses
Claim loss analytics	Analyzing claims data to reduce loss ratios and improve pricing/reserving
Competitive analytics	Analyzing changing market dynamics as it relates to market competition
Concentration analytics	Identifying the potential for a group of exposures to move together in an unfavourable direction and ensuring capital adequacy
Connected home	A residence that uses internet-connected devices to enable the remote monitoring and management of appliances and systems, such as lighting and heating
Customer lifetime value (CLV) analytics	Analyzing the total amount of money a customer is expected to spend in your business during their lifetime
Customer segmentation	Dividing a company’s customers into groups that reflect similarity among customers in each group
Diffusion analytics	Analyzing the speed of market adoption
Distribution segmentation	Segmenting the sales and distribution force to match with customer segmentation
Distribution value management	Optimizing the deployment of agents by geography to maximize profitability

Term	Definition
Experience analytics	Trying to better understand customers, and improving their buying or service experience through data-driven approaches
Feature construction	A process that builds intermediate features from the original descriptors in a dataset. The aim is to build more efficient features for a machine data-mining task
Feature engineering	The process of transforming raw data into features that better represent the underlying problem to the predictive model
Feature selection	The process of selecting a subset of relevant features (variables, predictors) for use in model construction. Often used to prevent overfitting and the curse of dimensionality
Flat files	A file that typically contains one record per line, in which fields are separated either by a delimiter or have a fixed length
Fraud analytics	Detecting possible fraudulent claims using internal and external sources of data
Free-form text	Text that is found in sentences and plain language. Examples would include emails and transcripts
FTE	Full-time equivalent – a unit that indicates the workload of an employed person
GLM	Generalized linear model – a flexible generalization of ordinary linear regression that allows for response variables that have error distribution models other than a normal distribution
Growth analytics	Finding growth opportunities as they relate to acquisition, conversion, and cross-selling
Imputation	The process of replacing missing data with substituted values. For example, mean imputation would involve replacing missing values with the mean of the available values
Log transform	A method to make highly skewed distributions less skewed, whereby you take logarithms of the response
Marketing spend/mix analytics	Optimizing marketing and ad spending. Some examples include optimizing cost-per-impressions or optimizing spending allocations for different mediums
Monitoring – dashboard of actual vs expected	Testing model forecasts against actual data on an ongoing basis to see if they are well calibrated
Multi-channel optimization	Optimizing how best to tailor different sales channels to drive customer service experiences
One-hot encoding	A method to deal with categorical data by generating dummy variables
Outliers	A data point that differs significantly from other observations, often due to variability in measurement or experimental error
Policy flow analytics	Analyzing the flow of policies during the new business/enrolment process given resource and information constraints

Term	Definition
Policyholder behaviour models	Modeling policyholder behaviour to improve pricing, hedging, and reserving
Portfolio optimization	Optimizing the portfolio of investments given a set strategy and resource constraints
Pricing analytics	Pricing products and product bundles to optimize market share under given risk appetites and combined ratio constraints
Product design analytics	Using data and analytics to inform product design decisions and help in the initial pricing of said products
Product profitability analytics	Determining product profitability across different segments, geographies, and distribution channels
Program planning & execution analytics	Managing IT throughput under changing business demands and resource constraints. For example, predicting project schedule delays
Recruitment analytics	Identifying predictors of superior performance and using that to drive recruiting efforts
Reserving analysis	Leveraging analytics to help set claims reserves and IBNR
Retention analytics	Trying to understand the causes of churn and improve customer retention
Return analytics	Maximizing Rate of Return (RoR) and other return-based indicators over a given time horizon
Scaling (e.g., normalization)	The creation of shifted and scaled versions of statistics in a way that eliminates the effects of certain gross influences
Scenario analytics	Scenario testing of micro- and macro-economic drivers and shocks
Solvency models	Modeling capital requirements under different micro- and macro-economic scenarios
Technical staff	Staff primarily preoccupied with the technical demands of building the model
UBI	Usage-based insurance – leveraging external devices such as dongles, cell phones, or other means to price auto insurance based on real-time behaviour
Underwriting analytics	Quantifying and classifying a customer’s risk through data-driven insights
Unstructured data	Data that are not organized in a pre-defined manner. For example, a field containing unstructured data may contain a combination of text, PDF files, video, audio, images, email messages, and web-page data

Appendix 4 – Works cited

Adam, Tim. Daniel Kahneman: "Clearly AI is going to win. How people are going to adjust is a fascinating problem." *The Guardian*, 2021. www.theguardian.com/books/2021/may/16/daniel-kahneman-clearly-ai-is-going-to-win-how-people-are-going-to-adjust-is-a-fascinating-problem-thinking-fast-and-slow.

Adriano, Lyle. DBRS: "Bancassurance" model fails to take off in North America. *Insurance Business Canada*, 2019. www.insurancebusinessmag.com/ca/news/breaking-news/dbrs-bancassurance-model-fails-to-take-off-in-north-america-166980.aspx.

Ageas UK. App-based insurance cover. 2016. www.ageas.co.uk/press-releases/2016-press-releases/app-based-insurance-cover-back-me-up---powered-by-ageas/.

AIcrowd. Insurance pricing game: Challenges. Accessed September 10, 2021. www.aicrowd.com/challenges/insurance-pricing-game.

AIcrowd. *Insurance Pricing Game Townhall*. YouTube video, 2:14:41. 2021. www.youtube.com/watch?v=GkU2IqZu1gA.

Alderighi, Marco, Alberto A. Gaggero, and Claudio A. Piga. *The Hidden Side Of Dynamic Pricing In Airline Markets*. Munich Personal RePEc Archive, 2016. https://mpra.ub.uni-muenchen.de/71674/1/MPRA_paper_71674.pdf.

Algorithmia. *2020 State of Enterprise Machine Learning*. 2019. https://info.algorithmia.com/hubfs/2019/Whitepapers/The-State-of-Enterprise-ML-2020/Algorithmia_2020_State_of_Enterprise_ML.pdf.

Allianz. Allianz: Blockchain technology successfully piloted by Allianz Risk Transfer and Nephila for catastrophe swap. 2016. www.allianz.com/en/press/news/commitment/sponsorship/160615-blockchain-technology-successfully-piloted.html.

Allianz Global Corporate & Specialty and The Value Group. The predictive power of ESG for insurance. 2018. www.agcs.allianz.com/news-and-insights/expert-risk-articles/the-predictive-power-of-esg-for-insurance.html.

Ang, Ponora, Sébastien Richemont, and Xin Jia Wang. Insurance coverage during a pandemic and force majeure. Fasken, 2020. www.fasken.com/en/knowledge/2020/03/27-covid19-assurance-pandemie-et-force-majeure.

Aon. Aon completes acquisition of CoverWallet, the leading digital insurance platform for small and medium-sized businesses. 2020. https://aon.mediaroom.com/2019-01-07-Aon-completes-acquisition-of-CoverWallet-the-leading-digital-insurance-platform-for-small-and-medium-sized-businesses.

Apollo Insurance Solutions. Canadian insurtech Apollo closes $13.5 million Series A financing round. Cision Newswire, 2021. www.newswire.ca/news-releases/canadian-insurtech-apollo-closes-13-5-million-series-a-financing-round-855112967.html.

Ashby, Dennis, and Claus T. Jensen. *APIs for Dummies*. John Wiley & Sons and IBM, 2018. www.ibm.com/downloads/cas/GJ5QVQ7X.

Aviva Canada. Ride sharing insurance quotes. Accessed September 9, 2021. www.aviva.ca/en/find-insurance/add-ons/ride-sharing/.

AWS Amazon. Amazon S3: Object storage built to store and retrieve any amount of data from anywhere. Accessed September 10, 2021. https://aws.amazon.com/s3/.

Azevedo, Mary Ann, and Alex Wilhelm. Proptech startup States Title, now Doma, going public via SPAC in $3b deal. *TechCrunch*, 2021. https://techcrunch.com/2021/03/02/proptech-startup-states-title-now-doma-going-public-via-spac-in-3b-deal/.

Bappalige, Sachin P. An Introduction to Apache Hadoop for big data. Opensource.com, 2014. https://opensource.com/life/14/8/intro-apache-hadoop-big-data.

Barry, Laurence, and Arthur Charpentier. *Personalization as a Promise: Can Big Data Change the Practice of Insurance?* PARI, 2019. www.chaire-pari.fr/wp-content/uploads/2019/12/WP-17-Telematics.pdf.

Becker, Gregor, Anne Dreller, Anna Güntner, and Johannes-Tobias Lorenz. Behavioral science in insurance: Nudges improve decision making. McKinsey and Company, 2020. www.mckinsey.com/industries/financial-services/our-insights/insurance-blog/behavioral-science-in-insurance-nudges-improve-decision-making.

Ben-Hutta, Gabriella. NYDIG raises $100 million. *CoverageR*, 2021. https://coverager.com/nydig-raises-100-million/.

Bhandari, Aniruddha. Hadoop Ecosystem: Hadoop for big data and data engineering. *Analytics Vidhya*, 2020. www.analyticsvidhya.com/blog/2020/10/introduction-hadoop-ecosystem/.

Binder, Stephan, Ulrike Deetjen, Simon Kaesler, Jörg Mußhoff, and Felix Schollmeier. Moving to a user-first, omnichannel approach. McKinsey and Company, 2021. www.mckinsey.com/industries/financial-services/our-insights/moving-to-a-user-first-omnichannel-approach.

Bommadevara, Nagendra, Björn Münstermann, Sanaya Nagpal, and Ulrike Vogelgesang. Scale matters … to an extent: Playing the scale game in insurance. McKinsey & Company, 2021. www.mckinsey.com/industries/financial-services/our-insights/scale-matters-to-an-extent-playing-the-scale-game-in-insurance.

Bon. What is on-demand insurance? *Insurance Marketer*, 2021. www.theinsurancem.com/what-is-on-demand-insurance/.

Boulton, Clint. Insurance firm banks on change management in digital overhaul. CIO United States, 2018. www.cio.com/article/3267641/insurance-firm-banks-on-change-management-in-digital-overhaul.html.

Breiman, Leo. Statistical modeling: The two cultures (with comments and a rejoinder by the author). Statist. Sci., 2001. https://doi.org/10.1214/ss/1009213726.

Breuel, Cristiano. ML Ops: Machine learning as an engineering discipline. *Towards Data Science*, 2020. https://towardsdatascience.com/ml-ops-machine-learning-as-an-engineering-discipline-b86ca4874a3f.

Brockbank, Nicole. Fewer Toronto restaurants cancelled business licences in first year of Covid-19 than normal. CBCnews, CBC/Radio Canada, March 29, 2021. www.cbc.ca/news/canada/toronto/toronto-business-licences-covid-1.5965568.

Brown, Eileen. Next Insurance launches Facebook Messenger chatbot to replace the insurance agent. *Social Business*, 2017. www.zdnet.com/article/next-insurance-launches-facebook-messenger-chatbot-to-replace-the-insurance-agent/.

Calvert, Scott. Rise in car crash deaths prompts new seat-belt push. *Wall Street Journal*, 2021. www.wsj.com/articles/rise-in-car-crash-deaths-prompts-new-seat-belt-push-11627637400.

Capgemini. Capgemini perspectives: Cloud native comes of age in insurance. 2018. www.capgemini.com/article/cloud-native-comes-of-age-in-insurance/.

Carpe Data. The Hartford taps new data sources for small business underwriting. 2020. https://carpe.io/the-hartford-taps-carpe-data-for-small-business-data/.

CarrierHQ. Aon: CarrierHQ's Small Fleet Advantage adjustable rate insurance for trucking wins 2021 Celent Model Insurer Award for data, analytics, and AI. Cision PR Newswire, 2021. www.prnewswire.com/news-releases/aon---

carrierhqs-small-fleet-advantage-adjustable-rate-insurance-for-trucking-wins-2021-celent-model-insurer-award-for-data-analytics-and-ai-301250532.html.

Casualty Actuarial Society. Charles A. Hachemeister Prize. Accessed September 10, 2021. www.casact.org/about/awards-prizes-scholarships/charles-hachemeister-prize.

Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rüdiger Wirth. *Step-by-Step Data Mining Guide*. SPSS, 2020. www.the-modeling-agency.com/crisp-dm.pdf.

Charpentier, Arthur. *Insurance: Risk Pooling and Price Segmentation*. ESSEC Paris, 2017. http://freakonometrics.free.fr/slides-essec-2017.pdf.

Clague, Chris. The perfect time for tech in insurance. The Economist Intelligence Unit, 2020. https://eiuperspectives.economist.com/financial-services/perfect-time-tech-insurance.

Coalition Against Insurance Fraud. *2020 Insurer SIU Benchmarking Study: Insurers Finding Stability in Their Anti-Fraud Units*. 2020. https://insurancefraud.org/wp-content/uploads/Benchmarking-Study-Summary.pdf.

Co-operators Group. New insurtech partnership to provide on-demand insurance in Canada. 2018. https://newsreleases.cooperators.ca/2018-07-18-New-Insurtech-Partnership-to-Provide-On-Demand-Insurance-in-Canada.

Dangerfield, Katie. "Inevitable" 9.0 earthquake, tsunami will hit Canada's West Coast: Expert. *Global News*, 2020. https://globalnews.ca/news/3981536/tsunami-earthquake-canada-the-big-one/.

Daninhirsch, Hilary. Some fleets turn to cameras to help mitigate rising insurance costs. *Transport Topics*, 2020. www.ttnews.com/articles/some-fleets-turn-cameras-help-mitigate-rising-insurance-costs.

Dastin, Jeffrey. Amazon scraps secret AI recruiting tool that showed bias against women. Reuters, 2018. www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G.

Delcea, Ramona. Discussion paper on blockchain and smart contracts in insurance: EIOPA invites comments. European Insurance and Occupational Pensions Authority, 2021. www.eiopa.europa.eu/content/discussion-paper-blockchain-and-smart-contracts-insurance-eiopa-invites-comments-0_en.

Deloitte. *Speed to Market: Part of the Insurance series – Benefits of a New Policy Administration System: Why Going Live Is Not Enough*. 2015. www2.deloitte.com/content/dam/Deloitte/us/Documents/financial-services/us-cons-policy-admin-systems-speed-to-market-042415.pdf.

Demarest, George. Four phases of operating big data. *CIO Review*. Accessed September 14, 2021. https://bigdata.cioreview.com/cxoinsight/four-phases-of-operationalizing-big-data-nid-15251-cid-15.html.

Dhieb, Najmeddine, et al. A secure AI-driven architecture for automated insurance systems: Fraud detection and risk measurement. *IEEE Access*, 2020. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9046765.

Diffey, William, Malcolm Cleugh, Laura Hobern, and Ed Harrison on behalf of members of the TORP Working Party. *Can You Trust Your Reserving? Reserving Validation Under Covid-19*. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/system/files/field/document/Can%20you%20trust%20your%20reserving%20-%20Reserving%20validation%20under%20Covid-19%20-%20v1.3.pdf.

DMZ. Aviva and DMZ are working to make Toronto the insurtech capital of Canada. 2019. https://dmz.ryerson.ca/partner_profiles/aviva/.

Došilović, Filip Karlo, Mario Brčić, and Nikica Hlupić. *Explainable Artificial Intelligence: A Survey*. IEEE, 2018. https://ieeexplore.ieee.org/document/8400040.

Dsouza, Jason. What is a GPU and do you need one in deep learning? *Towards Data Science*, 2020. https://towardsdatascience.com/what-is-a-gpu-and-do-you-need-one-in-deep-learning-718b9597aa0d.

EdX team. 9 top programming languages for data science. EdX Blog, 2021. https://blog.edx.org/9-top-programming-languages-for-data-science.

Environments. Reproducible environments. Accessed September 10, 2021. https://environments.rstudio.com/.

European Insurance and Occupational Pensions Authority. *Big Data Analytics in Motor and Health Insurance: A Thematic Review*. 2019. https://register.eiopa.europa.eu/Publications/EIOPA_BigDataAnalytics_ThematicReview_April2019.pdf.

Evans, Nadine. IoT and the reduction of claims – *Canadian Underwriter*. Eddy Solutions, 2019. https://eddysolutions.com/iot-and-the-reduction-of-claim-canadian-underwriter/.

Fannin, Brian A. *COVID-19: The Property-Casualty Perspective*. Casualty Actuarial Society, 2020. www.casact.org/sites/default/files/2021-03/COVID-19_The_PC_Perspective_3-27-2020.pdf.

FAT/ML. Fairness, Accountability, and Transparency in Machine Learning. Accessed October 12, 2021. www.fatml.org/.

FICO Community. Explainable Machine Learning Challenge. Accessed September 10, 2021. https://community.fico.com/s/explainable-machine-learning-challenge.

Financial Conduct Authority. *Call for Inputs on Big Data in Retail General Insurance*. 2016. www.fca.org.uk/publication/feedback/fs16-05.pdf.

Financial Conduct Authority. *General Insurance Pricing Practices*. 2020. www.fca.org.uk/publication/market-studies/ms18-1-3.pdf.

Financial Post. Canadian economy suffers biggest contraction since the Great Depression. 2021. https://financialpost.com/news/economy/canadian-press-newsalert-canadian-economy-contracted-5-4-per-cent-in-2020.

Financial Stability Board. *Artificial Intelligence and Machine Learning in Financial Services: Market Developments and Financial Stability Implications*. 2017. www.fsb.org/wp-content/uploads/P011117.pdf.

Finextra. AXA and Trov bring “on demand” insurance to UK. 2016. www.finextra.com/newsarticle/29804/axa-and-trov-bring-on-demand-insurance-to-uk.

Flavelle, Christopher. Scorched, parched and now uninsurable: Climate change hits wine country. *New York Times*, 2021. www.nytimes.com/2021/07/18/climate/napa-wine-heat-hot-weather.html.

Formstack. How to A/B test your forms for maximum conversion. Accessed September 14, 2021. www.formstack.com/resources/guide-ab-test-web-forms-maximum-conversion.

Friedland, Jacqueline. *Survey of Canadian Actuaries on ML in Reserving*. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/system/files/field/document/MLR_CanadaSurvey.pdf.

Galov, Gerashe. Hadoop filesystem at Twitter. Twitter (blog), 2015. https://blog.twitter.com/engineering/en_us/a/2015/hadoop-filesystem-at-twitter.

Gambrill, David. How insurance company websites just beat out brokers in a U.S. consumer satisfaction study. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/insurance/why-insurance-company-websites-just-beat-out-brokers-in-a-u-s-consumer-satisfaction-study-1004179552/.

Gardiner, Mark. A car insurance claim estimate before the tow truck is called. *New York Times*, 2020. www.nytimes.com/2020/09/17/business/car-insurance-claim-estimate-artificial-intelligence.html.

Garg, Amit, Davide Grande, Gloria Macías-Lizaso, and Christoph Sporleder. Analytics in banking: Time to realize the value. McKinsey and Company, 2017. www.mckinsey.com/industries/financial-services/our-insights/analytics-in-banking-time-to-realize-the-value.

Geddes Baribeau, Annmarie. Domestic perils: 2020 – a pivotal year for homeowners insurance. *Actuarial Review*, 2021. https://ar.casact.org/domestic-perils-2020-a-pivotal-year-for-homeowners-insurance/.

Geddes Baribeau, Annmarie. Getting personal – Can IoT do for homeowners insurance what telematics did for auto coverage? *Actuarial Review*, 2021. https://ar.casact.org/getting-personal-can-iot-do-for-homeowners-insurance-what-telematics-did-for-auto-coverage/.

Geddes Baribeau, Annmarie. Insurers enjoy benefits from data modeling the claims process. *Actuarial Review*, 2020. https://ar.casact.org/insurers-enjoy-benefits-from-data-modeling-the-claims-process/.

Gill, Navdeep, Patrick Hall, Kim Montgomery, and Nicholas Schmidt. 2020. A responsible machine learning workflow with focus on interpretable models, post-hoc explanation, and discrimination testing. *Information*, 2020. https://doi.org/10.3390/info11030137.

Goldberg, Jeff. The 3 pillars of on-demand insurance. Insurance Thought Leadership, 2018. www.insurancethoughtleadership.com/the-3-pillars-of-on-demand-insurance/.

Goldburd, Mark, Anand Khare, Dan Tevet, and Dmitriy Guller. *Generalized Linear Models for Insurance Rating*, 2nd edition. Casualty Actuarial Society, 2020. www.casact.org/sites/default/files/2021-03/8_GLM.pdf.

Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. *Explaining and Harnessing Adversarial Examples*. Google, 2015. https://arxiv.org/pdf/1412.6572.pdf.Google Cloud. Allstate: Helping agents build better relationships with customers. Accessed September 14, 2021. https://cloud.google.com/customers/allstate.

Google Cloud. Cloud Tensor Processing Units (TPUs). Accessed September 10, 2021. https://cloud.google.com/tpu/docs/tpus. https://cloud.google.com/tpu/docs/tpus.

Google Cloud. DevOps tech: Continuous testing. Accessed September 14, 2021. https://cloud.google.com/architecture/devops/devops-tech-test-automation.

Google Cloud. MLOps: Continuous delivery and automation pipelines in machine learning. 2020. https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning.

Grzadkowska, Alicja. Cancellation of Tokyo Olympics could cripple the insurance industry. *Insurance Business Canada*, 2021. www.insurancebusinessmag.com/ca/news/columns/cancellation-of-tokyo-olympics-could-cripple-the-insurance-industry-246029.aspx.

Gunning, David, and David Aha. DARPA's explainable artificial intelligence (XAI) Program. *AI Magazine*, 2019. https://doi.org/10.1609/aimag.v40i2.2850.

Guszcza, Jim. The last-mile problem: How data science and behavioral science can work together. Deloitte Insights, 2015. www2.deloitte.com/us/en/insights/deloitte-review/issue-16/behavioral-economics-predictive-analytics.html.

Han, Basil. Improving fraudulent claims detection with AI. AI Singapore, 2021. https://aisingapore.org/2021/07/improving-fraudulent-claims-detection-with-ai/.

Hancock, Matt. *Data Science Ethical Framework*. Cabinet Office United Kingdom, 2016. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/524298/Data_science_ethics_framework_v1.0_for_publication__1_.pdf.

Harrison, Nick, and Deborah O'Neill. If your company isn't good at analytics, it's not ready for AI. *Harvard Business Review*, 2017. https://hbr.org/2017/06/if-your-company-isnt-good-at-analytics-its-not-ready-for-ai.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, 2nd edition. Stanford Web Spring, 2017. https://web.stanford.edu/~hastie/ElemStatLearn/printings/ESLII_print12_toc.pdf.

Hope, Bradley, and Nicole Friedman. Climate change is forcing the insurance industry to recalculate. *Wall Street Journal*, 2018. https://web.archive.org/web/20181207130059/www.wsj.com/graphics/climate-change-forcing-insurance-industry-recalculate/.

Huang, Wanwan. *The Development of WeChat Marketing and Distribution of Insurance Products in China*. Society of Actuaries, 2018. www.soa.org/globalassets/assets/files/resources/research-report/2018/wechat-marketing-distribution.pdf.

Huzinga, Neil. Nudge theory and insurtech – Happy bedfellows? Insurance-Canada.ca (blog), 2018. www.insurance-canada.ca/2018/01/11/nudge-theory-insurtech-happy-bedfellows/.

IBM Services. IFFCO Tokio General Insurance Company Limited: Improving customer experience with smarter solutions. 2020. www.ibm.com/case-studies/iffco-tokio-ibm-services-ai.

Institute and Faculty of Actuaries. General insurance machine learning in reserving. 2020. www.actuaries.org.uk/practice-areas/general-insurance/research-working-parties/general-insurance-machine-learning-reserving.

Institute and Faculty of Actuaries. Pragmatic stochastic reserving. Accessed January 26, 2022. www.actuaries.org.uk/practice-areas/general-insurance/disbanded-research-working-parties/pragmatic-stochastic-reserving.

Institute and Faculty of Actuaries and Royal Statistical Society. *A Guide for Ethical Data Science*. 2019. www.actuaries.org.uk/system/files/field/document/An%20Ethical%20Charter%20for%20Date%20Science%20WEB%20FINAL.PDF.

Insurance-Canada.ca. Aon and Zesty.ai revolutionize underwriting with property data solution powered by artificial intelligence. 2019. www.insurance-canada.ca/2019/03/13/aon-zesty-ai-property-data-solution/.

Insurance-Canada.ca. Intact partners with Snapsheet to offer photo claims estimating service. 2019. www.insurance-canada.ca/2019/06/21/intact-snapsheet-photo-claims-estimating/.

Insurance-Canada.ca. Trov technology enables a new wave of consumer brands to offer digital renters insurance. 2021. www.insurance-canada.ca/2021/04/08/trov-enables-new-brands-digital-renters-insurance/.

Insurance Information Institute. Background on: Insurance fraud. 2021. www.iii.org/article/background-on-insurance-fraud.

Insurance Journal. Internet of Things devices increase risk of cyber attacks on industrial sector: Lloyd's. 2021. www.insurancejournal.com/news/international/2021/02/17/601582.htm.

Insurance Journal. Liberty Mutual giving consumers a voice in insurance via Amazon's Alexa. 2016. www.insurancejournal.com/news/national/2016/09/13/426162.htm.

Jamal, Salma, Stefano Canto, Ross Fernwood, Claudio Giancaterino, Munir Hiabu, Lorenzo Invernizzi, Tetiana Korzhynska, Zachary Martin, and Hong Shen. *Machine Learning and Traditional Methods Synergy in Non-Life Reserving*. ASTIN, 2018. www.actuaries.org/IAA/Documents/ASTIN/ASTIN_MLTMS%20Report_SJAMAL.pdf.

James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning with Applications in R*, 2nd edition. Stanford Web, 2021. www.statlearning.com/.

Josefowicz, Matthew, and Harry Huberty. Speed to market for property/casualty insurers. AiteNovarica, 2021. https://novarica.com/speed-to-market-for-property-casualty-insurers/.

Karakan, Burak. Python vs R for data science. *Towards Data Science*, 2020. https://towardsdatascience.com/python-vs-r-for-data-science-6a83e4541000.

Kost, Danielle. You're right! You are working longer and attending more meetings. Harvard Business School Working Knowledge, 2020. https://hbswk.hbs.edu/item/you-re-right-you-are-working-longer-and-attending-more-meetings.

Krizsan, Erika. Insuretech boosts China's online insurance market. *Insurance Factory*, 2020. https://insurance-factory.eu/insuretech-boosts-chinas-online-insurance-market/.

Kuo, Kevin. *DeepTriangle: A Deep Learning Approach to Loss Reserving*. Cornell University (arXiv.org), 2019. https://arxiv.org/pdf/1804.09253.pdf.

Kurmelovs, Royce. Climate change could put insurance out of reach for many Australians. *The Guardian*, 2021. www.theguardian.com/australia-news/2021/mar/02/climate-change-could-put-insurance-out-of-reach-for-many-australians.

Labram, Alex. The machine learning landscape. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/news-and-insights/news/machine-learning-landscape.

Lamb, Evelyn. Review: *Weapons of Math Destruction*. *Scientific American*, 2016. https://blogs.scientificamerican.com/roots-of-unity/review-weapons-of-math-destruction/.

Leroy, Sophie, and Theresa M. Glomb. A plan for managing (constant) interruptions at Work. *Harvard Business Review*, 2020. https://hbr.org/2020/06/a-plan-for-managing-constant-interruptions-at-work.

LexisNexis Risk Solutions. 2019 study results: How U.S. insurance carriers are using artificial intelligence and machine learning. 2021. https://risk.lexisnexis.com/insights-resources/research/state-of-ai-ml-in-the-insurance-industry.

Li, Joel, Rolly Molisho, and Harrison Jones. An introduction to AI ethics and regulation. *Seeing Beyond Risk*, 2021. www.seeingbeyondrisk.ca/2021/09/ai-ethics-and-regulation-in-insurance-actuaries-uniquely-positioned-for-success/.

Liang, Chen, Ziqi Liu, Bin Liu, Jun Zhou, Xiaolong Li, Shuang Yang, and Yuan Qi. *Uncovering Insurance Fraud Conspiracy with Network Learning*. Cornell University (arXiv.org), 2020. https://arxiv.org/pdf/2002.12789v1.pdf.

Liao, Xiyue, Guoqiang Chen, Ben Ku, Rahul Narula, and Janet Duncan. Text mining methods applied to insurance company customer calls: A case study. *North American Actuarial Journal*, 2020. www.semanticscholar.org/paper/Text-Mining-Methods-Applied-to-Insurance-Company-A-Liao-Chen/9bc53ae539fa6a96ea974f77768544b545aabb9a.

Light, Donald. *Building a First-to-Market Pay as You Drive Offering*. Guidewire, 2021. https://explore.guidewire.com/c/report-celent-buildi?x=-QNqMK.

Livak, Paul. 4 components to a modern insurance data strategy. *Digital Insurance*, 2020. www.dig-in.com/opinion/pwc-insurance-4-key-components-to-a-data-strategy.

Lorenz, Jinjer. Farmers Edge and Munich Re announce strategic partnership to implement large-scale parametric weather insurance solutions. Farmers Edge, 2020. www.farmersedge.ca/farmers-edge-and-munich-re-announce-strategic-partnership-to-implement-large-scale-parametric-weather-insurance-solutions/.

Ludwig, Sarah. Credit scores in America perpetuate racial injustice. *The Guardian*, 2015. www.theguardian.com/commentisfree/2015/oct/13/your-credit-score-is-racist-heres-why.

Lyons, Carol. Does Canada need a terrorism risk insurance scheme? McMillan, 2015. https://mcmillan.ca/insights/does-canada-need-a-terrorism-risk-insurance-scheme-2/.

Macías-Lizaso Miranda, Gloria. Building an effective analytics organization. McKinsey and Company, 2018. www.mckinsey.com/industries/financial-services/our-insights/building-an-effective-analytics-organization.

Mainelli, Michael, and Bernard Manson. *Chain Reaction: How Blockchain Technology Might Transform Wholesale Insurance*. PwC Global, 2016. www.pwc.com/gx/en/financial-services/pdf/how-blockchain-tecnology-might-transform-insurance.pdf.

Mayorga, Wilson, and Diego Torres. *A Practical Model for Pricing Optimization in Car Insurance*. ASTIN/AFIR-ERM Colloquium, Panama, 2017. www.actuaries.org/panama2017/docs/papers/3b_ASTIN_Paper_Mayorga.pdf.

McGuire, Grainne, and Jacky Poon. ML modelling on triangles: A worked example. Institute and Faculty of Actuaries, 2021. https://institute-and-faculty-of-actuaries.github.io/mlr-blog/post/f-mlr3example/.

Meckbach, Greg. Aviva rolls out automated claim notification at collision reporting centres. *Canadian Underwriter*, 2019. www.canadianunderwriter.ca/insurance/this-insurer-rolling-out-automated-claim-notification-at-collision-reporting-centres-1004159316/.

Meckbach, Greg. New industry anti-fraud group gets its first CEO. *Canadian Underwriter*, 2021. www.canadianunderwriter.ca/associations/new-industry-anti-fraud-group-gets-its-first-ceo-1004210351/.

Meckbach, Greg. The tough question for insurers withdrawing coverage from coal. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/climate-change/the-tough-question-for-insurers-withdrawing-coverage-from-coal-1004198475/.

Meckbach, Greg. Who is funding many of these insurance mergers. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/mergers-and-aqcuisitions/who-is-funding-many-of-these-insurance-mergers-1004200880/.

Meckbach, Greg. Why Uber Canada dropped Intact as its insurance provider. *Canadian Underwriter*, 2020. www.canadianunderwriter.ca/insurance/why-uber-canada-dropped-intact-as-its-insurance-provider-1004196996/.

Mei, Maarssen. *Insurance 2020 & Beyond: The Current Agenda of the CEO*. PwC Netherlands, 2017. www.ag-ai.nl/view/34908-JH+Lobregt+Insurance+2020+%26+beyond.pdf.

Microsoft Docs. Identify guiding principles for responsible AI – State Farm case study. Accessed September 14, 2021. https://docs.microsoft.com/en-us/learn/modules/responsible-ai-principles7-responsible-ai-case-study.

Micu, Eliade. *Territorial Ratemaking*. Casualty Actuarial Society and Eagle Eye Analytics, 2012. https://cas.confex.com/cas/rpms12/webprogram/Presentation/Session4723/Terr%20Ratemaking%20EEA%20v2.pdf.

Miehe, Pierre, Judith Lutz, Accroche-com, and Fabrice Taillieu. *Non-Life Reserving Practices*. ASTIN Working Party on Non-life Reserving Practices, 2016. www.actuaries.org/ASTIN/Documents/ASTIN_WP_NL_Reserving_Report1.0_2016-06-15.pdf.

MikeRayMSFT, julieMSFT, cawrites, icoric, mindlessroman, markingmyname, CarlRabeler, pmasl, PRMerger16, MashaMSFT, WilliamAntonRohm, craigg-msft and edmacauley. *Char and varchar (Transact-SQL)*. SQL Server/Microsoft Docs, 2019. https://docs.microsoft.com/en-us/sql/t-sql/data-types/char-and-varchar-transact-sql?view=sql-server-ver15.

Mims, Christopher. Self-driving cars could be decades away, no matter what Elon Musk said. *Wall Street Journal*, 2021. www.wsj.com/articles/self-driving-cars-could-be-decades-away-no-matter-what-elon-musk-said-11622865615.

Molnar, Christoph. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*. Github, 2021. https://christophm.github.io/interpretable-ml-book/.

Morrison, Sara. A disturbing, viral Twitter thread reveals how AI-powered insurance can go wrong. *Vox*, 2021. www.vox.com/recode/22455140/lemonade-insurance-ai-twitter.

MuleSoft. What is an API? (Application Programming Interface). Accessed September 14, 2021. www.mulesoft.com/resources/api/what-is-an-api.

Müller, Katja, Hato Schmeiser, and Joël Wagner. Insurance claims fraud: Optimal auditing strategies in insurance companies. *Variance Journal*, 2016. www.casact.org/abstract/insurance-claims-fraud-optimal-auditing-strategies-insurance-companies.

Nadarajah, Indrani. Auto insurance fraud. Insurance Institute of Canada, 2018. www.insuranceinstitute.ca/en/cipsociety/information-services/advantage-monthly/0718-insurance-fraud.

National Association of Insurance Commissioners. Chatbots. 2020. https://content.naic.org/cipr_topics/topic_chatbots.htm.

National Association of Insurance Commissioners. On-demand insurance. 2021. https://content.naic.org/cipr_topics/topic_ondemand_insurance.htm.

National Insurance Crime Bureau. Slip & fall incidents rise according to the National Insurance Crime Bureau. 2021. www.nicb.org/news/news-releases/slip-fall-incidents-rise-according-national-insurance-crime-bureau.

Novarica. Insurer speed to market depends on process as much as technology. Insurance-Canada.ca, 2019. www.insurance-canada.ca/2019/03/26/novarica-speed-to-market-process.

Nuruzzaman, Mohammad, and Omar Khadeer Hussain. IntelliBot: A dialogue-based chatbot for the insurance industry. *Science Direct*, 2020. www.sciencedirect.com/science/article/abs/pii/S0950705120301933.

Nuruzzaman, Mohammad, and Omar Khadeer Hussain. *IntelliBot: A Domain-Specific Chatbot for the Insurance Industry* (awarded by University of New South Wales, Business, 2020). www.unsworks.unsw.edu.au/primo-explore/fulldisplay/unsworks_72771/UNSWORKS.

Olah, Christopher. Understanding LSTM networks. Colah's blog (Github), 2015. https://colah.github.io/posts/2015-08-Understanding-LSTMs/.

Papers With Code. The latest in machine learning. Accessed September 10, 2021. https://paperswithcode.com/.

Parameshwaran, Reni, Himadri Sikhar Pramanik, Sayantan Datta, and Ujjwal Bunkar. *On-Demand Insurance: Challenges and Opportunities for Large Insurance Carriers*. TATA Consulting Services, 2019. www.tcs.com/content/dam/tcs/pdf/Industries/insurance/rise-of-on-demand-insurance.pdf.

Pearson. Storing and retrieving images in JDBC. InformIT, 2002. www.informit.com/articles/article.aspx?p=25280.

Perkins, Steven, Hazel Davis, and Valerie du Preez. *Practical Data Science for Actuarial Tasks*. Institute and Faculty of Actuaries, 2020. www.actuaries.org.uk/system/files/field/document/Practical%20Data%20Science%20for%20Actuarial%20Tasks%20v1.8.pdf.

Picozzi, Jeff. What APIs mean for an open and connected insurance industry. Red Hat Blog, 2020. www.redhat.com/en/blog/what-apis-mean-open-and-connected-insurance-industry.

Plumer, Sloan, and Scott Busse. How insurance carriers are modernizing to cloud based analytics. LinkedIn (article), PwC Advisory United States, 2020. www.linkedin.com/pulse/how-insurance-carriers-modernizing-cloud-based-analytics-sloan-plumer/.

Plumer, Sloan, Imran Ilyas, Josh Knipp, Yasir Safdar, and Tirath Desai. Eight considerations for insurance carriers migrating to the cloud. LinkedIn (article), PwC Advisory United States, 2020. www.linkedin.com/pulse/eight-considerations-insurance-carriers-migrating-cloud-sloan-plumer/?trackingId=4ER%2BMP1wTuyXxfsyNcIq5Q%3D%3D.

Podder, Sohini. Home insurance agency Hippo to go public in $5b SPAC. *Insurance Journal*, 2021. www.insurancejournal.com/news/national/2021/03/04/603732.htm.

Press, Gil. AI stats news: Chatbots lead to 80% sales decline, satisfied customers and fewer employees. *Forbes*, 2019. www.forbes.com/sites/gilpress/2019/09/25/ai-stats-news-chatbots-lead-to-80-sales-decline-satisfied-customers-and-fewer-employees/?sh=5335ba948e05.

PwC Canada. The retail landscape of the future: Canadian Consumer Insights 2021, Pulse 1. 2021. www.pwc.com/ca/en/industries/consumer-markets/consumer-insights-2021.html.

PwC United States. Ethical AI: Tensions and trade-offs. 2019. https://www.pwc.com.au/digitalpulse/ethical-artificial-intelligence-tensions-trade-offs.html

PwC Canada. Understanding the Canadian consumer of the moment: Canadian Consumer Insights 2021, Pulse 2. 2021. www.pwc.com/ca/en/industries/consumer-markets/consumer-insights-2021-pulse-2.html.

PwC Global. *A Practical Guide to Responsible Artificial Intelligence (AI)*. 2019. www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.pdf.

PwC Global. Insurance trends 2020: Moving from resilience to reinvention will help insurers succeed in uncertain times. 2020. www.pwc.com/gx/en/ceo-agenda/ceosurvey/2020/trends/insurance.html.

PwC Hong Kong. *Adding It All Up: Modern Rating Systems for P&C Carriers*. 2015. www.pwchk.com/en/migration/pdf/modern-rating-system-apr2015.pdf.

PwC Hong Kong. *Insurance Fraud Analytics*. 2017. www.pwccn.com/en/risk-assurance/publications/insurance-fraud-analytics.pdf.

PwC Legal Estonia. *Blockchain: The $5 Billion Opportunity for Reinsurers*. 2016. www.pwc.com/ee/et/publications/pub/blockchain-for-reinsurers.pdf.

PwC United Kingdom. *Explainable AI: Driving Business Value Through Greater Understanding*. Accessed September 10, 2021. www.pwc.co.uk/audit-assurance/assets/explainable-ai.pdf.

PwC United States. 2019 AI predictions: Six AI priorities you can't afford to ignore. 2019. https://web.archive.org/web/20211109223336/www.pwc.com/us/en/services/consulting/library/artificial-intelligence-predictions-2019.html.

PwC United States. How can US insurers become more innovative? Look to Asia. 2020. www.pwc.com/us/en/industries/insurance/library/insights-from-asia.html.

PwC United States. Insurance claims estimator uses AI for efficiency. 2017. www.pwc.com/us/en/library/case-studies/auto-insurance-ai-analytics.html.

PwC United States. Moment of truth: Why insurance carriers should rethink their customer service models now. 2020. www.pwc.com/us/en/industries/insurance/library/carriers-customer-service-models-covid-19.html.

PwC United States. Top insurance industry issues in 2021: Talent strategies for today's insurers. 2021. www.pwc.com/us/en/industries/insurance/library/top-issues/talent-strategies.html.

Property and Casualty Insurance Compensation Corporation. *Insolvency Protection for Home, Automobile and Business Insurance Customers, Annual Report 2019*. 2020. www.pacicc.ca/wp-content/uploads/2020/04/PACICC-2019-Annual-Report-ENG.pdf.

Property and Casualty Insurance Compensation Corporation. *Solvency Matters: A Quarterly Report on Solvency Issues Affecting P&C Insurers In Canada*. 2021. www.pacicc.ca/wp-content/uploads/2021/06/Solvency_Matters_14_June.pdf.

Radečić, Dario. PyTorch: Switching to the GPU. *Towards Data Science*, 2020. https://towardsdatascience.com/pytorch-switching-to-the-gpu-a7c0b21e8a99.

Rao, Anand, Jamie Yoder, and Scott Busse. *AI in Insurance: Hype or Reality? The Digital Insurer*. PwC, 2016. www.the-digital-insurer.com/wp-content/uploads/2016/06/716-pwc-top-issues-artificial-intelligence.pdf.

Rao, Anand S., and Ilana Golbin. What is fair when it comes to AI bias? *Strategy + Business*, 2019. www.strategy-business.com/article/What-is-fair-when-it-comes-to-AI-bias.

Rao, Anand S., and Gerard Verweij. Sizing the prize: What's the real value of AI for your business and how can you capitalise? PwC Global, 2017. www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html.

REST API Tutorial. HTTP methods. Accessed September 14, 2021. https://restfulapi.net/http-methods/.

Richman, Ronald. *AI in Actuarial Science*. Presented at the Actuarial Society of South Africa's 2018 Convention, October 24–25, Cape Town. www.actuarialsociety.org.za/wp-content/uploads/2018/10/2018-Richman-FIN.pdf.

Ricker, Thomas. Watch a drone hack a room full of smart lightbulbs from outside the window. *The Verge*, 2016. www.theverge.com/2016/11/3/13507126/iot-drone-hack.

Rogati, Monica. The AI hierarchy of needs. *Hackernoon*, 2017. https://hackernoon.com/the-ai-hierarchy-of-needs-18f111fcc007.

Sandhu, Angat, Steven Chen, Ajit Rochlani, Jun Hao Tay, and Bella Thamrin. *Insurance Redefined: A Roadmap for Insurers and Insurtechs*. Oliver Wyman and Singapore Fintech Association, 2020. www.oliverwyman.com/content/dam/oliver-wyman/v2/publications/2020/dec/insurance-redefined.pdf.

Sawers, Paul. Pay-per-mile car insurance company Metromile raises $90 million to automate the claims process. *VentureBeat*, 2018. https://venturebeat.com/2018/07/24/pay-per-mile-car-insurance-company-metromile-raises-90-million-to-automate-the-claims-process/.

Schreiber, Daniel. AI can vanquish bias: Algorithms we can't understand can make insurance fairer. Lemonade (blog). Accessed September 10, 2021. www.lemonade.com/blog/ai-can-vanquish-bias/.

Schwartz, Arthur J. *Price Optimization and Insurance Regulation with Examples and Calculations*. Spring Meeting of the Casualty Actuarial Society in Colorado Springs, CO, May 2015. www.casact.org/sites/default/files/presentation/spring_2015_handouts_c-21.pdf.

Schwartz, Oscar. In 2016, Microsoft's racist chatbot revealed the dangers of online conversation: The bot learned language from people on Twitter – but it also learned values. *IEEE Spectrum*, 2019. https://spectrum.ieee.org/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation.

Scism, Leslie. Some California homeowners can get coverage again after wildfires. *Wall Street Journal*, 2021. www.wsj.com/articles/some-california-homeowners-can-get-coverage-again-after-wildfires-11623589200.

Semenovich, Dimitri. *Applications of Convex Optimization in Premium Rating*. Casualty Actuarial Society, 2013. www.casact.org/sites/default/files/database/forum_13spforum_semenovich.pdf.

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, and Dharshan Kumaran. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 2018. https://science.sciencemag.org/content/362/6419/1140.

Skiba, Michael, Jeffrey G. Rapattoni, and Chris McKibbin. *Secrets to Combating Insurance Fraud with Data Analytics: Three Insurance Executives Offer a Global Perspective*. Casualty Actuarial Society, 2018. www.casact.org/sites/default/files/presentation/spring_2019_presentations_g-2_rapattoni_1.pdf.

Society of Actuaries. Ethical & Responsible Use of Data & Predictive Models Certificate Program. Accessed October 5, 2021. www.soa.org/programs/ethical-responsible-data-certificate/?utm_medium=Email&utm_source=SNWArticle&utm_campaign=ERUCert_2021&utm_content=2021-08-25.

Somasundaram, Srinivasan, Akshat Kant, and Prakhar Maheshwari. *The Future of Chatbots in Insurance*. Cognizant, 2019. www.cognizant.com/whitepapers/the-future-of-chatbots-in-insurance-codex4122.pdf.

Spedicato, Giorgio, Christophe Dutang, and Leonardo Petrini. *Machine Learning Methods to Perform Pricing Optimization: A Comparison with Standard Generalized Linear Models*. Casualty Actuarial Society, 2018. www.casact.org/abstract/machine-learning-methods-perform-pricing-optimization-comparison-standard-generalized.

Stack Overflow. Stack Overflow Developer Survey 2021. 2021. https://insights.stackoverflow.com/survey/2021#technology.

Synced. Tree boosting with XGBoost: Why does XGBoost win “every” machine learning competition? 2017. https://syncedreview.com/2017/10/22/tree-boosting-with-xgboost-why-does-xgboost-win-every-machine-learning-competition/.

Tang, Tjun, Michelle Hu, and Angelo Candreia. Why Chinese insurers lead the way in digital innovation. Boston Consulting Group, 2018. www.bcg.com/en-ca/publications/2018/chinese-insurers-digital-innovation.

Taulli, Tom. Lemonade IPO shows the power of AI (artificial intelligence). *Forbes*, 2020. www.forbes.com/sites/tomtaulli/2020/07/03/lemonade-ipo-shows-the-power-of-ai-artificial-intelligence/?sh=8263f053aebb.

Taulli, Tom. What you need to know about dark data. *Forbes*, 2019. www.forbes.com/sites/tomtaulli/2019/10/27/what-you-need-to-know-about-dark-data/?sh=770d4bdf2c79.

Teradata. What are the 5 V’s of Big Data? Accessed March 1, 2022. https://www.teradata.com/Glossary/What-are-the-5-V-s-of-Big-Data.

Tractable. MS&AD to use Tractable’s AI across Japan to accelerate recovery from auto accidents. Cision PR Newswire, 2020. www.prnewswire.com/news-releases/msad-to-use-tractables-ai-across-japan-to-accelerate-recovery-from-auto-accidents-301162486.html.

Trufla Technology. Trufla Technology celebrates Bullfrog Insurance’s innovative new website. *Canadian Underwriter*, 2018. www.canadianunderwriter.ca/inspress/trufla-technology-celebrates-bullfrog-insurances-innovative-new-website/.

VMware Security and Compliance Blog. Amid Covid-19, global orgs see a 148% spike in ransomware attacks; finance industry heavily targeted. 2020. https://blogs.vmware.com/security/2020/04/amid-covid-19-global-orgs-see-a-148-spike-in-ransomware-attacks-finance-industry-heavily-targeted.html.

Weed, Julie. Wearable tech that tells drowsy truckers it’s time to pull over. *New York Times*, 2020. www.nytimes.com/2020/02/06/business/drowsy-driving-truckers.html.

Whittaker, Zack. Security flaws in a popular smart home hub let hackers unlock front doors. *TechCrunch*, 2019. https://techcrunch.com/2019/07/02/smart-home-hub-flaws-unlock-doors/.

Wikimedia Commons. Phases of the CRISP-DM reference model. 2012. https://commons.wikimedia.org/wiki/File:CRISP-DM_Process_Diagram.png

Willis Towers Watson. 2019/2020 P&C Insurance Advanced Analytics Survey Report (North America): Fields of dreams – Three areas dominate the field of insurers’ aspirations for advanced analytics. 2020. www.willistowerswatson.com/en-CA/Insights/2020/01/fields-of-dreams-three-areas-dominate-the-field-of-insurers-aspirations-for-advanced-analytics.

Willis Towers Watson. Quarterly Insurtech Briefing Q2 2021. 2021. https://www.datocms-assets.com/24091/1627554491-wtw-quarterly-insurtech-briefing-q2-20212.pdf.

Wired. Amazon's Snowmobile is actually a truck hauling a huge hard drive. 2016. www.wired.com/2016/12/amazons-snowmobile-actually-truck-hauling-huge-hard-drive/.

Wood, Miranda. Bank of China officially launches insurance blockchain. *Ledger Insights Enterprise Blockchain News*, 2019. www.ledgerinsights.com/bank-of-china-insurance-blockchain/.

Wuthrich, Mario V. *Machine Learning in Individual Claims Reserving*. Swiss Finance Institute Research Paper No. 16-67, SSRN, 2016. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2867897.

Wuthrich, Mario V. *Neural Networks Applied to Chain-Ladder Reserving*. SSRN, 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2966126.

Yokoi-Ara, Mamiko. *The Impact of Big Data and Artificial Intelligence (AI) in the Insurance Sector*. Organisation for Economic Co-operation and Development, 2020. www.oecd.org/finance/The-Impact-Big-Data-AI-Insurance-Sector.pdf.

Zeist. Achmea launches Canadian online insurance proposition in partnership with Fairfax. Achmea, 2018. https://news.achmea.nl/achmea-launches-canadian-online-insurance-proposition-in-partnership-with-fairfax/.