

L'utilisation de l'analytique prédictive dans l'industrie canadienne de l'assurance-vie



L'utilisation de l'analytique prédictive dans l'industrie canadienne de l'assurance-vie

AUTEURS

Jean-Yves Rioux, FICA, FSA, CERA, Deloitte
Arthur Da Silva, AICA, FSA, Deloitte
Harrison Jones, ASA, Deloitte
Hadi Saleh, Deloitte

COMMANDITAIRES

Institut canadien des actuaires
Society of Actuaries

Mise en garde et avis de non-responsabilité

Les opinions exprimées et les conclusions tirées sont celles des auteurs et elles ne représentent pas une position ou une opinion officielle de l'Institut canadien des actuaires et de la Society of Actuaries ou de leurs membres. Ces organismes ne font aucune déclaration et n'offrent aucune garantie quant à l'exactitude de l'information.

Tous droits réservés © 2019 Institut canadien des actuaires et Society of Actuaries.

TABLE DES MATIÈRES

Introduction.....	4
Utilisation du rapport	5
Sommaire	6
Identification des points forts et des faiblesses	7
Méthodologie	9
1. Sondage.....	9
1.1 Population observée	9
1.2 Processus d'échantillonnage.....	9
1.3 Conception du sondage	9
2. Recherche sur les applications de modélisation prédictive à l'extérieur de l'industrie canadienne de l'assurance-vie	10
3. Structure du rapport	10
Constatations relatives aux applications d'analytique prédictive	10
4. Constatations du sondage.....	10
4.1 Applications actuelles.....	10
4.2 Applications futures	16
5. Recherche	18
5.1 Ventes et marketing.....	18
5.2 Souscription	19
5.3 Tarification, études d'expérience et établissement de réserves.....	20
5.4 Gestion de la distribution.....	22
5.5 Gestion des polices en vigueur.....	22
5.6 Gestion des sinistres et détection de la fraude	23
5.7 Autres	24
Constatations du sondage.....	29
6.1 Données	29
6.2 Entreprise.....	39
6.3 Leadership.....	43
6.4 Cibles.....	49
6.5 Analystes.....	51
Identification des points forts et des faiblesses.....	60
Références.....	61
Remerciement des participants et des personnes interviewées	64
Annexe A : Questionnaire du sondage	65
Annexe B : Définitions.....	74

L'utilisation de l'analytique prédictive dans l'industrie canadienne de l'assurance-vie

Introduction

La disponibilité croissante des mégadonnées et le recours à l'analytique prédictive modifient le mode de fonctionnement des assureurs. L'Institut canadien des actuaires (ICA) et la Society of Actuaries (SOA) ont désigné l'analytique prédictive comme une priorité stratégique.

Bien que les assureurs IARD aient recours à l'analytique prédictive depuis bon nombre d'années et en soient les pionniers, d'autres assureurs commencent à l'utiliser dans leurs activités courantes et dans leurs pratiques de gestion des risques. Même si les techniques de modélisation prédictive sont déjà utilisées depuis quelques années, leurs applications sont de plus en plus vastes et plus novatrices. Les connaissances connexes représentent maintenant une exigence officielle pour les professionnels en actuariat qui souhaitent obtenir le titre d'associé.

La présente étude constitue un autre exemple de l'investissement de l'ICA et de la SOA dans le domaine de la modélisation prédictive. Son but consiste à examiner comment l'industrie canadienne de l'assurance-vie utilise la modélisation prédictive et à déterminer les possibilités d'amélioration. Par conséquent, elle mettra l'accent sur la prestation de renseignements sur les applications dans l'industrie canadienne de l'assurance-vie par rapport à celles d'autres industries et elle traitera brièvement des tendances susceptibles d'influer sur la façon dont les sociétés effectueront des analyses.

Aux fins de la présente étude, l'industrie canadienne de l'assurance-vie s'entend des activités des sociétés d'assurance-vie en ce qui concerne tous les produits vendus par des assureurs-vie canadiens, y compris l'assurance-vie, l'assurance contre les accidents et la maladie, l'assurance-invalidité et les rentes. Les produits d'assurances IARD sont exclus de l'étude.

Dans la présente étude, les expressions « modélisation prédictive » et « analytique prédictive » sont utilisées indifféremment et elles désignent le recours aux statistiques pour prédire des résultats.

Utilisation du rapport

La présente étude a pour but de sonder et de comparer l'utilisation de la modélisation prédictive dans l'industrie canadienne de l'assurance-vie. Les services de Deloitte ont été retenus par l'ICA et la SOA pour réaliser le sondage et préparer le rapport. Deloitte n'est pas responsable des conséquences de l'utilisation du présent rapport à d'autres fins.

Nul élément de ce rapport ne peut être utilisé ou présenté sans y faire référence.

À noter au sujet de ce rapport :

- Quinze entités ont été sélectionnées pour participer au sondage. Cet échantillon comprend une combinaison d'assureurs, de réassureurs et de filiales d'assurances d'une banque qui représentaient une grande partie des assureurs de personnes et des réassureurs.
- Le préparateur a supposé que les réponses fournies dans le cadre du sondage étaient exactes. Les résultats présentés dans cette étude reposent sur ce que les participants ont bien voulu partager et, dans certains cas, sur les estimations des participants si ces derniers ne connaissaient pas la réponse exacte à une question.
- Pour les quelques questions auxquelles certains participants n'ont pas répondu, les statistiques ont été rajustées afin que les pourcentages soient calculés en fonction du plus petit nombre de participants au sondage.
- Les statistiques sommaires, comme les moyennes, sont fondées sur le nombre de participants plutôt qu'une mesure du passif ou des primes. Les pourcentages et les cotes d'évaluation ont été calculés de cette façon, sous forme de moyenne.

Sommaire

Cette recherche a été préparée en sondant les pratiques actuelles de l'industrie canadienne de l'assurance-vie (par entrevue auprès de 15 assureurs, réassureurs et filiales d'assurances appartenant à une banque) et en complétant ces constatations par des recherches sur les applications de la modélisation prédictive utilisées à l'extérieur de l'industrie canadienne de l'assurance-vie (les assurances IARD canadiennes et les assurances étrangères, de même que des industries autres que les assurances).

Nous avons mis au point un cadre pour classer et regrouper les applications de nature semblable couvrant différents aspects opérationnels ou domaines fonctionnels des assureurs, ce qui a permis de clarifier les discussions et les rapports sur les applications.

Dans le cadre de notre sondage, nous avons demandé aux participants de déterminer la valeur de leurs initiatives d'analytique, et les efforts déployés pour les mettre en œuvre. Les répondants ont indiqué que les applications qui ont la plus grande valeur se rapportent à la souscription simplifiée, à la détection de la fraude, au marketing ciblé et à la gestion de la conservation des polices en vigueur. Le deuxième niveau d'applications jugées de valeur modeste comprenait la segmentation des nouveaux clients, le triage des applications, la souscription accélérée et automatisée, les ventes croisées et les ventes en amont, la tarification, la segmentation des polices en vigueur, les études d'expérience fondées sur la déchéance, les gains d'efficacité opérationnelle et l'optimisation de la gestion des sinistres.

La recherche a permis de cerner les applications à l'extérieur de l'industrie canadienne de l'assurance-vie qui pourraient être pertinentes pour l'industrie, notamment l'optimisation du marketing, la conception des produits, la souscription cognitive, le traitement automatisé des données, la granularité accrue de la tarification, le traitement opérationnel, les stratégies de distribution (y compris le recrutement), les ventes par un robot-conseiller, l'amélioration de l'expérience client et l'analytique de l'effectif.

En outre, nous avons relevé de vastes tendances influant sur les applications de l'analytique et les processus connexes, et nous avons identifié des éléments que les assureurs devraient surveiller, par exemple :

- Le raffermissement des exigences relatives à la protection des renseignements personnels et à la transparence/au dilemme en matière d'éthique;
- Le flou entre l'automatisation et la modélisation;
- L'utilisation accrue des aspects cognitifs (p. ex., l'interprétation d'images, la détection des sentiments vocaux);
- Les assistants virtuels;
- L'accent sur l'expérience client.

D'après les réponses au sondage, des améliorations peuvent être apportées aux données utilisées pour l'analytique prédictive en ce qui concerne la centralisation, la qualité des données et les procédures d'accès aux données. En outre, les participants n'ont pas encore sensiblement mis à profit les nouvelles technologies (aux fins de l'augmentation des données), et c'est un point que les assureurs commencent à exploiter.

Nous constatons, d'après les réponses, que des progrès sont réalisés en ce qui concerne les politiques de gouvernance et de normalisation. Il semble exister un écart entre les participants de petite et moyenne tailles et ceux de grande taille en ce qui concerne la gouvernance des données (plus particulièrement la mise à jour des données, leur exactitude et les normes). Il est probable que le même écart s'applique à ces groupes de taille dans le reste de l'industrie. De même, l'écart doit être comblé au chapitre de la gouvernance de la modélisation (notamment la validation du modèle, l'examen des codes et le contrôle des versions). Seulement cinq répondants sur 15 ont indiqué qu'ils avaient des politiques de mise à niveau des logiciels, ce qui est sensiblement inférieur à ce qui est peut-être justifié.

Même si 78 % des décisions finales relatives à l'analytique prédictive sont prises par des membres de la série des C ou plus, seulement le cinquième des membres du conseil d'administration, des chefs de la direction et des premiers vice-présidents participent aux décisions relatives à l'analytique prédictive (cela nous semble faible et nous nous attendons à une participation accrue à l'analytique prédictive qui est utilisée pour la prise de décisions).

Lorsque l'analytique n'a pas la priorité absolue, les principales raisons invoquées par les participants comprennent la qualité des données, les problèmes d'agrégation des données et d'autres priorités opérationnelles concurrentes.

Il n'est donc pas étonnant de constater que les participants de grande taille exécutent en moyenne plus de projets et qu'ils en prévoient davantage que les participants de moyenne ou de petite taille.

Comme prévu, les répondants de grande taille accordent beaucoup plus d'importance à l'affectation d'équivalent à temps plein (ETP) et ils ont tendance à effectuer de l'analytique pendant une période beaucoup plus longue que les répondants des autres groupes de taille.

La composition des équipes d'analytique est telle que, dans l'ensemble, près de 50 % des ressources en analytique prédictive sont des « experts opérationnels » plutôt que des statisticiens/architectes des données. Parmi ces experts opérationnels, 75 % sont des actuaires qui, par conséquent, participent de façon importante à l'analytique.

En ce qui a trait aux capacités, les répondants de grande et de moyenne tailles ont évalué leurs propres capacités en matière de connaissances opérationnelles comme étant inférieures aux capacités technologiques et/ou aux connaissances en statistiques/analytique. Cela peut indiquer qu'il faut investir dans la formation des techniciens au sujet des opérations ou dans le recrutement des personnes possédant des capacités d'analytique. Les répondants de petite taille estiment que leurs capacités technologiques sont inférieures à leurs connaissances opérationnelles, ce qui indique la nécessité d'investir dans la technologie et la formation connexe.

Identification des points forts et des faiblesses

La plupart des participants au sondage ne disposent pas d'un dépôt de données centralisé (que ce soit entièrement ou partiellement centralisé), et de nombreux participants, plus particulièrement ceux de moyenne et de petite tailles, ne considèrent pas que leurs données sont particulièrement complètes ou exactes. Il est important de compter sur un vaste dépôt de données claires, complètes et exactes pour profiter du pouvoir de l'analytique, en particulier dans tous les domaines fonctionnels (p. ex., le marketing, le risque, l'évaluation).

La plupart des répondants ont envisagé ou prévoient d'envisager l'utilisation de l'analytique prédictive dans le cadre d'études d'expérience. Ce processus est riche en données et les systèmes de données sont déjà en place, de sorte que le processus est un candidat naturel pour les travaux d'analytique.

La plupart des répondants, particulièrement ceux qui ne relèvent pas de grandes organisations, déclarent de nombreuses lacunes dans les types et la portée des applications, particulièrement en ce qui concerne le marketing, la gestion de la conservation, la concordance des clients et de la distribution et même la souscription accélérée/automatisée.

La plupart des participants au sondage n'ont pas songé à utiliser l'analytique prédictive pour améliorer les opérations internes (p. ex., l'analyse de l'effectif, le recours au traitement de la langue naturelle (TLN) et à la génération de la langue naturelle (GLN) pour accélérer les processus et réduire les taux d'erreur).

Outre les répondants de grande taille et quelques répondants de moyenne taille, l'analytique n'a pas été jugée une priorité absolue. De nombreux participants au sondage n'obtiennent pas un appui solide de la part de leurs dirigeants aux fins du parrainage des initiatives d'analytique dans leur entreprise.

De nombreux répondants ont indiqué qu'ils n'ont pas songé à un contexte de développement normalisé assorti de procédures d'essai automatisées. Seulement la moitié des participants au sondage ont envisagé des techniques d'apprentissage machine ou d'apprentissage profond.

Les participants au sondage ont éprouvé de la difficulté à recruter et à maintenir en poste des experts de l'analytique qui répondent aux besoins opérationnels. En effet, la plupart de ces experts (p. ex., les statisticiens et les scientifiques des données) possèdent de solides compétences techniques, mais n'ont pas les connaissances opérationnelles pertinentes ou la capacité de communiquer des idées à des auditoires non techniques à l'échelle de l'entreprise. Les actuaires possèdent habituellement de solides compétences techniques, comprennent bien les activités et sont en mesure de communiquer efficacement à l'échelle de l'organisation. Ils devraient donc être bien placés pour approfondir leurs connaissances techniques afin de couvrir l'analytique prédictive et ses applications.

Méthodologie

Les travaux sous-jacents de ce projet ont suivi deux axes principaux :

- Un sondage sur les pratiques actuelles de l'industrie canadienne de l'assurance-vie;
- Des recherches sur les applications de la modélisation prédictive à l'extérieur de l'industrie canadienne de l'assurance-vie.

La méthodologie appliquée à chacun de ces deux axes est décrite ci-après.

1. Sondage

1.1 Population observée

Quinze entités ont été sélectionnées pour participer au sondage. Cet échantillon comprend une combinaison d'assureurs, de réassureurs et de filiales d'assurances de banques qui représentaient une grande partie des assureurs de personnes et des réassureurs. Les sociétés d'assurance de petite taille et les entreprises en démarrage insurtech n'ont pas été prises en compte dans le présent sondage.

Les participants au sondage ont été classés en trois groupes selon les données sur les primes brutes souscrites. Il en a résulté le classement des participants en trois sociétés de grande taille, sept de moyenne taille et cinq de petite taille. Le cas échéant, nous commentons toute différence dans les réponses des différents groupes.

1.2 Processus d'échantillonnage

Les participants au sondage ont été choisis en fonction de la taille globale des sociétés d'assurance de personnes et des réassureurs, d'après les provisions techniques totales pour les polices d'assurance canadiennes autres que d'assurances IARD déclarées au Bureau du surintendant des institutions financières (BSIF) et à l'Autorité des marchés financiers (AMF). Ce choix a été influencé et approuvé par les membres du Groupe chargé de la surveillance de ce projet.

1.3 Conception du sondage

Les questions du sondage ont été structurées en fonction du cadre « DELTA », popularisé par le professeur Thomas Davenport de Harvard. Ce cadre mesure la culture, l'état de préparation et la maturité d'une organisation en matière d'analytique. DELTA est un acronyme dont les lettres d'appel couvrent les éléments suivants : Données, Entreprise, Leadership, cibles (*Targets* en anglais) et Analystes. Le recours à ce cadre a fait en sorte que le sondage a porté sur chacun des éléments pertinents.

Le sondage a pris la forme d'une série d'entrevues avec les participants. Cette approche a été choisie puisque l'on ne décrit pas tous l'analytique de la même façon ou en utilisant la même terminologie, et pour donner l'occasion de demander des précisions.

Le sondage a été effectué en grande partie en mai 2018. Le questionnaire figure à l'annexe A du présent rapport.

2. Recherche sur les applications de modélisation prédictive à l'extérieur de l'industrie canadienne de l'assurance-vie

La deuxième composante importante de l'étude consistait à fournir des exemples d'applications de modélisation prédictive utilisées dans d'autres industries, y compris les assurances IARD, et dans des sociétés d'assurance-vie étrangères. Pour ce faire, nous avons examiné la documentation existante et nous avons interviewé des experts en la matière. Nous discuterons également des tendances émergentes cernées.

3. Structure du rapport

Le présent rapport a pour but d'examiner comment l'industrie canadienne de l'assurance-vie utilise la modélisation prédictive et de déterminer les possibilités d'amélioration. Il est donc structuré en deux grandes sections :

- Les constatations relatives aux applications de modélisation prédictive (basé sur le sondage et la recherche);
- Les constatations du sondage.

Constatations relatives aux applications d'analytique prédictive

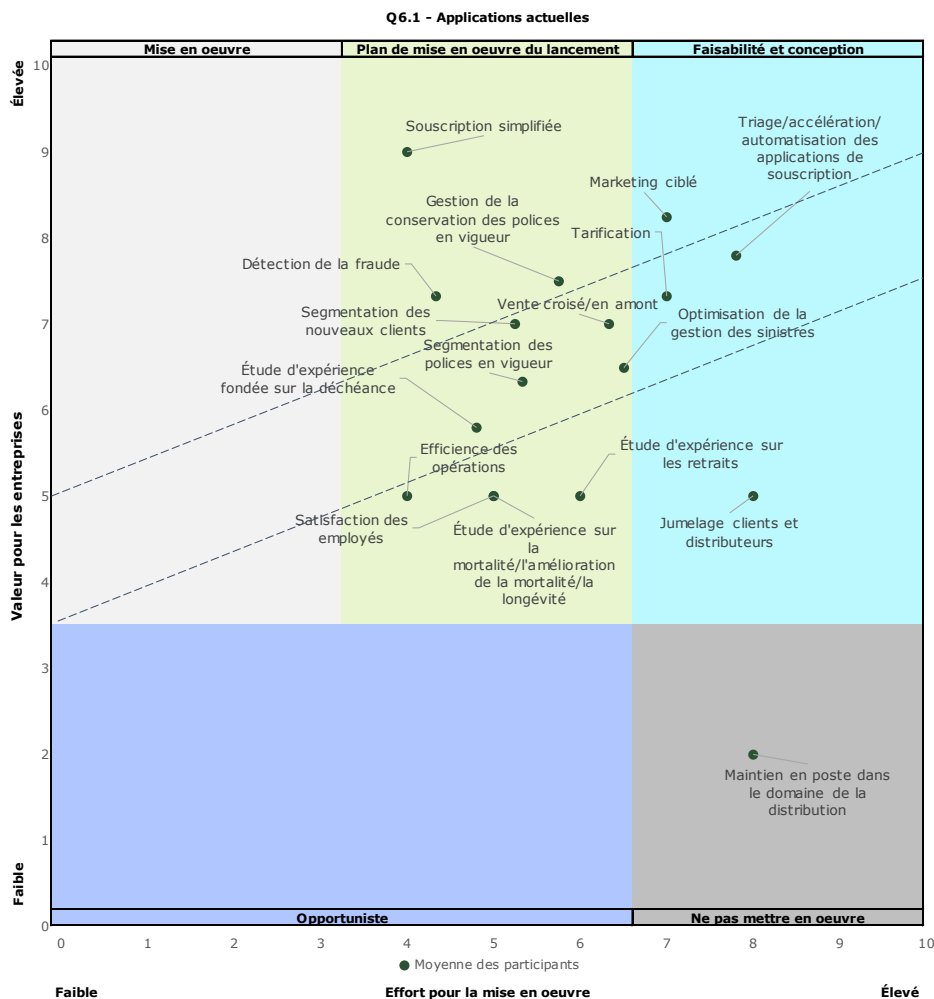
Les participants ont évalué la valeur de leurs initiatives d'analytique ainsi que les efforts déployés pour les mettre en œuvre, et ils ont conclu que bon nombre de ces initiatives présentaient une grande valeur par rapport aux efforts requis pour les mettre en œuvre.

Dans la présente section, nous examinons comment les participants au sondage utilisent et utiliseront la modélisation prédictive à l'aide des constatations du sondage. Nous étudions également d'éventuels points à améliorer à la suite d'autres recherches sur les applications émergentes des assureurs IARD et d'autres secteurs.

4. Constatations du sondage

4.1 Applications actuelles

Le tableau qui suit illustre les applications actuelles indiquées par les participants sur une grille d'effort/valeur. Le pointage pour l'effort est illustré sur l'axe des x, de sorte que les applications qui figurent dans la partie gauche de la grille sont plus faciles à mettre en œuvre. La cote de valeur est illustrée sur l'axe des y, de sorte que les applications figurant dans la partie supérieure de la grille sont plus utiles pour l'exécution. Les cotes d'effort et de valeur sont fondées sur le jugement des répondants.



Afin de classer et d'établir l'ordre de priorité des applications, nous avons attribué une cote de valeur « V » combinée à la cote d'effort « E » selon la moyenne des évaluations des répondants. En postulant que la valeur était plus importante que l'effort et que des applications de grande valeur seraient traitées malgré l'effort requis pour la mise en œuvre, nous avons accordé deux fois plus de poids à la cote de valeur qu'à la cote d'effort. Cela nous a permis d'élaborer un cadre de priorisation fondé sur la note relative « 2V-E ». Il s'agit de deux lignes qui séparent les applications et qui les attribuent aux groupes de priorité élevée (au-dessus de la ligne supérieure), moyenne (entre les lignes) ou faible (sous la ligne inférieure). Ils sont énoncés ci-dessous :

Applications à priorité élevée :

Note	Application	Description
2V-E=14 V=9 E=4	Souscription simplifiée	Utiliser des modèles prédictifs pour simplifier le processus d'application de la souscription (la souscription minimale est appliquée).
2V-E=10,3 V=7,3 E=4,3	Détection de la fraude	Élaborer des modèles prédictifs pour déterminer les sources potentielles de fraude dans les activités de certains partenaires de distribution, portefeuilles de produits, ou opérations ou politiques particulières.
2V-E=9,5 V=8,3 E=7	Marketing ciblé	Renforcer les efforts internes existants pour déterminer les occasions idéales de nouvelles ventes d'assurance.
2V-E=9,3 V=7,5 E=5,8	Gestion de la conservation des polices en vigueur	Distinguer la propension au rachat pour chaque contrat et déterminer le meilleur suivi d'offre.

Applications à priorité moyenne :

Note	Application	Description
2V-E=8,8 V=7 E=5,3	Segmentation des nouveaux clients	Établir le profil de la clientèle potentielle et déterminer les créneaux préférés, la capacité de vente croisée, l'affinité du produit, etc., souvent en utilisant des données de tiers.
2V-E=7,8 V=7,8 E=7,8	Triage/accélération/automatisation des applications de souscription	Solution conçue pour exécuter la totalité ou une partie des fonctions de présélection relevant habituellement des souscripteurs et qui vise ainsi à réduire la main-d'œuvre, le temps et les données nécessaires pour déposer une proposition d'assurance.
2V-E=7,7 V=7 E=6,3	Vente croisée/vente en amont	Identifier les clients optimaux de polices en vigueur pour des offres ciblées de protection d'assurance-vie supplémentaire ou de produits additionnels.
2V-E=7,7 V=7,3 E=7	Tarification	Utiliser des modèles prédictifs pour découvrir et utiliser de nouvelles variables de tarification.
2V-E=7,3 V=6,3 E=5,3	Segmentation des polices en vigueur	Établir le profil de la clientèle des polices en vigueur et déterminer les créneaux préférés, la capacité d'effectuer des ventes croisées, de préciser l'affinité des produits, etc., souvent en utilisant des données de tiers.
2V-E=6,8 V=5,8 E=4,8	Étude d'expérience fondée sur la déchéance	Utiliser des modèles prédictifs pour comprendre et modéliser les comportements en matière de déchéance.
2V-E=6,5 V=6,5 E=6,5	Optimisation de la gestion des sinistres	Optimiser les ressources et les efforts d'intervention en gestion de cas, habituellement pour les demandes de prestations d'invalidité.

Applications à priorité faible :

Note	Application	Description
2V-E=6 V=5 E=4	Efficiencce des opérations	Rationaliser les processus d'assurance (nouvelles polices, souscription, actuariat) afin de réduire ou d'éliminer les procédures inefficaces.
2V-E=5 V=5 E=5	Satisfaction des employés	Utiliser des modèles prédictifs pour comprendre la satisfaction des employés.
2V-E=5 V=5 E=5	Étude d'expérience sur la mortalité/l'amélioration de la mortalité/la longévité	Utiliser des modèles prédictifs pour comprendre et modéliser les résultats de mortalité et d'amélioration de la mortalité.
2V-E=4 V=5 E=6	Étude d'expérience sur les retraits	Utiliser des modèles prédictifs pour comprendre et modéliser les comportements par rapport aux retraits.
2V-E=2 V=5 E=8	Jumelage clients et distributeurs	Utiliser des modèles prédictifs pour établir des règles fondées sur un système afin de jumeler des partenaires de distribution compatibles avec des clients nouveaux ou clients sans agents ou avec des produits à vendre.
2V-E=-4 V=2 E=8	Maintien en poste dans le domaine de la distribution	Cerner les possibilités d'améliorer le maintien en poste des agents.

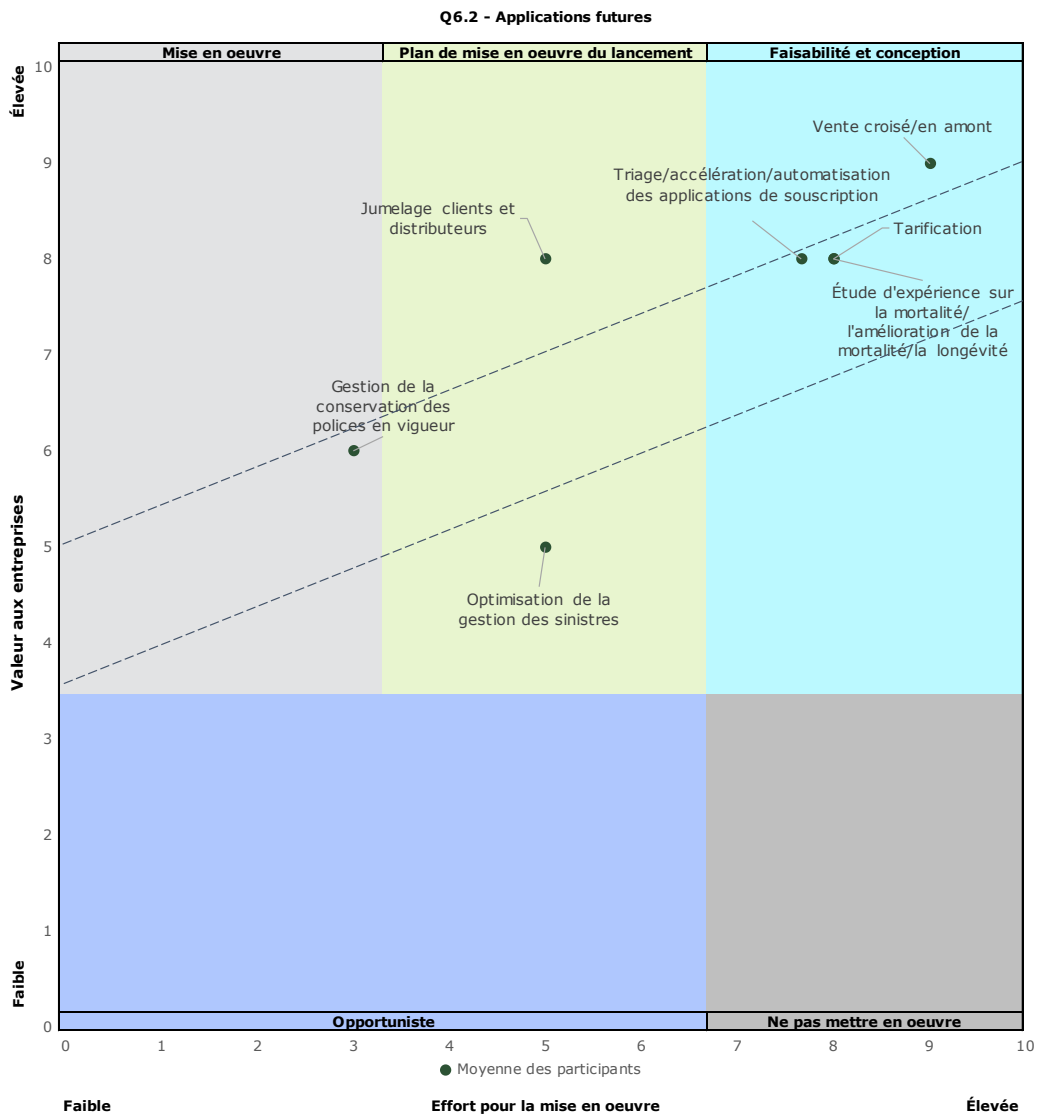
Nombre de fois où les participants ont mentionné qu'une application est actuelle :

Priorité	Application	Nombre
Élevée	Souscription simplifiée	1
	Détection de la fraude	6
	Marketing ciblé	4
	Gestion de la conservation des polices en vigueur	4
Moyenne	Segmentation des nouveaux clients	4
	Triage/accélération/automatisation des applications de souscription	5
	Vente croisée/vente en amont	6
	Tarifification	3
	Segmentation des polices en vigueur	3
	Études d'expérience fondée sur la déchéance	5
	Optimisation de la gestion des sinistres	2
Faible	Efficiency des opérations	3
	Satisfaction des employés	1
	Étude d'expérience sur la mortalité/l'amélioration de la mortalité/longévité	7
	Étude d'expérience sur les retraits	1
	Jumelage clients et distributeurs	1
	Conservation dans le domaine de la distribution	1

Il est peut-être étonnant de constater que la gestion des sinistres ait reçu une priorité moyenne et qu'elle ait été déclarée par seulement deux répondants, car d'après notre expérience, il s'agit d'une application d'analytique très populaire. Nous soupçonnons que d'autres initiatives sont en cours dans ce domaine et que le personnel avec lequel nous avons communiqué n'était pas au courant de ces initiatives, point de vue que partagent les membres du Groupe chargé de la surveillance du projet.

4.2 Applications futures

Le graphique qui suit illustre les applications futures mentionnées par les participants dans une grille d'effort/de valeur.



À l'aide de la méthode déjà décrite, nous avons déterminé les applications suivantes par priorité.

Applications à priorité élevée :

Note	Champ d'application	Description
2V-E=11 V=8 E=5	Jumelage clients et distributeurs	Utiliser des modèles prédictifs pour établir des règles fondées sur un système afin de jumeler des partenaires de distribution compatibles avec des clients nouveaux ou clients sans agents ou avec des produits à vendre.
2V-E=9 V=9 E=9	Vente croisée/vente en amont	Identifier les clients optimaux de polices en vigueur pour des offres ciblées de couverture d'assurance-vie supplémentaire ou de produits additionnels.
2V-E=9 V=6 E=3	Gestion de la conservation des polices en vigueur	Distinguer la propension au rachat pour des contrats individuels et déterminer le meilleur suivi d'offre.

Applications à priorité moyenne :

Note	Application	Description
2V-E=8,3 V=8 E=7,7	Triage/accélération/automatisation des applications de souscription	Solution conçue pour exécuter la totalité ou une partie des fonctions de présélection habituellement effectuées par les souscripteurs et qui vise ainsi à réduire la main-d'œuvre, le temps et les données nécessaires pour déposer une proposition d'assurance.
2V-E=8 V=8 E=8	Étude d'expérience sur la mortalité/l'amélioration de la mortalité/la longévité	Utiliser des modèles prédictifs pour comprendre et modéliser les résultats de mortalité et d'amélioration de la mortalité.
2V-E=8 V=8 E=8	Tarifification	Utiliser des modèles prédictifs pour découvrir et utiliser de nouvelles variables de tarification.

Applications à priorité faible :

Note	Application	Description
2V-E=5 V=5 E=5	Optimisation de la gestion des sinistres	Optimiser les ressources et les efforts d'intervention en gestion de cas, habituellement pour les demandes de prestations d'invalidité.

Nombre de fois où les participants ont mentionné qu'une application est future :

Priorité	Application	Nombre
Élevée	Jumelage clients et distributeurs	1
	Vente croisée/vente en amont	1
	Gestion de la conservation des polices en vigueur	1
Moyenne	Triage/accélération/automatisation des applications de souscription	3
	Étude d'expérience sur la mortalité/l'amélioration de la mortalité/la longévité	1
	Tarifification	1
Faible	Optimisation de la gestion des sinistres	1

5. Recherche

Étant donné que l'identification des applications au moyen de l'analytique constituait l'un des principaux objectifs de cette mission, nous avons complété le résumé des résultats du sondage par des recherches sur les applications utilisées à l'étranger, à l'extérieur de l'industrie de l'assurance-vie et par des entrevues avec des experts de l'analytique. Afin d'expliquer les diverses applications trouvées dans le cadre de notre examen de la documentation et de nos entrevues, nous avons jugé utile d'élaborer un vaste cadre de catégories pour que chaque application puisse être associée à une catégorie plus large. Dans la pratique, ces applications peuvent recouper plusieurs catégories. Les applications qui ne sont pas couvertes par l'assurance-vie canadienne, actuellement ou dans un proche avenir, sont détaillées au besoin. Ces constatations sont décrites ci-dessous.

5.1 Ventes et marketing

L'utilisation de modèles prédictifs pour les ventes et le marketing ne se limite pas à l'industrie de l'assurance. Il est courant d'utiliser la modélisation pour comprendre et prédire le comportement des clients (en utilisant des techniques de regroupement), évaluer et prioriser les occasions d'affaires, recommander des produits et services personnalisés, cibler les bons clients et éclairer les stratégies globales de marketing.

La gamme des applications identifiées dans la recherche pour cette catégorie comprend :

- Segmentation/profil des clients – identification des caractéristiques individuelles corrélées avec la décision d'achat (comme l'affinité du produit, la propension à acheter, la propension à renouveler, les profils de santé et les notes reliées au mode de vie);
- Vente croisée de produits supplémentaires ou vente en amont de protection additionnelle au moment de la vente à un nouveau client;
- Marketing ciblé/production de pistes et priorisation;
- Valeur à vie du client (VVC);
- Efficacité de la campagne de marketing, optimisation des dépenses de marketing, élaboration de stratégies de marketing;
- Élaboration de produits – élaboration de nouveaux produits qui tirent profit d'une grande quantité de données (par exemple, l'assurance fondée sur l'utilisation et la télématique), prévision de la demande de produits de la clientèle fondée sur des données internes et externes;
- Recommandations de produits adaptés.

Parmi les éléments qui précèdent, les participants au sondage n'ont pas utilisé les applications qui suivent de façon prédictive, au moment présent et dans un proche avenir :

Valeur à vie du client

Un seul des participants au sondage a indiqué avoir utilisé des applications de VVC. La VVC est une mesure qui calcule la valeur qu'un client apporte à une entreprise, soit par une augmentation des revenus, soit par une diminution des coûts opérationnels. L'une des principales distinctions entre la VVC et d'autres mesures réside dans la prise en compte de la valeur potentielle d'un client pour l'entreprise tout au long de sa vie. En utilisant la VVC, un assureur peut structurer ses décisions d'affaires de manière à cibler des clients très rentables dans de nombreuses branches d'assurance.

Efficacité et optimisation du marketing

De nombreuses mesures peuvent être prises pour accroître l'efficacité des campagnes de marketing, des messages et des dépenses. La rétroaction des clients fondée sur les appels de soutien et les messages dans les médias sociaux peut être utilisée pour élaborer de nouvelles campagnes de marketing au moyen de l'analytique de textes et du TLN. Ces données peuvent servir à étudier l'attrait du produit, à comprendre la satisfaction de la clientèle et à modifier les campagnes de promotion des produits.

Élaboration de produits

L'analytique prédictive peut servir à analyser les données internes (notamment les interactions numériques avec les clients, les appels vocaux, les courriels et les données de vente existantes) et externes (par exemple, les médias sociaux, les données démographiques) afin de prédire les produits et caractéristiques que le marché valorisera et souhaitera dans l'avenir. Voici un exemple connexe : Netflix élabore des critères de réussite pour son contenu d'origine en se fondant sur les données d'utilisation, les métadonnées et les données de l'écosystème, comme les tendances dans les nouvelles.

Recommandations de produits adaptés

Les recommandations de produits adaptés consistent à cibler les bons clients au bon moment avec le bon contenu. Ce processus met à contribution à la fois les données internes (comme les données sur les achats, les données sur les interactions avec les clients et les données sur les interactions numériques) et les données externes (comme les données sur le mode de vie, les comportements d'achat, les médias sociaux) pour comprendre les comportements d'achat afin de faire correspondre les recommandations de produits et de services aux besoins prévus des clients.

5.2 Souscription

Le cadre répétitif et fondé sur des règles de la plupart des polices de souscription est une option de la modélisation prédictive. L'entreprise peut créer un algorithme qui fournit une prévision à utiliser pour éliminer les tests inutiles fondés sur les réponses fournies par le client, accélérer le processus d'émission et relever les limites d'émission sans devoir effectuer des tests. L'analytique peut également servir à améliorer l'information au moyen de données de tiers, comme la base de données du Medical Information Bureau (Bureau de renseignements médicaux) pour l'indication des fumeurs. Si le modèle prédictif est bien construit, la rapidité et l'exactitude accrues sont perçues comme présentant un rendement élevé. Puisque la loi canadienne sur la discrimination est en vigueur, il existe d'éventuels problèmes au chapitre des restrictions réglementaires (par rapport aux données utilisées aux fins de la tarification et de la souscription) et du risque d'atteinte à la réputation (lorsque le public résiste aux pratiques de souscription).

La gamme des applications identifiées dans la recherche pour cette catégorie comprend :

- Segmentation améliorée des risques – élargissement des critères pour la sélection d’une catégorie de souscription (c.-à-d. un programme d’escompte), découverte de nouveaux prédicteurs en matière de santé à l’aide de sources de données nouvelles et de rechange (p. ex. tests génétiques 23andMe);
- Souscription simplifiée/rationalisation du processus d’application;
- Triage des applications et priorisation des ressources de souscription;
- Souscription accélérée – réduction des exigences intrusives de souscription, décisions automatisées sur les exigences de souscription pour chaque demandeur, modèles de propension des fumeurs (prévision du statut de tabagisme d’un demandeur sans vérification des fluides), traitement direct;
- Optimisation de la souscription – hausser les limites d’émission sans tests;
- Priorisation des ressources de souscription;
- Automatisation cognitive pour reproduire les décisions de souscription;
- Traitement automatisé des données sur la souscription (p. ex., à l’aide de la vision par ordinateur/TLN).

Parmi les éléments qui précèdent, les participants au sondage n’utilisent pas ou ne prévoient pas d’utiliser les applications qui suivent dans un proche avenir :

Automatisation cognitive pour reproduire les décisions de souscription

L’intelligence artificielle (IA) peut être structurée pour détecter les tendances dans la prise de décisions humaines de manière à pouvoir suggérer ou prendre des décisions que les humains devraient normalement prendre dans des scénarios plus simples. Une telle technologie pourrait être utilisée pour réduire la prise de décisions humaines dans ces scénarios afin de permettre aux gens de se concentrer sur des décisions plus complexes faisant appel à un jugement humain plus compliqué. Elle pourrait également servir de complément à un tel processus.

Automatisation du traitement des données de souscription

L’analyse du texte, le TLN et l’analyse de l’image peuvent alléger le fardeau que représente l’extraction d’information à partir de données de souscription non structurées (p. ex., demandes sur papier, rapport de médecin traitant) et son exportation vers un format structuré pour accélérer le processus de souscription.

5.3 Tarification, études d’expérience et établissement de réserves

Tarification – Les revenus futurs d’une société d’assurance dépendent directement de la tarification des produits et des volumes d’affaires, de sorte qu’il est de la plus haute importance de fixer le prix qui convient au produit. Une tarification plus granulaire et plus précise peut améliorer la stabilité financière de l’entreprise tout en maintenant des taux concurrentiels avec l’industrie. Les restrictions réglementaires constituent un obstacle majeur à l’utilisation généralisée de ces modèles prédictifs. Naturellement, les organismes de réglementation ne sont pas à l’aise d’appliquer un modèle inexplicable à la tarification des consommateurs.

Les applications de tarification potentielles dans cette catégorie comprennent :

- L’exactitude et/ou la granularité accrues des prix grâce à l’identification de nouvelles variables de tarification et à l’identification de cotes externes;
- La réduction du délai d’exécution du processus de tarification au moyen de l’exploration textuelle/TLN et de la GLN.

Parmi les éléments qui précèdent, les participants au sondage n’utilisaient pas et ne prévoyaient pas d’utiliser les applications qui suivent dans un proche avenir :

Exactitude et granularité accrues des prix

L'intégration de variables supplémentaires, y compris une gamme de variables non traditionnelles, peut aider à accroître la granularité des prix, dans le but d'améliorer l'exactitude des prix. Les variables non traditionnelles peuvent comprendre des renseignements sur le mode de vie (p. ex. passe-temps, comportement en matière de magasinage en ligne) provenant de fournisseurs tiers externes, une cote de crédit (selon l'information sur le mode de paiement, l'historique du compte, les faillites/privileges, les recouvrements) ou des approximations de la cote de crédit (p. ex., historique des paiements des services de télécommunications). En outre, l'exactitude des prix peut être améliorée en introduisant des interactions variables pour tenir compte des corrélations entre les variables.

Réduction du délai d'exécution du processus de tarification au moyen de l'exploration textuelle/du TLN et de la GLN

Le délai d'exécution pour la tarification peut être réduit à l'aide de technologies modernes, comme l'exploration de textes et le TLN, afin de récupérer l'information clé de documents textuels non structurés (comme l'information sur les produits et les documents sur la réassurance facultative). La GLN peut servir à automatiser de façon intelligente la production de documents sur la tarification, notamment des notes de service et des courriels d'approbation.

Études d'expérience et constitution de réserves – Le processus d'étude d'expérience est riche en données et est exécuté depuis de nombreuses années, de sorte que l'expertise est en place. Il est donc compréhensible qu'il y ait eu une vague récente d'utilisation des modèles prédictifs dans ce domaine pour déterminer s'ils peuvent fournir des prévisions plus exactes que les méthodes traditionnelles. Dans les cas où le modèle prédictif est optimal, de nouvelles hypothèses peuvent être appliquées aux réserves selon un modèle prédictif.

Les applications possibles dans cette catégorie pour les études d'expérience et la constitution de réserves comprennent :

- La détermination de facteurs d'expérience additionnels (p. ex., des données supplémentaires sur la souscription, des données sur les tiers externes) ou de variables corrélées aux facteurs existants des événements liés aux sinistres (p. ex., cotes de risque liées à la consommation de médicaments d'ordonnance, notes sur le mode de vie);
- L'amélioration des estimations relatives aux facteurs d'expérience existants (p. ex., la mortalité, le comportement des souscripteurs);
- Des estimations granulaires des réserves (p. ex. à l'échelle des membres) et des réserves économiques et estimation du capital.

Parmi les éléments qui précèdent, les participants au sondage n'utilisaient pas et ne prévoyaient pas d'utiliser les applications qui suivent dans un proche avenir :

Des estimations granulaires des réserves et du capital

Les réserves qui ont traditionnellement été évaluées à de faibles niveaux de granularité.

5.4 Gestion de la distribution

L'analytique prédictive peut être utilisée pour optimiser l'effectif des ventes en comprenant mieux les comportements et les facteurs de réussite.

Les applications possibles dans cette catégorie comprennent :

- Surveillance du rendement, gestion et maintien en poste des conseillers;
- Recrutement et sélection de nouveaux agents;
- Jumelage/optimisation des clients en matière de distribution;
- Stratégie de distribution;
- Ventes par un robot-conseiller.

On a fait référence à l'analytique de la distribution qui semblait générique; il est donc possible que certaines des applications les plus précises mentionnées ici aient été mises en œuvre par les participants au sondage, mais puisqu'elles ne sont pas très largement utilisées, elles sont décrites ici. Parmi les éléments qui précèdent, les participants au sondage n'utilisaient pas et ne prévoyaient pas d'utiliser les applications qui suivent dans un proche avenir :

Recrutement et sélection de nouveaux agents

L'analytique prédictive peut servir à cerner les traits de caractère des conseillers très productifs, puis à recruter de nouveaux conseillers.

Stratégie de distribution

Les assureurs peuvent mettre à contribution des données supplémentaires, internes et externes, pour approfondir leurs connaissances de la clientèle et ouvrir de nouveaux modes de distribution (p. ex., WeChat en Chine).

Robot-conseiller

Une autre méthode de distribution pour remplacer les conseillers en ressources humaines est connue sous l'appellation « robot-conseiller ». Ces conseillers peuvent interagir avec les clients dans un langage naturel et formuler des recommandations fondées sur les besoins des clients dans des circonstances plus simples.

5.5 Gestion des polices en vigueur

L'analytique prédictive peut servir à améliorer l'expérience des clients, à identifier les clients les plus intéressants et à élaborer des initiatives en lien avec eux et à optimiser les ventes ciblées vers ces personnes.

Les applications possibles dans cette catégorie comprennent :

- Segmentation de la clientèle des polices en vigueur – différencier les polices en vigueur d'après les notes attribuées à la santé/au risque plutôt que d'après la souscription, le mode de vie et l'affinité des produits;
- Gestion du maintien en poste – concevoir des stratégies de maintien en poste (p. ex., conversion ciblée, programmes de loyauté ciblés, meilleure offre) pour viser des clients de grande valeur avec propension au rachat;
- Gestion des polices en vigueur après déchéance;
- Modification des polices non garanties (ajustables);
- Vente croisée/offre de produits ciblés, vente en amont de protection supplémentaire (élément discuté à la section Ventes et marketing);
- Amélioration de l'expérience client.

Parmi les éléments qui précèdent, les participants au sondage n'ont pas utilisé les applications qui suivent de façon prédictive, au moment présent et dans un proche avenir :

Gestion des polices en vigueur après déchéance

Pour les clients qui ont déjà été frappés de déchéance, les données sur la raison de la déchéance peuvent être utilisées pour prévoir/offrir d'autres produits ou garanties qui pourraient intéresser le client perdu.

Modification des polices non garanties (ajustables)

Des techniques de modélisation prédictive pourraient être utilisées pour les polices à caractéristiques ajustables (comme les primes ajustables et les taux de crédit) à l'aide de sources de données supplémentaires.

Amélioration de l'expérience client

Fournir des services adaptés et des renseignements pertinents aux clients éclairés par des modèles prédictifs.

5.6 Gestion des sinistres et détection de la fraude

Gestion des sinistres – Le processus de gestion des sinistres peut être très actif et une partie de ce processus peut être répétitive. Cela signifie que l'automatisation peut prendre la forme d'un modèle prédictif. Ces modèles ont souvent pour but de traiter et de gérer les sinistres touchant des polices en vigueur. Même si l'élément humain ne peut être entièrement éliminé, des modèles peuvent être utilisés pour appuyer le processus et améliorer l'information disponible. Les applications modernes d'apprentissage machine peuvent améliorer la reconnaissance du contenu, établir des priorités plus intelligentes et même accroître la satisfaction de la clientèle en réduisant sensiblement le temps de réponse.

Les applications possibles dans cette catégorie comprennent :

- Le triage de la gestion des sinistres/traitement préliminaire qui déclenche différents niveaux de gestion des sinistres, de priorisation des ressources relatives aux sinistres et ce qu'il faut faire/ce qu'il faut cesser de faire avec les notes relatives aux sinistres (p. ex., la notation des sinistres en invalidité);
- Segmentation des sinistres – déterminer le type de demande de règlement et les mesures requises;
- Gains d'efficacité opérationnelle dans le traitement des sinistres.

Parmi les éléments qui précèdent, les participants au sondage n'utilisaient pas et ne prévoyaient pas d'utiliser les applications qui suivent dans un proche avenir :

Gains d'efficacité opérationnelle au titre du traitement des sinistres

Le TLN peut servir à extraire les principaux attributs (p. ex., la nature d'une blessure, la profession) des notes non structurées sur les sinistres qui peuvent être analysées afin d'améliorer les données sur le traitement des demandes et qui servent à la modélisation prédictive pour repérer de façon proactive les demandeurs à risque au titre des demandes de prestations prolongées.

De plus, le traitement automatisé des sinistres peut se faire au moyen d'assistants virtuels (p. ex., Trov, Aetna's Ann) qui interagissent en langage naturel. Par exemple, un modèle de demande de règlement recommandé peut être fourni pour la présentation d'une réclamation à partir d'une demande de renseignements du client.

Détection de la fraude – La fraude d’assurance est devenue un sujet et une application populaires de modélisation prédictive. Bien qu’il soit difficile de connaître l’ampleur de l’impact de la fraude pour chaque assureur, il ne fait aucun doute que la réduction de la fraude aura une incidence importante et positive sur les entreprises et les consommateurs. La mise en œuvre précise d’un modèle de détection de la fraude peut être très complexe et elle repose sur un ensemble solide de données sous-jacentes.

Les applications possibles dans cette catégorie comprennent :

- Fraude relative à une demande – détection au moyen de modèles de propension des fumeurs à réduire les fausses déclarations comme celles des non-fumeurs pour obtenir un rabais;
- Réponses proactives aux demandes frauduleuses;
- Fraude liée aux sinistres – demandes de prestations de décès contrefaits, demandes de prestations d’invalidité exagérées ou frauduleuses, et autres fraudes pouvant faire l’objet d’une enquête à l’aide de sources de données externes (p. ex., médias sociaux, conditions météorologiques);
- Gains d’efficacité opérationnelle au chapitre de la détection de la fraude.

Parmi les éléments qui précèdent, les participants au sondage n’utilisaient pas et ne prévoyaient pas d’utiliser les applications qui suivent dans un proche avenir :

Réponses proactives aux demandes frauduleuses

La fraude relative à une demande (p. ex., des applications fictives créées par les agents afin d’obtenir des bénéfices frauduleux) peut faire l’objet d’une enquête proactive avant le versement de prestations au moyen de la détection des tendances et des anomalies (p. ex., volume inhabituel de demandes, correspondance avec des listes de fraudes connues).

Gains d’efficacité opérationnelle au chapitre de la détection de la fraude

Des gains d’efficacité peuvent être obtenus dans le cadre d’une enquête portant sur les données structurées. Par exemple, Lemonade utilise près de 20 algorithmes d’apprentissage machine pour faciliter la détection de la fraude dans les demandes de règlement reçues. Les algorithmes reposent sur des images et des vidéos, et une réponse est fournie en quelques minutes.

5.7 Autres

La modélisation prédictive peut servir à appuyer les opérations de façon plus générale. Cela s’applique également pour les sociétés d’assurance et autres sociétés qui ne sont pas des sociétés d’assurance. La plupart de ces autres applications n’ont pas été prises en compte par les participants au sondage. Les applications englobent l’analytique de l’effectif, la gestion des risques, les placements et les échanges commerciaux, et l’amélioration des processus.

Les applications possibles dans cette catégorie comprennent :

Analytique de l'effectif

L'analytique prédictive peut être utilisée pour gérer les employés d'une entreprise à l'aide de méthodes comme :

- Maintenir des employés en poste en leur soumettant de bonnes offres;
- Prévoir les facteurs de mobilisation des employés et l'incidence des politiques en matière de ressources humaines sur le rendement;
- Déterminer où investir dans l'immobilier (bureaux), en prédisant où les talents futurs sont susceptibles de résider et le potentiel de croissance de ces talents à divers endroits;
- Améliorer le processus de recrutement en intégrant l'information sur les médias sociaux pour déterminer les meilleurs candidats et les facteurs de rendement élevé des employés.

Gestion des risques

L'analytique prédictive peut servir à la gestion des risques des sociétés dans divers contextes :

- Sécurité – dégager les facteurs déterminants des problèmes de santé et de sécurité des travailleurs et concevoir des politiques pour gérer ces risques;
- Risque lié à l'image de marque – cerner et surveiller le risque lié à l'image de marque en extrayant les tendances systématiques à partir de niveaux élevés d'interférence dans des ensembles de données non structurées comme les médias sociaux;
- Modélisation du risque de crédit – identifier les facteurs de crédit dans le risque de crédit à l'intérieur des titres de placement, à l'aide d'une variété de données de remplacement (plutôt que des cotes de crédit et des données internes sur les paiements), comme l'information sur les entreprises de télécommunications et les opérations, l'activité dans les médias sociaux, les modèles de paiement des factures autres que le crédit pour prédire la capacité de remboursement de leurs clients;
- Évaluation proactive de la qualité des données – repérer proactivement les anomalies et les signaler aux modélisateurs et(ou) aux fournisseurs de données; cerner les lacunes et les incohérences dans les données.

Placements et échanges commerciaux

L'analytique prédictive peut être utilisée pour les placements et les opérations des sociétés dans divers contextes :

- Trouver des signaux pour des rendements plus élevés (et non corrélés) aux fins de la construction de portefeuilles;
- Échange d'algorithmes à des fins de gestion de l'actif-passif ou de couverture pour réduire le nombre d'employés comme il a été constaté dans des sociétés de gestion de portefeuille (p. ex., BlackRock).

Amélioration des processus

L'analyse prédictive et l'IA ont été utilisées pour modifier les processus dans le but d'améliorer l'expérience client et d'accroître l'efficacité des opérations, qui peuvent être appliquées aux assureurs de la façon suivante :

- Utiliser des algorithmes de communication directe pour diagnostiquer plus rapidement les besoins des clients;
- Messages de rétroaction des clients qui utilisent le TLN pour mieux prédire quels vendeurs créeront de mauvaises expériences (p. ex. eBay/Airbnb);

- Amélioration de la recherche sur le site Web et de la présentation de l'information – utilisation de l'IA et de l'apprentissage machine combinés avec le TLN pour prédire l'intention d'un client ou segmenter les clients en groupes qui peuvent être mieux servis avec des renseignements personnalisés. Des assistants de magasinage virtuel peuvent être utilisés pour recommander des produits (p. ex., utilisation d'IBM Watson par North Face);
- TLN et analytique de texte pour améliorer le service à la clientèle par courriel ou assistants virtuels/agent conversationnel en tout temps qui peuvent interagir en utilisant un langage naturel. Par exemple, les départements de services peuvent utiliser le jumelage automatique des réponses aux demandes des clients en fonction des questions antérieures des clients ou des données historiques sur les services;
- Analytique du discours sur les appels des clients pour étudier les sentiments et les principales préoccupations et recommander les prochaines mesures à prendre;
- Faire en sorte que les clients comprennent mieux leurs propres risques (p. ex., MyFitnessPal), conseils proactifs sur les comportements sains (p. ex., fiche UA d'Under Armour);
- Automatisation intelligente des processus répétitifs, comme les ressources humaines (RH) et la technologie de l'information (TI), dans le but de réduire le besoin d'intervention humaine et les taux d'erreur;
- Automatisation de l'analyse des données non structurées (p. ex. image, vidéo, documents numérisés);
- Gestion et négociation de contrats – extraction et collecte de données pour déterminer des clauses et des conditions précises à l'aide de la technologie d'analytique textuelle (p. ex., JP Morgan utilise le « COIN » pour les ententes de prêt commercial);
- Prévisions – nouveaux plans d'activité, prévisions de la demande de ressources internes (personnes, processus, technologie).

En outre, nous avons relevé de vastes tendances influant sur les applications de l'analytique et les processus connexes, et nous avons identifié des éléments que les assureurs devraient surveiller, par exemple :

- Le raffermissement des exigences relatives à la protection des renseignements personnels et à la transparence/au dilemme en matière d'éthique;
- Le flou entre l'automatisation et la modélisation où les modèles se mettent à jour automatiquement;
- L'utilisation accrue des aspects cognitifs (p. ex. l'interprétation d'images, la détection des sentiments exprimés par la voix);
- Les assistants virtuels;
- L'accent sur l'expérience client.

Résumé des constatations relatives aux applications de l'analytique prédictive

Les participants au sondage ont calculé la valeur de leurs initiatives d'analytique, ainsi que les efforts déployés pour les mettre en œuvre, et ils ont constaté que nombre de ces applications ont produit une grande valeur par rapport aux efforts requis pour les mettre en œuvre.

Un cadre de classement des applications a été élaboré pour grouper les applications de nature semblable et ce cadre couvrirait les catégories suivantes :

- Vente et marketing;
- Souscription;
- Tarification, études d'expérience et constitution de réserves;
- Gestion de la distribution;
- Gestion des polices en vigueur;
- Gestion des sinistres et détection de la fraude;
- Autre.

À l'heure actuelle, les sociétés d'assurance-vie canadiennes (les participants au sondage) accordent une grande importance aux applications actuelles suivantes :

- [Souscription] simplifiée;
- [Gestion des sinistres et détection de la fraude] Détection de la fraude;
- [Ventes et marketing] Marketing ciblé;
- [Gestion des polices en vigueur] Gestion de la conservation des polices en vigueur.

De plus, ils valorisent actuellement modérément les applications actuelles suivantes :

- [Souscription] accélérée/automatisée;
- [Ventes et marketing] Vente croisée/Vente en amont;
- [Tarification, études d'expérience et constitution de réserves] Tarification;
- [Gestion des polices en vigueur] Segmentation des polices en vigueur;
- [Ventes et marketing] Segmentation des nouveaux clients;
- [Tarification, études d'expérience et constitution de réserves] Étude d'expérience de la déchéance.

Voici les principales applications qui pourraient être examinées par d'autres sociétés d'assurance-vie canadiennes :

- [Ventes et marketing] Valeur à vie du client (même si elle a été mentionnée par un seul répondant);
- [Ventes et marketing] Efficacité et optimisation du marketing;
- [Ventes et marketing] Élaboration de produits;
- [Ventes et marketing] Recommandations de produits adaptés;
- [Souscription] Automatisation cognitive pour reproduire les décisions de souscription;
- [Souscription] Traitement automatisé des données sur la souscription;
- [Tarification, études d'expérience et constitution de réserves] Meilleure exactitude et granularité de la tarification;
- [Tarification, études d'expérience et constitution de réserves] Optimisation dynamique des prix;
- [Tarification, études d'expérience et constitution de réserves] Réduction des délais d'exécution des processus à l'aide de l'exploration textuelle/TLN et GLN;
- [Tarification, études d'expérience et constitution de réserves] Estimation granulaire des réserves et du capital;
- [Gestion de la distribution] Recrutement/sélection de nouveaux agents;
- [Gestion de la distribution] Stratégie de distribution;

- *[Gestion de la distribution] Robot-conseiller;*
- *[Gestion des polices en vigueur] Gestion des polices en vigueur après déchéance;*
- *[Gestion des polices en vigueur] Modifications apportées aux polices non garanties (ajustables);*
- *[Gestion des polices en vigueur] Amélioration de l'expérience client;*
- *[Gestion des sinistres et détection de la fraude] Segmentation des sinistres;*
- *[Gestion des sinistres et détection de la fraude] Gains d'efficacité opérationnelle dans le traitement des sinistres;*
- *[Gestion des sinistres et détection de la fraude] Réponses proactives aux demandes frauduleuses;*
- *[Gestion des sinistres et détection de la fraude] Gains d'efficacité opérationnelle en matière de détection de la fraude;*
- *[Autres] Analytique de l'effectif (p. ex., maintien en poste des employés, engagement des employés, recrutement à l'aide de données de tiers);*
- *[Autres] Gestion des risques (p. ex., facteurs de sécurité des travailleurs, surveillance du risque de l'image marque, évaluation proactive des données);*
- *[Autres] Placements/Échanges commerciaux (p. ex., détermination des facteurs de rendement, échanges d'algorithmes);*
- *[Autres] Amélioration des processus (p. ex., traitement de la voix, processus de TLN).*

En outre, nous avons relevé de vastes tendances influant sur les applications de l'analytique et les processus connexes, et nous avons identifié des éléments que les assureurs devraient surveiller, par exemple :

- Le raffermissement des exigences relatives à la protection des renseignements personnels et à la transparence/dilemme en matière d'éthique;
- Le flou entre l'automatisation et la modélisation;
- L'utilisation accrue des aspects cognitifs (p. ex. interprétation de l'image, détection des sentiments vocaux);
- Les assistants virtuels;
- L'accent sur l'expérience client.

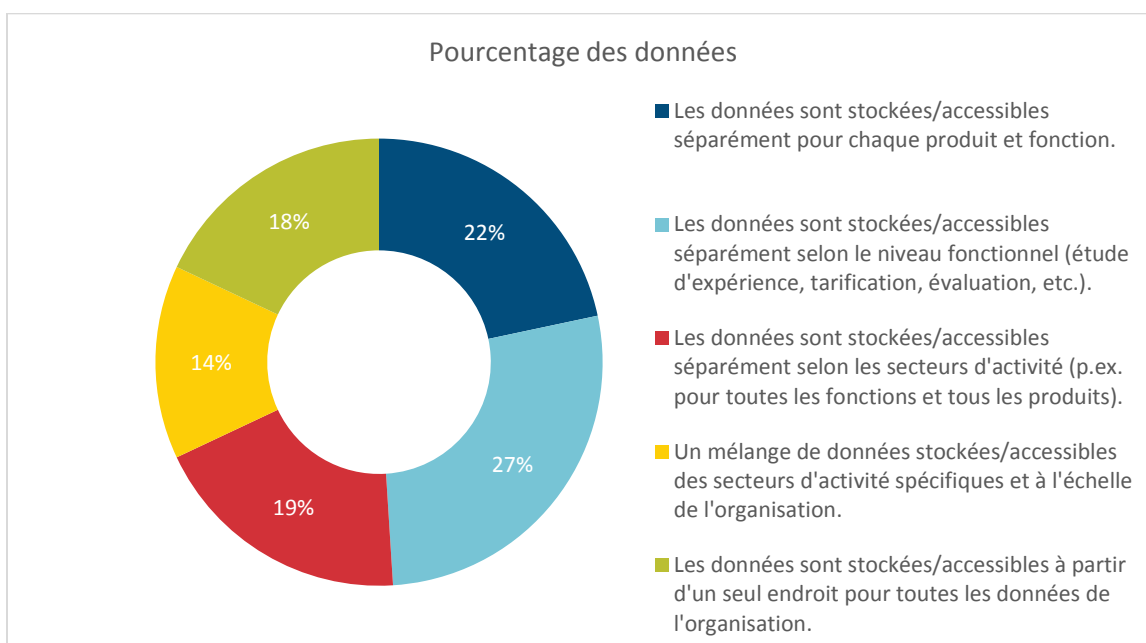
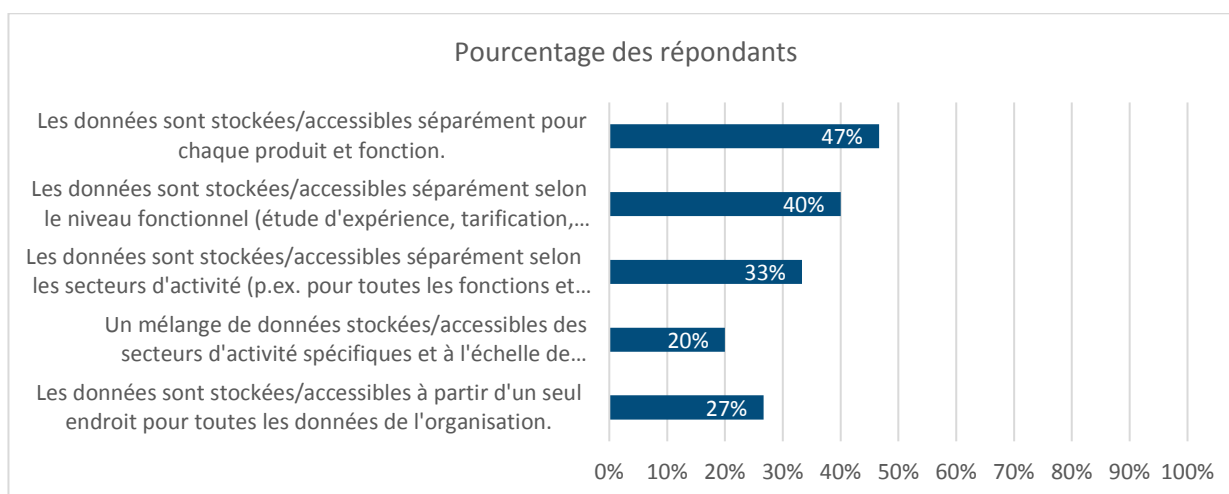
Constatations du sondage

Dans la présente section, nous décrivons les questions et les réponses du questionnaire, groupées dans les principaux thèmes.

6.1 Données

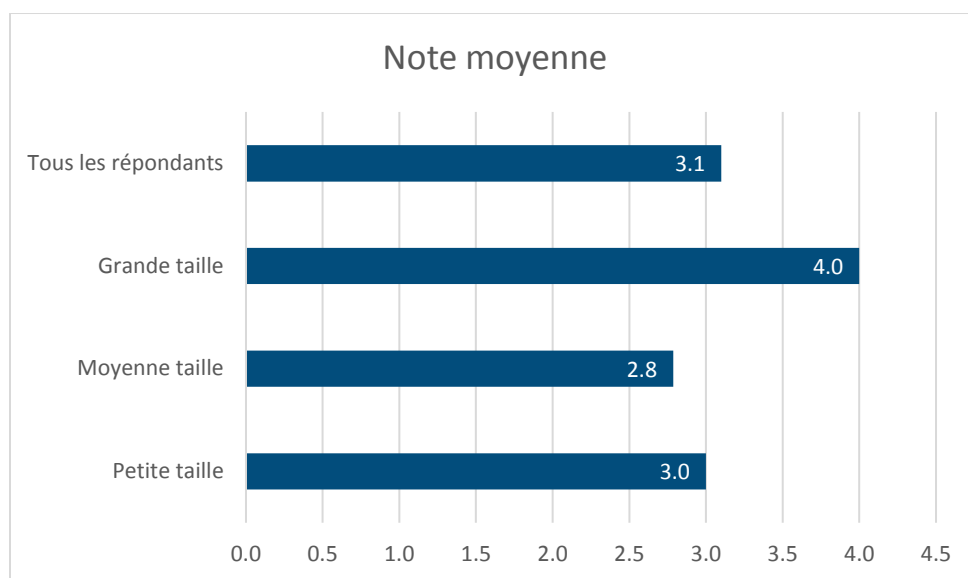
Les données constituent le point de départ de tous les travaux d'analytique d'une organisation. Les données sont-elles propres, facilement accessibles et efficaces?

Question 3.1 – Qu'est-ce qui décrit le mieux la centralisation de vos données (autoévaluées par les répondants)?



- Les participants au sondage semblent divisés sur la façon dont les données sont organisées à l'échelle de l'entreprise.
- Près de la moitié (47 %) des répondants ont des données segmentées au niveau des produits et des fonctions.
- Environ le tiers (27 %) des répondants ont indiqué avoir des données centralisées au niveau de l'organisation.
- Dans l'ensemble, 18 % de toutes les données appartenant aux répondants au sondage sont centralisées. D'après notre expérience, la centralisation des données est plus difficile à réaliser pour une entreprise de grande taille ayant de nombreux systèmes existants que pour une entreprise de petite taille.
- Au total, 68 % des données sont décentralisées à l'échelle de l'unité fonctionnelle ou à un niveau inférieur. Nous constatons également que les réassureurs disposent de données moins centralisées que les assureurs.
- À partir de notre expérience, nous avons constaté que la pratique exemplaire consiste à réduire le nombre de points d'accès et à augmenter la normalisation en centralisant les données. Il est ainsi plus facile pour les utilisateurs des données d'extraire l'information dont ils ont besoin, ce qui peut s'étendre à plusieurs secteurs d'activité ou fonctions. Toutefois, les répondants au sondage ont également indiqué qu'au moyen de diverses initiatives technologiques, acquisitions et priorités, il peut être difficile d'harmoniser toutes les données. La complexité de la structure et du stockage des données suit la croissance de l'entreprise.

Question 3.2 – Sur une échelle de 1 à 5, comment les utilisateurs évalueraient-ils l'exhaustivité et l'exactitude de vos données? [1 = plus bas/5 = plus élevé]

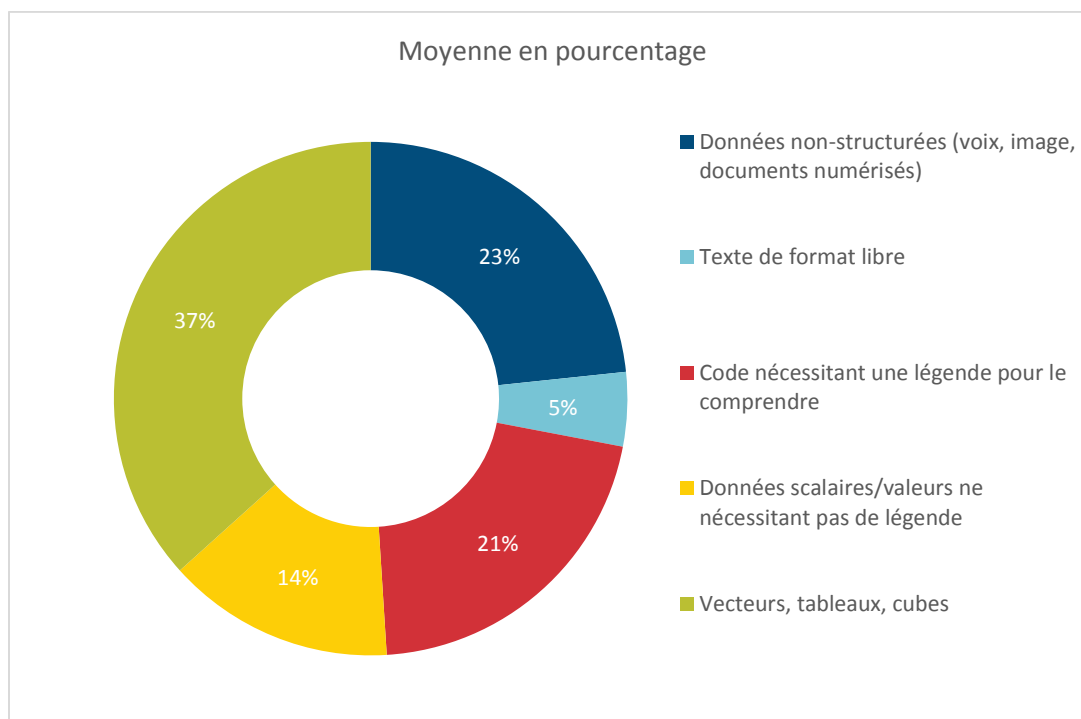


- Les participants au sondage ont indiqué que leurs données étaient complètes et exactes avec une note moyenne de 3,1 sur 5. Bien qu'il ne s'agisse pas d'une mauvaise note, elle montre également qu'il y a beaucoup de place à l'amélioration. Aucun répondant n'a répondu qu'il considérerait ses données comme parfaites (5 sur 5), ce qui indique que chaque répondant croit que l'exhaustivité et l'exactitude de ses données peuvent être améliorées.
- Il existe une différence minime entre la note moyenne des participants de moyenne taille et celle des participants de petite taille. Les participants de grande taille se sont attribués une note moyenne

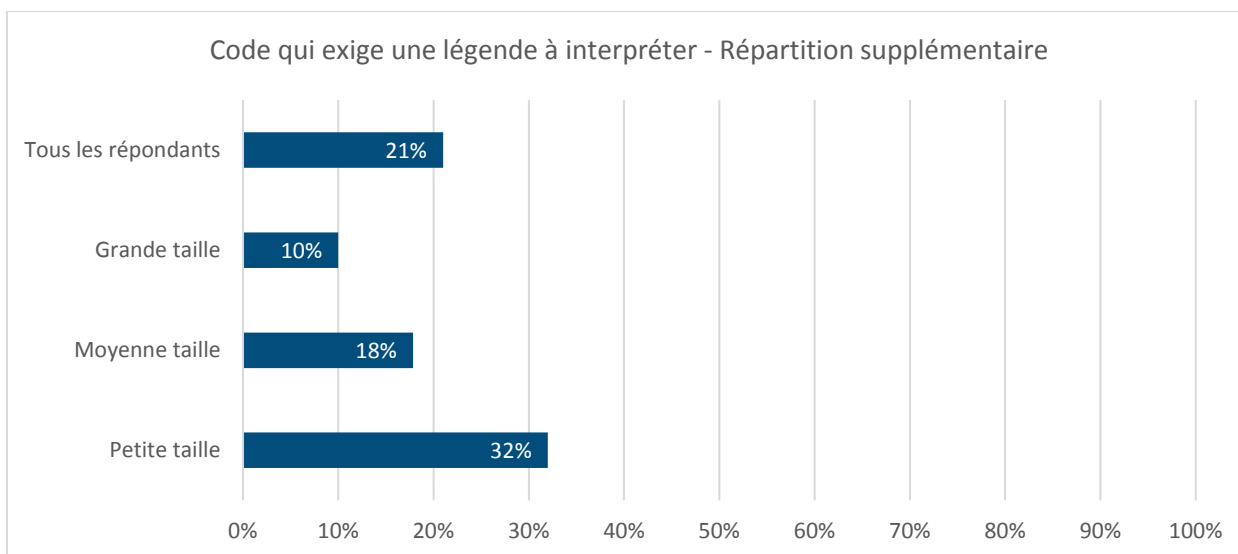
beaucoup plus élevée (4) que les autres répondants (2,8 et 3,0 pour les répondants de moyenne et petite taille respectivement).

- Il n’y a qu’une légère différence entre la cote attribuée aux assureurs (3,14 sur 5) et celle accordée aux réassureurs (3,00 sur 5).

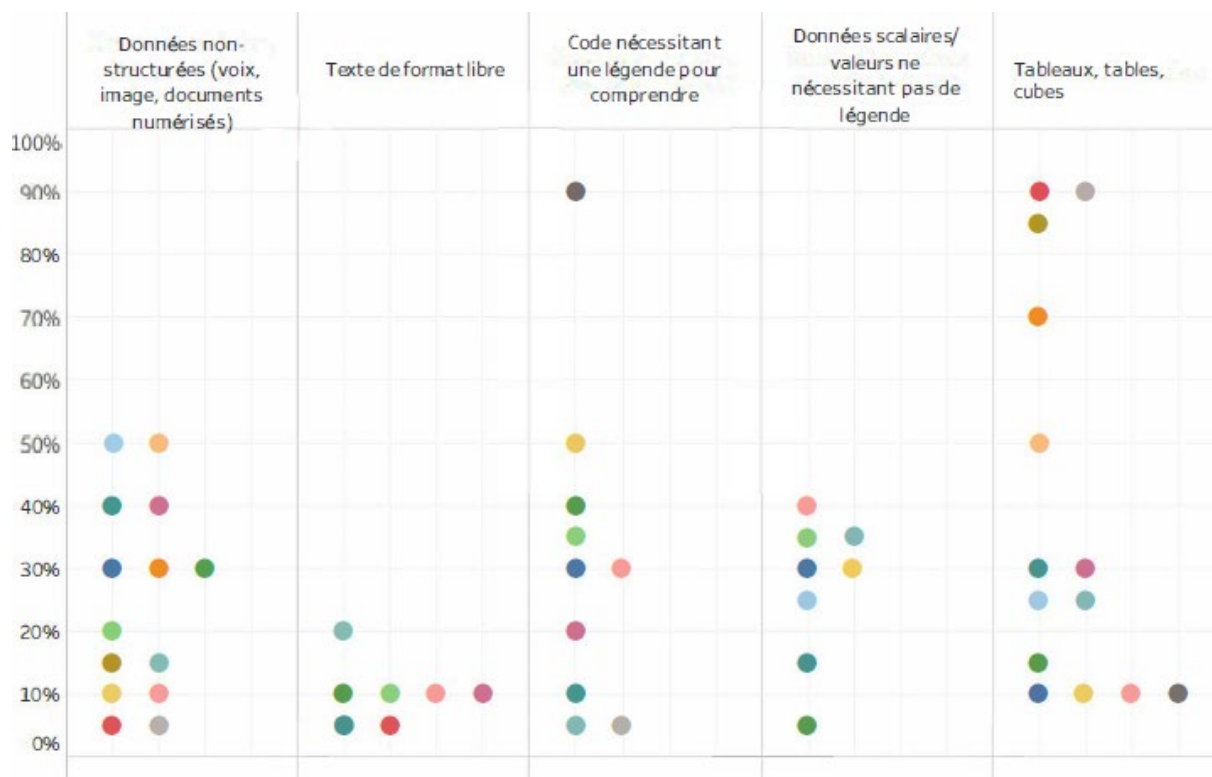
Question 3.3 – Indiquez un pourcentage approximatif de vos données dans chacune des catégories suivantes



- Quel que soit le groupe de taille, il existe une quantité appréciable de données non structurées, dont 23 % sont de ce type.
- Seulement 5 % des données ont été identifiées comme du texte de format libre.
- Par conséquent, moins d’un tiers (28 %) des répondants gèrent des données en format texte non structuré ou en format libre.



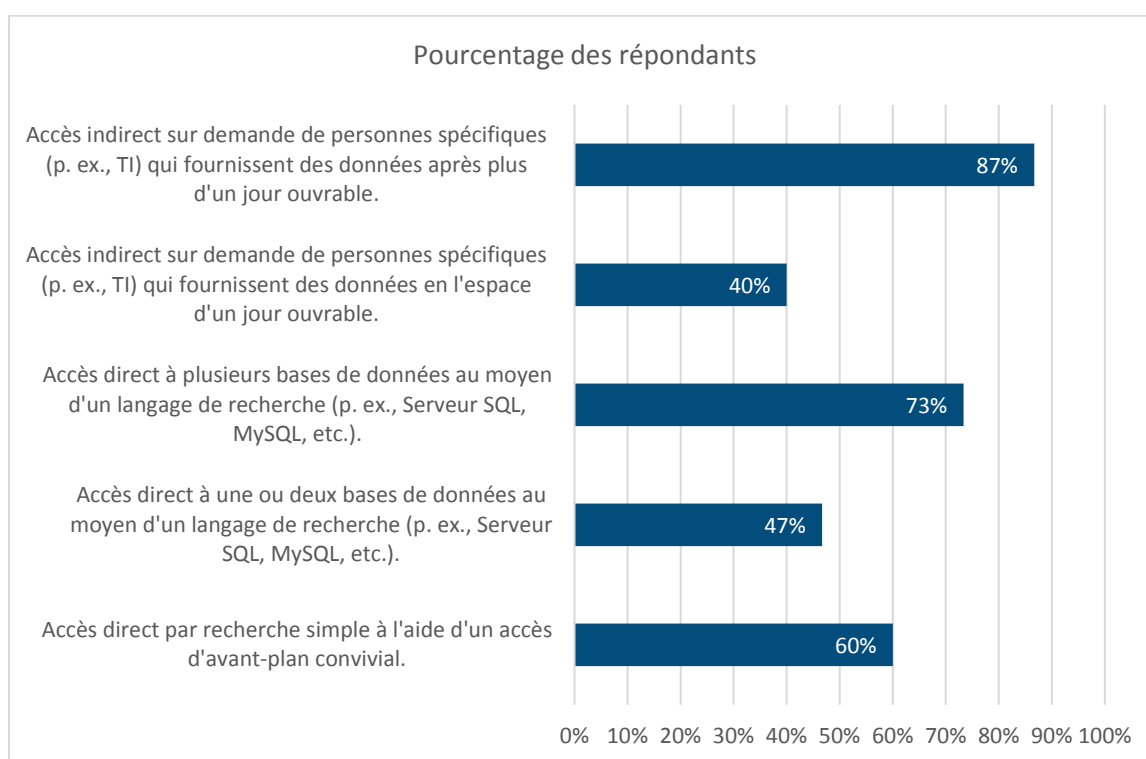
- Environ le cinquième (21 %) des données sont stockées sous forme de code qui nécessite une légende à interpréter. On note une augmentation marquée du pourcentage de ce type de données, lorsque la taille du groupe diminue chez les répondants.



Chaque couleur distincte correspond à l'un des répondants au sondage.

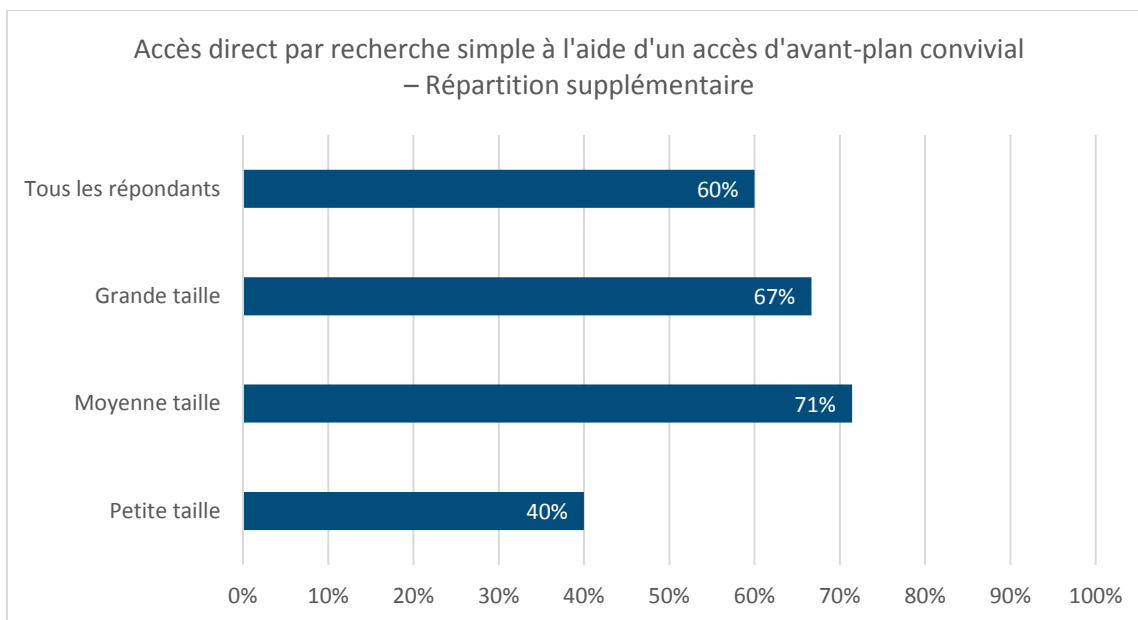
- Les participants au sondage stockent une vaste gamme de types de données et les proportions varient sensiblement entre les répondants.
- D'après notre expérience, la vaste gamme de types de données est à la base des défis auxquels sont confrontées les entreprises qui doivent gérer plusieurs sources de données dont la structure sous-jacente est différente (ou dont il n'y a pas de structure).
- Le graphique précédent montre que pour certains participants, le pourcentage élevé de données non structurées est élevé. Au plan historique, il s'agirait de données telles que les formulaires de demande de règlement et divers documents d'immatriculation (comme la photo d'un permis de conduire). Ces données peuvent être difficiles à traiter et ne peuvent pas être utilisées pour la modélisation prédictive. Toutefois, la question demeure de savoir si ces données mènent ou non à une mine d'information susceptible d'améliorer le pouvoir prédictif.

Question 3.4 – Comment les données sont-elles accessibles à l'échelle de l'organisation? (Choisir toutes les réponses pertinentes)



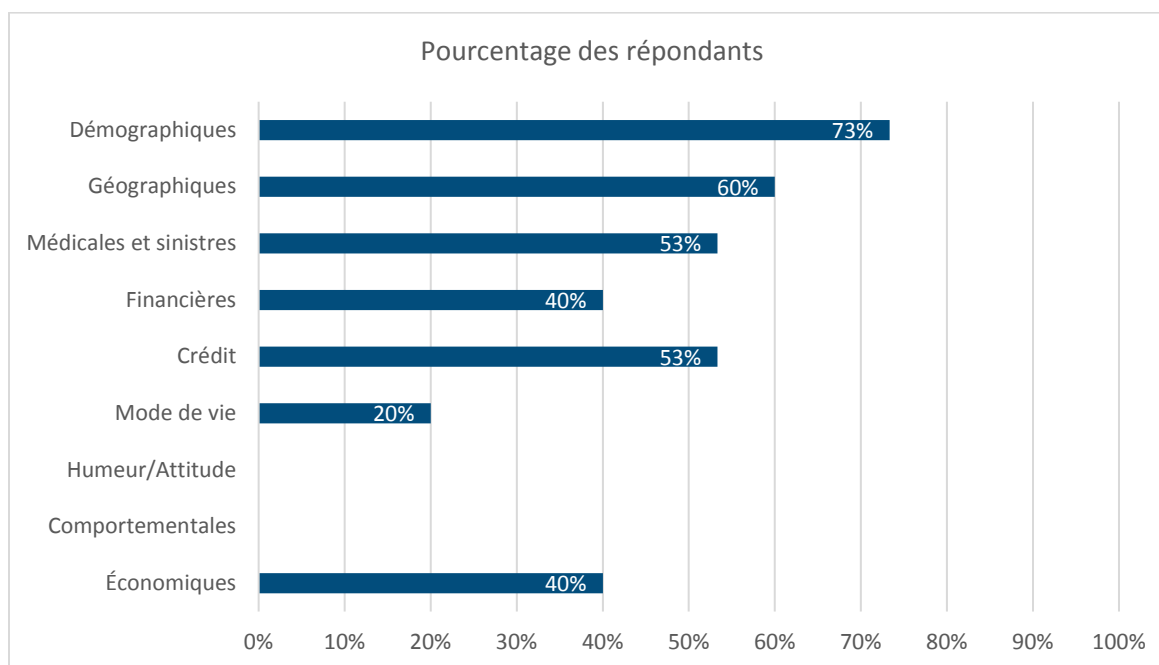
- La majorité des participants au sondage demandent encore au moins certaines données par l'entremise de personnes particulières qui ne sont pas membres de leur groupe d'analytique (comme une équipe de TI distincte).
- Parmi ces répondants qui demandent l'aide d'une autre source, la plupart (87 %) ont un temps d'attente supérieur à un jour ouvrable (comparativement à 40 % qui attendent un jour ou moins). Tous les répondants de grande et de moyenne tailles appliquent un tel processus retardé, contrairement à une fraction des répondants de petite taille.
- L'accès indirect retardé pourrait présenter des problèmes opérationnels, en particulier dans les situations où l'objectif consiste à accélérer/automatiser l'extraction et la production de rapports.

- Selon notre expérience, les pratiques exemplaires d'une entreprise dépendront des utilisateurs qui ont besoin d'accéder aux données. Pour les entreprises qui comptent sur des experts en gestion et en extraction de bases de données, l'accès direct au moyen d'un langage de recherche est souvent le processus le plus efficace. Si l'entreprise n'a pas ces experts, un point d'accès d'avant-plan serait idéal.
- Il convient également de souligner que la conservation de multiples bases de données peut être une nécessité, mais elle peut aussi engendrer des complications lorsqu'on tente de joindre plusieurs sources.



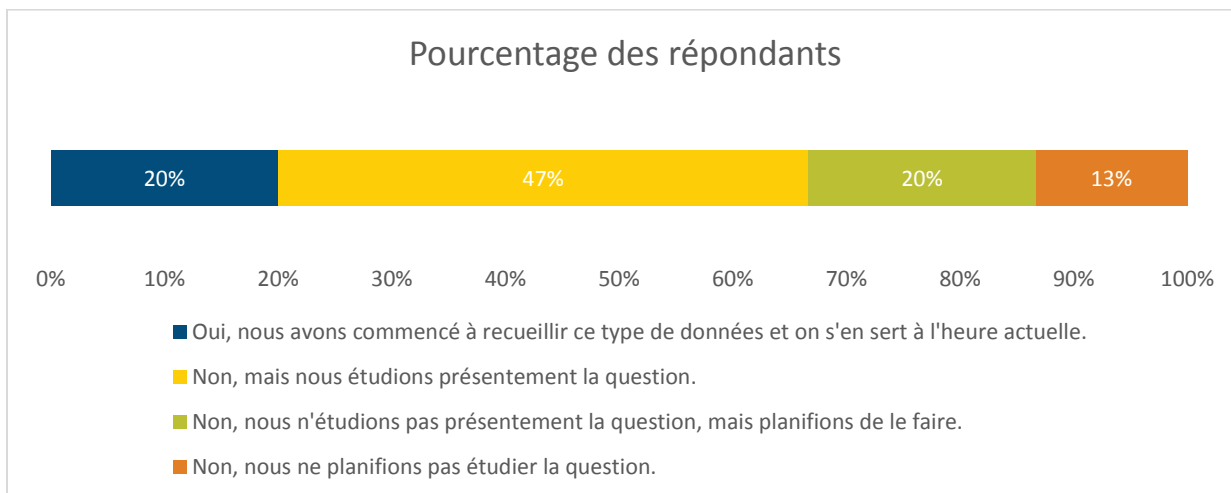
- L'accès direct par recherche simple à l'aide d'un accès d'avant-plan convivial existe davantage chez les répondants de grande taille (67 %) et de moyenne taille (71 %) que chez les répondants de petite taille (40 %).

Question 3.5 – Quels types de données de tiers votre organisation utilise-t-elle actuellement pour l’analytique? (Choisir toutes les réponses pertinentes)



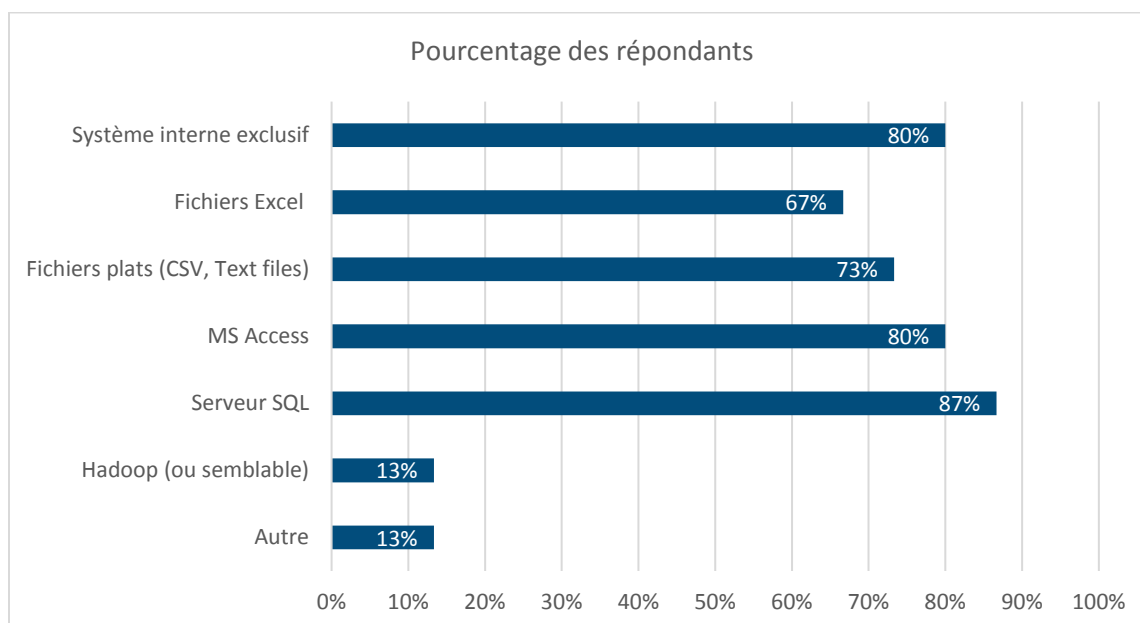
- Parmi les répondants qui ont recours à l’analytique prédictive, tous ont indiqué qu’ils utilisent actuellement des sources de données de tiers pour compléter leurs propres données. Il s’agit d’un signal fort indiquant que les participants au sondage sont prêts à investir dans leurs propres capacités d’analytique et qu’ils font confiance à ces sources tierces.
- Les données démographiques se classent comme le type de données de tiers le plus populaire parmi les répondants à 73 % selon le nombre d’entreprises, suivies des données géographiques à 60 %, des règlements et des données médicales (53 %), du crédit (53 %), et des données financières (40 %) et économiques (40 %).
- Aucun répondant n’a indiqué avoir utilisé des données de tiers pour l’humeur/l’attitude ou le comportement.
- Pour les sinistres et les données médicales, le fait que les répondants de grande taille n’aient pas tous indiqué utiliser ce type de données pourrait être interprété comme s’ils supposaient peut-être que leurs propres données sont suffisamment crédibles pour les applications actuelles.
- Les deux tiers des répondants de grande taille utilisent des données de tiers sur le mode de vie, comparativement à un seul répondant de petite taille.
- Aucun des répondants de petite taille n’utilise de données financières de tiers. Un seul répondant de grande taille utilise des données de crédit de tiers, alors que l’utilisation de ces données est plus répandue chez les répondants de petite et de moyenne tailles.
- De façon générale, plus le répondant est de grande taille, plus il a tendance à utiliser des données démographiques, géographiques, financières et économiques de tiers.
- En outre, les données sur le crédit et les données économiques sont davantage utilisées par les assureurs que par les réassureurs.
- En raison de l’analyse dans laquelle ces sources de données sont habituellement utilisées, nous pouvons déduire que les sources tierces sont principalement utilisées pour l’analyse des résultats (comme les taux de mortalité, les taux de déchéance et les taux de conversion).

Question 3.6 – Avez-vous commencé à recueillir des données au moyen de nouvelles technologies (p. ex., Fitbit) pour accroître l’utilisation des sources traditionnelles de données?



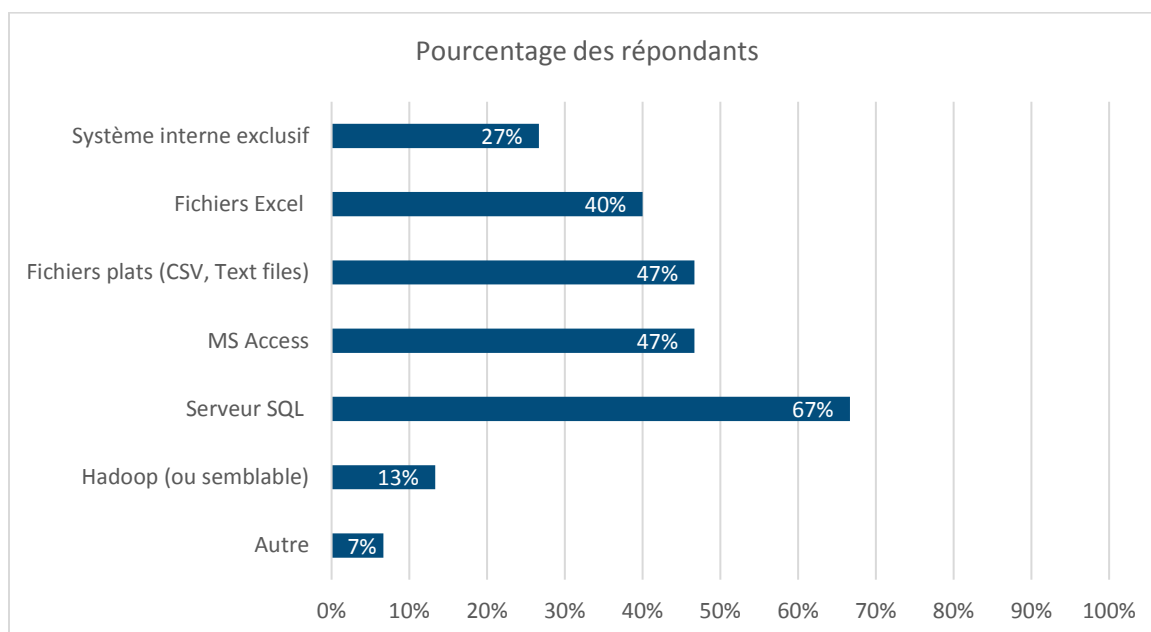
- Un répondant de chaque groupe de taille (grande, moyenne, petite) a indiqué avoir recueilli ce type de données et s’en sert à l’heure actuelle.
- Aucun participant au sondage n’a indiqué qu’il a recueilli des données et qu’il ne les utilise pas, ni n’a fait d’enquête et a décidé de ne pas recueillir ou utiliser ce type de données.
- Tous les répondants de grande taille qui n’ont pas encore recueilli de telles données étudient la valeur de l’augmentation des données grâce aux nouvelles technologies. Il en va de même pour tous les répondants de moyenne taille, sauf deux.
- De façon générale, la majorité des participants au sondage utilisent ou envisagent d’utiliser les données issues des nouvelles technologies. Cela confirme un sentiment croissant selon lequel d’autres sources de données pourraient être utiles aux assureurs et aux réassureurs. Toutefois, on ne sait trop pour le moment si l’utilisation de ces données procure un avantage concurrentiel à ces entreprises.

Question 3.7 – Comment stockez-vous vos données internes?



- Les plateformes les plus populaires pour stocker des données externes sont Serveur SQL (87 %), Système interne exclusif (80 %), MS Access (80 %), Fichiers plats (73 %) et Excel (67 %). La plateforme Hadoop a également été mentionnée.
- Parmi les autres logiciels mentionnés, citons DB2 et FileNet.

Question 3.8 – Comment stockez-vous vos données externes?



- Comme prévu, les plateformes de stockage utilisées pour les données internes et externes sont très différentes.
- Les plateformes les plus populaires pour stocker des données externes sont Serveur SQL (67 %), MS Access (47 %), Fichiers plats (47 %) et Excel (40 %). Des systèmes internes et Hadoop ont également été utilisés.
- Parmi les autres logiciels mentionnés, citons PDF et Word.
- Exemples d'autres logiciels : DB2, Oracle et FileNet.
- L'utilisation d'un des logiciels énumérés dans cette question présente des avantages et des inconvénients. La meilleure option pour une entreprise dépend encore une fois des utilisateurs finaux qui accèdent aux données. Cela dit, certains principes de gestion des données sont appliqués plus facilement à certains logiciels que d'autres. Par exemple :
 - Le Serveur SQL de Microsoft vous permet de créer des clés principales et étrangères qui relient les tableaux de votre base de données. Microsoft Excel ne comporte pas en soi cette fonction.
 - D'autres considérations comme la taille de vos données, les protocoles de sécurité, la convivialité et la vitesse de traitement dépendront grandement du logiciel que vous choisirez pour stocker vos données.
- Une grande partie des participants au sondage utilisent encore un logiciel « moins qu'idéal » pour gérer leurs données à l'échelle de l'entreprise. Cela comprendrait les données conservées spécifiquement dans Excel ou une autre extension de fichier plat (comme les fichiers csv et txt). Les pourcentages élevés d'autres logiciels « idéaux » indiquent que la plupart des entreprises les utilisent également, mais tout simplement pas pour toutes leurs sources de données.
- Les données de tiers sont stockées plus régulièrement dans une base de données relationnelle. Habituellement, les données de tiers se présentent déjà dans un format structuré, où la relation entre les tableaux est précisée, de sorte qu'il est plus facile de les stocker dans une base de données relationnelle. Toutefois, cela pourrait aussi indiquer que les entités qui achètent des données de tiers sont plus techniquement avancées dans la gestion des données et préfèrent donc les stocker dans une base de données relationnelle.
- Hadoop (ou un logiciel semblable) n'est encore utilisé que par un faible pourcentage des participants au sondage. Il se peut que ce pourcentage augmente au cours des prochaines années, car les participants au sondage semblent traiter un niveau élevé de données non structurées.

Sommaire – Données

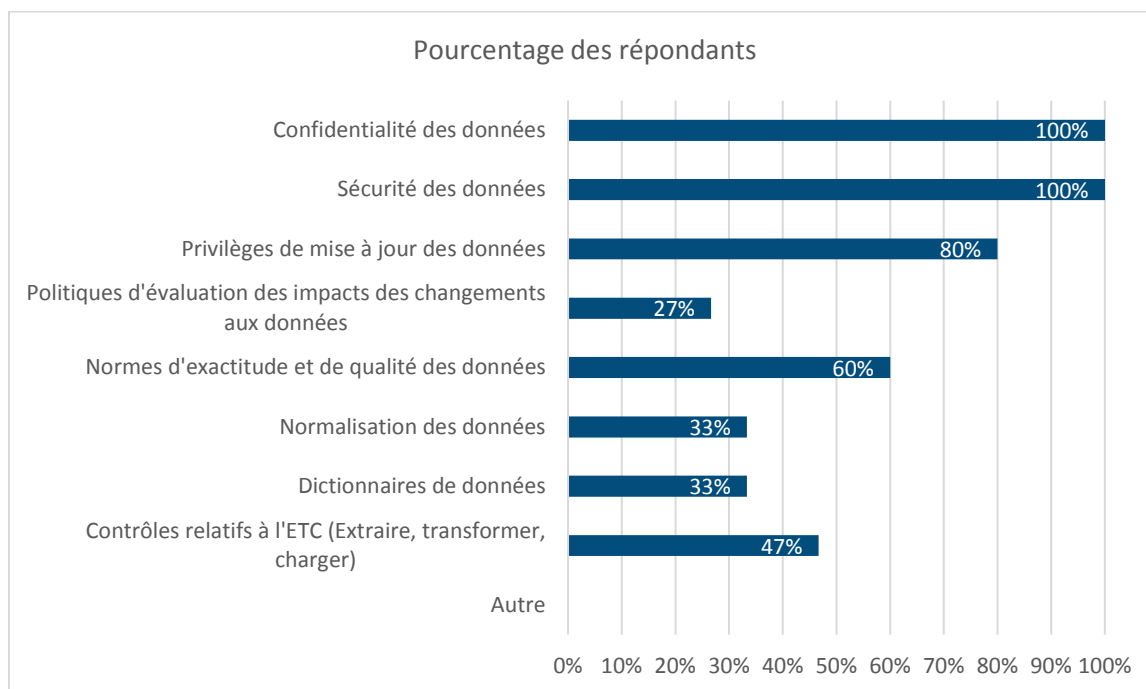
D'après les réponses au sondage, il y a lieu d'apporter des améliorations aux données aux fins de l'analytique prédictive en ce qui concerne la centralisation, la qualité des données et les procédures d'accès aux données pertinentes.

En outre, les participants n'ont pas encore sensiblement mis à profit les nouvelles technologies (aux fins de l'augmentation des données), et c'est un point que les assureurs commencent à exploiter.

6.2 Entreprise

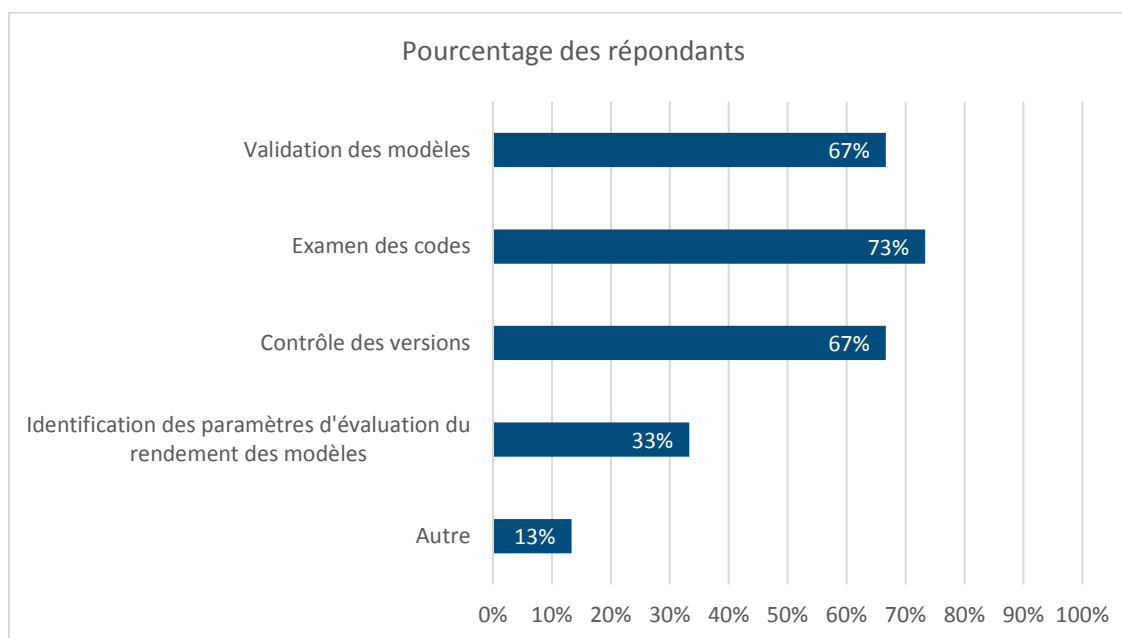
Il est important d'avoir le matériel, les logiciels et la structure nécessaires pour gérer les données. Comment les données sont-elles stockées, gérées et consultées par les analystes de l'organisation?

Question 4.1 – Sélectionnez les éléments et politiques de gouvernance qui s'appliquent aux données utilisées dans l'analytique prédictive pour votre organisation. (Choisir toutes les réponses pertinentes.)



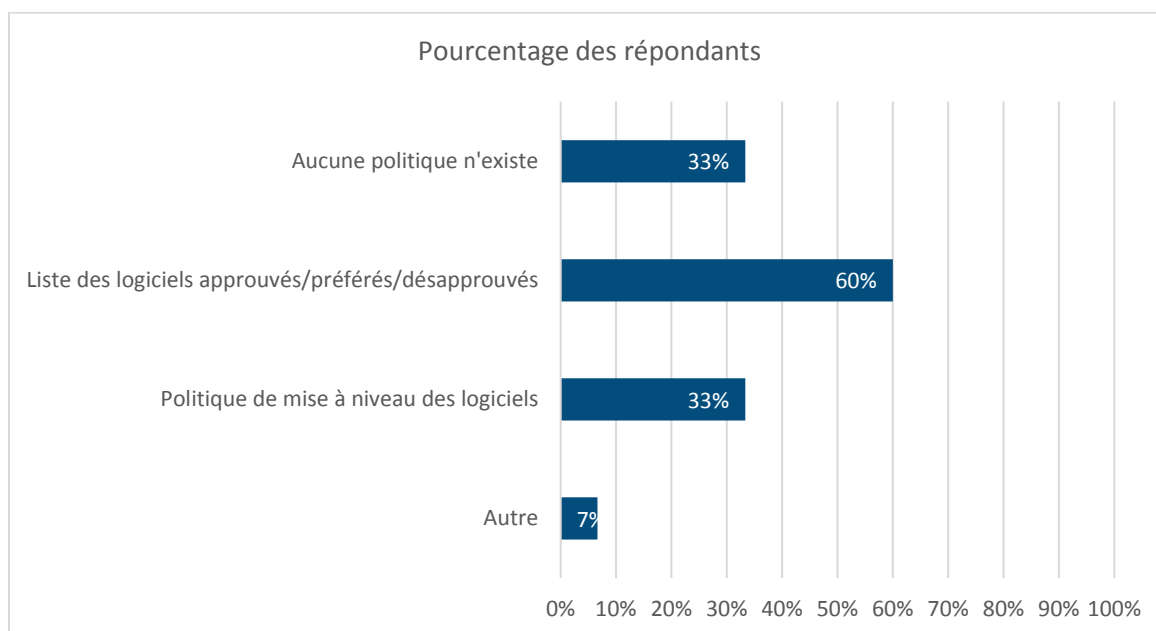
- Tous les répondants ont indiqué avoir des mesures de gouvernance et des politiques en matière de confidentialité des données et de sécurité des données. Parmi les autres politiques/mesures de gouvernance populaires en place, mentionnons les privilèges de mise à jour des données (80 %), les normes d'exactitude et de qualité des données (60 %) et les contrôles relatifs à l'ETC (47 %).
- Peu de répondants (environ le tiers) ont des mesures de gouvernance et des politiques en matière de normalisation, de production de dictionnaires de données et de gestion du changement de données.
- On met davantage l'accent sur les mesures de gouvernance et les politiques relatives aux privilèges de mise à jour des données, à l'exactitude des données et aux normes de qualité pour les répondants de plus grande taille.
- L'accent le plus marqué est mis sur la confidentialité des données et la sécurité des données, ce qui est logique compte tenu des points de données souvent traités par les assureurs et les réassureurs. Un risque élevé est associé à l'exposition de renseignements de nature délicate, et les participants au sondage prennent activement des mesures pour atténuer ce risque.
- On met moins l'accent sur les politiques de gouvernance qui peuvent faciliter la modélisation prédictive. Puisque le risque associé à ces politiques est moindre, il semble que certains répondants ne les promulguent pas. Cela dit, il existe toujours un risque associé au manque de connaissances concernant ses propres données. Il est considéré comme une pratique exemplaire de mettre en place des politiques pour chacun des éléments énumérés ci-dessus.

Question 4.2 – Sélectionnez les aspects et les politiques de gouvernance qui influent sur les efforts de modélisation de votre organisation. (Choisir toutes les réponses pertinentes.)



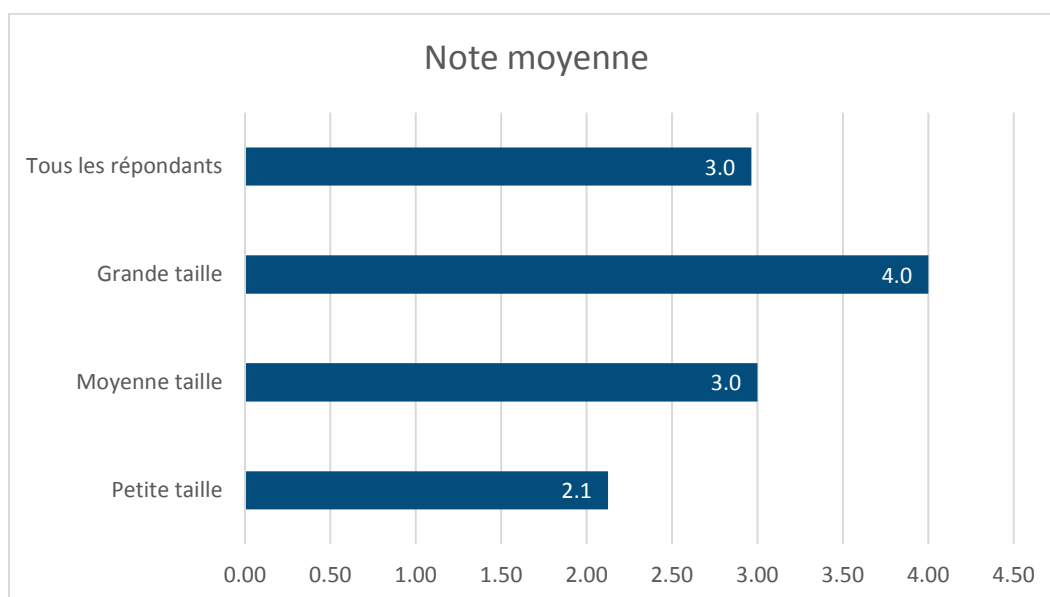
- Environ les deux tiers des participants au sondage ont des politiques de gouvernance pour leurs modèles prédictifs. Le tiers restant des participants au sondage n'effectuent que très peu de travaux de modélisation prédictive, de sorte que les politiques de gouvernance dans ce domaine sont moins prioritaires.
- Les mesures et les politiques de gouvernance les plus populaires ayant une incidence sur la modélisation comprennent l'examen des codes (73 %), la validation des modèles (67 %) et le contrôle des versions (67 %), qui sont tous importants pour le processus de modélisation.
- Parmi les autres réponses mentionnées, citons les normes de documentation, le suivi des problèmes et la documentation des codes.
- Parmi les résultats intéressants, mentionnons qu'environ le tiers seulement des participants au sondage ont une politique de gouvernance concernant les paramètres d'évaluation du rendement des modèles.
- Tous les répondants de grande taille ont des fonctions de validation des modèles, d'examen des codes et de contrôle des versions, mais seulement les deux tiers d'entre eux ont des paramètres d'évaluation du rendement définis. Nous sommes d'avis que la détermination de mesures d'évaluation appropriées et la définition des critères constituent une considération importante pour la sélection du modèle optimal et que les sociétés auraient avantage à adopter de telles politiques.

Question 4.3 – Sélectionnez les aspects et les politiques de gouvernance liés aux logiciels et à la technologie utilisés dans les applications d’analytique de votre organisation. (Choisir toutes les réponses pertinentes.)



- Les mesures de gouvernance et les politiques populaires en matière de logiciels et de technologie comprennent : la liste des logiciels approuvés (60 %) et la politique de mise à niveau des logiciels (33 %).
- Jusqu’à un tiers des répondants n’avaient ni mesures de gouvernance ni politiques en place concernant les logiciels et la technologie. Nous nous serions attendus à ce que très peu de répondants n’aient pas de telles politiques en place.
- D’autres mesures de gouvernance et politiques mentionnées par les répondants comprenaient les lignes directrices sur le déploiement et la politique sur le contexte de préproduction et de production.
- Les deux tiers des répondants de grande taille ont une liste de logiciels approuvés, préférés ou désapprouvés. Un seul répondant de grande taille, deux répondants de moyenne taille et deux répondants de petite taille ont des politiques de mise à niveau des logiciels. Ainsi, seulement cinq répondants sur 15 ont indiqué avoir des politiques de mise à niveau des logiciels.
- Dans l’ensemble, la gouvernance des logiciels est moindre que celle des données et des modèles. L’une des hypothèses qui expliquent cette situation réside dans le phénomène des logiciels de source ouverte (« open-source » en anglais). Comme l’ont démontré les questions ultérieures du sondage, les logiciels de source ouverte (R, par exemple) sont devenus les plus utilisés par les participants au sondage. Les politiques de gouvernance devraient être examinées fréquemment pour qu’elles demeurent à jour et suivent l’évolution des logiciels de source ouverte et de leurs progiciels.
- Une attention particulière devra être portée à l’avenir aux progiciels créés par l’utilisateur dans les logiciels de source ouverte. Bien que le code de base puisse être testé périodiquement, il se peut que le code écrit par l’utilisateur ne soit pas convenablement testé. Il conviendrait peut-être de déterminer s’il y aura un changement et que des entreprises commenceront à élaborer davantage de politiques de gouvernance sur les progiciels produits par les utilisateurs.

Question 4.4 – Si les fonctions d’analytique ne sont pas entièrement centralisées, veuillez évaluer la force de la coordination des compétences et des ressources au sein de l’organisation. [1 = extrêmement faible / 5 = extrêmement forte]



- Dans l’ensemble, les répondants ont attribué une note de 3,0 sur 5 à la force de la coordination des compétences et des ressources.
- Les répondants de grande taille ont accordé une note beaucoup plus élevée (4 sur 5) à la coordination des compétences et des ressources, suivis des répondants de moyenne taille (3 sur 5) et des répondants de petite taille (2,1 sur 5).
- La note globale donne à penser qu’il existe un certain niveau de coordination acceptable entre les participants au sondage, mais qu’il y a encore place à amélioration.

Sommaire – Entreprise

Nous constatons, d’après les réponses, que des progrès doivent être réalisés en ce qui concerne les politiques de gouvernance et de normalisation.

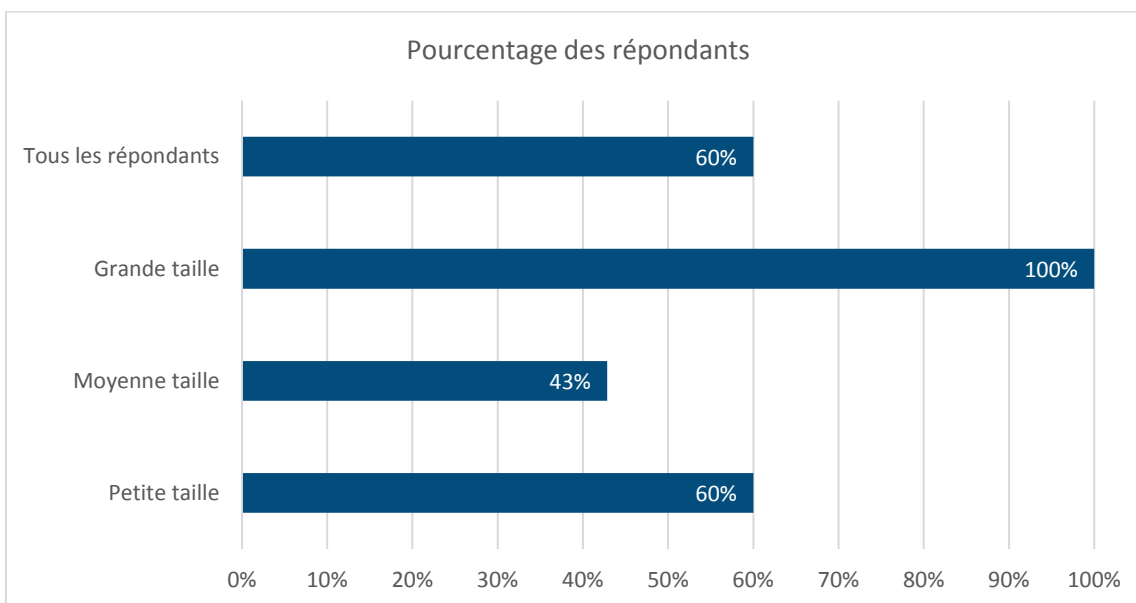
Les participants de petite et moyenne tailles doivent combler l’écart qui les sépare des participants de grande taille en ce qui concerne la gouvernance des données (plus particulièrement la mise à jour des données et leur exactitude, ainsi que les normes). Il est probable que le même écart s’applique à ces groupes de taille pour le reste du secteur. De même, l’écart doit être comblé au chapitre de la gouvernance de la modélisation (notamment la validation des modèles, l’examen des codes et le contrôle des versions).

Seulement cinq répondants sur 15 ont indiqué qu’ils avaient des politiques de mise à niveau des logiciels, ce qui est sensiblement inférieur à ce qui est peut-être justifié.

6.3 Leadership

Les cadres supérieurs souhaitent débloquer le pouvoir de l'analytique au sein de leurs organisations. Sont-ils disposés à recruter les bonnes personnes et à consacrer du temps à l'élaboration des systèmes et des processus pertinents?

Question 5.1 – Votre organisation compte-t-elle dans ses rangs un cadre supérieur chargé des données? Dans l'affirmative, quel est son titre?

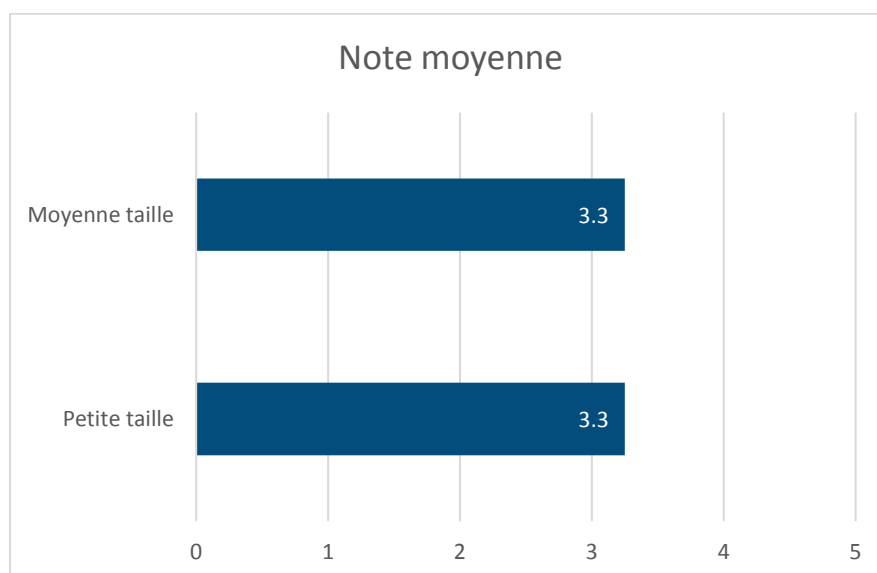


- Dans l'ensemble, 60 % des participants au sondage comptent un cadre supérieur chargé des données, et tous les répondants de grande taille ont indiqué l'existence d'un tel rôle. Il est étonnant de constater qu'un plus grand nombre de répondants de petite taille que de répondants de moyenne taille disposaient d'une telle fonction, même si on ne sait trop si ce cadre avait d'autres responsabilités (p. ex., le dirigeant principal de l'information).

Voici un résumé des titres des cadres supérieurs chargés des données :

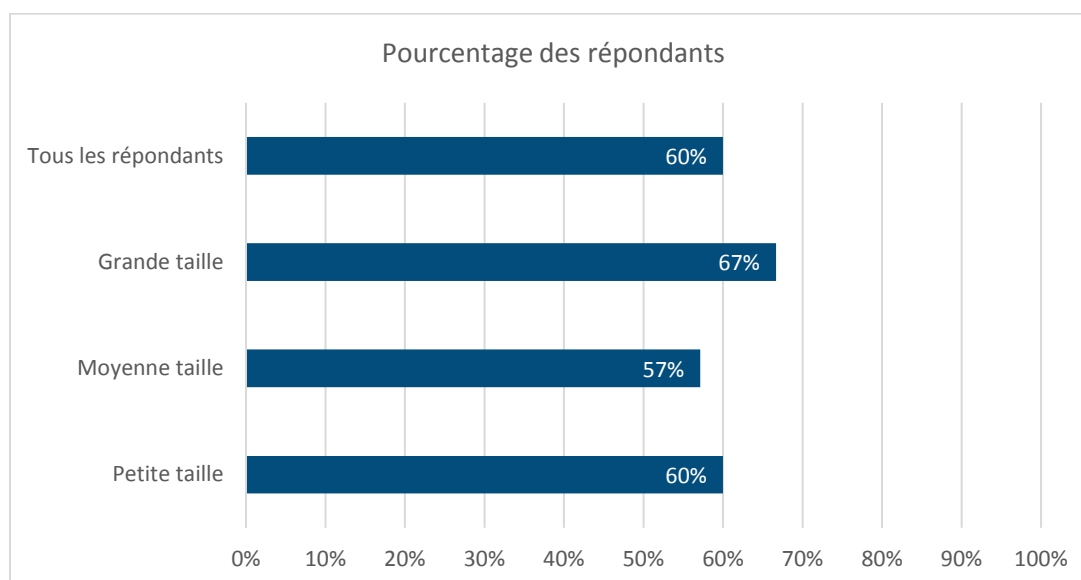
Titres	Nombre
Dirigeant principal des données	3
Vice-président en chef/premier vice-président	1
Vice-président	3
[le poste existe, mais le titre est inconnu]	1
Aucun	7

Question 5.1a – Si aucun cadre supérieur n’est chargé des données, dans quelle mesure la direction comprend-elle les données (qualité, structure, mesures de protection, etc.)? [1 = Aucune compréhension/ 5 = Excellente compréhension]



- Dans les cas où aucun cadre supérieur n’est chargé des données, la compréhension des données par les dirigeants était de 3,3 sur 5.
- D’après notre expérience, l’utilisation de tableaux de bord de haut niveau montrant l’état des données permet de mieux faire comprendre les données aux dirigeants.

Question 5.2 – Votre organisation compte-t-elle dans ses rangs un cadre supérieur chargé de l’analytique prédictive? Dans l’affirmative, quel est son titre?

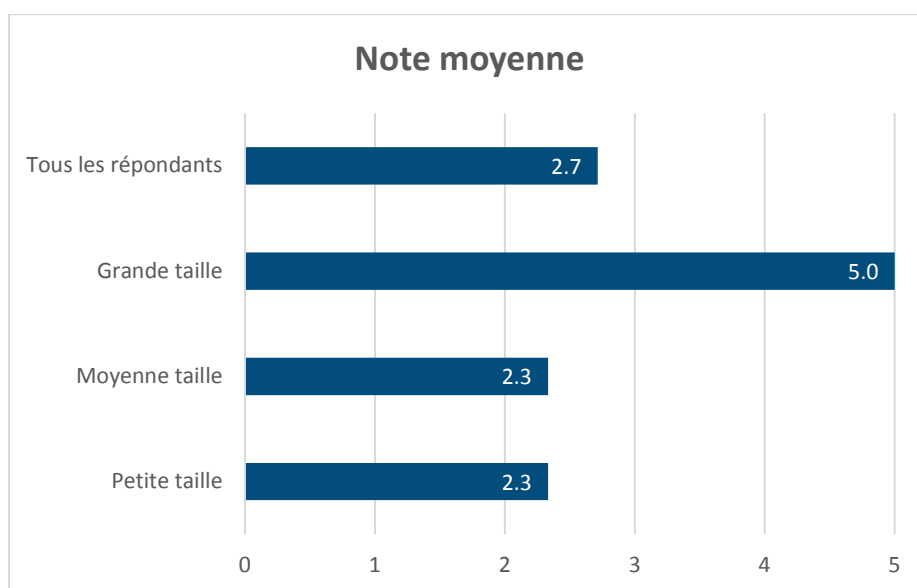


- Dans l’ensemble, 60 % des participants au sondage comptent un cadre supérieur chargé de l’analytique prédictive, et plus de 67 % des répondants de grande taille ont indiqué l’existence d’un tel rôle.

Voici un résumé des titres des cadres supérieurs chargés de l'analytique prédictive :

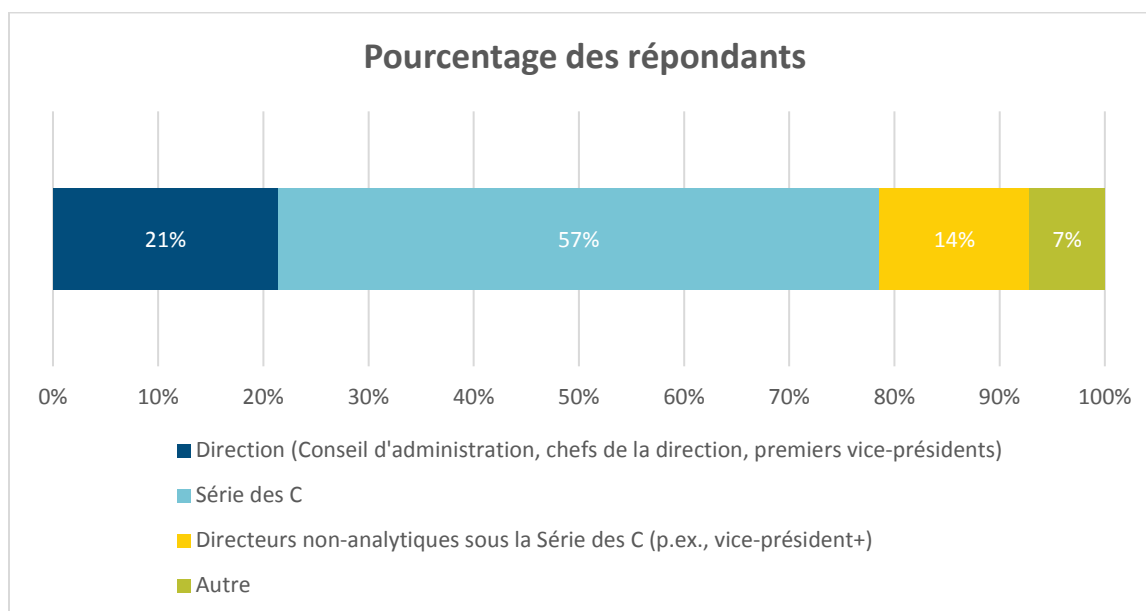
Titres	Nombre
Dirigeant principal de l'analyse	1
Vice-président en chef/premier vice-président	1
Premier vice-président	2
Vice-président	3
Directeur	1
[le poste existe, mais le titre est inconnu]	1
Aucun	6

Question 5.2a - Si aucun cadre supérieur n'est chargé de l'analytique prédictive, dans quelle mesure la direction comprend-elle ses possibilités de valeur ajoutée et les initiatives des concurrents? [1 = Aucune compréhension/5 = Excellente compréhension]



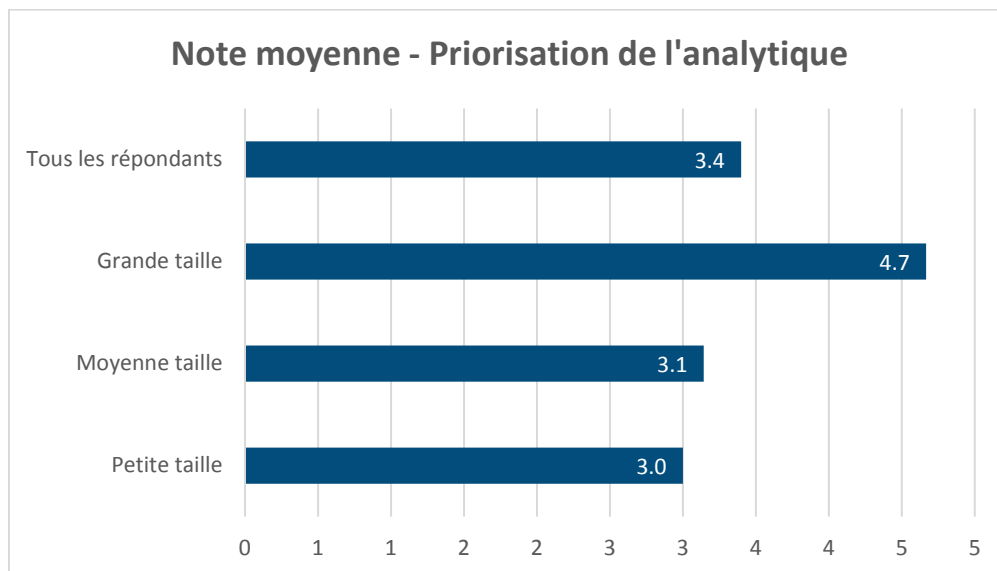
- Dans les cas où aucun cadre supérieur n'est chargé de l'analytique, la compréhension des possibilités de valeur ajoutée et des initiatives des concurrents par les dirigeants était de 2,7 sur 5.
- Les répondants de grande taille ont indiqué une compréhension parfaite.
- Cela sous-entend que, pour les sociétés autres que de grande taille, il peut parfois être difficile de faire reconnaître la nécessité/les résultats des modèles prédictifs de la part des cadres supérieurs, surtout aux premières étapes du processus d'analytique prédictive d'une société. De même, on pourrait y parvenir au moyen de tableaux de bord faciles à utiliser ou, dans le cas de la modélisation prédictive, de la conversion à un résultat quantifiable facile à comprendre par les cadres supérieurs.

Question 5.3 – Qui prend la décision finale sur la priorisation des initiatives d’analytique prédictive?



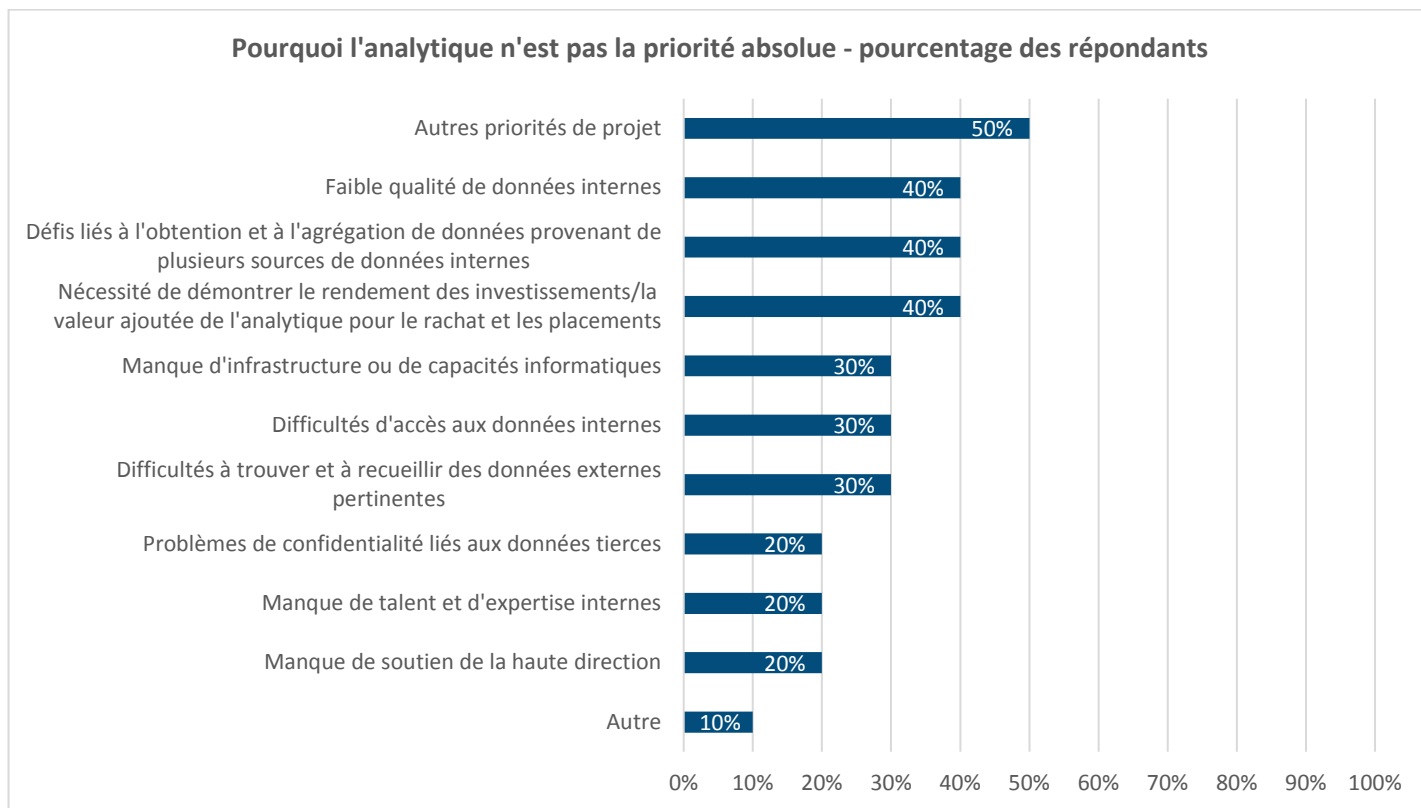
- La majorité des participants au sondage (78 %) ont indiqué que les cadres supérieurs (p. ex., Série des C ou de niveau plus élevé) prennent les décisions financières concernant les initiatives de modélisation prédictive. Près du cinquième (21 %) des conseils d’administration, des chefs de la direction et des premiers vice-présidents participent à ces décisions.
- Les deux tiers des décisions sont prises au niveau de la Série des C pour les répondants de grande et de moyenne tailles, tandis que toutes les décisions sont prises à ce niveau pour les répondants de petite taille.
- Bien qu’on puisse s’y attendre, on constate également un décalage entre les décisions financières et la compréhension générale des données et des pratiques de modélisation. Cela signifie probablement que l’on fait confiance aux cadres intermédiaires et l’on suppose qu’ils savent ce qu’ils font pour chaque initiative de modélisation prédictive.

Question 5.4 – Quelle est la priorité de l'analytique prédictive dans le cadre des projets de votre organisation? [1 = Faible priorité/5 = Investissement le plus important pour l'avenir]



- En moyenne, les répondants de grande taille ont indiqué que les initiatives d'analytique étaient hautement prioritaires. La priorité accordée par les répondants de petite et de moyenne tailles était plus faible et n'était pas sensiblement différente pour ces groupes de taille. Il convient de souligner qu'en moyenne, les réassureurs ont accordé une priorité sensiblement plus élevée (3,8) aux initiatives d'analytique que les assureurs (3,3).

Question 5.4a – Si l’analytique prédictive n’est pas la priorité absolue (p. ex., note de 5 à la question précédente), quelles sont les principales raisons? (Cochez toutes les réponses pertinentes.)



- Jusqu’à 10 répondants ont indiqué que l’analytique n’est pas la priorité absolue.
- Les raisons les plus courantes comprenaient l’existence d’autres priorités de projet (50 %), la faible qualité des données internes (40 %), les difficultés à obtenir des données de multiples sources internes (40 %) et la nécessité de démontrer le rendement des investissements (40 %).
- Le manque d’infrastructure ou de capacités informatiques, la difficulté d’accès aux données internes et la difficulté de recueillir des données externes ont également été mentionnés.
- Il ne semble pas exister de manque généralisé de soutien de la haute direction.

Sommaire – Leadership

60 % de tous les répondants disposent d’un cadre chargé de l’analytique des données et de l’analytique prédictive; tous les répondants de grande taille en ont un pour les données et les deux tiers pour l’analytique prédictive.

78 % des décisions finales relatives à l’analytique prédictive sont prises par les membres de la série des C ou de niveau supérieur.

Seulement le cinquième des membres du conseil d’administration, les chefs de la direction et les premiers vice-présidents participent aux décisions concernant l’analytique prédictive (cela nous semble faible et nous nous attendions à une plus grande participation pour quelque chose d’aussi important que l’analytique prédictive).

Les principales raisons de la réduction de l’ordre de priorité comprennent d’autres priorités opérationnelles concurrentes, la qualité des données, les difficultés liées à l’agrégation des données et la nécessité de prouver le rendement des investissements dans l’analytique.

6.4 Cibles

Les buts et cibles doivent être définis pour les travaux d'analytique. Les décisions stratégiques et les utilisateurs ont-ils été identifiés?

Question 6.1 – Pour quelles applications actuelles et passées utilisez-vous l'analytique? Les « applications » sont définies comme des cas où un modèle prédictif (ou ML, IA, etc.) a été étalonné à des fins de prédiction, pour identifier les facteurs/variables explicatives, etc.

Comment évaluez-vous chacune de ces applications en fonction des efforts perçus pour la mise en œuvre et de la valeur perçue pour l'entreprise?

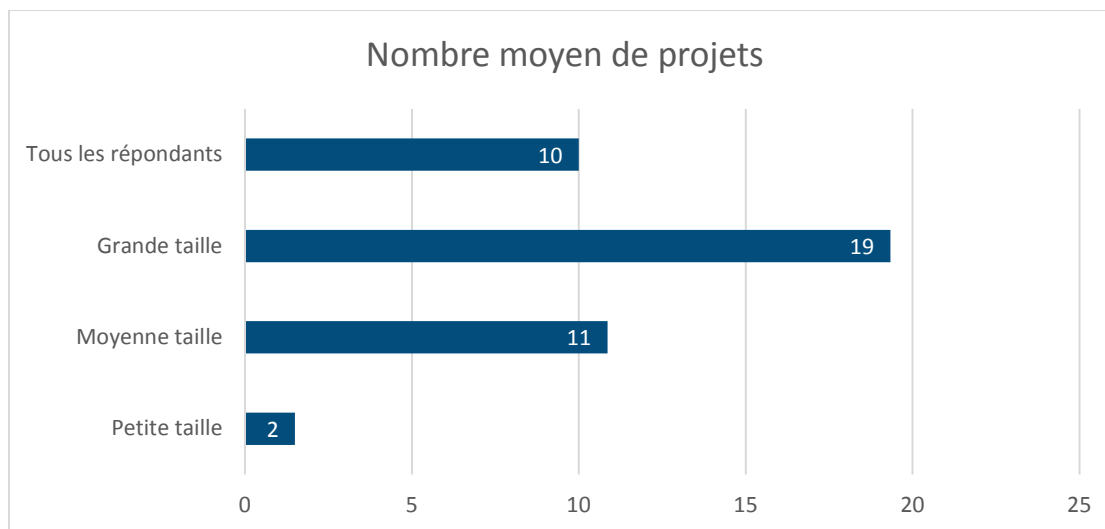
Voir le résumé aux pages 11 à 15.

Question 6.2 – Pour quelles applications prévoyez-vous utiliser l'analytique au cours de la prochaine année?

Comment évaluez-vous chacune de ces applications en fonction des efforts perçus pour la mise en œuvre et de la valeur perçue pour l'entreprise?

Voir le résumé aux pages 16 à 18.

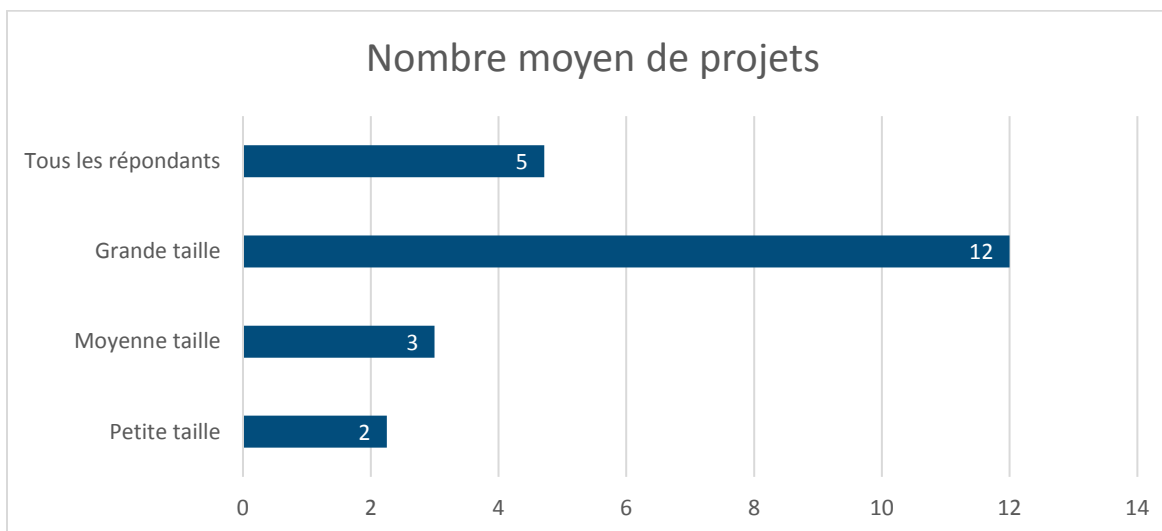
Question 6.3 – Environ combien de projets d'analytique avez-vous tentés au cours de la dernière année? [On a demandé aux participants de fournir le compte des applications/produits, tels que la police de 10 ans avec déchéance]



- Les participants au sondage ont exécuté en moyenne 10 projets au cours d'une année : 19 pour les participants de grande taille, 11 pour les participants de taille moyenne et deux pour les participants de petite taille. Un répondant de grande taille a indiqué un grand nombre de projets, ce qui a sensiblement majoré la moyenne pour cette catégorie.
- Le cinquième des répondants n'ont exécuté aucun projet au cours de la dernière année.
- Jusqu'à 42 % des répondants ont exécuté trois projets ou moins, y compris ceux qui n'en avaient aucun.
- Aucun des petits répondants n'a exécuté plus de trois projets.

- Il semble que les participants au sondage prennent part activement à une mission d'analytique prédictive, ou y sont encore relativement nouveaux, et qu'un petit nombre de répondants tentent d'accomplir une modeste quantité de travail dans le domaine.

Question 6.4 – Environ combien de nouvelles applications d'analytique qui n'ont pas été exécutées auparavant devraient être traitées au cours de la prochaine année?



Pour cette question, nous avons demandé aux répondants d'inclure de nouveaux types d'analytique et d'exclure l'application simple d'une initiative d'analytique antérieure à un autre groupe opérationnel.

- Tous les répondants sauf un ont des projets prévus pour l'avenir.
- Une moyenne totale de cinq projets est prévue pour l'ensemble des participants au sondage : 12 projets pour les participants de grande taille, trois pour les participants de moyenne taille et deux pour les participants de petite taille.
- La plupart des répondants auront plus d'un nouveau projet.

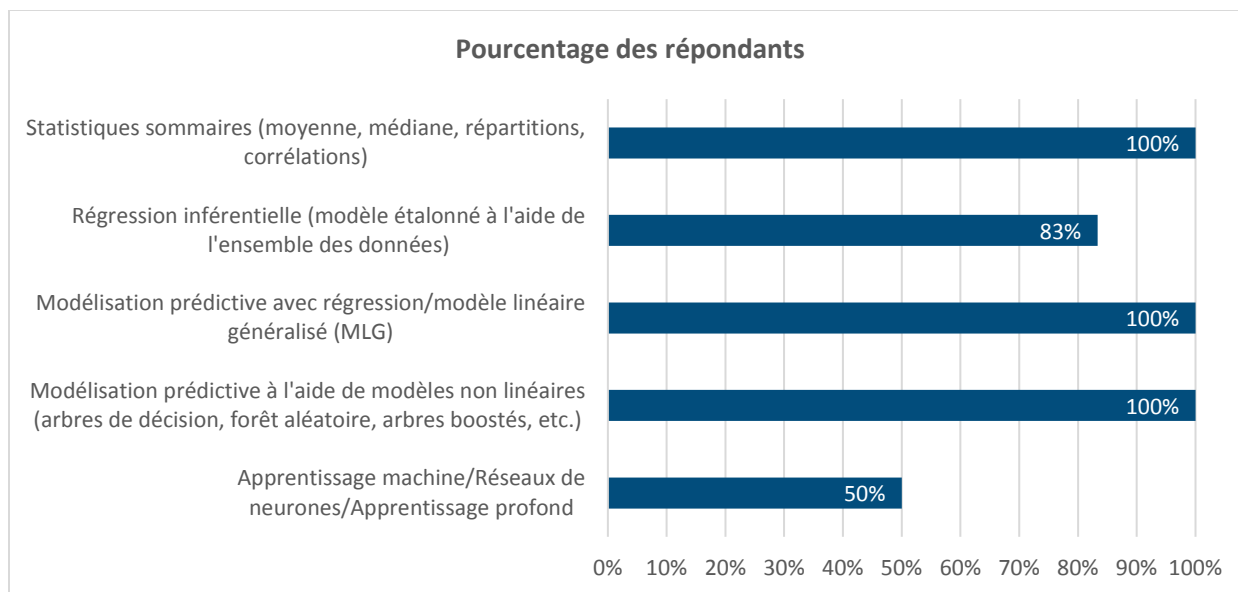
Sommaire – Cibles

Les participants de grande taille exécutent en moyenne plus de projets et prévoient plus de projets pour l'avenir que les participants de moyenne ou de petite taille.

6.5 Analystes

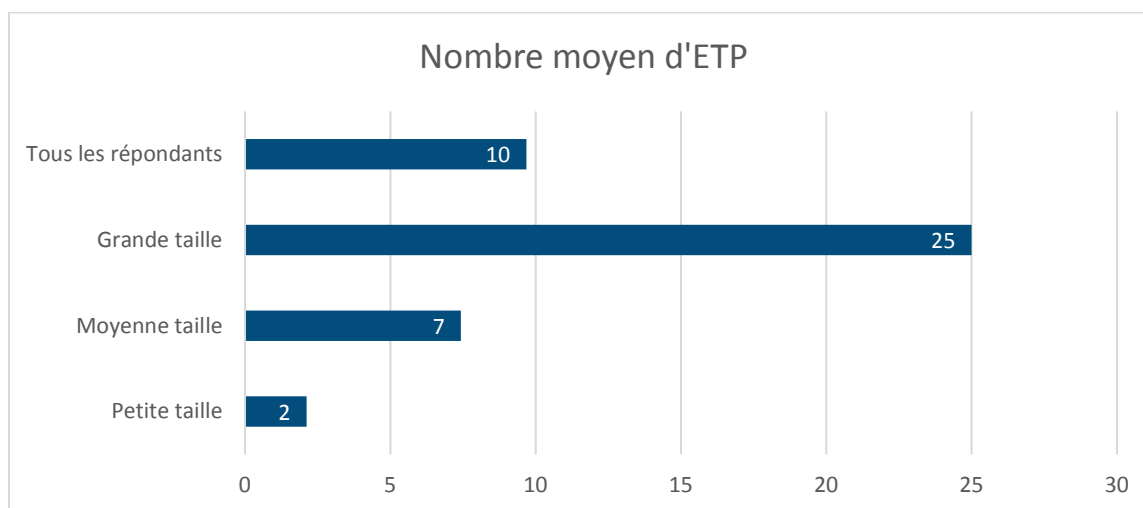
Les analystes sont différents les uns des autres et il est important de déterminer la bonne correspondance pour chaque organisation. A-t-on embauché des analystes qui répondent aux exigences opérationnelles pour réussir dans le cadre de leurs travaux d'analytique?

Question 7.1 – Quelles techniques sont actuellement utilisées?



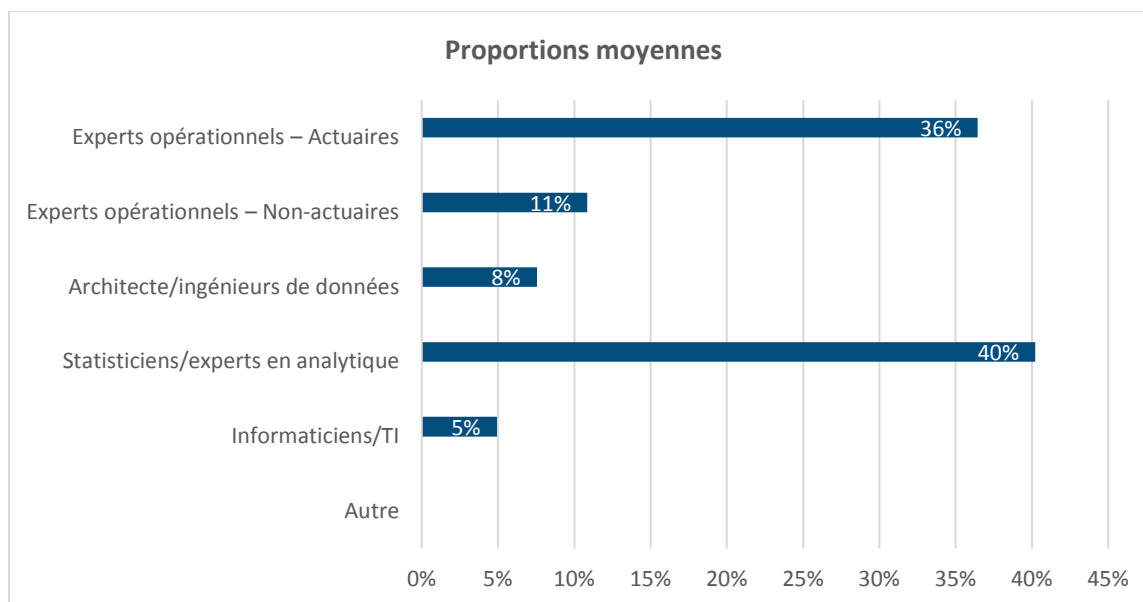
- Parmi ceux qui effectuent de l'analytique prédictive (cette question exclut trois répondants qui n'avaient pas encore effectué de modélisation analytique au moment de l'entrevue), tous les répondants ont indiqué qu'ils mettent en œuvre des statistiques sommaires, des MLG et des techniques d'arbre de décision.
- La plupart des participants au sondage ont recours à plusieurs approches, et cinq répondants ont indiqué toutes les techniques.
- Parmi les techniques les moins utilisées, mentionnons l'apprentissage machine, les réseaux de neurones et l'apprentissage profond, seulement 50 % des répondants ont utilisé de telles techniques.
- L'approche traditionnelle de compréhension des données consiste à produire un certain type de statistiques sommaires. Il n'est donc pas étonnant de constater que les sociétés l'aient largement inclus dans leur processus d'analytique, qui comprend également la compréhension des données.

Question 7.2 – Combien d'équivalents temps plein (ETP) travaillent à l'analytique?



- Le nombre moyen d'ETP affecté à l'analytique est 10, réparti ainsi : 25 pour les répondants de grande taille, sept pour les répondants de moyenne taille et deux pour les répondants de petite taille.
- Il ressort clairement de ce qui précède que les répondants de grande taille ont consacré plus de ressources à l'analytique, alors que certains répondants de moyenne ou petite taille n'en ont pas encore consacré.
- Les répondants qui sont des assureurs ont consacré plus de ressources que les réassureurs.

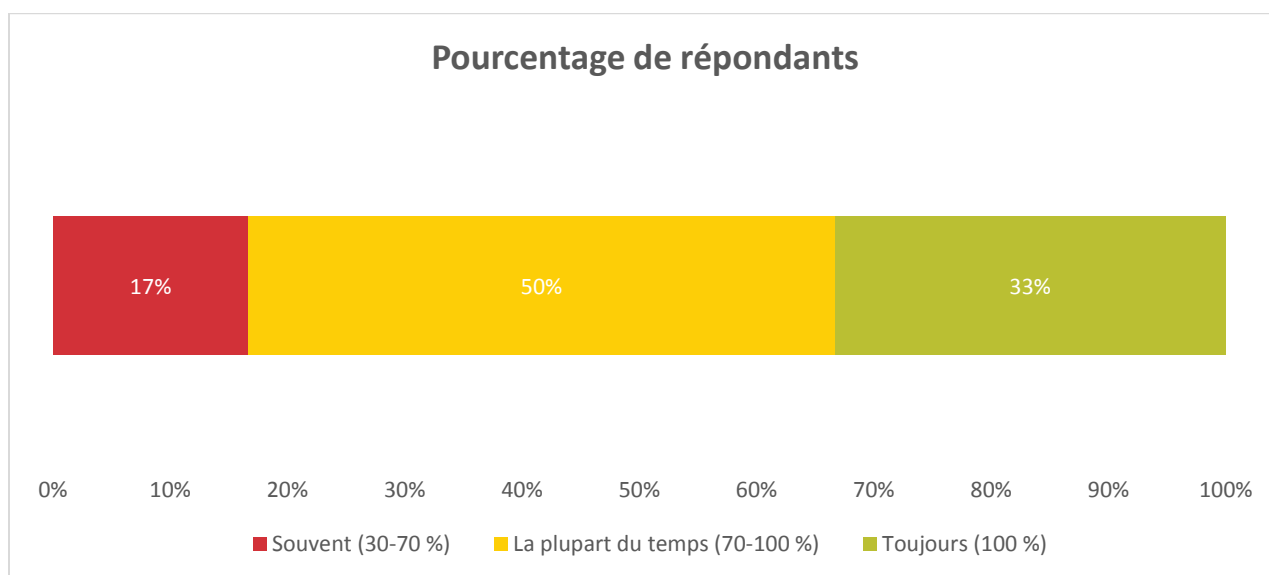
Question 7.3 – Quelles sont les proportions approximatives d'ETP dans chacune des catégories suivantes?



- Dans l'ensemble, près de 50 % des ETP sont des experts opérationnels comparativement aux architectes des données et aux statisticiens.
- Environ les trois quarts des experts opérationnels sont des actuaires.

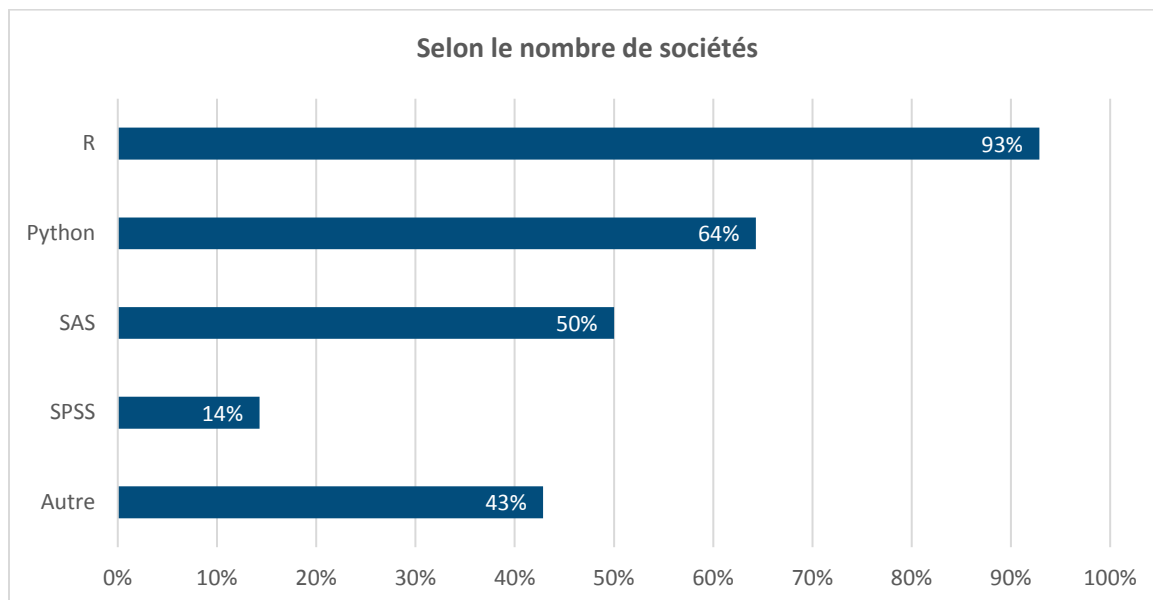
- Les participants de grande taille ont tendance à avoir recours à plus d'architectes de données/de statisticiens (78 %), tandis que les participants de petite taille ont tendance à avoir recours à plus d'experts opérationnels (actuariers/non-actuariers) (81 %).
- Les réassureurs ont tendance à faire appel à plus d'experts opérationnels que les assureurs (83 % contre 35 %, respectivement).
- Enfin, un recours limité à des informaticiens et à des spécialistes de la TI a lieu au moment du sondage (5 % des ressources seulement).

Question 7.4 – Dans chaque projet d'analytique prédictive, dans quelle mesure intégrez-vous les compétences des experts opérationnels (actuariers et non-actuariers), des statisticiens/experts en analytique et du personnel des TI?



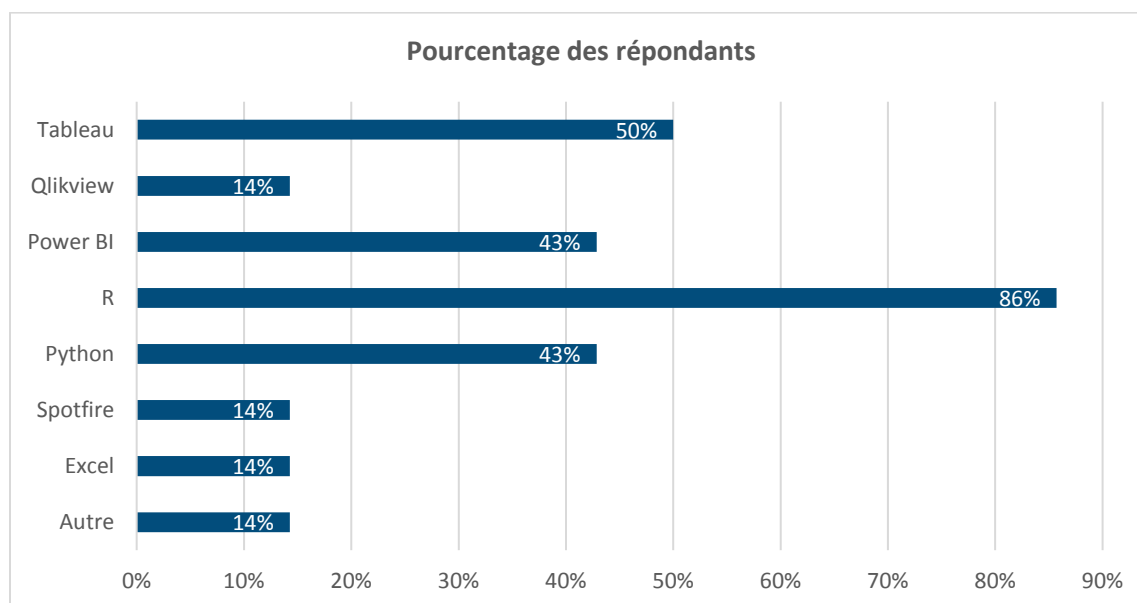
- Au total, 33 % des répondants ont indiqué avoir toujours eu recours à une équipe intégrée et 50 % ont indiqué qu'ils l'utilisaient la plupart du temps. Ainsi, 83 % des répondants ont indiqué qu'ils intègrent les compétences la plupart du temps ou plus souvent.
- Il en résulte une meilleure intégration des connaissances opérationnelles et une plus grande puissance prédictive en modélisation, de même qu'un taux de réussite plus élevé en résolution de problèmes opérationnels pertinents.

Question 7.5 – Quel logiciel utilisez-vous pour la modélisation de l'analytique prédictive? (Choisir toutes les réponses pertinentes.)



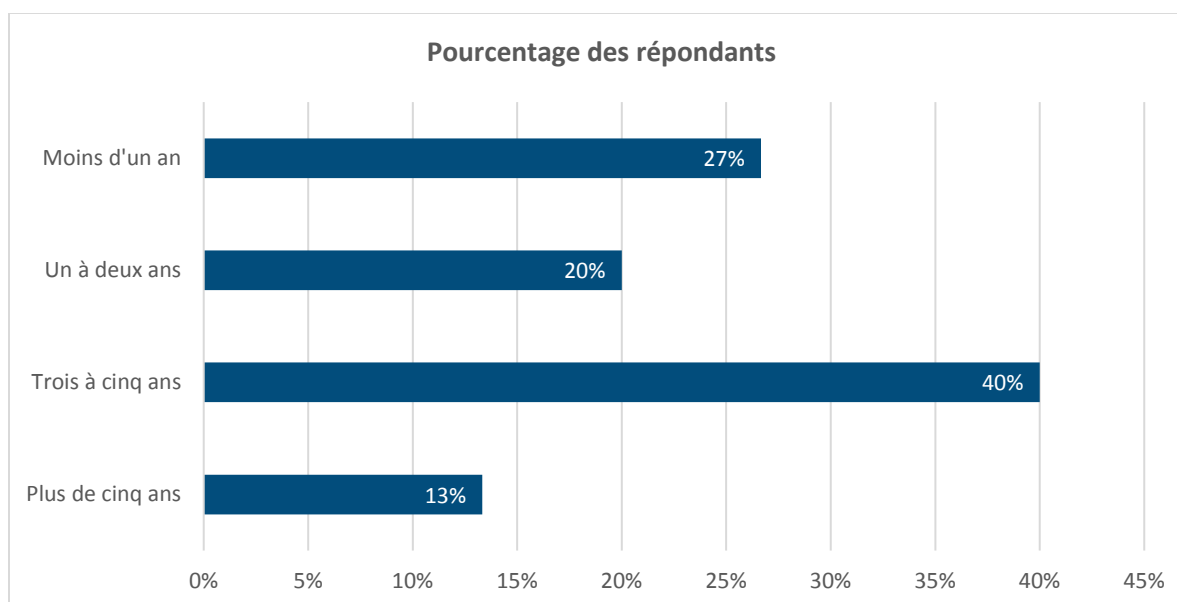
- Le logiciel de source ouverte R est très populaire auprès de 93 % des participants au sondage, soit 13 participants. Les autres logiciels populaires comprenaient Python (64 % ou neuf participants) et SAS (50 % ou sept participants).
- Les participants au sondage ont indiqué qu'ils utilisent d'autres logiciels comme Data Meer, MS Power BI, Statistica, SPSS, Matlab et IBM Watson.
- SAS est davantage utilisé par les participants de moyenne taille que par les participants de grande taille.

Question 7.6 – Quel logiciel utilisez-vous pour l’analyse et la visualisation des données exploratoires?
(Choisir toutes les réponses pertinentes.)



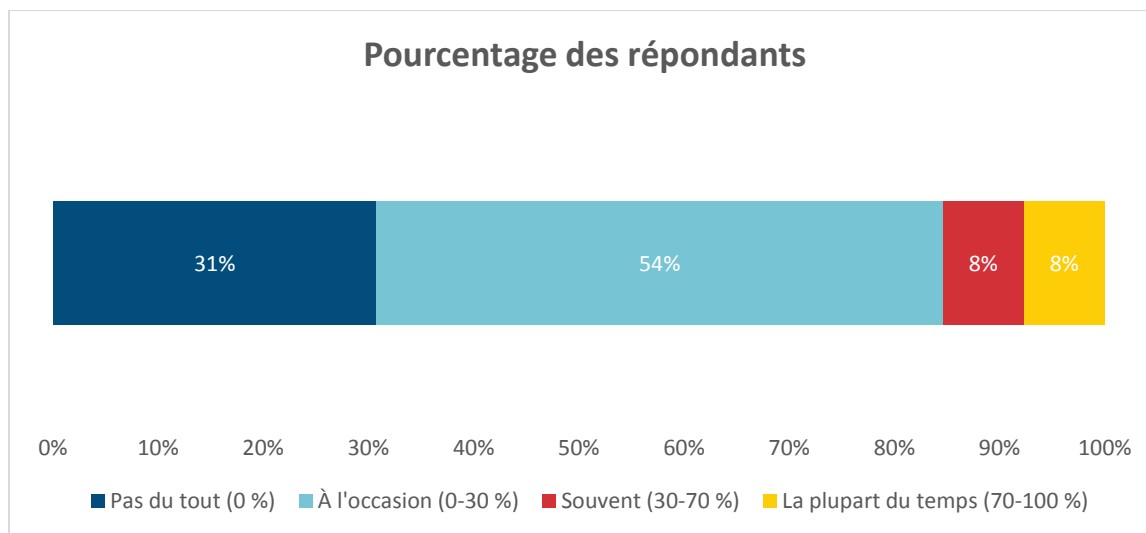
- R demeure le logiciel le plus populaire sur le plan de l’utilisation pour l’analyse et la visualisation, à 86 % ou 12 participants au sondage.
- Parmi les autres logiciels populaires au moment du sondage, mentionnons Tableau (50 %), Power BI (43 %) et Python (43 %).
- Un autre logiciel mentionné était SAS.

Question 7.7 – Depuis combien de temps votre entreprise effectue-t-elle la modélisation prédictive?



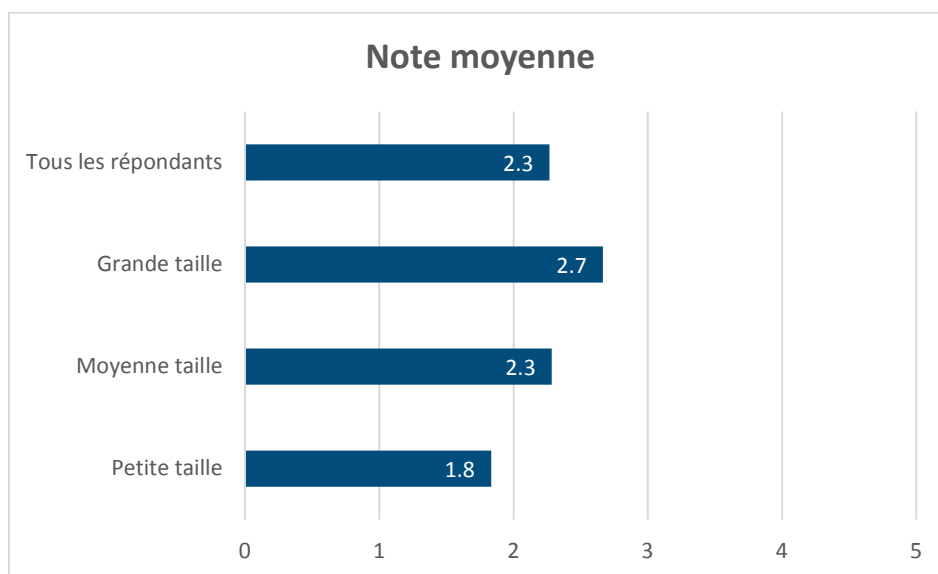
- Le nombre moyen d'années pendant lesquelles les participants au sondage ont effectué de l'analytique s'établissait à 3,3 ans.
- Les répondants de plus grande taille ont tendance à effectuer de l'analytique pendant plus d'années (en moyenne, quatre ans) et tous les répondants de plus grande taille en ont effectué pendant plus de trois ans (67 % entre trois et cinq ans et 33 % pendant plus de cinq ans).
- Aucun des répondants de petite taille n'a effectué des analyses pendant plus de cinq ans (80 % pendant moins de deux ans).

Question 7.8 – Lorsque vous exécutez une nouvelle application d'analytique qui n'a pas été effectuée antérieurement, dans quelle mesure faites-vous appel à des experts-conseils externes?



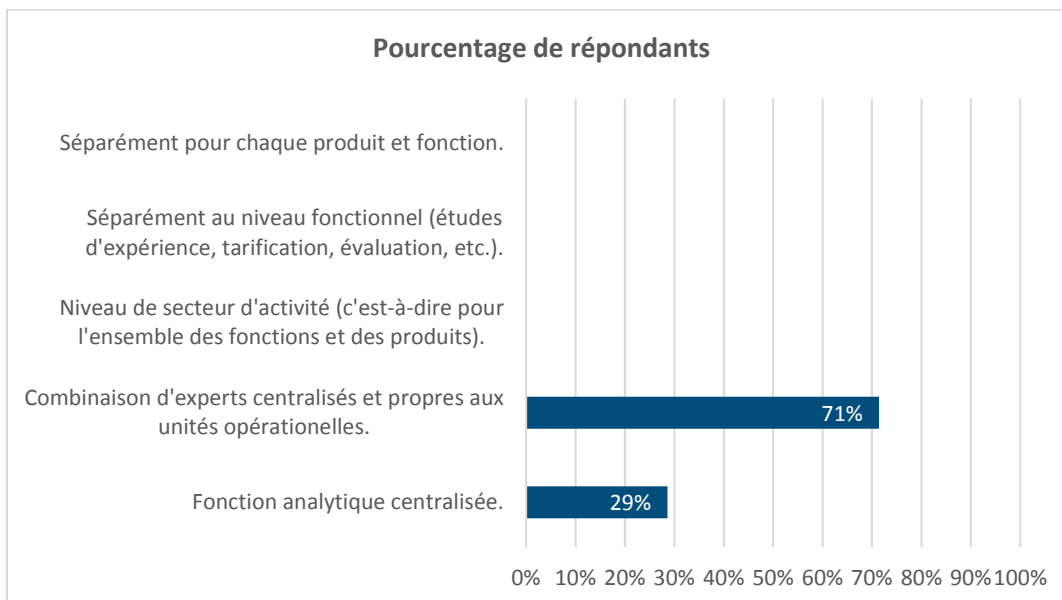
- Les participants de grande taille ne recrutent qu'à l'occasion des experts-conseils, tandis que les participants de petite taille sont plus susceptibles d'en recruter.
- Pour la majorité des participants au sondage, la fonction de modélisation prédictive est demeurée à l'interne pour la plupart de leurs travaux de modélisation, plutôt que d'être impartie à un expert-conseil/fournisseur de TI externes. Cela suppose que la modélisation prédictive devient une nécessité pour la trousse d'outils de l'actuaire. Les actuaires exécutent ces travaux et, compte tenu des tendances générales de l'industrie, il semble que ce sera une compétence utile dans l'avenir.
- Toutefois, près de 70 % des participants ont eu recours à des experts-conseils externes à l'occasion pour exécuter une nouvelle application d'analytique.

Question 7.9 – Dans quelle mesure est-il difficile de recruter et de maintenir en poste des experts de l’analytique pour des postes à l’interne? [1 = Extrêmement difficile/5 = Facile]



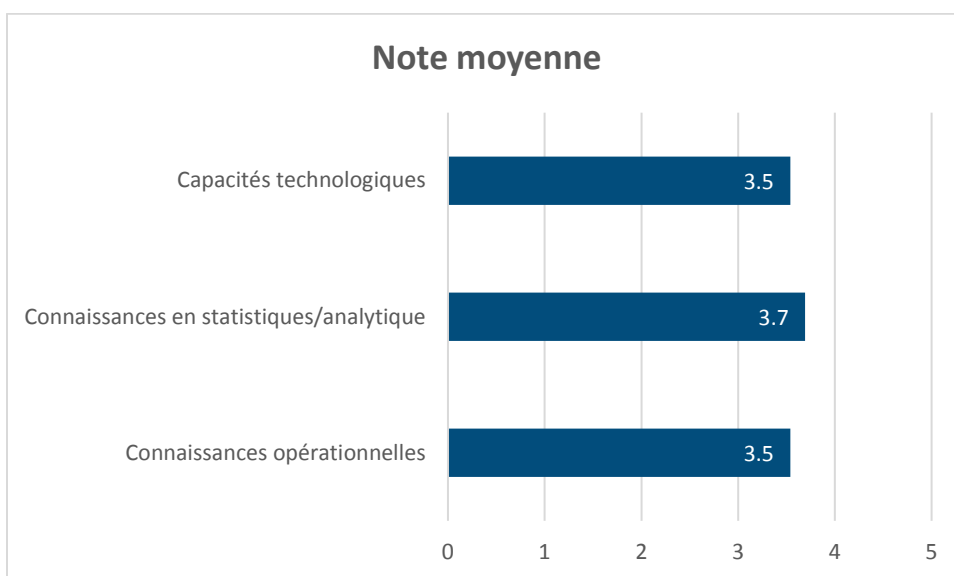
- Dans l’ensemble, il est difficile d’embaucher et de maintenir en poste des experts de l’analytique (note de 2,3). Le consensus est de déclarer qu’il est moyennement difficile ou difficile de trouver des talents pour la fonction de modélisation prédictive.
- On constate qu’il est plus facile pour les répondants de plus grande taille de recruter des experts de l’analytique que pour les répondants de plus petite taille.
- De nombreux participants ont indiqué qu’il est plus difficile de recruter les bonnes personnes (en établissant un juste équilibre de compétences techniques et de connaissances opérationnelles) que de les maintenir en poste.
- Bien que ce ne soit pas le cas des participants de grande taille, quelques répondants ont indiqué avoir de la difficulté à recruter des experts.

Question 7.10 – Comment vos capacités d’analytique sont-elles organisées dans l’ensemble de votre organisation?



- Au total, 71 % des répondants ont indiqué avoir eu recours à une combinaison d’experts centralisés et propres aux unités opérationnelles (spécialistes dans des groupes décentralisés), et 29 % ont déclaré avoir utilisé une fonction centralisée.
- Aucune différence importante n’a été constatée entre les divers groupes de taille.
- Les réassureurs ont tendance à être plus centralisés que les assureurs.
- Il ne semble pas y avoir beaucoup de séparation dans les fonctions d’analytique pour les participants. Il convient de souligner que même si les données ont été indiquées comme très segmentées, la fonction d’analytique ne l’est pas.

Question 7.11 – Comment classeriez-vous les aspects suivants de votre talent en analytique?



- Dans l'ensemble, une note de 3,5 a été accordée au talent en analytique pour les capacités technologiques, 3,7 pour les connaissances en analytique et 3,5 pour les connaissances opérationnelles.
- Les répondants de plus grande taille obtiennent généralement de meilleurs résultats (4,3, 4,7 et 3,3 respectivement).
- Les répondants de grande et de moyenne tailles ont accordé une note moindre aux connaissances opérationnelles qu'aux autres aspects, tandis que les répondants de petite taille ont donné une moins bonne note à leurs capacités technologiques qu'à leurs connaissances opérationnelles.

Sommaire – Analystes

Les répondants de grande taille accordent beaucoup plus d'importance à l'assignation d'ETP. Ils ont tendance à effectuer de l'analytique depuis beaucoup plus longtemps que les répondants des autres groupes de taille. Ces faits semblent indiquer que les répondants de plus grande taille y accordent une plus grande priorité que ceux des autres groupes de taille.

Dans l'ensemble, près de 50 % des ressources d'analytique prédictive sont des experts opérationnels par rapport à des architectes et statisticiens des données. Parmi ces experts opérationnels, 75 % sont des actuaires.

Les logiciels les plus populaires sont R et Python pour les travaux de modélisation, et R et Tableau pour les travaux de visualisation.

Les répondants de grande et de moyenne tailles ont donné une note moindre aux capacités en connaissances opérationnelles par rapport à d'autres aspects (capacités techniques et connaissances statistiques/analytiques), ce qui indique la nécessité d'investir dans la formation des gens au sujet de l'entreprise ou de recruter des personnes possédant des capacités d'analytique.

Les répondants de petite taille ont accordé une note plus basse à leurs capacités technologiques qu'à leurs connaissances opérationnelles, ce qui indique la nécessité d'investir dans la technologie et la formation connexe.

Identification des points forts et des faiblesses

- La plupart des participants au sondage ne disposent pas d'un dépôt de données centralisé (entièrement ou partiellement centralisé), et de nombreux participants, plus particulièrement ceux de moyenne et de petite tailles, ne considèrent pas que leurs données sont particulièrement complètes ou exactes (3 sur 5).
- La plupart des répondants ont envisagé ou prévoient d'envisager l'utilisation de l'analytique prédictive dans le cadre d'études d'expérience.
- La plupart des répondants, particulièrement ceux qui ne sont pas de grande taille, déclarent de nombreuses lacunes dans les types et la portée des applications, particulièrement en ce qui concerne le marketing, la gestion du maintien en poste, la concordance des clients et de la distribution et même la souscription accélérée/automatisée.
- La plupart des participants au sondage n'ont pas songé à utiliser l'analytique prédictive pour améliorer les opérations internes (p. ex., l'analytique de l'effectif, l'utilisation du TLN et de la GLN pour accélérer les processus et réduire les taux d'erreur).
- De nombreux répondants n'ont pas songé à un contexte de développement normalisé assorti de procédures d'essai automatisées.
- Outre les répondants de grande taille et quelques répondants de moyenne taille, l'analytique n'a pas été jugée prioritaire (de façon générale, 3 sur 5).
- De nombreux participants au sondage n'obtiennent pas un appui solide de la part des dirigeants aux fins du parrainage des initiatives d'analytique dans leur entreprise.
- Seulement la moitié des participants au sondage ont exploré des techniques d'apprentissage machine ou d'apprentissage profond.
- Les participants au sondage ont éprouvé de la difficulté à recruter et à maintenir en poste des experts de l'analytique qui répondent aux besoins opérationnels.
- La plupart des experts de l'analytique (p. ex. les statisticiens et les scientifiques des données) possèdent de solides compétences techniques, mais n'ont pas les connaissances opérationnelles pertinentes ou la capacité de communiquer des idées à des auditoires non techniques à l'échelle de l'entreprise.

Références

Domaine	Spécialisation	Nom	Titre de revue/périodique/conférence	Auteur	Date
Tous	Toutes	How Artificial Intelligence and Machine Learning Can Impact Market Design	National Bureau of Economic Research	Paul R. Milgrom, Steve Tadelis	Janv. 2018
Tous	Toutes	Predicting Emergency Room Frequent Flyers; Producing Actionable Insights from Predictive Models Built Upon Condensed Electronic Medical Records; Risk Segmentation: Application of Predictive Modeling in Life Underwriting	Predictive Analytics 2014 Call For Articles	Joseph Randazzo, J. Patrick Kinney, Sheamus Kee Parkes, Richard Xu, Minyu Cao, Scott Rushing	2014
Assurances – Tous	Toutes	Advanced Analytics for Insurance	Ernst & Young	Ernst & Young	2013
Assurances – Tous	Toutes	Anticipating Events – Using Member-level Predictive Models to Calculate IBNR Reserves	<i>The Actuary</i> Magazine	Anders Larson, Jack Leemhuis, et Michael Niemerg	Juillet 2018
Assurances – Tous	Toutes	Data Science Landscape in the Insurance Industry	ETH Zurich	Stefano Perfetti	Déc. 2017
Assurances – Tous	Toutes	Predictive Data Analytics for Claims in Insurance Industry	Infosys	Infosys	Juin 2018
Assurances – Tous	Toutes	Predictive Analytics White Paper	The Digital Insurer	Charles Nyce	2007
Assurances – Tous	Toutes	300 Years of Data Analytics in Life Insurance	Financial Services Forum, Actuaries Institute	Matt Ralph et Avanti Patki	Mai 2016
Assurances – de personnes	Toutes	Predictive Modeling Applications for Life and Annuity Pricing and Underwriting	SOA 2013 Life & Annuity Symposium	Jonathan P. Polon, Qichun (Richard) Xu	Mai 2013
Assurances – de personnes	Toutes	Predictive Analytics and Accelerated Underwriting Survey Report	SOA	Predictive Analytics and Accelerated Underwriting Subcommittee of the SOA Committee on Life Insurance Mortality and Underwriting Surveys	Mai 2017
Assurances – de personnes	Toutes	Report of the SOA Predictive Modeling Survey Subcommittee	SOA	SOA Predictive Modeling	Janv. 2012

				Survey Subcommittee	
Assurances – de personnes	Toutes	Comparing Policyholder Efficiency in Variable Annuity Lapses; Insurance Product Recommendation System; Machine Reserving: Integrating Machine Learning into Your Reserve Estimates; Variable Selection Using Parallel Random Forest for Mortality Prediction in Highly Imbalanced Data	Predictive Analytics 2016 Call For Essays	Jenny Jin, Vincent Embser, Kailan Shang, Dale Cap, Mahmoud Shehadeh, Rebecca Kokes, Guizhou Hu	2016
Assurances – de personnes	Toutes	Predictive Modeling for Life Insurance – Ways Life Insurers Can Participate in the Business Analytics Revolution	Deloitte Consulting	Mike Batty, Arun Tripathi, Alice Kroll, Chengsheng Peter Wu, David Moore, Chris Stehno, Lucas Lau, Jim Guszczka, Mitch Katcher	Avril 2010
Assurances – de personnes	Toutes	Application of Predictive Modeling Techniques in Measuring Policyholder Behavior in Variable Annuity Contracts	Insights (Towers Watson)	Guillaume Brière-Giroux, Jean-Felix Huet, Robert Spaul, Andy Staudt, David Weinsier	Avril 2010
Assurances – de personnes	Toutes	Predictive Analytics in Life Insurance	Advances in Predictive Analytic Conference	Jean-Yves Rioux, Kevin Pledge, Ian Bancroft, Eugene Wen	Déc. 2017
Assurances – de personnes	Toutes	Predictive Analytics in Life Insurance; Predictive Modeling with Prescription Histories	ACLI Annual Conference	Sam Nandi, Eric Carlson	Oct. 2017
Assurances – de personnes	Toutes	Predictive Analytics Global Survey Results – Still Room to Grow for Life & Health Insurers	Gen Re	Guizhou Hu	Oct. 2017
Assurances – de personnes	Toutes	Transforming the Life Insurance Industry Lifestyle based Analytics	Actuaries Summit	John King et Kim Cohen	Mai 2013
Assurances – de personnes	Toutes	Analytics: A Powerful Tool for the Life Insurance Industry Using Analytics to Acquire and Retain Customers	Capgemini	Capgemini	2011
Assurances – de personnes	Toutes	From Mystery to Mastery: Unlocking the Business Value of Artificial Intelligence in the Insurance Industry	Deloitte Digital	Deloitte Digital	Nov. 2017
Assurances – de personnes	Toutes	Predictive Analytics in Life Insurance – Predictive Modeling with Prescription Histories	ACLI Annual Conference	Sam Nandi et Eric Carlson	Oct. 2017

Assurances – de personnes	Toutes	Predictive Analytics and Accelerated Underwriting Follow-up Survey Report	SOA	Allen M. Klein, Roland P. Fawthrop, Gordon A. Gibbins, William M. Tilford, David N. Wylde	Mars 2018
Assurances – de personnes	Toutes	Modeling of Policyholder Behavior for Life Insurance and Annuity Products – A Survey and Literature Review	SOA, LIMRA	PwC (Jason Campbell, Michael Chan, Kate Li, Louis Lombardi, Lucian Lombardi, Marianne Purushotham)	2014
Assurances – de personnes	Toutes	Accuracy of Claims-Based Risk Scoring Models	SOA	Geof Hileman, Spenser Steele	Oct. 2016
Assurances – de personnes	Toutes	Insurers Flock to Analytics to Stretch Underwriting Limits	<i>Insurance Journal</i>	Alain Thériault	Mars 2017
Assurances – de personnes	Fraude	Apprentissage machine et sa mise en œuvre dans l'industrie de l'assurance pouvant prédire la propension à fumer du demandeur	Assemblée annuelle 2017 de l'ICA	Nitin Nayak, Swiss Re	2017
Assurances – de personnes	Soins de santé	10 Promising AI Applications in Health Care	<i>Harvard Business Review</i>	Brian Kalis, Matt Collier, Richard Fu	Mai 2018
Assurances – de personnes	Soins de santé	Predictive Modeling with Consumer Data	<i>The Actuary</i>	Ksenia Draaghtel	Nov. 2011
Assurances – de personnes	Soins de santé	Healthcare and Artificial Intelligence: Saving Lives and Costs	Datameer Blog Post	Samantha Leggat	Mai 2018
Assurances – IARD	Toutes	Predictive Modeling in P&C Insurance	Iowa Actuaries Club	Gilbert Korthals	Déc. 2016
Assurances – IARD	Toutes	2016 Property-Casualty Industry Observations and Market Trends in Analytics	BACE Fall Meeting	Ward Group (Aon)	Sept. 2016
Assurances – IARD	Toutes	Property and Casualty Insurance Predictive Analytics in SAS	SAS	Mei Najim	2017
Assurances et services bancaires	Toutes	Artificial Intelligence and Machine Learning in Financial Services	Financial Stability Board	Financial Stability Board	Nov. 2017
Autre	Toutes	Artificial Intelligence and Digital Banking	Mapa Research	Mapa Research	Nov. 2016
Autre	Toutes	Modélisation prédictive : Transformer les mégadonnées en grandes possibilités	ICA	Commission de l'ICA sur la modélisation prédictive	Juin 2018
Autre	Toutes	Predictive Analytics for Marketing – What's Possible and How it Works	Tech Emergence	Daniel Faggella	Juin 2018

Remerciement des participants et des personnes interviewées

Nous tenons à remercier les participants à ce sondage sans lesquels nous n'aurions pu obtenir une vue d'ensemble des pratiques actuelles et futures de l'industrie canadienne de l'assurance de personnes, et résumer les observations formulées dans le présent rapport. Les participants comprenaient :

- BMO Assurance
- Desjardins
- Empire Vie
- Great-West, Compagnie d'assurance-vie
- Industrielle Alliance
- Ivari
- Financière Manuvie
- Munich Re
- Optimum Re
- Partner Re
- RBC Assurances
- RGA
- SCOR
- SSQ Groupe financier
- Financière Sun Life

Nous remercions également les personnes qui ont accepté d'être interviewées pour la section de recherche :

- Denis Repin
- Shanil Ebrahim
- Peter Wu

Annexe A : Questionnaire du sondage

Profil

N°	Question	Réponse
1	Quelle entreprise représentez-vous?	Fournie automatiquement par Deloitte
1.1	Quel est votre nom?	
1.2	Quel est votre titre actuel?	
1.3	À quel service êtes-vous actuellement rattaché(e)?	

Généralités

2	Question	Réponse
2.1	Veuillez préciser les principaux problèmes que pose l'utilisation efficace de l'analytique prédictive et la façon dont vous les avez abordés.	
2.2	Que tente d'apprendre votre entreprise au moyen de l'analytique prédictive qu'elle ne peut apprendre par des méthodes traditionnelles?	

Données

3	Question	Réponse
3.1	Qu'est-ce qui décrit le mieux la centralisation de vos données? S'il s'agit de combinaisons, quelle est la proportion approximative de chacune?	<ul style="list-style-type: none"> a. Les données sont stockées/accessibles séparément pour chaque produit et fonction. b. Les données sont stockées/accessibles séparément au niveau fonctionnel (études d'expérience, tarification, évaluation, etc.). c. Les données sont stockées/accessibles séparément par secteur d'activité (c'est-à-dire pour toutes les fonctions et produits). d. Combinaison (c) et (e) – stockage/accès aux données par secteur d'activité et pour l'ensemble de l'entreprise. e. Toutes les données de l'organisation sont stockées/accessibles à un seul point.
3.2	Sur une échelle de 1 à 5, comment les utilisateurs évalueraient-ils l'exhaustivité et l'exactitude de vos données?	<p>1 = plus bas</p> <p>5 = plus élevé</p>

<p>3.3</p>	<p>Indiquez un pourcentage approximatif de vos données (à l'exception des assurances IARD, si les deux types d'assurances sont souscrites) dans chacune des catégories suivantes :</p>	<ul style="list-style-type: none"> a. Données non structurées (voix, image, documents numérisés). b. Texte de format libre. c. Code nécessitant une légende pour le comprendre d. Données scalaires/Valeurs ne nécessitant pas de légende. e. Vecteurs, tableaux, cubes.
<p>3.4</p>	<p>Comment les données sont-elles accessibles à l'échelle de l'organisation? (Choisir toutes les réponses pertinentes.)</p>	<ul style="list-style-type: none"> a. Accès indirect sur demande de personnes spécifiques (p. ex., TI) qui fournissent les données après plus d'un jour ouvrable. b. Accès indirect par demande à des personnes spécifiques (p. ex., TI) qui fournissent les données en l'espace d'un jour ouvrable. c. Accès direct à plusieurs bases de données au moyen d'un langage de recherche (p. ex., Serveur SQL, MySQL, etc.). d. Accès direct à une ou deux bases de données au moyen d'un ou deux langages de recherche (p. ex., Serveur SQL, MySQL, etc.). e. Accès direct par recherche simple à l'aide d'accès d'avant-plan convivial.
<p>3.5</p>	<p>Quels types de données de tiers votre organisation utilise-t-elle actuellement pour l'analytique? (Choisir toutes les réponses pertinentes.)</p>	<ul style="list-style-type: none"> a. Données démographiques. b. Données géographiques. c. Données sur les sinistres et données médicales. d. Données financières. e. Données sur le crédit. f. Données sur le style de vie. g. Données sur l'attitude/l'humeur. h. Données comportementales. i. Données économiques.
<p>3.6</p>	<p>Avez-vous commencé à recueillir des données au moyen de nouvelles technologies (p. ex., Fitbit) pour accroître l'utilisation des sources traditionnelles de données sur les applications/la souscription/la réassurance?</p>	<ul style="list-style-type: none"> a. Oui, nous avons commencé à recueillir les données et nous les utilisons maintenant. b. Oui, nous avons commencé à recueillir les données, mais nous ne les utilisons pas encore. c. Non, nous avons déjà examiné la question et nous avons décidé de ne pas recueillir/utiliser ce type de données.

		<ul style="list-style-type: none"> d. Non, mais nous étudions présentement la question. e. Non, nous n'étudions pas présentement la question. f. Non, nous ne planifions pas étudier la question.
3.7	Comment stockez-vous vos données internes? (Cochez toutes les réponses pertinentes.)	<ul style="list-style-type: none"> a. Système interne exclusif. b. Feuilles de calcul Excel. c. Fichiers plats (CSV, fichiers texte). d. MS Access. e. Serveur SQL. f. Hadoop (ou logiciel semblable). g. Autre (préciser).
3.8	Comment stockez-vous vos données externes? (Cochez toutes les réponses pertinentes)	<ul style="list-style-type: none"> a. Système interne exclusif. b. Feuilles de calcul Excel. c. Fichiers plats (CSV, fichiers texte). d. MS Access. e. Serveur SQL. f. Hadoop (ou logiciel semblable). g. Autre (préciser).

Entreprise

4	Question	Réponse
4.1	Sélectionnez les éléments et politiques de gouvernance qui s'appliquent aux données utilisées dans l'analytique prédictive pour votre organisation. (Cochez toutes les réponses pertinentes.)	<ul style="list-style-type: none"> a. Confidentialité des données. b. Sécurité des données. c. Privilèges de mise à jour des données. d. Politiques d'évaluation des impacts des changements apportés aux données. e. Normes d'exactitude/de qualité des données f. Normalisation des données. g. Dictionnaires de données. h. Contrôles relatifs à l'ETC (« Extraction, transformation et chargement »). i. Autre (préciser).
4.2	Sélectionnez les aspects et politiques de gouvernance qui influent sur les efforts de modélisation de votre organisation. (Cochez toutes les réponses pertinentes.)	<ul style="list-style-type: none"> j. Validation de modèles. k. Examen des codes. l. Contrôle des versions. m. Identification des paramètres d'évaluation du rendement des modèles. n. Autre (préciser).
4.3	Sélectionnez les aspects et politiques de gouvernance liés aux logiciels et à la technologie utilisés dans les applications d'analytique de votre organisation. (Cochez toutes les réponses pertinentes.)	<ul style="list-style-type: none"> a. Il n'existe aucune politique. b. Liste de logiciels approuvés/préférés/désapprouvés. c. Politique de mise à jour des logiciels. d. Autre (préciser).
4.4	Si la fonction d'analytique n'est pas entièrement centralisée, veuillez évaluer la force de la coordination des compétences et des ressources au sein de l'organisation.	<p>1 = extrêmement faible. 5 = extrêmement forte.</p>

Leadership

5	Question	Réponse
5.1	Votre organisation compte-t-elle dans ses rangs un cadre supérieur chargé des données? Dans l'affirmative, quel est son titre?	
5.1a	Si aucun cadre supérieur n'est chargé des données, dans quelle mesure la direction comprend-elle les données (qualité, structure, mesures de protection, etc.)?	1 = La direction comprend très peu les données. 5 = La direction comprend très bien les données.
5.2	Votre organisation compte-t-elle dans ses rangs un cadre supérieur chargé de l'analytique prédictive? Dans l'affirmative, quel est son titre?	
5.2a	Si aucun cadre supérieur n'est chargé de l'analytique prédictive, dans quelle mesure la direction comprend-elle ses possibilités de valeur ajoutée et les initiatives connexes des concurrents?	1 = La direction comprend très peu l'analytique prédictive. 5 = La direction comprend très bien l'analytique prédictive.
5.3	Qui prend la décision finale sur la priorisation des initiatives d'analytique prédictive?	a. La direction (le conseil d'administration, le chef de la direction, le premier vice-président). b. Les cadres de la Série des C. c. Les cadres chargés de l'analytique en deçà de la Série des C (p. ex., les vice-présidents). d. Les cadres non spécialisés en analytique en deçà de la Série des C (p. ex., les vice-présidents). e. Autre.
5.4	Quelle est la priorité de l'analytique prédictive dans le cadre des projets de votre organisation?	1 = Faible priorité. 5 = Investissement le plus important pour l'avenir.

5.4a	<p>Si l'analytique prédictive n'est pas la priorité absolue (note de 5 à la question précédente), quelles en sont les principales raisons? (Cochez toutes les réponses pertinentes.)</p>	<ul style="list-style-type: none"> a. Manque d'infrastructure ou de capacités en informatique. b. Faible qualité des données internes. c. Accès difficile aux données internes. d. Problèmes relatifs à l'obtention et à l'agrégation des données à partir de nombreuses sources de données. e. Difficulté à trouver et à recueillir des données externes pertinentes. f. Problèmes de protection des données de tiers. g. Manque de talent et d'expertise à l'interne. h. Manque de soutien de la direction. i. Autres (préciser).
-------------	--	--

Cibles

6	Question	Réponse
6.1	<p>Pour quelles applications actuelles et passées utilisez-vous l'analytique? Les « applications » sont définies comme des cas où un modèle prédictif (ou ML, IA, etc.) a été étalonné à des fins de prédiction, pour identifier les facteurs/variables explicatives, etc.</p> <p>Comment évaluez-vous chacune de ces applications en fonction des efforts perçus pour la mise en œuvre et de la valeur perçue pour l'entreprise?</p>	<p>Liste des applications</p> <p>Effort (faible, modéré et énorme) et valeur cotés :</p> <p>F-</p> <p>F</p> <p>F+</p> <p>M-</p> <p>M</p> <p>M+</p> <p>É-</p> <p>É</p> <p>É+</p>

6.2	<p>Pour quelles applications prévoyez-vous utiliser l'analytique au cours de la prochaine année?</p> <p>Comment évaluez-vous chacune de ces applications en fonction des efforts perçus pour la mise en œuvre et la valeur perçue pour l'entreprise?</p>	<p>Liste des applications</p> <p>Effort (faible, modéré et énorme) et valeur cotés :</p> <p>F-</p> <p>F</p> <p>F+</p> <p>M-</p> <p>M</p> <p>M+</p> <p>É-</p> <p>É</p> <p>É+</p>
6.3	<p>Environ combien de projets d'analytique avez-vous tentés au cours de la dernière année?</p>	
6.4	<p>Environ combien de nouvelles applications d'analytique qui n'ont pas été exécutées auparavant devraient être traitées au cours de la prochaine année (p. ex., produire l'analytique de la déchéance pour la 1^{re} fois)?</p>	

Analystes

7	Question	Réponse
7.1	Quelles techniques sont actuellement utilisées?	<ul style="list-style-type: none"> a. Statistiques sommaires (moyennes, médianes, distributions, corrélations). b. Régression inférentielle (modèle étalonné à l'aide de l'ensemble des données). c. Modélisation prédictive avec régression/MLG. d. Modélisation prédictive à l'aide de modèles non linéaires (arbres de décision, forêt aléatoire, arbres boostés, etc.). e. Apprentissage machine/réseaux de neurones/apprentissage profond.
7.2	Combien d'équivalents temps plein (ETP) travaillent à l'analytique?	
7.3	Quelles sont les proportions approximatives d'ETP dans chacune des catégories suivantes? Tenir compte des ressources empruntées du siège social, d'autres unités opérationnelles, etc.	<ul style="list-style-type: none"> a. Experts opérationnels – Actuariers. b. Experts opérationnels – Non-actuariers. c. Architectes/ingénieurs des données. d. Statisticiens/experts de l'analytique. e. Informaticiens/TI. f. Autres (préciser).
7.4	Dans chaque projet d'analytique prédictive, dans quelle mesure intégrez-vous les compétences des experts opérationnels (actuariat et non-actuariat), des statisticiens/experts en analytique et du personnel de la TI?	<ul style="list-style-type: none"> a. Pas du tout (0 %). b. À l'occasion (0-30 %). c. Souvent (30-70 %). d. La plupart du temps (70-100 %). e. Toujours (100 %).
7.5	Quel logiciel utilisez-vous pour la modélisation de l'analytique prédictive? (Choisir toutes les réponses pertinentes.)	<ul style="list-style-type: none"> a. R. b. Python. c. SAS. d. Autres (préciser).
7.6	Quel logiciel utilisez-vous pour l'analyse et la visualisation des données exploratoires? (Choisir toutes les réponses pertinentes.)	<ul style="list-style-type: none"> a. Tableau. b. Qlikview. c. Power BI. d. R. e. Python. f. Autres (préciser).
7.7	Depuis combien de temps votre entreprise effectue-t-elle la modélisation prédictive?	

7.8	Lorsque vous exécutez une nouvelle application d'analytique qui n'a pas été effectuée antérieurement, dans quelle mesure faites-vous appel à des experts-conseils externes?	<ul style="list-style-type: none"> a. Pas du tout (0 %). b. À l'occasion (0-30 %). c. Souvent (30-70 %). d. La plupart du temps (70-100 %). e. Toujours (100 %).
7.9	Dans quelle mesure est-il difficile de recruter et de maintenir en poste des experts de l'analytique pour des postes à l'interne?	<p>1 = Extrêmement difficile.</p> <p>5 = Facile.</p> <p>S.O. – Aucun poste à l'interne.</p>
7.10	Comment vos capacités d'analytique sont-elles organisées dans l'ensemble de votre organisation?	<ul style="list-style-type: none"> a. Séparément pour chaque produit et fonction. b. Séparément au niveau fonctionnel (études d'expérience, tarification, évaluation, etc.). c. Niveau de secteur d'activité (toutes les fonctions et tous les produits). d. Combinaison d'experts centralisés et des unités fonctionnelles. e. Fonction d'analytique centralisée.
7.11	<p>Comment classeriez-vous les aspects suivants de votre talent en analytique? :</p> <ul style="list-style-type: none"> a. Capacités technologiques; b. Connaissances des statistiques/de l'analytique; c. Connaissances opérationnelles. 	<p>1 = Faibles, jusqu'à 5 = Élevées.</p> <p>Je ne sais pas.</p> <p>S.O. – Aucun ETP affecté à l'analytique.</p>

Annexe B : Définitions

Aux fins du présent rapport, nous utiliserons les définitions suivantes :

Apprentissage profond : Dans le contexte du présent sondage, l'apprentissage profond correspond à l'amélioration d'un réseau de neurones artificielles au moyen de plusieurs couches de nœuds cachées.

Données externes : Données qui n'ont pas été saisies au départ par votre organisation et qui sont habituellement recueillies ou achetées d'un tiers. Le présent sondage utilise indifféremment les expressions « données externes » et « données de tiers ».

Texte de format libre : Données stockées sans restriction sur un formulaire. Par exemple, le prénom et le nom sont traités comme texte de format libre, car il n'existe pas d'ensemble prédéfini de valeurs.

Avant-plan : Point d'accès à une base de données (celle-ci parfois désignée « arrière-plan ») qui n'exige pas de connaissances en programmation et qui est habituellement convivial. Ainsi, divers membres du personnel de l'entreprise peuvent avoir accès aux mêmes données sans devoir utiliser un langage d'interrogation d'arrière-plan.

Modèle linéaire généralisé (MLG) : Le vaste groupe de modèles de régression linéaire qui permet à des variables dépendantes d'avoir une distribution normale des erreurs, y compris la méthode des moindres carrés ordinaire (MCO), la régression logistique (logit), la régression probit, la régression de Poisson, et bien d'autres.

Statistiques inférentielles : Utilisation de données historiques pour produire des inférences générales au sujet d'événements passés. Cette méthode se distingue de l'analytique prédictive du fait que l'on n'utilise pas les résultats pour prédire un événement futur. On utilise simplement les résultats de l'analytique pour mieux comprendre un événement passé.

Réseau de neurones : Dans le contexte du présent sondage, un réseau de neurones est un modèle théoriquement inspiré par le réseau de neurones biologiques du cerveau d'un animal.

Analytique prédictive : Application de techniques statistiques pour étalonner un modèle à l'aide de données historiques afin d'effectuer des prédictions au sujet de l'avenir ou d'événements par ailleurs inconnus. Dans le présent sondage, nous utilisons indifféremment les expressions « analytique prédictive », « modélisation prédictive » et « apprentissage machine ». Dans certains cas, le terme « analytique » a été utilisé par référence à la pratique élargie pour désigner la modélisation de données historiques, qui renferme des éléments d'« analytique prédictive ».

Données non structurées : Information qui n'a pas de modèle, de structure ou de lien définis avec les autres données de la base de données. Les données non structurées ne seraient pas stockées dans une base de données relationnelle où l'information peut être liée au moyen de clés, par exemple, des enregistrements vocaux, des images et des documents numérisés.