

LIVRE BLANC

.....



Du Big Data aux modèles prédictifs : où en sont les entreprises françaises ?

Sponsorisé par : Dataiku

Thierry Hamelin
January 2019



data
iku



IDC

ANALYZE THE FUTURE

OPINION IDC

A l'heure où les entreprises s'engagent massivement dans un processus de transformation numérique, l'analyse des données apparaît comme un facteur clé de réussite en la matière. Mais pour faire du traitement des données une véritable source de compétitivité, le recours au Big Data et aux modèles prédictifs basés sur l'intelligence artificielle (IA), apparaissent comme des prérequis pour optimiser les processus internes des organisations et générer de nouveaux revenus basés sur le développement de nouvelles offres.

Les deux prochaines années seront donc charnières pour les entreprises qui souhaitent renforcer leurs avantages concurrentiels. S'agissant plus particulièrement des modèles prédictifs, qui apparaissent déjà comme une priorité pour 26% des directions métiers, la maîtrise des coûts et des délais, ainsi que la capacité à faire travailler étroitement les métiers avec les spécialistes de la donnée seront des éléments importants que les entreprises devront prendre en compte dans le choix de leurs plates-formes de développement et leur organisation. C'est ce que montrent les résultats d'une enquête IDC France présentés dans ce document.



EXECUTIVE SUMMARY

Près de neuf directions métiers sur dix sont conscientes du rôle capital de la donnée. Ainsi, 55% considèrent l'analyse de leurs données comme importante, et 33% la qualifie de très importante. Toutefois, de nombreux challenges restent à relever en la matière.

En premier lieu, 41% des directions métiers déplorent un manque de stratégie claire de la part de leur entreprise en matière d'analyse des données. Cela souligne la nécessité d'une implication croissante du top management pour donner une orientation et éviter une approche en silo, conduisant à des initiatives déconnectées d'un département à l'autre.

Ensuite, les directions métiers attendent de leur entreprise des améliorations relatives à l'organisation autour du traitement des données. Les trois domaines concernés sont une meilleure prise en compte de leurs attentes, la capacité à repenser l'organisation pour en améliorer l'efficacité et la nécessité de disposer des compétences de spécialistes de la donnée. Preuve en est, dans 63% des organisations de 500 à 999 employés, on ne trouve ni Directeur de la donnée, ni responsable Business Intelligence, ni data scientist, ni data analyst, ni data miner. Cette part tombe à 37% dans les entreprises de 1 000 employés ou plus.

Par ailleurs, 46% des directions métiers se déclarent insatisfaites des délais nécessaires pour réaliser de nouvelles applications s'appuyant sur du traitement de données. En conséquence, 35% des responsables métiers sont demandeurs de plus d'échanges avec les équipes techniques et 34% souhaitent faire gagner leurs équipes en autonomie, en leur permettant d'effectuer des analyses par elles-mêmes, sans passer par les techniciens de la donnée.

Dans le même temps, les priorités des directions métiers font apparaître des besoins qui trouvent leurs réponses dans les apports du Big Data, à savoir la capacité à traiter très rapidement de gros volumes de données structurées et non-structurées. Enfin, construire des modèles prédictifs basés sur l'intelligence artificielle est une priorité d'ores et déjà partagée par 26% des métiers.

Face à tous ces enjeux, beaucoup reste à faire : 48% des entreprises étudiées recourent au Big Data et seulement 9% aux modèles prédictifs. Avec le besoin croissant des métiers de disposer d'analyses avancées pour anticiper les tendances à venir, la maîtrise des délais et des coûts vont devenir des points capitaux. Cela souligne l'importance pour les entreprises d'opter pour une plate-forme de développement permettant de faire étroitement travailler les équipes techniques avec les métiers pour bien s'aligner sur leurs besoins tout au long des projets et ainsi gagner en réactivité et efficacité. Un atout reconnu par les entreprises utilisatrices de la plate-forme Dataiku de développement de modèles prédictifs qu'IDC a eu l'occasion d'interroger, au même titre que ses fonctionnalités d'aide à la préparation des données.

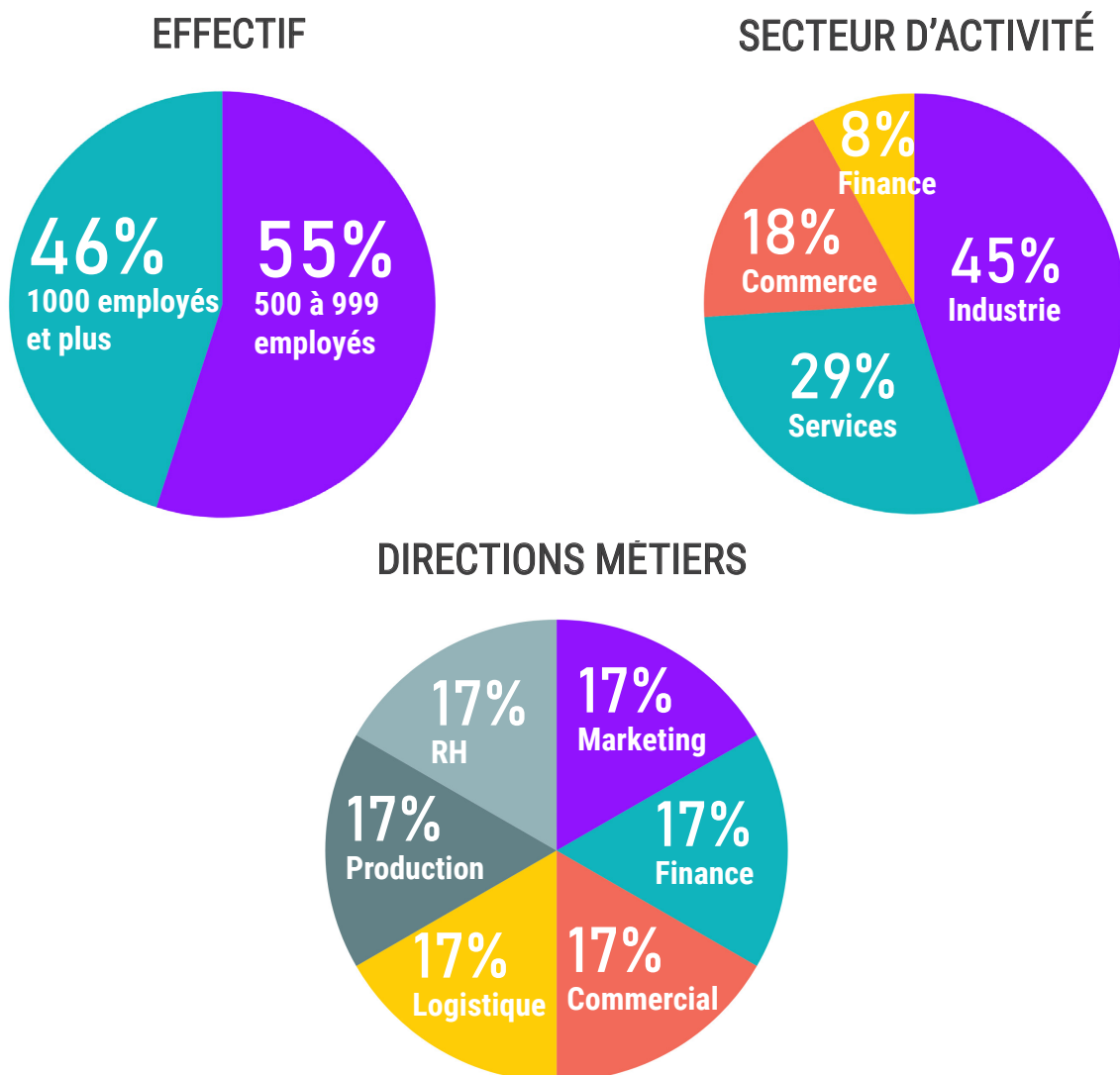
MÉTHODOLOGIE

Ce document repose sur une enquête téléphonique menée par IDC France entre octobre et novembre 2018 auprès de 150 directions métiers d'entreprises françaises du secteur privé, dont l'effectif est d'au moins 500 personnes.

Les résultats ont été redressés conformément aux statistiques de l'Insee, en termes d'effectif salarié et de secteur d'activité.

Cette enquête a été complétée par des entretiens approfondis conduits par les analystes d'IDC avec trois entreprises utilisatrices de la plate-forme de développement de modèles prédictifs de Dataiku.

Graphique 1 :
Structure de l'échantillon des entreprises étudiées



Source IDC France, enquête Analyse des données, novembre 2018, n=150

ANALYSE DES DONNÉES : LES MÉTIERS ATTENDENT DE NOMBREUX CHANGEMENTS DE LA PART DE LEUR ENTREPRISE

Comme le montre le graphique 2, plus de huit directions métiers sur dix sont demandeuses d'améliorations quant à l'approche de leur entreprise en matière d'analyse des données.

A l'heure où le traitement des données offre de nouvelles opportunités grâce au Big Data et à l'intelligence artificielle (IA), 41% des directions métiers déplorent un manque de stratégie claire de la part de leur entreprise. Cela souligne la nécessité d'une implication croissante du top management pour donner une orientation en matière d'analyse des données et éviter une approche en silo, conduisant à des initiatives déconnectées d'un département à l'autre. Dans le même esprit de désenclavement, quatre directions métiers sur dix attendent davantage de communications sur les différents projets menés à l'échelle de l'entreprise.

38% des responsables métiers souhaitent également que la direction informatique s'investisse davantage dans une veille sur les outils et les technologies afin de les aider dans leur réflexion autour de l'analyse des données.

Viennent ensuite des axes d'amélioration relatifs à l'organisation des entreprises autour du traitement des données. Les trois domaines concernés sont une meilleure prise en compte des attentes des métiers, la capacité à repenser l'organisation pour en améliorer l'efficacité et la nécessité de disposer de compétences de spécialistes de la données (Chief Data Officers, data scientists, data analysts, etc).



41%

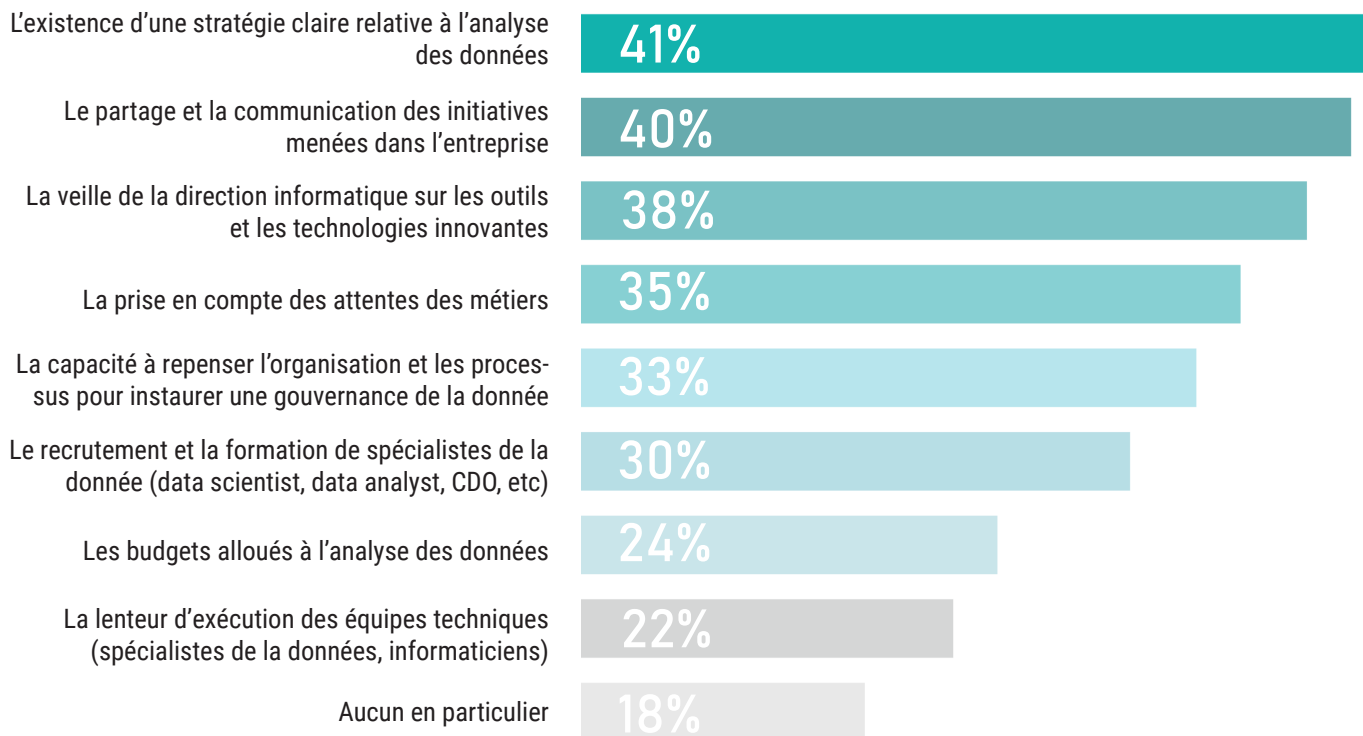
des directions métiers déplorent un manque de stratégie claire de la part de leur entreprise s'agissant de l'analyse des données

38%

des responsables métier souhaitent une implication plus importante de la direction informatique

Graphique 2 : Attentes des métiers en matière d'analyse des données

Q. Dans quels domaines trouvez-vous que l'approche de votre entreprise en matière d'analyse des données est perfectible ?



Source IDC France, enquête Analyse des données, novembre 2018, n=150

S'agissant du degré de satisfaction des métiers vis-à-vis de la conduite de leurs projets d'analyse des données, c'est surtout le délai pour développer de nouvelles applications qui pose problème (voir le graphique 3). Souvent confrontées à des contraintes fortes de time-to-market, 46% des directions se déclarent insatisfaites du temps nécessaire pour créer une application et 10% se disent même très insatisfaites.

Par effet induit, 29% des responsables métiers interrogés sont insatisfaits des coûts de développement, et même près d'un tiers d'entre eux se montrent très insatisfaits.

Avec le besoin croissant des métiers de disposer d'analyses avancées pour anticiper les tendances à venir, les délais et les coûts vont devenir des points capitaux. Cela souligne l'importance pour les entreprises d'opter pour une plate-forme de développement permettant de travailler étroitement avec les métiers pour bien s'aligner sur leurs besoins tout au long des projets et ainsi gagner en réactivité et efficacité.

46%

des directions sont insatisfaites du temps imparti pour créer une application

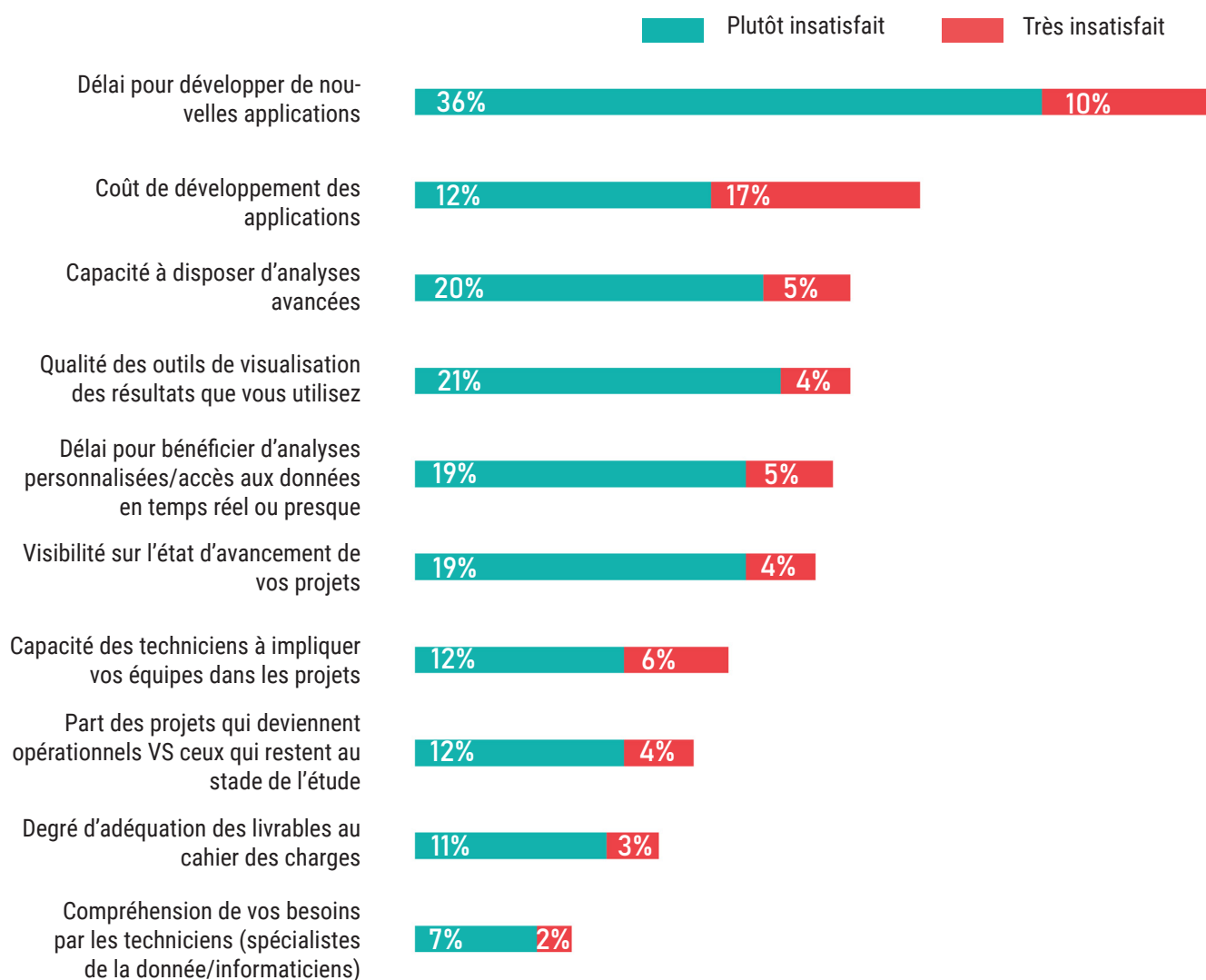
29%

des responsables métier sont insatisfaits des coûts de développement



Graphique 3 : Degré de satisfaction des métiers vis-à-vis de la conduite des projets d'analyse des données

Q. Concernant vos projets d'analyse des données, quel est votre degré de satisfaction dans les domaines suivants ?



Source IDC France, enquête Analyse des données, novembre 2018, n=150

LA QUALITÉ DES DONNÉES, LA VITESSE DE TRAITEMENT ET LES RELATIONS AVEC L'IT ARRIVENT EN TÊTE DES CHALLENGES À RELEVER CÔTÉ MÉTIERS

Comme le révèle le graphique 4, l'amélioration de la qualité des données arrive en tête des priorités des métiers. En effet, tous les efforts menés pour analyser les données sont réduits à néant si ces dernières sont incomplètes, voire erronées. Il faut donc procéder à un travail de préparation des données avant de les traiter. Cette phase étant souvent très chronophage en raison des volumes de données concernés, des logiciels automatisant ce processus ont fait leur apparition.

Cité par près d'un responsable métier sur deux, le traitement des données en temps réel ou quasi temps réel se hisse en deuxième position des priorités. La capacité à analyser des données "fraîches" représente en effet une vraie valeur ajoutée dans le pilotage des activités.

Ensuite, plusieurs priorités arrivent au coude à coude, parmi lesquelles les relations avec les techniciens de la donnée. 35% des directions métiers sont demandeuses de plus d'échanges avec les équipes techniques. Une situation qui tient largement au fait que le délai de réalisation des projets est souvent jugé trop long par les métiers, comme en atteste le graphique 3. Par ailleurs, 34% des responsables métiers souhaitent faire gagner leurs équipes en autonomie, en leur permettant d'effectuer des analyses par elles-mêmes, sans passer par les techniciens de la donnée. Cela traduit là encore la volonté des métiers de gagner en rapidité.

Au final, la lecture des priorités fait apparaître des besoins qui trouvent leurs réponses dans les apports du Big Data, à savoir la capacité à traiter rapidement de gros volumes de données structurées et non-structurées comme du texte. Enfin, construire des modèles prédictifs est une priorité d'ores et déjà partagée par 26% des directions métiers.



35%

des directions métiers sont demandeuses de plus d'échanges avec les équipes techniques

34%

des responsables métiers souhaitent faire gagner leurs équipes en autonomie

26%

des directions métiers font des modèles prédictifs une priorité

Graphique 4 : Priorités des métiers en matière d'analyse des données

Q. Quelles sont vos priorités en matière d'analyse des données pour les 2 ans à venir ?



Source IDC France, enquête Analyse des données, novembre 2018, n=150

REPENSER L'ORGANISATION APPARAÎT COMME UNE NÉCESSITÉ

Comme le révèle le graphique 5, dans la moitié des entreprises étudiées, les métiers passent par l'informatique, car il n'existe pas de service dédié à l'analyse des données en tant que tel. Et lorsqu'une équipe de spécialistes de la donnée existe, elle est en charge de répondre aux demandes de tous les métiers.

Au final, seules 11% des directions métiers disposent de spécialistes de la donnée qui leur sont dédiés et 12% sont totalement autonomes, puisqu'elles embarquent au sein de leurs équipes des compétences informatiques et statistiques.

12%

seulement des directions métiers sont totalement autonomes en matière de traitements de données

Graphique 5 : **Organisation des métiers en matière d'analyse des données**

Q. Comment êtes-vous organisé pour vos projets d'analyse des données ?

Vous passez par le service IT car il n'existe pas de service dédié à l'analyse de la donnée en tant que tel

50%

Vous passez par une équipe de spécialistes de la donnée en charge de répondre aux demandes de tous les métiers (marketing, finance, RH, etc)

28%

Vous passez par un ou plusieurs spécialistes de la donnée dédié à votre département

11%

Votre département est totalement autonome vis-à-vis de l'IT / il dispose de toutes les compétences techniques (informaticiens, spécialistes de la donnée) en interne pour mener ses projets de bout en bout

12%

Source IDC France, enquête Analyse des données, novembre 2018, n=150

A l'absence d'organisation optimisée dans les entreprises pour mener les projets d'analyse des données s'ajoute souvent une carence de compétences. Ainsi, moins d'un tiers des entreprises étudiée disposent d'un Directeur de la donnée, d'un responsable Business Intelligence ou encore d'un data scientist. Ce sont surtout les plus petites entreprises qui sont les moins bien loties. Dans 63% des organisations de 500 à 999 employés, aucun des profils listés dans le graphique 6 n'existe. Cette part tombe à 37% dans les entreprises de 1 000 employés ou plus.

Graphique 6 : **Existence de spécialistes de la donnée dans les entreprises**

Q. Parmi les profils suivants, quels sont ceux que l'on trouve dans votre entreprise ?

Un Chief Data Officer (ou Directeur de la donnée)

30%

Un Responsable Business Intelligence

31%

Un ou des data scientist(s)

29%

Un ou des data analyst(s)

41%

Un ou des data miner(s)

14%

Source IDC France, enquête Analyse des données, novembre 2018, n=150



BIG DATA ET INTELLIGENCE ARTIFICIELLE : BEAUCOUP RESTE À FAIRE

Si les priorités des métiers (graphique 4) font apparaître des besoins en termes de rapidité de traitement, de volume et de variété des données à analyser, qui trouvent leurs réponses dans les apports du Big Data, seules 48% des entreprises y recourent.

Ce sont les organisations de 1 000 employés et plus qui sont les plus équipées avec un taux de pénétration du Big Data de 59%, contre 39% dans les entreprises de 500 à 999 salariés. Ces dernières vont donc devoir réagir si elles ne veulent pas se voir distancer par les plus grandes structures. D'autant qu'elles partagent le même intérêt pour les modèles prédictifs basés sur le machine learning qui nécessitent de traiter de gros volumes de données.

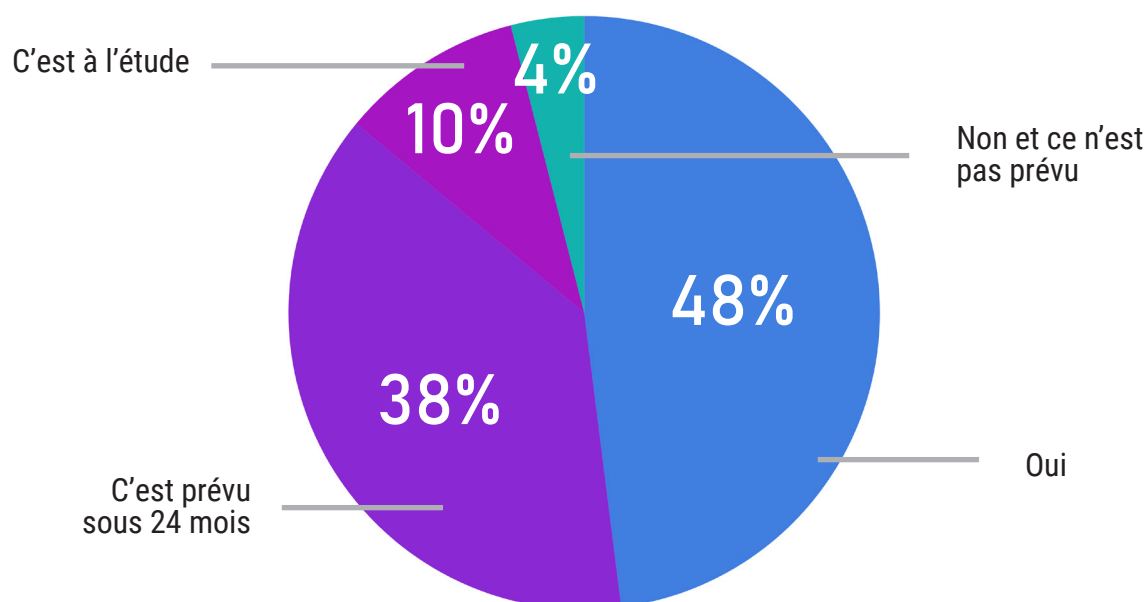


48%

des entreprises recourent au Big Data

Graphique 7 :
Recours au Big Data dans les entreprises

Q. Est-ce que votre entreprise recourt au Big Data ?



Source IDC France, enquête Analyse des données, novembre 2018, n=150

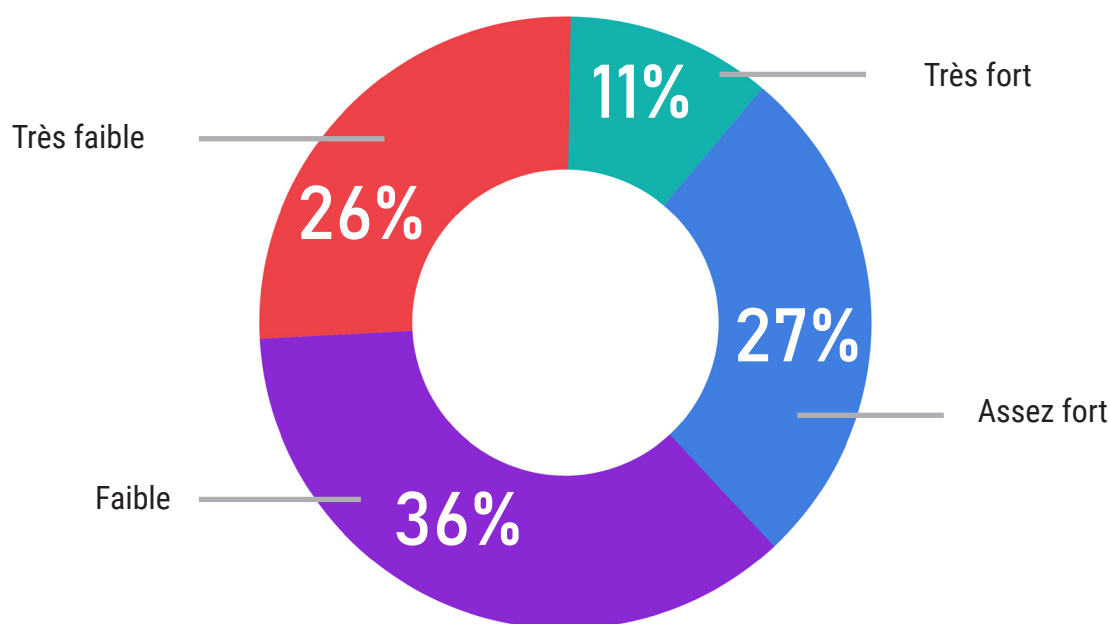
Au regard du potentiel qu'offrent les modèles prédictifs basés sur l'intelligence artificielle - et plus particulièrement le machine learning qui permet à l'algorithme de s'améliorer au cours du temps grâce à l'apport de nouvelles données - l'appétence qu'ils suscitent est relativement limité. En effet, 36% des entreprises n'y voient qu'un faible intérêt, voire un très faible intérêt pour 26% d'entre elles.



Les directions métiers sont donc peu incitées par leur entreprise à se lancer. Plutôt qu'un désintéressement, cette situation traduit surtout une méconnaissance de l'apport de ces technologies encore récentes. En effet, les initiatives menées par les organisations les plus avancées, décrites dans la suite de ce document, attestent des bénéfices que les directions métiers peuvent en retirer.

Graphique 8 : Intérêt accordé aux modèles prédictifs

Q. Quel intérêt accorde votre entreprise aux modèles prédictifs s'appuyant sur l'IA / le Machine Learning ?



Source IDC France, enquête Analyse des données, novembre 2018, n=150

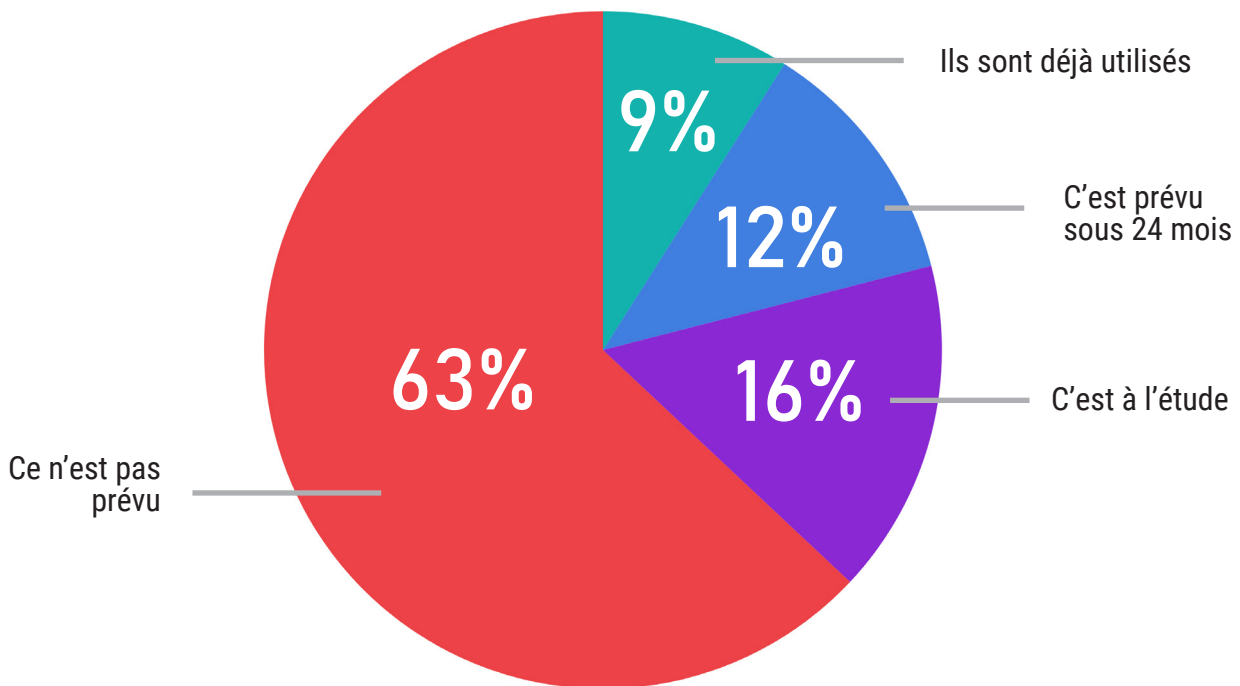
RECOURS AUX MODÈLES PRÉDICTIFS : QUELQUES ENTREPRISES SONT EN TRAIN DE CREUSER L'ÉCART

Comme le montre le graphique 9, le recours aux modèles prédictifs basés sur l'IA au sein des métiers reflète le degré d'intérêt des entreprises pour ces technologies. Ainsi, un peu moins d'une direction métier sur dix y recourt déjà, 28% envisagent ou ont d'ores et déjà prévu de les utiliser et enfin, 63% restent pour l'instant sans projet. Les entreprises d'au moins 1 000 employés sont les plus proactives. Elles sont 15% à utiliser les modèles prédictifs, contre 4% dans les organisations de 500 à 999 salariés.

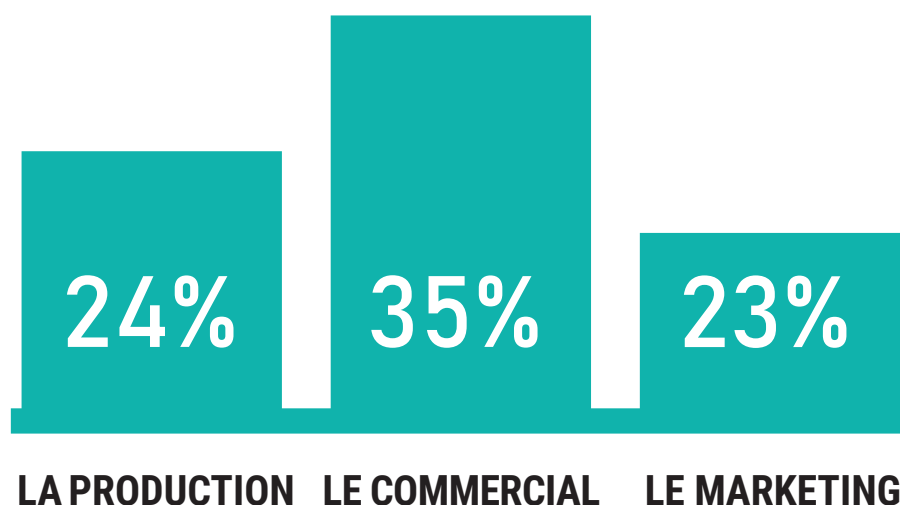
15%
des entreprises d'au moins
1000 employés utilisent déjà
des modèles prédictifs

Graphique 9 :
Recours aux modèles prédictifs

Q. Où en est votre département de l'utilisation de modèles prédictifs basés sur l'Intelligence Artificielle ?



Au final, les trois directions métier les plus enclines à recourir aux modèles prédictifs, que ce soit en termes de réalisations actuelles ou de projets à venir, sont :



Arrivent ensuite la logistique (22%), la finance (19%) et enfin les ressources humaines (5%). Par ailleurs, le tableau ci-dessous atteste de l'étendue des domaines d'application des modèles prédictifs au sein des métiers.

Applications associées aux modèles prédictifs citées par les directions métiers :

Commercial

- Identifier les leads
- Identifier les opportunités parmi les clients existants
- Actions à mener pour concrétiser une vente

Production

- Faire de la maintenance prédictive
- Accélérer les tests qualité
- Optimiser la productivité

Marketing

- Anticiper la consommation
- Anticiper le churn
- Optimiser le cross / l'up-selling
- Optimiser les opérations marketing

Finance

- Anticiper les risques de défaillance
- Identifier la fraude, de défaut de paiement
- Mieux estimer les revenus à venir

Logistique

- Prévenir les ruptures d'approvisionnement
- Anticiper les risques de défaillance des fournisseurs
- Optimiser les coûts d'approvisionnement

Ressources humaines

- Prévenir les risques de démission
- Améliorer le recrutement
- Améliorer la mobilité interne
- Améliorer la progression de carrière

Source IDC France, enquête Analyse des données, novembre 2018, n=55
(directions métiers chez lesquelles les modèles prédictifs sont utilisés, prévus ou à l'étude)

RETOURS D'EXPÉRIENCE D'UTILISATEURS DE LA SOLUTION DE DATAIKU

En complément de l'enquête menée auprès de 150 directions métiers, les analystes d'IDC ont recueilli les témoignages de trois entreprises utilisatrices de la plate-forme de Data Science centralisée de Dataiku au travers d'entretiens approfondis menés avec :

Un Associé de Roland Berger, société d'études d'origine allemande spécialisée dans le Conseil en Stratégie de 2 400 collaborateurs dans le monde

Le Responsable innovation de la BRED, banque filiale du groupe BPCE, comptant 5 500 collaborateurs dans le monde dont 4 000 en France

Le Responsable datlab de GRTgaz, transporteur exclusif du gaz en France, dont l'effectif est de 3 000 collaborateurs

Un équipement dicté par une explosion du volume de données

Au fur et à mesure des années, le volume de données recueilli par les trois entreprises interrogées a considérablement augmenté. Les outils utilisés jusque-là par les collaborateurs se sont révélés de plus en plus limités pour traiter cette masse d'informations. Les techniciens de la donnée de ces trois organisations ont alors recherché sur le marché un outil leur permettant de traiter rapidement, et le plus facilement possible, de gros volumes de données. Leur choix s'est au final porté sur Dataiku.

Un équipement dicté par une explosion du volume de données

Outre les capacités de traitement de gros volumes de données par l'outil, ce sont les fonctionnalités collaboratives qu'offre la solution Dataiku, et qui permettent de faire étroitement travailler les techniciens de la donnée avec les utilisateurs finaux (les directions métiers de l'entreprise), qui ont été appréciées.

L'autre atout mis en avant par les entreprises interrogées est la polyvalence de l'outil de développement de Dataiku qui s'adresse à deux populations :

- Les data scientists, côté développeurs, afin de créer des modèles prédictifs spécifiques ;
- Les data specialists, côté métiers, qui analysent des données, et qui peuvent utiliser des modèles pré-développés par Dataiku.

L'existence de modèles prêts à l'emploi offre ainsi la possibilité aux consultants de Roland Berger de "travailler de leur côté sur des modélisations et des analyses simples", tandis que l'équipe de data scientists peut "travailler sur des modélisations plus complexes et plus sophistiquées".

Pour GRTgaz, cette double approche permet à la fois d'avoir des résultats rapides en "lançant plusieurs modèles de machine learning en parallèle et en regardant celui qui donne le meilleur score" mais aussi de mener "des projets plus spécifiques avec les développements en Python".



L'outil offre de nombreuses fonctionnalités pour faciliter le travail de préparation des données.

GRTgaz

Les entreprises ont également souligné l'apport de la solution Dataiku pour faciliter la préparation des données. En effet, elles sont unanimes sur le fait que pour réaliser des modèles prédictifs, il faut des données de bonne qualité, mais que cela peut se révéler très consommateur de temps. Comme le décrit GRTgaz, "dans les projets Data science, on passe beaucoup de temps à raffiner de la donnée, analyser les valeurs manquantes, améliorer la qualité. Là, l'outil offre de nombreuses fonctionnalités pour faciliter ce travail de préparation. C'est un point très fort de l'outil."

Des clients très satisfaits de la solution Dataiku

Le bilan que font les clients de la plate-forme Dataiku est très positif. Elle est considérée comme 'la solution' pour le prédictif à la BRED. De son côté, le représentant de GRTgaz voit dans la solution Dataiku "un outil d'avenir de la Data Science" pour l'ensemble de la communauté Data. Il estime également que la solution "va générer des gains de temps significatifs sur les projets". Enfin, pour Roland Berger, la plate-forme Dataiku correspond "vraiment ce qu'ils recherchaient pour avoir une production plus fluide entre consultants et data scientists". A titre d'exemple, la plate-forme Dataiku a ainsi été utilisée pour développer des modèles prédictifs dans les domaines suivants :

- La maintenance prédictive de distributeurs de billets au sein de la BRED ;
- Des prévisions lors de conditions météorologiques exceptionnelles, au sein de GRTgaz ;
- Des modélisations de réseaux de distribution pour déterminer les endroits les plus pertinents pour de nouvelles implantations, chez Roland Berger.

Confortées dans leurs premières initiatives, les entreprises interrogées ont de fortes ambitions autour de l'utilisation de la solution Dataiku. Comme on l'explique chez Roland Berger, c'est un outil important puisqu'il "accompagne notre stratégie de proposition de nouveaux services".

A terme, les entreprises souhaitent impliquer l'ensemble des services métiers, parmi lesquels les Ressources Humaines, notamment pour "l'analyse d'évolution de carrière" comme au sein de la BRED ou les processus de recrutement chez GRTgaz.

Cette stratégie volontariste s'accompagne de séances de formation pour sensibiliser les métiers aux méthodes statistiques et aux modèles prédictifs. Comme on le résume à la BRED "plus on aura d'ambassadeurs, de témoignages positifs, plus on progressera dans l'analyse des données".



Les modèles prêts à l'emploi permettent aux consultants de travailler sur des analyses simples, tandis que les data scientists peuvent réaliser des modélisations plus sophistiquées

Roland Berger



Pour nous, la plate-forme Dataiku est la solution pour le prédictif

BRED

A propos d'IDC

IDC est un acteur majeur de la Recherche, du Conseil et de l'Évènementiel sur les marchés des Technologies de l'Information, des Télécommunications et des Technologies Grand Public. IDC aide les professionnels évoluant sur les marchés IT et les investisseurs à prendre des décisions stratégiques basées sur des données factuelles. Plus de 1100 analystes proposent leur expertise globale, régionale et locale sur les opportunités et les tendances technologies dans plus de 110 pays à travers le monde. Depuis plus de 50 ans, IDC propose des analyses stratégiques pour aider ses clients à atteindre leurs objectifs clés. IDC est une filiale de la société IDG, leader mondial du marché de l'information dédiée aux technologies de l'information.

IDC France

13 Rue Paul Valéry

75116 Paris, France

+33.1 56.26.26.66

Twitter: @IDCfrance

idc-community.com

www.idc.com / www.idc.fr

Copyright

External Publication of IDC Information and Data — Any IDC information that is to be used in advertising, press releases, or promotional materials requires prior written approval from the appropriate IDC Vice President or Country Manager. A draft of the proposed document should accompany any such request. IDC reserves the right to deny approval of external usage for any reason. Copyright 2019 IDC. Reproduction without written permission is completely forbidden.