

Le Data-Scientist

Étude de Marché



Pierre-Claver Diara
Matthieu Caron
Arnaud Cojez

Le 27 Octobre 2014



Le Data Scientist

Le métier de Data scientist est inspiré de la data science, un terme utilisé pour la première fois par William Cleveland dans un article programmatique paru en 2001,

Un data scientist est un haut responsable de gestion et d'analyse de données « grosses données » (appelées Big Data) dont les principaux secteurs d'activité sont la finance, les banques et assurances, le marketing et l'informatique statistique.

Ce travail allie statistique et programmation, dans le but de tirer des informations plus ou moins évidentes à partir de grandes quantités de données. Les informations obtenues vont de l'impact qu'a un simple clic sur une pub de grande marque pour l'entreprise, jusqu'à l'éventuelle toxicité d'un agent chimique pour le corps humain.

Qu'est-ce que le BIG DATA ?

La différence entre l'analyse « basique » de données et l'analyse type Big Data vient du fait que le jeu de données traitées est largement supérieur aux bases de données habituelles.

La quantité et la variété de ces données font que leurs représentations se font le plus souvent de façon visuelle. Cela permet aux Data Scientists d'identifier, analyser, et mettre en évidence des relations complexes entre différents facteurs, qui resteraient cachées sinon.

Les Data Scientists analysent/« minent » les données provenant de différentes sources : clics, capteurs, statuts sur les réseaux sociaux, transactions, signaux GPS... Leur mission est de dénicher des signes avant-coureurs d'événements ou d'effets de mode afin d'orienter certaines entreprises et les rendre plus compétitives.

L'explosion des données, intensifiée par une technologie en constante évolution, des outils d'analyses de données toujours plus innovants et la prolifération des systèmes de stockage en ligne (cloud etc.), a placé le Data Scientist dans le top des carrières informatiques les plus prisées.



Des profils encore rares

Il n'est pas vraiment possible d'énumérer toutes les qualités qu'un Data Scientist doit avoir. Cela reviendrait à décrire une entité surhumaine, tant ces qualités sont nombreuses. Effectivement, il n'existe pas de profil de Data Scientist « universel » mais plutôt des profils différents adaptés à chaque entreprise.

Cependant, de ces différents profils ressortent deux attributs majeurs :

- Un Data Scientist doit connaître les tenants et aboutissants du domaine dans lequel il devra travailler ;
- Il doit savoir se poser les bonnes questions, et faire prendre d'imagination dans la façon de collecter et d'interpréter les données récoltées ;

Le Data Scientist doit également posséder quelques compétences « techniques » :

- Connaissances en programmation ;
- Maîtrise des mathématiques et statistiques ;
- Connaissances en marketing.

Il maîtrise ainsi de nombreuses technologies comme Hadoop, Java, MapReduce, Bigtable, NoSQL

On voit donc que le Data Scientist n'est pas un simple statisticien. Il devra travailler avec des quantités massives de données non structurées et non nettoyées. Aux méthodes traditionnelles d'analyse statistique, il préférera les techniques de data mining, qui permettent de révéler des propriétés inconnues dans un jeu de données, ou celles du « machine learning » qui cherchent à prédire des comportements futurs sur la base de données connues

Ainsi, alors qu'un statisticien sera plus rigoureux dans ses méthodes ainsi que dans la sélection des jeux de données à analyser, le Data Scientist devra faire preuve de plus d'imagination et d'ouverture d'esprit afin de découvrir certaines tendances dans d'énormes paquets d'informations.



Formations

Malheureusement il n'y a pas beaucoup de formations uniquement dédiée au métier de Data Scientist mais avec certains diplômes spécialisés en statistique, en informatique ou en marketing permettent d'exercer ce métier.

La plupart des Data Scientists actuels sont en possession d'un BAC+5, dans le domaine informatique, mathématique, ou statistique. Bien que certaines formations accélérées commencent à éclore ça et là, il n'existe pas encore de cursus dédié à l'analyse des « grosses données », et c'est le parcours professionnel qui apportera les compétences nécessaires au métier de Data Scientist.

Perspective d'évolutions

La perspective d'évolutions est un des points négatifs de ce métier. En effet, un Data Scientist devra, s'il veut évoluer ou obtenir un meilleur salaire, changer d'agence et aller dans une plus grande entreprise. Cela est très certainement la conséquence d'un métier nouvellement créé. Cependant le métier est déjà promis à un bel avenir.

Présence sur le marché du travail

On prévoit à terme 30000 postes de Data Scientist à pourvoir en France. Cependant comme les personnes maîtrisant complètement les trois qualités importantes (programmation, mathématiques et marketing) sont un peu rares, on pourra certainement voir des entreprises créer des équipes de plusieurs personnes réunissant 2 des 3 compétences importantes.

Salaire :

Un Data Scientist peut gagner de 38 à 50 000 euros par an, ce métier fait partie des 6 les mieux rémunérés de 2014.



Exemple concret

Pour illustrer le rôle et l'importance de ce métier voici un exemple concret :

Il y a environ une dizaine d'années l'entreprise postale UPS réalisa une étude visant à améliorer ses stratégies de livraison.

Des analystes se mirent à étudier de nombreux paramètres sur les trajets empruntés quotidiennement. Au final les analystes trouvèrent que les trajets les plus économiques étaient ceux qui comportaient le moins de virages à gauche, en effet un virage à gauche face au trafic coûte du temps et du carburant. En 2011 cette vaste analyse a permis à l'entreprise postale d'économiser 30 millions de dollars en dépenses de carburant et d'éviter un rejet de CO2 équivalent à 5300 voitures.

Sources :

https://en.wikipedia.org/wiki/Data_science

<http://www.cidj.com/article-metier/data-scientist>

<http://www.journaldunet.com/solutions/expert/56293/les-data-scientists---nouvelles-rock-stars-de-l-it.shtml>

<http://www.itcareerfinder.com/it-careers/big-data-scientist.html>

http://www.alliancy.fr/article/formation-data-scientist-la-perle-rare_4030.html

<http://www.lemondeinformatique.fr/actualites/lire-salaires-entre-38-a-50-000-euros-pour-un-data-scientist-57950.html>