

DATA SCIENTIST



DATA SCIENTIST

Le *data scientist* est expert en mathématiques, statistique et informatique. Ce métier d'explorateur de la donnée relève, selon les profils – plutôt orientés mathématique ou informatique –, de la recherche, du développement d'algorithmes, voire de l'industrialisation de solutions.

Selon les cas, le *data scientist* peut travailler sur des problèmes plus ou moins appliqués ou faire de la recherche pour une entité publique ou privée. Dans le premier cas, avec des compétences plus orientées informatique, il code et développe des algorithmes avancés pour répondre aux enjeux et aux problématiques stratégiques de l'entreprise. En allant à la rencontre des métiers pour en définir les besoins et en collaborant avec les équipes de R&D, il apporte son expertise en techniques et algorithmes d'apprentissage automatique et en systèmes d'intelligence artificielle. Il intervient à toutes les étapes de la chaîne de données : définition du problème, collecte des données, nettoyage, mise en place des modèles prédictifs et création

des algorithmes. Il a d'excellentes capacités de communication et sait adapter son discours à chacun.

Concernant le *data scientist* orienté recherche, différents cas de figure existent. Certains chercheurs font progresser les techniques et technologies au cœur des différentes disciplines des data sciences (*machine learning*, apprentissage profond, etc.). D'autres utilisent les techniques de la data science, notamment l'intelligence artificielle, pour faire avancer les recherches dans leur propre discipline. D'autres encore mettent à profit les évolutions de la data science dans leur propre domaine d'activité : *designers*, juristes, sociologues, économistes...

L'EXPLORATEUR DE DONNÉES COMPLEXES

Au quotidien :

- Explorer de nouvelles pistes dans les ensembles de données et tester des hypothèses
- Créer de nouveaux modèles de données
- Concevoir des algorithmes à base de *machine learning* et les implémenter

Compétences :

- Maîtrise du *machine learning*, *deep Learning*, *Random Forests*, modèles de Markov cachés, *SVM*, régression, séries temporelles, traitement du signal... ;
- Maîtrise des outils *Big data* ;
- Maîtrise de *Python*, *R*, *Java*, *C*, *C++*, *Matlab*, écosystème *Hadoop*, *Spark*... ;
- Esprit d'équipe, communication, ouverture d'esprit, créativité, curiosité, transdisciplinarité ;
- Sensibilité aux enjeux business, notamment dans les secteurs comme le marketing, le web, la publicité...

Formation :

- Masters ou thèses en informatique, sciences cognitives, statistiques, mathématiques... ;
- Bac + 4-5 en mathématiques-statistiques ou informatique ;
- Formation *big data* ;
- Expérience professionnelle (*Data Analyst*).

