



R&D

Biométrie  
Data management

**leem**  
les entreprises  
du médicament

# Data scientist

## Ingénieur(e) base de données

L'ingénieur(e) base de données, aussi appelé(e) data scientist, valorise l'ensemble des données de l'entreprise pour l'aider dans sa prise de décisions et en faire un levier de création de valeur.

Il/elle analyse des masses de données hétérogènes, éventuellement non structurées, pour en extraire de la connaissance utile à l'optimisation des offres et services de l'entreprise. Il/elle possède une vision transverse et croise les données de différentes sources dispersées.



### Profil de recrutement :

Profils confirmés, avec une première expérience de 3 ans à minima.



### Formations

### Parcours recommandés :

- Formation de niveau Bac +5 : Master en économétrie ou informatique, statistiques
- Diplôme d'école d'ingénieurs avec spécialisation big data, ingénierie, application des masses de données
- Doctorat en informatique, statistiques, mathématiques, modélisation de données

Pour aller plus loin : [www.imfis.fr](http://www.imfis.fr)



### Passerelles métier :

**Au sein de la filière métier :**

- Data manager

**Hors filière métier :**

- Chargé(e) / responsable marketing
- Chargé(e) pharmaco-économie

Pour aller plus loin :

<http://www.macarrieredanslapharma.org>



### Autres appellations :

- Data analyst

## ACTIVITÉS

**Gestion de projets - Management d'équipe interne / externe**

- Participation à la définition des projets structurants de l'entreprise, en collaboration avec les métiers (et en particulier la R&D)
- Dimensionnement des projets et définition des choix techniques et méthodologiques des missions
- Pilotage du budget
- Management d'équipe interne ou externalisée

**Extraction, uniformisation et structuration des données**

- Collecte, sélection et validation des données pertinentes pour l'analyse
- Définition des solutions de stockage et de la structuration des données
- Conversion, codage et cartographie des données de consommation ou d'usage produit dans un format compréhensible par l'ensemble des collaborateurs/trices
- Amélioration de la qualité et enrichissement des bases de données
- Détermination des outils et méthodes d'acquisition de données depuis un ensemble de bases techniquement hétérogènes
- Conception de l'architecture d'un entrepôt de données décisionnelles (data warehouse)
- Garantie de la conformité et de la qualité des données dans son domaine d'activité, mise en place des processus de surveillance, mise en œuvre des améliorations appropriées dans son domaine

**Analyses prédictives et développement de la connaissance client**

- Mise en œuvre et garantie de la modélisation statistique des données
- Développement des algorithmes d'apprentissage (machine learning) et des scénarios prédictifs des comportements clients
- Optimisation de la segmentation client à l'aide des statistiques et données de consommation
- Étude et mise en place des meilleures solutions techniques pour gérer les grands volumes de données
- Conception des modèles de détection des insights et des opportunités de marché
- Test et contrôle de la qualité et de la cohérence des bases de données
- Accompagnement de l'entreprise dans le développement de leviers de création de valeur

# Data scientist

## Ingénieur(e) base de données



### Développement d'outils de support aux clients internes

- Participation à la mise en œuvre de la stratégie marketing de l'entreprise
- Analyse de l'ensemble des données pour développer des systèmes d'aide à la décision
- Participation au développement des indicateurs de performance de l'entreprise
- Fourniture aux chefs de produit des leviers statistiques décisionnels pour la conduite et l'analyse des campagnes de prospection
- Réalisation des études statistiques pour les clients internes ou la direction de l'entreprise
- Animation des ateliers d'expression des besoins internes et rédaction des cahiers des charges
- Rédaction de la spécification des besoins à destination des DSI ou de la maîtrise d'ouvrage
- Détermination des outils de reporting dynamiques et multidimensionnels (OLAP4)
- Présentation des résultats des études réalisées aux clients internes
- Formation des utilisateurs aux outils informatiques et décisionnels

### Veille technologique sur les outils de datamining – Archivage

- Réalisation d'une veille et d'une évaluation sur les nouvelles technologies et solutions logicielles d'analyse des données
- Recherche, expérimentation de nouvelles méthodes de modélisation et d'analyse des données
- Archivage, maintien à jour de la documentation technique

# COMPÉTENCES CLÉS

## ←|→ Transverses

- Savoir identifier et comprendre les problématiques économiques transversales de l'entreprise
- Avoir une forte capacité de conceptualisation
- Faire preuve de rigueur et de concentration pour assurer l'exactitude des calculs réalisés
- Savoir structurer ses méthodes de travail et son plan d'intervention
- Communiquer pour expliquer et convaincre
- Faire preuve de pédagogie pour expliciter sa méthode de travail aux non-experts
- Être curieux/euse pour suivre les nouvelles tendances et découvrir de nouveaux outils
- Maîtriser les données numériques et les chiffres car les volumes de données sont importants
- Être force de proposition pour proposer des recommandations aux équipes métiers
- Ecouter afin de recueillir avec précision les informations et besoins des clients internes et externes
- Avoir une appétence pour apprendre

## Anglais

- Comprendre des documents techniques ou scientifiques
- Exprimer ses idées et opinions avec précision et lier ses interventions à celles des interlocuteurs
- Ecrire sur des sujets techniques dans une présentation ou un rapport



## Métier

- Maîtriser les algorithmes d'apprentissage automatique (Machine Learning) et autres programmes d'analyse prédictive
- Comprendre le fonctionnement d'une Business Intelligence et les problématiques liées afin d'en faire évoluer les outils
- Maîtriser les outils de data management (SAS, SPSS, SAP Infinite Insight, Python, R, Excel, Access, etc.)
- Maîtriser les technologies HADOOP
- Connaître les sources / bases de données de santé
- Maîtriser des bases de données SQL et no-SQL.
- Avoir une bonne connaissance des réseaux de neurones et d'intelligence artificielle
- Avoir une bonne connaissance des outils de web analyse (Omniture, Google analytics etc.)
- Être capable de dialoguer avec les utilisateurs/trices et les opérationnels afin de comprendre leurs contraintes techniques et de coordonner les actions d'amélioration nécessaires afin d'optimiser leurs procédés
- Avoir de solides connaissances en marketing
- Intégrer dans son quotidien l'ensemble des éléments juridiques, de conformité et de sécurité relatifs au RGPD et à la bonne utilisation/exploitation des données
- Être capable d'identifier le besoin de maintenance à partir des informations fournies par l'appareil digital, d'intervenir sur les éléments techniques de premier niveau et/ou de solliciter le bon interlocuteur en cas de dysfonctionnement (salarié, prestataire, service, ...)

L'augmentation de la quantité et de la complexité des données entraîne un besoin en profils polyvalents, possédant plusieurs compétences, à la fois informatique / IT et scientifiques / médicales. Cette augmentation de données représente des opportunités de développement à exploiter pour les entreprises.

Le / la data scientist tend à collaborer de plus en plus étroitement avec la R&D du fait du potentiel des données massives en santé pour les projets de recherche, et notamment l'analyse prédictive. Les data scientists interviennent également très souvent en appui d'autres fonctions au sein des entreprises, comme le marketing. Compte tenu de la forte spécialisation nécessaire pour exercer ce métier, les opportunités d'embauche sont nombreuses et supérieures à la quantité de profils qualifiés disponibles sur le marché.

L'automatisation progressive d'une partie des tâches de l'ingénieur(e) base de données / data scientist a pour conséquence d'augmenter sa productivité.

Le profil de data scientist est très recherché, tout en étant régulièrement externalisé, notamment dans les PME.

