

# Formation Moyens et solutions pratiques pour le Big Data Analysis

## Présentation

Cette formation vous permettra de comprendre les enjeux et les apports du Big Data ainsi que les technologies pour le mettre en œuvre. Vous apprendrez à intégrer des volumétries massives de données structurées et non structurées via un ETL, puis à les analyser grâce à des modèles statistiques et des dashboards dynamiques.

Durée : 28,00 heures (4 jours)

Tarif INTRA : [Nous consulter](#)

## Objectifs de la formation

- Comprendre les concepts et l'apport du Big Data par rapport aux enjeux métiers
- Comprendre l'écosystème technologique nécessaire pour réaliser un projet Big Data
- Acquérir les compétences techniques pour gérer des flux de données complexes, non structurés et massifs
- Implémenter des modèles d'analyses statistiques pour répondre aux besoins métiers
- Appréhender un outil de data visualisation pour restituer des analyses dynamiques

## Prérequis

- Connaissances de base des modèles relationnels, des statistiques et des langages de programmation
- Connaissances de base des concepts de la Business Intelligence

## Public

- Dataminers
- Chargés d'études statistiques
- Développeurs
- Chefs de projet
- Consultants en informatique décisionnelle



## Programme de la formation

### Comprendre les concepts et les enjeux du Big Data

- Origines et définition du Big Data : la BI face à la croissance et à la diversité des données
- Les chiffres clés du marché dans le monde et en France
- Les enjeux du Big Data : ROI, organisation et confidentialité des données
- Un exemple d'architecture Big Data

### Les technologies du Big Data

- Description de l'architecture et des composants de la plateforme Hadoop
- Les modes de stockage (NoSQL, HDFS)
- Principes de fonctionnement de MapReduce
- Présentation des distributions principales du marché et des outils complémentaires (Hortonworks, Cloudera, MapR, Aster)
- Installer une plateforme Hadoop
- Présentation des technologies spécifiques pour le Big Data (Talend, Tableau, Qlikview ...)
- Exercice : installation d'une plateforme Big Data complète via Cloudera et ses composants.

### Les méthodes d'analyse des données pour le Big Data

- Les méthodes d'exploration
- Segmentation et classification
- Estimation et prédiction
- L'implémentation des modèles
- Exercice : mise en place d'analyses avec le logiciel R

### Gérer les données structurées et non structurées

- Principes de fonctionnement de Hadoop Distributed File System (HDFS)
- Importer des données externes vers HDFS
- Réaliser des requêtes SQL avec HIVE
- Utiliser PIG pour traiter la donnée
- Utiliser un ETL pour industrialiser la création de flux de données massives

- Présentation de Talend For Big Data.
- Exercice : implémentation de flux de données massives

### **Les méthodes d'analyse des données pour le Big Data**

- Les outils de restitution du marché
- Méthodologie de mise en forme des rapports
- Apport du Big Data pour le « Social Business »
- Mesurer l'e-réputation et la notoriété d'une marque
- Mesurer l'expérience et la satisfaction clients, optimiser le parcours client

## **Organisation**

### **Formateur**

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Riches de leur expérience sur le sujet, ils sauront accompagner vos collaborateurs dans leur montée en compétence.

### **Moyens pédagogiques et techniques**

- Apports des connaissances communes.
- Mises en situation sur le thème de la formation et des cas concrets.
- Méthodologie d'apprentissage attractive, interactive et participative.
- Equilibre théorie / pratique : 60 % / 40 %.
- Supports de cours fournis au format papier et/ou numérique.
- Ressources documentaires en ligne et références mises à disposition par le formateur.
- Pour les formations en présentiel dans les locaux mis à disposition, les apprenants sont accueillis dans une salle de cours équipée d'un réseau Wi-Fi, d'un tableau blanc ou paperboard. Un ordinateur avec les logiciels appropriés est mis à disposition (le cas échéant).

## Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation

- Recueil des besoins des apprenants afin de disposer des informations essentielles au bon déroulé de la formation (profil, niveau, attentes particulières...).
- Auto-positionnement des apprenants afin de mesurer le niveau de départ.

Tout au long de la formation

- Évaluation continue des acquis avec des questions orales, des exercices, des QCM, des cas pratiques ou mises en situation...

A la fin de la formation

- Auto-positionnement des apprenants afin de mesurer l'acquisition des compétences.
- Evaluation par le formateur des compétences acquises par les apprenants.
- Questionnaire de satisfaction à chaud afin de recueillir la satisfaction des apprenants à l'issue de la formation.
- Questionnaire de satisfaction à froid afin d'évaluer les apports ancrés de la formation et leurs mises en application au quotidien.

## Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.