

# Big Data Foundation

3 j (21 heures)

Ref : BDFO

## Public

Connaître les fondamentaux du Big Data, ses origines et ses caractéristiques  
Comprendre ce qu'est le Data Mining  
Appréhender les technologies les plus populaires du Big Data

## Pré-requis

Connaissances de base des architectures techniques

## Moyens pédagogiques

Exposé théorique et participatif, dispensé en FR, alternant : Présentation théorique, Cas d'études, exemples, discussion autour des cas organisationnels des participants

Sont fournis aux stagiaires : Support en FR présenté par un instructeur certifié, format papier prise de note (noir & blanc) ou en accès en ligne, Syllabus Big Data Foundation en EN, Exemple d'examen en EN

Un PC par stagiaire, connexion internet, télé connectée ou vidéoprojecteur

## Modalités de suivi et d'évaluation

Feuille de présence émargée par demi-journée par les stagiaires et le formateur

Exercices de mise en pratique ou quiz de connaissances tout au long de la formation permettant de mesurer la progression des stagiaires

Questionnaire d'évaluation de la satisfaction en fin de stage

Auto-évaluation des acquis de la formation par les stagiaires

Passage de l'examen de certification Big Data Foundation d'Exin en fin de formation (durée : 1h)

Attestation de fin de formation

L'explosion quantitative des données numériques a contraint de trouver de nouvelles manières de voir et d'analyser le monde. La nécessité d'avoir de nouveaux ordres de grandeur concernant la capture, la recherche, le partage, le stockage, l'analyse et la présentation de données a abouti à la naissance de ce que l'on appelle le « Big Data », considéré de nos jours comme une source de bouleversement profond de la société.

Cette formation « Big Data : Les fondamentaux » apporte aux participants les bases du Big Data et du Data Mining et prépare au passage de la certification « Big Data Foundation » d'EXIN.

Le passage de l'examen s'effectue le dernier jour, en ligne et en anglais. Cela consiste en un QCM d'une heure, comportant 40 questions. Un score minimum de 65% est requis pour réussir l'examen.

## Objectifs

Connaître les fondamentaux du Big Data, ses origines et ses caractéristiques

Comprendre ce qu'est le Data Mining

Appréhender les technologies les plus populaires du Big Data

## Programme détaillé

### INTRODUCTION

---

#### LES FONDAMENTAUX DU BIG DATA

---

- Un peu d'histoire...
- Définition, caractéristiques et bénéfices
- Les dimensions (volume, variété, vitesse)
- Les technologies les plus populaires (Hadoop, MongoDB)
- Ethique et vie privée

#### LES SOURCES DU BIG DATA

---

- Les données de l'entreprise (Oracle, SAP, Microsoft, Data Warehouses, données nonstructurées)
- Réseaux sociaux (Facebook, Twitter, autres)
- Données publiques (économiques, financières, comportementales, etc.)

#### DATA MINING

---

- Définition
- Les techniques descriptives et prédictives (Classification, association, segmentation, etc.)
- Exemples d'outils (Weka, KNIME, R)

#### ZOOM SUR HADOOP

---

- Présentation de Hadoop (Les principaux composants, les extensions)
- Installation et configuration
- MapReduce
- Data Processing avec Hadoop

#### ZOOM SUR MONGODB

---

- Présentation de MongoDB
- Réplication et sharding
- Ecosystèmes (langages, drivers, outils, intégration avec Hadoop)
- Installation et configuration
- Base de données de documents
- Modèle de données de documents

#### CERTIFICATION BIG DATA FOUNDATION

---

- Préparation à l'examen « Big Data Foundation »
- Passage de l'examen « Big Data Foundation »

