# BIG DATA, MÉTHODES ET SOLUTIONS

Formation proposée par BIZOÏ RAZVAN

Big Data

Ce stage 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
**5 Jour(s)**
▣MODALITÉS
**Présentiel / Distanciel**
♙PARTICIPANTS
**Jusqu'à 10 participants**
★LANGUE
**Français**
LIEU
**France entière**
♙CERTIFICATION
**Attestation de formation**
▦MISE À JOUR
**11/04/2023**

VOTRE FORMATEUR

BR **BIZOÏ RAZVAN**
Formateur professionnel
Conseiller Pédagogique / Formateur

**TARIFS INDICATIFS**

| | |
|---|---|
| Inter-entreprise | **Pas de dates planifiées** |
| Intra-entreprise | **Sur devis** |
| Individuelle | **Sur devis** |
| Durée | **5 Jour(s)** |

Tarifs indicatifs, sous réserve d'analyse du besoin et de devis.

## ◎Objectifs pédagogiques

- ✓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.

## ⌘Compétences visées

Big Data

## Public cible et prérequis

**Public cible**

- Dataminers, chargés d'études statistiques, développeurs, chefs de projet, consultants en informatique décisionnelle.

**Prérequis**

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

## Programme détaillé

Comprendre les concepts et les enjeux du Big Data
- Origines et définition du Big Data.
- Les chiffres clés du marché dans le monde et en France.
- Les enjeux du Big Data : ROI, organisation, 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, Spark, Storm...
- Principales distributions du marché (Hortonworks, Cloudera, MapR, Elastic Map Reduce, Biginsights).
- Installer une plateforme Hadoop.
- Les technologies du datascientist.
- Présentation des technologies spécifiques pour le Big Data (Tableau, Talend, Qlikview ...).
Travaux pratiques
Installation d'une plateforme Big Data Hadoop.

Gérer les données structurées et non structurées
- Principes de fonctionnement de Hadoop Distributed File System (HDFS).
- Importer des données externes vers HDFS.
- Utiliser SPARK pour traiter la donnée.
- Réaliser des requêtes SQL avec HIVE et SPARK.
- Le principe des ETL (Talend...).
- Gestion de streaming de données massive (NIFI, Kafka, Spark, Storm...)
Travaux pratiques
Implémentation de flux de données massives.

L'univers NoSQL
- Le théorème CAP
- NoSQL
- Les entrepôts clé-valeur
- Les bases orientées documents
- Les bases orientées colonnes
- Le déluge des données
- La diversité des sources
- La diversité des formats
Travaux pratiques
Mise en place une base de données NoSQL.

Technique et méthodes Big data analytics
- Machine Learning, une composante de l'intelligence artificielle.
- Découvrir les trois familles : Régression, Classification et Clustering.
- La préparation des données (data preparation, feature engineering).
- Générer des modèles en R ou Python.
- Ensemble Learning.
- Découvrir les outils du marché : Jupyter Notebook, Dataïku, DataBricks, Amazon Machine Learning...
Travaux pratiques
Mise en place d'analyses avec une des outils étudiés.

Data visualisation et cas d'usage concrets
- Définir le besoin de la data visualisation.
- Analyse et visualisation des données.
- Peut concerner tous les types de données dans la DataViz
- Les outils DataViz du marché.
Travaux pratiques
Installation et utilisation d'un outil de Data Visualisation pour constituer des analyses dynamiques.

Conclusion
- Ce qu'il faut retenir.
- Synthèse des bonnes pratiques.
- Bibliographie.

## ▣Méthodes pédagogiques & évaluation

**Méthodes pédagogiques**

- ✓ Apports théoriques et retours d'expérience
- ✓ Exercices pratiques et mises en situation
- ✓ Travaux de groupe et échanges entre pairs
- ✓ Études de cas concrets

**Modalités d'évaluation**

- ✓ Quiz de validation des acquis
- ✓ Mises en situation évaluées
- ✓ Projet fil rouge
- ✓ Feuille de route post-formation

## ♿Accessibilité

**Une formation accessible à tous**

Ce programme peut être adapté pour les personnes en situation de handicap. Contactez-nous pour étudier vos besoins spécifiques.