

# Data Analyst

1 boulevard du Trieux Pacé 35740

57 jours	12 places	du 02/06/2020 au 04/09/2020
----------	-----------	--------------------------------

Découvrez le programme de la session de formation de “Data analyst” organisée à Pace (35740) du 2 Juin au 4 septembre 2020. Cette formation “Data analyst” dure 57 jours et dispose de 12 places.

La pédagogie principalement pratique dispensée par nos formateurs “experts terrain” permettra au stagiaire à l’issue de la formation d’acquérir l’autonomie nécessaire pour mener à bien ses futures missions grâce à des travaux pratiques, travaux de groupe et un projet fil rouge. Apprendre à maîtriser bases de données, le décisionnel et les principaux outils associés et le Big Data, à travers les outils les plus utilisés. Apprendre à maîtriser bases de données, le décisionnel et les principaux outils associés et le Big Data, à travers les outils les plus utilisés.

.....

## Pré-requis

Force de proposition, Rigueur, Réactivité, Esprit analytique et de synthèse, Aisance rédactionnelle, Excellent relationnel, Sens de l’organisation.

Une première expérience d’un langage et la connaissance de l’algorithmie sont un plus. Être inscrit à Pôle Emploi comme demandeur d’emploi.

Des connaissances en statistiques sont obligatoires, la « bosse » des mathématiques également.

## Financement

Avec Pôle Emploi sous réserve d’être inscrit comme demandeur d’emploi

## Programme

- Présentation générale d’Hadoop
- Exemples d’utilisation dans différents secteurs
- Historique et chiffres clés
- Quand parle-t-on de Big Data ?
- L’écosystème d’Hadoop
- Le système de fichier HDFS
- Le paradigme MapReduce et l’utilisation à travers YARN

- Manipulation des données dans un Cluster Hadoop
- Hue : comment fonctionne cette interface Web ?
- Hive : pourquoi ce n'est pas une base de données ?
- Requête sur Hive
- Utilisation de HCatalog
- Utilisation avancée sur Hive
- Utilisation de fonctions utilisateurs
- Paramétrage de requête
- Pig : fonctionnement
- Programmation avec Pig Latin
- Utilisation du mode Local
- Utilisation de fonctions utilisateurs
- Tez : qu'est-ce que c'est ?
- Oozie : fonctionnement
- Création de workflows avec Oozie
- Manipulation des workflows
- Ajout d'éléments d'exploitation dans les workflows
- Ajout de conditions d'exécution
- Paramétrage des workflows
- Sqoop : à quoi ça sert ?
- Chargement des données depuis une base de données relationnelle
- Chargement des données depuis Hadoop
- Utilisation et paramétrage avancée
- Les particularités des distributions
- Impala et HAWQ

## **Salaire**

Salaire à définir

**A l'issue de cette formation,  
vous serez en mesure de :**