

Big Data Foundation

La formation Big Data Foundation fournit une compréhension du Big Data, des sources potentielles utilisées pour résoudre des problématiques d'entreprise telles que comment peut-on exploiter le Big Data ? Quelles sont les règles de la Gouvernance du Big Data ? Comment les technologies de Hadoop et MongoDB permettent d'obtenir des résultats ?

Objectifs

- Comprendre le phénomène Big Data et les notions qu'il recouvre
- Expliquer le Data Mining
- Appréhender les outils applicables au process Big Data
- Comprendre et installer les technologies du Big Data : Hadoop et MongoDB, Python, R, Redshift, Airflow, Sklearn, Google Cloud Machine Learning
- Réaliser les contraintes, les limites mais aussi l'intérêt du mode inductif permis par le Big Data
- Passer dans les conditions optimales la certification internationale «Big Data Foundation» d'exin /Cloud Credential Council

Prérequis et nombre minimum 4 et maximum 12 stagiaires

- Des connaissances en langue anglaise et une culture de la Gestion des Systèmes d'Information est conseillée pour bien comprendre l'ensemble de la formation. Se munir de sa pièce d'identité pour le passage de l'examen
- Les participants doivent être munis de leur ordinateur avec 4GB RAM et la connexion internet

Public

CDO (Chief Digital Office), Directions Métiers, Chef de Projet Big Data, Business Analyst, Ingénieur Analyst Big Data, Développeur, Administrateur Réseaux, Architecte IT, consultants et opérationnels.

Durée (Intégrant 1 pause matin et après-midi et pauses déjeuners), dates et lieux

3 journées : 9h00-17h00 - 21 heures (Travail personnel non inclus) - dates/lieux : <https://www.valnaos.com/>

Moyens : pédagogiques, techniques et encadrement

Pour la formation à distance en visioconférence : besoin d'un PC, internet et vidéocaméra. Projection du cours. Le support et l'examen sont en langue anglaise. La formation est en langue française. Le cours est théorique à 60% et 40% pratique avec des discussions, partage d'expérience et étude de cas sur une plateforme dédiée sur internet (Work-Labs). Un accès à notre plateforme permet de télécharger tous les documents utiles à la formation. Prévoir pendant la formation votre smartphone/pc pour vous connecter. L'accès Wifi vous est fourni en présentiel. Le formateur est un consultant terrain avec plus de 15 ans d'expérience et/ou certifié et accrédité pour délivrer la formation.

Moyens d'évaluation, déroulement de l'examen et validation de la formation

- Des quizz et un examen a blanc sont proposés pour évaluer vos connaissances, compétences.
- QCM de 40 questions en langue anglaise (les candidats peuvent amener un dictionnaire format papier) – examen version web : examen sur l'interface de l'organisme certifiant. Besoin d'un PC, internet, vidéocaméra. Pas de téléchargement d'application particulier. Résultat immédiat.
- Obtention du certificat à partir de 26 points. Un certificat de réalisation d'action de formation est délivré en fin de formation
- Durée : maximum : 60 minutes + 15 minutes pour les non anglophones

Programme

Introduction

- Les objectifs, l'agenda, les activités : discussion, théorie, les exercices et études de cas

Les Fondamentaux du Big Data

- Big Data : l'histoire – les caractéristiques - de la déduction à l'industrialisation
- Les 3 "V" : Volumétrie exponentielle - Vitesse augmentée - Variété des données
- Les Technologies Big Data Hadoop et MongoDB : les méthodes, les concepts, les utilisations
- Le privé et les codes éthiques : la conformité, les défis, l'approche
- Les projets Big Data : qui est impliqué ? qu'est ce qui est impliqué ?

Les Sources du Big Data : Dans les sources des entreprises - des média sociaux – Public : Identifier les différentes sources, Expliquer, comment accéder aux sources Data, Démontrer comment obtenir les data des sources

Data Mining : Introduction – les outils – les concepts

Les Technologies Big Data : vue en détail

- Hadoop : les fondamentaux – installer, configurer et MapReduce - MongoDB : les fondamentaux – installer, configurer, les bases de données, data Modelling

Use cases : REX du formateur sur la gestion d'équipes data dans 2 startups

- Technologies abordées : Python, R, Redshift, Airflow, Sklearn, Google Cloud Machine Learning
- Contexte et problématique métier des use cases (conception d'un produit de pricing, construction d'une plateforme analytics, Segmentation client et optimisation produit)
- Conception des produits et analyses
- Discussion des choix d'architecture technique
- Analyse des dimensions managériales et organisationnelles
- Bilan des succès et échecs

Révisions et passage de l'examen final en fin de formation (voir le déroulement de l'examen)