

# Data Engineering

Durée : 399 heures

## Description:

Cette formation a pour but d'acquérir les connaissances nécessaires permettant aux participants d'acquérir les connaissances nécessaires qui leur permettent de travailler en tant que Data Engineer ou Data Analyst sur Azure, et obtenir facilement les certificats : PL-300: Microsoft Power BI Data Analyst et DP-203: Data Engineering on Microsoft Azure

## PROGRAMME

### Module 1: Analyse et Visualisation des Données avec Power BI (10 jours)

- Introduction à la Business Intelligence
- Présentation de Power BI
- Importer et modéliser les données avec Power BI
- Visualiser les données avec Power BI
- Explorer les données avec Power BI
- Publier et Partager avec Power BI
- Préparation à l'examen DA-100 : Analyzing data with Microsoft Power BI

### Module 2: Microsoft BI (10 jours)

**Querying Microsoft SQL Server:** Cette formation permet d'apporter aux participants les compétences et connaissances nécessaires pour écrire des requêtes Transact-SQL pour Microsoft SQL Server

Cette formation est la base de toutes les possibilités offertes par SQL Server : administration et développement de la base de données et la Business Intelligence.

**Administering Microsoft SQL Server Databases:** Cette formation fournit les compétences et connaissances

nécessaires pour maintenir une base de données SQL Server . La formation met l’accent sur les fonctionnalités du produit et les outils liés à la maintenance des bases de données.

**Implementing A Data Warehouse With Microsoft SQL Server:** La formation Microsoft permet d'acquérir les compétences et connaissances pour mettre en oeuvre une plate-forme data warehouse pour supporter une solution BI. Les participants apprennent à créer une data warehouse avec SQL Server , à mettre en oeuvre ETL avec SQL Server Integration Services, valider et à « purifier » les données avec SQL Server Data Quality Services et SQL Server Master Data Services.

### Module 3: Azure Analysis Services (3 jours)

- Introduction
- Découvrir Azure Analysis Services
- Créer un serveur Azure Analysis Services
- Administration d'un serveur Azure Analysis Services
- Déploiement des Cubes dans Azure Analysis Services

### Module 4: Databricks (5 jours)

- Introduction
- Découvrir Azure Databricks
- Principes de base de l'architecture Spark
- Apprendre les bases de Python
- Découvrir PySpark et Spark-SQL
- Créer un espace de travail et un cluster Azure Databricks
- Créer et exécuter des notebooks Databricks
- Lire et écrire des données
- Utiliser des DataFrames dans Azure Databricks
- Sécurité et la protection des données dans Azure Databricks
- Structured Streaming : Traiter des données de streaming avec Azure Databricks
- Décrire les bonnes pratiques Azure Databricks

### Module 5: Azure Data Factory (5 jours)

- Introduction
- Découvrir Azure Data Factory
- Composants d'Azure Data Factory
- Transformation des données
- Gérer la sécurité Azure Data Factory
- Runtimes d'intégration
- Créer des services liés et les jeux de données
- Créer des pipelines et activités
- Monitor and Manage

### Module 6: Intégration continue déploiement continu CI/CD (3 jours)

- Introduction
- Décrire CI/CD
- GIT Version Control

### Module 7: Azure DevOps (3 jours)

- Introduction
- Découvrir Azure DevOps
- Contrôle de version d'un Azure Data Factory avec Azure DevOps
- Développement collaboratif
- Créer un pipeline de déploiement avec Azure Pipelines

### Module 8: Azure Logic Apps

- Introduction
- Présentation d'Azure Logic Apps
- Router et traiter des données automatiquement à l'aide de Logic Apps
- Découvrir des connecteurs pour Azure Logic Apps
- Créer et déployer des Logic Apps à l'aide des modèles Azure Resource Manager

### Mini projets/ Use cases (10 jours)

Travailler sur des mini projets sous formes d'ateliers pratiques.

### Préparation à l'examen PL-300 et DP-203 (5 jours)

Tests pratiques avec les types de questions figurant dans l'examen

## OBJECTIF :

- Analyser et Visualiser des données en utilisant l'outil Power BI.
- Apprendre à tirer parti de la puissance des clusters Apache Spark exécutés sur la plateforme Azure
- Databricks pour traiter de grandes charges de travail d'engineering données dans le cloud.
- Apprendre à créer et à gérer des pipelines de données dans le cloud avec Azure Data Factory
- Comprendre les concepts de CICD et développement collaboratif
- Gérer et déployer des applications/pipelines avec Azure DevOps
- Construire des flux ServerLess Logic Apps

## PREREQUIS

- Notions de base en programmation (boucles, affectation de variables, etc)
- Notions de base en Big Data

## MODALITE PEDAGOGIQUE

**Documents supports de formation projetés.**

- Exposés théoriques
- Etude de cas concrets
- Quiz en salle
- Mise à disposition en ligne de documents supports à la suite de la formation.

## MODALITE D'EVALUATION

- Pendant la formation : QCM, Mise en situation.

## DEBOUCHE

**Professionnel:** Data Analyst, Data engineer, Data Architect, Cloud data engineer, Machine learning engineer

## ACCESSIBILITÉ AUX PERSONNES EN SITUATION DE HANDICAP

**Formation accessible aux personnes en situation de handicap. Contacter notre référent handicap : Julina DESSALINES au 01 42 80 38 39 ou [j.dessalines@ibssc.com](mailto:j.dessalines@ibssc.com)**