

Microsoft Azure - Ingénierie de données

Exploitez vos données avec Microsoft Azure

 Présentiel ou en classe à distance

Durée : 4 jours (28 h)

+ activité à distance

+ certificat optionnel

Réf. : MSDP203

Prix inter : 2.665,00 € HT

Forfait intra : 8.380,00 € HT

Formation officielle



L'économiste et auteur britannique Ronald Coase a dit « torturer les données et elles avoueront tout ». Au regard de l'importance croissante que prend la maîtrise de la donnée dans le monde actuel, c'est n'est pas un hasard si l'on entend de plus en plus parler du data engineering, une discipline qui se présente dorénavant comme une branche à part entière des sciences de la data. L'objectif du data engineering est de sélectionner, trier, stocker, agencer, tester et sécuriser les données de manière à pouvoir garantir leur qualité, leur disponibilité et leur pertinence pour les équipes qui les exploiteront. A l'issue de cette formation de 4 jours, les participants maîtriseront l'ingénierie de données reposant sur les technologies de la plateforme Azure.

Cette formation prépare à la certification Microsoft Certified Azure Data Engineer Associate.

Les objectifs de la formation

- Savoir explorer les options de calcul et de stockage pour les charges de travail d'ingénierie des données dans Azure
- Apprendre à exécuter des requêtes interactives à l'aide de pools SQL sans serveur
- Effectuer l'exploration et la transformation des données dans Azure Databricks
- Comprendre comment explorer, transformer et charger des données dans l'entrepôt de données à l'aide d'Apache Spark
- Pouvoir intégrer et charger des données dans l'entrepôt de données
- Être capable de transformer les données avec Azure Data Factory ou Azure Synapse Pipelines
- Apprendre à intégrer les données des blocs-notes avec Azure Data Factory ou Azure Synapse Pipelines
- Pouvoir prendre en charge le traitement analytique transactionnel hybride (HTAP) avec Azure Synapse Link
- Savoir réaliser une sécurité de bout en bout avec Azure Synapse Analytics
- Être capable d'effectuer un traitement de flux en temps réel avec Stream Analytics
- Apprendre à créer une solution de traitement de flux avec Event Hubs et Azure Databricks

A qui s'adresse cette formation ?

Pour qui

- Professionnels des données, architectes de données et professionnels BI qui souhaitent en savoir plus sur l'ingénierie des données et la création de solutions analytiques à l'aide des technologies de plate-forme de

données existantes sur Microsoft Azure

- Analystes de données et data scientists qui travaillent avec des solutions analytiques basées sur Microsoft Azure

Prérequis

- Avoir suivi les formations "[Microsoft Azure - Notions fondamentales](#)" (MSAZ900) et "[Microsoft Azure - Principes fondamentaux des données](#)" (MSDP900) ou connaissance du cloud computing et des concepts de base des données et avoir une expérience pratique avec des solutions de données
- [Disposez-vous des compétences nécessaires pour suivre cette formation ? Testez-vous !](#)

CPF

Cette formation préparant au passage de la certification "Concevoir et mettre en oeuvre la gestion des données dans Microsoft Azure" est éligible au CPF (331067). Connectez-vous sur www.moncompteformation.gouv.fr

Programme

1 - Explorer les options de calcul et de stockage pour les charges de travail d'ingénierie des données

- Introduction à Azure Synapse Analytics
- Décrire Azure Databricks
- Introduction au stockage Azure Data Lake
- Décrire l'architecture Delta Lake
- Travailler avec des flux de données à l'aide d'Azure Stream Analytics

2 - Exécuter des requêtes interactives à l'aide de pools SQL sans serveur Azure Synapse Analytics

- Explorer les fonctionnalités des pools SQL sans serveur Azure Synapse
- Interroger des données dans le lac à l'aide de pools SQL sans serveur Azure Synapse
- Créer des objets de métadonnées dans des pools SQL sans serveur Azure Synapse
- Sécuriser les données et gérer les utilisateurs dans les pools SQL sans serveur Azure Synapse

3 - Exploration et transformation des données dans Azure Databricks

- Décrire Azure Databricks
- Lire et écrire des données dans Azure Databricks
- Utiliser DataFrames dans Azure Databricks
- Utiliser les méthodes avancées DataFrames dans Azure Databricks

4 - Explorer, transformer et charger des données dans l'entrepôt de données à l'aide d'Apache Spark

- Comprendre l'ingénierie Big Data avec Apache Spark dans Azure Synapse Analytics
- Intégrer des données avec des blocs-notes Apache Spark dans Azure Synapse Analytics
- Transformer des données avec DataFrames dans des pools Apache Spark dans Azure Synapse Analytics
- Intégrer des pools SQL et Apache Spark dans Azure Synapse Analytics

5 - Intégrer et charger des données dans l'entrepôt de données

- Utiliser les meilleures pratiques de chargement des données dans Azure Synapse Analytics
- Ingestion à l'échelle du pétaoctet avec Azure Data Factory

6 - Transformer les données avec Azure Data Factory ou Azure Synapse Pipelines

- Intégration de données avec Azure Data Factory ou Azure Synapse Pipelines
- Transformation sans code à grande échelle avec Azure Data Factory ou Azure Synapse Pipelines

7 - Orchestrer le mouvement et la transformation des données dans Azure Synapse Pipelines

- Orchestrer le mouvement et la transformation des données dans Azure Data Factory

8 - Sécurité de bout en bout avec Azure Synapse Analytics

- Sécuriser un entrepôt de données dans Azure Synapse Analytics
- Configurer et gérer les secrets dans Azure Key Vault
- Mettre en oeuvre des contrôles de conformité pour les données sensibles

9 - Prise en charge du traitement analytique transactionnel hybride (HTAP) avec Azure Synapse Link

- Concevoir un traitement transactionnel et analytique hybride à l'aide d'Azure Synapse Analytics
- Configurer Azure Synapse Link avec Azure Cosmos DB
- Interroger Azure Cosmos DB avec des pools Apache Spark
- Interroger Azure Cosmos DB avec des pools SQL sans serveur

10 - Traitement de flux en temps réel avec Stream Analytics

- Activer une messagerie fiable pour les applications Big Data à l'aide d'Azure Event Hubs
- Utiliser des flux de données à l'aide d'Azure Stream Analytics
- Ingérer des flux de données avec Azure Stream Analytics

11 - Créer une solution de traitement de flux avec Event Hubs et Azure Databricks

- Traiter les données de streaming avec le streaming structuré Azure Databricks

Certificat

Les points forts de la formation

- Une formation très pratique : les participants se familiarisent avec les modèles et les pratiques d'ingénierie des données à l'occasion de nombreux TP et mises en situation.
- Le partage de bonnes pratiques de la part de spécialistes de l'ingénierie de données en environnement Azure
- La qualité d'une formation officielle Microsoft (support de cours numérique en anglais).
- Cette formation préparant au passage de la certification "Concevoir et mettre en oeuvre la gestion des données dans Microsoft Azure" est éligible au CPF (331067). Connectez-vous sur www.moncompteformation.gouv.fr
- 92% des participants à cette formation se sont déclarés satisfaits ou très satisfaits au cours des 12 derniers mois.

Dates et villes 2024 - Référence MSDP203

Paris

du 10 juin au 13 juin
du 18 nov. au 21 nov.

du 22 juil. au 25 juil.

du 23 sept. au 26 sept. [Session garantie](#)

A distance

du 10 juin au 13 juin
du 18 nov. au 21 nov.

du 22 juil. au 25 juil.

du 23 sept. au 26 sept. [Session garantie](#)

Aix-en-Provence

du 10 juin au 13 juin

du 23 sept. au 26 sept.

Sophia Antipolis

du 10 juin au 13 juin

du 23 sept. au 26 sept.

Bordeaux

du 10 juin au 13 juin
du 18 nov. au 21 nov.

du 22 juil. au 25 juil.

du 23 sept. au 26 sept.

Strasbourg

du 10 juin au 13 juin

du 23 sept. au 26 sept.

Nantes

du 10 juin au 13 juin

du 22 juil. au 25 juil.

du 18 nov. au 21 nov.

Lyon

du 10 juin au 13 juin

du 23 sept. au 26 sept.

Rennes

du 10 juin au 13 juin

du 22 juil. au 25 juil.

du 18 nov. au 21 nov.

Toulouse

du 22 juil. au 25 juil.

du 18 nov. au 21 nov.

Lille

du 22 juil. au 25 juil.

du 18 nov. au 21 nov.

Rouen

du 22 juil. au 25 juil.

du 23 sept. au 26 sept.

du 18 nov. au 21 nov.