Devenir certifié Microsoft Azure Data Engineer (DP-203)

Date et durée

Code formation : DP-203-BIS
Durée : 5 jours
Nombre d'heures : 35 heures

Formation avec certification

Azure Data Engineer Associate

Description

Important : en raison de la mise à jour des certifications Microsoft, cette formation Microsoft DP-203 a été remplacée par la nouvelle formation Microsoft DP-700 : Analyser et visualiser les données avec Microsoft Azure

Objectifs

Au terme de cette formation **Microsoft Azure Data Engineer (DP-203)**, vous serez capable de valider les objectifs suivants :

- **Concevoir** des solutions de données fiables et évolutives dans Azure.
- **Intégrer et transformer** les données provenant de multiples sources pour répondre aux besoins analytiques.
- **Optimiser** les flux de données pour une meilleure performance et sécurité.
- **Automatiser et monitorer** les pipelines de données pour garantir la continuité des services.

Points forts

Formation certifiante, avec labs pratiques en français, pour maîtriser l'ingénierie de données sur Azure, menée par des experts Microsoft.
Garantie de certification : cette formation inclut le "Microsoft Exam Replay", ce qui vous permet de repasser l'examen gratuitement en cas d'échec à la première tentative.

Certification

Cette formation prépare au passage de l'examen DP-203, vous permettant d'obtenir la certification **Microsoft Certified: Azure Data Engineer Associate**. Cette certification valorise vos compétences en ingénierie de données sur Azure, attestant de votre expertise dans la mise en place de solutions de données sécurisées et performantes dans le cloud. Elle constitue une référence de choix pour les employeurs et est un atout pour progresser vers des rôles de Data Engineer avancés.

Modalités d'évaluation

Travaux Pratiques

Pré-requis

- Expérience professionnelle en bases de données ou en solutions de données.
- Connaissances en SQL et en plateformes cloud, idéalement Azure.

Public

- Ingénieurs de données cherchant à maîtriser Azure.
- Architectes cloud souhaitant approfondir l’ingénierie des données.
- Analystes de données et professionnels de la BI cherchant à optimiser leurs compétences.

Cette formation s'adresse aux profils suivants

Analyste de données
Architecte informatique / SI
Ingénieur système

Programme

Les fondamentaux de Microsoft Azure DP-900*(1 jour)*

Module 1 : découverte et compréhension des fondements de donnée Azure

- L'analyse des principaux concepts de données.
- Les rôles et les responsabilités de chacun dans le traitement des données.
- Les bases de données relationnelles et non relationnelles.
- Les principes de l'analyse de données.

À la fin de ce module N°1, vous acquérez les compétences suivantes :

- disposer de notions de base sur les services de données dans le cloud ;
- reconnaître et expliquer les principaux types de données, notamment les données relationnelles, non relationnelles, les mégadonnées et les données analytiques ;
- décrire le fonctionnement de la technologie Azure.

Module 2 : l’analyse des données relationnelles Azure

- Présentation des services de bases de données relationnelles.
- Création et déploiement des services de bases de données relationnelles.
- Interrogation de données relationnelles via des requêtes.

À la fin de ce module N°2, vous acquérez les compétences suivantes :

- présenter et expliquer le fonctionnement des services de données relationnelles ;
- décrire les étapes de création et de déploiement des bases de données relationnelles ;
- effectuer des requêtes sur des données relationnelles en utilisant des applications de Cloud Data.

Module 3 : l’analyse des données non-relationnelles Azure

- Présentation des services de bases de données non-relationnelles.
- Création et déploiement des services de bases de données non-relationnelles.
- Gestion des entrepôts de données non-relationnels.

À la fin de ce module N°3, vous acquérez les compétences suivantes :

- présenter et expliquer le fonctionnement des services de données relationnelles ;
- décrire les étapes de création et de déploiement des bases de données relationnelles ;
- gérer les data warehouses non-relationnelles.

Module 4 : l'analyse moderne des entrepôts de données Azure

- Présentation des différents éléments d'un entrepôt de données moderne.
- Découverte du processus d'ingestion de données.
- Exploration du stockage et du traitement des données.
- Création de tables et saisie de données avec Power BI.

À la fin de ce module N°4, vous acquérez les compétences suivantes :

- exposer les différentes possibilités de traitement pour la création de solutions d'analyse de données ;
- présenter les services Azure Synapse Analytics, Azure Databricks et Azure HDInsight ;
- expliquer la fonction de Microsoft Power BI, notamment ses modules et leur interaction avec Azure.

Data Engineer sur Microsoft Azure (DP-203)*(4 jour)*

Module 1 : Exploration des services de données sur Azure

- Introduction aux concepts de base de l’ingénierie des données et à l’écosystème Azure.
- Présentation de Synapse Analytics, Azure Data Lake, et Azure Databricks.
- Différenciation des types de données, structures et flux de données.
- **Lab :** Configuration initiale d’un environnement de données sur Azure, création et gestion d’un lac de données.

Module 2 : Ingestion de données sur Azure

- Orchestration et ingestion des données via Azure Data Factory.
- Utilisation d’Azure Synapse pour connecter et ingérer des données de sources multiples.
- Exploration des différents connecteurs de données et techniques de transformation.
- **Lab :** Création de pipelines de données dans Data Factory, ingestion de données à partir de bases SQL et de sources de données non relationnelles.

Module 3 : Transformation des données avec Azure

- Introduction à Azure Databricks pour la transformation des données.
- Utilisation de Spark SQL et de DataFrames pour manipuler les données.
- Automatisation des transformations dans les pipelines de données.
- **Lab :** Transformation de données avec Databricks et intégration avec Data Factory.

Module 4 : Intégration et gestion des flux de données

- Combinaison des flux de données avec Azure Synapse et Data Lake.
- Création de tables externes et gestion des partitions dans Synapse.
- Suivi des pipelines et débogage d’erreurs dans les flux de données.
- **Lab :** Création de flux de données complexes, gestion des tables de partition et débogage dans Azure Synapse.

Module 5 : Optimisation des performances et gestion des coûts

- Optimisation des requêtes et analyse des performances des pipelines.
- Gestion des coûts dans les services de données Azure.
- Utilisation d’Azure Monitor et d’Application Insights pour surveiller les performances.
- **Lab :** Optimisation des requêtes et surveillance de l’utilisation des ressources pour un contrôle des coûts.

Module 6 : Mise en place de la sécurité et conformité des données

- Application de stratégies de sécurité et de conformité dans les services Azure.
- Gestion des accès via Azure Active Directory et des contrôles d’accès conditionnels.
- Implémentation du cryptage des données au repos et en transit.
- **Lab :** Configuration des accès sécurisés, application de contrôles d’accès et de cryptage des données dans Azure Synapse et Data Lake.

Module 7 : Automatisation et gestion de la haute disponibilité

- Automatisation des tâches récurrentes dans les pipelines de données.
- Stratégies de haute disponibilité et de reprise après sinistre (HADR) pour les solutions de données.
- Configuration de la géo-réplication et des sauvegardes de données.
- **Lab :** Mise en place de la géo-réplication, sauvegarde et restauration, et configuration des options de haute disponibilité.

Microsoft®, Microsoft Azure Active Directory®, Microsoft 365®, Microsoft Entra® et Power Platform® sont des marques déposées ou des marques commerciales de Microsoft Corporation aux États-Unis et dans d'autres pays.