

Formation Microsoft Azure - Ingénierie de données

Référence : **MSDP203**

Durée : **4 jours**

Certification : **DP-203**

CONNAISSANCES PREALABLES

- Avoir suivi les formations "Microsoft Azure - Notions fondamentales" (MSAZ900) et "Microsoft Azure - Notions fondamentales 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.

PROFIL DES STAGIAIRES

- 1-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. • 2-Analystes de données et data scientists qui travaillent avec des solutions analytiques basées sur Microsoft Azure.

OBJECTIFS

- Savoir explorer les options de calcul et de stockage pour les charges de travail d'ingénierie des données dans Azure. • Être capable de concevoir et mettre en œuvre la couche de diffusion. • Pouvoir comprendre les considérations d'ingénierie des données. • Apprendre à exécuter des requêtes interactives à l'aide de pools SQL sans serveur. • Comprendre comment explorer, transformer et charger des données dans l'entrepôt de données à l'aide d'Apache Spark. • Effectuer l'exploration et la transformation des données dans Azure Databricks. • 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 optimiser les performances des requêtes avec des pools SQL dédiés dans Azure Synapse. • Comprendre comment analyser et optimiser le stockage de l'entrepôt de données. • 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. • Comprendre comment créer des rapports à l'aide de l'intégration de Power BI avec Azure Synapse Analytics. • Pouvoir effectuer des processus d'apprentissage automatique intégrés dans Azure Synapse Analytics.

CERTIFICATION PREPAREE

- Data Engineering on Microsoft Azure

METHODES PEDAGOGIQUES

- Mise à disposition d'un poste de travail par stagiaire
- Remise d'une documentation pédagogique papier ou numérique pendant le stage
- La formation est constituée d'apports théoriques, d'exercices pratiques, de réflexions et de retours d'expérience
- Le suivi de cette formation donne lieu à la signature d'une feuille d'émargement

FORMATEUR

Consultant-Formateur expert Azure

METHODE D'EVALUATION DES ACQUIS

- Auto-évaluation des acquis par le stagiaire via un questionnaire
- Attestation de fin de stage adressée avec la facture

CONTENU DU COURS

Explorer les options de calcul et de stockage pour les charges de travail

- 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

Concevoir et mettre en œuvre la couche de service

- Concevoir un schéma multidimensionnel pour optimiser les charges de travail analytiques
- Transformation sans code à grande échelle avec Azure Data Factory
- Remplir les dimensions à évolution lente dans les pipelines Azure Synapse Analytics

Considérer l'ingénierie de données pour les fichiers sources

- Concevoir un entrepôt de données moderne à l'aide d'Azure Synapse Analytics
- Sécuriser un entrepôt de données dans Azure Synapse Analytics

Exécuter des requêtes interactives à l'aide de pools SQL sans serveur Azure

- Découvrir 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

Explorer, transformer et charger des données dans l'entrepôt de données

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

Explorer et transformer des données dans Azure Databricks

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

Ingérer et charger des données dans l'entrepôt de données

- Utiliser les meilleures pratiques de chargement de données dans Azure Synapse Analytics

- Ingestion à l'échelle du pétaoctet avec Azure Data Factory

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

Orchestrer le mouvement et la transformation des données dans Azure

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

Optimiser les performances des requêtes avec des pools SQL dédiés dans Azure

- Optimiser les performances des requêtes d'entrepôt de données dans Azure Synapse Analytics
- Comprendre les fonctionnalités pour les développeurs de l'entrepôt de données d'Azure Synapse Analytics

Analyser et Optimiser le stockage de l'entrepôt de données

- Analyser et optimiser le stockage de l'entrepôt de données dans Azure Synapse Analytics

Prendre en charge le traitement analytique transactionnel hybride (HATP)

- 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

Sécuriser 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 œuvre des contrôles de conformité pour les données sensibles

Traiter des flux en temps réel avec Stream Analytics

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

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

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

Créer des rapports à l'aide de l'intégration de Power BI avec Azure Synapse

- Créer des rapports avec Power BI à l'aide de son intégration avec Azure Synapse Analytics

Effectuer des processus d'apprentissage automatique intégrés dans Azure

- Utiliser le processus d'apprentissage automatique intégré dans Azure Synapse Analytics

Certification Microsoft Data Engineering on Microsoft Azure

- Cette formation prépare au passage de la certification Microsoft Data Engineering on Microsoft Azure (DP-203)