

AZURE

## Ingénierie de données

4 jours (28h00) | MW0007 | Num form : form-237 | Perfectionnement / Avancé

INFORMATIQUE / BASE DE DONNÉES & DÉCISIONNEL / AZURE

### À l'issue de ce stage vous serez capable de :

- 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 oeuvre 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

### Niveau requis :

- Connaissance du cloud computing et des concepts de base des données et avoir une expérience pratique avec des solutions de données

### **Public concerné :**

- 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

### **Programme :**

#### **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

#### **CONCEVOIR ET METTRE EN OEUVRE 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

#### **CONSIDÉRATIONS D'INGÉNIERIE DES DONNÉES POUR LES FICHIERS SOURCE**

- 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 SYNAPSE ANALYTICS**

- 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 À L'AIDE D'APACHE SPARK**

- 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

## **EXPLORATION ET TRANSFORMATION 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 SYNAPSE PIPELINES**

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

## **OPTIMISER LES PERFORMANCES DES REQUÊTES AVEC DES POOLS SQL DÉDIÉS DANS AZURE SYNAPSE**

- 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

## **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

## **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

## **TRAITEMENT DE FLUX EN TEMPS RÉEL AVEC STREAM ANALYTICS**

- Activez 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 ANALYTICS**

- 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 SYNAPSE ANALYTICS**

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