



Data Engineering on Microsoft Azure (DP-203)

Description

Ce cours **Data Engineering on Microsoft Azure (DP-203)** vous permettra de découvrir les modèles et les pratiques d'ingénierie des données dans le cadre de solutions analytiques en temps réel et par lots utilisant les technologies de la plate-forme de données Azure.

Vous découvrirez les technologies de calcul et de stockage de base qui sont utilisées pour construire une solution analytique, et explorerez ensuite comment concevoir des couches de service analytiques et se concentrer sur les considérations d'ingénierie des données pour travailler avec des fichiers sources.

Vous apprendrez à travers ce cours Data Engineering on Microsoft Azure, à explorer de manière interactive les données stockées dans des fichiers dans un lac de données, les différentes techniques d'ingestion qui peuvent être utilisées pour charger des données à l'aide de la fonctionnalité Apache Spark présente dans Azure Synapse Analytics ou Azure Databricks, ou comment ingérer à l'aide d'Azure Data Factory ou des pipelines Azure Synapse.

Vous aborderez également les différentes façons de transformer les données à l'aide des mêmes technologies que celles utilisées pour l'acquisition des données et apprendront à surveiller et à analyser les performances des systèmes analytiques afin d'optimiser les performances des charges de données ou des requêtes émises sur les systèmes. Enfin, vous comprendrez l'importance de la mise en œuvre de la sécurité pour garantir la protection des données au repos ou en transit.

Contenu du cours

Module 1 : Explorez les options de calcul et de stockage pour les charges de travail en ingénierie des données

- Leçon 1 : Introduction à Azure Synapse Analytics
- Leçon 2 : Décrire Azure Databricks
- Leçon 3 : Introduction au stockage Azure Data Lake
- Leçon 4 : Décrire l'architecture du lac Delta
- Leçon 5 : Travaillez avec les flux de données en utilisant Azure Stream Analytics

Module 2 : Concevoir et implémenter la couche de service

- Leçon 1 : Concevoir un schéma multidimensionnel pour optimiser les charges de travail analytiques
- Leçon 2 : Transformation sans code à grande échelle avec Azure Data Factory

- Leçon 3 : Peuplez les dimensions changeantes lentement dans les pipelines Azure Synapse Analytics

Module 3 : Considérations d'ingénierie des données pour les fichiers sources

- Leçon 1 : Concevoir un entrepôt de données moderne à l'aide d'Azure Synapse Analytics
- Leçon 2 : Sécuriser un entrepôt de données dans Azure Synapse Analytics

Module 4 : Exécutez des requêtes interactives à l'aide des pools SQL sans serveur Azure Synapse Analytics

- Leçon 1 : Explorer les fonctionnalités de pools SQL sans serveur Azure Synapse
- Leçon 2 : Données de requête dans le lac à l'aide des pools SQL sans serveur Azure Synapse
- Leçon 3 : Créer des objets de métadonnées dans les pools SQL sans serveur Azure Synapse
- Leçon 4 : Sécuriser les données et gérer les utilisateurs dans les pools SQL sans serveur Azure Synapse

Module 5 : Explorez, transformez et chargez les données dans l'entrepôt de données à l'aide d'Apache Spark

- Leçon 1 : Comprendre l'ingénierie big data avec Apache Spark dans Azure Synapse Analytics
- Leçon 2 : Ingérer des données Apache Spark notebooks dans Azure Synapse Analytics
- Leçon 3 : Transformer les données avec DataFrames dans Apache Spark Pools dans Azure Synapse Analytics
- Leçon 4 : Intégrer les pools SQL et Apache Spark dans Azure Synapse Analytics

Module 6 : Exploration et transformation des données dans Azure Databricks

- Leçon 1 : Décrire Azure Databricks
- Leçon 2 : Lire et écrire des données dans Azure Databricks
- Leçon 3 : Travailler avec DataFrames dans Azure Databricks
- Leçon 4 : Travailler avec Dataframes méthodes avancées dans Azure Databricks

Module 7 : Ingérer et charger des données dans l'entrepôt de données

- Leçon 1 : Utiliser les meilleures pratiques de chargement de données dans Azure Synapse Analytics
- Leçon 2 : Ingestion à l'échelle des pétaoctets avec Azure Data Factory

Module 8 : Transformez les données avec Azure Data Factory ou Azure Synapse Pipelines

- Leçon 1 : Intégration de données avec Azure Data Factory ou Azure Synapse Pipelines
- Leçon 2 : Transformation sans code à grande échelle avec Azure Data Factory ou Azure Synapse Pipelines

Module 9 : Orchestrer le mouvement et la transformation des données dans Azure Synapse Pipelines

- Leçon 1 : Orchestrer le mouvement et la transformation des données dans Azure Data Factory

Module 10 : Optimisez les performances des requêtes avec des pools SQL dédiés dans Azure Synapse

- Leçon 1 : Optimiser les performances des requêtes d'entrepôt de données dans Azure Synapse Analytics
- Leçon 2 : Comprendre les fonctionnalités des développeurs d'entrepôts de données d'Azure Synapse Analytics

Module 11 : Analyser et optimiser le stockage des entrepôts de données

- Leçon 1 : Analyser et optimiser le stockage des entrepôts de données dans Azure Synapse Analytics

Module 12 : Supporter Hybrid Transactional Analytical Processing (HTAP) avec Azure Synapse Link

- Leçon 1 : Concevoir un traitement transactionnel et analytique hybride à l'aide d'Azure Synapse Analytics
- Leçon 2 : Configurer Azure Synapse Link avec Azure Cosmos DB
- Leçon 3 : Requête Azure Cosmos DB avec pools Apache Spark
- Leçon 4 : Requête Azure Cosmos DB avec pools SQL sans serveur

Module 13 : Sécurité de bout en bout avec Azure Synapse Analytics

- Leçon 1 : Sécuriser un entrepôt de données dans Azure Synapse Analytics
- Leçon 2 : Configurer et gérer les secrets dans Azure Key Vault
- Leçon 3 : Mettre en place des contrôles de conformité pour les données sensibles

Module 14 : Traitement en temps réel des flux avec Stream Analytics

- Leçon 1 : Activer des messageries fiables pour les applications Big Data à l'aide des Hubs d'événements Azure
- Leçon 2 : Travailler avec les flux de données en utilisant Azure Stream Analytics
- Leçon 3 : Ingérer des flux de données avec Azure Stream Analytics

Module 15 : Créer une solution de traitement des flux avec des concentrateurs d'événements et des databricks Azure

- Leçon 1 : Traiter les données de streaming avec le streaming structuré Azure Databricks

Module 16 : Créez des rapports à l'aide de l'intégration Power BI avec Azure Synapse Analytics

- Leçon 1 : Créer des rapports avec Power BI en utilisant son intégration avec Azure Synapse Analytics

Module 17 : Effectuer des processus intégrés d'apprentissage automatique dans Azure Synapse Analytics

- Leçon 1 : Utilisez le processus d'apprentissage automatique intégré dans Azure Synapse Analytics

Lab / Exercices

Laboratoires officiels Microsoft

- Lab 1 : Explorer les options de calcul et de stockage pour les charges de travail en ingénierie des données
- Lab 2 : Conception et mise en œuvre de la couche de service
- Lab 3 : Considérations d'ingénierie des données
- Lab 4 : Exécutez des requêtes interactives à l'aide de pools SQL sans serveur
- Lab 5 : Explorez, transformez et chargez les données dans l'entrepôt de données à l'aide d'Apache Spark
- Lab 4 : Exploration et transformation de données dans Azure Databricks
- Lab 5 : Ingérer et charger des données dans l'entrepôt de données
- Lab 6 : Transformer les données avec Azure Data Factory ou Azure Synapse Pipelines
- Lab 7 : Orchestrer le mouvement et la transformation des données dans Azure Synapse Pipelines
- Lab 8 : Optimisez les performances des requêtes avec des pools SQL dédiés dans Azure Synapse
- Lab 9 : Analyser et optimiser le stockage des entrepôts de données
- Lab 10 : Support Hybrid Transactional Analytical Processing (HTAP) avec Azure Synapse Link
- Lab 11 : Sécurité de bout en bout avec Azure Synapse Analytics
- Lab 12 : Traitement en temps réel des flux avec Stream Analytics

- Lab 13 : Créer une solution de traitement de flux avec Event Hubs et Azure Databricks
- Lab 14 : Créer des rapports à l'aide de l'intégration Power BI avec Azure Synapse Analytics
- Lab 15 : Effectuer des processus intégrés d'apprentissage automatique dans Azure Synapse Analytics

Documentation

- Accès à Microsoft Learn (contenu d'apprentissage en ligne)

Examen

- Ce cours prépare à la certification DP-203 : Data Engineering on Microsoft Azure. Si vous souhaitez passer cet examen, merci de contacter notre secrétariat qui vous communiquera son prix et s'occupera de toutes les démarches administratives nécessaires pour vous.

Profils des participants

- Les 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 plateforme de données existantes sur Microsoft Azure
- Les analystes de données et data scientists qui travaillent avec des solutions analytiques basées sur Microsoft Azure

Connaissances Préalables

- Avoir une connaissance de cloud computing et des concepts de données de base
- Avoir une expérience professionnelle avec des solutions de données
- Avoir suivi les formations ci-dessous : [Microsoft Azure Fundamentals \(AZ-900\)](#) et [Microsoft Azure Data Fundamentals \(DP-900\)](#)

Objectifs

- Explorer les options de calcul et de stockage pour les charges de travail d'ingénierie des données dans Azure
- Concevoir et implémenter la couche de service
- Comprendre les considérations liées à l'ingénierie des données
- Exécuter des requêtes interactives à l'aide de pools SQL sans serveur
- Explorer, transformer et charger les données dans l'entrepôt de données à l'aide d'Apache Spark
- Effectuer l'exploration et la transformation de données dans Azure Databricks
- Ingérer et charger des données dans l'entrepôt de données
- Transformer les données avec Azure Data Factory ou Azure Synapse Pipelines
- Intégrer les données des ordinateurs portables à Azure Data Factory ou Azure Synapse Pipelines
- Optimiser les performances des requêtes avec des pools SQL dédiés dans Azure Synapse
- Analyser et optimiser le stockage des entrepôts de données
- Supporter Hybrid Transactional Analytical Processing (HTAP) avec Azure Synapse Link
- Effectuer la sécurité de bout en bout avec Azure Synapse Analytics
- Effectuer le traitement en temps réel des flux avec Stream Analytics
- Créer une solution de traitement de flux avec des hubs d'événements et des databricks Azure
- Créer des rapports à l'aide de l'intégration Power BI avec Azure Synapse Analytics
- Effectuer des processus intégrés d'apprentissage automatique dans Azure Synapse Analytics

Description

Data Engineering on Microsoft Azure (DP-203)

Niveau

Intermédiaire

Prix de l'inscription en Présentiel (CHF)

3200

Prix de l'inscription en Virtuel (CHF)

3000

Durée (Nombre de Jours)

4

Reference

DP-203T00