

DP-203T00 Data Engineering on Microsoft Azure

Czas trwania: 4 dni

Dla kogo:

Szkolenie skierowane do analityków biznesowych, architektów danych oraz inżynierów danych, którzy chcą poznać technologie wykorzystywane w Azure przy tworzeniu i przechowywaniu danych, ich analizie oraz zapewnieniu bezpieczeństwa.

Wymagania wstępne:

- podstawy platformy Azure
- zrozumienie technologii magazynu Azure
- znajomość języka C # lub Python

Szkolenie przygotowuje do egzaminu:

DP-203: Data Engineering on Microsoft Azure

Zakres tematyczny:

1. Zasoby obliczeniowe i magazynowanie danych
 - wprowadzenie do Azure Synapse Analytics
 - omówienie Azure Databricks
 - magazyn Azure Data Lake - wprowadzenie
 - architektura Data Lake
2. Zapytania interakcyjne przy użyciu bezserwerowych pul SQL Azure Synapse Analytics
 - możliwości bez serwerowych pul SQL Azure Synapse Analytics
 - zapytania Data Lake przy użyciu bezserwerowych pul SQL Azure Synapse Analytics
 - metadane w bezserwerowych pulach SQL Azure Synapse Analytics
 - bezpieczeństwo i zarządzanie użytkownikami w bezserwerowych pulach SQL Azure Synapse Analytics
3. Eksploracja i transformacja danych w Azure Databricks
 - omówienie Azure Databricks
 - odczyt i zapis danych w Azure Databricks
 - wykorzystanie DataFrames w Azure Databricks
 - zaawansowane metody wykorzystania DataFrames w Azure Databricks
4. Eksploracja, przekształcanie danych oraz zasilanie hurtowni danych za pomocą Apache Spark
 - omówienie Apache Spark w usłudze Azure Synapse Analytics
 - pozyskiwanie danych za pomocą notesów Apache Spark w usłudze Azure Synapse Analytics
 - transformowanie danych z DataFrames w Apache Spark Pools
 - integracja SQL i pul Apache Spark w Azure Synapse Analytics
5. Hurtownie danych - zasilanie oraz pozyskiwanie informacji
 - zasilanie danymi - dobre praktyki w Azure Synapse Analytics
 - pozyskiwanie petabajtowych ilości danych przy pomocy Azure Data Factory
6. Transformowanie danych w Azure Data Factory oraz Azure Synapse Pipelines
 - integrowanie danych w Azure Data Factory oraz Azure Synapse Pipelines
 - bezkodowe transformacje danych w Azure Data Factory oraz Azure Synapse Pipelines
7. Organizacja przenoszenia i transformacji danych w Azure Synapse Pipelines Azure Synapse Analytics - bezpieczeństwo
 - zabezpieczanie magazynu danych w usłudze Azure Synapse Analytics
 - konfigurowanie wpisów tajnych i zarządzanie nimi w Azure Key Vault
 - wdrożenie kontroli zgodności z wymaganiami dla danych wrażliwych

8. Azure Synapse Link - Obsługa hybrydowego przetwarzania transakcyjnego i analitycznego (Hybrid Transactional Analytical Processing - HTAP)
 - projektowanie hybrydowego przetwarzania transakcyjnego i analitycznego w Azure Synapse Analytics
 - konfigurowanie Azure Synapse Link z Azure Cosmos DB
 - pule Apache Spark - zapytania do Azure Cosmos DB
 - bezserwerowe pule SQL - zapytania do Azure Cosmos DB
9. Przetwarzanie strumieni w czasie rzeczywistym za pomocą Stream Analytics
 - przesyłanie wiadomości dla aplikacji Big Data za pomocą usługi Azure Event Hubs
 - Azure Stream Analytics - strumień danych
 - Azure Stream Analytics - pozyskiwanie strumieni danych
10. Tworzenie rozwiązań przetwarzania strumieniowego za pomocą Event Hubs i Azure Databricks
 - przetwarzanie danych strumieniowych za pomocą ustrukturyzowanego przesyłania strumieniowego Azure Databricks