

Implementing a Lakehouse with Microsoft Fabric

Szkolenie Implementing a Lakehouse with Microsoft Fabric wprowadza uczestników w koncepcję Lakehouse – nowoczesnego podejścia do zarządzania i analizy danych, łączącego elastyczność jezior danych z możliwościami klasycznego magazynu danych.

Podczas zajęć uczestnicy poznają możliwości Microsoft Fabric w zakresie pozyskiwania, przetwarzania i organizowania danych z wykorzystaniem Apache Spark, Delta Lake, Dataflows Gen2 oraz Data Factory. Szkolenie skupia się na praktycznych aspektach inżynierii danych, takich jak wersjonowanie, orkiestracja przepływów danych i budowa skalowalnych środowisk analitycznych.

Dobierz formułę

szkolenia: <https://www.altkomakademia.pl/dobierz-formule-szkolenia-do-celu-poziomu-i-efektu-ktory-chcesz-osiagnac/>



Odbiorcy szkolenia

Szkolenie jest przeznaczone dla specjalistów zajmujących się danymi, którzy chcą poznać możliwości Microsoft Fabric w zakresie przetwarzania i integracji danych w architekturze Lakehouse.

W szczególności dedykowane jest dla:

- inżynierów danych,
- analityków danych pracujących z dużymi zbiorami danych,
- specjalistów Power BI, którzy chcą rozszerzyć swoje kompetencje o warstwę inżynierii danych,
- architektów danych i rozwiązań analitycznych,
- członków zespołów projektujących środowiska Lakehouse w chmurze Microsoft.



Korzyści

- Zrozumienie koncepcji Lakehouse – poznasz strukturę i zastosowania architektury Lakehouse w środowisku Microsoft Fabric.
- Wykorzystanie Apache Spark – nauczysz się przetwarzać, transformować i analizować dane rozproszone przy użyciu Spark DataFrames i Spark SQL.
- Praca z Delta Lake – dowiesz się, jak tworzyć i zarządzać tabelami Delta Lake, korzystać z wersjonowania danych i funkcji „time travel”.
- Pozyskiwanie i integracja danych – opanujesz użycie przepływów danych (Dataflows Gen2) oraz potoków Data Factory do wprowadzania i przetwarzania danych.
- Tworzenie nowoczesnych rozwiązań analitycznych – nauczysz się budować zintegrowane, wydajne środowiska danych w Microsoft Fabric, gotowe do dalszej analizy w Power BI lub innych narzędziach.



Program szkolenia

1. Wprowadzenie do analizy danych przy użyciu Microsoft Fabric
2. Rozpoczęcie pracy z lakehouse w Microsoft Fabric
 - Opis podstawowych funkcji i możliwości lakehouse w Microsoft Fabric.
 - Tworzenie lakehouse.
 - Wprowadzanie danych do plików i tabel w lakehouse.
 - Zapytania do tabel lakehouse za pomocą SQL.
3. Korzystanie z Apache Sparka w Microsoft Fabric
 - Konfiguracja Sparka w przestrzeni roboczej Microsoft Fabric.
 - Identyfikacja odpowiednich scenariuszy dla notatników Sparka i zadań Sparka.
 - Użycie ramek danych Sparka do analizy i przekształcania danych.
 - Użycie Spark SQL do zapytań danych w tabelach i widokach.
 - Wizualizacja danych w notatniku Sparka.
4. Praca z tabelami Delta Lake w Microsoft Fabric
 - Zrozumienie technologii Delta Lake i tabel delta w Microsoft Fabric.
 - Tworzenie i zarządzanie tabelami delta za pomocą Sparka.
 - Użycie Sparka do zapytań i przekształcania danych w tabelach delta.
 - Wykorzystywanie tabel delta w strumieniowaniu z użyciem Sparka.
5. Wprowadzanie danych za pomocą Dataflows Gen2 w Microsoft Fabric
 - Opis możliwości Dataflow w Microsoft Fabric.
 - Tworzenie rozwiązań Dataflow do wprowadzania i przekształcania danych.
 - Włączanie Dataflow do potoku przetwarzania.
6. Korzystanie z potoków Data Factory w Microsoft Fabric
 - Opis możliwości potoków w Microsoft Fabric.
 - Użycie aktywności kopiowania danych w potoku.
 - Tworzenie potoków na podstawie predefiniowanych szablonów.
 - Uruchamianie i monitorowanie potoków.

7. Tworzenie architektury medalionowej w Microsoft Fabric

- Opis architektury medalionowej.
- Użycie architektury medalionowej.
- Zarządzanie architekturą medalionową w MS Fabric.



Oczekiwane przygotowanie uczestnika

- Znajomość podstawowych pojęć i usług związanych bazami danych
- Znajomość podstaw języków Python i SQL (a przynajmniej jednego z nich).
- Ukończenie ścieżki szkoleniowej DP-900

Potrzebna będzie umiejętność korzystania z angielskojęzycznej dokumentacji (w tym opis laboratoriów).

Kurs DP-601 należy do ścieżek:

- Data Analyst
- Data Engineer
- Data Scientist

Dla zwiększenia komfortu pracy oraz efektywności szkolenia zalecamy skorzystanie z dodatkowego ekranu. Brak dodatkowego ekranu nie jest przeciwwskazaniem do udziału w szkoleniu, ale w znaczący sposób wpływa na komfort pracy podczas zajęć.

Informacje oraz wymagania dotyczące uczestniczenia w szkoleniach w formule zdalnej dostępne na:

<https://www.altkomakademia.pl/distance-learning/#FAQ>



Szkolenie obejmuje

* podręcznik w formie elektronicznej dostępny na platformie: <https://learn.microsoft.com/pl-pl/training/>

* dostęp do portalu słuchacza AltKom Akademii

Produkt zawiera:

* wykład i prezentacja produktu (50%)

* ćwiczenia (50%)

Główne narzędzia dydaktyczne obejmują prezentacje PowerPoint, demonstracje na środowisku testowym, praktyczne środowiska laboratoryjne i zasoby Microsoft Learn



Język

- Szkolenie: polski
- Materiały: angielski

Czas trwania

1 dni / 7 godzin

Metoda egzaminacyjna