

# Implement data engineering solutions using Azure Databricks

Opanuj data engineering z Azure Databricks i Unity Catalog. Ten kurs przechodzi od podstawowego ustawienia do wdrażania produkcyjnego, obejmując konfigurację środowiska i zarządzanie na poziomie korporacyjnym. Nauczysz się budować potoki danych, wdrażać zabezpieczenia w Unity Catalog i wdrażać zoptymalizowane obciążenia. Na koniec będziesz posiadać praktyczne umiejętności, aby wdrożyć, zabezpieczyć i utrzymać skalowalne rozwiązania Lakehouse, które spełniają rygorystyczne wymagania przedsiębiorstwa.



## Odbiorcy szkolenia

Szkolenie przeznaczone jest dla:

- Inżynierów danych



## Korzyści

Zdobyte umiejętności obejmują:

- Nauczysz się, jak tworzyć i optymalizować obiekty bazy danych dla nowoczesnych platform SQL.
- Nauczysz się projektować inteligentne wyszukiwanie z wykorzystaniem pełnotekstowego i wektorowego wyszukiwania, integrować modele AI i osadzenia oraz budować rozwiązania Retrieval Augmented Generation (RAG) w całości w T-SQL.
- Przyspieszysz także swój proces tworzenia dzięki narzędziom wspomaganym AI, takim jak GitHub Copilot i Fabric Copilot.



## Program szkolenia

1. Konfiguracja środowiska Azure Databricks
  - Poznaj Azure Databricks
  - Architektura Azure Databricks
  - Integracja Azure Databricks
  - Konfiguracja zasobów obliczeniowych Azure Databricks
  - Twórz i organizuj obiekty w Unity Catalog
2. Zabezpieczaj i zarządzaj obiektami w Unity Catalog w Azure Databricks
  - Bezpieczeństwo obiektów katalogu Unity
  - Zarządzanie obiektami Unity Catalog
3. Przygotowanie i przetwarzanie danych za pomocą Azure Databricks
  - Projektuj i wdrażaj modelowanie danych za pomocą Azure Databricks
  - Pobieranie danych do Unity Catalog
  - Oczyszczanie, przekształcanie i ładowanie danych do Unity Catalog
  - Implementuj i zarządzaj wymogami jakości danych za pomocą Azure Databricks
4. Wdrażaj i utrzymuj potoki danych oraz obciążenia z Azure Databricks
  - Projektuj i implementuj potoki danych za pomocą Azure Databricks
  - Implementuj Lakeflow Jobs z Azure Databricks
  - Implementuj procesy cyklu życia i tworzenia w Azure Databricks
  - Monitoruj, rozwiążuj problemy i optymalizuj obciążenia w Azure Databricks



## Oczekiwane przygotowanie uczestnika

- Grupą docelową są inżynierowie danych, którzy mają podstawową wiedzę na temat koncepcji analizy danych, podstawową wiedzę o przechowywaniu w chmurze oraz zasady organizacji danych. Powinni czuć się pewnie w pracy z SQL i mieć doświadczenie w Pythonie, w tym z notebookami, do zadań związanych z data engineering.
- Oczekuje się, że uczestnicy będą dobrze rozumieć przestrzenie robocze Azure Databricks i Unity Catalog, a także znać wzorce dostępu do danych oraz podstawowe koncepcje inżynierii danych i hurtowni danych. Ponadto powinni mieć podstawową wiedzę z zakresu bezpieczeństwa Azure, w tym Microsoft Entra ID, oraz znać podstawy kontroli wersji Git-a.



## Szkolenie obejmuje

- \* podręcznik w formie elektronicznej dostępny na platformie: <https://learn.microsoft.com/pl-pl/training/>
- \* dostęp do portalu słuchacza Altkom Akademii



Czas trwania

4 dni / 28 godzin

## Język

- Wykład: polski
- Materiały: angielski