

# Databricks Fundamentals

Szkolenie Databricks Fundamentals to pierwszy krok w spójnej ścieżce szkoleniowej Fundamental → Explorer → Lakehouse → Transformation. Uczestnicy poznają środowisko Databricks, nauczą się poruszać po platformie, korzystać z podstawowych narzędzi oraz wykonywać proste operacje na danych. Szkolenie obejmuje także elementy zarządzania kosztami, co pozwala od początku zrozumieć, jak efektywnie pracować w chmurze.



## Odbiorcy szkolenia

Szkolenie przeznaczone jest dla osób rozpoczynających pracę z platformą Databricks: analityków danych, inżynierów danych, specjalistów BI oraz osób technicznych, które chcą poznać podstawowe funkcje środowiska.



## Korzyści

Korzyści dla uczestników

- poznają podstawowe elementy środowiska Databricks
- zrozumieją rolę Unity Catalog i Delta Lake jako fundamentów pracy w Databricks
- potrafią wczytywać dane, wykonywać proste transformacje i eksploracje
- umieją korzystać z Workflows do prostych automatyzacji
- mają świadomość podstawowych zasad zarządzania kosztami w chmurze
- są przygotowani do kolejnego etapu ścieżki - szkolenia Databricks Explorer



## Program szkolenia

1. Poznanie platformy Databricks

- Workspace, notebooki, klastry, DBFS
  - Podstawowe pojęcia i architektura platformy
  - Unity Catalog – katalogi, schematy, podstawy uprawnień
  - Delta Lake jako format domyślny
2. Cost management
- Typy klastrów: Jobs vs All-purpose
  - Autoscaling i wpływ na koszty
  - Podstawowe zasady optymalnego korzystania z zasobów
3. GUI i narzędzia
- Interfejs Databricks – obsługa notebooków, markdown, miks SQL i Python
  - Tworzenie prostych wizualizacji w GUI
  - Asystent AI w notebookach – jak wspiera użytkownika
4. Proste operacje na danych
- Wczytywanie danych z plików CSV i JSON
  - Eksploracja danych: display(), show(), describe(), summary()
  - Podstawowe transformacje w PySpark (withColumn, when, cast)
5. Workflows i automatyzacja
- Tworzenie prostych zadań w Workflows GUI
  - Parametryzacja, harmonogramy, uruchamianie cykliczne
6. Projekt końcowy
- Załadowanie danych do Delta Lake, wykonanie prostych transformacji i uruchomienie ich w zaplanowanym workflow



## Oczekiwane przygotowanie uczestnika

- Podstawowa znajomość SQL
- Ogólna znajomość koncepcji baz danych i analizy danych
- Podstawowa wiedza o chmurze (opcjonalnie)



## Szkolenie obejmuje

\* dostęp do portalu słuchacza Altkom Akademii

Metoda szkolenia:

Szkolenie prowadzone jest w środowisku Databricks w chmurze. Każdy uczestnik otrzymuje własny workspace z dostępem do Unity Catalog, SQL Editor, Notebooków oraz katalogu z danymi testowymi.

Uwaga: W trakcie wybranych szkoleń Akademia zapewnia Uczestnikom dostęp do płatnych narzędzi i usług niezbędnych do realizacji programu szkolenia (w szczególności usług Azure, Fabric oraz innych

zasobów chmurowych). Uczestnik zobowiązany jest do korzystania z udostępnionych zasobów wyłącznie w zakresie wynikającym z agendy szkolenia oraz zgodnie z poleceniami trenera. Wykorzystywanie tych zasobów w sposób wykraczający poza zakres szkoleniowy, w szczególności do realizacji projektów prywatnych, testowania dodatkowych usług, modyfikowania konfiguracji środowiska lub podejmowania działań generujących koszty niezwiązane z realizacją ćwiczeń – skutkować będzie obciążeniem Uczestnika kosztami przekraczającymi przewidziany limit, na podstawie wydatków poniesionych przez Akademię (<https://www.altkomakademia.pl/ogolne-warunki-uczestnictwa-w-szkoleniach/>)



## Język

Wykład: polski

Materiały: angielski

## Metoda egzaminacyjna

Szkolenie nie zawiera formalnego egzaminu. Zaleca się jednak podejście do egzaminu certyfikacyjnego Databricks Certified Data Engineer Associate jako kolejny krok rozwojowy.

## Czas trwania

1 dni / 7 godzin

## Opis egzaminu

Szkolenie nie zawiera formalnego egzaminu. Zaleca się jednak podejście do egzaminu certyfikacyjnego Databricks Certified Data Engineer Associate jako kolejny krok rozwojowy.