

Opleiding : Databricks voor data engineers

Praktijkcursus - 3d - 21u00 - Ref. DKI

Prijs : 2030 € V.B.

NEW

In drie dagen biedt deze opleiding data engineers de sleutels om het volledige potentieel van Azure Databricks te benutten binnen een moderne cloudomgeving. Van het aanmaken en beheren van notebooks tot het orkestreren van workflows, inclusief het gebruik van Python, SQL en Datastore. De opleiding legt de nadruk op prestaties, automatisering en datavalorisatie ter ondersteuning van grootschalige, geïndustrialiseerde dataprojecten.

Pedagogische doelstellingen

Aan het einde van de training is de deelnemer in staat om:

- ✓ De verschillende functies van Databricks begrijpen en gebruiken
- ✓ Notebooks maken en manipuleren
- ✓ Taken instellen
- ✓ Python en SQL gebruiken met Databricks
- ✓ De Databricks Datastore begrijpen en gebruiken
- ✓ Dashboards maken met SQL Analytics
- ✓ Databricks en Power BI verbinden

Doelgroep

Data engineers, medewerkers die betrokken zijn bij het opzetten, automatiseren en industrialiseren van datapijplijnen op moderne cloudplatforms.

Voorafgaande vereisten

Goede kennis van SQL en gegevensmanipulatie. Basiskennis van Python.

Praktische modaliteiten

Praktisch werk

Discussies, oefeningen en praktisch werk.

Leer methodes

Actief

Opleidingsprogramma

DEELNEMERS

Data engineers, medewerkers die betrokken zijn bij het opzetten, automatiseren en industrialiseren van datapijplijnen op moderne cloudplatforms.

VOORAFGAANDE VEREISTEN

Goede kennis van SQL en gegevensmanipulatie. Basiskennis van Python.

VAARDIGHEDEN VAN DE CURSUSLEIDER

De deskundigen die de cursus leiden zijn specialisten op het betreffende vakgebied. Zij werden geselecteerd door onze pedagogische teams zowel om hun vakkennis als hun pedagogische vaardigheden voor elke cursus die zij geven. Zij hebben minstens vijf tot tien jaar ervaring in hun vakgebied en oefenen of oefenden verantwoordelijke bedrijfsfuncties uit.

BEOORDELINGSMODALITEITEN

De cursusleider beoordeelt de pedagogische vooruitgang van de deelnemer gedurende de gehele cursus aan de hand van meerkeuzevragen, praktijksituaties, praktische opdrachten, ...
De deelnemer legt ook van tevoren en naderhand een test af ter bevestiging van de verworven kennis.

1 Inleiding

- Azure: de belangrijkste principes van de cloud, presentatie en voordelen

2 Databricks: overzicht

- Presentatie van Azure Databricks: concepten (data lakehouse, ETL, datawarehouse)
- Gegevensbeheer: concepten van Unity Catalog
- Voordelen van Databricks

3 Werkplekken

- Azure Databricks bronnen
- Een werkruimte maken
- De werkruimte verbinden met externe gegevensbronnen

Praktisch werk

Maak een workspace en verbind deze met externe gegevensbronnen.

4 Verschillende soorten bronnen: clusters en SQL-magazijn

- Presentatie
- De verschillende soorten berekeningen
- Het creëren van een cluster
- Beheer van clusters

Praktisch werk

De verschillende soorten berekeningen identificeren, een cluster maken en beheren.

5 Eenheid Catalogus

- Een catalogus maken
- Verbinding maken met gegevens uit Unity Catalog
- Een tabel toevoegen aan het catalogusschema
- Autorisaties definiëren voor een catalogustabel
- Focus: verbinding maken met SQL-databases
- Andere manieren om gegevens aan de catalogus toe te voegen

Praktisch werk

Een catalogus maken en er gegevens aan koppelen. Voeg een tabel toe, definieer autorisaties en leer hoe je verbinding maakt met SQL databases.

6 Databricks: notatieblokken

- De doelstellingen en voordelen van notebooks
- Het celprincipe
- Bestellingen
- Magische commando's
- PySpark programmeren, SQL...
- Cellen gebruiken en resultaten weergeven
- Notebooks exporteren en delen

Praktisch werk

Databricks verbinden met opslagtools (ADLS, SQL database, API), gegevens bevragen, transformeren en opslaan in Databricks. De verschillende mogelijke formaten begrijpen (Parquet, Delta, externe tabel, enz.).

PEDAGOGISCHE EN TECHNISCHE MIDDELEN

- De gebruikte pedagogische middelen en cursusmethoden zijn voornamelijk: audiovisuele hulpmiddelen, documentatie en cursusmateriaal, praktische oefeningen en correcties van de oefeningen voor praktijkstages, casestudies of reële voorbeelden voor de seminars.
- Na afloop van de stages of seminars verstrekt ORSYS de deelnemers een evaluatievragenlijst over de cursus die vervolgens door onze pedagogische teams wordt geanalyseerd.
- Na afloop van de cursus wordt een presentielijst per halve dag verstrekt, evenals een verklaring van de afronding van de cursus indien de stagiair alle sessies heeft bijgewoond.

TOEGANGSMODALITEITEN EN TERMIJNEN

De inschrijving dient 24 uur voor aanvang van de cursus plaatsgevonden te hebben.

TOEGANKELIJKHEID VOOR MINDERVALIDEN

Is voor u speciale toegankelijkheid vereist? Neem contact op met mevr. FOSSE, contactpersoon voor mindervaliden, via het adres psh-accueil@ORSYS.fr om uw verzoek en de haalbaarheid daarvan zo goed mogelijk te bestuderen.

7 Principe van vonk en delta

- Introductie en kennismaking met Spark SQL
- Deltameer concepten
- Beheer levenscyclus tabel: delta samenvoegen, verwijderen/invoegen

8 Procesorkestratie

- Notebooks oproepen vanuit Azure data factory en configuratie
- Databricks workflows gebruiken
- Databricks notebooks opzetten en configureren met behulp van een workflow
- Inleiding tot Databricks Delta Live Tabellen

Praktisch werk

Omgaan met Databricks workflows, Databricks notebooks opzetten en configureren vanuit een workflow.

9 SQL Databricks

- Databricks SQL: beheer
- SQL-editor
- Een verzoek maken en uitvoeren
- Een verzoek plannen
- Waarschuwingen maken
- Queryresultaten delen
- De beschikbare opties voor queryresultaten: exporteren, weergeven, weergavefilters, resultaatparameters, enz.

Praktisch werk

Maak een query en deel de resultaten.

Data en plaats

PARIS LA DÉFENSE

2026 : 30 maa., 29 juni, 28 sep., 16 dec.