

Úvod do Microsoft Fabric

Kód kurzu: GOC680

Kurz je určený dátovým profesionálom, ktorí sa chcú zoznámiť s platformou Microsoft Fabric, ktorá ponúka moderné, jednotné a plne integrované prostredie pre zber, uloženie a následnú analýzu vašich dát. Na školení pochopíte kľúčové princípy platformy Microsoft Fabric a vyskúšate si na praktických cvičeniach celú škálu podporovaných spôsobov použitia Microsoft Fabric tak, aby ste si vytvorili ucelenú predstavu o možnostiach tejto platformy a mohli ste následne absolvovať nadväzujúce školenie špecializované na konkrétnu oblasť, ktorá bude preberaná veľmi do hĺbky na rozdiel od úvodného predstavenia na tomto školení. Vyskúšate si všetko od uloženia a dotazovania dát v Data Lake či Data Warehouse s využitím jazyka SQL či Python cez implementáciu extrakcie a načítanie dát s využitím Dataflows Gen2, Pipelines a Notebooks až po tvorbu sémantického modelu a základných reportov v Power BI, ktoré sa dokáže po novom pripojiť priamo na dáta uložené v Data Lake. Stranou nezostanú ani ukážky spracovania real-time dát, použitie strojového učenia a súvisiace možnosti zabezpečenia dát.

Čo Vás naučíme

- Zoznámite sa s prostredím Microsoft Fabric
- Naučíte sa ukladať dáta v úložiskách Lakehouse a Data Warehouse a zabezpečiť k nim prístupy
- Vyskúšate si implementovať základné ETL procesy pre načítanie a transformáciu dát
- Zistíte, ako dotazovať a analyzovať uložené dáta pomocou jazyka T-SQL či Apache Spark
- Vyskúšate si prezentovať uložené dáta za pomoci reportov v Power BI
- Dozviete sa, ako vytvárať, trénovať a skórovať modely strojového učenia
- Pochopíte, aké máte možnosti spracovania dát v reálnom čase

Pre koho je kurz určený

- Dátovým inžinierom, ktorí sa chcú naučiť navrhovať úložiska dát a súvisiace ETL/ELT procesy v prostredí služby Microsoft Fabric
- Dátovým analytikom, ktorí budú navrhovať dátové modely v službe Microsoft Fabric za účelom následnej analýzy a vizualizácie uložených dát v Microsoft Power BI
- Dátovým a BI architektom, ktorí chcú pochopiť, ako navrhovať architektúru moderného riešenia pre analýzu dát postaveného na službe Microsoft Fabric

Požadované vstupné znalosti

- Znalosti základných princípov návrhu dátových skladov a implementácie ETL procesov na úrovni kurzu GOC645
- Základná znalosť jazyka SQL na úrovni kurzu GOC277, pokročilá znalosť jazyka T-SQL na úrovni kurzu GOC278 výhodou
- Odporúčaná základná znalosť dátových služieb v Microsoft Azure na úrovni kurzu MOC DP-900
- Odporúčaná základná znalosť analýzy dát v jazyku Python/Apache Spark

Osnova kurzu

1 Predstavenie služby Microsoft Fabric

- Zoznámite sa s kľúčovými konceptmi, službami a komponentmi, ktoré ponúka platforma Microsoft Fabric
- Dozviete sa, aké role v tíme pracujú s ktorými službami v Microsoft Fabric
- Zistíte, aké sú možnosti správy a zabezpečenia tohto prostredia
- Pochopíte, ako funguje pridelenie výpočtovej kapacity a aké sú možnosti nákupu tejto služby

2 Uloženie dát v Data Lake

- Zoznámite sa s úložiskom OneLake a jeho kľúčovými princípmi
- Pochopíte rozdiely medzi Data Lake a Data Warehouse
- Naučíte sa vytvárať vlastný Lakehouse a načítateľ doňho dáta
- Naučíte sa pracovať s Delta Lake tabuľkami
- Dozviete sa, ako zapisovať dáta do Lakehouse a aké máte možnosti ich analýzy
- Zistíte, ako využívať shortcuts, aby nedochádzalo k duplicitnému uloženiu dát

GOPAS Praha

Kodaňská 1441/46
101 00 Praha 10
Tel.: +420 234 064 900-3
info@gopas.cz

GOPAS Brno

Nové sady 996/25
602 00 Brno
Tel.: +420 542 422 111
info@gopas.cz

GOPAS Bratislava

Dr. Vladimíra Clementisa 10
Bratislava, 821 02
Tel.: +421 248 282 701-2
info@gopas.sk



Copyright © 2020 GOPAS, a.s.,
All rights reserved

Úvod do Microsoft Fabric

- Dozviete sa, ako vhodne organizovať Lakehouse s využitím medallion architektúry, aby bola zaistená zodpovedajúca kvalita uložených dát

3 Načítanie dát do Microsoft Fabric

- Zoznámite sa s možnosťami použitia Data Factory v prostredí Microsoft Fabric
- Pochopíte princípy Dataflow Gen2
- Naučíte sa vytvoriť základné Dataflow pre načítanie dát
- Zoznámite sa s možnosťami tvorby Pipelines pre orchestráciu procesu načítania a transformácie dát
- Naučíte sa pripojiť k externým dátovým zdrojom a načítať dáta pomocou Apache Spark a notebookov
- Dozviete sa, ako efektívne ukladať dáta do Lakehouse a Delta tabuliek

4 Uloženie dát v Data Warehouse

- Zoznámite sa s kľúčovými vlastnosťami dátových skladov v Microsoft Fabric
- Dozviete sa základné princípy návrhu dátových skladov, tabuliek dimenzií a faktov
- Naučíte sa definovať vzťahy medzi tabuľkami propagované do sémantického modelu
- Dozviete sa, aké máte možnosti pre načítanie dát do tabuliek dimenzií a tabuliek faktov
- Vyskúšate si načítať dáta do dátového skladu
- Naučíte sa nad dátovým skladoom spúšťať dotazy v jazyku T-SQL
- Dozviete sa, ako pracovať s Visual Query Editorom
- Vyskúšate si pripojiť sa cez SQL Server Management Studio
- Dozviete sa, ako monitorovať bežiacie dotazy pomocou Dynamic Management Views a ako sledovať časté a náročné dotazy pomocou Query Insights
- Naučíte sa nastaviť oprávnenia v jazyku T-SQL na úrovni objektov dátového skladu

5 Analýza a vizualizácia dát v Microsoft Fabric

- Zoznámite sa s rôznymi možnosťami analýzy uložených dát
- Zistíte, aké máte možnosti analýzy dát s využitím jazyka SQL
- Pochopíte, ako vytvárať Notebooky za účelom analýzy a vizualizácie dát s využitím Apache Spark
- Dozviete sa, ako definovať nad dátami sémantické modely, správne previazať dáta a využiť jazyk DAX pre výpočty nad dátami
- Objavíte rôzne možnosti vizualizácie dát a tvorby reportov
- Vyskúšate si vytvoriť vlastné reporty nad dátami s využitím Power BI

6 Strojové učenie v Microsoft Fabric

- Dozviete sa, aké máte možnosti tvorby, tréningu a skórovania modelov strojového učenia v prostredí Microsoft Fabric
- Vyskúšate si vytvoriť jednoduchý model strojového učenia a vytrénovať ho nad dátami uloženými v Data Lake

7 Uloženie a spracovanie real-time dát v Microsoft Fabric

- Zoznámite sa s možnosťami spracovania udalostí v reálnom čase v Microsoft Fabric
- Vyskúšate si pracovať s Real-Time Hubmi, ukladať dáta do úložiska Eventhouse a dotazovať ich jazykom KQL
- Zistíte, ako vizualizovať dáta v reálnom čase

8 Automatizácia upozornení nad dátami

- Zoznámite sa so službou Data Activator
- Naučíte sa automatizovať detekcie prekročenia hraníc a vzorov v dátach za účelom generovania upozornení a spúšťania vlastných akcií