

Introduction to Microsoft Fabric

Course code: GOC680

Kurz je určen pro datové profesionály, kteří se chtějí seznámit s platformou Microsoft Fabric, která nabízí moderní, jednotné a plně integrované prostředí pro sběr, uložení a následnou analýzu Vašich dat. Na školení pochopíte klíčové principy platformy Microsoft Fabric a vyzkoušíte si na praktických cvičeních celou šíři podporovaných způsobů použití Microsoft Fabric, tak, abyste si udělali ucelenou představu o možnostech této platformy a mohli jste následně absolvovat navazující školení specializované na konkrétní oblast, kde bude daná oblast probrána velmi do hloubky, na rozdíl od úvodního představení v tomto školení. Vyzkoušíte si od uložení a dotazování dat v Data Lake či Data Warehouse s využitím jazyka SQL či Python, přes implementaci extrakce a načítání dat s využitím Dataflows Gen2, Pipelines a Notebooks, až po tvorbu sémantického modelu a základních reportů v Power BI, které se dokáže nově připojit přímo na data uložená v Data Lake. Stranou nezůstanou ani ukázky zpracování real-time dat, použití strojového učení a související možnosti zabezpečení dat.

Co Vás naučíme

- Seznámíte se s prostředím Microsoft Fabric
- Naučíte se ukládat data v úložištích Lakehouse a Data Warehouse a zabezpečit k nim přístupy
- Vyzkoušíte si implementovat základní ETL procesy pro načtení a transformaci dat
- Zjistíte, jak dotazovat a analyzovat uložená data s pomocí jazyka T-SQL či Apache Spark
- Vyzkoušíte si prezentovat uložená data za pomoci reportů v Power BI
- Dozvíte se, jak vytvářet, trénovat a skórovat modely strojového učení
- Pochopíte, jaké máte možnosti zpracování dat v reálném čase

Pro koho je kurz určen

- Datovým inženýrům, kteří se chtějí naučit navrhovat úložiště dat a související ETL/ELT procesy v prostředí služby Microsoft Fabric.
- Datovým analytikům, kteří budou navrhovat datové modely ve službě Microsoft Fabric za účelem následné analýzy a vizualizace uložených dat v Microsoft Power BI.
- Datovým a BI architektům, kteří chtějí pochopit, jak navrhovat architekturu moderního řešení pro analýzu dat postaveného na službě Microsoft Fabric.

Požadované vstupní znalosti

- Znalosti základních principů návrhu datových skladů a implementace ETL procesů na úrovni kurzu GOC645.
- Základní znalost jazyka SQL na úrovni kurzu GOC277, pokročilá znalost jazyka T-SQL na úrovni kurzu GOC278 výhodou
- Doporučena základní znalost datových služeb v Microsoft Azure na úrovni kurzu MOC DP-900.
- Doporučena základní znalost analýzy dat v jazyce Python / Apache Spark

Osnova kurzu

1 Představení služby Microsoft Fabric

- Seznámíte se s klíčovými koncepty, službami a komponentami, které nabízí platforma Microsoft Fabric
- Dozvíte se, jaké role v týmu pracují s kterými službami v Microsoft Fabric
- Zjistíte, jaké jsou možnosti správy a zabezpečení tohoto prostředí
- Pochopíte, jak funguje přidělení výpočetní kapacity a jaké jsou možnosti nákupu této služby

2 Uložení dat v Data Lake

- Seznámíte se s úložištěm OneLake a jeho klíčovými principy
- Pochopíte rozdíly mezi Data Lake a Data Warehouse
- Naučíte se vytvářet vlastní Lakehouse a načítat do něj data
- Naučíte se pracovat s Delta Lake tabulkami
- Dozvíte se, jak zapisovat data do Lakehouse a jaké máte možnosti jejich analýzy
- Zjistíte, jak využívat shortcuts, aby nedocházelo k duplicitnímu uložení dat

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

Introduction to Microsoft Fabric

- Dozvíte se, jak vhodně organizovat Lakehouse s využitím medallion architektury, aby byla zajištěna odpovídající kvalita uložených dat

3 Načítání dat do Microsoft Fabric

- Seznámíte se s možnostmi použití Data Factory v prostředí Microsoft Fabric
- Pochopíte principy Dataflow Gen2
- Naučíte se vytvořit základní Dataflow pro načtení dat
- Seznámíte se s možnostmi tvorby Pipelines pro orchestraci procesu načítání a transformace dat
- Naučíte se připojit k externím datovým zdrojům a načítat dat s pomocí Apache Spark a notebooků
- Dozvíte se, jak efektivně ukládat data do Lakehouse a Delta tabulek

4 Uložení dat v Data Warehouse

- Seznámíte se s klíčovými vlastnostmi datových skladů v Microsoft Fabric
- Dozvíte se základní principy návrhu datových skladů, tabulek dimenzí a faktů
- Naučíte se definovat vztahy mezi tabulkami propagované do sémantického modelu
- Dozvíte se, jaké máte možnosti pro načítání dat do tabulek dimenzí a tabulek faktů
- Vyzkoušíte si načíst data do datového skladu
- Naučíte se nad datovým skladem spouštět dotazy v jazyce T-SQL
- Dozvíte se, jak pracovat s Visual Query Editorem
- Vyzkoušíte si připojit se přes SQL Server Management Studio
- Dozvíte se, jak monitorovat běžící dotazy s pomocí Dynamic Management Views a jak sledovat časté a náročné dotazy s pomocí Query Insights
- Naučíte se nastavit oprávnění v jazyce T-SQL na úrovni objektů datového skladu

5 Analýza a vizualizace dat v Microsoft Fabric

- Seznámíte s různými možnostmi analýzy uložených dat
- Zjistíte, jaké máte možnosti analýzy dat s využitím jazyka SQL
- Pochopíte, jak vytvářet Notebooky za účelem analýzy a vizualizace dat s využitím Apache Spark
- Dozvíte se, jak definovat nad daty sémantické modely, správně provázat data a využít jazyk DAX pro výpočty nad daty
- Objevíte různé možnosti vizualizace dat a tvorby reportů
- Vyzkoušíte si vytvořit vlastní reporty nad daty s využitím Power BI

6 Strojové učení v Microsoft Fabric

- Dozvíte se, jaké máte možnosti tvorby, tréninku a skórování modelů strojového učení v prostředí Microsoft Fabric
- Vyzkoušíte si vytvořit jednoduchý model strojového učení a vytrénovat jej nad daty uloženými v Data Lake

7 Uložení a zpracování real-time dat v Microsoft Fabric

- Seznámíte se s možnostmi zpracování událostí v reálném čase v Microsoft Fabric
- Vyzkoušíte si pracovat s Real-Time Huby, ukládat data do úložiště Eventhouse a dotazovat je jazykem KQL
- Zjistíte, jak vizualizovat data v reálném čase

8 Automatizace upozornění nad daty

- Seznámíte se se službou Data Activator
- Naučíte se automatizovat detekce překročení hranic a vzorů v datech za účelem generování upozornění a spouštění vlastních akcí