DP-600: Microsoft Fabric Analytics Engineer

Microsoft - Data & AI

- **Nível:** Avançado
- **Duração:** 28h

Sobre o curso

This course covers methods and practices for implementing and managing enterprise-scale data analytics solutions using Microsoft Fabric.

Students will build on existing analytics experience and will learn how to use Microsoft Fabric components, including lakehouses, data warehouses, notebooks, dataflows, data pipelines, and semantic models, to create and deploy analytics assets.

Destinatários

The primary audience for this course is data professionals with experience in data modeling, extraction, and analytics. DP-600 is designed for professionals who want to use Microsoft Fabric to create and deploy enterprise-scale data analytics solutions.

Pré-requisitos

This course is best suited for those who have the PL-300 certification or similar expertise in using Power BI for data transformation, modeling, visualization, and sharing. Also, learners should have prior experience in building and deploying data analytics solutions at the enterprise level.

Programa

- Ingest Data with Dataflows Gen2 in Microsoft Fabric
- Ingest data with Spark and Microsoft Fabric notebooks

- Use Data Factory pipelines in Microsoft Fabric
- Get started with lakehouses in Microsoft Fabric
- Organize a Fabric lakehouse using medallion architecture design
- Use Apache Spark in Microsoft Fabric
- Work with Delta Lake tables in Microsoft Fabric
- Get started with data warehouses in Microsoft Fabric
- Load data into a Microsoft Fabric data warehouse
- Query a data warehouse in Microsoft Fabric
- Monitor a Microsoft Fabric data warehouse
- Understand scalability in Power BI
- Create Power BI model relationships
- Use tools to optimize Power BI performance
- Enforce Power BI model security

Ingest Data with Dataflows Gen2 in Microsoft Fabric

Data ingestion is crucial in analytics. Microsoft Fabric's Data Factory offers Dataflows (Gen2) for visually creating multi-step data ingestion and transformation using Power Query Online.

In this module, you'll learn how to:

- Describe Dataflow (Gen2) capabilities in Microsoft Fabric
- Create Dataflow (Gen2) solutions to ingest and transform data
- Include a Dataflow (Gen2) in a pipeline
- Introduction
- Understand Dataflows (Gen2) in Microsoft Fabric
- Explore Dataflows (Gen2) in Microsoft Fabric
- Integrate Dataflows (Gen2) and Pipelines in Microsoft Fabric
- Exercise – Create and use a Dataflow (Gen2) in Microsoft Fabric
- Knowledge check
- Summary

Ingest data with Spark and Microsoft Fabric notebooks

Discover how to use Apache Spark and Python for data ingestion into a Microsoft Fabric lakehouse. Fabric notebooks provide a scalable and systematic solution.

By the end of this module, you'll be able to:

- Ingest external data to Fabric lakehouses using Spark
- Configure external source authentication and optimization
- Load data into lakehouse as files or as Delta tables
- Introduction
- Connect to data with Spark
- Write data into a lakehouse
- Consider uses for ingested data
- Exercise – Ingest data with Spark and Microsoft Fabric notebooks
- Knowledge check
- Summary

Use Data Factory pipelines in Microsoft Fabric

Microsoft Fabric includes Data Factory capabilities, including the ability to create pipelines that orchestrate data ingestion and transformation tasks.

In this module, you'll learn how to:

- Describe pipeline capabilities in Microsoft Fabric
- Use the Copy Data activity in a pipeline
- Create pipelines based on predefined templates
- Run and monitor pipelines

- Introduction
- Understand pipelines
- Use the Copy Data activity
- Use pipeline templates
- Run and monitor pipelines
- Exercise – Ingest data with a pipeline
- Knowledge check
- Summary

Get started with lakehouses in Microsoft Fabric

Lakehouses merge data lake storage flexibility with data warehouse analytics. Microsoft Fabric offers a lakehouse solution for comprehensive analytics on a single SaaS platform.

In this module, you'll learn how to:

- Describe core features and capabilities of lakehouses in Microsoft Fabric
- Create a lakehouse
- Ingest data into files and tables in a lakehouse
- Query lakehouse tables with SQL

- Introduction
- Explore the Microsoft Fabric Lakehouse
- Work with Microsoft Fabric Lakehouses
- Explore and transform data in a lakehouse
- Exercise – Create and ingest data with a Microsoft Fabric Lakehouse
- Knowledge check

- Summary

Organize a Fabric lakehouse using medallion architecture design

Explore the potential of the medallion architecture design in Microsoft Fabric. Organize and transform your data across Bronze, Silver, and Gold layers of a lakehouse for optimized analytics.

In this module, you'll learn how to:

- Describe the principles of using the medallion architecture in data management.
- Apply the medallion architecture framework within the Microsoft Fabric environment.
- Analyze data stored in the lakehouse using DirectLake in Power BI.
- Describe best practices for ensuring the security and governance of data stored in the medallion architecture.

- Introduction
- Describe medallion architecture
- Implement a medallion architecture in Fabric
- Query and report on data in your Fabric lakehouse
- Considerations for managing your lakehouse

- Exercise – Organize your Fabric lakehouse using a medallion architecture
- Knowledge check
- Summary

Use Apache Spark in Microsoft Fabric

Apache Spark is a core technology for large-scale data analytics. Microsoft Fabric provides support for Spark clusters, enabling you to analyze and process data in a Lakehouse at scale.

In this module, you'll learn how to:

- Configure Spark in a Microsoft Fabric workspace
- Identify suitable scenarios for Spark notebooks and Spark jobs
- Use Spark dataframes to analyze and transform data
- Use Spark SQL to query data in tables and views
- Visualize data in a Spark notebook

- Introduction
- Prepare to use Apache Spark

- Run Spark code
- Work with data in a Spark dataframe
- Work with data using Spark SQL
- Visualize data in a Spark notebook
- Exercise – Analyze data with Apache Spark
- Knowledge check
- Summary

Work with Delta Lake tables in Microsoft Fabric

Tables in a Microsoft Fabric lakehouse are based on the Delta Lake storage format commonly used in Apache Spark. By using the enhanced capabilities of delta tables, you can create advanced analytics solutions.

In this module, you'll learn how to:

- Understand Delta Lake and delta tables in Microsoft Fabric
- Create and manage delta tables using Spark
- Use Spark to query and transform data in delta tables
- Use delta tables with Spark structured streaming

- Introduction
- Understand Delta Lake
- Create delta tables
- Work with delta tables in Spark
- Use delta tables with streaming data
- Exercise – Use delta tables in Apache Spark
- Knowledge check
- Summary

Get started with data warehouses in Microsoft Fabric

Data warehouses are analytical stores built on a relational schema to support SQL queries. Microsoft Fabric enables you to create a relational data warehouse in your workspace and integrate it easily with other elements of your end-to-end analytics solution.

In this module, you'll learn how to:

- Describe data warehouses in Fabric
- Understand a data warehouse vs a data Lakehouse

- Work with data warehouses in Fabric
- Create and manage datasets within a data warehouse
- Introduction
- Understand data warehouse fundamentals
- Understand data warehouses in Fabric
- Query and transform data
- Prepare data for analysis and reporting
- Secure and monitor your data warehouse
- Exercise – Analyze data in a data warehouse
- Knowledge check
- Summary

Load data into a Microsoft Fabric data warehouse

Data warehouse in Microsoft Fabric is a comprehensive platform for data and analytics, featuring advanced query processing and full transactional T-SQL capabilities for easy data management and analysis.

In this module, you'll:

- Learn different strategies to load data into a data warehouse in Microsoft Fabric.
- Learn how to build a data pipeline to load a warehouse in Microsoft Fabric.
- Learn how to load data in a warehouse using T-SQL.
- Learn how to load and transform data with dataflow (Gen 2).

- Introduction
- Explore data load strategies
- Use data pipelines to load a warehouse
- Load data using T-SQL
- Load and transform data with Dataflow Gen2
- Exercise: Load data into a warehouse in Microsoft Fabric
- Knowledge check
- Summary

Query a data warehouse in Microsoft Fabric

Data warehouse in Microsoft Fabric is a comprehensive platform for data and analytics, featuring advanced query processing and full transactional T-SQL capabilities for easy data management and analysis.

In this module, you'll:

- Use SQL query editor to query a data warehouse.
- Explore how visual query editor works.
- Learn how to connect and query a data warehouse using SQL Server Management Studio.

Monitor a Microsoft Fabric data warehouse

A data warehouse is a vital component of an enterprise analytics solution. It’s important to learn how to monitor a data warehouse so you can better understand the activity that occurs in it.

After completing this module, you’ll be able to:

- Monitor capacity unit usage with the Microsoft Fabric Capacity Metrics app.
- Monitor current activity in the data warehouse with dynamic management views.
- Monitor querying trends with query insights views.

Understand scalability in Power BI

Scalable data models enable enterprise-scale analytics in Power BI. Implement data modeling best practices, use large dataset storage format, and practice building a star schema to design analytics solutions that can scale.

By the end of this module, you'll be able to:

- Describe the importance of building scalable data models
- Implement Power BI data modeling best practices
- Use the Power BI large dataset storage format

Create Power BI model relationships

Power BI model relationships form the basis of a tabular model. Define Power BI model relationships, set up relationships, recognize DAX relationship functions, and describe relationship evaluation.

By the end of this module, you'll be able to:

- Understand how model relationship work.
- Set up relationships.
- Use DAX relationship functions.
- Understand relationship evaluation.

- Introduction
- Understand model relationships
- Set up relationships
- Use DAX relationship functions
- Understand relationship evaluation
- Exercise: Work with model relationships

- Knowledge check

- Summary