A guide to achieving a DP-203 Certification

For years now, we’ve created highly specialized technologies to help us aggregate and analyze disparate data sources and formats. These use-case specific solutions are instrumental in harnessing data, but they have created numerous silos and niches, making subsequent data collection efforts inefficient and disjointed.

With Azure Synapse Analytics, every data engineer has access to a unified experience that enables them to leverage all of their data to unlock powerful insights. This guide will show you how Azure Synapse brings together data integration, enterprise data warehousing, and big data analytics at cloud scale.

By committing less than an hour each day, you’ll be able to better understand how to ingest different data sources, transform the data, and optimize for analytics all within a single platform. Each week you’ll watch a video on foundational concepts of Azure data engineering, complete a step-by-step training, and try what you’ve learned. When all is said and done, you will have the expertise you need to successfully complete your DP-203 certification.
There's so much to learn about Data Engineering on Azure. Don't worry—we’ve curated an easy-to-understand path to drive you towards certification in only 4 weeks. Each week you'll watch a video on foundational concepts, learn from step-by-step training, and try skills for yourself with a self-guided exercise. Click on the icon to jump to that week's training.

Before you begin, click here to prepare for this path.

**WEEK 1**
**Explore your data**
Learn about key data engineering concepts and tools, and how to easily get started querying your data lake with Azure Synapse

**WEEK 2**
**Optimize your data warehouse**
Find out how to create and augment an enterprise data warehouse using Azure Synapse

**WEEK 3**
**Transform your data**
Discover how to use Apache Spark for Azure Synapse to simplify the ingestion and transformation of your data

**WEEK 4**
**Streamline your data pipelines**
Discover how to use Azure Synapse to create efficient pipelines at cloud scale with only one application

**Complete certification exam**
Prepare the learning path

Make sure to complete the following tasks before you hit the road. Then use the navigation bar to return to your Learning Path and start your training or click ahead to move to week 1.

- Sign up for the 30 days challenge and after completing all required modules, you’ll be eligible to receive 50 percent off a Microsoft certification exam*.
- Start an Azure Free Trial or use your existing Azure subscription. You’ll need it for hands-on learning.

*Visit Terms and Conditions to learn more.
Week 1
Explore your data

With the amount of data being amassed today, it is often hard to figure out what data you have, let alone what is useful. Azure Synapse is a central place where you can view and interact with the data across your entire data estate. Instead of code switching between languages, programs and technologies, Azure Synapse helps you explore the breadth of your data lake and easily query your data.

Spend this first week learning the essentials. Learn how to assess your entire data estate, determine where pertinent data resides, build a simple data warehouse, and easily query data in the data lake using serverless SQL pools. By the end of the week, you’ll be able to transform your data in the lake, secure it, and understand how to use this newly formatted and queried data to drive meaningful insights.

Click ahead to access this week’s trainings >
Analyze data with a serverless SQL pool

In this tutorial, you'll learn how to analyze a data lake with serverless SQL pool.
Week 2
Optimize your data warehouse

The purpose of an enterprise data warehouse is to consolidate your data at scale so you can derive meaningful insights. But ingesting incongruent data types across silos can be mind-numbingly complex and often presents serious performance constraints. Azure Synapse helps you optimize your enterprise data warehouse at scale in one service—enabling you to integrate structured and unstructured data, transform the data, and serve insights.

This week we’ll build upon our basic data warehouse from week one, and you’ll see how much time and effort can be saved by scaling your data warehouse with analytical architecture patterns. We’ll focus on how to optimize your data and load data in an enterprise data warehouse to query and streamline system performance. Ultimately, you’ll learn how powerful the SQL and Apache Spark integrations are in Azure Synapse, and understand how to manage, secure, monitor, and analyze storage in a modern data warehouse.
Click the box to launch each module. Once completed, be sure to check the box to easily track your progress.

**W A T C H**

- **5 mins**
  - Flexibility of SQL pools on Azure Synapse
    Discover how to use T-SQL to explore and analyze data in Azure Synapse

**L E A R N**

- **6 hrs, 52 mins**
  - Work with Data Warehouses using Azure Synapse
    Explore the tools and techniques that can be used to work with enterprise data warehouses productively and securely within Azure Synapse

**T R Y**

- **5 mins**
  - Analyze data with dedicated SQL pools
    Learn how to use T-SQL to explore and analyze data in Azure Synapse

Use the navigation bar to return to your Learning Path and preview next week or click ahead to move to week 3.
Week 3

Transform your data

Because it’s a crowd favorite, we have fully integrated Apache Spark for Azure Synapse. The integration is complete with security and fully managed provisioning, ultimately simplifying the ingestion and transformation of your data. To make collaboration with data scientists easier, we’ve also created Apache Spark notebooks that have live code, visualizations, and narrative text to run quick experiments and derive preliminary insights.

This week, we’ll focus on how to ingest your data with Apache Spark notebooks inside of Azure Synapse, then learn how transform that data, all while integrating your SQL and Apache Spark pools. We’ll cap the week off by learning how to monitor and manage our data engineering workloads.

Click ahead to access this week's trainings >
Week 3

2 hrs 26 mins
Perform data engineering with Azure Synapse Apache Spark Pools
Learn how to use Apache Spark Pools, which enable you to boost the performance of big-data analytic applications

1 hr.
Analyze with Apache Spark
Discover the basic steps to load and analyze data with Apache Spark for Azure Synapse

WATCH

LEARN

Try

Click the box to launch each module. Once completed, be sure to check the box to easily track your progress.

Use the navigation bar to return to your Learning Path and preview next week or click ahead to move to week 4.
Week 4
Streamline Your Data Pipelines

In the past, moving data from different parts of the business proved difficult: pipelines were incredibly complex and costly, requiring multiple languages, time intensive scripts, and huge amounts of bandwidth. Azure Synapse enables you to create efficient pipelines at cloud scale with only one application. In minutes, you can connect to external data storage services, explore those files, create a pipeline with a data flow connecting to the outside source, and route the pertinent data to your warehouse or lake.

This week we’ll break down just how easy it is to build and manage data pipelines in the cloud using Azure Synapse. We’ll perform operational analytics with Azure Synapse Link for Azure Cosmos DB. We’ll end with an exercise that will show the culmination of your knowledge in which you’ll integrate pipelines into your data lakes and data warehouse using Synapse.
Click the box to launch each module. Once completed, be sure to check the box to easily track your progress.

**WATCH**

**5 mins**
**Build Data Pipelines with Ease**
Learn how to easily create an Azure Synapse pipeline to facilitate data integration

**2 hrs 26 mins**
**Work with Hybrid Transactional and Analytical Processing**
Perform operational analytics using Azure Synapse link for Azure Cosmos DB

**LEARN**

**5 hrs 48 mins**
**Data integration at scale with Azure Synapse**
Learn how to create and manage data pipelines in the cloud using Azure Synapse and Azure Data Factory

**TRY**

**5 mins**
**Integrate with Pipelines**
Learn how to integrate pipelines and activities using Synapse

Use the navigation bar to click ahead to complete your learning path.
Congratulations you’ve completed your training! Now, you’re ready to perform data engineering on Microsoft Azure.

**Schedule your D-203 certification exam**
Show the world you’re equipped with the core skills for working on Azure.

**Engage with us on our Azure Synapse Analytics Page**
Get the latest updates and best practices from the Product Management team.

**Check out our Azure Data Scientist Learning Journey**
Use this 30-day learning path to grow your skills and prepare for the Azure Data Scientist Associate certification.