# DP200 - Implementing an Azure Data Solution

## Course Content

## Course Description:

In this course, the students will implement various data platform technologies into solutions that are in line with business and technical requirements including on-premises, cloud, and hybrid data scenarios incorporating both relational and No-SQL data. They will also learn how to process data using a range of technologies and languages for both streaming and batch data.

The students will also explore how to implement data security including authentication, authorization, data policies and standards. They will also define and implement data solution monitoring for both the data storage and data processing activities. Finally, they will manage and troubleshoot Azure data solutions which includes the optimization and disaster recovery of big data, batch processing and streaming data solutions.

## Target Student:

The primary audience for this course is data professionals, data architects, and business intelligence professionals who want to learn about the data platform technologies that exist on Microsoft Azure. The secondary audience for this course is individuals who develop applications that deliver content from the data platform technologies that exist on Microsoft Azure.

## Prerequisites:

In addition to their professional experience, students who take this training should have technical knowledge equivalent to the following courses:

- Azure Fundamentals

## Topics:

**Module 1: Azure for the Data Engineer**

This module explores how the world of data has evolved and how cloud data platform technologies are providing new opportunities for business to explore their data in different ways. The student will gain an overview of the various data platform technologies that are available, and how a Data Engineers role and responsibilities has evolved to work in this new world to an organization benefit

**Lessons**

- Explain the evolving world of data
- Survey the services in the Azure Data Platform
- Identify the tasks that are performed by a Data Engineer
- Describe the use cases for the cloud in a Case Study

## Module 2: Working with Data Storage

This module teaches the variety of ways to store data in Azure. The Student will learn the basics of storage management in Azure, how to create a Storage Account, and how to choose the right model for the data you want to store in the cloud. They will also understand how data lake storage can be created to support a wide variety of big data analytics solutions with minimal effort.

**Lessons**

- Choose a data storage approach in Azure
- Create an Azure Storage Account
- Explain Azure Data Lake storage
- Upload data into Azure Data Lake

## Module 3: Enabling Team Based Data Science with Azure Databricks

This module introduces students to Azure Databricks and how a Data Engineer works with it to enable an organization to perform Team Data Science projects. They will learn the fundamentals of Azure Databricks and Apache Spark notebooks; how to provision the service and workspaces and learn how to perform data preparation task that can contribute to the data science project.

**Lessons**

- Explain Azure Databricks
- Work with Azure Databricks
- Read data with Azure Databricks
- Perform transformations with Azure Databricks

## Module 4: Building Globally Distributed Databases with Cosmos DB

In this module, students will learn how to work with NoSQL data using Azure Cosmos DB. They will learn how to provision the

service, and how they can load and interrogate data in the service using Visual Studio Code extensions, and the Azure Cosmos DB .NET Core SDK. They will also learn how to configure the availability options so that users are able to access the data from anywhere in the world.

**Lessons**

- Create an Azure Cosmos DB database built to scale
- Insert and query data in your Azure Cosmos DB database
- Build a .NET Core app for Cosmos DB in Visual Studio Code
- Distribute your data globally with Azure Cosmos DB

## Module 5: Working with Relational Data Stores in the Cloud

In this module, students will explore the Azure relational data platform options including SQL Database and SQL Data Warehouse. The student will be able explain why they would choose one service over another, and how to provision, connect and manage each of the services.

**Lessons**

- Use Azure SQL Database
- Describe Azure SQL Data Warehouse
- Creating and Querying an Azure SQL Data Warehouse
- Use PolyBase to Load Data into Azure SQL Data Warehouse

## Module 6: Performing Real-Time Analytics with Stream Analytics

In this module, students will learn the concepts of event processing and streaming data and how this applies to Events Hubs and Azure

Stream Analytics. The students will then set up a stream analytics job to stream data and learn how to query the incoming data to perform analysis of the data. Finally, you will learn how to manage and monitor running jobs.

**Lessons**
- Explain data streams and event processing
- Data Ingestion with Event Hubs
- Processing Data with Stream Analytics Jobs

## Module 7: Orchestrating Data Movement with Azure Data Factory

In this module, students will learn how Azure Data factory can be used to orchestrate the data movement and transformation from a wide range of data platform technologies. They will be able to explain the capabilities of the technology and be able to set up an end to end data pipeline that ingests and transforms data.

**Lessons**
- Explain how Azure Data Factory works
- Azure Data Factory Components
- Azure Data Factory and Databricks

## Module 8: Securing Azure Data Platforms

In this module, students will learn how Azure provides a multi-layered security model to protect your data. The students will explore how security can range from setting up secure networks and access keys, to defining permission through to monitoring across a range of data stores.

**Lessons**
- An introduction to security
- Key security components
- Securing Storage Accounts and Data Lake Storage
- Securing Data Stores
- Securing Streaming Data

## Module 9: Monitoring and Troubleshooting Data Storage and Processing

In this module, the student will get an overview of the range of monitoring capabilities that are available to provide operational support should there be issue with a data platform architecture. They will explore the common data storage and data processing issues. Finally, disaster recovery options are revealed to ensure business continuity.

**Lessons**
- Explain the monitoring capabilities that are available
- Troubleshoot common data storage issues
- Troubleshoot common data processing issues
- Manage disaster recovery