DP-203 Weekly Study Plan (6 Weeks)

This 6-week plan prepares you for DP-203: Data Engineering on Microsoft Azure. It covers modern data warehouses, star schemas, ETL/ELT pipelines, orchestration, and best practices using Azure Data Factory, Synapse, and Databricks.

Week 1: Data Storage & Ingestion

- Learn Azure Data Lake Storage Gen2, OneLake concepts.
- Study batch vs streaming ingestion approaches.
- Hands-on: Ingest CSV and JSON data into ADLS using ADF.

Week 2: Data Transformation & Processing

- Learn Azure Data Factory pipelines and mapping data flows.
- Explore Databricks basics (PySpark, notebooks).
- Hands-on: Build a simple ETL job in ADF + Databricks.

Week 3: Data Modeling & Warehousing

- Review star vs snowflake schema design.
- Study Synapse dedicated vs serverless SQL pools.
- Hands-on: Create fact/dim tables in Synapse, populate with ETL.

Week 4: Streaming & Orchestration

- Learn Azure Stream Analytics basics.
- Understand event-driven pipelines (Event Hubs, Kafka).
- Hands-on: Create a streaming pipeline from Event Hub to Synapse.

Week 5: Security & Optimization

- Study Managed Identity, RBAC, and data masking.
- Learn about partitioning, indexing, and query optimization in Synapse.
- Hands-on: Secure your pipeline and test performance improvements.

Week 6: Review & Practice Exams

- Take at least 2 full-length practice exams.
- Revisit weak areas (ADF, Synapse, Databricks).
- Final lab: Build a mini end-to-end pipeline (ingest → transform → model → visualize in Power BI).