**Daily Data Engineering Plan (10 Weeks)**

**Week 1 — SQL Foundations & Data Modeling**

**Goal:** Be fluent in basic T-SQL and understand OLTP vs OLAP.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Install SQL Server Developer + SSMS + VS Code. Connect to your instance. |
| 2 | Learn SELECT, WHERE, ORDER BY, TOP, DISTINCT. Practice on system tables in SSMS. |
| 3 | Learn JOINs (INNER, LEFT, RIGHT, FULL). Practice joining 2–3 tables. |
| 4 | Learn GROUP BY, HAVING, aggregates (SUM, AVG, COUNT). |
| 5 | Create your first database, tables, and insert data manually. |
| 6 | Use Draw.io to design an ERD for a small OLTP system; then design a star schema. |
| 7 | **Review:** Write 10 sample queries; commit ERD to GitHub. |

**Week 2 — Advanced SQL & Performance**

**Goal:** Master stored procs, functions, indexing, optimization.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Learn stored procedures, parameters, and execution. |
| 2 | Learn scalar & table-valued functions. |
| 3 | Create triggers for INSERT/UPDATE events. |
| 4 | Learn indexing: clustered vs non-clustered. Create indexes and test query speed. |
| 5 | Use execution plans in SSMS to find slow queries and optimize them. |
| 6 | Learn CTEs and window functions (ROW\_NUMBER, RANK). |
| 7 | **Review:** Optimize 3 queries; push scripts to GitHub. |

**Week 3 — ETL with SSIS**

**Goal:** Build your first ETL pipeline.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Install SSDT in Visual Studio. Create a new Integration Services Project. |
| 2 | Learn Control Flow vs Data Flow in SSIS. |
| 3 | Load CSV into staging tables (stg\_Customers, etc.). |
| 4 | Add transformations: data type conversion, Derived Columns. |
| 5 | Implement Slowly Changing Dimension Type 1. |
| 6 | Implement SCD Type 2; add error handling and logging. |
| 7 | **Review:** Build a complete CSV → DW ETL package and commit to GitHub. |

**Week 4 — Data Warehousing & OLAP with SSAS**

**Goal:** Build a cube/model for analytics.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Learn OLAP concepts: dimensions, measures, hierarchies. |
| 2 | Create an SSAS Tabular model from RetailDW. |
| 3 | Add relationships and basic DAX measures. |
| 4 | Create hierarchies (Year → Month → Day). |
| 5 | Add KPIs and time intelligence. |
| 6 | Deploy the model; connect with Excel to browse. |
| 7 | **Review:** Build a Product Sales Analysis model; commit project files. |

**Week 5 — Reporting & Visualization**

**Goal:** Learn SSRS + Power BI.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Install/configure SSRS locally. |
| 2 | Create a simple SSRS report (table). |
| 3 | Add parameters, drill-down reports, and charts. |
| 4 | Install Power BI Desktop; connect to SQL Server. |
| 5 | Create a dashboard with slicers and visuals. |
| 6 | Publish to Power BI Service (if account available). |
| 7 | **Review:** Export SSRS reports and Power BI PBIX to GitHub. |

**Week 6 — Python for Data Engineering**

**Goal:** Automate ETL with Python.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Install Python + VS Code extensions. |
| 2 | Learn Pandas basics: reading CSV, filtering, grouping. |
| 3 | Connect to SQL Server using SQLAlchemy. |
| 4 | Load CSV data into SQL Server with Python. |
| 5 | Call a REST API and save the JSON response to SQL Server. |
| 6 | Create a Python ETL that cleans data and appends to DW. |
| 7 | **Review:** Push Python ETL script to GitHub. |

**Week 7 — Big Data Basics**

**Goal:** Process large data in Spark.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Set up Databricks Community Edition. |
| 2 | Load a CSV into Spark DataFrame. |
| 3 | Transform and aggregate with Spark SQL. |
| 4 | Save Spark results back to CSV. |
| 5 | Connect Spark to SQL Server (JDBC). |
| 6 | Learn Hive basics; run queries. |
| 7 | **Review:** Document process in GitHub. |

**Week 8 — Cloud Platforms**

**Goal:** Learn cloud ingestion & storage.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Create free Azure account; set up Azure SQL DB. |
| 2 | Upload data to Azure Blob Storage. |
| 3 | Use Azure Data Factory to copy data to Azure SQL DB. |
| 4 | Learn AWS S3 basics; upload/download files. |
| 5 | Query data in Google BigQuery. |
| 6 | Compare Azure vs AWS vs GCP for data engineering. |
| 7 | **Review:** Cloud pipeline notes to GitHub. |

**Week 9 — CI/CD & Orchestration**

**Goal:** Automate deployment and scheduling.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Install Git; create GitHub repo for your project. |
| 2 | Learn branching, commits, and pull requests. |
| 3 | Set up Azure DevOps pipeline for database scripts. |
| 4 | Install Apache Airflow locally. |
| 5 | Create a simple Airflow DAG to run Python ETL daily. |
| 6 | Test the DAG; add logging and alerts. |
| 7 | **Review:** Push DAG and CI/CD configs to GitHub. |

**Week 10 — Capstone Project**

**Goal:** Build a full data engineering pipeline.

|  |  |
| --- | --- |
| **Day** | **Tasks** |
| 1 | Design project scope & ERD. |
| 2 | Set up source ingestion: CSV + API + Cloud storage. |
| 3 | Build SSIS package for CSV ingestion. |
| 4 | Build Python script for API ingestion. |
| 5 | Transform and load into SQL DW. |
| 6 | Build SSAS model + Power BI dashboard. |
| 7 | Document everything in GitHub with README + diagrams. |