# 🚀 2-Month Data Engineering Learning Plan

## 📅 \*\*Daily Learning Tracker\*\*

| Day | Topic | Completed (✅/❌) |

|---|---|---|

| 1-2 | Python Basics (Data Types, Loops, Functions) | ❌ |

| 3-4 | Pandas & NumPy | ❌ |

| 5-6 | SQL Basics (SELECT, JOINS, CTEs) | ❌ |

| 7 | SQL Optimization & Mini-Project | ❌ |

| 8-9 | Data Modeling & Warehousing | ❌ |

| 10-11 | ETL Concepts & Airflow | ❌ |

| 12-13 | Cloud Storage & Data Pipelines | ❌ |

| 14 | ETL Pipeline Mini-Project | ❌ |

| 15-16 | Hadoop & Spark Basics | ❌ |

| 17-18 | Cloud Data Engineering | ❌ |

| 19-20 | Batch vs. Streaming Processing | ❌ |

| 21 | Batch Processing Pipeline Project | ❌ |

| 22-23 | Kafka Basics | ❌ |

| 24-25 | Spark Streaming | ❌ |

| 26-27 | Data Pipeline Automation (Docker, Kubernetes) | ❌ |

| 28 | Real-time Streaming Pipeline Project | ❌ |

| 29-30 | Docker & Kubernetes | ❌ |

| 31-32 | Terraform & IaC | ❌ |

| 33-34 | Git & CI/CD | ❌ |

| 35 | CI/CD Data Pipeline Project | ❌ |

| 36-38 | Project 1 – Batch Processing Pipeline | ❌ |

| 39-41 | Project 2 – Real-time Streaming Pipeline | ❌ |

| 42 | Documentation & GitHub Updates | ❌ |

| 43-44 | System Design Basics | ❌ |

| 45-46 | Data Structures & Algorithms | ❌ |

| 47-48 | Mock Interviews | ❌ |

| 49-50 | Resume & LinkedIn Optimization | ❌ |

| 51-52 | Job Applications & Networking | ❌ |

| 53-54 | Final Interview Prep | ❌ |

| 55-56 | Job Interviews | ❌ |

## 📌 \*\*Resources & Learning Links\*\*

- \*\*Python & SQL\*\*: [freeCodeCamp](https://www.freecodecamp.org/), [LeetCode](https://leetcode.com/)

- \*\*Data Warehousing\*\*: [Star & Snowflake Schema](https://www.youtube.com/watch?v=kyjTRi1mIhI)

- \*\*ETL & Airflow\*\*: [Airflow Documentation](https://airflow.apache.org/)

- \*\*Big Data\*\*: [Spark Tutorial](https://spark.apache.org/docs/latest/)

- \*\*Streaming\*\*: [Kafka Guide](https://kafka.apache.org/documentation/)

- \*\*Cloud\*\*: [AWS Data Engineering](https://aws.amazon.com/data-engineering/)

- \*\*CI/CD\*\*: [Docker & Kubernetes](https://www.docker.com/)

## 📁 \*\*Projects for Portfolio\*\*

1. \*\*ETL Pipeline\*\*: API -> Transform with Pandas -> Store in PostgreSQL

2. \*\*Batch Processing Pipeline\*\*: Spark Job to Process Large CSV files -> Load to BigQuery

3. \*\*Real-time Streaming Pipeline\*\*: Kafka -> Spark Streaming -> Database

4. \*\*CI/CD Data Pipeline\*\*: Automate Data Pipeline Deployment with Docker & Terraform

## 🎯 \*\*Interview Preparation Checklist\*\*

✅ SQL Queries on LeetCode (50 Problems)

✅ Python Coding Challenges (30 Problems)

✅ System Design Basics (2-3 Case Studies)

✅ Mock Interviews (2-3 Sessions)

✅ Resume & LinkedIn Updated

## 📂 \*\*Job Application Tracker\*\*

| Company | Role | Application Date | Status |

|---|---|---|---|

| XYZ Corp | Data Engineer | MM/DD/YYYY | Pending |

| ABC Tech | Data Engineer | MM/DD/YYYY | Interview Scheduled |