

GCP Data Engineering (45-Day Plan)

1. **Module 1: Introduction to Cloud Computing (2 Days)**
2. **Module 2: Introduction to Google Cloud Platform (GCP) (3 Days)**
3. **Module 3: SQL and Relational Database Systems with MySQL (3 Days)**
4. **Module 4: Python for Data Engineering (5 Days)**
5. **Module 5: Storage and Core GCP Services (6 Days)**
6. **Module 6: Data Warehouse Concepts and BigQuery (6 Days)**
7. **Module 7: Batch and Streaming Data (4 Days)**
8. **Module 8: Data Pipeline Orchestration (5 Days)**
9. **Module 9: Data Security and Privacy (2 Days)**
10. **Module 10: Data Visualization (2 Days)**
11. **Module 11: Advanced Data Engineering Concepts (4 Days)**
12. **Module 12: Final Wrap-Up (3 Days)**

Syllabus in Detail for 45 Days

Module 1: Introduction to Cloud Computing (2 Days)

- **Overview:**
 - Who can learn this course
 - Introduction to Cloud Computing
 - Benefits of Cloud Computing
 - Roles and Responsibilities of Cloud Data Engineer
 - Overview of Cloud Platforms (GCP/AWS/Azure)
 - Comparison among the popular Cloud Platforms
 - Understanding cloud service providers
- **Core Concepts:**
 - IaaS, PaaS, SaaS
 - Google Compute Engine:
 - Virtual machine instances
 - Creating and managing VM instances
 - Networking and security

Module 2: Introduction to Google Cloud Platform (GCP) (3 Days)

- GCP Overview and Architecture
- Advantages of Google Cloud Platform
- Google Network Infrastructure: Points of Presence/Data Centers/Regions/Zones
- Setting up Individual Accounts on GCP
- Navigating the GCP Console
- GCP Project Management, Credits, and Billing
- Basic Linux Commands
- Accessing GCP Services:
 - Google Cloud Shell (Shell Scripting)
 - Google Cloud SDK
- Introduction to Key GCP Services:
 - BigQuery, DataProc, DataFlow, DataFusion, Cloud Composer, Cloud SQL, Cloud Functions

Module 3: SQL and Relational Database Systems with MySQL (3 Days)

- SQL Fundamentals
- DML and DDL Statements
- Queries and Sub-Queries
- Joins and Keys (Primary, Foreign, Unique)
- Aggregate Functions
- Window Functions
- Introduction to Common Table Expressions (CTE)

Module 4: Python for Data Engineering (5 Days)

- Python Basics:
 - Syntax, Code Structure, Data Types, Collections (List, Tuple, Dictionary)
 - Operators, Conditional Statements, Loops, Functions, Exception Handling
 - File Operations
- Introduction to Pandas:
 - Data structures in Pandas (Series, DataFrames)
 - Operations on DataFrames

- Overview of Python Modules for Data Engineering
- **Module 5: Storage and Core GCP Services (6 Days)**
- Introduction to Google Cloud Storage (GCS):
 - Create/Delete/Upload Buckets and Files using Web UI, gsutil, and Python
 - Handling multiple files in GCS using Python
 - Data Processing in GCS using Pandas
 - Data Conversions and Validation
- Overview of Cloud SQL, Bigtable, Datastore, and Data Lakehouse

Module 6: Data Warehouse Concepts and BigQuery (6 Days)

- **Data Warehouse Basics:**
 - Introduction to DWH (Data Warehouse)
 - OLTP vs. OLAP
 - Dimension and Fact Tables
- **BigQuery Fundamentals:**
 - CRUD Operations
 - Loading Data into BigQuery Tables
 - Partitioned and Clustered Tables
 - External Tables and Queries
 - Integration with Python
 - Advanced SQL in BigQuery
 - BigQuery Views and Materialized Views

Module 7: Batch and Streaming Data (4 Days)

- Introduction to Batch and Streaming:
 - Cloud Dataflow
 - Cloud Datafusion
 - Pub/Sub
 - Apache Kafka

Module 8: Data Pipeline Orchestration (5 Days)

- Introduction to Google Cloud Composer:
 - Airflow Basics and Architecture
 - Deploying and Running Airflow DAGs
 - Integration with Dataproc and BigQuery
 - Managing Variables and Job Dependencies

Module 9: Data Security and Privacy (2 Days)

- IAM Roles and Permissions
- Data Encryption and Key Management
- Data Privacy Regulations (GDPR, CCPA)

Module 10: Data Visualization (2 Days)

- Introduction to Google Data Studio and Looker Studio
- Data Visualization Techniques

Module 11: Advanced Data Engineering Concepts (4 Days)

- Big Data Fundamentals
- Hadoop Basics and Architecture
- ETL and ELT Concepts

Module 12: Final Wrap-Up (3 Days)

- Mock Interviews
- Resume Preparation
- Job Application Assistance

This curriculum provides a comprehensive foundation for aspiring Cloud Data Engineers, combining in-depth theoretical understanding with hands-on experience in leveraging GCP tools and technologies for data engineering excellence.