Data Engineering & Cloud Computing Roadmap

This roadmap will guide you through the stages of learning Data Engineering and Cloud Computing. It includes free resources for each stage to help you acquire the necessary skills and knowledge. The roadmap is broken down into several stages, each focusing on different aspects of Data Engineering, Cloud Computing, and related tools.

# 🌱 Stage 1: Core Foundations

## 🔹 Programming (Python)

Free Resources:

- Python.org Documentation (Official docs, great for learning Python basics): https://docs.python.org/3/tutorial/

- Automate the Boring Stuff with Python (Book/Free course, focuses on practical Python applications): https://automatetheboringstuff.com/

- freeCodeCamp (Python section, great interactive lessons): https://www.freecodecamp.org/

## 🔹 SQL & Databases

Free Resources:

- SQLZoo (Interactive SQL tutorials): https://sqlzoo.net/

- Mode Analytics SQL Tutorial (Interactive lessons with practical exercises): https://mode.com/sql-tutorial/

- LeetCode SQL Problems (Great for practicing SQL problems): https://leetcode.com/problemset/database/

- Khan Academy SQL (Beginner-friendly): https://www.khanacademy.org/computing/computer-programming/sql

# 🏗️ Stage 2: Data Engineering Basics

## 🔹 Data Manipulation & Analysis

Free Resources:

- Kaggle Python and Pandas Course (Free course with hands-on exercises): https://www.kaggle.com/learn/pandas

- DataCamp’s Introduction to Pandas (First chapter free): https://www.datacamp.com/courses/pandas-foundations

- Pandas Documentation (Official documentation): https://pandas.pydata.org/pandas-docs/stable/

## 🔹 ETL Pipelines

Free Resources:

- ETL with Python (Blog): https://realpython.com/python-etl/

- Airflow Documentation (Official Airflow docs and tutorials): https://airflow.apache.org/docs/apache-airflow/stable/

- Building ETL Pipelines (YouTube): https://www.youtube.com/watch?v=XTpXfmgJl3w

## 🔹 Data Storage Formats

Free Resources:

- Data Formats (Medium Blog): https://medium.com/@saeed.akbari96/understanding-the-difference-between-parquet-avro-and-orc-file-formats-6c2b9d0c79f2

## 🔹 Version Control

Free Resources:

- Git & GitHub Course (freeCodeCamp): https://www.freecodecamp.org/news/git-and-github-for-beginners/

- GitHub Learning Lab (Interactive GitHub tutorials): https://lab.github.com/

# ☁️ Stage 3: Cloud & Big Data Tools

## 🔹 Intro to Cloud Platforms

Free Resources:

- AWS Training (Free introductory courses on AWS services): https://aws.amazon.com/training/

- Google Cloud Training (Free training resources from Google Cloud): https://cloud.google.com/training

- Microsoft Learn - Azure (Free Azure courses): https://learn.microsoft.com/en-us/training/azure/

## 🔹 Big Data Ecosystem

Free Resources:

- Intro to Big Data (Coursera): https://www.coursera.org/learn/big-data-introduction

- Apache Spark YouTube: https://www.youtube.com/watch?v=I\_zvByOY\_JA

- Learn Hadoop (TutorialsPoint): https://www.tutorialspoint.com/hadoop/index.htm

# 🧰 Stage 4: Advanced Data Engineering Tools

## 🔹 Orchestration & Workflow

Free Resources:

- Airflow Documentation: https://airflow.apache.org/docs/apache-airflow/stable/

- dbt Documentation: https://docs.getdbt.com/docs/introduction

## 🔹 Data Warehousing

Free Resources:

- Google BigQuery (Coursera): https://www.coursera.org/learn/google-cloud-bigquery

- Snowflake Essentials: https://www.snowflake.com/webinars/

- AWS Redshift Basics: https://aws.amazon.com/redshift/getting-started/

## 🔹 Containerization & DevOps Basics

Free Resources:

- Docker for Beginners: https://docker-curriculum.com/

- Kubernetes Basics: https://kubernetes.io/docs/tutorials/kubernetes-basics/

- CI/CD Tutorial (GitLab): https://docs.gitlab.com/ee/ci/

# 🚀 Stage 5: Cloud Engineering Specialization

## 🔹 Infrastructure as Code (IaC)

Free Resources:

- Terraform Getting Started: https://learn.hashicorp.com/collections/terraform/aws-get-started

- AWS CloudFormation (AWS Docs): https://aws.amazon.com/cloudformation/

## 🔹 Serverless & Event-Driven Systems

Free Resources:

- AWS Lambda (AWS Docs): https://aws.amazon.com/lambda/

- Google Cloud Functions (Docs): https://cloud.google.com/functions/docs/

## 🔹 Monitoring & Security

Free Resources:

- AWS CloudWatch Documentation: https://aws.amazon.com/cloudwatch/

- Azure Monitor (Docs): https://learn.microsoft.com/en-us/azure/azure-monitor/

- Google Cloud Security: https://cloud.google.com/security

# 📦 Stage 6: Projects & Portfolio

## 🔹 Free Resources to Help Build Projects

- Kaggle: https://www.kaggle.com/

- GitHub: https://github.com/

- Streamlit: https://streamlit.io/

- Google Colab: https://colab.research.google.com/

# 🧠 Additional Learning Platforms

- Coursera: https://www.coursera.org/

- edX: https://www.edx.org/

- YouTube (Search for data engineering-specific tutorials or cloud and DevOps guides): https://www.youtube.com/