PREREQUISITES

Python and **OOP**

DBMS

Linux

DSA, **Leetcode**

SQL

NoSQL (basics)

BIG DATA

Spark

Real Time

Kafka

Flink (basics)

Beam (basics)

ORCHESTRATION

Airflow

Dagster
(optional)

CLOUD (any 1)

AWS

GCP

Azure

Data Warehouse / Lakehouse

Data Modeling

Snowflake

Lakehouse and **Iceberg** (basics)

DBT (basics)

DATA LINEAGE AND CATALOGING

YouTube Video

Open Source: **Marquez**

Proprietary: **Collibra**

Look into only 1 of these 2

AI

YouTube video with 5 best free courses

DEVOPS

IaC - **Terraform** (basics)

Docker

CICD - **Jenkins**

Monitoring and Logging

Kubernetes (basics)

PROJECTS

Link-1
Link-2
Link-3