Rohith R - Detailed Weekly Interview Prep Plan (Sept 2025 - Feb 2026)

This document provides a detailed 22-week interview preparation plan for Data Engineer, Backend Engineer, and Cloud roles.

■ Roles You Can Target

- Data Engineer (Primary Role - ETL, Airflow, AWS, Big Data)
- Backend Developer (Python/AWS/FastAPI)
- Cloud Data Engineer / Big Data Engineer
- ML/AI Engineer (secondary option, bonus skills with Movenet, PyVista)

■ Weekly Prep Calendar

Month 1 (Sept 2025) → Foundations

- Week 1: Python Deep Dive - OOP, decorators, generators, context managers, multiprocessing; LeetCode Easy; S3 file processor project.
- Week 2: SQL Mastery - Joins, CTEs, window functions, optimizations; LeetCode SQL + Athena practice.
- Week 3: AWS Core - S3, Athena, Lambda, API Gateway, DynamoDB, ECS/EC2; Deploy FastAPI on Lambda.
- Week 4: DSA Brush-up - Sorting, searching, linked lists, recursion; LeetCode Easy-Medium daily.

Month 2 (Oct 2025) → Data Engineering Core

- Week 5: Airflow Basics - DAGs, operators, sensors, scheduling.
- Week 6: Airflow Advanced - Plugins, custom operators, AWS integration.
- Week 7: ETL Concepts - Partitioning, bucketing, Parquet/ORC; batch vs streaming.
- Week 8: AWS Big Data - EMR, Glue, Redshift basics; Glue job with Athena data.

Month 3 (Nov 2025) → Backend & Scaling

- Week 9: FastAPI & Microservices - CRUD APIs, Pydantic, JWT, async APIs.
- Week 10: System Design Basics - REST API design, caching, load balancing, SNS/SQS pub-sub.
- Week 11: Docker & Kubernetes - Dockerfile, docker-compose, pods, deployments.
- Week 12: CI/CD & Infra - GitHub Actions, Jenkins basics, Terraform for AWS.

Month 4 (Dec 2025) → Advanced Data Engineering

- Week 13: PySpark Foundations - RDDs, DataFrames, transformations, joins.
- Week 14: PySpark Advanced - Optimizations, Spark Streaming basics.
- Week 15: Kafka Basics - Producers, consumers, topics, partitions.
- Week 16: Databricks & Snowflake - Delta Lake, clusters, Snowflake warehouses.

Month 5 (Jan 2026) → Interview Readiness

- Week 17: Project Consolidation - End-to-end batch + streaming pipelines.
- Week 18: Mock Interviews (Technical) - Python, SQL, Airflow, Backend.
- Week 19: Mock Interviews (System Design) - Data pipeline & realtime design.
- Week 20: Resume & GitHub Polish - Update resume, upload projects, diagrams.
- Week 21: Behavioral Prep - STAR method, project storytelling.
- Week 22: Final Revision - AWS, PySpark, SQL, Python patterns.

■ Final Big Projects

- Project 1: Retail Analytics Pipeline (Batch ETL) - Airflow, PySpark, AWS, Snowflake. Raw data in S3 → Airflow DAG → PySpark on EMR → Snowflake for analytics.
- Project 2: Real-Time Order Tracking (Streaming) - FastAPI, Kafka, Spark Streaming, Databricks, Snowflake. Orders → Kafka → Spark Streaming → Snowflake → API/Dashboard.

■ Strategy

- Sept–Oct: Focus on foundations + build batch pipeline project.
- Nov–Dec: Work on streaming pipeline project + backend/system design.
- Jan: Mock interviews, polish GitHub, finalize resume, prepare stories.

■ Key Tech to Master

- Apache Airflow (advanced DAGs, plugins)
- PySpark (batch + streaming)
- Kafka (stream processing)
- Snowflake & Databricks (modern DE stack)
- AWS (Athena, S3, Glue, EMR, Lambda, ECS, DynamoDB)
- FastAPI & System Design
- Terraform, Docker, Kubernetes