

Umair

Karachi, Pakistan | +923494804814 | umairnawaz7228@gmail.com | linkedin.com/in/umairnawaz7228/ | github.com/umair7228 | umair-portfolio-web.vercel.app

EDUCATION

Saylani Mass IT Training (SMIT)

Diploma in Cloud Data Engineering | Among in Top 5 Students

Nov 2023 - Dec 2024

SKILLS

Languages Python (Pandas, Numpy, Matplotlib) · PySpark · SQL · TypeScript · JavaScript

Softwares AWS · Snowflake · Databricks · Docker · Apache Spark · Kafka · Airflow · Web Scrapping · BeautifulSoup · Selenium · Git · Bash · React.js · Next.js

EXPERIENCE

Data Engineer

CURVE DIGITAL SOLUTIONS PVT. LTD.

May 2025 – Present

Karachi, Pakistan

- Optimized a large-scale ETL pipeline, reducing execution time by over 80% (from 2 minutes 30 seconds to 30 seconds) through efficient SQL restructuring, use of Polars, and in-memory data transformations in Python.
- Developed real-time data streaming pipelines using Apache Kafka, enabling event-driven architecture and reducing data delivery latency.
- Accelerated big data processing using Apache Spark by tuning job configurations, applying custom partitioning, and leveraging caching, resulting in a 50% improvement in performance.
- Implemented workflow orchestration with Apache Airflow, achieving 100% automation of recurring pipelines with built-in failure handling and alerting.
- Collaborated cross-functionally with data scientists to deploy machine learning models into production-ready pipelines, enhancing model delivery speed and scalability.

Cloud Data Engineer

SAYLANI TECH

Jan 2024 – May 2025

Karachi, Pakistan

- Built and optimized ETL pipelines using Apache Airflow and Spark, automating data ingestion and transformation for scalable analytics.
- Designed and implemented data models in Snowflake, enhancing query performance and enabling real-time business insights.
- Developed cloud-native data solutions on AWS, leveraging S3, Lambda, and Kafka for efficient data processing.
- Integrated and processed diverse datasets from multiple sources, ensuring high data quality and reliability through validation pipelines.
- Collaborated with cross-functional teams to deliver end-to-end data solutions, improving accessibility for analytics and reporting.

PROJECTS

Data Warehousing & ETL Pipeline for an Online Learning Platform

| Freelance Project - [Github Repo](#)

Feb 2025 - Mar 2025

- Built a Snowflake data warehouse with AWS S3 & Apache Spark for student analytics.
- Automated ETL pipeline using Airflow & Snowpipe for real-time data ingestion.
- Created Power BI dashboards to track enrollments, completion rates, and engagement.

Pandemic Insights: ETL Pipeline with AWS Glue, Athena & Redshift

| Academic Project - [Github Repo](#)

Nov 2024 - Dec 2024

- Developed a scalable ETL pipeline to process and analyze COVID-19 data from multiple sources.
- Designed fact and dimension tables for optimized data querying in Amazon Redshift.
- Used AWS Glue and Athena for data transformation and insights generation.

Event-Driven Data Processing System

| Academic Project - [Github Repo](#)

Nov 2024 - Nov 2024

- Designed an event-driven architecture using AWS SNS, SQS, Lambda, and S3.
- Enabled real-time data processing through event triggers and message queues.
- Optimized data handling and response times for dynamic event flows.

Real-Time Stock Market Data Pipeline

| Academic Project - [Github Repo](#)

Oct 2024 - Nov 2024

- Developed a real-time stock data pipeline with Apache Kafka and AWS.
- Integrated Glue, Athena, and Crawlers for scalable storage and queries.
- Ensured timely access to market data for analytics.

Real-Time Data Pipeline with SCD Implementation

| Academic Project - [Github Repo](#)

Oct 2024 - Oct 2024

- Created a synthetic data pipeline using Python and Apache NiFi for Amazon S3.
- Implemented Slowly Changing Dimensions (SCD) Type 1 & 2 for data evolution tracking.
- Used Snowpipe for seamless data ingestion into Snowflake for efficient updates.

COURSES & CERTIFICATIONS

AWS Certified Data Engineer - Associate

Udemy

Data Engineering | Data Warehousing

Snowflake

Associate Data Engineering in SQL | Python

DataCamp