

KHALED ALRUWITA

Data Analysis | Data Engineer

 alruwita.k@gmail.com  +966566200419  kr.sa  linkedin.com/in/khaled-alruwita/  Github  Riyadh

Summary

Data Engineer and Data Analyst with strong experience in designing end-to-end data pipelines, building scalable data platforms, and automating ETL workflows using Python and SQL. Proven in delivering reliable, high-quality datasets through data warehousing, orchestration, and observability practices. Skilled at translating complex data requirements into production-ready solutions that support analytics.

Experience

Data Engineer Trainee, SITE

08/2025 – Present

- Built an end-to-end data engineering toolkit that automated SQL validation and optimization, reducing manual workload by ~80% and providing executives and leadership with reliable visibility into data quality, lineage, and overall transformation health.
- Built a fully automated daily ETL pipeline transforming raw data into cleansed layers, dashboards, and reports.
- Collaborated with the data modeling team to design and refine dimensional models and warehouse schemas.

Data Analyst Intern, Keeta Technology Arabia Limited

06/2025 – 08/2025

- Performed data cleaning, validation, and quality checks on large datasets.
- Ensured compliance with internal data management processes and reporting standards.

Education

Bachelor's Degree in Computer Information Systems, Prince Sattam Bin Abdulaziz

08/2021 – 05/2025

Certifications

Certified Data Management Professional (CDMP®)

Associate DAMA International

Business Analysis Fundamentals

Tuwaiq Academy

Projects

Traffic Simulation Project, [Final Year Project]

Developed a comprehensive traffic simulation at Prince Sattam bin Abdulaziz University (PSAU) campus to optimize vehicular flow under different scenarios. The project evaluated key traffic metrics such as stop time, waiting time, and vehicle throughput across scenarios like introducing new traffic lights and adding campus entry gates. Results were analyzed using decision tree algorithms to support data-driven traffic management decisions.

- The project was awarded Best Research Project at the university level for the year 2025

Warehouse & ETL Platform

Built an end-to-end ETL pipeline for used-car listings using Python, SQL, PySpark, Airflow, and PostgreSQL. Implemented bronze-silver-gold layers, star-schema modeling, automated orchestration, data-quality checks with Great Expectations, and interactive dashboards for analytics and impact analysis.

Skills

Back-End

Python , MongoDB , Flask , Django , PostgreSQL , Airflow

Front-End

HTML, CSS, JavaScript, Streamlit

Cloud Computing

AWS (EC2, S3, RDS, Lambda), Alibaba Cloud (ECS), Oracle Cloud Infrastructure

Data Analysis/Visualization

Power BI , Tableau , Pandas , PySpark

DevOps

Docker , Nginx , Linux , Git , Cloud Deployment

Soft Skills

- Problem Solving
- Communication Skills
- Multitasking

- Adaptability
- Time Management
- Reliability