NAVEED JOKHIO

Cloud Data Engineer|Software Engineering Undergraduate

📞 0304-6618347 ✉ naveedjokhio243@gmail.com 🔗 linkedin.com/in/naveed-jokhio

📍 Karachi, Pakistan

SUMMARY

Cloud Data Engineer with hands-on experience building ETL pipelines using Python, AWS (S3, Glue, Lambda), and SQL. Experienced in transforming real-world datasets into analytics-ready formats and optimizing data workflows through internships and projects to support decision-making and reporting

EXPERIENCE

Data Engineering / Analytics Intern
Data Science Internship (DataEngineering Focus) HexSoftwares 1 month

📅 01/2026 📍 Remote

Internship focusing on data engineering principles and hands-on experience

- Performed data extraction, cleaning, and transformation on real-world datasets using Python
- Built ETL workflows to load structured data into SQL databases, improving data processing efficiency by 30% and reducing manual data handling errors by 25%
- Implemented data validation and quality checks to ensure accuracy and consistency
- Prepared analytics-ready datasets supporting reporting and business insights
- Collaborated in a remote, team-based development environment following data engineering best practices

Data Engineering / Analytics Intern
Data Science Internship
Arch technologies 2 month internship

📅 01/2026 📍 Remote

Internship providing training in data science and data analytics

- Processed and pre-cleaned structured datasets using Python and SQL
- Assisted in transforming raw data into analysis-ready formats, increasing data usability by 35% and accelerating reporting turnaround time by 20%
- Supported data pipeline tasks and documentation under mentorship

EDUCATION

Bachelor of Software Engineering
Sindh Madressatul Islam University SMIU

📅 08/2023 - 08/2027 📍 Karachi, Pakistan

Pre-Engineering
Govt.Boys Degree College Naudero

📅 08/2019 - 08/2021 📍 Naudero, Pakistan

CERTIFICATIONS

- Data Analytics
- GenAI-Powered Data Analytics
- Introduction to Python
- Python Programming
- Data Science
- Snowflake Hands-On Essentials

SKILLS

- **Languages:** Python, SQL, PySpark
- **Cloud:** AWS (S3, EC2, Glue, Lambda)
- **Data Warehousing:** Snowflake
- **ETL/ELT:** Airflow, dbt

PROJECTS

ETL Pipeline – Wikipedia Banking Data

Cloud Data Engineering Project

- Designed and implemented an end-to-end ETL pipeline using Python for real-world financial data processing
- Extracted structured data from Wikipedia using Beautiful Soup and Requests
- Transformed and enriched datasets using Pandas, including multi-currency conversions (GBP, EUR, INR)
- Loaded cleaned datasets into CSV files and SQLite database for efficient storage and querying
- Developed SQL query functions to analyze processed data directly from the database
- Implemented logging mechanisms to monitor pipeline execution, improve reliability, and track failures

ETL Pipeline – Kaggle Orders Data (SQL Server)

Data Engineering Project

- Developed a complete ETL pipeline using a real-world dataset from Kaggle
- Performed data cleaning and schema optimization by removing redundant columns and refining data structures
- Loaded transformed datasets into SQL Server, ensuring analytics-ready and structured storage
- Executed analytical SQL queries to extract actionable business insights from orders data
- Strengthened expertise in data pipelines, data quality assurance, and SQL-based analytics aligned with industry best practices