

Jeiya Kumari

Data Engineer | Data Analyst

Email: jeiyakumari@gmail.com | Phone: +92-327-2815521 | Location: Karachi, Pakistan

Portfolio: <https://k-jeiya.github.io/Jeiya-Portfolio/> | GitHub: <https://github.com/K-Jeiya>

LinkedIn: <https://www.linkedin.com/in/jeiyakumari/>

Professional Summary

Detail-oriented **Data Engineer** with hands-on experience in designing ETL pipelines, orchestrating workflows, and building analytical dashboards. Proficient in **Python, SQL, and cloud data tools** including Airflow, PySpark, Power BI, and Streamlit. Experienced in transforming raw datasets into actionable insights and scalable data systems using both on-premises and cloud environments (**AWS, Azure**). Strong foundation in data modeling, automation, and performance optimization with a focus on clarity, maintainability, and accuracy.

Key Skills

- **Data Engineering:** ETL Pipelines · Data Modeling · OLTP/OLAP · Orchestration (Airflow) · Batch Processing (PySpark)
 - **Programming:** Python · SQL · Java · Bash
 - **Databases:** MySQL · PostgreSQL · MongoDB
 - **Tools & Cloud:** Airflow · Streamlit · Power BI · Docker · AWS · Google Cloud (GCP) · Azure · Firebase · Postman
 - **Data Analysis & Visualization:** Pandas · NumPy · Matplotlib · Power BI · Excel · Google Sheets
 - **Other Skills:** API Integration · Version Control (Git) · Machine Learning (Basic) · Collaboration & Communication
-

Professional Experience

Data Management Intern – National Expansion Plan of NICs | Feb 2024 – May 2024

- Cleaned, validated, and transformed startup data using **Python and Excel**.
 - Built automated spreadsheets for tracking and performance reporting.
 - Created visual dashboards to support data-driven decisions.
-

Projects

- **Complaints ETL Pipeline (Airflow · MySQL · Google Sheets)**

[!\[\]\(b4eeff342f60cc7bcd67d869b4fedca2_img.jpg\) GitHub Repository](#)

Designed an end-to-end ETL pipeline for the Consumer Financial Protection Bureau (CFPB) API using **Apache Airflow**.

Fetched and transformed data for all U.S. states, stored results in **MySQL and Google Sheets**, and automated email notifications.

- **Watchpoint Simulator (Final Year Project – Unity · AI · Behavior Trees)**

Simulated crowd and security behaviors using Unity and behavior trees.

Implemented phone-based item ordering, customer shopping logic, and data-driven flows.

- **Hubway Bikes Data Pipeline (PySpark · MongoDB)**
[!\[\]\(694fcb4611893e9db5249daba48abfc1_img.jpg\) GitHub Repository](#)
Built a data engineering workflow to join and aggregate trip and station datasets in **PySpark**, adding state and month-year derivations.
Loaded final transformed data into **MongoDB** for storage and BI analysis.
 - **Live Poll App (Kafka · Streamlit · Python)**
[!\[\]\(8ec8d5dc48934930a762fecf6ecbe179_img.jpg\) GitHub Repository](#)
Developed a real-time polling app using **Kafka** and **Streamlit**, visualizing live responses with interactive bar charts and metrics.
Simulated poll data with Python producer and displayed dynamic updates in dashboard.
 - **GenAI Startup Idea App (Streamlit · Gemini API)**
[!\[\]\(c34a15e67573dae8fbb88f4cbfb0f2e9_img.jpg\) Live Demo](#)
Designed a multi-page Streamlit app for the National Incubation Center, integrating the Gemini API to generate startup plans based on user themes and keywords.
 - **News to Humor LLM App (Python · Flask · LLM API)**
[!\[\]\(41f06fdeabb4e5a71d06fe8f32a46127_img.jpg\) GitHub Repository](#)
Created a REST API application that fetches news articles from NewsAPI and transforms content into humorous versions using LLM APIs (Groq / Hugging Face).
 - **School Management System (Java & Database Project)**
[!\[\]\(18eb66208e65404cce5042d73cf0a851_img.jpg\) GitHub Repository](#)
Developed a school database system using Java, Integrated student records and fee management.
-

Education

Bachelor of Science in Computer Science | Sindh Madressatul Islam University
CGPA: 3.5 | 2021 – 2025

Higher Secondary Certificate (Pre-Engineering) | Govt. Girls Degree College
A1 Grade | 2019 – 2021

Certifications

- Data Engineering (CDE Batch 9 – Karachi.AI @ NUST) – Completed
 - Microsoft Power BI Data Analyst (Datacamp) – In Progress
 - Google Data Analytics (Coursera)
 - Python for Data Science, AI & Development (Coursera)
 - Web & Mobile Application Development (SMIT)
-

Soft Skills

Analytical Thinking · Clear Communication · Team Collaboration · Independent Task Ownership