Muhammad Sufyaan Khateeb

+91 9591135172 | sufyaankhateeb@gmail.com | LinkedIn | Github | Bengaluru, India

SUMMARY

Data Engineer with 3+ years designing and operating production ETL/data-pipeline systems in **Python and SQL** at scale. Built event-driven pipeline orchestration with **Kafka** and **Airflow**, optimized **BigQuery** data models and partitioning to cut query cost by 99%, and integrated **AI-driven agents** into data tooling to remove manual pipeline work. Experienced with cloud data platforms (**AWS**, **GCP**), relational databases (**PostgreSQL**), and building validation/automation workflows that raise data quality and reliability across teams.

WORK EXPERIENCE

Wells Fargo June 2025 – Present

Software Engineer Bengaluru, India

- Designed and built an **event-driven ETL/data-pipeline orchestration platform in Python** that automated complex data-pipeline migrations and integrated **AI-driven agents**, reducing manual developer effort for new data ingestions by **95%**.
- Built end-to-end data pipelines using **Kafka** for reliable eventing/logging and **Airflow** for orchestration; introduced partitioning and clustering in **BigQuery** to optimize storage/compute cost, cutting query spend by **99%**.
- Designed source/destination data configuration and **SQL** templating with automated **data validation checks**, enabling teams to onboard new data sources 4x faster and reducing migration errors.
- Built monitoring, alerting, and automated remediation workflows for pipeline failures, reducing mean time to resolution (MTTR) from days to hours and improving overall data-pipeline reliability.
- Partnered cross-functionally with product and SRE teams to translate data and reporting requirements into pipeline design, improving deployment reliability.
- **Technologies:** Python, SQL, Kafka, Airflow, GCP BigQuery, GCP Cloud Functions, Spring Boot, AWS, AI agents

CAST Software September 2023 – June 2025

Junior Full Stack Software Engineer Bengaluru, India

- Led design and rollout of an enterprise SSO/OAuth2 authentication framework across distributed microservices, improving secure onboarding and reducing authentication incidents by **>80%**.
- Engineered a high-performance reverse proxy and RESTful API gateway backend that reduced end-to-end API latency by **40%**, improving customer experience for enterprise deployments.
- Contributed to the architecture of a distributed job scheduler with retries and circuit-breaker patterns for fault tolerance, raising successful job completion rates by **30–40%**.
- Queried and modeled data in **PostgreSQL** to support license validation and feature-level access controls for **100+ enterprise users**.
- Proactively remediated critical CVEs and hardened container images to meet enterprise security requirements for customers including the **US Army**, **US Air Force**, and **Google Cloud**.
- **Technologies:** Java, Spring Boot, PostgreSQL, SQL, OAuth2, Docker, Kubernetes, Kafka, AWS

CAST Software January 2023 – August 2023

Software Engineering Intern Bengaluru, India

- Developed a multi-threaded Java log-parsing pipeline that processed large-scale logs in parallel, reducing analysis time for engineering diagnostics.
- Built a VS Code extension to accelerate developer workflows and standardized onboarding for the engineering team.
- **Technologies:** Spring Boot, SQL, RabbitMQ, Kafka, Jenkins

TECHNICAL SKILLS

Languages: Python, SQL, Go, Java, JavaScript, TypeScript, Bash/ZSH, C++
Data Engineering: ETL/ELT Pipeline Design, Kafka, Airflow, BigQuery, PostgreSQL, MongoDB, Data Validation & Quality Automation, Data Modeling
Cloud & AI/ML: AWS, GCP (Cloud Functions, BigQuery), AI Agent Integration, LLM-Assisted Tooling
Practices: System Design, Distributed Systems, Event-Driven Architecture, Reliability Engineering, Observability, Agile
Backend, Containers & DevOps: Spring Boot, Node.js, Microservices, RESTful APIs, Docker, Kubernetes, CI/CD (Jenkins, GitHub Actions), Terraform (basic)

EDUCATION

M.S. Ramaiah Institute of Technology July 2019 – July 2023

Bachelor of Engineering in Information Science Bengaluru, India

- CGPA: 8.82/10.0
- Relevant coursework: Data Structures & Algorithms, DBMS, Cloud Computing, Machine Learning