

Waleed Ahmed

Data Engineering Student

waleedahmed.ca
hello@waleedahmed.ca
linkedin.com/in/waleed-ahmed
github.com/TRAP33ZOID
Toronto, ON, Canada
437-450-9195

RELEVANT EXPERIENCE

Founder — ZoidAI — Sept 2025 – Present

- Architected and deployed a multi-tenant SaaS data pipeline processing **document ingestion, embedding generation, and vector storage** using PostgreSQL with pgvector, LangChain text splitters, and chunking strategies for **RAG-based retrieval systems**
- Designed PostgreSQL schema with **Row-Level Security (RLS)** policies to enforce tenant isolation across **5+ tables**, including documents, call logs, credit transactions, and organization members with role-based access control
- Built real-time analytics dashboards tracking **call volume, cost metrics, and usage patterns** by implementing custom SQL queries with window functions, aggregations, and joins across normalized tables
- Developed **webhook processing pipeline** handling **Stripe payment events** and voice call data streams, implementing idempotent handlers and transaction management to ensure data consistency across distributed systems
- Optimized database performance by implementing proper indexing strategies on high-cardinality columns, reducing query latency by **30-40%** for analytics workloads

Data Analyst Intern — Rijpen — Nov 2023 – Dec 2023

- Built **ETL pipelines** using **Python pandas** to transform and clean insurance policy data from multiple sources, handling missing values, type conversions, and schema normalization for **several thousand records**
- Designed and implemented **PostgreSQL database schema** with proper normalization (3NF) for storing policy, customer, and claims data, writing complex queries with CTEs and window functions
- Created interactive **Tableau dashboards** visualizing policy trends, customer demographics, and claims patterns, enabling stakeholders to identify high-risk segments and optimize pricing strategies
- Performed statistical analysis including hypothesis testing and regression modeling to identify key drivers of policy renewals, presenting findings to cross-functional teams

TECHNICAL PROJECTS

Distributed Analytics Platform — Python, Apache Spark, Databricks, Docker

- Designed a **Databricks-based analytics platform** processing **financial transaction data**, implementing **PySpark jobs** with DataFrame transformations, aggregations, and window functions for time-series analysis
- Built **data quality checks** using Great Expectations, validating schema constraints, null values, and business rules before writing to production tables, reducing downstream errors by **25-35%**
- Implemented incremental data loading patterns using Delta Lake's **MERGE operations**, efficiently updating **dimension and fact tables** while maintaining full audit history
- Containerized Spark jobs with **Docker** and orchestrated workflows using **Airflow DAGs**, managing task dependencies and scheduling for daily batch processing

PROFESSIONAL SUMMARY

Computer Engineering student with hands on experience building production analytics systems and data pipelines. Developed multi-tenant data architecture with PostgreSQL, implemented real-time streaming pipelines with AWS services, and built scalable ETL workflows using Python and Spark. Proficient in designing normalized database schemas, optimizing query performance, and creating analytics dashboards. Actively exploring distributed systems and modern data stack technologies through personal projects and open-source contributions. Seeking to apply data engineering skills to help Ripple improve financial systems through innovative crypto solutions.

SKILLS

Programming Languages

Python, SQL, TypeScript

Data Platforms & Tools

Databricks, Apache Spark, PostgreSQL, pgvector, Delta Lake

Cloud & Infrastructure

AWS (S3, Lambda, Kinesis, Athena, API Gateway), Docker, Serverless Framework

Data Engineering

ETL/ELT Pipelines, Data Warehousing, Stream Processing, Data Modeling, Schema Design, Query Optimization

APIs & Integration

RESTful APIs, WebSockets, Webhook Processing, Server-side Integration, OpenAPI

Frameworks & Libraries

pandas, NumPy, PySpark, LangChain, Flask, Next.js, React

Analytics & Visualization

Tableau, SQL Analytics, Statistical Modeling, Data Quality Testing

EDUCATION

Bachelor of Computer Engineering — Toronto Metropolitan University — May 2026

Diploma of Data Science — Lighthouse Labs — Oct 2023