

Atif Awan

DATA ENGINEER – Python, SQL, Apache Airflow

✉ ateefcollectionsinc@gmail.com

☎ (+1) 905-933-7878

📍 [St. Catharines](#)

🌐 [LinkedIn](#)

SKILLS

- **Languages & Scripting:** Python, SQL, Java, Bash, DAX, JavaScript for building data pipelines, scripting, APIs.
- **ETL & Data Engineering:** Apache Airflow, Spark, dbt, SSIS, Informatica, Talend, Azure Factory, and Databricks.
- **Databases & Warehousing:** SQL Server, Redshift, Snowflake, Athena, BigQuery, PostgreSQL, MySQL, Synapse.
- **BI & Cloud Tools:** Power BI, Tableau, Looker, Seaborn, AWS S3/Glue/Lambda, Azure Synapse, GCP BigQuery.
- **DevOps & Testing:** Git, Docker, Jenkins, CircleCI, Pytest, Great Expectations, CloudWatch, Terraform, and IaC.

WORK EXPERIENCE

Data Scientist

Brainzaa

January 2020 – Present

Remote – Mississauga

- Developed 50+ ingestion flows and batch jobs using Python, dbt, Spark, and Airflow on Databricks, optimizing Azure Synapse data loads, reducing Spark latency by 30% across scheduled orchestration jobs over mixed volume data inputs.
- Constructed 20+ feature stores and schema-managed pipelines within Databricks notebooks, verifying 100+ metrics via CI/CD hooks and GitHub workflows, maintaining thresholds by enforcing 15 checkpoints using custom drift flags.
- Produced A/B datasets from 500K+ records using synthetic sampling, post-stratified summaries, churn tables, Created across Redshift and BigQuery with optimization factors computed from matrix models and 10 segmentation loops.
- Automated retraining and deployment via MLflow, Terraform, GitHub Actions on Azure, GCP using IaC with 5 APIs.

Data Analyst

Entrycloud

June 2018 – December 2024

St. Catharines

- Extracted CRM logs and customer event trails using SQL, built 50+ views to quantify retention factors, and designed metrics for churn prediction across trend surfaces using key patterns collected from over one million monthly records.
- Designed 15 Power BI dashboards with row-level DAX and slicer logic, delivering pipeline metrics from customer and usage logs; ensured access restrictions, calculated KPIs align with constraints for sales, billing, and performance metrics.
- Refined 30+ report generation scripts using Python, NumPy, and Pandas with validations tied to triggers; replaced stored procedures with modular query scripts across ETL stages to remove excess load steps from daily refresh cycles.
- Authored datasets by linking structured and semi-structured sources; logged 100+ anomalies and stored outputs.

INTERNSHIP

Software Engineer Intern (Data Engineering Focus)

Meridian Credit Union

July 2019 – July 2020

Ontario

- Optimized 30+ SSIS ETL pipelines using SQL Server, reducing data mart refresh latency by 60% under SLA constraints.
- Deployed Dockerized Java microservices via Terraform on AWS, automating 95% infra provisioning with DevOps team.
- Developed REST APIs with embedded audit logs, enabling 100% secure data access for analytics and reporting tools.
- Documented 12+ data workflows, API specs, and infra diagrams, increasing inter-team integration velocity by 40%.
- Collaborated in Agile sprints and tuned backend queries, reducing execution time by 55% across 8 key modules.

PROJECTS

Motor Match

Cross-Platform Mobile App

- Composed mobile UIs via React Native, added routing for 10+ screens, delivered store-ready builds for both platforms.
- Coded image upload, login control, swipe compare, Redux logic to handle 1000+ updates from real-time form edits.
- Enabled shortlisting flows and chats by testing asset filters, resolving 20+ permission blocks on camera/location.

Housing Price Prediction

Machine Learning Model – Python

- Built smart regression pipelines with NumPy, Pandas; mapped 5 features like lot size, cooling, and bedroom count.
- Cleaned 10K+ housing rows, neatly plotted trends via Matplotlib, compared RMSE and R² for tree vs linear models.
- Coordinated 500+ model queries by scaling all inputs, predicting values, and respecting housing type constraints.

EDUCATION

Bachelor of Arts – Sociology

Brock University, St. Catharines, Ontario

September 2019 – October 2025

Certificate in Data Science

Intelle College, Mississauga

May 2023 – November 2023

CERTIFICATIONS

- **Microsoft Certified: Power BI Data Analyst Associate**