

Hazem Mohamed

hazem.horus.7@gmail.com | 438-409-1091 | [linkedin](#) | [github](#)

WORK EXPERIENCE

Azure Data Engineer

Sep 2024 – Apr 2025

IRCC (Immigration, Refugee and Citizenship Canada)

Montreal, Canada

- Engineered ETL/ELT pipelines using Azure Data Factory, Python, and Azure Functions to ingest and transform 1TB+ of structured and unstructured data into Azure Data Lake Storage (ADLS), reducing storage costs by 25%.
- Designed Power BI dashboards with modeled operational datasets to track data metrics and KPIs across 2,500+ users.
- Built Python data validation pipelines across 40+ federal datasets, enforcing PII compliance and governance standards.
- Built Azure DevOps/Terraform CI/CD for data pipelines across dev/staging/prod, cutting deploys from 2hr to 15min.
- Designed dimensional data models for 20+ federal assets, powering 12+ downstream Power BI reports.
- Profiled datasets in Python/Pandas, surfacing pipeline bottlenecks and quality gaps that drove 4 infrastructure decisions.

Data Engineer

Jan 2023 – Aug 2023

SEG Products

Montreal, Canada

- Built Azure DevOps CI/CD for pipeline deployment and schema migration, cutting releases from 3 days to 1 hour
- Built 8+ Power BI dashboards on SQL Server data models, replacing weekly manual sales/ops reporting.
- Automated Python/REST API ingestion across 5+ source systems, cutting manual data workload 40%.
- Executed 3 E2E SQL Server warehouse migrations with ETL and rollback, achieving minimal downtime.
- Redesigned product catalog star schema ERDs, improving query performance 60% and analytics reliability.
- Developed Python validation and anomaly flagging scripts, catching 95%+ of ingestion errors pre-reporting.

PROJECTS

Facial Expression Recognition | *PyTorch, CNNs, OpenCV*

- Trained a real-time facial expression classifier using convolutional neural networks in PyTorch; applied data augmentation, K-fold cross-validation, bias evaluation, and latency-aware inference optimization for accurate real-time predictions.

AI Invoice-to-Statements Web App | *LLMs, NLP, Python, REST APIs*

- Built an end-to-end application leveraging LLM pipelines and NLP to parse heterogeneous invoice formats, extract structured fields, and generate standardized financial statements automatically.

RFID-Tracked Sample Management | *Python, Pandas, Automation*

- Engineered a Python-automated tracking system with RFID integration to classify and log test samples, reducing manual data entry and improving retrieval time by 56%.

Canadian Household Purchasing Power Pipeline | *Python, SQL, Pandas, AWS*

- Engineered an ETL pipeline harmonizing 110 years of Canadian income data across 4 sources, integrating 2,112+ household records with group-based imputation for 1,708 missing values and Z-score outlier detection.

EDUCATION

Concordia University

Montreal, Canada

Bachelor of Engineering, Software Engineering

Sep 2021 – Apr 2026

- Member of the Institute for Co-operative Education

Vanier College

Montreal, Canada

Health Sciences College Diploma (Honors)

Sep 2019 – May 2021

SKILLS

Languages: Python, SQL, PySpark, Bash, C#, JavaScript, TypeScript, Git

Azure Data: Data Factory (ADF), Data Lake Storage (ADLS), Synapse Analytics, Functions, Logic Apps, Databricks

Data Engineering: ETL/ELT Pipelines, Apache Airflow, Data Modeling/warehousing, Star Schema, Dimensional Design

Databases: SQL Server (SSMS), PostgreSQL, MySQL, T-SQL, Query Optimization

DevOps & IaC: Terraform, Docker, Kubernetes, CI/CD (YAML), PowerShell

Practices: Agile/Scrum, Secure Data Handling (PII), Model Evaluation, Disaster Recovery