

Samson Tran

Samson.Tran@hotmail.ca ❖ (647) 562-8730 ❖ Toronto, ON

PROFESSIONAL SUMMARY

Experienced data engineer with over 3 years developing robust ETL pipelines using languages such as Python, SQL, and Spark. My work experience in big data using various cloud technologies, like Databricks and AWS, to support avionics systems and aircraft engine performance demonstrates my ability to transform data and gather insights.

WORK EXPERIENCE

Pratt & Whitney Canada (contract under Belcan Inc.)

Dec. 2022 – Present

Data Engineer / Scientist

Mississauga, ON

- Build large-scale ETL pipelines using cloud platforms (Databricks) and on-premises systems (SQL Server, SSIS) to ingest and transform batch and streaming data for PowerBI dashboard reporting
- Develop software applications using Python and SQL to automate aviation performance monitoring, while performing code reviews and testing to ensure high-quality and maintainable code
- Develop and optimize PySpark code for big data fleet investigations, processing historical data using Databricks and Delta Lake to efficiently summarize key engine health and operational metrics
- Research and implement best data modeling strategies to build a new cloud data warehouse in Delta Lake, and upgrade existing on-prem SQL database. Focus on optimum storage costs and end-user query performance
- Present and document proposed ETL pipeline workflows and database architecture to communicate project strategy to upper management and lead alignment among all team members
- Manage the onboarding and training of new data engineers, scientists and analysts, concentrating on the development of critical technical skills to improve and expedite data processing and analysis capabilities

Pratt & Whitney Canada (contract under University of Toronto)

May 2022 – August 2022

Data Science Intern

Mississauga, ON

- Developed Python scripts to implement time-series forecasting algorithms for predicting engine health performance, tested using Pytest and deployed in AWS Lambda
- Developed efficient SQL scripts to extract and visualize statistical results from large engine fleet datasets, ensuring accurate results through validation and improving performance
- Managed version control of Python scripts using Git, including branching, merging, and resolving conflicts, to ensure smooth collaboration and maintain code integrity across the development team

Simpson Gumpertz & Heger

August 2019 – August 2021

Associate Project Consultant

San Francisco, CA

- Managed schedules, progress, and budgets for multiple large-scale projects simultaneously by working closely with cross-functional teams and clients to ensure adherence to design specifications and documents

TECHNICAL SKILLS

- | | | |
|--------------------------------|--------------|----------------|
| ■ Python | ■ SQL | ■ PySpark |
| ■ AWS (Glue, Redshift, Lambda) | ■ Databricks | ■ ETL pipeline |

EDUCATION

University of Toronto

June, 2022

Master of Engineering (*Industrial Engineering with an Emphasis in Analytics*)

GPA: 3.98/4.0

- Relevant Coursework: Cloud-Based Data Analytics (Big Data), Data Analytics and Science, Machine Learning, AI in Finance, Financial Management for Engineers

University of Toronto

June, 2019

Bachelor of Applied Science