

Jeremy Gonsalves

Jeremygonsalves98@gmail.com | 647-229-2249 | 290 Adelaide St W, Toronto ON | <https://jeremygonsalves.com/>

Data Engineer / Business Intelligence Engineer with 3+ years of experience building cloud-based data pipelines, data warehouses, and analytics solutions using Python, SQL, Databricks, Airflow, and Power BI. Proven track record of improving data quality, ETL performance, and executive reporting, supporting stakeholder decision-making across financial and consumer data products. Strong background in applied mathematics, automation, and scalable data architecture.

EDUCATION

Applied Mathematics Engineering (BAsC), Queen's University, Kingston, ON

September 2019 - April 2024

- Specialized in Mathematics and Mechanical Engineering with First Class Honours (3.51 cGPA).

WORK EXPERIENCE

Business Intelligence Analyst, TD Bank, Toronto, ON

June 2025 - Present

- Designed and deployed scalable **ETL/ELT pipelines** in **Python, SQL, and Databricks**, reducing manual reporting effort by **20%** and improving data quality and reliability for Chequing and Saving accounts analytics.
- Collaborated with data engineering and business stakeholders to automate analytics workflows, improving **data** delivery SLAs by 30% and accelerating executive KPI reporting in Power BI.
- Developed a high-volume **PySpark ETL pipeline** ingesting 703K+ customer records, enabling tracking of \$1.5B in growth over 3 months and identifying 334K+ campaign-eligible customers, while enforcing data validation checks to maintain high data quality standards.

Analytics Engineer, Credit Sesame, Toronto, ON

July 2024 - May 2025

- Engineered and optimized **ETL/ELT** workflows in **Python, SQL, and Airflow**, integrating data from **AWS Redshift** and other sources into a centralized data warehouse, improving revenue reporting accuracy and efficiency by **20%**, measured through reductions in manual reconciliation time, report refresh latency, and post-release data discrepancies across monthly revenue reports.
- Reduced **ETL** processing time by 30% by developing **Airflow** DAGs and production-level **Python** scripts to automate weekly data pulls for executive-level reports on metrics like total new users enrolled, revenue from fees, and subscription revenue. Leveraged Git version control, ensuring efficient script deployment and enhancements.
- Developed ad hoc **SQL** queries and custom **Python** scripts to extract data based on key metrics such as total funding users, fraud status, and other critical KPIs, improving intraday performance monitoring and enabling real-time analysis to support business needs and guide strategic decisions.

Lead Developer & Founder, Pipeline AI, Toronto, ON

May 2025 - Present

- Founded and built a scalable **SaaS** data automation platform using **Python, Airflow, and AWS Redshift**, delivering **ELT**-based data pipelines that transform raw webhook and application data into analytics-ready datasets for health professionals.
- Led full-cycle stakeholder management, working directly with physiotherapists and clinic operators to perform requirements gathering, and translate business workflows into automated analytics solutions.
- Architected end-to-end business intelligence solution leveraging Slack API, Google Workspace integration, and **AWS EC2** deployment with SSL certificates, comprehensive logging, and 24/7 monitoring - delivering automated performance analytics that reduces manual reporting time by 90%.

Software Developer Co-op, Ontario Teachers' Pension Plan, Toronto, ON

January 2023 - August 2023

- Promoted from a software developer intern to a full-time Analyst position.
- Reduced data processing time by 30% by developing **Power BI** reports that integrated data from **SQL, Excel, and APIs**. Extracted data via an **API** into a **JSON** file, efficiently transported and utilized for generating dynamic insights on resource hiring timelines, defect analyses, and hiring trends. Improved data accuracy to 97% by validating and refining **Python** scripts, shared and version-controlled through **GitHub**.
- Presented critical metric updates to IT executives using **Power BI**. Enhanced executive decision-making and improved client satisfaction by providing detailed analysis of retention rates and ticket volume.

Teaching Assistant, Queen's University, Kingston, ON

September 2021 - April 2024

- Directed interactive tutoring sessions for 100+ engineering students in Calculus to foster essential learning skills.

Founder, Jeremy's Tutoring Academy, Aurora, ON

September 2020 - June 2025

- Founded and managed a successful tutoring business which generated over \$20,000 in profit across North America.
- Promoted the Academy and maintained continuous student engagement through online educational videos.

PERSONAL PROJECTS

Real-Time Stock Prediction System | *Python, Machine Learning*

June 2025 - December 2025

- Engineered an ensemble trading model using **Python**, combining **LSTM** neural networks, **XGBoost**, and sentiment analysis, processing 50+ technical indicators and real-time Reddit and news API data, to forecast stock prices 4 hours ahead, delivering a 42% accuracy out of 100 trials.
- Developed a real-time trading alert system automated via **crontab**, generating daily buy/sell signals at noon, distributing **Slack API** notifications with interactive charts; achieved 70% confidence thresholds by validating across Yahoo Finance data, sentiment feeds from Alpha vantage and Reddit, and market microstructure analytics.

NHL Draft Prediction | *Python, Machine Learning*

June 2024 - March 2025

- Leading a research initiative using **Python**, integrating advanced libraries such as **TensorFlow** and **NLTK** to develop predictive models, including **Support Vector Machines** and **Linear Regression**. These models forecast NHL team draft selections based on qualitative data from sports analysts' commentaries.
- Achieved an 82% accuracy rate in predicting 200 candidates for the 2023 NHL draft year, validating the model's performance and relevance.

Financial Tracker | *Python, Machine Learning, Excel*

September 2024 - December 2024

- Automated statement aggregation with an **Airflow DAG** that scans downloads for new statements, consolidates data, and uses an **XGBoost** model to classify transactions - reducing categorization time by 80% and achieving 95% accuracy on 3,000+ transactions.
- Built an interactive dashboard in **Excel** to visualize categorized spending across multiple categories, enabling data-driven insights into personal financial management, reducing unnecessary costs and saving 20% of personal budget.

Autonomous Aquaculture Cage Navigation Capstone | *MATLAB, Simulink*

September 2023 - April 2024

- Developed a dynamic model for an autonomous aquaculture cage using principles from marine hydrodynamics and maneuvering theory. This model was implemented in **MATLAB** and **Simulink**, accurately simulating the cage's behaviour under various oceanic conditions, including noiseless and noisy environments.
- Designed a Linear Quadratic Regulator (LQR) controller that optimized the cage's path-following ability, achieving a deviation of no more than 2% from the desired path. The controller was rigorously tested on circular, linear, and octagonal paths, demonstrating high accuracy and stability, successfully converging the Algebraic Riccati Equation (ARE) in MATLAB.

VOLUNTEER EXPERIENCE

Data Analytics Specialist, Queen's Sports Analytics Organization (QSAO), Kingston, ON **September 2023-April 2024**

- Analyzed sports datasets using Python, web scraping, and statistical models and stored insights in SQL.
- Used machine learning and Large Language Models (LLMs) for predictive analytics to ensure data accuracy.

Project Manager-Youth Innovation Success, Enactus Queen's University, Kingston, ON **September 2023-April 2024**

- Led 20+ students to empower local youth with entrepreneurial skills. Managed high school teaching roles.

SKILLS & HOBBIES

- **Technical Skills:** Python, SQL, PySpark, Airflow, Databricks, Azure, AWS, Power BI, Looker, Excel, ETL/ELT, Data Warehousing, Data Modeling, Dimensional Modeling, APIs, Git control.
- **Additional Information:** Triathlete, Chess (1200 Elo), Language learner: French (Fluent), Spanish (Intermediate)