Sebastian Gonzalez

📞 778-227-4225 ✉ sebastiangm.dev@gmail.com 🌐 www.sebastiangm.com 🔗 linkedin.com/in/sebastian-sgm/

Professional Experience

Spring Financial   **February 2025 – July 2025**
Business Analyst & Data Operations Associate   *Vancouver, Canada*
- Built Python ingestion pipelines for credit-risk datasets from AWS S3, PostgreSQL, and partner APIs, standardized schemas, and validation for downstream ML.
- Collaborated closely with Data Science teams to define training/inference data contracts, feature readiness checks, and performance monitoring dashboards.
- Implemented automated data-quality tests (nulls, ranges, referential integrity) and alerting to ensure model/data reliability.
- Drove sprint planning and backlog hygiene to align business priorities with data deliverables.

Canfor   **June 2024 – February 2025**
Credit & BI Analyst   *Vancouver, Canada*
- Developed and maintained near real-time Power BI dashboards (credit utilization, aging, bad debt KPIs) for a $50M+ portfolio.
- Built ELT from JD Edwards (Oracle) to Azure SQL via ADF/Python with validations to protect reporting SLAs.
- Reduced Days Sales Outstanding (DSO) by 8 days through recurring analyses and operational triggers informed by modeled metrics.

Autofin — Auto Financing and Credit   **March 2021 – January 2022**
Data Engineer   *Santiago, Chile*
- Built ETL pipelines with Apache Airflow and Apache Spark to integrate CRM, loan applications, and credit-bureau feeds.
- Managed an AWS S3 data lake and curated Bronze, Silver, and Gold layers, which improved query speed and data availability.
- Implemented Git-based version control and CI/CD to promote changes safely and reduce deployment risk.

TANNER Financial Services   **January 2018 – February 2021**
Data Scientist   *Santiago, Chile*
- Maintained Airflow and Spark ETL to unify operational data, and engineered features and gold tables that supported modeling and customer segmentation.
- Automated routine workflows, reduced manual processing by 20%, and shortened cycle time for A/B tests and model updates.

Projects

Phishing Detection Data Pipeline 🔗 | *Python, FastAPI, AWS (S3/ECR/ECS), Docker, GitHub Actions, MLflow*
- Delivered a real-time API that scores suspicious URLs, cutting analyst triage time from minutes to seconds (FastAPI on ECS/ECR, artifacts in S3).
- Built an end-to-end MLOps workflow with data-quality checks and drift monitoring, and released with Docker, GitHub Actions, and MLflow.

Retail Analytics Pipeline (Bronze → Silver → Gold) 🔗 | *ADF, ADLS Gen2, Databricks (Spark/Delta), Synapse, Key Vault*
- Built an end-to-end analytics path from raw files to Delta tables to daily marts in Synapse, so teams could track orders, customers, and revenue without ad-hoc exports.
- Defined a simple star schema and basic data-quality rules, which lowered maintenance effort and kept dashboard queries fast.

Technical Skills

Data Engineering & Orchestration: Airflow; Spark/PySpark (Azure Databricks); dbt; ETL/ELT
Cloud & Warehousing: AWS (S3, Lambda, EMR, ECS/ECR); Azure (Data Factory, ADLS Gen2, Synapse, Azure SQL); Docker; Snowflake; CI/CD (GitHub Actions)
Modeling & Storage: Star/Dimensional; PostgreSQL/MySQL/MongoDB
Programming & BI: Python (pandas, FastAPI), SQL; Power BI, Tableau
Dev Practices: Git; Agile/Scrum; MLflow

Education

Douglas College   **2022 – 2023**
Postgraduate Diploma in Computer and Information Systems   *Vancouver, Canada*

European University of Madrid   **2016 – 2017**
Master of Business Administration (MBA)   *Madrid, Spain*

Andres Bello University   **2012 – 2015**
BSc in Economics (Concentration in Finance and Statistics)   *Santiago, Chile*