Rafael A. Espinoza

Cambridge, MA | (915) 222-9927 | rafa3l.espinoza@gmail.com | linkedin.com/in/rafael-espinoza-2001/

EXECUTIVE SUMMARY

Data Analytics Consultant with experience building and operating **production-grade data pipelines and analytics systems** across public and private sector clients. Strong foundation in **Python and SQL**, with hands-on ownership of data ingestion, validation, modeling, and delivery under strict accuracy, regulatory, and business constraints.

WORK EXPERIENCE

Data Analytics Consultant Feb 2025 – Present

Guidehouse *Boston, MA*

- Designed and executed **automated batch data workflows** processing **3M+ records per month** across 9 concurrent client engagements, reducing **data delivery turnaround by 70%** for audit and compliance workflows
- Implemented **confidence-based entity matching and validation logic** across multiple authoritative sources (state vital records, death master file, large-scale obituary data), achieving **99% accuracy** and supporting regulatory compliance
- **Customized matching thresholds and confidence scoring** per client to balance false positives and false negatives in alignment with business and regulatory requirements
- Built and maintained **SQL-based transformations and data quality checks**, standardizing multi-source data and reducing data errors by **30%**
- Delivered **validated, audit-ready datasets** with confidence and reconciliation metrics, enabling compliance teams to proactively address high-risk records

TECHNICAL PROJECTS

StreamWatch | *Python, SQL, Airflow, LightGBM, Google Cloud Storage, Streamlit*

- Built an **end-to-end forecasting pipeline** predicting weekly TV show popularity across multiple horizons, from ingestion through model training and serving
- Orchestrated automated data preparation and model retraining using **Dockerized Airflow DAGs**, publishing versioned datasets and models to Google Cloud Storage for reproducible deployments
- Designed a **training–serving contract** using immutable artifacts and a manifest-driven interface, reducing deployment errors by **30%**
- Monitored model performance (RMSE, R^2) across forecast horizons to assess stability and decision reliability
- Developed an **interactive Streamlit dashboard** for forecast visualization and stakeholder adoption

Smart Cabbie: Chicago Taxi Analytics | *BigQuery SQL, Python, Tableau*

- Queried and transformed **10M+ Chicago taxi trip records** using BigQuery SQL and Python-based ETL pipelines
- Modeled trip-level and aggregate metrics to analyze **demand patterns, revenue efficiency, and tipping behavior**
- Developed interactive Tableau dashboards and KPIs to support data-driven operational insights

TECHNICAL SKILLS

Languages – Python, SQL, Scala, Java
Technologies – Airflow, Docker, ETL pipelines, SQL-based data modeling, BigQuery, Spark, Pandas, NumPy, Tableau, Streamlit, Google Cloud Platform, Git

EDUCATION

Boston College Chestnut Hill, MA

B.S. Computer Science, Minor in Mathematics *May 2024*

ADDITIONAL INFORMATION

Certifications – Microsoft Career Essentials in Data Analysis (Oct 2024); LinkedIn: Advanced SQL for Data Scientists (Oct 2024), Tableau Essential Training (Oct 2024), Python Data Structures and Algorithms (Feb 2023)
Languages – English (Bilingual), Spanish (Native)
Interests – Outdoors, hiking, live music, pickleball, weightlifting, running