

# Mario Paulin Vega

mpaulinv@gmail.com    linkedin.com/in/mario-paulin-vega  
github.com/mpaulinv  
Unión 28, Miguel Hidalgo, CDMX, Mexico

## Education

---

### **The University of Texas at Austin**

MSc in Data Science, December 2024

MSc in Economics

### **ITAM**

Master in Economic Theory

Bachelor in Economics, Finance Track, With Honors

## Professional Experience

---

### **Team Lead**

*CitiBanamex, Anti-Money Laundering Division, Mexico City, Mexico*

June 2020 – May 2024

- Led a team of 3 data scientists in validating a regulatory model inventory of around 100 models across Mexico, LATAM, the US, and Canada, ensuring full compliance with audit and regulatory timelines and achieving zero findings from Internal Audit and external regulators.
- Partnered with Business, Technology, and Audit teams to integrate validation insights into the compliance process, enhancing model quality, data integrity, and strategic monitoring decisions.
- Developed and implemented process improvement initiatives and methodologies to streamline global operations.
- Mentored and trained team members, enhancing performance and fostering a culture of continuous improvement.
- Contributed to projects that involved code reviews, data reconciliation, and verification of algorithm specifications.
- Presented an initiative to senior compliance management to reduce by 15% time spent on Manager Control Assessment across the Bank.

### **Sr. Statistician**

*CitiBanamex, Anti-Money Laundering Division, Mexico City, Mexico*

June 2017 – May 2020

- Prepared validation reports on predictive and rule-based models for regulatory compliance, achieving zero observations from Internal Audit and external regulators.
- Conducted data extraction, code review, replication, and quality assessments for models in production and test environments.
- Applied statistical methodologies such as Logistic Regression, Decision Trees, Random Forest, Lasso, and K-means clustering to transactional and client profile data.
- Executed a global project automating data input stability assessment, reducing time spent by 50%.

## Independent Consultant

*Queretaro, Mexico*

August 2016 – May 2017

- Designed and implemented automated data analysis scripts for client presentations on structural transformation.
- Developed econometric models to analyze labor transition probabilities in Colombia.

## Academic Research and Teaching

*The University of Texas at Austin & CIE, Mexico City*

2011 – 2016

- Over 100 hours teaching undergraduate and M.A. economics.
- Analyzed export dynamics using firm-level panel data and econometric models.

## Impact Evaluation Analyst

*CONEVAL, Mexico City, Mexico*

2009 – 2010

- Evaluated social programs using DiD, RDD, matching, and panel data methods.
- Estimated treatment effects and led user satisfaction assessment projects.

## Skills

---

**Programming:** Python (Pandas, Numpy, Scikit-learn, Pytorch), R, SAS, Stata

**Big Data and Cloud:** SQL, Hive, Spark, Hadoop, GCP

**Data Visualization:** ggplot2, Matplotlib, Seaborn, Power BI

**Machine Learning:** Regression, Clustering, Decision Trees, Random Forest, Neural Networks

**Statistical Analysis:** Hypothesis Testing, Causal Inference, Econometric Methods

## Certifications and Courses

---

Harnessing Language Models: Your Path to NLP Expert (2024)

Certified Anti-money Laundering Specialist – ACAMS (2021)

## Languages

---

Spanish (Native), English (Fluent)

## Projects

---

### Cryptocurrency AML Detection 2025

Machine learning models for Bitcoin transaction classification using graph neural networks on the Elliptic dataset, combining AML domain expertise with deep learning techniques for financial crime detection.

### Clinical Risk Prediction 2025

Developed ensemble models for cardiovascular risk assessment using clinical indicators, demonstrating healthcare analytics capabilities and model interpretability for medical decision support.

### Chess Analytics Platform 2025

Built an end-to-end analytics dashboard for chess game analysis using the Lichess API, showcasing full-stack data science skills from data ingestion to interactive visualization.