Mauro Florez

Houston, TX | (346) 610-0770 | linkedin.com/in/mauroefr | mauroflorez.github.io

EDUCATION

Rice University, *Houston, TX* — May 2025
PhD in Statistics — GPA: 3.94/4.00

Rice University, *Houston, TX* — August 2024
MA in Statistics — GPA: 3.94/4.00

Universidad Nacional de Colombia, *Bogotá, Colombia (COL)* — June 2019
BS in Statistics — Ranked #1 in class

Universidad Sergio Arboleda, *Bogotá, COL* — September 2017
BS in Mathematics — Honors: 75% tuition waiver scholarship

SKILLS

Quantitative Skills: Data Analysis, Statistical Modeling, Machine Learning, Bayesian Statistics, Computational Statistics
Computer Skills: R, Python, Matlab, Tableau, SQL, Latex, MS Office
Language Skills: Fluent in English, Fluent in Spanish, Beginner in Italian

SELECTED DATA ANALYSIS PROJECTS

A Multivariate Model for Analysis of Correlated Count Data, Rice University, *Houston, TX* — Aug. 2022 - May 2023

- Designed and implemented a Bayesian model to analyze correlated count data, capable of handling data with any type of dispersion and outperforming traditional Negative Binomial and Poisson models.
- Authored a research paper in the Journal of Quantitative Analysis in Sports, showcasing the model's advantages and its practical application to real-world sports data.
- Developed and published the R package *MultRegCM* on the CRAN repository, enhancing accessibility and reproducibility.

Soccer Betting Model - Machine Learning Model, Rice University, *Houston, TX* — Jan. 2021 - May 2021

- Developed and applied machine learning models to predict the number of cards a referee shows in soccer games.
- Employed web scraping techniques in R to gather data on soccer games across 20 years.
- Demonstrated that the proposed Neural Network model outperforms commonly used models in predicting the number of cards in soccer games. Attaining a positive predicted value of 75.23%.

Lung Disease Risk Prediction in a Coal Mine, Universidad del Rosario, *Bogotá, COL* — May 2017 - Sep. 2017

- Formulated a Survival Model to predict the risk of contracting a lung disease in one of the largest coal mines in America
- Conducted the cleaning data process of the data and proposed a methodology for the imputation of missing data in the longitudinal study comprising more than 300 workers spanning over 20 years
- Partnered with interdisciplinary teams to gather and analyze relevant data, ensuring the validity and accuracy of the implemented predictive model

SELECTED WORK EXPERIENCE

Instructor - Rice University, *Houston, TX* — Jun. 2022 - Aug. 2022

- Taught probability and statistics course (STAT 310), teaching 26 students through prepared and recorded lessons, adapting teaching methods to accommodate various learning styles, and ensuring a comprehensive understanding of the material.
- Rated as “Outstanding” instructor by 65% of students
- Developed strong communication and presentation skills by effectively conveying complex statistical concepts to diverse groups of students

Data Analyst - Department of Science, Technology and Innovation (Minciencias), *Bogotá, COL* — May 2019 - Jun. 2020

- Maintained and updated institutional information databases in SQL to ensure accuracy and relevance
- Collaborated in the collection, consolidation, and refinement of data for precise and up-to-date reporting
- Designed and implemented information dashboards on the Tableau platform for data analysis, supporting internal and external decision-making needs of the organization

SELECTED PUBLICATIONS & PROJECTS

- **Florez, M.**, Gottard, A., Guindani, M., Vannucci, M. (2024). Bayesian Mixed Graphical Model. Manuscript in preparation.
- **Florez, M.,** Guindani, M. & Vannucci, M. (2024). Bayesian bivariate Conway–Maxwell–Poisson regression model for correlated count data in sports. Journal of Quantitative Analysis in Sports. aop.
- Otálora-Otálora, B. A., **Florez, M.**, López-Kleine, L., Canas Arboleda, A., Grajales Urrego, D. M., & Rojas, A. (2019). Joint transcriptomic analysis of lung cancer and other lung diseases. Frontiers in Genetics, 10, 1260