Jaime Andrés Castañeda
Data Scientist
jacasta2@gmail.com | +57 (300) 222-4658 | Colombia | linkedin.com/in/jaime-data-science

Summary of Qualifications

- Researcher with about 15 years of experience applying analytics to study decision-making in supply chains
- Data science and machine learning training with Correlation One and DeepLearning.AI using Python and SQL
- Rigorous, organized and ready to give a hand to my teammates and learn from them

Skills

- **Programming languages:** Python, SQL, Java
- **Analytics:** data wrangling, data visualization, descriptive and inferential statistics, econometrics, machine learning
- **Tools:** Python (pandas, numpy, matplotlib, seaborn, scikit-learn, keras, FastAPI, streamlit), Postgres, Stata, SPSS, Excel, Power BI, Git, Qualtrics, LimeSurvey

Relevant Experience

Data Scientist, *Colombian ICT Ministry*   Mar 2023 – Present | Remote

- Build a composite index for prioritizing internet connectivity initiatives for schools using multiple techniques such as principal components analysis and factor analysis
- Coordinate the creation of a data governance policy for the team to streamline the development of analytics-related projects
- Draft guidelines and develop templates to support the implementation of the data governance policy (e.g., guidelines and template for metadata documentation)

Data Science Fellow, *Correlation One*   Mar 2022 – Jul 2022 | Remote

- Collaborated with five teammates to analyze motorcycle accidents in Bogotá for prioritizing road safety operations, pulling accident data from a GIS web service into a local Postgres database using Python and SQL
- Analyzed more than 200,000 rows of accident data, creating heatmap visualizations based on crosstabs that relate accident type, accident severity and vehicles involved using SQL, Python and Seaborn
- Developed an unsupervised machine learning (ML) model by engineering continuous and categorical features using the K-Prototype algorithm in Python, defining prioritization levels for road safety operations

Researcher, *Universidad del Rosario*   Apr 2016 – Jan 2022 | Bogotá, Colombia

- Co-led a funded research project that evaluated the role of expediting prices in inventory replenishment, implementing an in-person and online behavioral experiment with 400+ participants
- Examined more than 25,000 rows of replenishment decisions, creating customized box plot visualizations of the decisions grouped by expediting prices and averaged over subjects using Python and Seaborn
- Fit several supervised ML models using random-effects panel data models in Python, providing insight into how expediting prices impact replenishment decisions

Additional Experience

- Deployed a Streamlit app that compares the influence of a pair of NBA teammates' stats on their team's winning chances using logistic regression and SHAP values — 2023
- Implemented an API endpoint for my data science fellowship project using FastAPI to allow users to retrieve predictions from the clustering model — 2023
- Built a data pipeline module for my data science fellowship project to update the local database with monthly accident data from the GIS web service — 2023
- Evaluate LaCardio's covid-19 vaccination process using simulation and optimization, finding process configurations that can reduce patient waiting times by ~50% — 2021
- Simulated the operation of a cross-docking facility for BAT Colombia, finding operational configurations that allow to dispatch the fleet in less than half an hour — 2019

Education

Ph.D., Economics, *Università della Svizzera italiana*   2013 | Lugano, Switzerland

Personal website