Lina Florez

florez.lina0123@gmail.com ❖ linaflorez.github.io ❖ Pittsburgh, PA 15206

SUMMARY

A highly motivated, versatile data scientist with a strong focus on data analysis, machine learning, and data visualization using Python, SQL, and Tableau, looking forward to making meaningful contributions to an organization's data-driven mission.

SKILLS

Programming Languages: Python • SQL/MySQL • Javascript (including Google Apps Script) • Tableau • Google Sheets
Technical Skills: Statistical Analysis • Machine Learning • Data Visualization • Data Mining
Packages & Tools: Numpy • Pandas • Matplotlib • bash • Scipy • Jupyter • scikit-learn
Natural Languages: Spanish (fluent), English (fluent)

EXPERIENCE

Rapid TPC **June 2023 – Present**
Software and Production Operations Intern *Pittsburgh, PA*

- Spearheaded the creation and development of a FLASK API seamlessly integrating with Google Apps Script to automate production blueprint generation, resulting in enhanced efficiency and robust functionality.
- Facilitated recurring meetings to ensure the seamless integration of additional features into the production templates.
- Coordinated the integration of the improved blueprints with concurrent production runs.

University of Pittsburgh **August 2020 – April 2023**
Graduate Research Assistant *Pittsburgh, PA*

- Analyzed and manipulated multi-dimensional data of nearby galaxies to forecast observations of distant objects
- Processed, analyzed, and visualized spectroscopic data from thousands of galaxies in the LEGA-C astronomical survey
- Thrived in collaborative environments to learn, address challenges, and implement strategies that produced successful outcomes
- Instructed over 300 students in Introduction to Astrophysics as a teaching assistant

Princeton University **June 2017 – May 2020**
Research Assistant *Princeton, NJ*

- Optimized the merging and analysis of astrophysical databases using pandas and SQL, to explore quantitative correlations between the visible and X-ray properties of black holes
- Performed principal components analysis (PCA) to model sky emission and optimize foreground removal algorithms
- Communicated complex research methodologies and results for multiple presentations to diverse audiences

PROJECTS

DESI-ML

- Developed predictive models for the Dark Energy Survey Instrument (DESI) using Random Forest and XGBoost
- Applied machine learning regression models to DESI spectra to predict physical characteristics of galaxies
- Future developments include regressing more parameters to determine which models produce the best overall predictions

Sephora Analysis

- Analyzed Sephora merchandising data using SQL and pandas to determine rising brands and product types
- Currently working on a product recommender system using Python
- Future work includes performing a sentiment analysis of various brands and product types using natural language processing (NPL)

EDUCATION

University of Pittsburgh **August 2020 – April 2023**
Master of Sciences, Physics *Pittsburgh, PA*
Relevant Coursework: Astrostatistics, Classical & Quantum Mechanics, Thermodynamics & Statistical Mechanics

University of Illinois at Urbana-Champaign **August 2016 – May 2020**
Bachelor of Science, Astronomy *Urbana, IL*
Relevant Coursework: Calculus III, Linear Algebra, Computing in Astronomy, Computing in Physics