Miguel Rivera-Lanas

mriveral@andrew.cmu.edu | (703) 981-0554 | Website: mriveralanas.com | GitHub: rivera-lanasm

Data Engineer passionate about reproducible and scalable research methods with 5 years of experience in asset management and academic research settings. Works effectively in autonomous and project management roles. Experience designing and implementing cloud-based data science solutions, including web scraping, feature engineering, data warehousing, exploratory data analysis, machine learning, and data visualization. Passionate about public policy research and public interest technology.

EDUCATION

Carnegie Mellon University, Heinz College, Pittsburgh, PA   **May 2025**
M.S. in Public Policy & Management – Data Analytics

University of Pennsylvania, Philadelphia, PA   **May 2017**
B.A. in Economics, Minor in Statistics, International Relations, (GPA: 3.5/4)

CORE COMPETENCIES

Scripting: Python, R, Scala, Bash **Devops:** Git, Airflow, Docker, Jenkins **Cloud:** AWS (S3, EC2, Lambda, SNS, SES, Glue, Athena) **Database:** PostgreSQL, MongoDB **Distributed Computing:** Spark **Machine Learning:** Sklearn, SciPy, TensorFlow, **Web Development:** Flask, D3, Hugo, Plotly/Dash **Web Scraping:** Selenium **Data Visualization:** Tableau, PowerBI

PROFESSIONAL EXPERIENCE

UPenn Computational Social Science Lab   **Philadelphia, PA**
Research Data Engineer   **Apr. 2022 – Aug. 2023**
- Standardized AWS cloud, on-premises compute cluster, and research software development best practices.
- Consulted post-doctoral researchers on designing reproducible and scalable data processing and analysis methods, focused on online and TV news media NLP based topic modeling.
- Implemented AWS based research pipeline for >50 TB database for unstructured TV transcript text data.
- Co-authored a research paper with other post docs on simulating human activity on YouTube.
- Managed and delegated work for research assistants across a variety of lab research projects.

Point72 Asset Management   **New York City, NY**
Long/Short Energy & Industrials Portfolio Data Scientist   **Sept. 2019 - Apr. 2022**
- Managed and evaluated large, "non-traditional" datasets, including device geolocation, credit card data, crude oil well production data, web-scraped data, and others, as sources for leading indicators of company key performance metrics.
- Informed investment decisions using exploratory analysis, anomaly detection, regression models, and data visualizations alongside team of eight equity analysts and portfolio manager.
- Designed event driven ETL job orchestration using Apache Airflow, AWS (EC2, Lambda, SNS), and Databricks.

AllianceBernstein Asset Management   **New York City, NY**
Investment Research Associate, Fixed Income   **June 2017 - August 2019**
- Contributed to portfolio construction and risk discussions with written weekly data analysis and commentary and presented at monthly research review meetings.
- Developed and maintained novel market liquidity signals from streaming bid/offer orderbook database, utilized by traders to assess real time supply and demand levels in subsets of the market.

PUBLICATIONS

- H. Hosseinmardi, **M. Rivera-Lanas**, A. Ghasemian, M. Ribeiro, R. West, D. J. Watts "Causally Estimating the Effect of YouTube's Recommender System Using Counterfactual Bots", *Submitted.* **See paper [here](#)**

SELECTED PROJECTS

Volunteer Developer — Statisticians without Borders   **Nov. 2020 - March 2021**
- Team developed AWS based web dashboard used by municipal health organizations in India evaluating COVID-19 related epidemiological model results. Work on behalf of non-profit, India Excellence Forum (IEF). **See citation in resulting project article [here](#).**