Rohan Weyerbrock Arora

rohanarora@gwu.edu | linkedin.com/in/rohanwarora | rohanwarora.github.io/portfolio

Education

The George Washington University, Honors Program – BS in Data Science & Economics, Minor in History
Summa Cum Laude, Economics Special Honors, Dean's List (Every Semester)
Relevant Coursework: Data Mining, Bayesian Statistics, Data Warehousing, Econometrics, Financial Economics, Multivariable Calculus, Applied Time Series Analysis, Linear Algebra, Data Visualization, Business Statistics

Skills

Languages: Python, R, SQL, Excel, Tableau, Looker, Power BI, dbt, STATA, MongoDB
Libraries/Packages: NumPy, Pandas, Matplotlib, Seaborn, Plotly, Statsmodel, LangChain, Sckit-learn, dplyr, ggplot2, tidyr, shiny, RJAGS, MCMC, astsa, Jira, Kanban

Experience

Data Analytics Engineer Intern, CarGurus – Boston, MA   June 2024 – August 2024
- Built and maintained 20+ data models using dbt, Snowflake, and Looker. Optimized SQL queries to handle datasets exceeding 1 million rows, ensuring high performance and reliability in data retrieval and predictive analytics.
- Managed data warehouse access for over 500 users in Snowflake, utilizing Metaplane to improve big data resource allocation, reducing costs by 10% through efficient usage.
- Overhauled 10+ explores in Looker for self-service data models, implementing best practices for data modeling and documentation, and ensuring accuracy with Spectacles testing.
- Designed end-to-end big data pipelines in Airflow, supporting decision-making and strategic KPI tracking.

Data Scientist Year-Round Intern, Federal Reserve Board – Washington, DC   May 2023 – May 2024
- Created and updated Tableau and Power BI dashboards that visualized $10M+ in expenses, enabling data-driven decisions to align with business objectives and uncovering $10K+ in cost savings.
- Developed a SARIMA time series model in Python for a $500K+ annual travel budget, using Monte Carlo simulations to boost performance and achieve a 16% increase in prediction accuracy, enhancing decision-making.
- Engineered a job recommendation algorithm by scraping 150+ internal job descriptions, applying NLP techniques, and using TF-IDF with Cosine Similarity to deliver personalized recommendations.

Honors Research Assistant, George Washington University – Washington, DC   Oct 2022 – May 2023
- Constructed a panel dataset for 3,143 counties over 22 years, using R and STATA for quasi-experimental and regression analyses to study unionization's effect on U.S. voting access.
- Analyzed a 53-year dataset of 100+ countries with Python to study the U.S. impact on global economic liberalization.
- Conducted an in-depth review and synthesis of relevant literature, leveraging research skills to validate analysis.

Success Coach, AmeriCorps – St. Louis, MO   August 2020 – July 2021
- Mentored 13 sixth graders 1-on-1, supporting literacy, mathematics goals, and social development through regular sessions, contributing to a 100% graduation rate, with one receiving the Beyond School Student of the Year award.
- Liaised with caretakers and teachers to create personalized learning materials and strategies.
- Recruited and trained 40 volunteers from local colleges to allow tutoring of more students needing academic support.

Projects

Personalized Book Recommendation System with RAG   **GitHub**
- Engineered a personalized book recommendation system in Python using several LLMS, employing LangChain.
- Utilized zero-shot classification and sentiment analysis with Hugging Face to enhance recommendation accuracy.

From Tweets to FOREX: How Trump Shaped International Financial Markets   **GitHub**
- Analyzed 56,000+ tweets in Python to study the effect of Trump on short-term foreign exchange market volatility.
- Applied LLM tweet classification and regressions to analyze correlations with tweet topics and currency fluctuations.