

Samuel Harris

Washington, DC | samharris@gwu.edu | 412-527-9194

<https://www.linkedin.com/in/samuel-harris-sh/> | <https://github.com/harrissamuel/projects>

EDUCATION

The George Washington University BS in Data Science and Mathematics

Washington, DC
May 2025

- GPA: 3.94
- Relevant Coursework: Data Mining, Data Visualization, Data Warehousing, Multivariable Calculus, Linear Algebra, Mathematical Statistics, Engineering Computations, Numerical Analysis and Modern Physics

TECHNICAL SKILLS

Languages: Python | R/RStudio | SQL and NoSQL

Packages/Libraries: Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, Selenium, ggplot2, tidyr, dplyr, plotly

Skills: Supervised Learning (Regression Models/Random Forests), Unsupervised Learning (Clustering), Data Mining, Data Visualization, Data Cleaning, Data Wrangling, Web Scraping

EXPERIENCE

GW Data Science for Sustainable Development Hub Data Scientist

Washington DC
January 2023 - Present

- Research team member of GW's Data Science for Sustainable Development (DSSD) chapter. DSSD provides pro bono data science services to mission-driven organizations operating in the development and humanitarian sphere
- Developed data visualizations in Python (utilizing packages like Pandas, NumPy, and Matplotlib) for energy and water trends in hundreds of buildings in Washington DC for an ongoing project to develop localized ecolabels for DC

GW Hatchet Student Newspaper Metro Reporter

Washington DC
October 2021 - December 2021

- Interviewed GW professor Frank Sesno about 30 years of experience with GW Alumnus Colin Powell
- Reported on construction noise complaints from 10+ GW students at District and Lafayette Halls

TECHNICAL PROJECTS

Underrated Fantasy Football Players 2022 Creator

Washington DC
June 2023 - August 2023

- Deployed linear regression with K-fold cross-validation to predict fantasy football points per game for wide receivers and running backs for a personal supervised learning project in Python using Pandas, Seaborn, and Scikit-Learn
- Determined the difference between predicted point values and real point values to rank fantasy players from overrated to underrated, thereby enhancing insights into making optimal player selections

Baseball Over the Last 40 Years Author

Washington DC
March 2023 - May 2023

- Analyzed trends among baseball statistics like strikeouts and home runs over the last 40 years applying regression in R for a Data Visualization project. Project included interactive and animated visualizations using ggplot2 and plotly

Costs of Climate Change Author

Washington DC
November 2021 - December 2021

- Explored the relationship between climate change and disaster costs in different regions of the US for a Data Science project in R. Project included heat maps of all 3000+ counties in the US and statistical plots using ggplot2

AWARDS

- Outstanding Academic Achievement Award - awarded to the top 2% of students by GPA within each school with over 60 credits earned