

Samuel Harris

Pittsburgh, PA | harrissamuel130@gmail.com | 412-527-9194
<https://www.linkedin.com/in/samuel-harris-sh/> | <https://github.com/harrissamuel/projects>

EDUCATION

The George Washington University

Washington, DC

B.S. in Data Science and Mathematics, Minor in Statistics

May 2025

- GPA: 3.95 / 4.0 | *Dean's List (All Semesters)* | *Graduated Summa Cum Laude*
- Relevant Coursework: Data Mining, Data Visualization, Data Warehousing, Multivariable Calculus, Linear Algebra, Mathematical Statistics, Real Analysis, Engineering Computations, Bayesian Statistics, and Modern Physics

TECHNICAL SKILLS

Languages: Python | R/RStudio | SQL and NoSQL (MongoDB and Neo4j) | Power BI | SAS

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

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

EXPERIENCE

GW Data Science for Sustainable Development Hub

Washington, DC

Head of Data Science, Research

January 2023 - May 2025

- Led research team of GW's Data Science for Sustainable Development (DSSD) chapter. DSSD is a nonprofit providing pro bono data science services to mission-driven organizations in the development and humanitarian sphere
- Directed development of a Python dashboard using Pandas, NumPy, and Plotly to analyze energy and water trends across 100+ university buildings in Washington, DC, supporting a sustainability review of seven local universities

Amtrak

Pittsburgh, PA

Voluntary Safety Reporting Intern

May 2024 - August 2024

- Analyzed safety incident data from frontline rail workers leveraging Power BI to create visualizations for the monthly report and briefing to the Chief Safety Officer. Produced safety metrics by department and type of incident
- Built two data pipelines to isolate safety incidents of interest for legal and marketing purposes. Generated an additional dataset to formulate an infographic design promoting the diversity of Amtrak's safety reporting culture

TECHNICAL PROJECTS

World Disasters in Context

Washington, DC

Author

January 2024 - May 2024

- Placed over 15,000 world disaster events since 2000 in the context of development indicators by country for capstone project in Python. Developed a risk score methodology to map the most severe disaster types by country

Underrated Fantasy Football Players

Washington, DC

Creator

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

Washington, DC

Creator

March 2023 - May 2023

- Examined 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 Plotly and ggplot2

HONORS AND AWARDS

- Outstanding Academic Achievement Award (2023)— Awarded to top 2% of students by GPA with 60+ credits
- Member, Pi Mu Epsilon— National Mathematics Honor Society