Sophia Hughes

Boston, MA • hughes.so@northeastern.edu • 978-764-5538 • linkedin.com/in/sophiahughes1105/
Availability: July – December 2026

Education

Northeastern University, Boston, MA
Candidate for Bachelor of Science - Major: Data Science and Economics — Expected May 2028
PlusOne Pathway for Master of Science - Master: Economics (Data Science concentration) — Expected May 2029
Honors: GPA 4.0/4.0 • Dean's List (2024 – Present) • Omicron Delta Epsilon
Relevant Coursework: Advanced Programming with Data; Foundations of Data Science; Applied Econometrics; Mathematics and Statistics for Economics; Macroeconomic Theory; Microeconomic Theory; Discrete Structures; Information Visualization; Introduction to Databases

Technical Skills

Programming Languages and Tools: Python, R, Git, SQL, Microsoft Office, Google Suites
Data Science: Pandas, NumPy, Scikit-learn, Matplotlib, JSON
Core Skills: Predictive Modeling, Feature Engineering, Data Cleaning, Statistical Analysis
Languages: English (Native), Spanish (Intermediate/Beginner)

Projects

Who Will Go Pro? College Hockey Predictive Modeling • Lead Data Scientist October 2025 – December 2025
Technologies: Python (Pandas, Matplotlib, NumPy, Scikit-learn)
- Led a four-person team to collect, standardize, and analyze 15,000+ rows of NCAA Division I women's hockey skater and goalie data (2019–2025) using Python and Scikit-learn to engineer performance features and build predictive models of professional advancement.
- Developed and validated classification models achieving up to 94% accuracy and 100% precision for goalies and 90% accuracy with 92% precision for skaters, using model outputs to assess Northeastern Women's Hockey players' likelihood of advancing to the PWHL.

NBA Data Analysis: Growth of the Three-Point Shot • Data Science Analyst February 2025 – April 2025
Technologies: Python (Pandas, Matplotlib)
- Processed and analyzed NBA game-level data (1979–2025) using Python and Pandas to examine long-term trends in three-point shooting and its relationship to team success.
- Built multi-decade visual analyses with Matplotlib (including trend lines, stacked bar charts, and z-score comparisons of championship teams versus league averages), identifying that teams with above-average three-point performance were more likely to achieve postseason success, particularly in the modern NBA era.

The Gini Coefficient and the Environment • Econometric Analyst February 2025 – April 2025
Technologies: STATA
- Developed and evaluated an OLS regression model across 169 countries to analyze how environmental and political factors relate to income inequality.
- Identified GDP (PPP) and renewable energy share as statistically significant predictors of the Gini coefficient

Experience

Teaching Assistant for CS1800 Discrete Structures May 2025 – Present
Khoury College of Computer Sciences • Boston, MA
- Lead weekly office hours, facilitate review sessions for a high-enrollment discrete mathematics course (up to 550 students), strengthening student understanding of logic, induction, and proof techniques.
- Provide individualized academic support through homework assistance and exam grading using Gradescope and Piazza, addressing diverse learning needs.
- Support and guide hundreds of students across three semesters, building consistent regular attendance base for office hours, and earning repeated TA appointments based on subject mastery and instructional reliability.

Activities & Interests

NEU Sports Analytics Club, NEU Chaarg, Basketball (player), NBA, MLB, NFL, Guitar, Cello, Traveling, Sewing