Maria C. Sanchez Linares

Saint Paul, MN 55105 | msanche2@macalester.edu | LinkedIn

EDUCATION

Macalester College, St. Paul, MN | **Bachelor of Arts Expected** May 2025 | **GPA:** 3.50/4.00

Data Science Major, Economics Minor | **Honors:** Kofi Annan Scholarship for Academic Excellence

Relevant Coursework: Statistical Machine Learning, Statistical Modeling, Database Management Systems, Econometrics, Linear Algebra, Multivariate Calculus I & II, Probability, Object Oriented Programming.

RELEVANT EXPERIENCE

Education Analytics Inc. | *Research Analyst Intern* June 2024 – May 2025

- Conducted a 3-month independent research project on teacher turnover rates across 2,276 Wisconsin public schools, utilizing panel data to calculate metrics, perform trend analysis, and generate summary statistics to identify underlying causes and their impact student outcomes.
- Collaborated with a team of analysts on ad-hoc tasks, including data cleaning, quality control and investigation analysis, to develop Student Performance Measures (SPM) ratings for Pasco County using district assessments and state value-added models.
- Developed data visualizations for NewSchools Venture Fund's annual presentation, analyzing the correlation between socio-emotional learning goals and academic outcomes across their school portfolio.
- Collected and processed enrollment and demographic data on Washington, D.C. public schools to develop metrics informing improvement and investment decisions for the Michael & Susan Dell Foundation.

Macalester College | *Data Science Capstone Project* September 2024 — December 2024

- Adopted Latent Dirichlet Allocation (LDA) probabilistic model to conduct a topic modeling analysis.
- Extracted and pre-processed 5,000+ newspaper articles using `pdftools` package and natural language processing (NLP) tools lemmatization and tokenization in R.
- Calculated TF-IDF coefficients to evaluate the importance of different words in the corpus and applied a 0.01 score threshold to reduce noise in the dataset.
- Performed 4 model fitness tests —Griffiths2004, Deveaud2014, CaoJuan2009, Arun2010— to determine the optimal number of topics (k=6) for LDA model.
- Classified articles using LDA-generated Gamma values an conducted a 3-year time-series analysis, tracking topic prominence and identifying key narratives on Venezuela-Colombia media discourse.

PREDOC & Opportunity Insights | *Student Researcher* June 2023 – August 2023

- Conducted a research project on social mobility under the mentorship of Prof. Bruich from the Harvard University Department of Economics, engaging in coursework and weekly advisory meetings.
- Designed and implemented an empirical model to analyze US-wide tax revenue and census tract-level demographic data, leveraging STATA for data wrangling and statistical modelling.
- Co-authored a research paper, interpreting linear regression results and summarizing key findings.
- Chosen as one of the 5 best students out of 50 to present my research findings at a conference hosted by the Federal Reserve Board of Governors.

UVA School of Data Science Data Justice Academy | *Research Fellow* May 2024 – August 2023

- Analyzed different data collection methods and developed a comprehensive literature review on the invisibility and misrepresentation of transgender people in official health statistics.
- Employed R tools to curate a list of all federally supported survey data and designed a set of data visualizations displaying the proportion of surveys collecting specific information across different SOGI areas.
- Contributed research findings to a report being developed by the UVA Center for Global Health Equity.

Starr Hill Pathways | *Student Mentor* June 2023 – July 2023

- Taught Intro to Sports Analytics workshops to 30 middle schoolers at a STEM career exploration program.
- Expanded the program's curriculum to incorporate the achievements and perspectives of female athletes, enriching the students' educational experience through the design of a new workshop.

SKILLS

- **Languages**: Spanish (native speaker), English (fluent).
- **Programming Languages:** R, Python, Java, Stata, SQL, SPSS, LaTeX.
- **Software & Tools:** Microsoft Office Suite, Google Suit, Tableau, Plotly, Git, Snowflake, Preset.