

EDUCATION

University of California, Santa Barbara: 3.99 GPA

Santa Barbara, CA

B.S. Statistics and Data Science | B.A. Theater (Design) | Minor: Language and Speech Technologies *September 2022 - June 2026*

Relevant Coursework: Adv. Computational Linguistics: Text Processing, Adv. Linear Algebra, Data Structures and Algorithms, Predictive Modeling in Linguistics, Probability and Statistics (3 courses), Statistical Machine Learning, Uncertainty Quantification

Selected Presentations:

- Chatteraj, Siddharth, Shahil Patel, and Simon Todd. 2024. Deepening understanding of frequency effects in exemplar models of language change [poster]. Conference. *CAMP[7]: 7th Annual California Meeting on Psycholinguistics*. UC San Diego, CA
 - Authored a simulation-based study comparing decay and overwriting architectures in exemplar models using R. Visualized value and variance trajectories across 10,000+ iterations to reveal parameter-sensitive frequency effects
- Nunez, Jesse, Madhav Viswesvaran, Ishaan Avinash, Calais Waring, Siddharth Chatteraj, Anna Lin. 2023. SlangSwitch [presentation and poster]. *UCSB Data Science Project Showcase*. Conference. UC Santa Barbara, CA.
 - Developed a regional slang-to-formal text translator using Styleformer, natural language processing, and web-scraped slang corpora. Applied Regex-based tagging and built an interactive demo to demonstrate sociolinguistic style transfer

Selected Tools and Languages: Adobe Suite, ArcGIS, Fusion, HTML/JavaScript/CSS, Python, R, SQL, QuickSight, Vectorworks

WORK EXPERIENCE

Amazon

Seattle, WA

Business Intelligence Engineer Intern

June 2025 - September 2025

- Saved 92 hours per week of manual analysis by building an internal AI agent that analyzes return cost per unit via SQL and schema analysis and then forecasts global trends with time-series models and ML, informing policy decisions on return centers and backhaul carriers in 3–5 minutes. From learnings, developed two automated Cline AI-agent-creation workflows, enabling engineers to generate working agents in 5 minutes and cutting development time by 20% by starting at testing and refinement
- Investigated unreconciled EU returns by building SQL queries and automating ETL pipelines, analyzing backhaul carrier data to identify fulfillment inefficiencies and automate reporting; Designed QuickSight dashboards tracking weekly, monthly, and quarterly transportation return metrics to support business planning and operational insights

Savie Health

Remote/Lompoc, CA

Revenue Manager

February 2025 - September 2025

- Developed forecasting algorithms to predict monthly Medi-Cal revenue with 95% accuracy and ensure timely payments
- Built Python SARIMA time series models to forecast patient attendance based on operating changes and revenue from Medi-Cal

Data Analyst

August 2024 - February 2025

- Designed and automated monthly reports to help the organization earn over \$65,000 in grants and make hiring decisions

Data Analysis/Marketing Intern

July 2022 - August 2024

- Created promotional graphics and e-blasts in English and Spanish to advertise the startup nonprofit and helped secure over \$100,000 in funding and 1500 monthly social media engagements through email marketing and data reports for grant requests

San Francisco District Attorney's Office

San Francisco, CA

San Francisco Sentencing Community Initiatives Data Fellow

June 2023 - September 2023

- Created a regeneratable temporal and geospatial data memo for distribution to community members across the city using HTML, JavaScript, and Python Plotly to visualize crime trends in San Francisco supervisor and police districts
- Analyzed several databases of over 200,000 county crime records and visualized crime location trends using Geopandas, R, Python Plotly, and ArcGIS for grant requests. From learnings, wrote a 21-page guide on how to analyze and map geospatial data

UCSB Department of Computer Science

Santa Barbara, CA

Lead Undergraduate Learning Assistant (ULA)

April 2023 - Present

- Served as Lead ULA for UCSB's introductory data science course; mentored teaching staff, designed LLM-integrated assessments, and supported instruction on topics including hypothesis testing, bootstrapping, regression, KNN, and classification
- Tutored weekly sections of 30+ students and evaluated 100–200 assignments per week, offering individualized support through office hours and grading presentations, code, and statistical analyses.

Computational Psycholinguistics of Listening and Speaking Lab

Santa Barbara, CA

Undergraduate Researcher

January 2023 - Present

- Built models of simple dynamical systems of language change in Python by deriving and implementing formulas for statistical model analysis. Optimized parallel process and file writing runtimes from 120 minutes to 30 seconds
- Developed an SQL tool to query word frequency and score data from 1000 relational databases and determined best samples by calculating Jensen-Shannon divergence 1000 times for 30 classifications of data using kernel density curves

Daily Nexus

Santa Barbara, CA

Data Editor and Games Section Founder

September 2022 - Present

- Oversaw the production of over 65 data-driven articles and mentored more than 200 writers on data storytelling, sourcing, and visualization, leading to multiple statewide awards: 1st Place for Best Multimedia Package (2025), 2nd Place for Best Interactive Graphic (2024), and 1st Place for Best Infographic (2023) from the California College Media Association
- Founded the Games section and launched its website, making the Nexus one of the first college newspapers with a dedicated games page. Designed interactive crosswords with JavaScript and taught workshops on data visualization, NLP, and web scraping
- Led management of The Labyrinth, a public dashboard visualizing 16 years of UCSB course data—over 40,000 classes—using D3.js. The tool is now accessible to UCSB's 30,000 students to explore grade distributions by instructor, quarter, and department