SAMREENA SIDDIQUI

samreena@berkeley.edu || LinkedIn || Portfolio || (408) 835 - 3864

EDUCATION

University of California, Berkeley **Graduation: May 2026**
B.A. *Data Science*, B.A. *Cognitive Science* | GPA 3.96

Relevant Coursework
Artificial Intelligence, Data Structures, Foundations of Data Science, Linear Algebra and Differential Equations, Principles & Techniques of Data Science, Probability of Data Science, The Structure and Interpretation of Computer Programs

Foothill High School, Pleasanton CA **Graduation: June 2022**
GPA: 4.0 (unweighted)

EXPERIENCE

Jacob's Ladder Therapies **San Jose, CA**
Data Science Intern *June 2024 - August 2024*

- Developed and implemented advanced data visualization and analytical models on client company review datasets, resulting in a 20% improvement in predictive accuracy for service performance and operational efficiency.
- Optimized survey design using statistical techniques [Correlation, Factor Analysis, Frequency Distribution], increasing response rate by 15%, and applied data analysis to generate actionable insights that informed business strategies for over 100 clients.
- Powered data mining and analysis on 200+ legacy surveys, uncovering key trends and patterns that drove data-driven decision-making in marketing and project management across the organization.

The Gomez Lab || Computational and Molecular Biology **Berkeley, CA**
Undergraduate Researcher *May 2023 - April 2024*

- Utilized R for RNA sequencing, converting biological files (FASTA, FASTQ) into amino acid sequencing data.
- Shadowed and was mentored by in lab computational biologist.
- Designed experiments to assess the strength of neuronal connections, conducting extensive genotyping and PCR reactions to analyze results.

Neurotechnology at Berkeley || Software **Berkeley, CA**
Dementia Detection *August 2022 - Present*

- Actively developing machine learning algorithms [Decision-trees, Logistic regression, PCA] aimed at predicting the presence of dementia utilizing voice biomarkers.
- Engaging in the process of cleaning voice recording data to effectively compress and diminish noise, facilitating subsequent analysis and further experimentation.
- Conducting comprehensive literature reviews, identifying key variables utilized in published papers to inform the foundation of project.

UC Berkeley College of Computing, Data Science, and Society **Berkeley, CA**
Data Science Academic Tutor *August 2024 - Present*

- Collaborate with data science instructor and team of 20 other tutors to support a class of 1600 students, providing one-on-one tutoring to 10 students per section on data science fundametnals, including linear regression, distributions (hyper-geometric, binomial, normal) and classifiers.
- Conduct lab sections and grade students based on accuracy, ensuring comprehensive understanding of course material.
- Facilitate weekly office hours to address individual student questions using personalized worksheets and mini-lectures.

LEADERSHIP AND CLUB WORK

Big Data @ Berkeley **Berkeley, CA**
Data Science Content Creator *January 2024 - Present*

- Developed a comprehensive data science course, slated for release on EdX by Decemeber 2024, covering fundamental principles such as logistic regression, clustering, exploratory data analysis (EDA), and other essential topics in data science.
- Utilized DeepNote to create coding assignments, incorporating HTML and Python along with associated packages such as Pandas, NumPy, SciKit, and more.
- Produced and published lecture videos for global dissemination, enabling access to educational content for students worldwide.

Brain Exercise Initiative **Berkeley, CA**
President *August 2022 - Present*

- Expanded club membership from 25 to 60 active participants within a single semester, fostering a more vibrant and engaged community.
- Initiated and broadened connections with Senior Center homes, increasing partnerships from 2 to 5 establishments.
- Conducted weekly team meetings, facilitating discussion, creating weekly summaries, and coordinating tasks for other executive members.
- Volunteering with seniors with dementia to enhance cognitive abilities through reading, writing, and conversation.

SKILLS AND INTERESTS

Coding Languages: Python: NumPy, Pandas, Matplotlib, Seaborn, SciKit,, R, Java, SQL, HTML & CSS

Skills: Data Visualization, Hypothesis Testing, Data Wrangling and Clearning, Data Mining, Excel, PowerPoint, Problem Solving,

Interests: Machine Learning, Big Data Analytics, AI, Deep Learning, Health Informatics, Behavioral Data Analytics; Avid guitar and tennis player