

# Statement of Purpose

“This data is key in detecting down-syndrome early enough to terminate pregnancy,” my manager explained on the first day at my recent internship. As someone who was misdiagnosed with down-syndrome, issues like these drove my desire to improve technological shortcomings through data. I chose data science because my strengths lie in the harmonious intersection of computer science, creativity, and their applications to real world problems like preventing a misdiagnosis. Ultimately, Berkeley’s MEng EECS Program is my top choice because the curriculum, research opportunities, and emphasis on industry, innovation, and leadership align with my goal to become an industry leader in applying data science towards improving healthcare, biotechnology, and other domains.

**Engineering Efforts.** Beyond my coursework, I sought to apply my knowledge to real world problems. During my first internship at Chan Zuckerberg Biohub, I developed database querying and visualization tools to help researchers identify significant genes which contribute to virus (Ebola, COVID variants, Dengue, etc) transmission. Consequently, my contributions helped researchers better understand how dangerous viruses infect humans. Through this I learned about contributing directly to ongoing research with individuals from innovative institutions and how I can positively impact society through data science. Understanding viral transmission was invaluable, and using data to investigate important problems like these inspired me to pursue a masters.

Building upon this, at my most recent internship at Bio-Rad, I built systems to protect the integrity of a limited amount of maternal data, ensuring that False Positives/Negatives for down-syndrome and other conditions were properly recorded. Maintaining these misclassifications was vital as each data point was also an affected family. A single error equated to potentially terminating a healthy pregnancy and vice versa. I live knowing I could have been a terminated pregnancy and therefore feel strongly for data driven technology like Bio-Rad’s. Having completed my internship, my influence will be folds greater when my work eventually culminates in a working aneuploidy test for future mothers and families. Reflecting on this perspective, I am committed to applying data science to problems of great downstream magnitude like my misdiagnosis, and I plan to pursue graduate school to expand my skills for this purpose.

**Leadership.** Collectively, these industry experiences have helped me share the value of data science. As an instructional assistant serving over 1500 students, I’ve empowered students from creating exam and homework questions, holding office hours, and creating helpful guides on coding. Consequently, I’ve provided a sandbox for students to make mistakes and witnessed their growth over time. Relating to my misdiagnosis, I wanted to foster students to accurately represent and utilize data to ensure data driven tools are reliable. Another aspect that excites me is the Fung Institute’s attention to detail on leadership and career development; specifically the intertwining of management courses in the curriculum and domain specific capstone project will help me develop my engineering experiences with others. Overall, I would like to learn how to become a leader in the field through embracing a challenging, diverse curriculum, and I

believe Fung's close knit culture and leadership emphasis fits my academic background and goals perfectly.

**Future Plans & Fit.** To prepare for a graduate curriculum, I cross enrolled in graduate classes such as Deep Learning, Recommender Systems, and Computer Vision. At Berkeley, I hope to take a deeper dive in ML methodology to further develop my knowledge towards healthcare; however, it is only one aspect as to why Berkeley's environment provides me a foundation to achieve my career goals. To preface, I enrolled in CS188 during a summer internship to apply its topics to my project and familiarize myself with Berkeley's academic environment. Unexpectedly, I realized the faculty placed an importance on ethical ML/AI - a culture of integrity integral to why I pursued data science. My current research experience also aligns with this culture; to prepare for potential graduate research, I designed a regression model and robust accessibility metric, advised by Professor Zaslavsky to fairly score regions on how equitable public transit infrastructure is. Additionally, this fall I am involved in Professor Hu's group centered around LLM reasoning to continue my initiative to make ML/AI more reliable.

Similarly, Berkeley's focus on innovation perfectly aligns with my engineering experience in data science for healthcare/biotechnology. My future research interests are increasing accessibility, reliability, and equity in healthcare ML/AI. Berkeley captivates me because there are many labs working on problems that resonate with my misdiagnosis. For example, I would like to engage in research similar to Dr. Petersen's "Adaptive Designs for HIV Control" (CTML) and especially Professor Irene Chen's initiative on making healthcare ML unbiased and equitable (Chen Lab). In Chen's PhD thesis she approached questions such as "how can we mitigate algorithmic bias in healthcare?" and "how can we train models to account for these biases?" Having access to these opportunities is important because they resonate with my goal of making ML more reliable and equitable for societal impact. Knowing faculty like Professor Chen are investigating problems similar to my previous experiences is a strong reason alone as to why Berkeley excites me intellectually. This institution provides the curriculum, resources via the Spring Capstone, and unparalleled faculty for multidisciplinary research problems like these, and will allow me to learn from minds like Chen and achieve as a graduate student and future data scientist.

Ultimately, my goals lie beyond a masters in the field. In the future I envision myself leading an R&D team to solve problems like those at Bio-Rad and Biohub through data. Graduate school at Berkeley would provide an environment where I can continue to mature as I take courses at the cusp of innovation and provide me access to research opportunities that resonate with similar problems to my misdiagnosis. Reflecting on my contributions to virus transmission research, prenatal down-syndrome tests, and my future career goals, I am thrilled to pursue the MEng EECS Program to continue growing as a student, data scientist, and leader as I transition to industry.