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 its emphasis on innovation, leadership preparation, and industry applications through the capstone 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 professional master's program.

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 the program's focus on professional development through challenging computer science and leadership coursework will prepare me 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 will help me develop as a leader in an engineering environment. Overall, I highly value mentorship, and believe Fung's unique, close

knit culture and leadership emphasis complement my academic background and goals of becoming a leader in data science.

Future Plans. With respect to my background, I view the MEng program as not only advanced coursework, but an opportunity to enrich the learning community at Berkeley. Previously, I enrolled in CS188 during a summer internship to apply its topics to my project and familiarize myself with Berkeley's academic environment. Through this, I discovered the faculty's emphasis on ethical ML/AI - a culture of integrity embedded in why I pursued data science. My current research also aligns with this value. Specifically, I designed a regression model and accessibility metric, advised by Professor Zaslavsky to fairly evaluate public transit equity across regions. Additionally, this fall I am involved in Professor Hu's group centered around LLM reasoning to continue my initiative on ML/AI reliability. Collectively, these involvements helped me envision my future contributions at Berkeley and beyond. To prepare for a graduate curriculum, I cross enrolled in graduate classes such as Deep Learning, Recommender Systems, and Computer Vision. Likewise, at Berkeley I plan to dive further in ML methodology and apply it towards healthcare.

Specifically, the program's core around innovation closely aligns with my engineering experience in data science for healthcare and equity. This captivates me because there are many groups working on problems that resonate with my misdiagnosis. For example, I plan to use the capstone opportunity to apply my technical skills to problems similar to Professor Irene Chen's initiative on making healthcare ML unbiased and equitable (Chen Lab). In Chen's PhD thesis, she addressed questions such as "How can we mitigate algorithmic bias in healthcare?" and "How can we train models to account for these biases?" Consequently, the capstone project will provide me a platform to approach problems like those at Chen Lab and Bio-Rad to improve healthcare ML systems while receiving valuable mentorship. Additionally, although an industry focused program, having access to a network of opportunities like Chen Lab is important because it will open more doors for me to contribute to reliable and equitable ML. Moreover, knowing faculty like Professor Chen are investigating problems similar to my previous experiences is a strong reason why Berkeley excites me intellectually. Overall, the Fung Institute provides the curriculum, resources via the capstone, and will allow me to learn from unparalleled minds like Chen to apply my skills 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. The Fung Institute would provide an environment where I can continue to mature as a leader, take courses at the cusp of innovation, and contribute to solving problems similar 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.