Statement of Purpose - UC San Diego CSE
Sean Kross - smk240@gmail.com

My aspiration is to spend the next several years pursuing a PhD in computer science in the field of human-computer interaction. I'm interested in online education systems, cybersecurity, and software engineering.

I've had the opportunity to participate in the development of several online courses and I've had the privilege of meeting students who were able to use their experience from a course in order to improve their lives and careers. My motivation stems from the exhilarating feeling of bringing knowledge and opportunities to these students. In addition, it's important to me to see that cybersecurity tools are accessible, easy to use, and distributed equitably as online interactions become ubiquitous in everyone's lives.

For the past two years I've worked with Professors Jeff Leek, Roger Peng, and Brian Caffo in the department of biostatistics in the Johns Hopkins Bloomberg School of Public Health. I helped them develop the Data Science Specialization on Coursera.org, a series of nine month-long online courses which teach programming, statistics, and communication skills. Each course runs every month and since the courses were launched over four million students have enrolled. I coauthored a book called Developing Data Products for a course with the same name in the specialization. This book details how to implement web applications and data visualisations as interfaces for data analyses that students produce in earlier courses.

We created a version of the specialization for managers of data scientists, and for this specialization I wrote a screenplay to be performed by data scientists working at Zillow.com. I created an online course in partnership with DataCamp.com where students watched the videos of Zillow data scientists giving them feedback, and then students could choose one of several ways to respond to the data scienctist. Students navigate a path of videos and quizzes where their decisions send them down different forks in the path. Through this system students were able to simulate the experience of managing data scientists.

To complement these online courses I coauthored a software package for R called Swirl which transforms the R programming console into an interactive learning environment. Swirl prints messages to the student's R console which instructs them in lessons about programming and data analysis. Swirl can assess correctness for code entered into the console, R scripts, and multiple choice questions. Swirl can also display visualizations and readings to the student. Swirl has several built-in courses though anyone can write a Swirl course using plain text and R code. I also wrote an R package called Swirlify which provides tools for building Swirl courses. Swirl has been downloaded over one million times.

Recently I've been working with Professor Michelle Mazurek at the University of Maryland in order to research the relationship between negative cybersecurity outcomes for individuals, their computer security advice sources, and their socioeconomic status. Our analysis of census representative surveys has shown that certain advice sources put individuals more at risk of having a negative experience online regardless of their socioeconomic status. I hope to work with her group this year in order to design an experiment where we can test the observations we've made about online security behavior.

If admitted to your graduate program I would be most interested in working with Professor Philip Guo because of his work in developing online learning environments and data science tools. I'm also interested in the DNA sequence assembly work of Professor Pavel Pevzner. I am open to pursuing other areas of research as well, however I am sure that the PhD program at the University of California San Diego would be the perfect environment for me to explore my curiosities.