**CMSI 4072 Project Proposal**

Description:

My Sample Size is a website that aids researchers with statistical calculations in order to design statistically reproducible experiments. The following tools are present on the website: Sample size, statistical power, power plots, simulations, random number generator, and data analysis. Additionally, the main feature is the design guide – a step by step questionnaire that walks users through the experiment they want to design. The end of the design guide provides the user with a PDF report compiling together all the data the user provided, in addition to the statistical information generated by the website. The website is designed for pre-clinical researchers looking to perform medical experiments such as the effects of certain drugs, for this reason the design guide features a fill-in walkthrough and tooltips related to this topic. The goal is to make the website as easy to use as possible to someone who is not a statistical expert. The website runs on Amazon Web Services and relies mainly on Python and JavaScript. Additionally, there are some important dependencies included in the project such as Weasyprint which is used to produce the all-important PDF report at the end of the design guide, which relies on an external dependency called GTK. The project also uses some services for user authentication and previously for charging users for subscriptions: Stripe, and Auth0.

Justification:

This project is a good fit for myself as I am doing a minor in statistics and data science, therefore the website which focuses on statistics is a good crossover between my studies at LMU. Additionally, this project provides me with valuable experience for going into the tech industry. Previous to taking on this project I had no experience with website development, as I opted for the mobile development practicum instead. Furthermore, this project also provides me with experience working with an existing project and having to understand how everything works. Normally, in LMU’s coding assignments we are given direct instruction on what parts of the files to interact with rather than having to understand how it is that everything works together. Furthermore, the website is necessary due to its nature in aiding individuals in an important industry. 50% of the annual US preclinical research spend is wasted on non-reproducible experiments ($28.2 Billion), this is due to improper experiment design from a statistical perspective. Researchers should not be expected to be masters of statistics as well as knowledgeable in the field of medicine – it is for this reason that My Sample Size is necessary; in order to alleviate the difficulty placed upon the researchers, and in turn lead to less wasted time and money on experiments whose data is not reliable.