Surprises:
For this first technical milestone, the goal was to generate a plot of normalized fixation probability versus population size under a specific set of population assumptions and hopefully replicate the sign inversion effect obtained in [1]. To begin, I implemented a population model as outlined in the paper. Then, for a range of thirty-two distinct population sizes, I estimated the fixation probability over 1000 trials; however, the resulting plot did not manifest the desired sign inversion effect. For a long time I suspected that my model implementation was faulty, but in the end, it turned out that the 1000 trials was insufficient. Variance effects were overpowering the underlying signal. Thus, after I re-ran the simulations for each population size over $10^6$ trials instead of only 1000, I was able to achieve a matching plot.

Still, the increased iterations came with its challenges. The $10^3$ factor increase implied that implementation efficiency was now non-negligible. My implementation for this milestone followed a logically clear approach since correctness was the first priority and forwent any algorithmic optimizations. Consequently, when I ran the simulations on an Andrew Unix machine directly, I was getting `Write Failed: broken pipe` errors due to timeout. Ultimately, the workaround that I employed was to partition the work, e.g. break the $10^6$ trials into ten $10^5$ blocks, but even then, the total running time was significant.

What Has Been Accomplished So Far:
As prefaced in the Surprises section, I successfully implemented a Wright-Fisher model in Python to simulate the population for a 100-fold mutator with 99 fitness-affecting loci and 1 mutation rate modifier where the former can generate beneficial and deleterious mutations. Running the model over various population sizes yielded the sign inversion effect studied in [1]. In addition, I read “Chapter 1: Genetic Variation” and “Chapter 2: Genetic Drift” in Population Genetics: A Concise Guide by John H. Gillespie to bolster my background in population genetics.

Meeting Your Milestone:
I believe that I accomplished the goal set out for my first technical milestone. The predominant motivation for this milestone was to expose me to population modeling and give me a sense of the kind of work being done. Indeed, I feel that after going through this exercise of replicating the work done in [1], I have a much better sense of the things I can expect moving forward. Furthermore, I think that the background reading will prove helpful.

Major Changes:
I do not anticipate any major changes to the goals or implementation of my project since my written proposal.

Revisions to the 15-400 Milestones:
Given that my first technical milestone was completed without any delays, my milestones for 15-400 remain unchanged.

Resources Needed:
In the spirit of what was discussed in the Surprises section, I may want to look into offloading the computation onto AWS for example if the running time becomes a hindrance to the research progress.
References