Data Science Reflection: Modeling and Simulations in Public Health, Winter 2021

COVID-19 had completely changed my college experience in every way possible. Classes had adopted an either virtual or hybrid teaching method and public gatherings were limited. Masking and vaccinations were also enforced for all students and employees. With all this occurring, I thought it would there would be no better time to take a data science class by the name of “Modeling and Simulations in Public Health”. The goal was to model the effects of vaccinations and masking in elementary and middle schools in Virginia. The reason being that masking had been a contentious issue all throughout the pandemic and vaccines for children had just been approved at that time. Our research team had decided that this would be interesting to model and see what the outcome would be.

I had up until this point only ever used R to program. The class was structured to use Python though, which prompted me to learn the language. With this came having to learn basic syntax, importing of packages as well as how to work in Jupyter Notebook. Simple commands such as indexing, deleting rows or columns, and creating functions had to be learned. All of this was done within the context of Python packages such as pandas or numpy. The statistical tests that I knew how to do in R, such as regression, correlation, and p-tests, had to be relearned in Python. I also learned how basic machine learning works and implemented that in several projects, one of which was to determine survivors of the Titanic. Data visualization was the last topic I learned which included using packages such as matplotlib or seaborn.

All these skills eventually culminated in the creation of a SIRED model which predicted the severity of a COVID-19 outbreak in Virginia schools. Using key assumptions such as transmission rate, incubation period and recovery rate, we were able to see how COVID-19 spread when vaccinations were mandated, when masks were mandated, and when both were mandated. What was found was that a vaccination rate of 60% would significantly decrease the risk of an outbreak both in terms of total victims as well as time of outbreak. We also found that surgical masks and KN95 masks were nearly identical in preventing spread in middle and high schools. These findings were then placed on a poster which was presented to peers with similar projects surrounding COVID-19. These Python skills were later improved through a bootcamp taken outside of Washington and Lee University.