Week 1 Writing

The prompt I choose is about how my understanding of the work of statisticians changes after taking statistics courses in university.

My perception towards data was narrow before. I thought the data can be a list of numbers or a form with indices and the work of statisticians can be as simple as calculating the mean of the numbers and plotting the trend. Now I know that the data can be much more complicated and how the data is represented varies a lot. The data we obtain from the original source like real world can be messy and not structured well. Statisticians always need to preprocess the data first before going deep with it. For example, some data can lie in a lower-dimensional manifold in a high-dimensional space. Then applying a dimension reduction is necessary and would make the following work more efficient.

I thought the work of statisticians more like using tools to show the data nicely instead of a rigorous process in the field of science before. Now I know that the work of statisticians can be rigorous in some way after I learn the probability theory and the model building process. Statisticians have been doing quite a lot of research in random variables and came up with many wonderful theorems like central limit theorem. All such theoretical foundations provide enough support to the work of statisticians and make it more convincible.

In general, studying statistics in university makes me realize the remarkable power of data with the help of statistical models. We can make strong inferences based on the evidence observed using probabilistic models like Bayesian Network. We can learn mappings from inputs to targets to make predictions through different regressions or multilayer neural network. We can do classifications using random forest or create unsupervised classifications for the data with similarities using k-means or generative models. Those wonderful materials make me understand that the work of statisticians is significant and powerful at any time.