Antonio da Silva

STAT 331

Final Portfolio – Essays

**The power of efficiency**

The term “efficiency” means, in relation to data science, minimizing the use of functions and inputs while also being robust, so the code can run faster and reduce the likelihood of errors. I showed efficiencies in lots of places during my time in Stat 331. First, I was efficient in Lab 3 (study subjects) by only using the function (filter) once, and filtered through many columns with their specifications by just separating this using a coma. Next, another efficiency I utilized was in challenge 3 (1) by using the “:” in a function to select more than 1 column.

The definition of efficiency is the fastest and easiest way of doing something. Efficiency in data science is demonstrated by minimizing lines of code to portray the same result. This is important because efficient code will save the user time and memory when the code is run which in a industry format will save the company lots of time and money. Increasing margins in a company is a very important KPI metric that efficiency is vital for the metric. I had an “a-ha” moment with this during my challenge 9 when I made a function to manipulate several columns at once while only having to write the function out once. Initially, I was manipulating each column separately with its own function, but after some research in the textbook and slides, I realized that I could use “:” to condense the columns using across function and make a function to adjust all the columns at once. I learned through this process that there is no need to ever repeat code more than once whether it is for data wrangling or other coding types. This has helped me realize that I can implement this efficient mindset into more than just coding, but even making breakfast. The goal being to make a process the fastest, cheapest (memory cost, ingredient cost, etc), and highest quality.