Homework 1: Titanic

I began by looking over the Titanic “train.csv” data set and determining which variables (columns) would be the most likely to factor into any given passenger’s survival. From the given variables in the data set, I narrowed it down to three:

1. Class number (1-3)
2. Age (0.42-80)
3. Gender (Male or Female)

These variables were selected using the historical context the ship was in during the disaster.

Looking at the Class variable, first class passengers were likely to survive as they were closer to the top of the ship and likely to be the first to receive access to lifeboats. The second class passengers were the next likely to receive access to life boats, but were more at risk to environmental danger as they were lower in the ship as well as having to wait for the first class passengers to be saved so their likelihood of survival is not high but certainly not as low as it could be. The third-class passengers had the lowest survival likelihood as they were on the lowest passenger decks of the ship and the possibility of them being killed in the initial crash was significantly higher than the other passengers. Also because of the fact that they were so deep within the ship, it was very unlikely they would receive access to a lifeboat and survival became more based on luck than anything else. When Looking at age, because of the standing social maritime rule “Women and children first”, children were likely to survive over adults. I defined children as any age under 15. It was also likely that elderly women survived as older women were likely in first class and older men died in their place. When looking at gender, the aforementioned maritime rule also comes into factor in immediately deciding the survival of men in a disaster situation. This rule made gender the most statistically significant factor for survival as most of the men aboard the ship died because of this rule.

My approach for determining survival for the data set was to assign every passenger a Survival Score of 100 and lowering it according to the unique values they had within each variable. Since no passenger was guaranteed survival, the program only lowered the score from the ceiling of 100. My findings from this experiment is that approximately 38% of passengers survived and that numpy does not pass individual elements to check equality when presented with the = = operation. This crippling logical problem didn’t permit the program to run appropriately, but the logic is present in the code and will run appropriately once the logic is corrected within the syntax.