James Patalan

DCS530 Data Exploration and Analysis

Professor Matthew Metzger

8/15/23

Final Project Reflection: Diabetes

The outcome of my EDA was somewhat conclusive, but the results could have been a lot better. I only was able to explain 27.7% of relationships between my chosen variables and whether a person had diabetes or not. I definitely think that I missed something, and perhaps should have taken a closer look at some of the variables I did not choose to test, such as Skin Thickness. I assumed that the greatest indicators of whether a person would have diabetes would a Plasma Glucose aka blood sugar, BMI and Blood Pressure. However, there seems to be another indicator that was not among the chosen variables. One challenge that I faced was attempting to use the **imgkit** python library to generate a .png file of the analysis that I conducted. I was unable to get it to work and unfortunately had to resort to copying and pasting the generated table directly into my presentation. Given more time, I would have liked to make a more visually appealing chart to look at. I also struggled a little with interpreting the results of the regression analysis and had to research in order to understand what was going on.