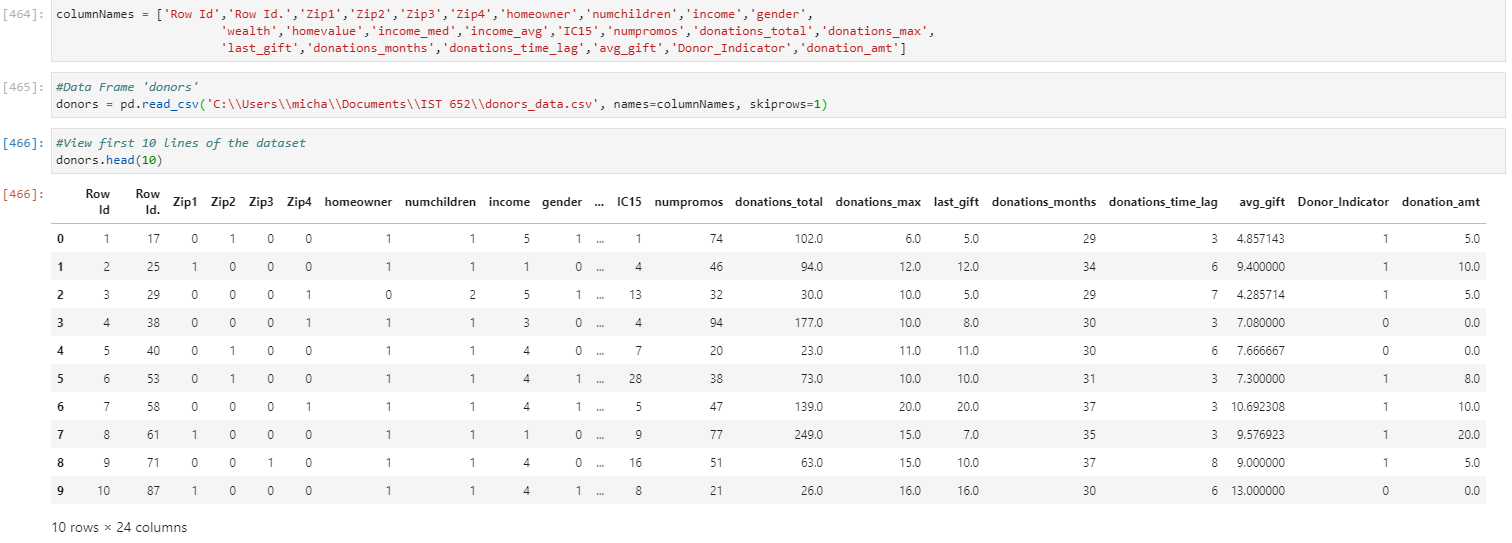
**Homework 1**

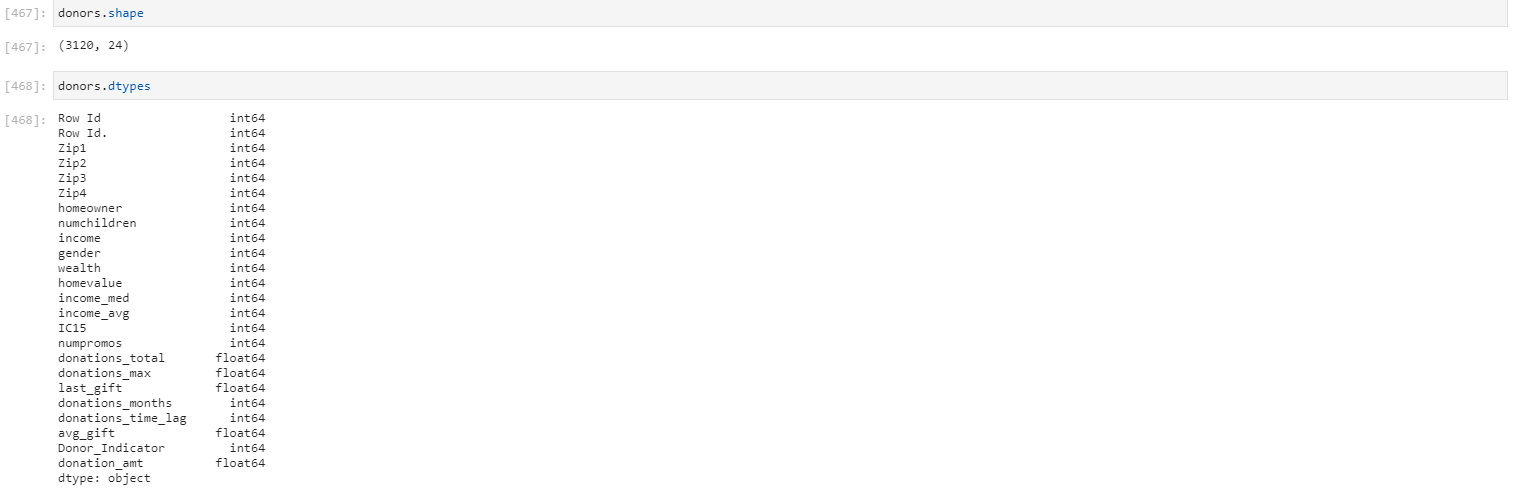
**IST 652**

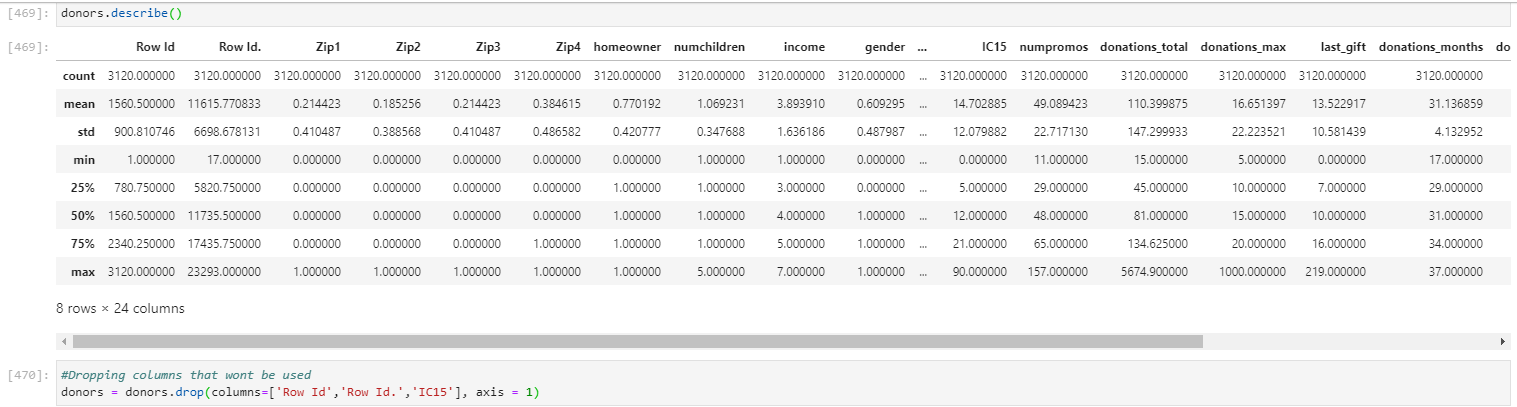
**David Doman**

**Source, Data Cleaning & Exploration:**

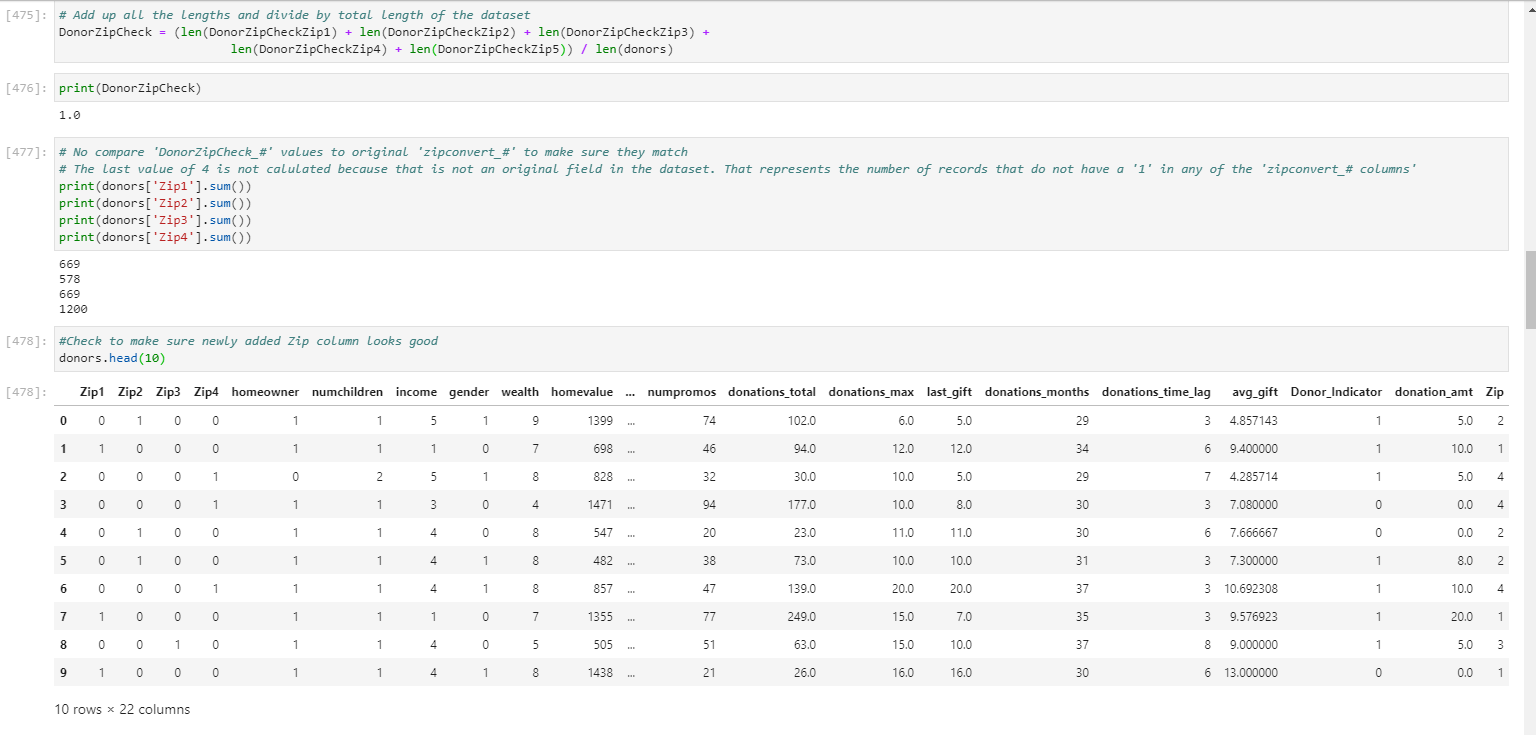
I utilized the donor dataset that was provided through this course. There was not much cleaning that needed to be done. A few columns were removed that were not necessary for analysis but besides that there was not much additional data cleaning that was done. I also decided to combine the multiple zip fields in order to perform analysis on this attribute. These cleaning techniques and creation of a new zip field will be shown later in this report.



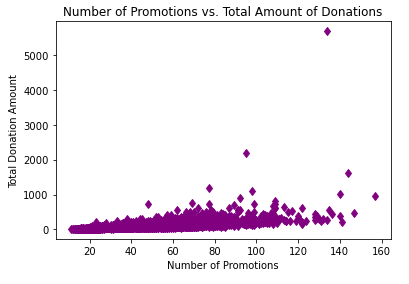






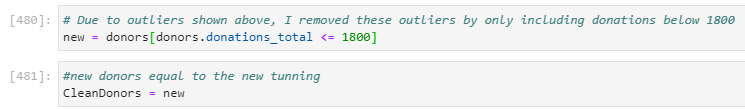


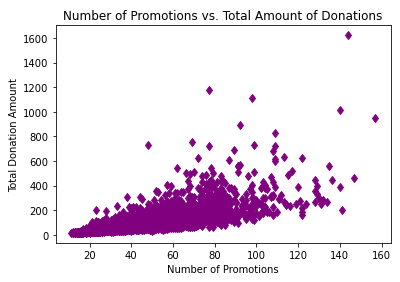
**Question 1: Do the number of promotions impact the Total Amount of donations?**

****

After running this first plot I found that outliers were making it difficult to analyze all of the other data. I removed the outliers in the next step.

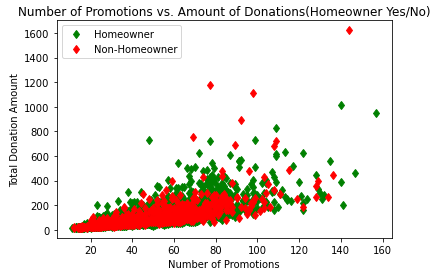
**Question 1 (Contd.): To get a better look at the data, I excluded the outliers that had a total value of over $1,800**

****

****

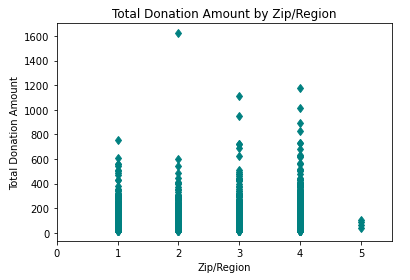
It is clear to see from above that the more promotions, the more total donations are made. This makes sense from a practical standpoint. If one receives more money from a promotion they are more likely to donate.

**Question 1 - Extended: Does being a Homeowner affect whether promotions lead to more total donations**

****

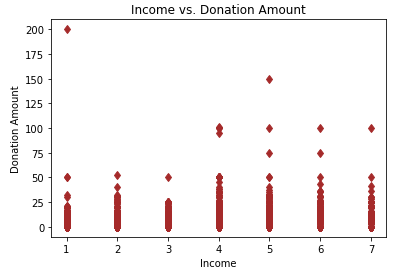
This plot is very similar to the previous but this takes into account the split between Homeowners and Non-Homeowners. Up to 60 promotions it seems pretty similar between Homeowners and Non-Homeowners but beyond that it is clear that Homeowners tend to donate more than Non-Homeowners.

**Question 2: Is there a relationship between Zip/Regions and Total Amount Donated?**

****

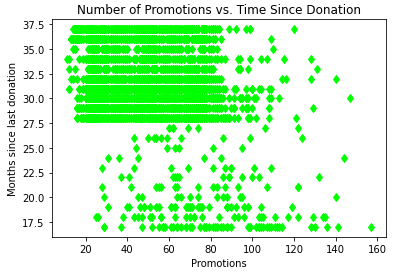
As shown above, Zip/Region 3 & Zip/Region 4 have the highest total donation amount. There was one donation from Zip/Region 2 that was higher than any other. Without more information on these zips/regions, we cannot make much more sense of what this means.

**Question 3: Are Income and Donation Amount correlated?**

****

Shown above, more income does correlate with higher donation amounts. This makes sense but what is interesting is the highest donation amount came from someone in the lowest income bracket. This can be attributed to a possible data error or maybe just a person that truly believes in donation even if they don’t have a lot of income themselves.

**Question 4: Does number of promotions influence the time since the last donation?**

****

Shown above describes that if you have less promotions you are less likely to donate. This is shown by the big cluster shown in the top left corner of the plot above. As promotions increase, the months since the last donation decrease.

**Final Conclusion:**

A lot of conclusions were able to be made from this analysis. First, when the number of promotions increase, so do the total amount of donations. Outliers were present in this dataset so they needed to be removed to portray an accurate and explainable story of the data. Furthermore, Homeowners tend to donate at higher rates than Non-Homeowners. Higher Income also correlated to higher amounts of donations, which is what would be expected. Finally, those with fewer promotions seem to have bigger gaps in between donations being made.