**Outcome of your EDA**

The outcome of the EDA was that hypothesis 3 had the highest correlation and likelihood of being most accurately defined. Hypothesis 1 and 2 I cannot accurately accept as fully true with the dataset used. Further information and sources would need to be gathered to further investigate any correlations in variable types.

**What do you feel was missed during the analysis?**

One variable I feel is needed from another data source to test the other hypotheses is the time since the last transaction. If there was a larger time between the transactions, then it was less likely to be fraud if the distance between is a little higher than in other circumstances. Further datasets will be needed with additional variables to fully prove or deny the hypotheses.

**Were there any variables you felt could have helped in the analysis?**

One variable that could have helped in the analysis or understanding is the time since the last transaction. This could provide more indications of fraud or non-fraud transactions. There is a lot of outside factors that could come into play with fraud not included in the dataset that is causing a less than accurate correlation between variables.

**Were there any assumptions you felt were incorrect?**

I don’t think any of the assumptions are 100% incorrect. I believe they are correct to a certain degree or percentage of the time. This could be tested if further outside variables were considered. All the hypotheses could be an indication of fraud in the correct scenarios or added to other triggers.

**What challenges did you face, what did you not fully understand?**

One challenge of this dataset was the number of variables that were true or false only. This increased the difficulty in running the Python programs in an understandable way. I believe to more accurately prove or disprove these hypotheses, more variables would need added to a dataset like this one used.