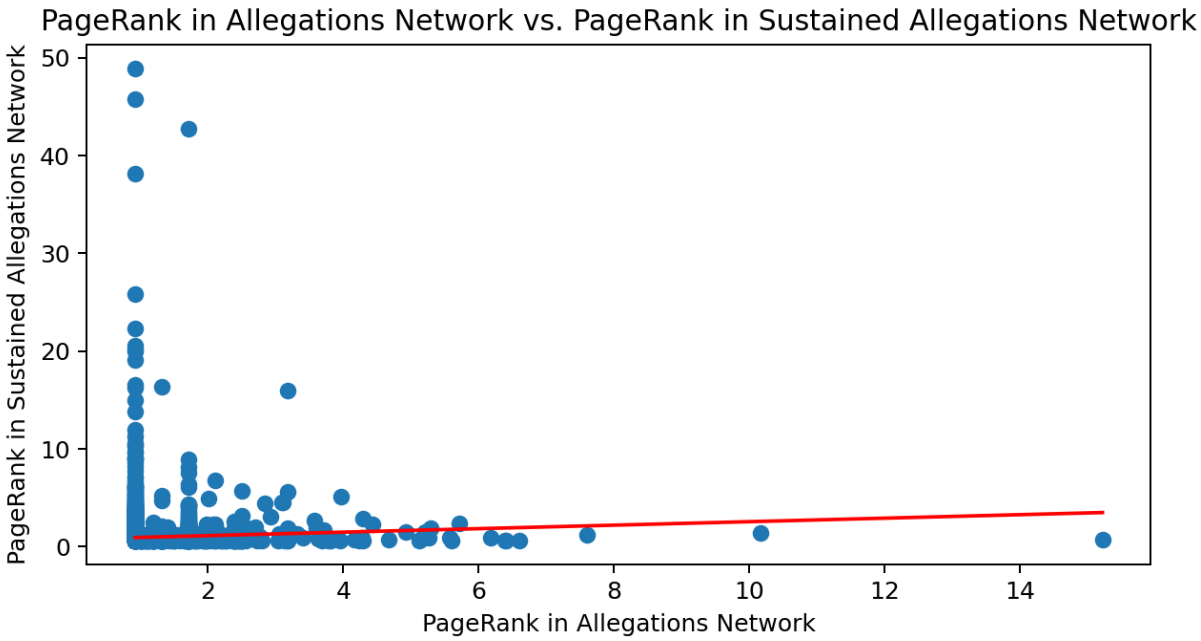
Questions:

- How does an officer's pagerank in the allegations network correlate with their position in the sustained al legations network? How does this correlation compare to just looking at an officer's number of allegation s vs. number of sustained allegations?
- How are officer node degrees in the Allegations network correlated with sustained allegations?

I created two graphs to answer the above questions. I chose to only include officers with ten or more allegations because I did not want the data to be clouded by officers with small numbers of allegations. The Allegation Network consisted of edges between officers that were in the same allegation while the Sustained Allegation Network only included edges from sustained allegations.

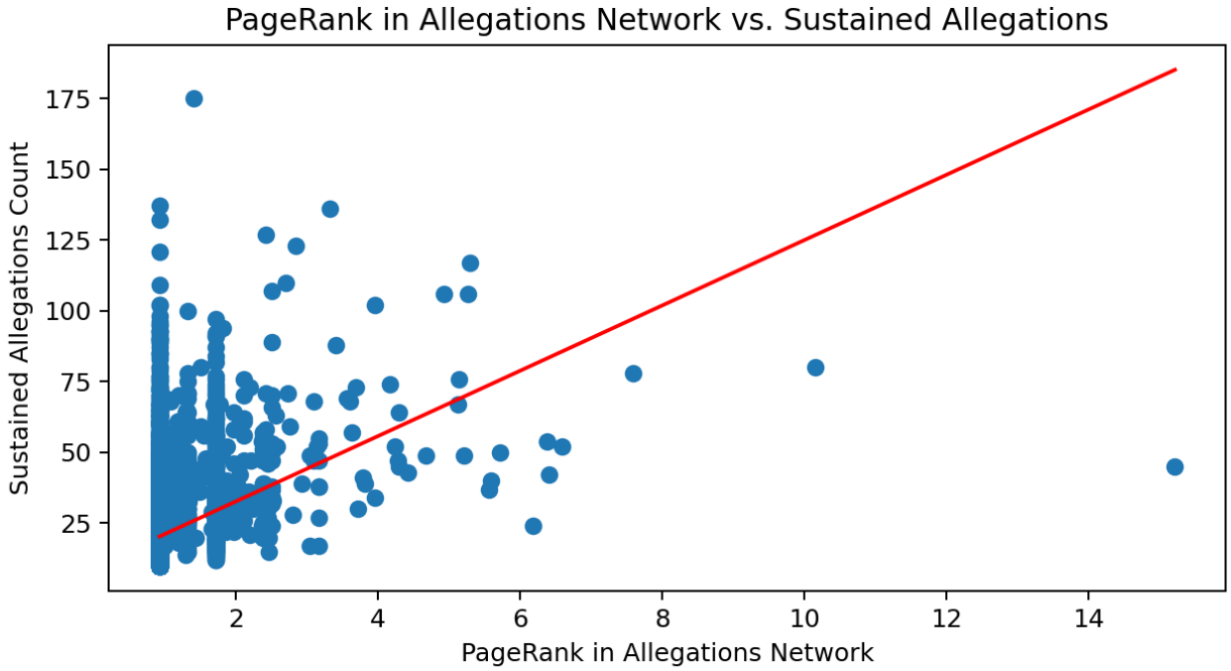
Below are the results from running the PageRank algorithm on both the Allegation Network and the Sustained Allegation Network. It is apparent from the graph and the statistical tests that there is no meaningful correlation between the PageRanks in either network for a given officer. A lack of correlation here shows that an officer's PageRank centrality in one network has no bearing on their ranking in another network, so trying to predict an officer's centrality in the Sustained Allegations Network is not possible with just their centrality in the Allegations Network, making it difficult to predict which officers will become central in the Sustained Allegations Network.




```
R Squared: 0.0025645108615160706
Pearson Correlation Coefficient 0.05064099980762659
Spearman Correlation Coefficient 0.10570854388338251
```

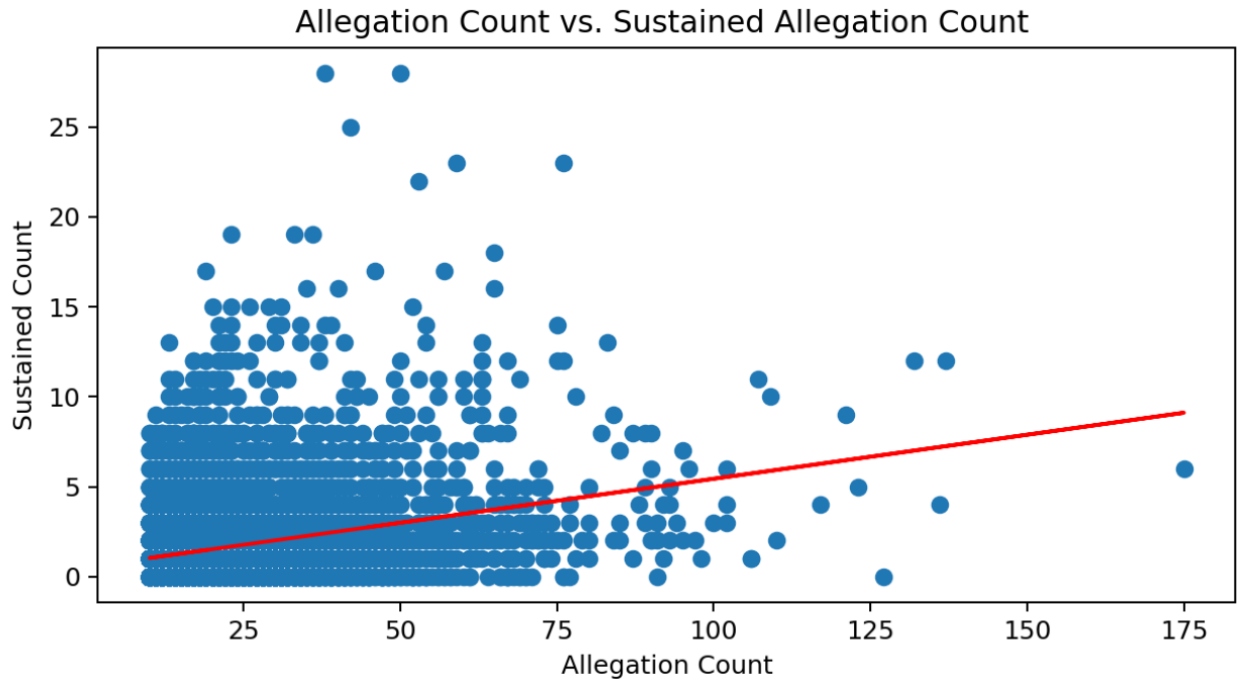
Below are the results of graphing the PageRank in the Allegations Network against the number of Sustained Allegations against a given officer. As we can see, there is a moderate positive correlation between the variables. More importantly this correlation is higher than the correlation between

allegations and sustained allegations for a given officer, which can be seen two graphs below in the graph titled *Allegation Count vs. Sustained Allegation Count*. This is important because it shows that the PageRank centrality of an officer in the allegations network can be more predictive of the number of allegations this officer has sustained and possible will sustain in the future when compared to predicting an officer's sustained allegation likelihood using just a fit from the relationship between allegation counts and sustained allegation counts. This improvement shows that graph analytics can help in the predictive power of which officers may or may not commit misconduct that will be sustained.



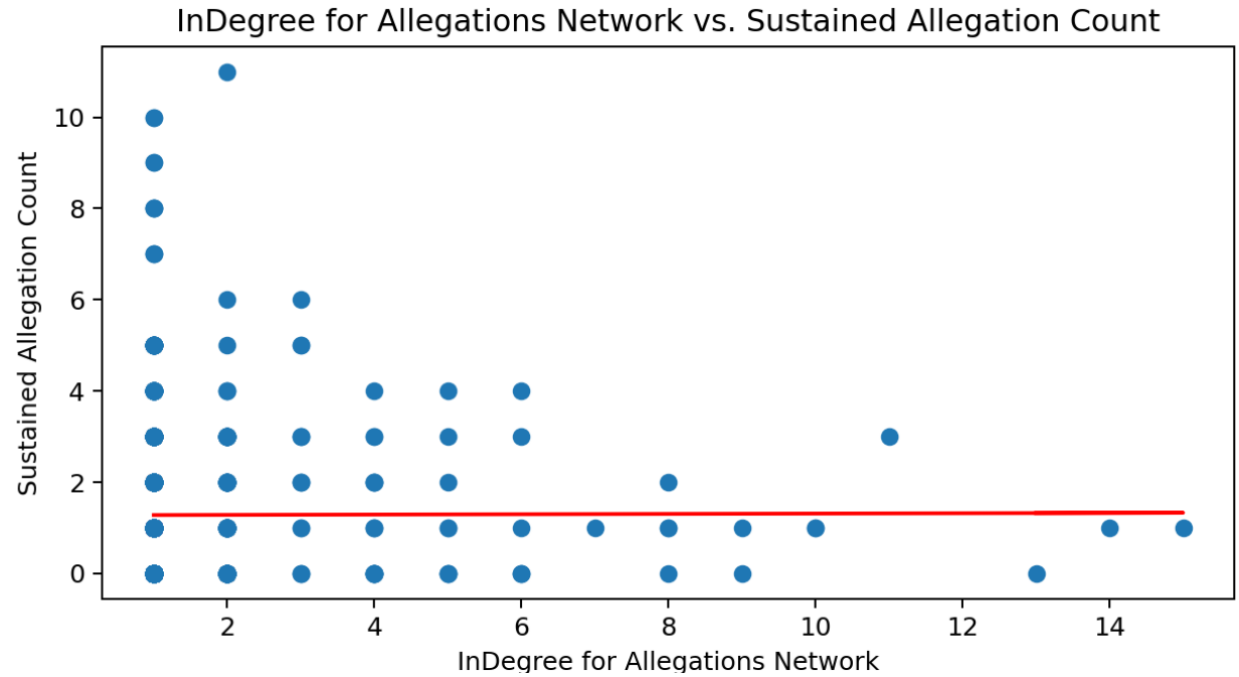

```
R Squared: 0.10905566180371778
Pearson Correlation Coefficient 0.3302357669964258
Spearman Correlation Coefficient 0.3219471033631776
```

The below graph of allegation count against sustained allegation count serves as a baseline for the other relationships so that it is easier to see which methods of graph analysis will lead to better predictions of which officers will commit misconduct. These variables have a moderate positive correlation as it would make sense that officers that have more allegations against them would also probably have more sustained allegations as well.



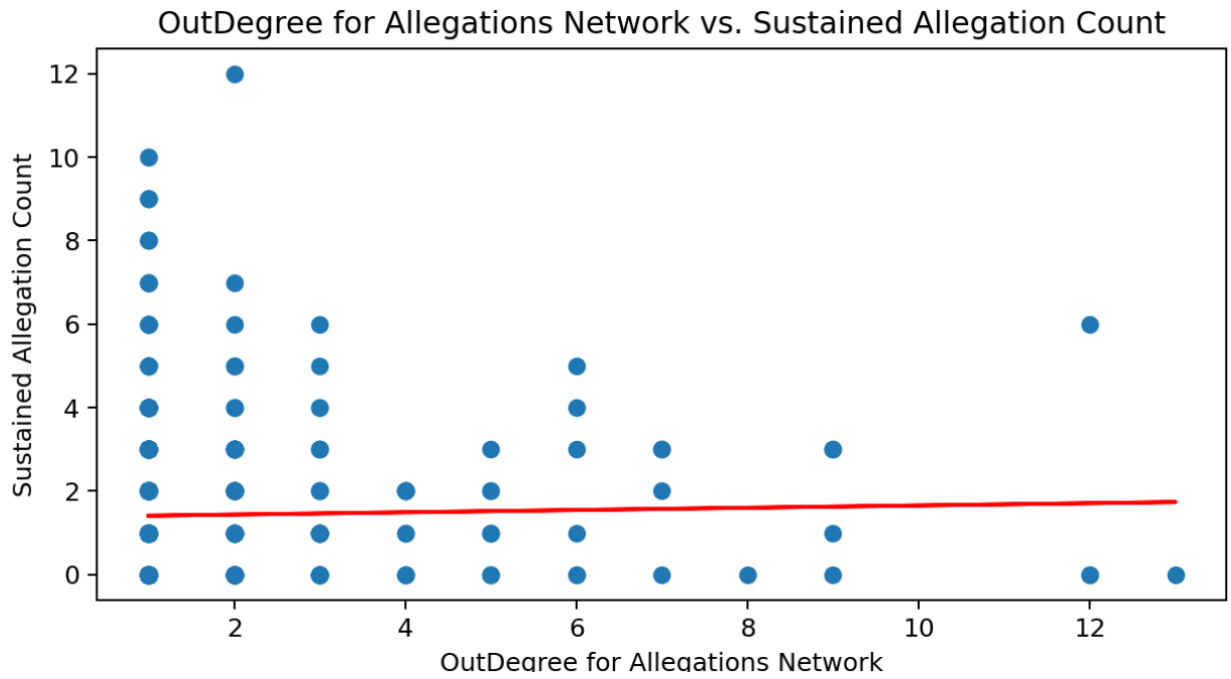

```
R Squared: 0.08547698751276421
Pearson Correlation Coefficient 0.29236447717321
Spearman Correlation Coefficient 0.24444004127077873
```

The below graph shows the relationship between the InDegrees for the Allegations Network against the Sustained allegations count. Here we see that no correlation exists between these variables.



The below graph shows the relationship between the OutDegrees for the Allegations Network against the Sustained allegations count. Here we see that no correlation exists between these variables.

```
R Squared: 1.1176062804581476e-05
Pearson Correlation Coefficient 0.003343061890629783
Spearman Correlation Coefficient 0.06041296471188047
```




```
R Squared: 0.0004728404178542964
Pearson Correlation Coefficient 0.021744894063993422
Spearman Correlation Coefficient 0.050600088410802956
```

Ultimately, the degrees of a given officer node are not useful for predicting whether they will have allegations sustained against them. However, their PageRank in the Allegations network has a moderate positive correlation that is higher than the correlation between allegations and sustained allegations, showing that there is good reason to believe that utilizing PageRank centrality in a graph of co-accused police officers may be more predictive of which officers receive misconduct than just linearly fitting a graph of allegations vs. sustained allegations.