My name is Gabrielle Bartomeo and I'll be going over my final project.

For the final project, it was required to apply the techniques we have learned in this class to a topic of our choosing. I decided to focus on history and pop culture by examining the characters from Marvel's Thor comics and their inspirations from the Old Norse texts of the Elder Edda and the Younger Edda. I was looking to identify:

One: What alignments did the mythological beings that inspired Thor have based off of text analysis - Good, Bad, or Neutral?

Two: How did these alignments found as a result of text analysis compare to what is typically accepted to be the alignments? How about Marvel's?

And Three: Were the beings in these texts more likely to interact with those of their own alignment, or another alignment?

In order to complete this project, seven libraries were used. Requests was used for bringing in the text file with the Eddas. The library arr ee was used for cleaning up the text and finding matches. En Ell Tee Kay has a sentence tokenizer and a sentiment intensity analyzer with which the majority of the project relied on. Making data frames and performing functions on them was done using pandas. Mat plot lib's pie plot and network ex were used to visualize the data. Num Pie had a few useful functions in it, so it was used as well.

The Marvel Wikia's data was used to gather the alignments of the characters in Thor. This database was scraped by the company 538 and put into a Cee Ess Vee in their Github. Pandas was used to bring in the CSV, and Re was used to clean it up and pick out only the characters we needed. As you can see, we decided we would only be working with a set of twenty one characters and the mythological beings that inspired them.

The raw text of the two Eddas was then gone through to be broken apart into sentences. Due to the archaic nature of the language used in these documents, a total of seventy eight words had to be added to the lexicon before sentiment analysis could be performed. Using En Ell Tee Kay's sentiment intensity analyzer, every sentence was given a score of positive, neutral, or negative if there was one of the twenty one beings mentioned. A score called positive negative was calculated where the negative score was subtracted from the positive score to see if a more polarized score alignment might be more accurate. A z-score was then calculated for each sentence's positive, neutral, or negative score by analyzing all the other sentences' values alongside it. It resulted in a database that is being sampled for your viewing at this moment.

This here is the first ten rows of the database created using the previous database. The rows labeled average positive, neutral, and negative were created by averaging all the sentences' sentiment scores for a given being. The positive negative score assigned to each sentence before was also average. The z-scores were averaged as well. Lastly, the average positive, neutral, and negative scores created for this chart were then given z-scores themselves, and were called the normalized average sentiment score. These scores were then used to give an alignment to each of the beings.

The average positive negative score's alignment assignment was based off of whether the score was above or below 0. Below zero, Bad; above zero, Good. The z alignment was based off of whichever value was the highest of the average z-scores. Lastly, the normalized alignment was based on the highest of the normalized average sentiment scores. The only true anomaly in these scores is Freyja who was categorized as Bad due to the number of insults she's received. These alignments were then compared against the commonly accepted alignments for each of the beings, which resulted in an accuracy for each method of assigning alignments. While the average of the z-scores and the normalized average sentiment scores had the same rate of accuracy, ninety point forty seven percent, the latter's alignments were used for the network analysis performed.

This graph displays every sentence that mentions at least two of the mythological beings from the Eddas. Each of the nodes are colored corresponding to alignment, save the sentences, which are all green. From this, what is most immediately visible is the number of sentences that each of these beings are in. Specifically, Odin, Thor, and Loki have the most number of mentions by far. Using this, network made a projected bipartite graph that examined the relationships between each of the mythological beings.

This is the projected bipartite graph. As you can see, the nodes are again colored by alignment, but this time, they are also shaped by their status of being – monster, jotunn, or God.

Based off of the number of neighbors each being had, it was more likely for Gods to have some sort of relationship, where as jotunn were the most isolated. While monsters had a variety in the number of relationships, overall, they tended to have more relationships than jotunn but less than Gods.

The number of relationships these beings had to each other, taking account into alignment of each neighbor, was then calculated. Each being's results were then combined for their given alignment to produce the first chart, which counts the number of unique neighbors for each of the alignments. Each of these totals was then compared against the total number of beings in said alignment to produce a percentage for each alignment. Lastly was the correlation of this chart, which is the final chart. Despite the percentages calculated before, those of the Neutral alignment were far more likely to interact with those who were Neutral or Bad, whereas those who were Good kept to themselves, and were unlikely to interact with another alignment.

In conclusion, the results of this project found the alignments of the mythological beings with over ninety percent accuracy with one major exception, Freyja. When it came to alignments, Marvel missed the ball, but this is likely attributable to comic books having characters that are more polarized in the first place. The Normalized Average Sentiment Score was the best way to identify the alignment of a given monster, jotunn, or deity. There was a positive correlation between the alignment a being was and what alignments they would tend to interact with, usually favoring their own over others.