The questions: Which episodes tended to do better? Building off of that, which season is the most popular season?

The primary purpose to analyze imdb reviews was to understand which seasons and episodes had the highest ratings. Another question that followed was which season was the most popular season. To answer these questions, we pulled data from the imdb api. In the imdb data we pulled the ratings per episode in each season as well as the number of votes that each episode received. Finally, to find out the most popular season, we analyzed the average ratings per season as well as the average amount of votes overall. Based on the data, most ratings and votes tend to increase towards the end of the season. Overall, season 4 was the most popular season by rating but season 6 had the highest votes overall. Which then leaves the question, which one is actually the most popular season? When looking into the the votes of both season 4 and season 6, they were pretty close. But the last two episodes of season 6 received higher votes.

To retrieve this data, we pulled from the imdb api and extracted the rating score and the number of votes. Then with this data, we created a dataframe for each season and then plotted the graphs.

Some hiccups, that were encountered initially was converting the string values into integers or floating numbers so that the y axis of the plots are actually arranged in order.





