

Could Chandler Bing *be* any more popular?

Project Summary

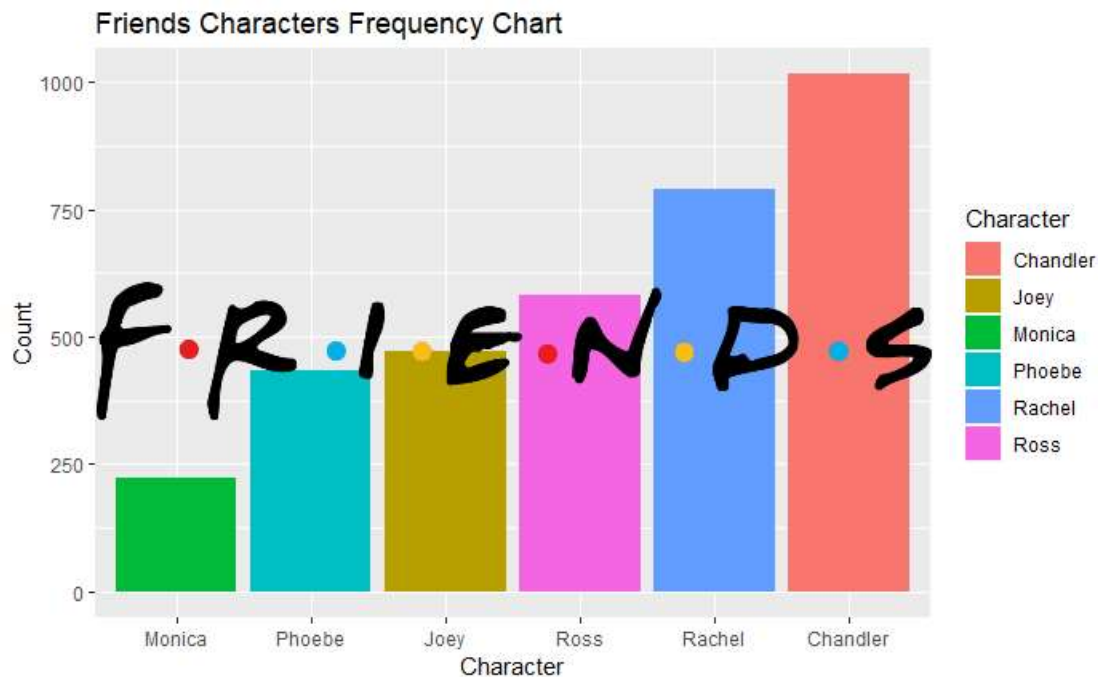
There is always a debate among the Friends serial fans who among the lead six actors is the most likable character. For the final project I am going to investigate which character out of the six main characters in Friends series is famous among twitter users. Most of the poll outcomes find Chandler to be the most famous character and I intend to find whether the same opinion is true among twitter users even now. This type of analysis has a wide set of audience from Ad companies to fashion industries. They can benefit from understanding which character resonates well with the public and utilize that knowledge in ad campaigns or in designing fashion apparel/merchandise based out of that character.

Data Collection

Data is collected using the Twitter API method using Python with the help of identifying tweets containing the full name of the characters. This way we collected 6 json files that contains all the tweet mentions of these character in the last week. The data is cleaned to remove duplicates and retweets. By the end of data collection process, there are 3516 tweets ready for analysis.

Analysis

First a Chi Goodness of Fit test is performed in the data with the counts of all the characters against expected proportion as $1/6$ for each of them. The test returned a p value $< 2.2e-16$ which helped us to reject the hypothesis that the tweets are equally distributed among them.



Next, to further confirm that Chandler gets tweeted about more than the other character, a two-sample t-test is performed between Chandler and each of the other characters. The two-sample t test is performed to identify if mean of Chandler tweets is greater than the other characters thereby proving his popularity in twitter.

The hypothesis for the two-sample t test with alternative = greater is as follows

Could Chandler Bing *be* any more popular?

H0: $\mu_1 - \mu_2 = 0$

Ha: $\mu_1 - \mu_2 > 0$

Conclusion

A series of 5 tests are performed and the results are tabulated as follows

Character 1	Character 2	pValue
Chandler	Ross	0.000375307
Chandler	Joey	0.000124333
Chandler	Monica	5.22E-06
Chandler	Phoebe	7.51E-05
Chandler	Rachel	0.04808094

From the results, we can say there is statistically significant evidence that the frequency of tweets about Chandler is greater than the tweets towards others. Among others Rachel comes close to Chandler in terms of popularity.