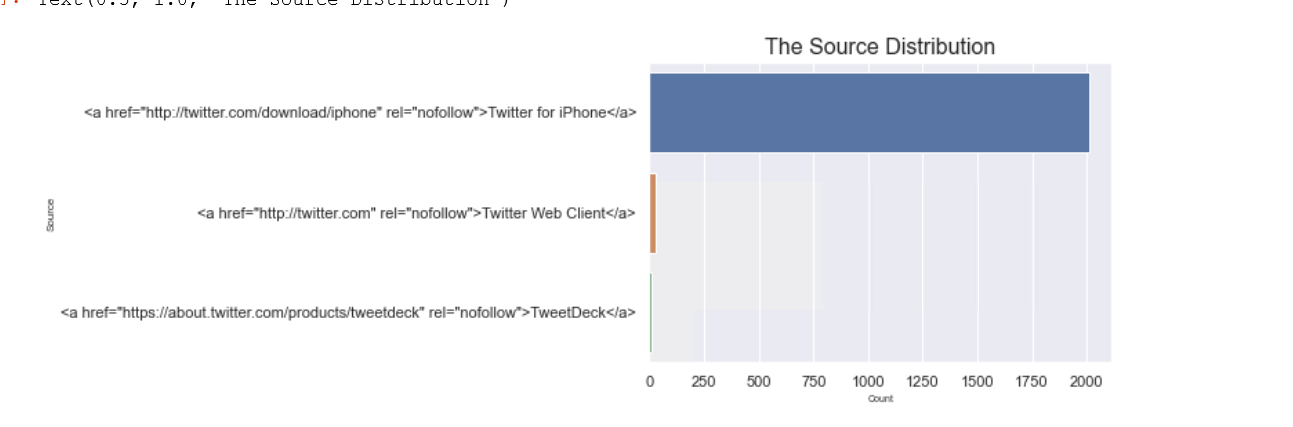
**Summary of the Analysis and insights**

This is the last phase, however the most important part for wrangling process to be able to get some insights that we have gathered from the previous phases.

1. The first question that I wanted to get the answer is the biggest source of the tweets? In this assessment I have taken the plot as follows

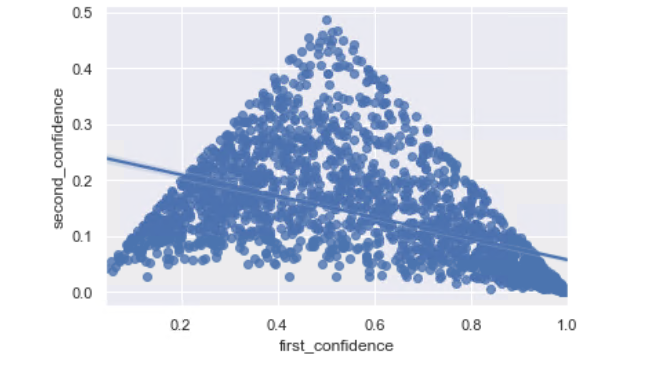


The highest source of the tweets came from the IPhone app of the twitter. This Is interesting insight on how the users have adopted to the technology and almost 2000 plus tweets out of 2053 range of the data index. This is evident from the graph.

1. Does the first and second prediction increase all the time? What is the relationship between them?

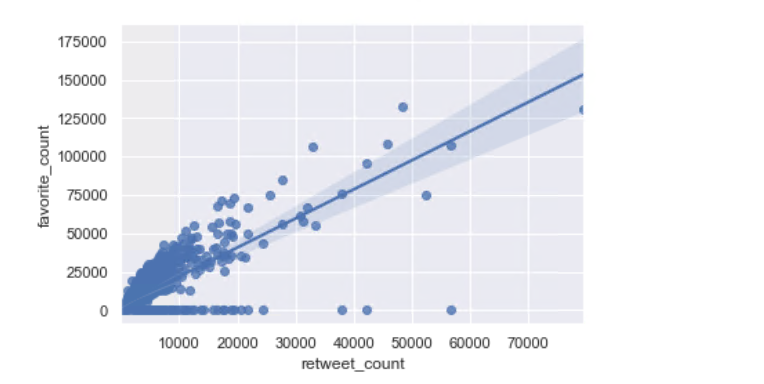
I chose a simple scatter plot to get the relationship and it shows the first confidence and second confidence have a negative relationship as per the graph below. The correlation coefficient

-0.5114049885418799.



1. Do the most favorite ones are retweeted more?

Looking at the data for the relationship between the tweet count and the retweet count of the dataset. Looking at the following plot



The coefficient relationship 0.7899404711971294 between the retweet\_count and favorite\_count swows are strong positive relationship between the two two variables.

So the categories that have been the most favorite have been tweeted again the most.

**Conclusion**

**As a part of this process of data insights it has been observed that iphone is the biggest source of all the tweets and the most favorite categories have been retweeted the most. Looking at the 2nd and 1st prediction trend they are not always increasing. It’s in inverse (negative) relationship.**