



# **Insights/ Results Report**

By Areej AlZahrani

# Storing

Will store two csv files. The first one "twitter\_archive\_master.csv" is to hold major columns and important for analysis. The second "tweet\_details\_master.csv" will have detailed info like the 2nd, 3rd, and tweet text info.

# Analyzing

There are few questions I would like to answer regarding this data. Here are the three questions presented with their findings:

1. Is the tweet source related to high rating?

The relationship between the source and rating is that tweets from TweetDeck have the highest rating mean. Vine and Web sources tend to have lower ratings while tweets from iphone are average.

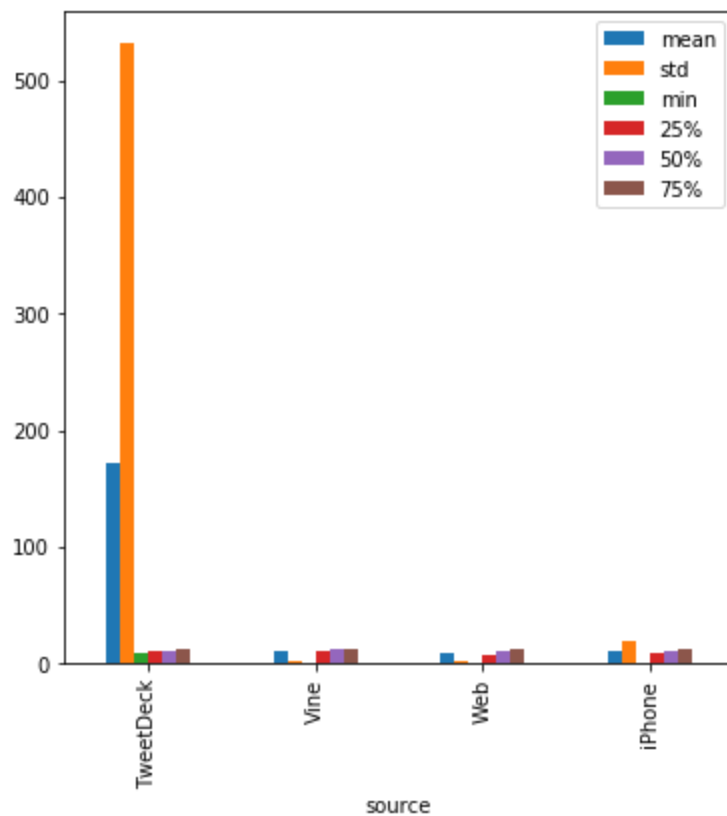
2. Does high retweets mean high rating?

As you can see having high rating is not necessary to lead to more retweets since the group of medium retweets have highest rating avg.

3. What is the relation between the first prediction and having dog info i.e., the dog name, stage is not null?

as you can see having unknown name or stage of the dog leads to better predictions! although the sample size is large but still having known name and stage don't give good prediction confidence.

# Visualizing



This bar chart shows the relationship between the source and rating is that tweets from TweetDeck have the highest rating mean. Vine and Web sources tend to have lower ratings while tweets from iPhone are average.