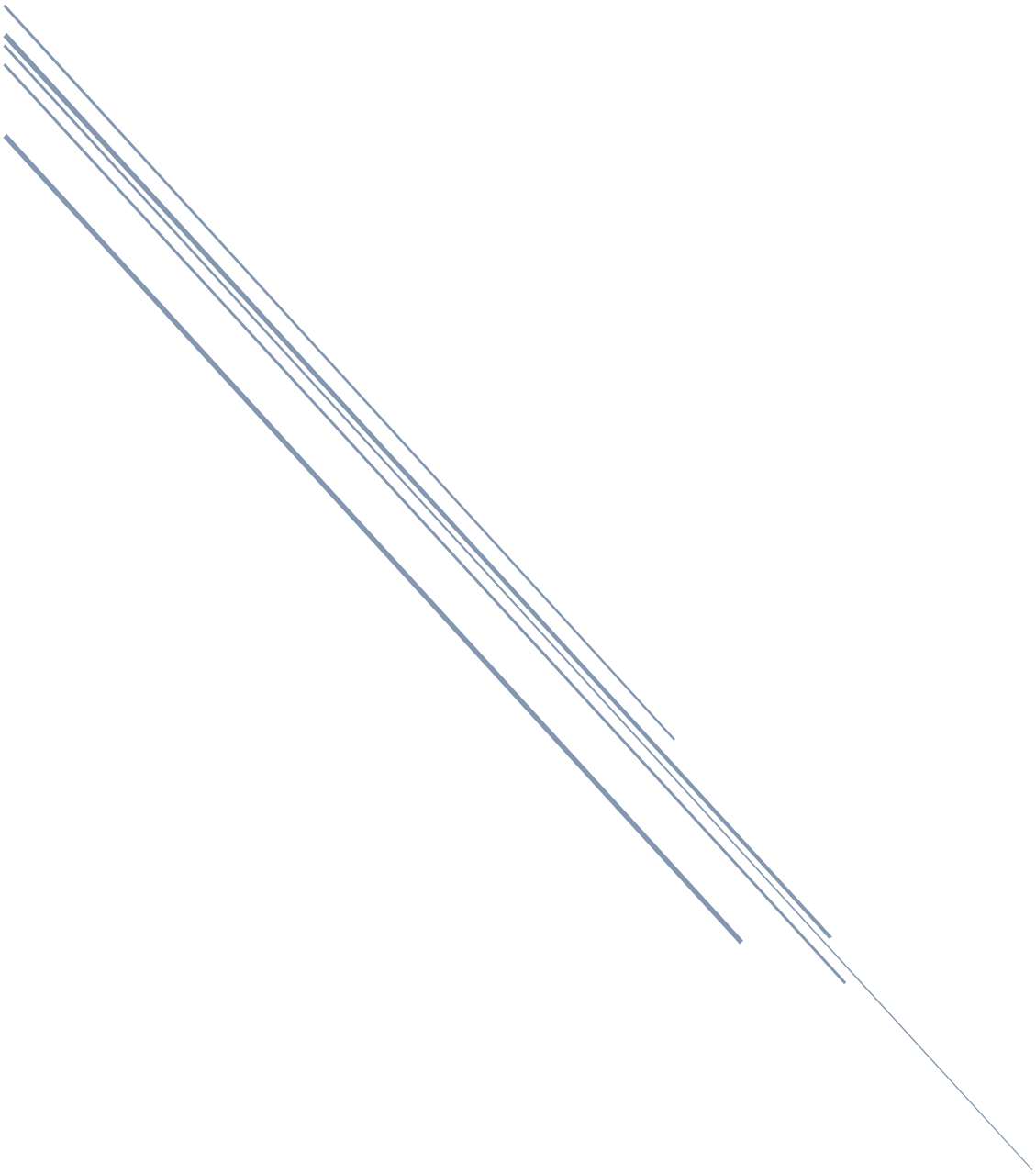
WRANGLE ACT REPORT

Dalal AlAli



Udacity-Data Analyst Nano Degree program

Project five: Wrangle and Analyze Data

Introduction:

In this project we are going to wrangle and analyze the tweet archive of Twitter user @dog_rates. This account @dog_rates is a Twitter account that rates dogs. The account opened in 2015 by a man who called Matt Nelson, and has received international attention. By Using Python and its libraries,I will gather data from many sources then,I will assess these data and its quality and tidiness issues, then,I will clean it. This is called data wrangling. And I will document my wrangling efforts in a Jupyter Notebook.

Insights:

1-I represented the most retweeted tweet in the dataset

2-I represented the most favorite tweet in the dataset?

3- I represented the average tweet retweet ?

4- I represented the average tweet favorite ?

Visualization:

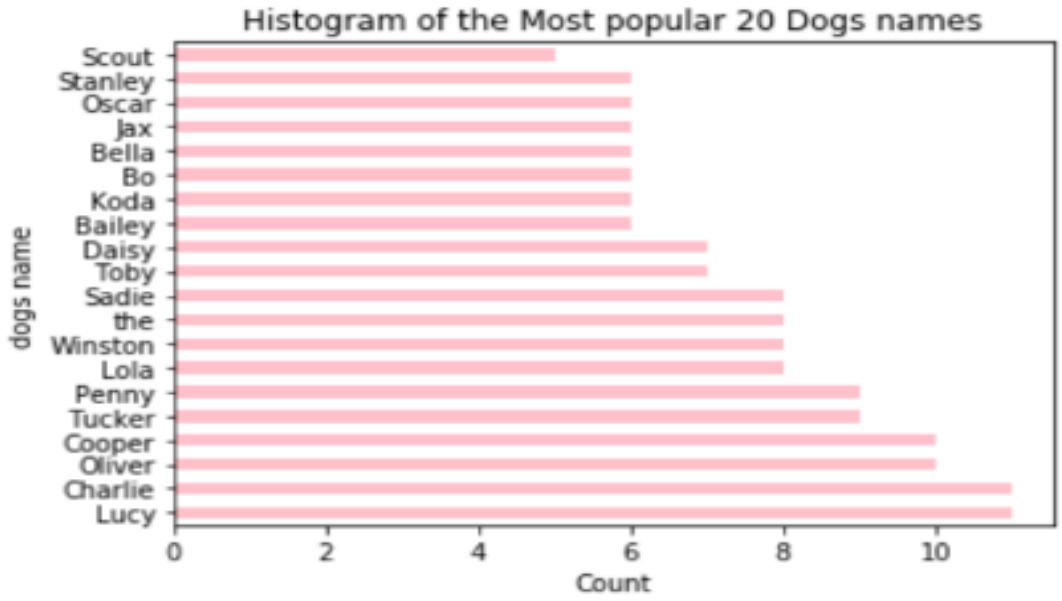
1-Show relationship between favorites and retweets?

tweets_json_clean.plot(x = 'retweet_count', y= 'favorite_count' ,color='pink', kind = 'scatter', figsize= (6, 4))

plt.title('Relatioship between favorites and retweets');

plt.show()

by this code I showed the most 20 popular dogs names.



This Histogram represent the most 20 popular dogs names which are Lucy, Charlie, Oliver, Cooper, Tucker, Light, Stormy, Nigel, Carter, Leonidas ,etc.

2-What is realitionship between favorites and retweets?

```
tweets_json_clean.plot(x = 'retweet_count', y= 'favorite_count' ,color='pink', kind = 'scatter', figsize= (6, 4))

plt.title('Relatioship between favorites and retweets');

plt.show()
```

by this code I showed the relationship between favorites count and retweets count.

This scatter graph shows the relationship between favorites count and retweets count and we can see that both retweet count and favorite count are increasing.