ANALYSIS REPORT

ON

WeRateDogs

TWITTER

DATA

INTRODUCTION

This report briefly shares results from an analysis performed on data gathered from WeRateDogs twitter account. **WeRateDogs** is a twitter account that was started in 2015 by a college student named Matt Nelson.It rates people's dogs with a humorous comment about the dog. *{Excerpt from Wikipedia}*

The main purpose of this project being to perform data wrangling; we therefore gather this data from the **WeRateDogs** twitter account and assess it then finally clean it to make it easy to run the analysis needed for this report.

INSIGHTS AND VISUALIZATIONS.

The datasets had a lot of cleaning to be done and fact is, there're a lot of insights to draw from

the analysis, but I will just share three insights and visualize one of the insights.

I gathered the following insights:

1. Pupper is the most popular dog_stage

Of the four dog stages, Pupper was the most popular stage leading with 212 dogs. This means that most of the dogs on our dataset belong to Pupper dog stage and very few come from the category of floofer. As it also clearly shows we had a couple of null values, which means the dogs were not categorized.

```
df_merged.dog_stage.value_counts()
```
Out[72]:

```
None      1994
pupper     212
doggo       74
puppo       23
floofer      8
Name: dog_stage, dtype: int64
```

2. Twitter for iPhone is the most popular source for dog_tweets

People use different gadgets to post their tweets. From our dataset it is clear most of the dog lovers use iPhones.

```
df_merged.source.value_counts()
```
Out[73]:

```
Twitter for iPhone    2269
Twitter Web Client      29
TweetDeck               13
Name: source, dtype: int64
```

3. Cooper is the most popular dog's name

It is evident there're high chances that most dogs out here share names. Dogs are pets so definitely people would easily have common pet names. Cooper represents the most popular dog's name in our dataset, closely followed by Oliver, Charlie and Lucy.
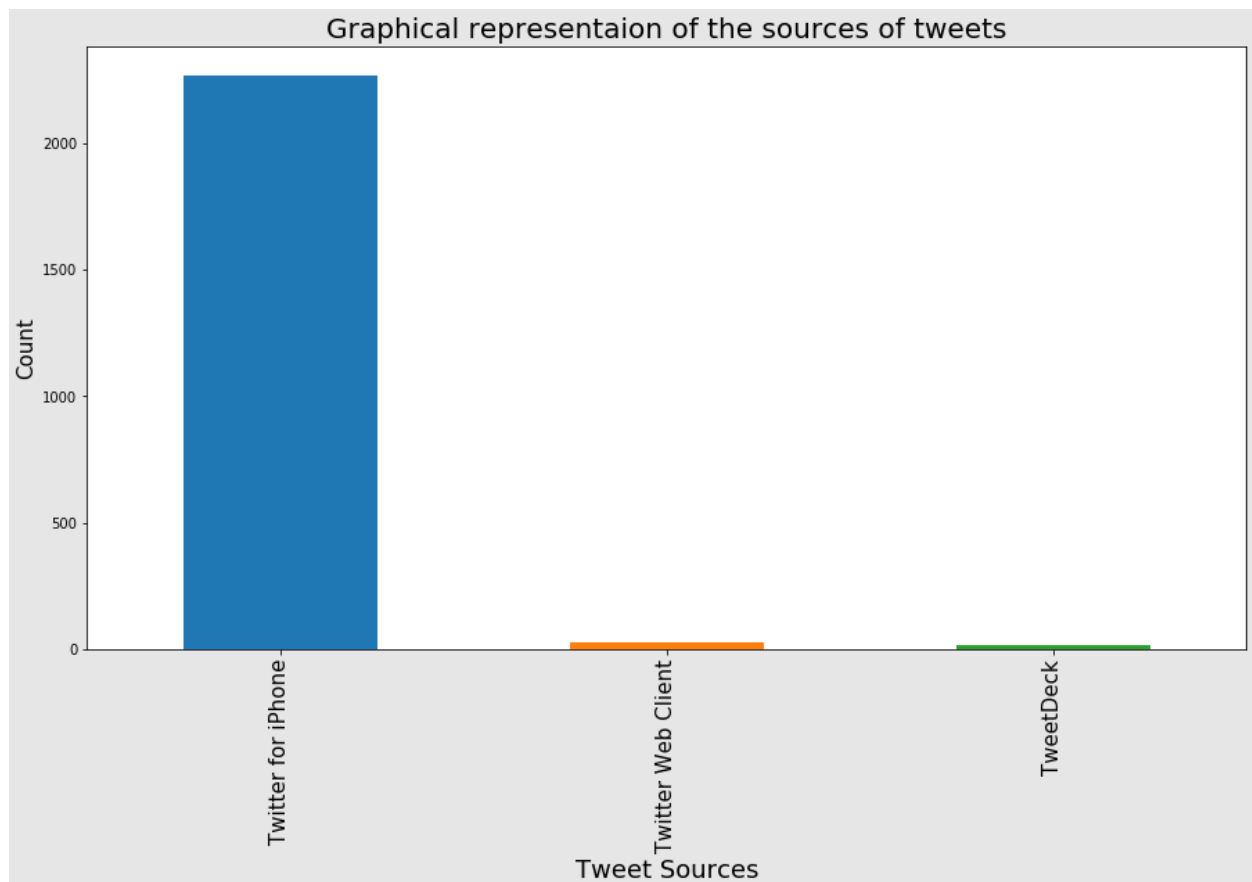
```
df_merged.name.value_counts().head()
```

```
NaN        111
Cooper      13
Oliver      12
Charlie     12
Lucy        11
Name: name, dtype: int64
```

Visualization of the most popular source of tweets.

The analysis clearly shows that "Twitter for iPhone" was the most popular source of tweets with a count of 2269 compared to the other two sources that had less than 50 people using them

Figure 1.0: Graph of the top most popular tweet sources