

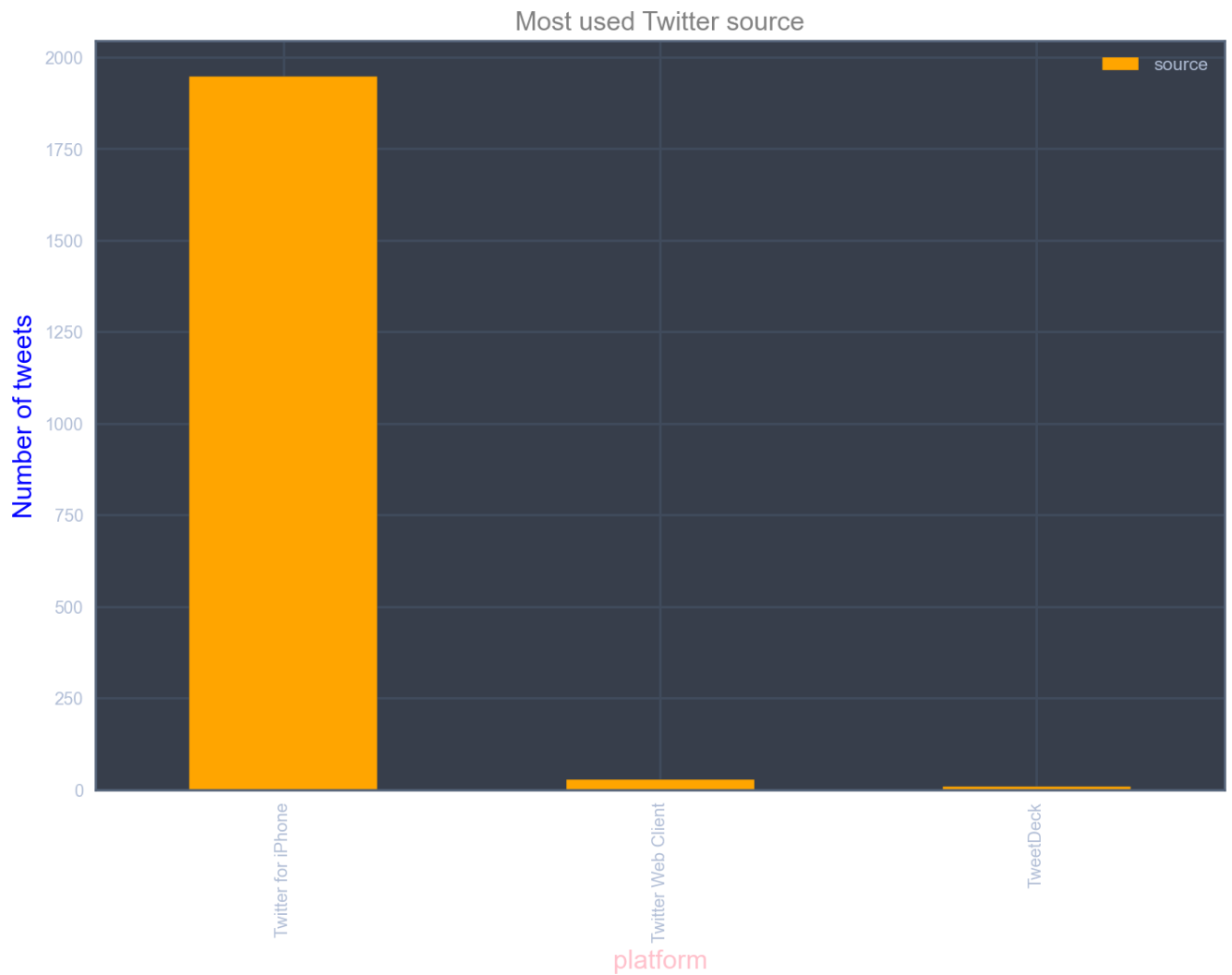
Introduction

Scarcely do Real-world data comes clean and this project on Data analysis of tweets by twitter user WeRateDogs is no exception, The dataset image_prediction.tsv, twitter_archive and twitter_data, was wrangled cleaned and merged to form one DataFrame used for the final analysis (analyzing and visualizing) of the User(WeRateDogs)'s Tweet. Basically the Twitter user @dog_rates, also known as WeRateDogs rates people's dogs with a humorous comment about the dog. they have a huge following of over 9 million and the following are visualization and observation obtained from the analysis based on available data

Introduction

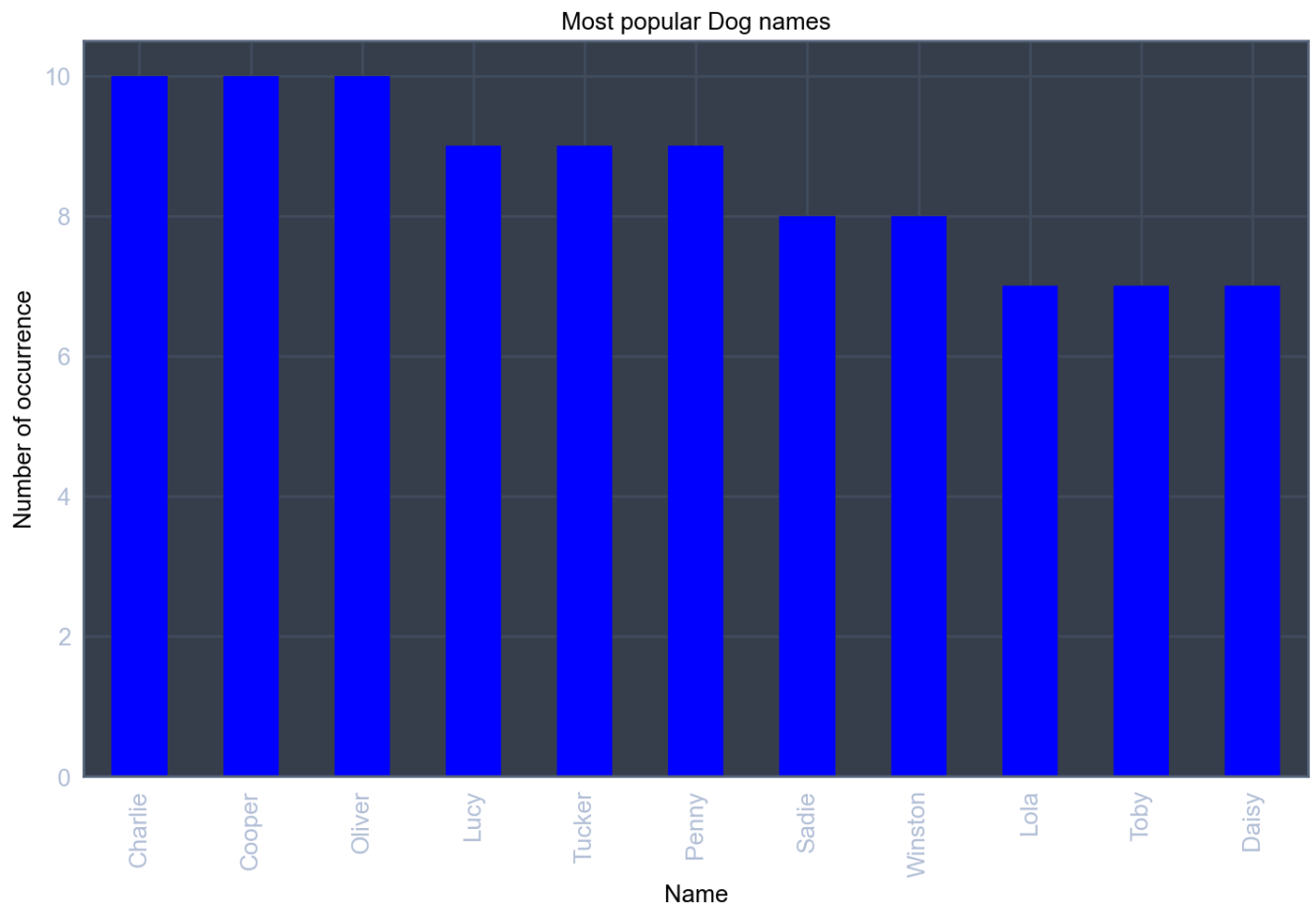


Followers preferred Device for Tweeting



we could clearly see that twitter for the iPhone is miles ahead in terms of followers device usage when compared to other devices (web/desktop client and twitter deck), but a very striking/ disturbing issue is the no records for the Android phone.

Common Dogs Names



although there were a lot of unnamed dog, users tend to pick "charlie, cooper, oliver" as their dog names hence the popularity amongst dog owners, with lucy and tucker completing the top 5.

Popular Dog Breeds

```
twitter_archive_master['p1'].value_counts()
```

```
golden_retriever 139
```

```
Labrador_retriever 92
```

```
Pembroke 88
```

```
Chihuahua 79
```

```
pug 54
```

```
...
```

```
nail 1
```

```
timber_wolf 1
```

```
espresso 1
```

```
bib 1
```

```
desktop_computer 1
```

```
Name: p1, Length: 374, dtype: int64
```

From the data dog breed value count, shows that the Retriever Breed is undisputed with the Golden retriever (the more hairy/ furry type) edging its Labrador retriever while the Pembroke, Chihuahua, pug comes in 3rd, 4th and 5th respectively.