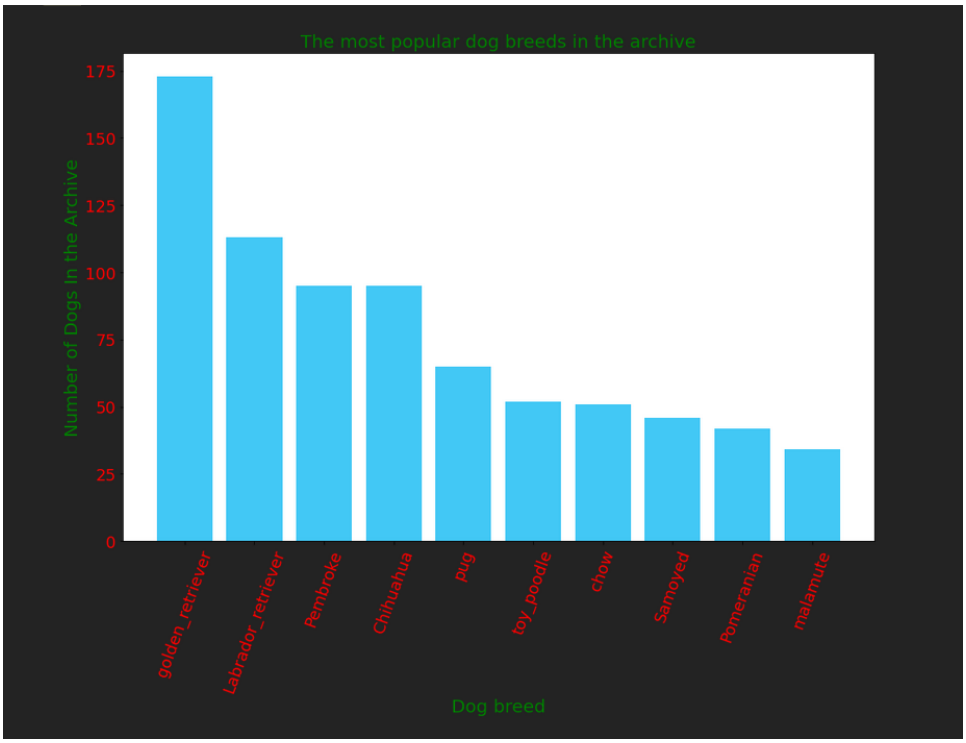
Finidings from the “WeRateDogs” account on twitter.

If we see that account as a good sample to the house dogs population outside, we could get a very interesting conclusions from this dataset.

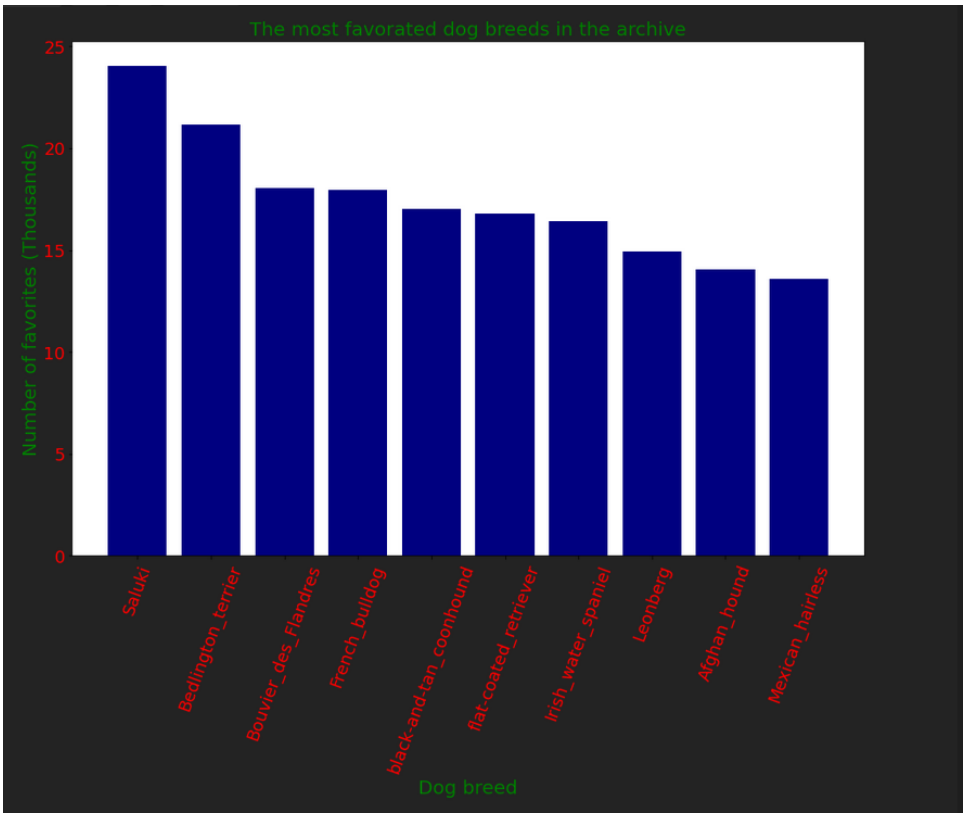
We could for example to see what are the popular dogs on the dataset, and it reasonable to say that there are the most popular dogs in USA and maybe other parts of the world too.

So after we cleaned all the data we did our explorations, and this are the most popular dogs in the dataset:

golden_retriever	173
Labrador_retriever	113
Pembroke	95
Chihuahua	95
pug	65
toy_poodle	52
chow	51
Samoyed	46
Pomeranian	42
malamute	34



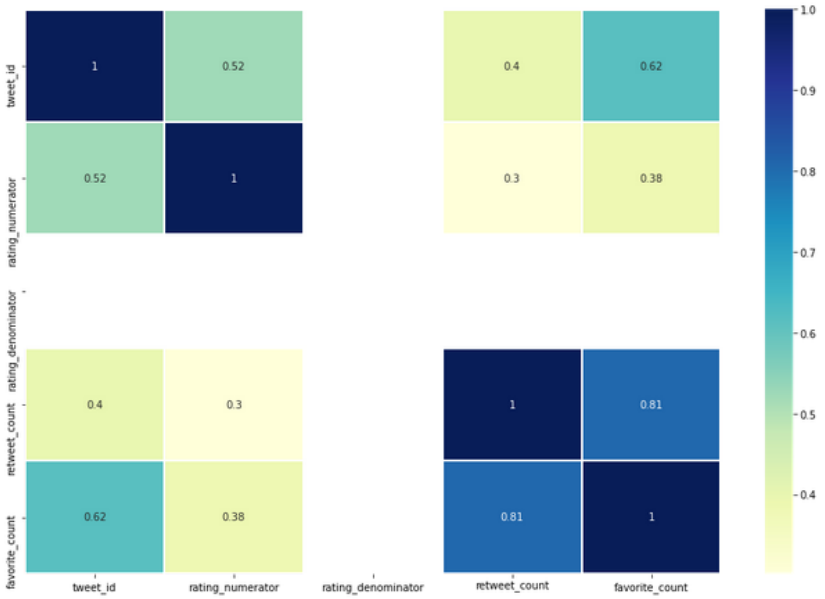
And what are the dogs breeds who gets the most of the favorites ?



(We see intersting thing – there are different breeds on the two plots ! The most popular isn’t the most favorite)

Than we asked ourselves is there a correlation between the rating on the site to the number of favorites ? Or the number of retweets ?

So we did a heat map:

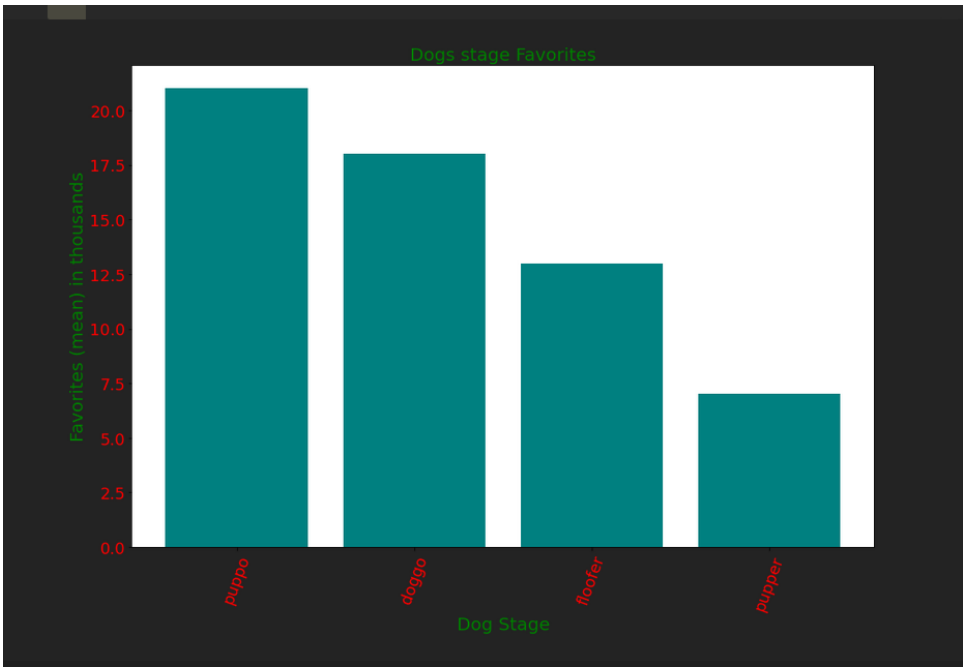


1.4.1 seems that the only correlation here is between re-tweets and favorites count.

From the map we see that the only correlation here is between retweets and favorite count (0.81 Pearson correlation). So, the rating of the account “we rate dogs” isn’t affect too much ….

Next we asked Is there a different in favorites to dogs in different "stage" ?

And we see that the staging has some impact :



But with this we must say that there are only few dogs with staging:

```
doggo        69
floofer       7
pupper      172
puppo        22
(270 dogs from 1750 – 15%)
```