

# wrangle\_act-Copy2

December 18, 2019

## 0.0.1 Getting some insights from the final cleaned data frame

```
In [48]: df.name.value_counts()
```

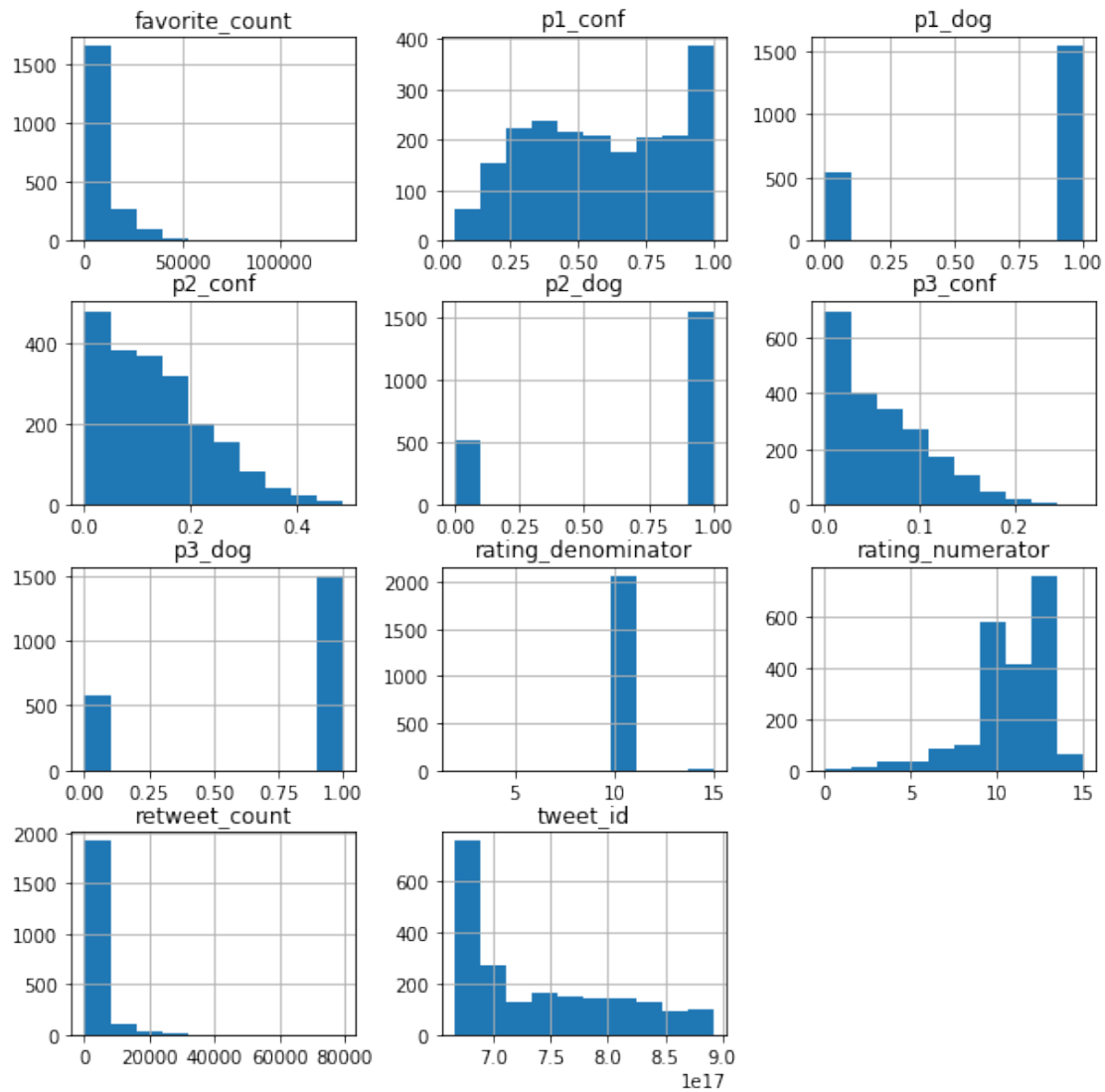
```
Out[48]: None          577
         a             55
         Charlie       11
         Penny         10
         Lucy          10
         Oliver        10
         Cooper        10
         Tucker        10
         Lola          8
         Winston       8
         Sadie         8
         Bo            8
         an            7
         Daisy         7
         the           7
         Toby          7
         Bella         6
         Koda          6
         Milo          6
         Scout         6
         Dave          6
         Jax           6
         Bailey        6
         Rusty         6
         Stanley       6
         Buddy         5
         Louis         5
         Chester       5
         Larry         5
         Leo           5
         ...
         Katie         1
         Tupawc        1
         Zara          1
```

Fletcher	1
Emma	1
Rey	1
Bloop	1
Cora	1
Dudley	1
Bertson	1
Joey	1
Gunner	1
Edgar	1
Anakin	1
Sonny	1
Devón	1
Jebberson	1
Lili	1
Harrison	1
Jeffri	1
Jeb	1
Cecil	1
Ralphie	1
JD	1
Ebby	1
Ridley	1
Bronte	1
Lulu	1
Bruno	1
Beebop	1

Name: name, Length: 936, dtype: int64

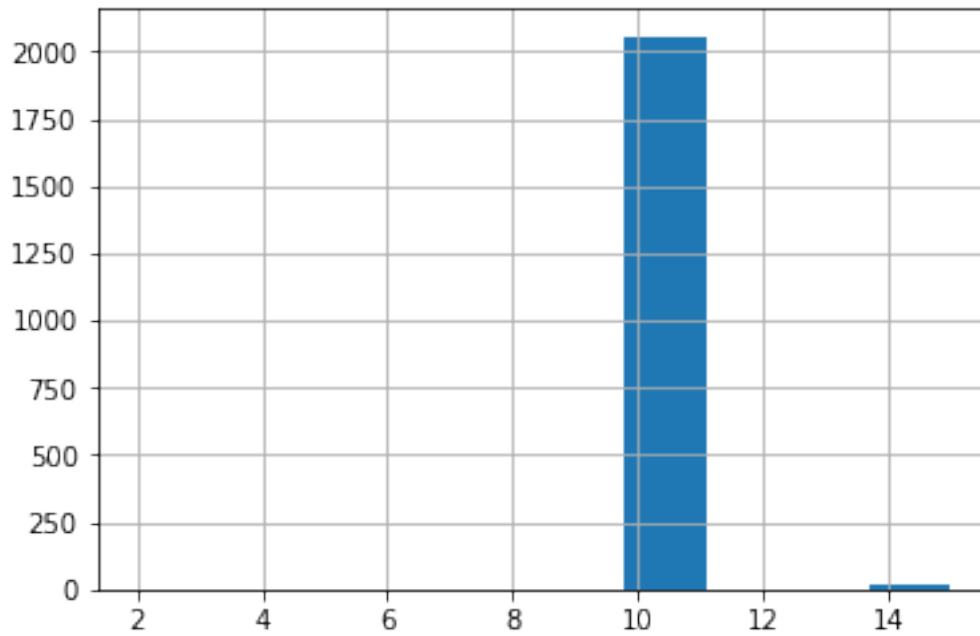
```
In [53]: df.hist(figsize=(10,10))
```

```
Out[53]: array([[<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aae191710>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadd224e0>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadc4d55c0>],
[<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadc8e5c0>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadc4a5c0>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadc4a5f8>],
[<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadc2bf60>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadb5f60>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadb9bf60>],
[<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadb837b8>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadb64cc0>,
<matplotlib.axes._subplots.AxesSubplot object at 0x7f1aadb21e80>]], dtype=object)
```



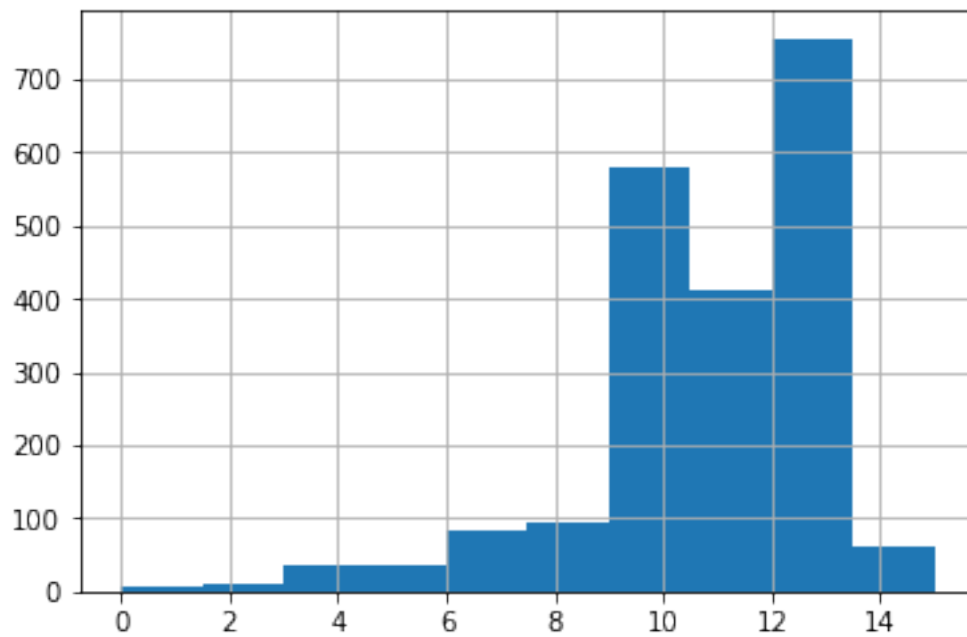
```
In [54]: df.rating_denominator.hist()
```

```
Out[54]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aadae4940>
```



```
In [55]: df.rating_numerator.hist()
```

```
Out[55]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aad975748>
```



# 1 Insights

- 1) Almost all the dog ratings denominators values are between 10~11
- 2) The dog ratings denominators vary on different values but mostly from the range of 9~13 values
- 3) Some people give over rating values for their dogs which should be rounded as we did here.
- 4) Either some people don't name their dogs or they see it's not necessary to mention.
- 5) There is a positive relation between retweet count and favorite count most probably people who favorite a tweet also retweet it