We rate dogs data quality issues

Archive

1. Re-tweets are included in the dataset
2. Replies are included in the dataset
3. Dog classification names are “NONE” should be nan
4. Some names are missing in the name column
5. Numerator and denominator are off – check this

Predictions

1. Check Data types and correct if necessary

Tidiness

1. Drop unnecessary tables
2. Fix column headings to make everything more readable to a human
3. Lower case everything.