The WeRateDogs Twitter archive contains basic tweet data for all 5000+ of their tweets. Project had three data source one is the image prediction file in TSV format which was downloaded from the online url provided in the project instructions. The tweet data was available in The Enhanced Twitter Archive in the CS, which was provided. The third dataset was a json file downloaded using the instruction in the project (please note that I had issue with getting the developer account and I am not a big user of social media platforms specially Twitter, so I have used the csv file as provided).

The idea for this data wrangle is to show 2 Tidiness issues and 8 data quality issues. In the end some useful insights were derived along with the corresponding visualization for purpose of this project. Of course there can be more visualizations, cleanup and insights.