Quality Issues	Tidiness Issues
Twitter_archive	
1- Delete columns that won't be used for analysis. **2- The timestamp has an incorrect datatype - is an object, should be DateTime.** **3- some of the gathered tweets are replies and should be removed.** **4- some of the gathered tweets are retweets.** **5- some dogs have more than one category assigned.** **6- Correct denominators other than 10.** **7- float ratings have been incorrectly read from the text of tweet.** **8- we have 639 expanded urls which contain more than one url address.**	**1- Dog classification (doggo, floofer, pupper or puppo) should be in one column.**
Image Predictions	
9- the dataset has 2075 entries, while twitter archive dataset has 2356 entries. **10- column names are confusing and do not give much information about the content.**	**2- dataset should be merged with the twitter archive dataset.**

11- dog breeds contain underscores, and have different case formatting.

12- only 2075 images have been classified as dog images for top prediction.

Twitter API Data

13- twitter archive dataset has 2356 entries, while twitter API data has 2354.

3- dataset should be merged with the twitter archive dataset.