

Wrangle_Report

By Dhuha Baqarish

25/7/2019

1) Gathering data:

In the data wrangling project I have gathered from tow different resources the first on was from a csv file and I used `pd.read_csv`. Then I download a file using request library and URL programmatically from the internet and store it in `image_prediction`.

2) Assessing data:

In this process of the data wrangling I assessed the both data set programmatically and visually to explore and understanding the structure and finding the issues related to the quality and tidiness of the both data set

3) Cleaning data:

In this process I deal with the issues found in the `twitter_archived` and `image_prediction` for example renaming the columns and dropping the repeated columns and fixing the data type the comining the tow data set into one and saving them in a CSV file.

Insights:

- 1- The most dog type tweeted is “pupper”
- 2- The image prediction table stores the result of a classification of dog breeds through a neural network. And the most predicted dog is golden retriever