# Quality Match

Priyadharshini Ravi

# Annotators

1. Gather insights about the annotators:

a. How many annotators did contribute to the dataset?

Ans : 22 annotators

# Time

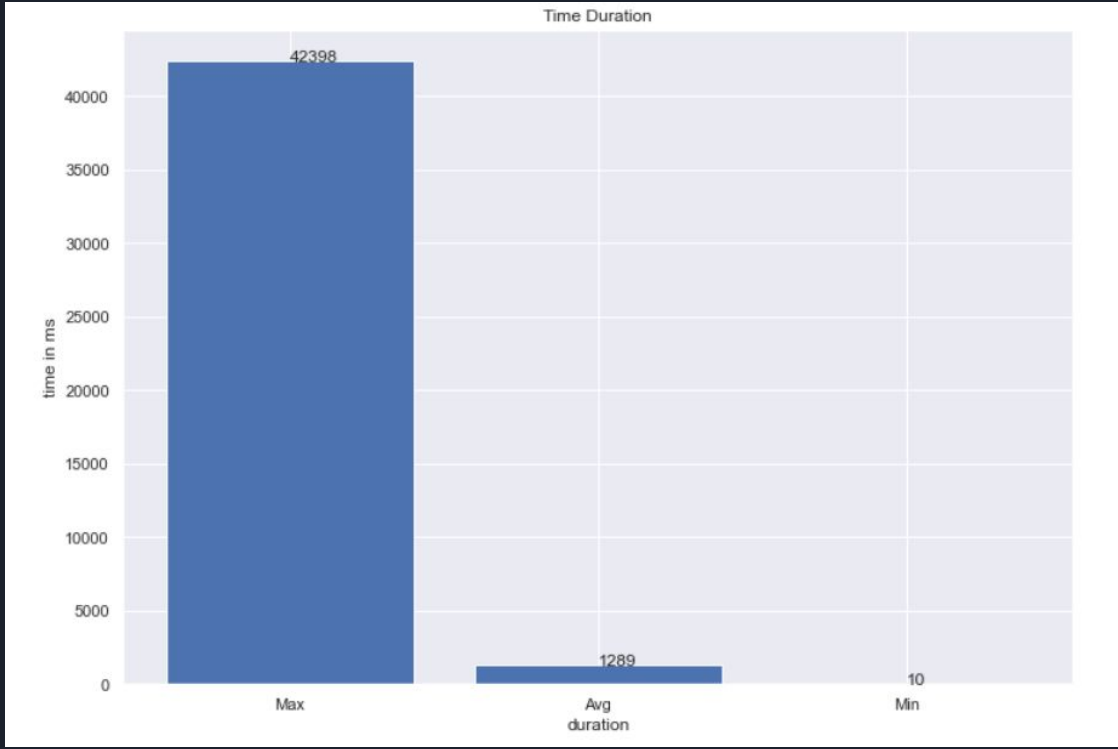
b. What are the average, min and max annotation times (durations)? Feel free to add visual representations here such as graphs if you like.

Maximum time : 42398ms

Minimum time : 10ms (ignoring negative and zero)

Average time : 1289ms

# Time


Time Duration
42398
1289
10
40000
35000
30000
25000
20000
15000
10000
5000
0
time in ms
Max
Avg
Min
duration

# Task

c. Did all annotators produce the same amount of results, or are there differences?

Ans: Since each annotators was assigned different no of task, the task completed is also different.

annotator_01 : 1280,annotator_02 : 7596,annotator_03 : 630,annotator_04 : 6421,

annotator_05 : 3475,annotator_06 : 5337,annotator_07 : 2175,annotator_08 : 6537

annotator_09 : 4860,annotator_10 : 315,annotator_11 : 6436,annotator_12 : 6210,annotator_13 : 7078

annotator_14 : 1725,annotator_15 : 6088,annotator_16 : 5061,annotator_17 : 3485,annotator_18 : 5170

annotator_19 : 170,annotator_20 : 6126,annotator_21 : 2950,annotator_22 : 1745

# Disagreement among annotators

d. Are there questions for which annotators highly disagree?

Considering, if there are equal number of ‘yes’ and ‘no’ answers from all annotators to a particular task id.  There are 68 task id which has disagreements among annotators

# Cant_solve or Corrupt

2. Besides picking yes or no the annotators had the chance to tell if the data were corrupted or if they for any reason were not able to solve the task. These are fields 'cant_solve' and 'corrupt_data' given in the task_output.

a. How often does each occur in the project and do you see a trend within the annotators that made use of these options?

the no of occurances of cant_solve : 17

the no of occurances of corrupt_data : 4

Annotator_08 and Annotator_18 have used both the option of cant_solve and corrupt_data

# Reference Set

3. Is the reference set balanced? Please demonstrate via numbers and visualizations.

Ans: the reference set is balanced, for class is_bicycle there are almost equal number of True and False.

df1['is_bicycle'].value_counts()

True : 4586

False : 4501

Name: is_bicycle, dtype: int64

# Good and Bad Annotators

4. Using the reference set, can you identify good and bad annotators? Please use statistics and visualizations. Feel free to get creative.

Good annotators : annotator_01,annotator_15,annotator_20,annotator_14,annotator_19

Bad annotators :annotator_10,annotator_08,annotator_07,annotator_06,annotator_03

# Good and Bad Annotators