

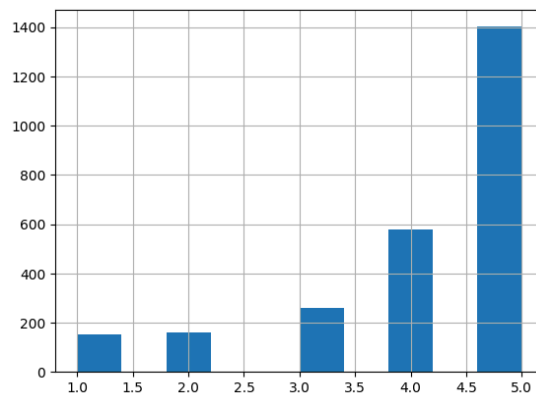
Charlottesville Restaurant Reviews Data Appendix

This appendix contains further information on the Charlottesville restaurant reviews data set (final_restaurant_df) used to conduct a sentiment analysis on the impact of seasonality and age of restaurants. The unit of observation in this data set is single review, with information on the data, rating, message, author reviews, restaurant, sentiment, and week day.

Variables: description of variables included in each review.

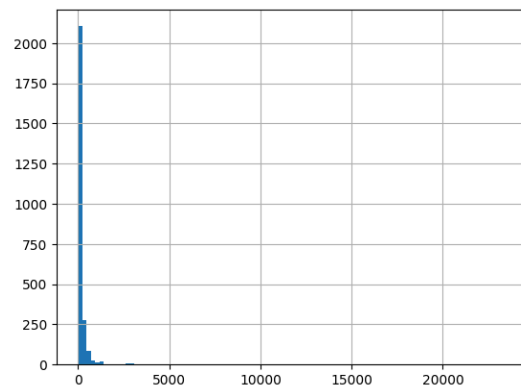
- Date: date the review was published on in yyyy-mm-dd format. Frequency table and distribution are non descriptive because there are 1556 unique values.
- Rating: score designated by reviewer between 0.0 and 5.0.

count	2554.000000
mean	4.140564
std	1.193161
min	1.000000
25%	4.000000
50%	5.000000
75%	5.000000
max	5.000000

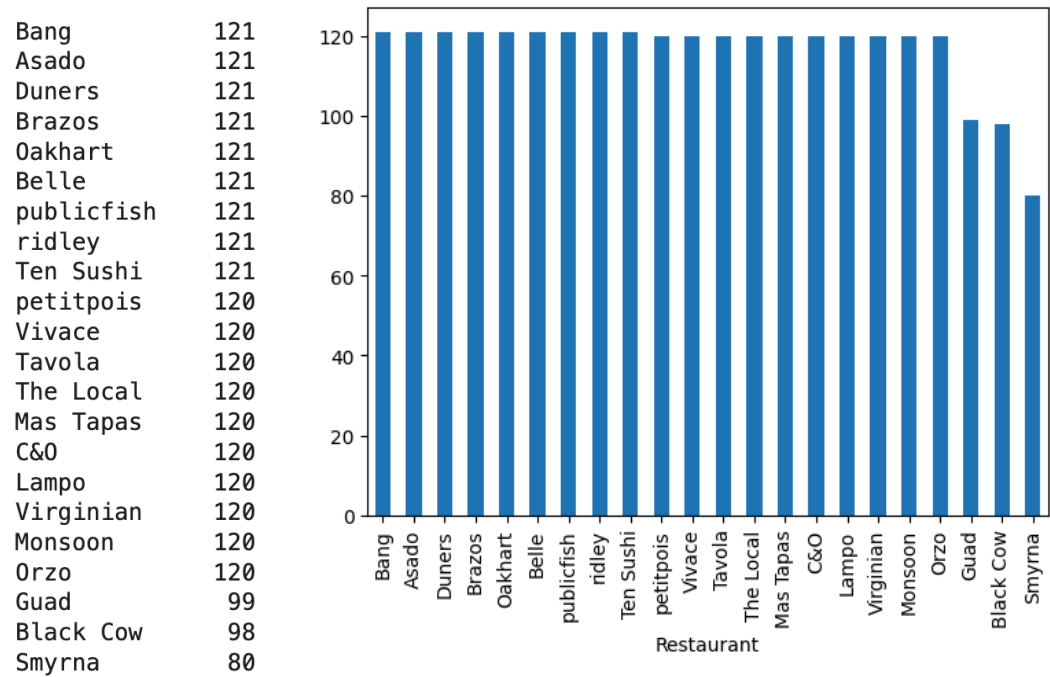


- Message: content of review in paragraph format. Frequency table and distribution are non descriptive because there are 2554 unique values.
- Author Reviews: number of reviews reviewer has published on Yelp.

count	2554.000000
mean	169.932263
std	736.260483
min	1.000000
25%	11.000000
50%	42.000000
75%	168.000000
max	23666.000000

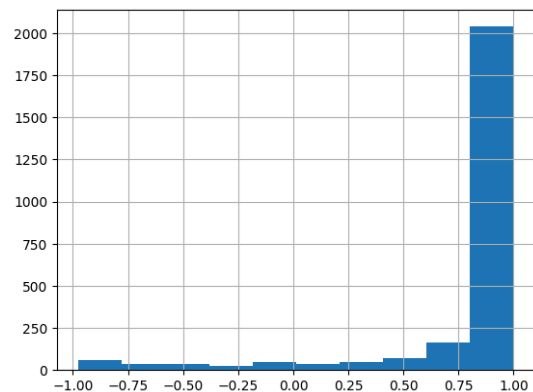


- Restaurant: name of restaurant.



- Sentiment: mood score for review between -1 and 1.

count	2566.000000
mean	0.781497
std	0.423942
min	-0.974200
25%	0.858400
50%	0.947800
75%	0.980975
max	0.999500



- Day of Week: day of the week the review was published on.

Sunday	501
Saturday	401
Monday	368
Tuesday	337
Friday	333
Thursday	324
Wednesday	290

