

The dataset comes from Kaggle, and contains restaurant reviews from 5/31/2016 to 5/25/2019. There are 10,000 observations, with each observation containing data for a single restaurant review. All variables begin as object data types.

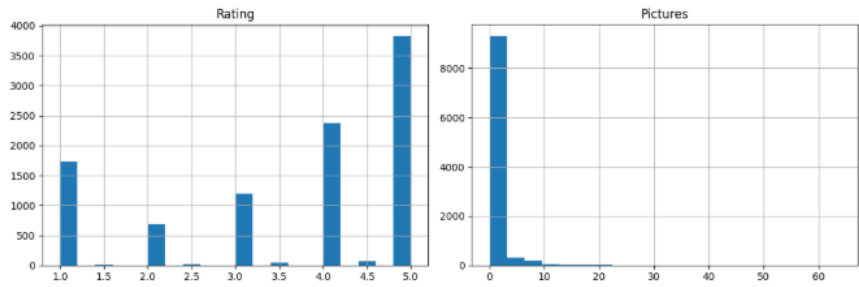
There were no outliers or transformations. Observations with missing review and rating values were dropped, as well as observations with non-numeric rating values. Columns titled "7514", "Time", and "Reviewer" were dropped. A variable called "sentiment" was constructed that creates "positive", "negative", or "neutral" sentiment values based on a review's rating. Rating was coerced to a numeric value, and Review was coerced to a string. The review text was also cleaned.

The original Restaurant reviews.csv contains:

Column	Description	Potential Responses	Missing responses
Restaurant	Restaurant name	Beyond Flavours, Flechazo, The Fisherman's Wharf	10000(0)
Reviewer	Reviewer Name	Dileep, Akash Thorat, Tony	10000(38)
Review	Review text	"Really healthy and tasty food. Quantity is a bit at the lower end for the given price, hence expensive."	10000(45)
Rating	Rating value	1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5	10000(38)
Metadata	Number of reviews that the reviewer has made and the number of followers the reviewer has	6 Reviews, 1 Follower	10000(38)
Time	Date and time of review	5/25/2019 3:54:00 PM	10000(38)
Pictures	Number of pictures included in the review	0, 6, 8	10000(0)

	Restaurant	Review	Rating	Metadata	Pictures	Sentiment
count	9954	9954	9954.000000	9954	9954.000000	9954
unique	100	9363	NaN	2477	NaN	3
top	Beyond Flavours	good	NaN	1 Review	NaN	Positive
freq	100	237	NaN	916	NaN	6315
mean	NaN	NaN	3.600261	NaN	0.751959	NaN
std	NaN	NaN	1.483573	NaN	2.575819	NaN
min	NaN	NaN	1.000000	NaN	0.000000	NaN
25%	NaN	NaN	3.000000	NaN	0.000000	NaN
50%	NaN	NaN	4.000000	NaN	0.000000	NaN
75%	NaN	NaN	5.000000	NaN	0.000000	NaN
max	NaN	NaN	5.000000	NaN	64.000000	NaN

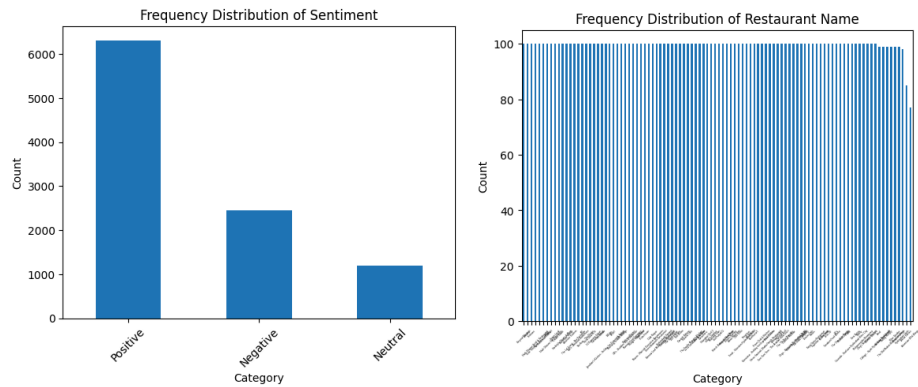
Summary Statistics:



Histograms:

Frequency table for Restaurant:		
Restaurant		
Beyond Flavours	0.010046	
Paradise	0.010046	
Flechazo	0.010046	
Shah Ghouse Hotel & Restaurant	0.010046	
Over The Moon Brew Company	0.010046	
...		
Olive Garden	0.009946	
Domino's Pizza	0.009946	
Hyderabad Chefs	0.009845	
Arena Eleven	0.008539	
American Wild Wings	0.007736	
Name: proportion, Length: 100, dtype: float64		
Frequency table for Sentiment:		
Sentiment		
Positive	0.634418	
Negative	0.245831	
Neutral	0.119751	

Frequency Tables:



Frequency Bar Charts: