

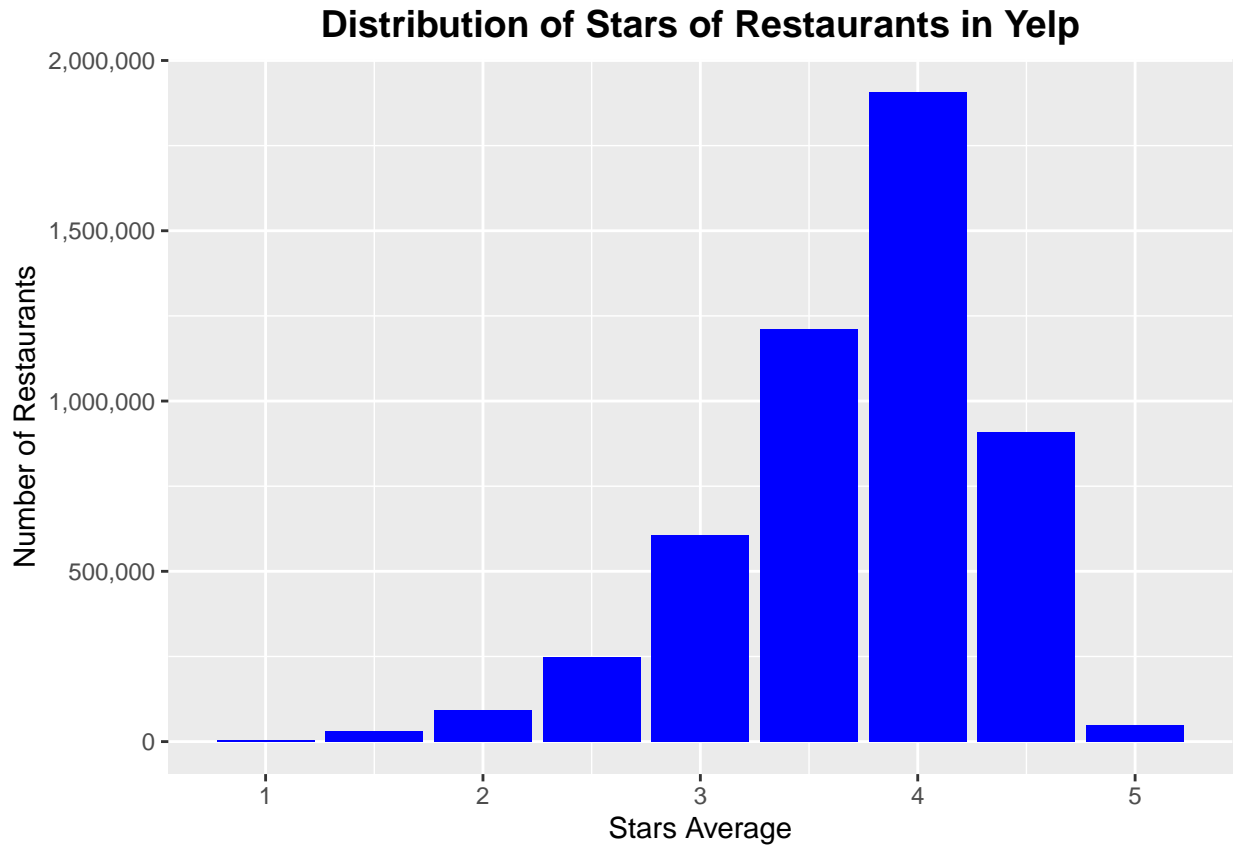
# Yelp Visualizations

Angelicaqj

3/27/2021

Data is collected from <https://www.yelp.com/dataset/documentation/main>. The yelp data is from 2019, and it is centered in metropolitan areas in part of North America. It provides 6 JSON files about business, review, user, check-in, tip, and photos, where each file is composed of the main JSON object. The dataset is approximately 10GB in size, it contains strings with punctuation (quotes and commas), and it is composed of more than one type of related data.

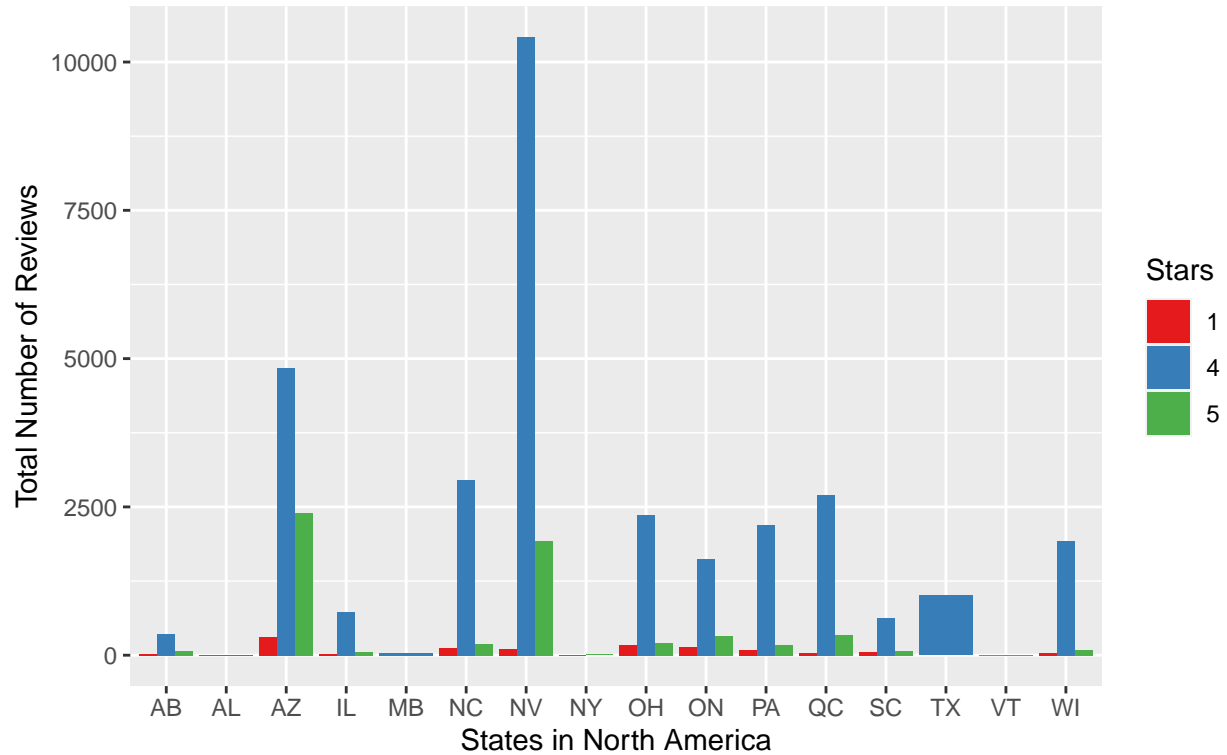
1. What is the average stars regarding all the restaurants on yelp in 2019?



The data visualization illustrates that the average stars is 4.0 regarding all the restaurants added to the yelp dataset in 2019. That means most of the users give a good review in general.

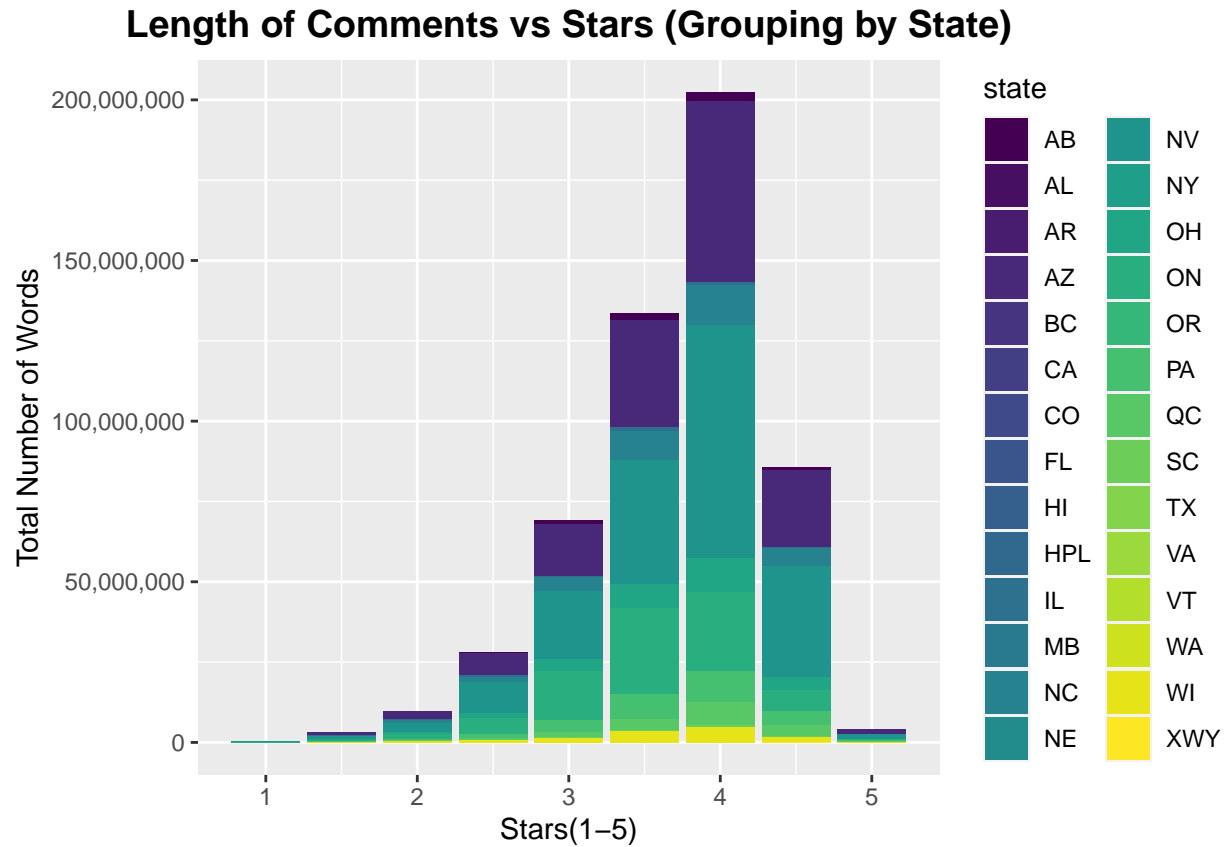
2. How the star average is comparing with the highest or lowest stars in any state in North America regarding the total number of reviews?

### Distribution of Comparing the Average Stars with Lowest/Highest Stars of all Restaurants in Yelp by States



The graph is showing that the state of Nevada has the most reviews on the star average (4.0), by obtaining more than 10,000 reviews of all restaurants in Yelp at that location. Also, Arizona State is the one with more reviews on the highest star (5.0) with approximately 2,500 reviews. In general, star average predominates in all the states that are representing the Yelp dataset.

3. How the length of the comments affects the rating stars category on yelp?



This visualization is confirming that the length of the comments is one of the reasons that users are giving star average (4.0) on yelp. It is reflected on the text or the comments they leave on the reviews.

On the other hand, the state as a categorical variable, it can give a scale related to the stars and the total number of words in all restaurants. It shows again that Nevada and Arizona states have the most length of comments by stars.