



# Analyzing Distribution of Zomato Restaurant Ratings and Prices Across the World

Richard Chun, Tara Jawahar, Shravan Ramamurthy, Grant Xu  
Instructor: Matey Neykov (36-315)

## Background

This dataset compiles information from the Zomato API regarding different cuisines and restaurants in countries around the world. It is organized by restaurant, with each restaurant having a unique Restaurant ID for a total of 9550 restaurants, or 9550 rows, recorded. There are 20 variables for each restaurant including the Restaurant ID, 12 of which are categorical and 8 of which are continuous. The continuous variables include Latitude, Longitude, and Average Cost, with the remaining categorical variables

Variables of interest that we can further analyze include Country Code, Cuisine, Price Range, Rating, Latitude and Longitude, and Number of Votes. Country Code and Cuisine can give us information about the type of food and where in the world the restaurant is located. The Price Range variable also helps us differentiate the types of restaurants and compare within these categories as well as compare across categories. An interesting note is that the Currency variable is not normalized to a standard currency (e.g. USD), and instead uses the local currency. Furthermore, we can differentiate between Aggregate Ratings and Votes, as Votes influence Aggregate Ratings. For instance, we would prefer a Restaurant that has 4.5 stars with 1000 votes than a restaurant with 5 stars and 3 votes.

## Exploratory Data Analysis

Frequency of Food Cuisines

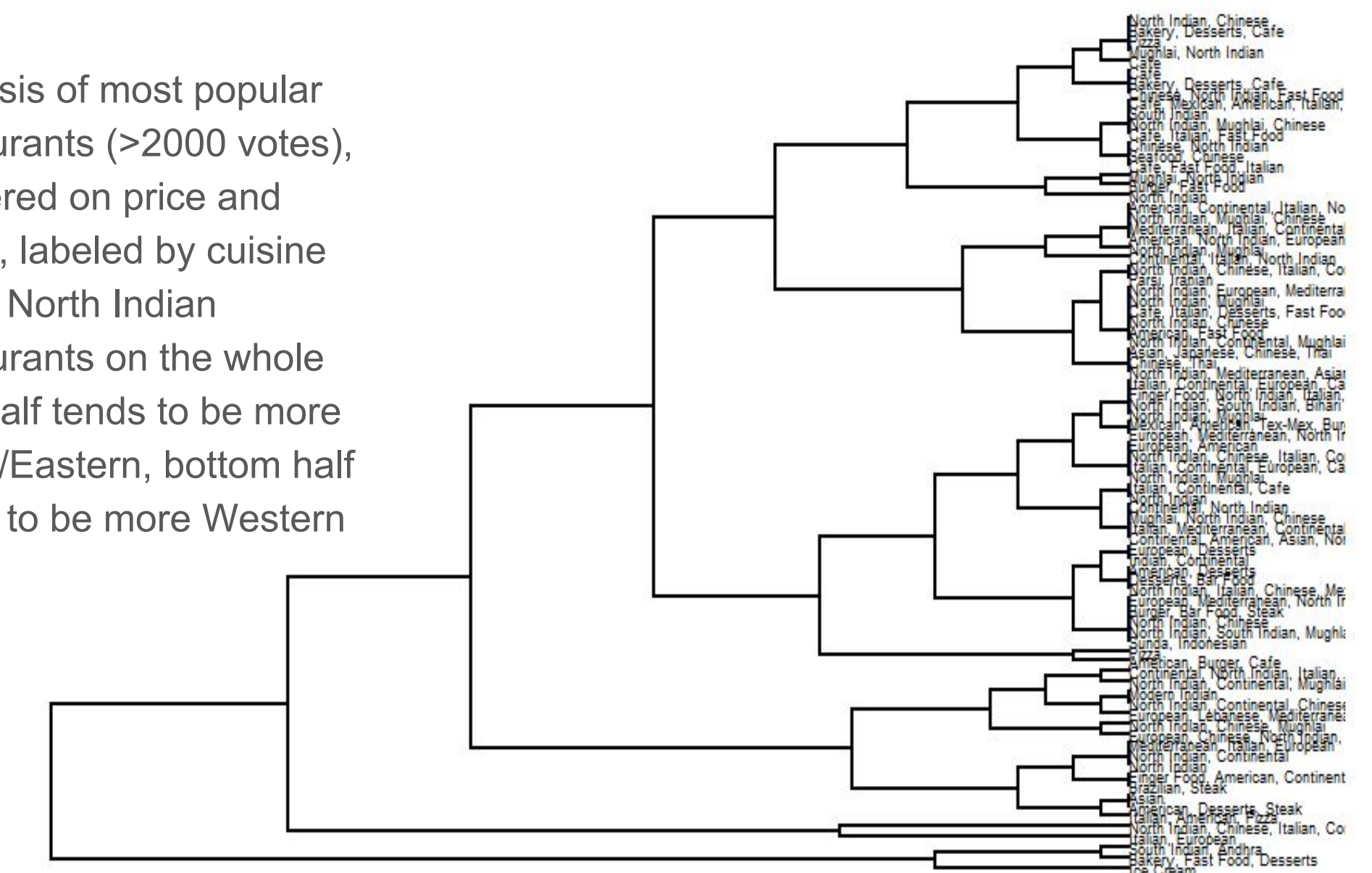


- Cuisines are varied, with the most frequent cuisine being North Indian by a significant margin
- Price Range and Rating do not seem to be independent because of extreme residual values



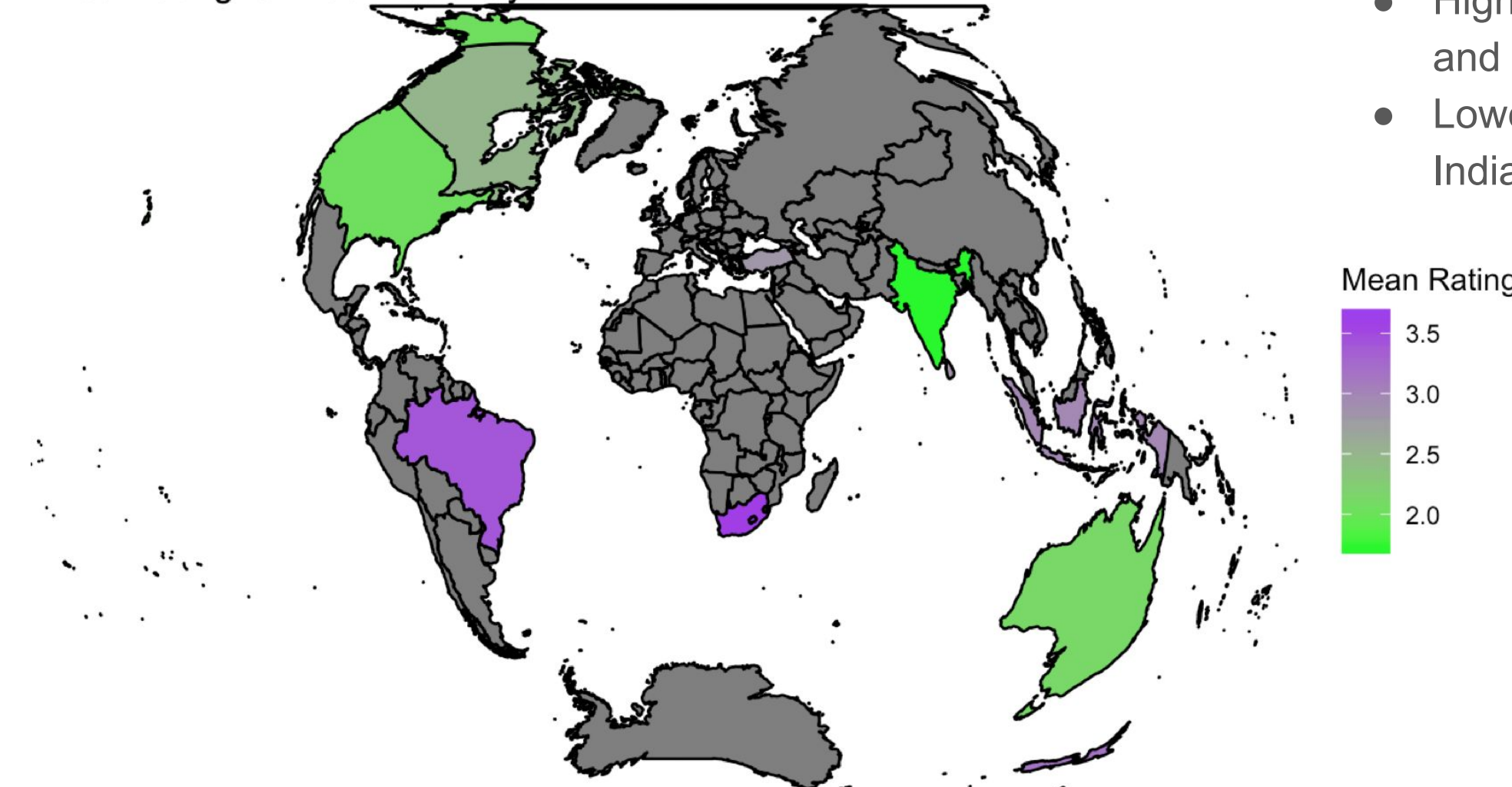
Cuisines of Popular Restaurants  
Clustered by Price and Rating

- Analysis of most popular restaurants (>2000 votes), clustered on price and rating, labeled by cuisine
- Many North Indian restaurants on the whole
- Top half tends to be more Asian/Eastern, bottom half tends to be more Western



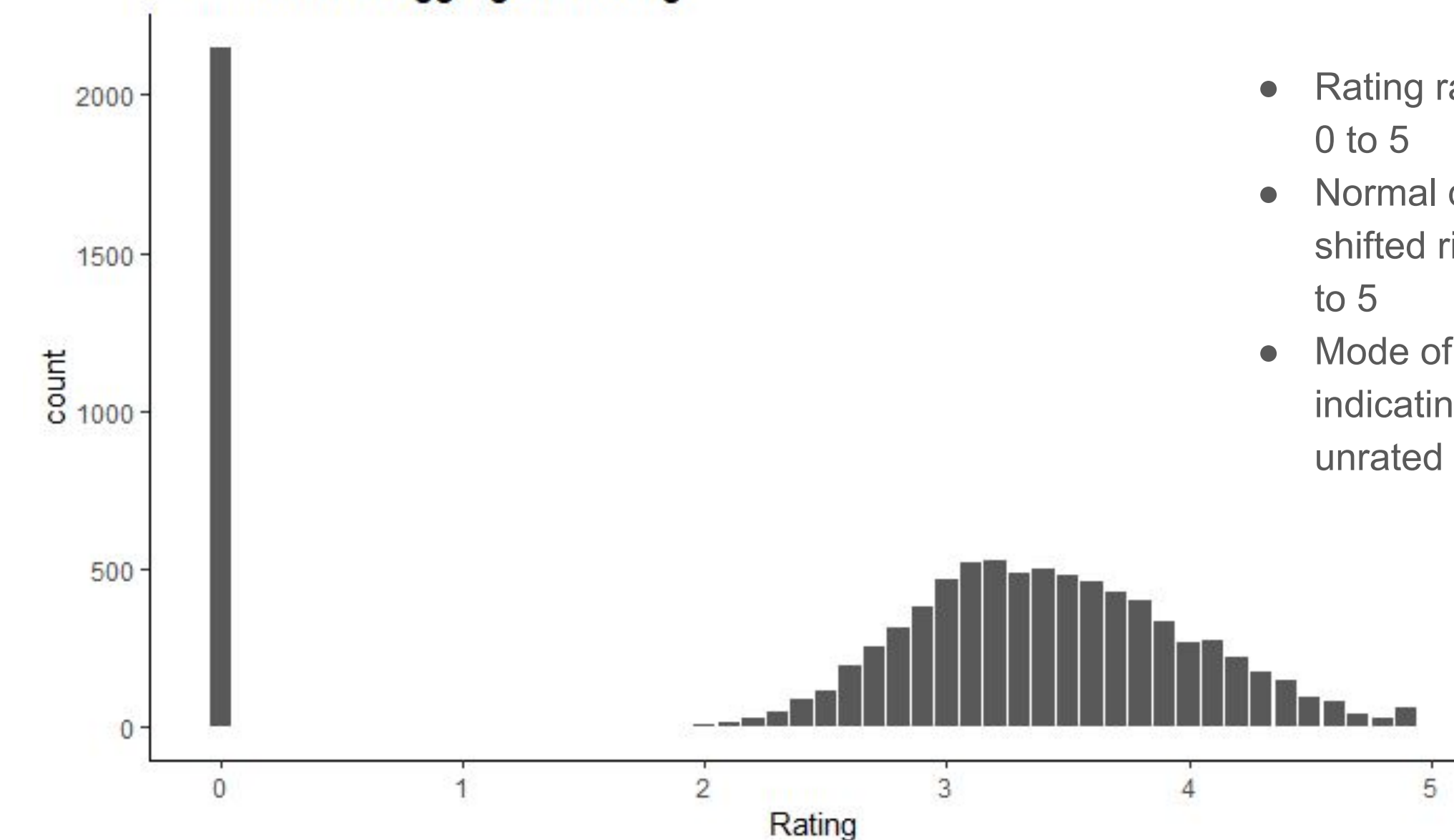
## Overview of World Rating

Price Rating for Each Country



- Prices ranges from 0 to 4
- Highest ratings are in Brazil and South Africa
- Lowest ratings in the US, India, and Australia

Distribution of Aggregate Rating



- Rating ranges from 0 to 5
- Normal distribution shifted right from 2 to 5
- Mode of ratings is 0, indicating many unrated restaurants.

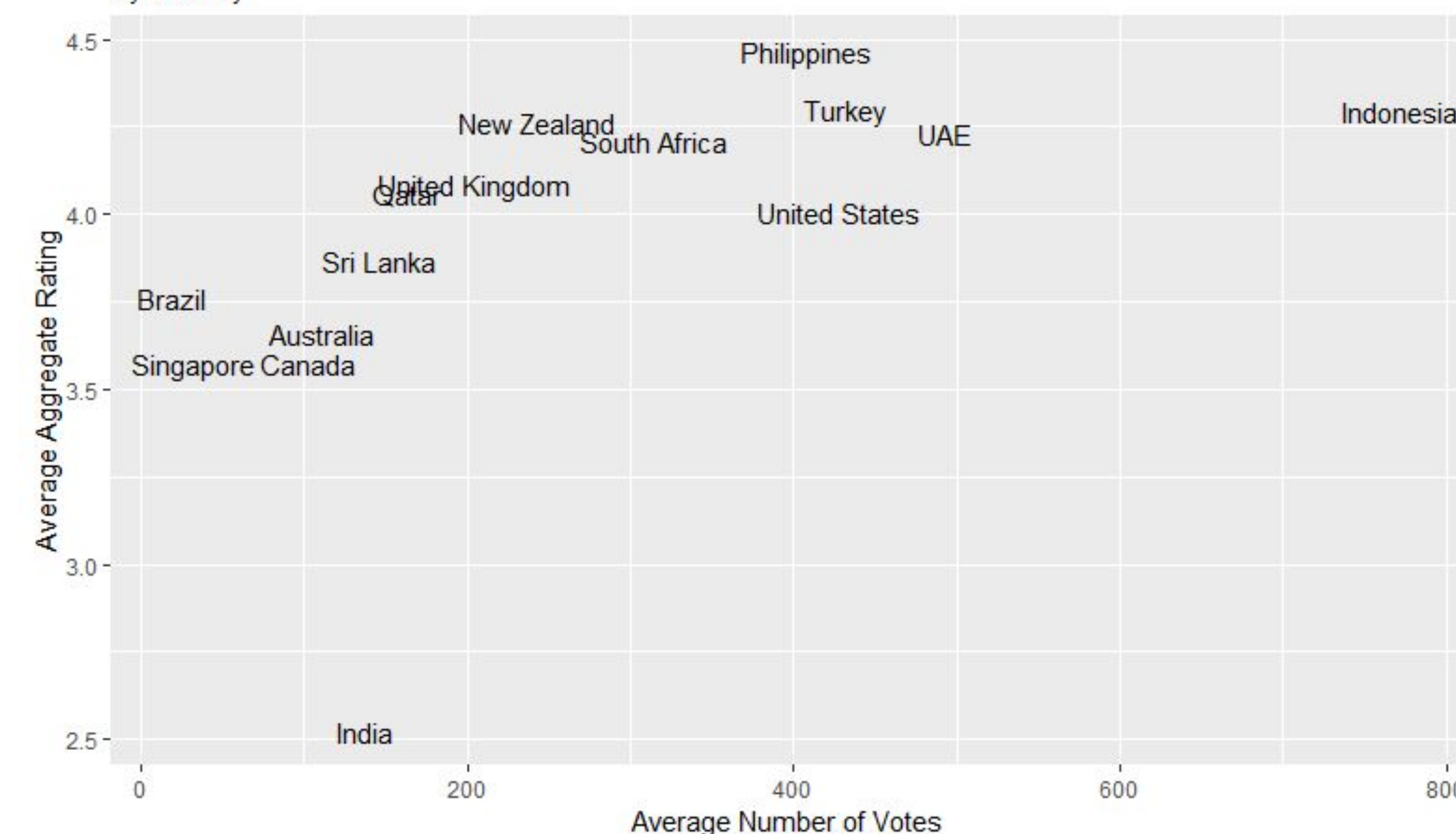
## Effect of Other Factors on Restaurant Ratings

Availability of Online Delivery and Table Booking  
By Aggregate Rating



- Online Delivery appears to be correlated with Rating, while Reservations is not correlated with Rating
- Outlier of 1.8 rating

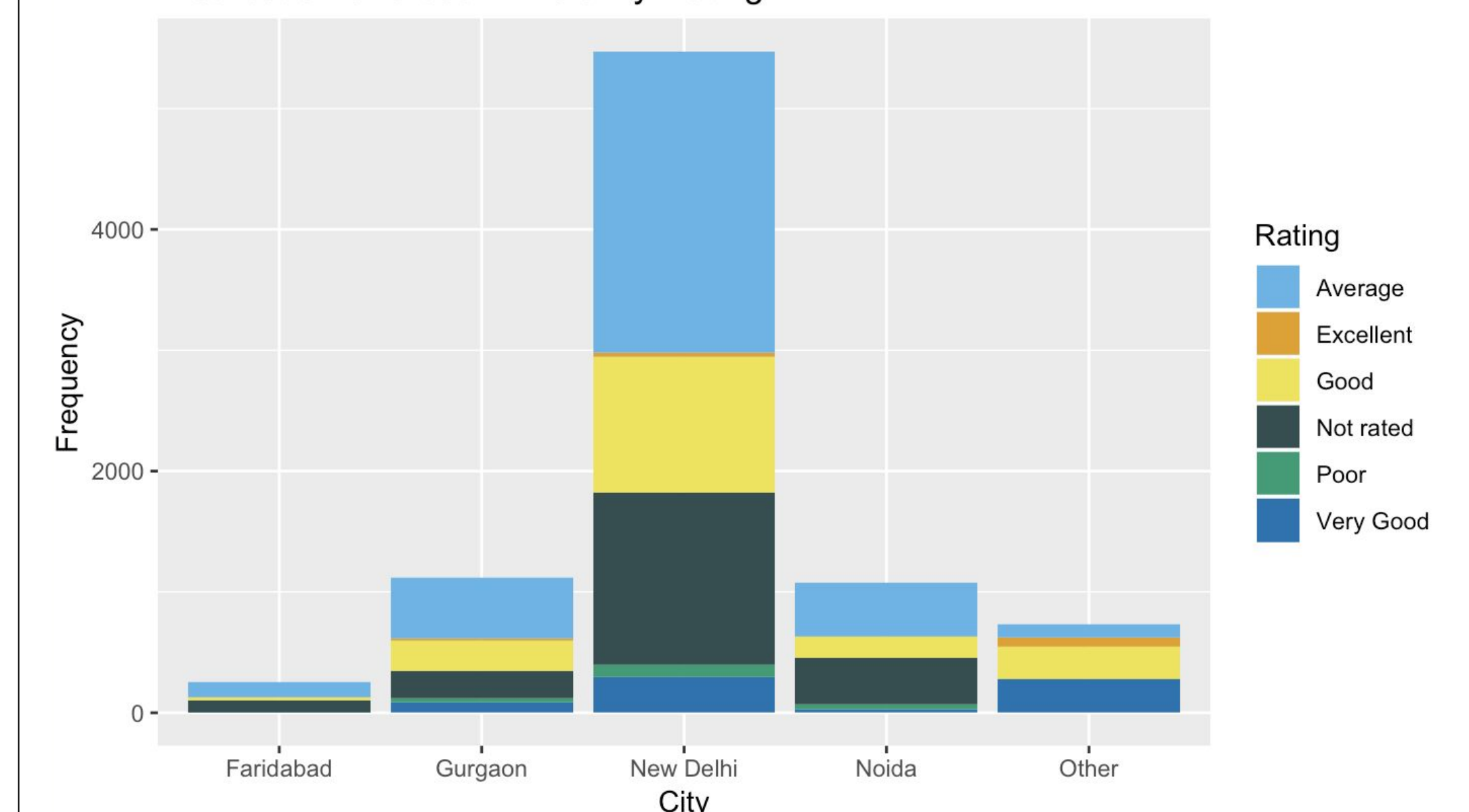
Average Number of Votes and Average Aggregate Rating  
By Country



- Shows usage of Zomato and average rating by country
- Outliers are India, with low rating, and Indonesia, with high rating and number of votes

## Focusing on Restaurants in India

Distribution of Cities in India By Rating



- There appears to be very few Excellent ratings, with many Average, Good, and NA ratings.
- New Delhi has the highest count, and the highest proportion of average ratings.