Task 1: Data Download and Descriptive Analysis

Q 1: Which types of restaurants are the most popular in each city?

Define and describe what “popularity” means.

Solution: In my opinion, a popular restaurant is the one which satisfies the following conditions:

- If the restaurant review rating > average review rating of all the restaurants in the city
- No of check-ins of the restaurant > average number of check-ins of all restaurants in the city

The following are the most popular restaurants from my findings:

Pittsburg:

American Restaurants and Restaurants serving Sandwiches

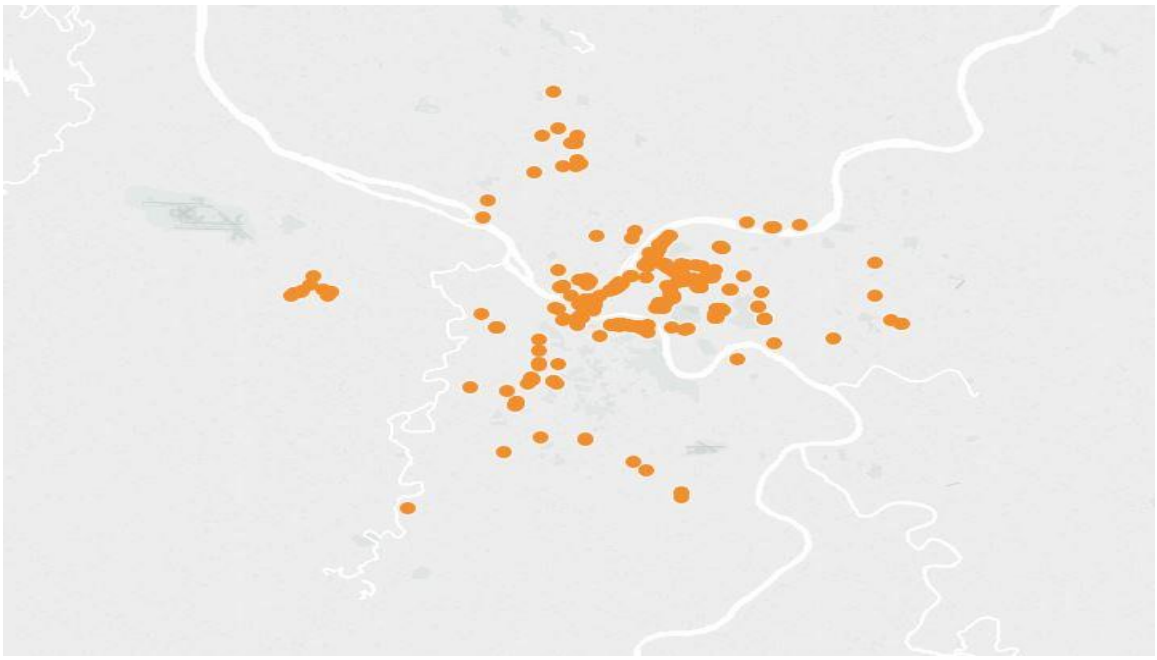
Charlotte:

American(New), Sandwiches and restaurants serving breakfast and brunch

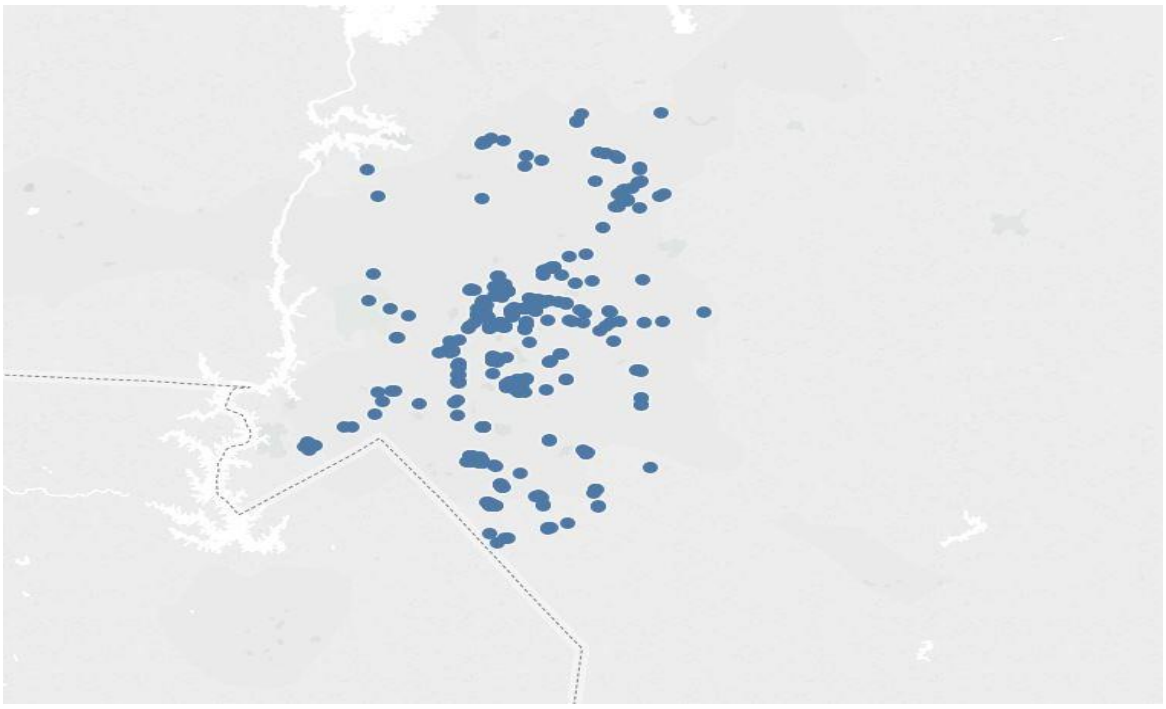
Q 2: Popularity of each city, visualize distribution on the map

Solution: The following are the visualizations of popular restaurants of each city on their respective city maps.

Pittsburg:



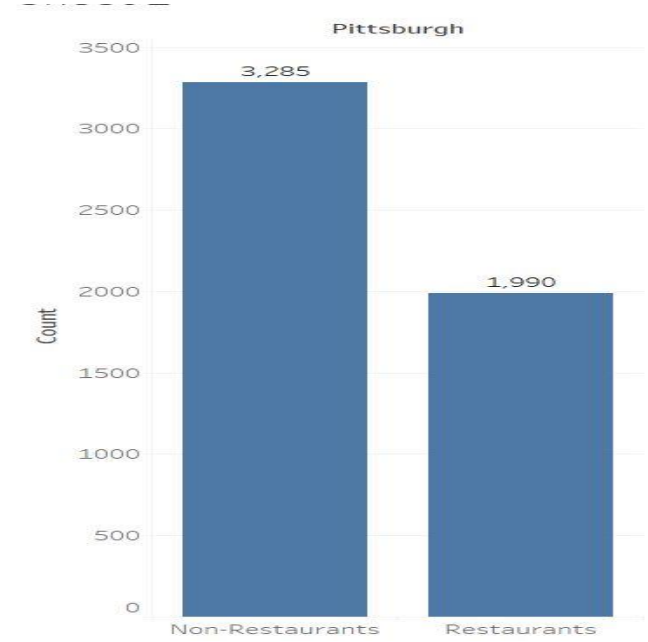
Charlotte:



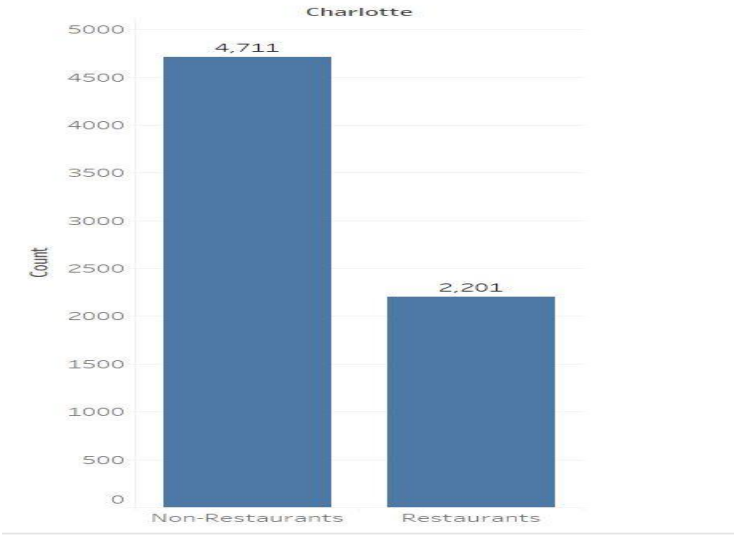
Q 3: One more interesting finding through descriptive analysis

Solution: In every city, there exists a good number of businesses. My focus here is to come up with a bar graph which depicts the count of restaurants and non-restaurants.

Pittsburg:



Charlotte:



Task 2: Text Processing & LDA (Latent Dirichlet Allocation)

Q 1: What words are most frequently used to describe Chinese restaurants?

Solution: Frequently used words to describe Chinese restaurants are place, chicken, food and so on.

Q 2: What are the major themes/topics in the reviews of Chinese restaurants?

Solution:

topic #0: 0.025*”die” + 0.024"und” + 0.023*”da” + 0.014*”nicht”+ 0.012*”der”

topic #1: 0.017*”chicken” + 0.016*”good” + 0.013*”food” + 0.011*”rice” + 0.011*”order”

topic #2: 0.016*”place” + 0.025*”food” + 0.012*”s” + 0.011*”order” + 0.011*”time”

topic #3: 0.009*”bar” + 0.016*”sushi” + 0.013*”great” + 0.011*”buffet” + 0.008*”drink” ..