Location Decision of a Museum Service Firm

Problem Description:

Wave, Inc is a Museum service firm in Seattle, Washington, specializing in exhibit design, display supplies, exhibition services, and restoration. For the past 15 years, Wave has grown from a small service firm offering services in exhibition design to a full-blown museum services firm with an extensive service provision. They've also grown in the number of employees, revenue, and service offering. Wave has been able to differentiate themselves from their competitors and is now the leading museum service firm in Seattle. While their services are sought after by various Museums across the United States, Wave has no businesses in any other state or city apart from Seattle, Washington.
In the last Management meeting, the firm decided to expand its business to 20 more cities in the US. *The management is interested in finding their target cities: cities that are more likely to lead to higher revenues for the firm. These cities should have a significant number of high rating museums. Their most preferred cities should also have a relatively higher population, with a considerable number of ratings.*

In this project, we will analyze different cities in the US and cluster the cities based on the number of museums in the city, the average rating of the museum in each city, the average number of ratings received by museums in the city, the average number of likes received by each museum in the city, and the average number of tips received by museums in each city. This will help Wave to determine the cities with a large number of highly-rated museums to help it to make an optimal location choice decision.

Data

We will use data from different sources. First, we will scrap the city name and state of the US cities with at least 100,000 people from Wikipedia.
Next, we will search for the top rating museums in each city with Foursquare.com API. Finally, we will do a Foursquare API venue search to find the rating, number of rating (ratingCount), number of likes (likesCount), and number of tips(tipsCount) of each museum.

For each city, we will average the data to find the number of high ranking museums, the average rating of the museum in each city, the average number of ratings received by museums in the city, the average number of likes received by each museum in the city, and the average number of tips received by museums in each city.