Battle of Neighborhoods–Tokyo

May 1, 2020

1 Introduction of Problem and Background

Today tourism is one of the most important parts of the economy. Tourists always want to travel to as many places as possible. On the other hand, there are too many amazing and exciting countries or cities to visit. Due to the limit of time or money, we have to make choices. Even after the travel destinations are determined, usually, there are still much more spots worth to visiting than tourists could visit.

Then the information about the destinations would be very helpful for travels to make a good travel plan. For example, by clustering places in a city into several clusters, travelers can choose more dissimilar places to visit. This is also valuable for the travel agencies, since it would help to attract customers.

In this capstone project, I plan to choose Tokyo as the travel destination and choose the top 9 sights suggested by Google and use the resource of Fouresquare to get the top venues around these sights and then make a travel plan based on the analysis of the data.

2 Data Description

Here I choose Tokyo (Japan) as the destination to visit. First, I search "Toky+touris+places" on Google. Then by clicking "top sights in Tokyo", I get the Top 9 Sight in Tokyo. By scrapping the website, I got the coordinates about the 9 sights in Table 1

Table 1: **Top 9 Sights in Tokyo, Japan**

ID	name	latitude	longitude
1	Tokyo Skytree	35.7101	139.811
2	Senso-ji	35.7148	139.797
3	Tokyo Tower	35.6586	139.745
4	Meiji Jingu	35.6764	139.699
5	Tokyo Disneyland	35.6329	139.88
6	Odaiba	35.6206	139.781
7	Ueno Park	35.7155	139.774
8	Imperial Palace	35.6852	139.753
9	Shinjuku Gyoen National Garden	35.6852	139.71

Next, by using the resource of Foursquare, I will get data about the venues around each sight. Then by getting the top 10 common venues in these places and clustering the top 9 sights into 4 clusters, a travel plan will be made.