Eat, Pray and Love Across ContinentsTitle

Heba Mushtaha

Introduction

Traveling around the continents to the largest city in each continent. We choose the largest city in each continent as a start, and then we can evolve our dataset with more cities and data. The project shall compare neighbourhoods across these cities and shall further recommend which city to visit first according to further criteria that the user sets.

7/31/2019

Problem

- **Part 1:** If you decide to take a tour across continents' largest cities (largest city in each continent), which continent/city shall you start with according to the city's neighborhoods and venues?

-

- **Part 2:** If you decide to take a tour across continents' largest cities (largest city in each continent), which continent/city shall you start with according to your specific preferences (i.e. food, touristic destinations...etc.)?

7/31/2019

Analytic Approach

- **Part 1: Battle of Neighborhood** - Analyzing city's neighborhoods and their venues to get more insight about the city's neighborhoods clusters. We shall cluster the neighborhoods (equal clusters) in each city and compare and you choose the city of your preference according to the clustering results.

- **Part 2: Recommendation System** - Analyzing each city's neighborhoods venues and classifying these venues under three starting classifications (Eat, Pray, Love). You shall set your preference on a scale of 1 to 5 for these categories, and accordingly the recommender shall recommend the cities to visit in sequence.

Data

Working with six (6) cities at each continent, the cities within the scope of this project are:
- North America: Mexico City, Mexico.
- South America: Sao Paulo, Brazil.
- Asia: Tokyo, Japan.
- Australia: Sydney, Australia.
- Europe: Paris, France.
- Africa: Greater Cairo, Egypt.

Data Requirements

- Neighborhoods in each city (we used the central neighborhoods only due to size, processing and time limitations).

- Venues and venues categories in each neighborhood.

- User's preferences

Data Collection

- Neighborhoods in each city collected in csv

- Venues and venues categories in each neighborhood (we used FourSquare APIs).

- User's preferences (this is set on a scale of 1 to 5)

Data Understanding & Preparation

- For All: Cleaned the dataframes for redundancy and inaccurate data.

- For the Recommender System: we further cleaned the data and dropped venues categories whose count is less than 10

Methodology/Modeling

Clustering

- Divided the neighborhoods in each city into 5 clusters (with the same *limit* and *radius* passed to all), with venues and venues categories in each cluster.

- Add all the neighborhoods' clusters to one dataframe so that we have one dataframe consisting of all neighborhoods (with venues and venues categories in each city) in the 6 cities.

Classification

- Classified each venue category as either Eat/Pray/Love

- Add all the results of venues lying under Eat/Pray/Love in one dataset for comparison and using in the recommender system.

Recommendation System

- Using the output of the classification method above, we now set the user's preference for Eat/Pray/Love.

- Add the user's preferences to the user's profile; using the output of the classification method and content-based recommendation system the user gets recommendations for the cities to visit by order.
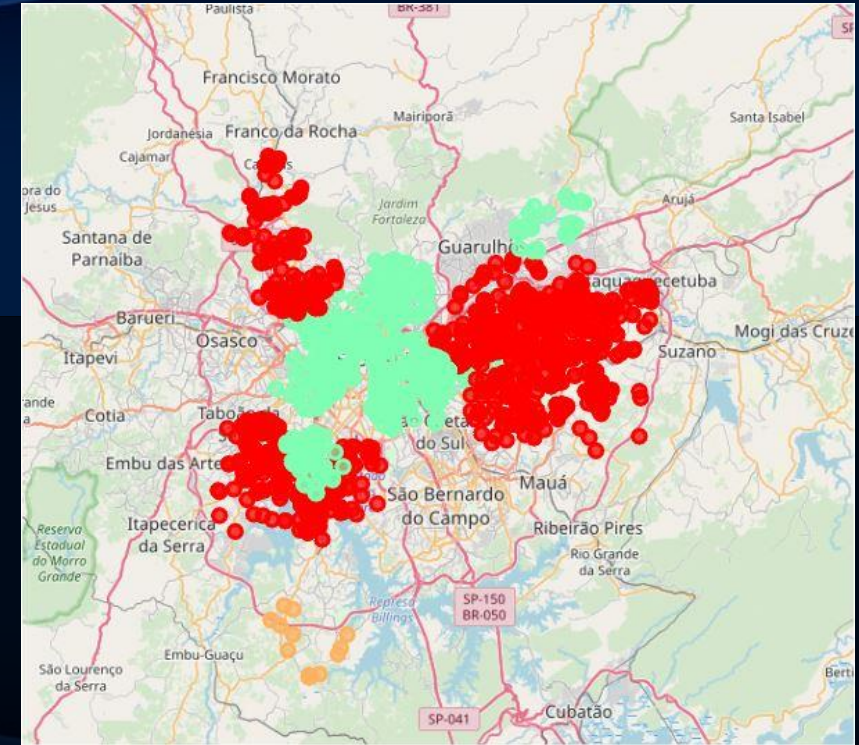
- We also repeated this exercise for many users.

7/31/2019

Results

Clustering

Analyzing 6 cities with 5 clusters each. We added the results to a dataframe so that we can compare clusters.
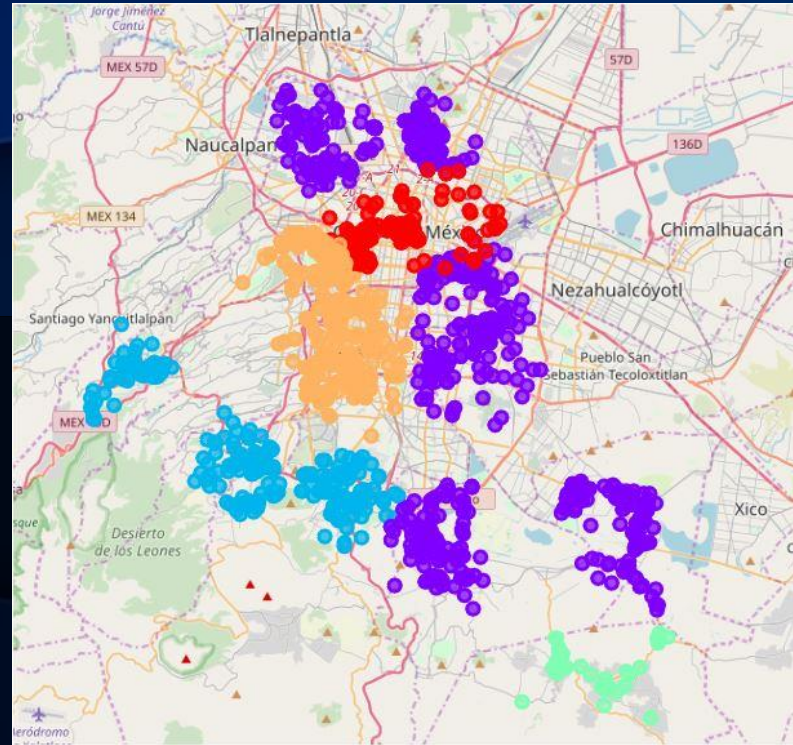
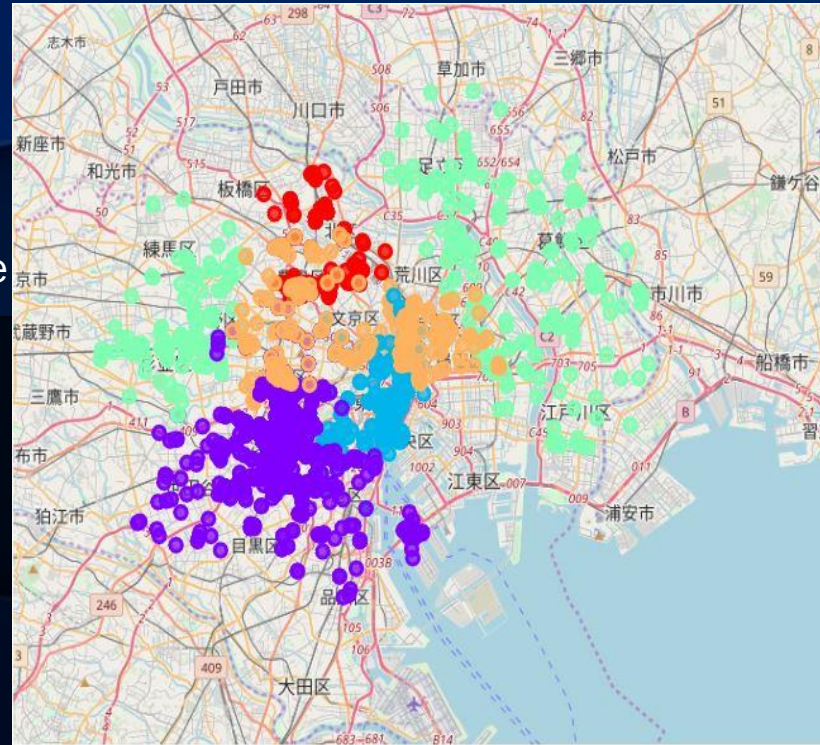North America: Mexico City, Mexico - Clusters
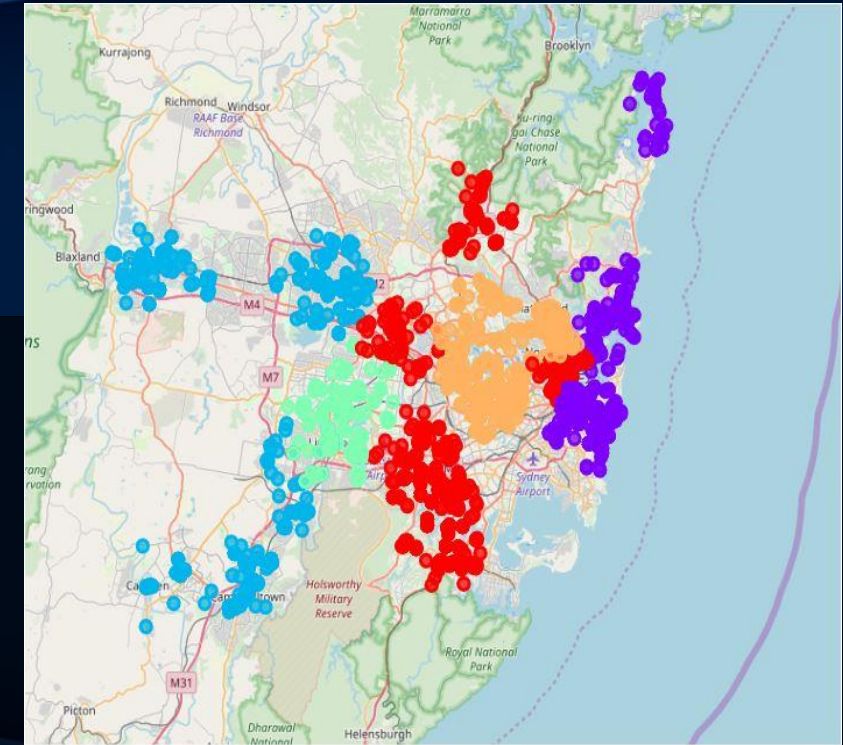

South America: Sao Paulo, Brazil- Clusters

Results

Clustering

Analyzing 6 cities with 5 clusters each. We added the results to a dataframe so that we can compare clusters.
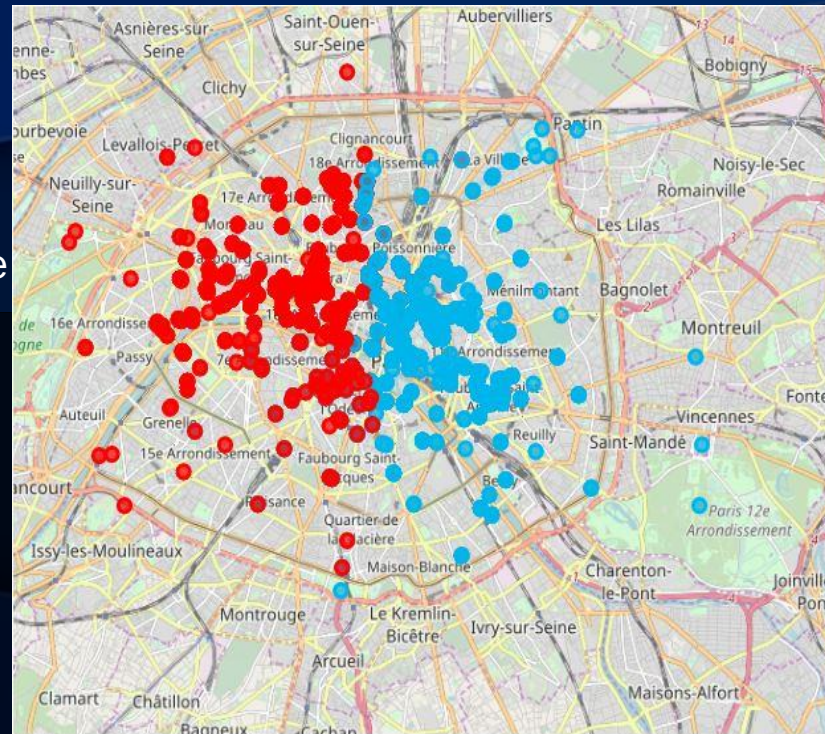

Asia: Tokyo, Japan - Clusters

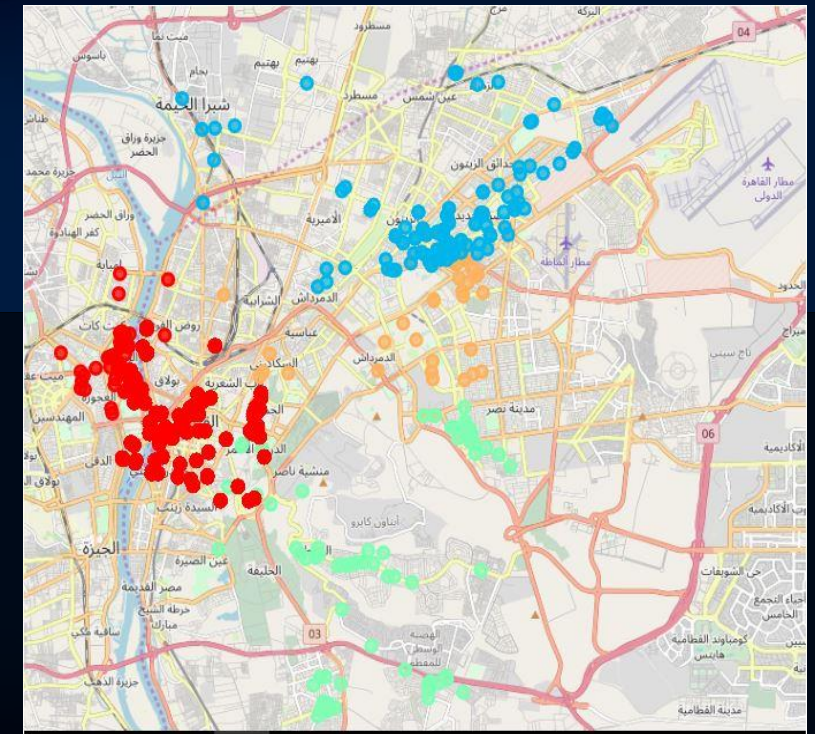Australia: Sydney, Australia Clusters

Results

Clustering

Analyzing 6 cities with 5 clusters each. We added the results to a dataframe so that we can compare clusters.

Europe: Paris, France- Clusters

Africa: Greater Cairo, Egypt Clusters
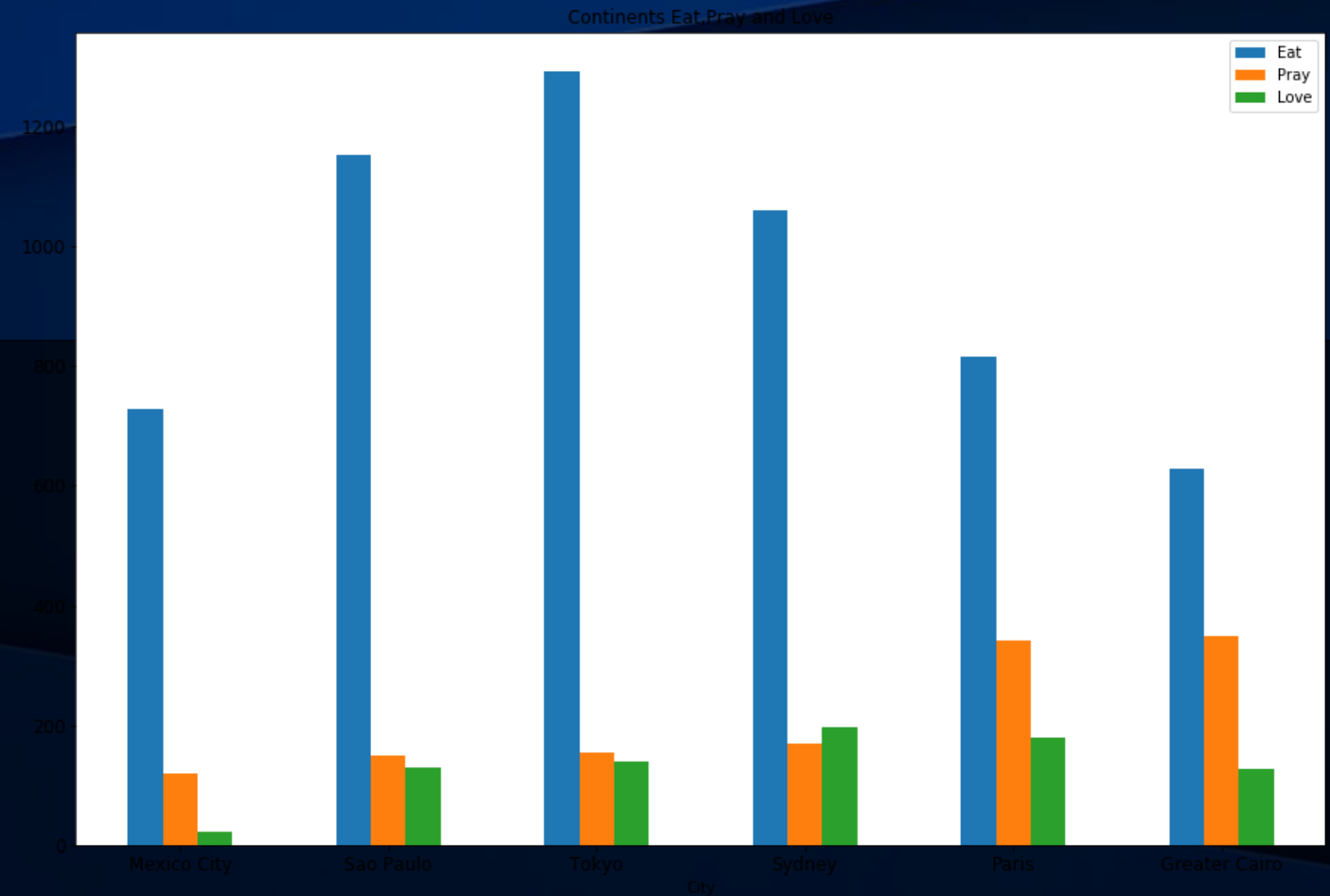
Results

Clustering

- Clusters in all neighborhoods in all cities are added to one dataframe

- Arranged the dataframe descending according to number of venues (my preference)

- Results for cities with largest number of venues were Sao Paulo then Greater Cairo.

- Resulting dataframe is displayed on the notebook (not here as there are 30 clusters – 6 cities,

  5 clusters each)

Results

Classification

From the results of classifying venues by Eat, Pray and Love, the following results are generated:

|	Continent Name	Country	City	Eat	Pray	Love
0	North America	Mexico	Mexico City	728	120	23
1	South America	Brazil	Sao Paulo	1153	148	130
2	Asia	Japan	Tokyo	1292	153	140
3	Australia	Australia	Sydney	1061	169	197
4	Europe	France	Paris	816	341	178
5	Africa	Egypt	Greater Cairo	629	348	126

Continents Eat,Pray and Love

Results

Recommendation System – Single User

From the results of the recommender based on user's preferences

	Continent Name	Country	City	Eat	Pray	Love	MyScore
2	Asia	Japan	Tokyo	1292	153	140	433.416667
4	Europe	France	Paris	816	341	178	405.416667
3	Australia	Australia	Sydney	1061	169	197	401.333333
1	South America	Brazil	Sao Paulo	1153	148	130	393.250000
5	Africa	Egypt	Greater Cairo	629	348	126	344.250000
0	North America	Mexico	Mexico City	728	120	23	239.666667

```
According to your user profile and your rating for Eat, Pray and Love, You can visit the following cities by the following order:
------------------------------------------------------------------------------------------------
2               Tokyo in Japan at Asia
4              Paris in France at Europe
3           Sydney in Australia at Australia
1        Sao Paulo in Brazil at South America
5          Greater Cairo in Egypt at Africa
0     Mexico City in Mexico at North America
```
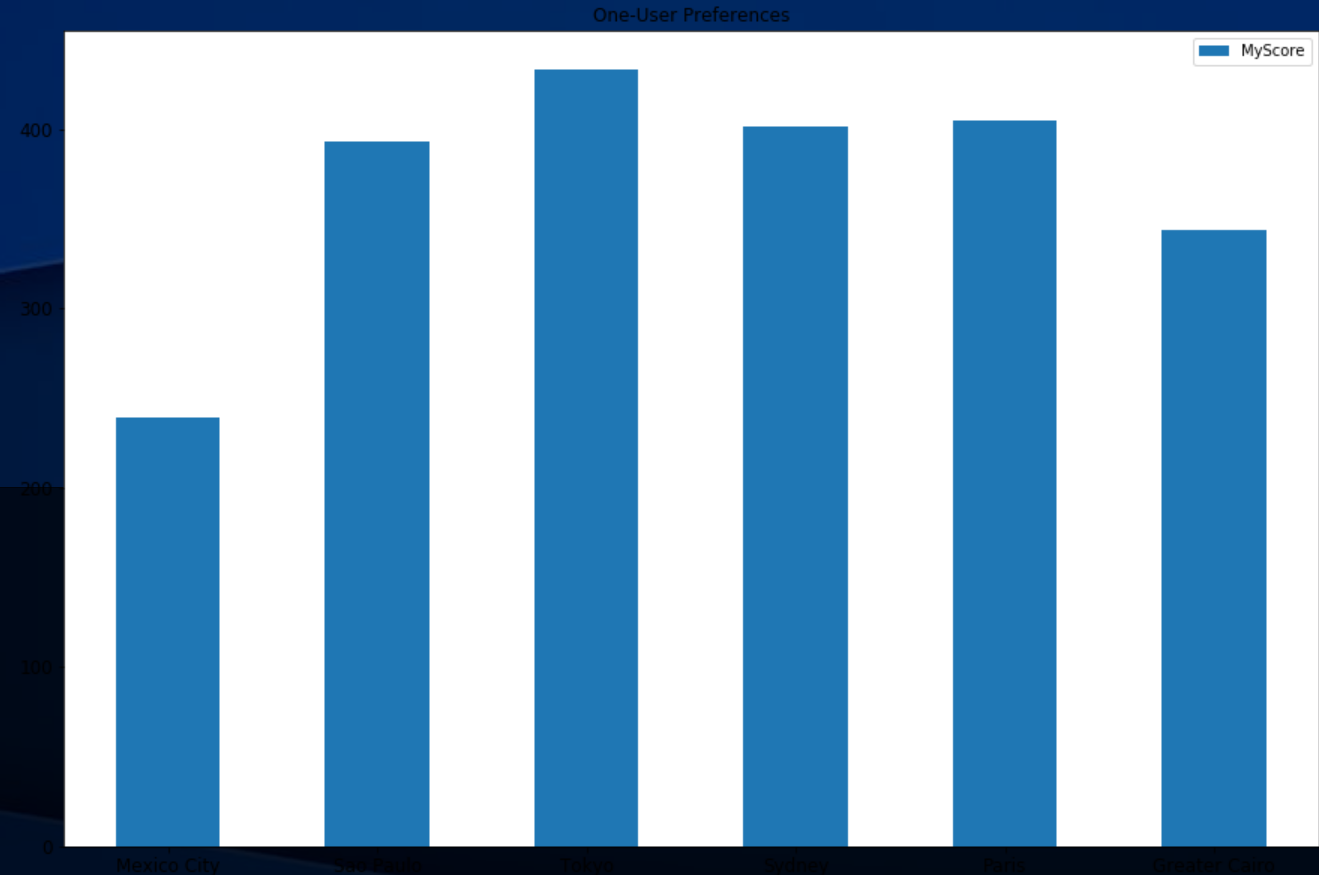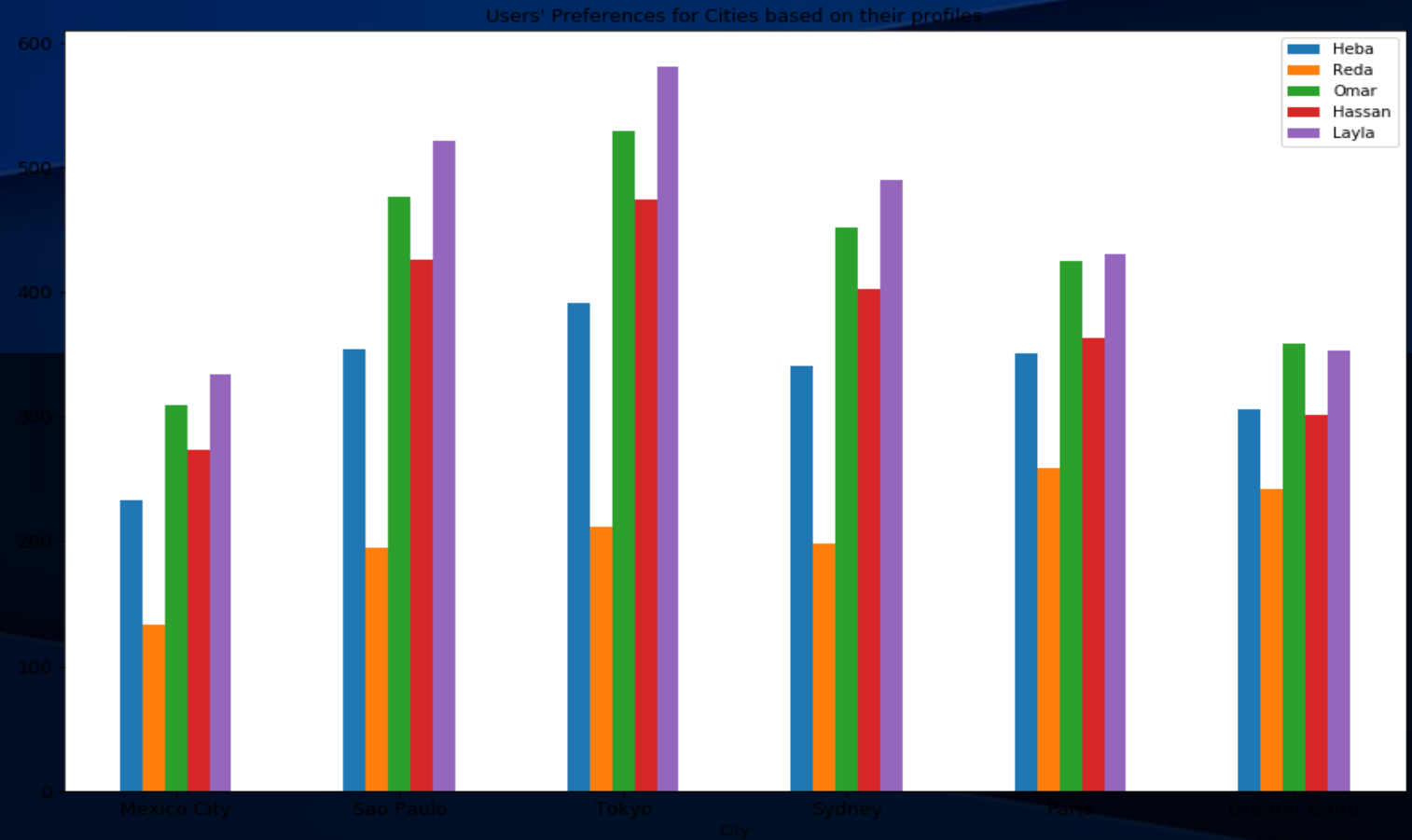
Results

Recommendation System – Multi-User

For Multi –User Recommendation System, the following diagram also shall be used by recommendation system to group people of similar preferences and consequently recommend other cities.

Users' Preferences for Cities based on their profiles

Conclusion

Battle of Neighborhood

Now You have analysis for each city's neighborhood's clusters. Choose which one suits you more!

Conclusion

Recommendation System

Set your preferences and we will RECOMMEND your route across the continents

7/31/2019

Discussions

You can compare the first city you chose from the clustering exercise and the first city recommended from the recommender system and you might be surprised!

Discussions

From the Battle of Neighborhoods, I choose the cluster with the highest # of venues and the city was Sao Paulo, Brazil then Greater Cairo!

My preferences for Eat, Pray and Love was 3,5,4 respectively – the city to start with from recommender was Tokyo then Paris!

7/31/2019

Thank you!