Izmir and Thessaloniki Similarity :
A Data Approach

Introduction

- "Many cities in different European countries are remarkably similar. Thessaloniki and Izmir are such towns, closely linked by history and culture" [1]

- "Thessaloniki known as Izmir's sibling" [2]

- "Izmir is similar to Thessaloniki; Thessaloniki to Izmir" [3]

- "The "brotherhood" between Karsiyaka and Kordelio, Menemen and Menemeni, bears the exchange." [4]

Data Description

- Finding parts of the cities (municipalities) in Wikipedia.

- Finding coordinates of municipalities with API request via geopy.geocode Python library

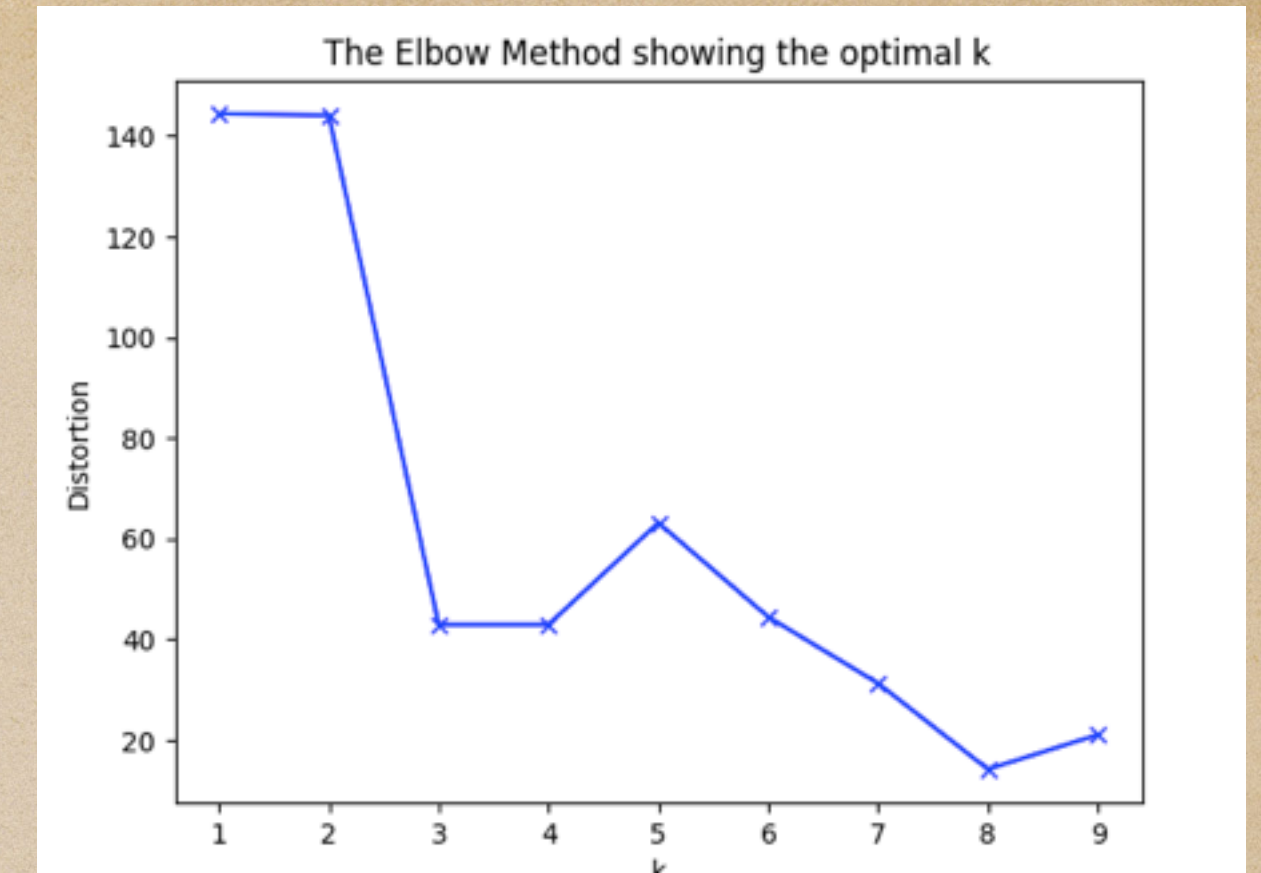- Finding the venues nearby the municipalities via Foursquare API.

Methodology

- Firstly, builded a dataframe which includes the mucipalities of the two cities and their coordinates. Totally, 44 municipalities were included in this study.

- Secondly, found nearby 12 venues for each municipality. There are 147 different types of venues for 44 municipalities.
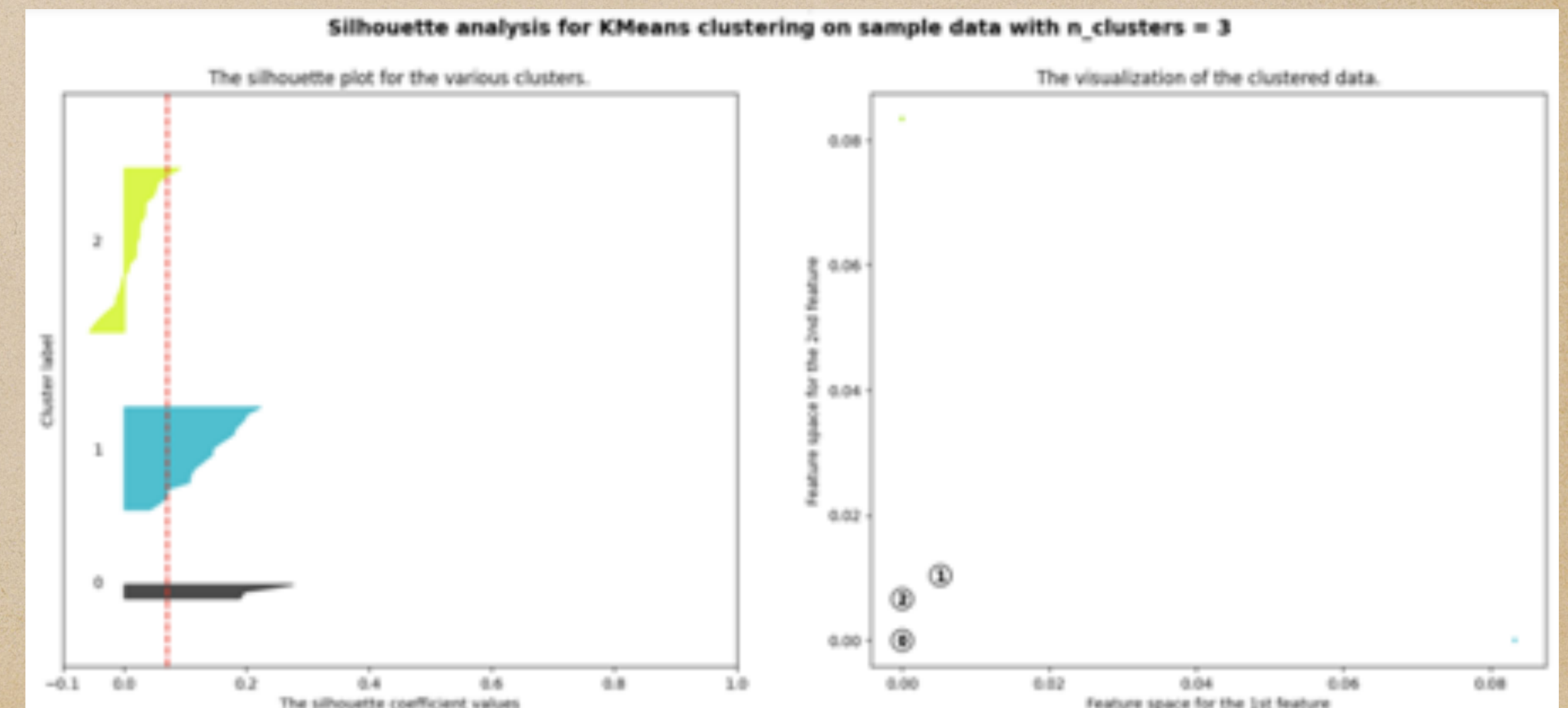
Methodology

The Elbow Method showing the optimal k

- Clustering is suitable for this study
  due to this study is a descriptive
  analysis. In this case, it is appropriate
  to use the k-means method since all
  the data we are working on is
  numerical. At this point, determining
  the number of k clusters is a critical
  problem. Elbow and Silhouette
  methods showed that creating 3
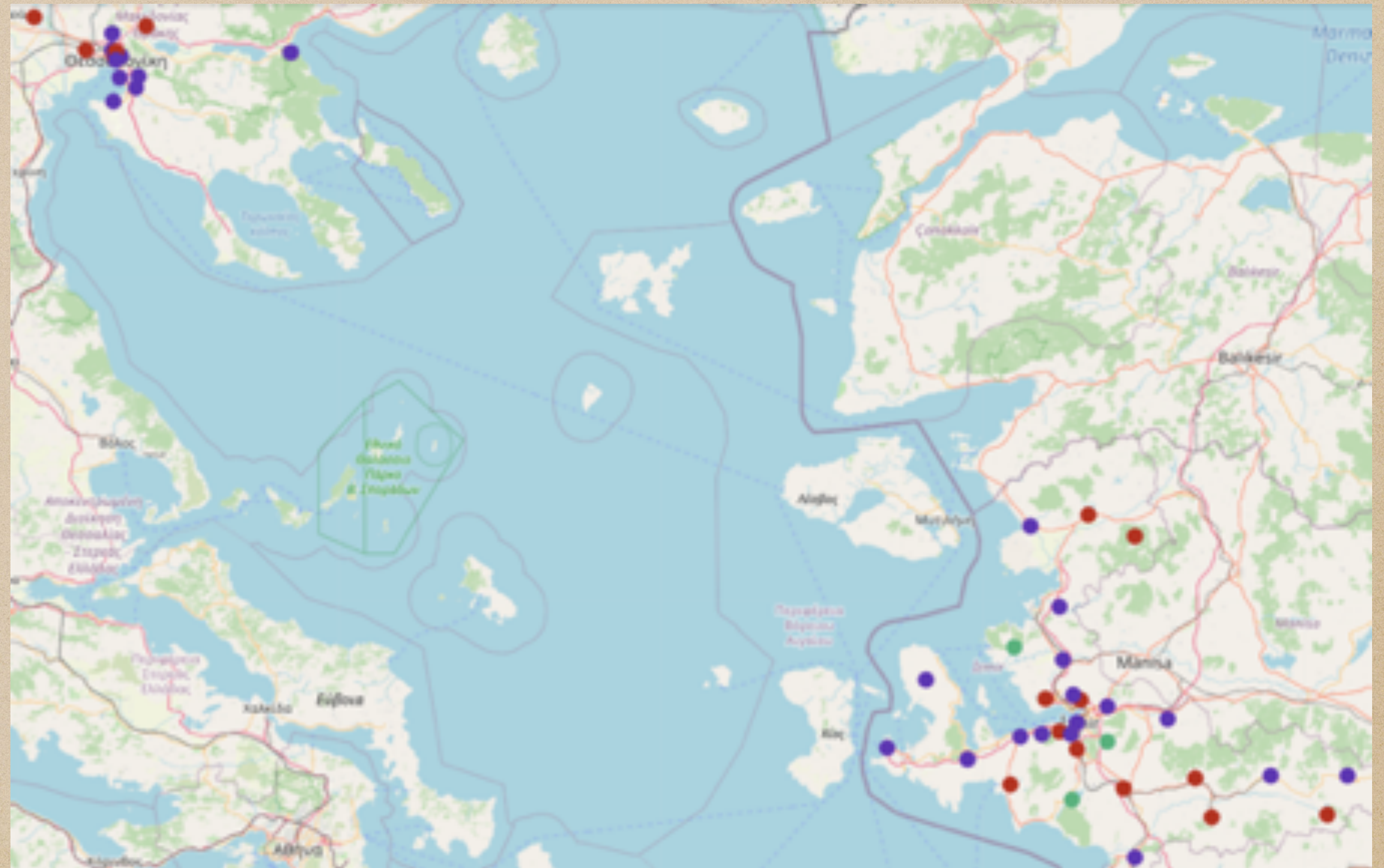 clusters is an optimal solution.

```
Automatically created module for IPython interactive environment
For n_clusters = 2 The average silhouette_score is : 0.05681855542631025
For n_clusters = 3 The average silhouette_score is : 0.07132255354444246
For n_clusters = 4 The average silhouette_score is : 0.06571251285844346
For n_clusters = 5 The average silhouette_score is : 0.06047416545871338
For n_clusters = 6 The average silhouette_score is : 0.05516778078102227
For n_clusters = 7 The average silhouette_score is : 0.04095650047489114
For n_clusters = 8 The average silhouette_score is : 0.04641864029249359
For n_clusters = 9 The average silhouette_score is : 0.05584052661265438
```

Silhouette analysis for KMeans clustering on sample data with n_clusters = 3

Results

- Clustered municipalities figured out with a folium map. The map shows that there are no seperated two cities. Only the first cluster has municipalities from Izmir but the other two clusters include municipalities from the two cities.

Results

- The municipalities of the first cluster are Buca, Foça and Menderes from Izmir. Breakfast spot, Farm and Steakhouse are very popular in these regions. Basketball Court, Butcher, Fishing Spot, Go Kart Track, Gozleme Place, Lake, Rock Climbing Spot venues are identifiers for these regions.

Results

- The second cluster contains 10 regions from Izmir such as Balçova, Bergama, Seferihisar and contains 5 regions from Thessaloniki such as Sindos and Lagkadas. Cafe, Bar and Gym center are very popular in these regions. Airport, Brewery, Bus Station, Camera Store, Dance Studio, Jewelry Store, Kids Store, Military Base, Monument / Landmark, Museum, Paper / Office Supplies Store, Pet Store, Pizza Place, Public Art, Rental Car Location, Shop & Service, Snack Place, Spa, Train Station, Turkish Coffeehouse, Wings Joint, Zoo venues are identifiers for these regions.

Results

- The third cluster contains 15 regions from Izmir such as Çeşme and Konak and it contains 9 regions from Thessaloniki such as Kalamaria and Thermi. Restaurant and Hotel are very popular in these regions. Auto Garage, Bakery, Bay, Beach, Beach Bar, Bed & Breakfast,Bistro, Boarding House, Bookstore, Borek Place, Botanical Garden, Burger Joint, Campground, Castle, Church, College Quad, Construction & Landscaping, Convenience Store, Country Dance Club, Cupcake Shop, Daycare, Deli / Bodega, Dive Bar, Donut Shop, Electronics Store, Fair, Flower Shop, Food Court, Food Truck, Frozen Yogurt Shop, Gastropub, Gym Pool, Health Food Store, Italian Restaurant, Kebab Restaurant, Lounge, Market, Mediterranean Restaurant,Mens Store, Meyhane, Meze Restaurant, Modern Greek Restaurant, Music Store, Other Great Outdoors, Paintball Field, Pastry Shop, Pedestrian Plaza, Pharmacy, Seafood Restaurant, Soccer Field, Stadium, Surf Spot, Sushi Restaurant, Söğüş Place, Tennis Court, Theater, Theme Restaurant, Vineyard, Waterfront, Wine Bar venues are identifiers for these regions.

Discussion

- This study shows that there are no completely seperate two cities. Many municipalities are similar. It shows that the two cities still have similar culture and lifestyle.Only the municipalities of the first cluster completely belong to Izmir.

- These municipalities far away the Aegean sea, maybe their culture and lifestyle effected by this issue.

- On the other hand, Menemen from Izmir and Menemeni(Ampelokipoi) from Thessaloniki are located in different clusters.

Conclusion

- a6. This approach can be usefull for considering the similarity of the different locations.

- On the other hand, Foursquare API can provide the data with venue count limit and/or radius.

- Foursquare can provide very few venue data for small places such as Karaburun and Stavros. If the scope is expanded, more detailed results can be obtained, but this study was not expanded enough to investigate the best result, as it aimed to present only one approach.

Referances

- [1] https://www.etranslation.gr/commonalities-across-borders-observations-thessaloniki-greece-izmir-turkey/

- [2] https://eksisozluk.com/selanik-vs-izmir--4345888

- [3] https://www.hurriyet.com.tr/yazarlar/deniz-sipahi/izmir-selanik-e-benzer-selanik-de-izmir-e-28146202

- [4] https://m.bianet.org/biamag/kent/216180-kentlerin-kardesligi-ve-selanik-izmir

Thanks...