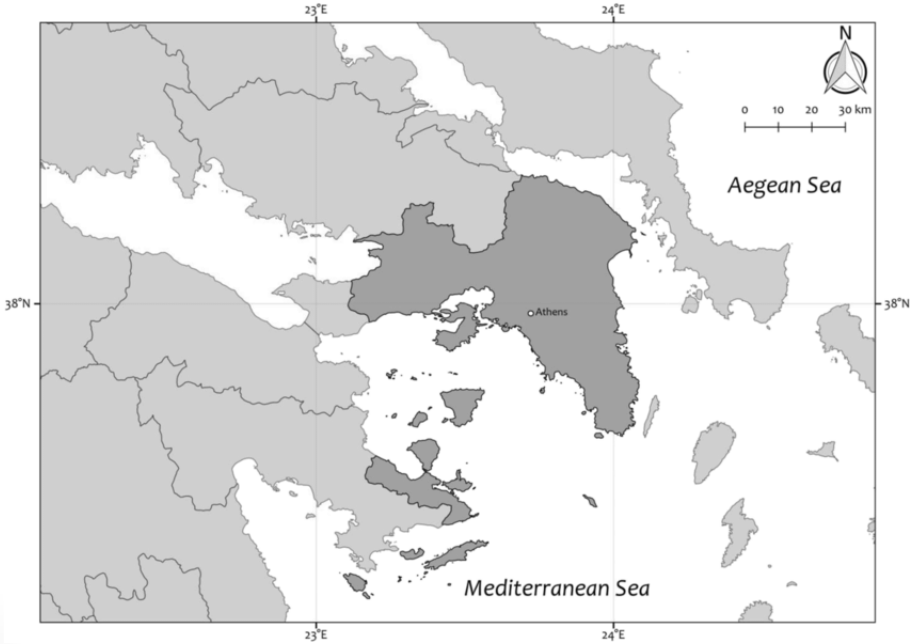
Applied Data Science Capstone

by IBM



Finding the best place to open a restaurant in Attika, Greece

Konstantinos Anathreptakis

June, 2021

Introduction

- **Background:** Explore Attica neighborhoods Latitude and Longitude data in order to find the best place to open a restaurant.
- **Problems:** Which locations restaurants are the most frequently venues. Which Neighborhoods exhibit the same characteristics based on venues.
- **Interest:** Businesses in the food industry as well advertising companies. Knowledge of the location with the most frequently restaurants venues.

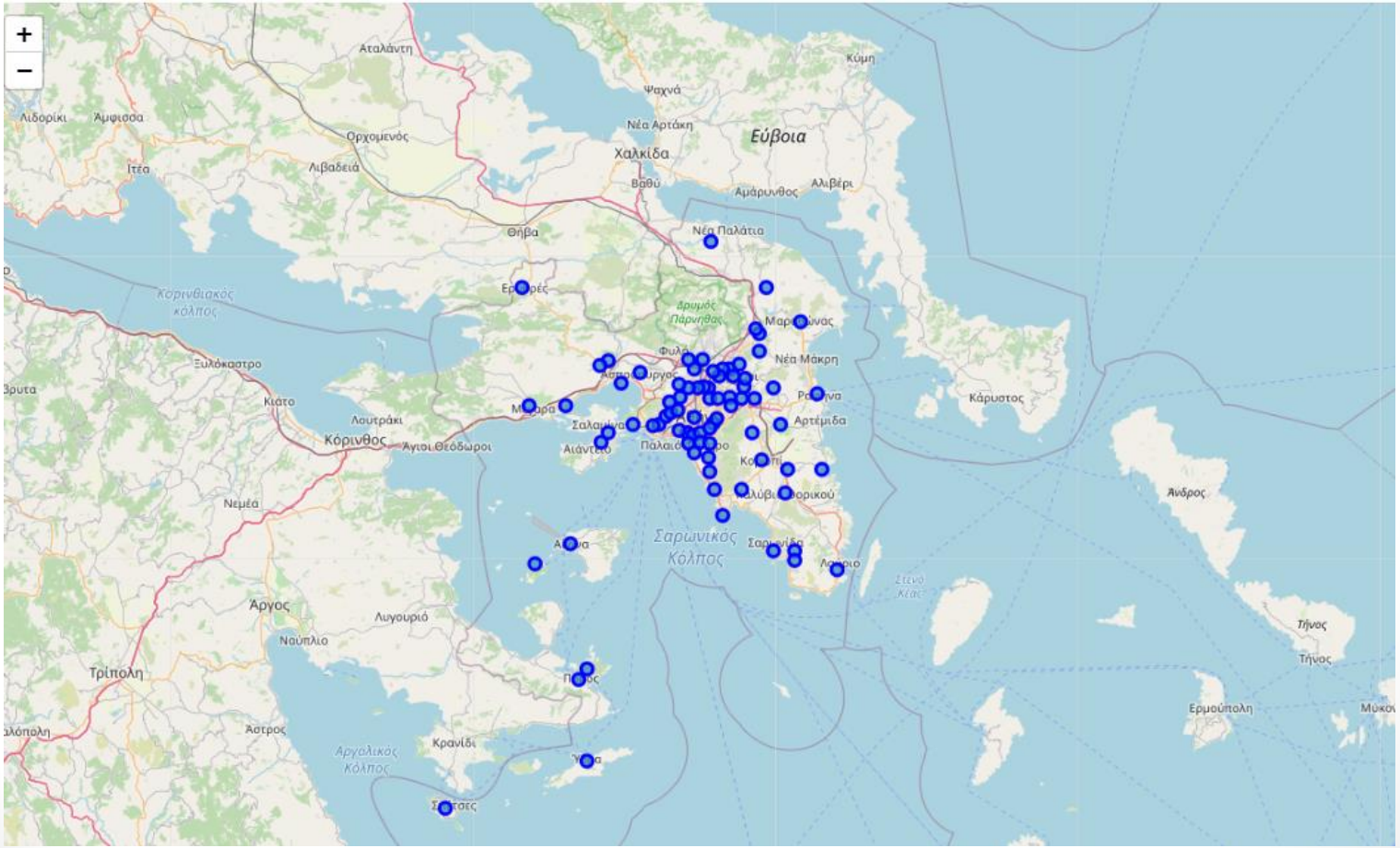
Data acquisition and cleaning

- **Data sources : D**ata was obtained from the web and Foursquare. The initial dataset was downloaded from (https://simplemaps.com/data/gr-cities).

- **Data cleaning:** Data downloaded in csv file from the web and cleaned and then combined into dataframe.

- **Features:** The number neighborhoods in Attika are 82. Number o features will be 198, as the number of venues. Data will be downloaded form Foursquare API.

Exploratory Data Analysis – Visualization Map with locations



Exploratory Data Analysis –Neighborhood with top 10 venues

Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Acharnes	Café	Mobile Phone Shop	Supermarket	Creperie	Souvlaki Shop	Taverna	Seafood Restaurant	Plaza	Cosmetics Shop	Dance Studio
Agia Paraskevi	Bakery	Pharmacy	Pizza Place	Clothing Store	Café	Plaza	Baby Store	Cosmetics Shop	Coffee Shop	Furniture / Home Store
Agia Varvara	Grocery Store	Fish Taverna	Café	Optical Shop	Greek Restaurant	Restaurant	Bakery	Bus Stop	Fast Food Restaurant	Betting Shop
Agkistri	Hotel	Nightclub	Hotel Bar	Cocktail Bar	Harbor / Marina	Café	Greek Restaurant	Women's Store	Farmers Market	Food & Drink Shop
Aigaleo	Café	Bar	Meze Restaurant	Coffee Shop	Burger Joint	Souvlaki Shop	Mobile Phone Shop	Donut Shop	Bakery	Snack Place

Exploratory Data Analysis – Number of restaurants in every Neighborhood (top 10)

Neighborhood	Number_of_Restraurants
ydra	21
Nea Filadelfeia	19
Argyroupoli	17
Kaisariani	15
Chalandri	15
Palaia Fokaia	14
Ilioupoli	14
Melissia	12
Aigina	12
alimos	12

Number_of_Restraurants: Venues containing the words Restaurant, Taverna and Souvlaki

Exploratory Data Analysis – Which Boroughs contain restaurant in top 3 venues

Neighborhood	1st Most Common Venue
Dafni	Greek Restaurant
Elefsina	Cretan Restaurant
Kalyvia Thorikou	Taverna
Koropi	Meze Restaurant
Mandra	Taverna
Melissia	Greek Restaurant
Oropos	Greek Restaurant
Palaia Fokaia	Seafood Restaurant
Porto Rafti	Fish Taverna
Zografos	Greek Restaurant
agios Dimitrios	Souvlaki Shop
ydra	Greek Restaurant

Neighborhood	2nd Most Common Venue
Agia Varvara	Fish Taverna
Aigina	Greek Restaurant
Chaidari	Greek Restaurant
Dafni	Meze Restaurant
Galatas	Greek Restaurant
Galatsi	Greek Restaurant
Gerakas	Grilled Meat Restaurant
Ilioupoli	Meze Restaurant
Kaisariani	Meze Restaurant
Kallithea	Souvlaki Shop
Kalyvia Thorikou	Greek Restaurant
Kifisia	Japanese Restaurant
Megara	Grilled Meat Restaurant
Moschato	Greek Restaurant
Nea Filadelfeia	Souvlaki Shop
Nea Peramos	Greek Restaurant
Paiania	Greek Restaurant
Palaia Fokaia	Greek Restaurant
Peristeri	Souvlaki Shop
Petroupoli	Greek Restaurant
Piraeus	Souvlaki Shop
Porto Rafti	Greek Restaurant
alimos	Greek Restaurant

Neighborhood	3rd Most Common Venue
Aigaleo	Meze Restaurant
Anavyssos	Greek Restaurant
Argyroupoli	Souvlaki Shop
Ilioupoli	Kebab Restaurant
Kalyvia Thorikou	Souvlaki Shop
Kapandriti	Seafood Restaurant
Kitsi	Greek Restaurant
Mandra	Souvlaki Shop
Metamorfosi	Greek Restaurant
Nea Filadelfeia	Meze Restaurant
Perama	Meze Restaurant
Saronida	Seafood Restaurant
Vyronas	Souvlaki Shop
Zefyri	Souvlaki Shop
agios Stefanos	Souvlaki Shop

Clustering model

- **Clustering**: Is a type of unsupervised machine learning that is used to organize data points into similar groups called clusters.
- **Method:** K-means is an algorithm that repeatedly partitions observations into a fixed number, k, of non-overlapping clusters
- **Parameters:** In this project we assume that k is 5. The number of features for the Clustering algorithm will be 198.

Clusters

Cluster 1

Agia Paraskevi	Kalyvia Thorikou	Pefki	agios Dimitrios
Agia Varvara	Keratsini	Perama	agios Stefanos
Agkistri	Kifisia	Peristeri	alimos
Argyroupoli	Kitsi	Piraeus	ano Liosia
Athens	Koropi	Poros	anoixi
Chalandri	Mandra	Porto Rafti	ilion
Cholargos	Marathonas	Psychiko	Kallithea
Dafni	Marousi	Rafina	Palaia Fokaia
Dionysos	Megara	Saronida	agioi Anargyroi
Elefsina	Melissia	Spata	Kaisariani
Erythres	Metamorfosi	Voula	Ntrafi
Gerakas	Nea Filadelfeia	Vouliagmeni	Zografos
Glyfada	Nea Ionia	Vrilissia	
Ilioupoli	Nea Smyrni	Zefyri	

Cluster 2

Mobile Phone Shop	Kapandriti	Paiania	Galatas
Aigaleo	Korydallos	Palaio Faliro	Galatsi
Aigina	Kythira	Petroupoli	Irakleio
Anavyssos	Lykovrysi	Vyronas	Moschato
Aspropyrgos	Magoula	ydra	Nea Peramos
Chaidari	Markopoulo	Nikaia	
Mobile Phone Shop	Kapandriti	Paiania	

Cluster 3

Lavrio

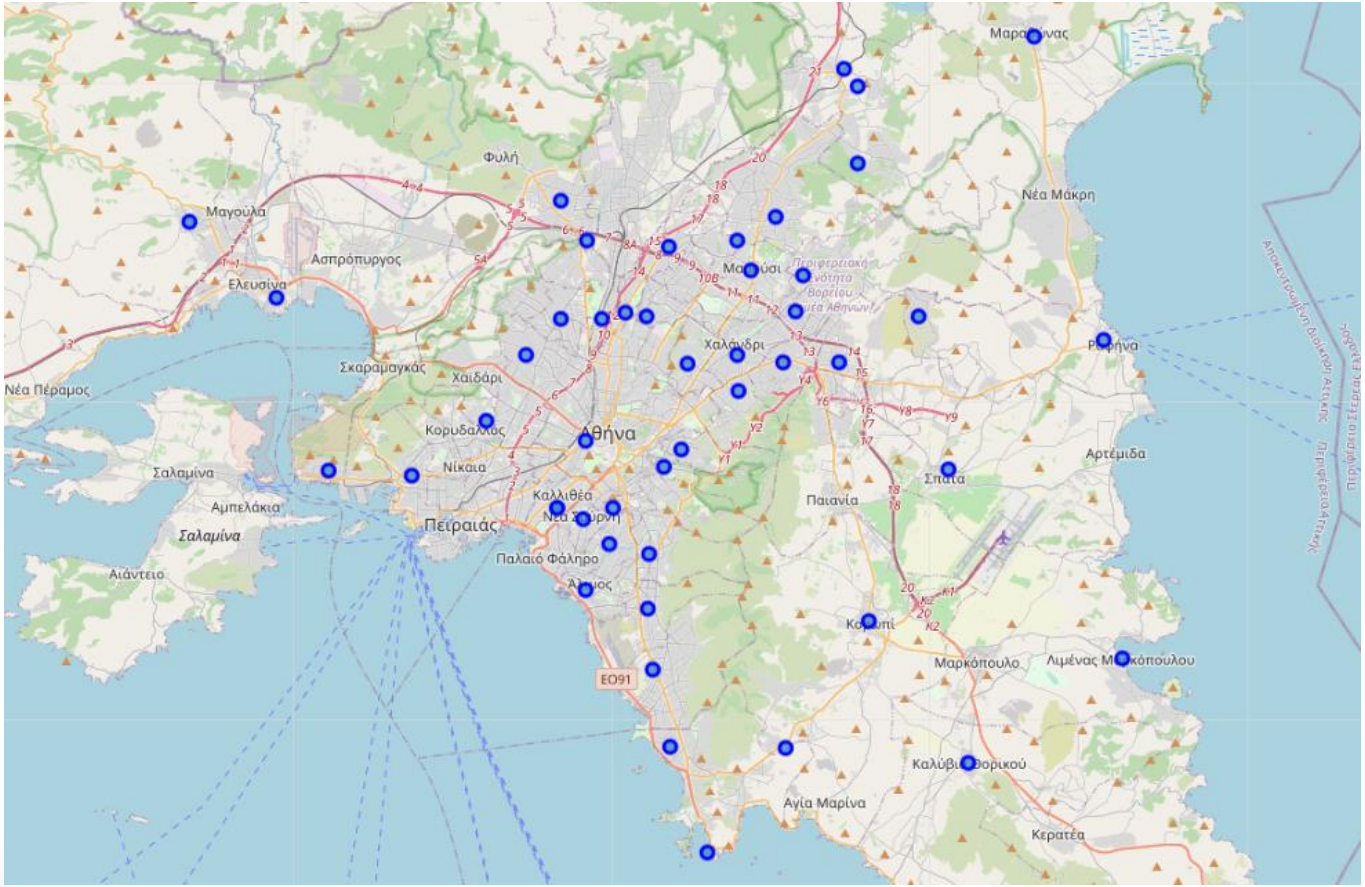
Cluster 4

Spetses

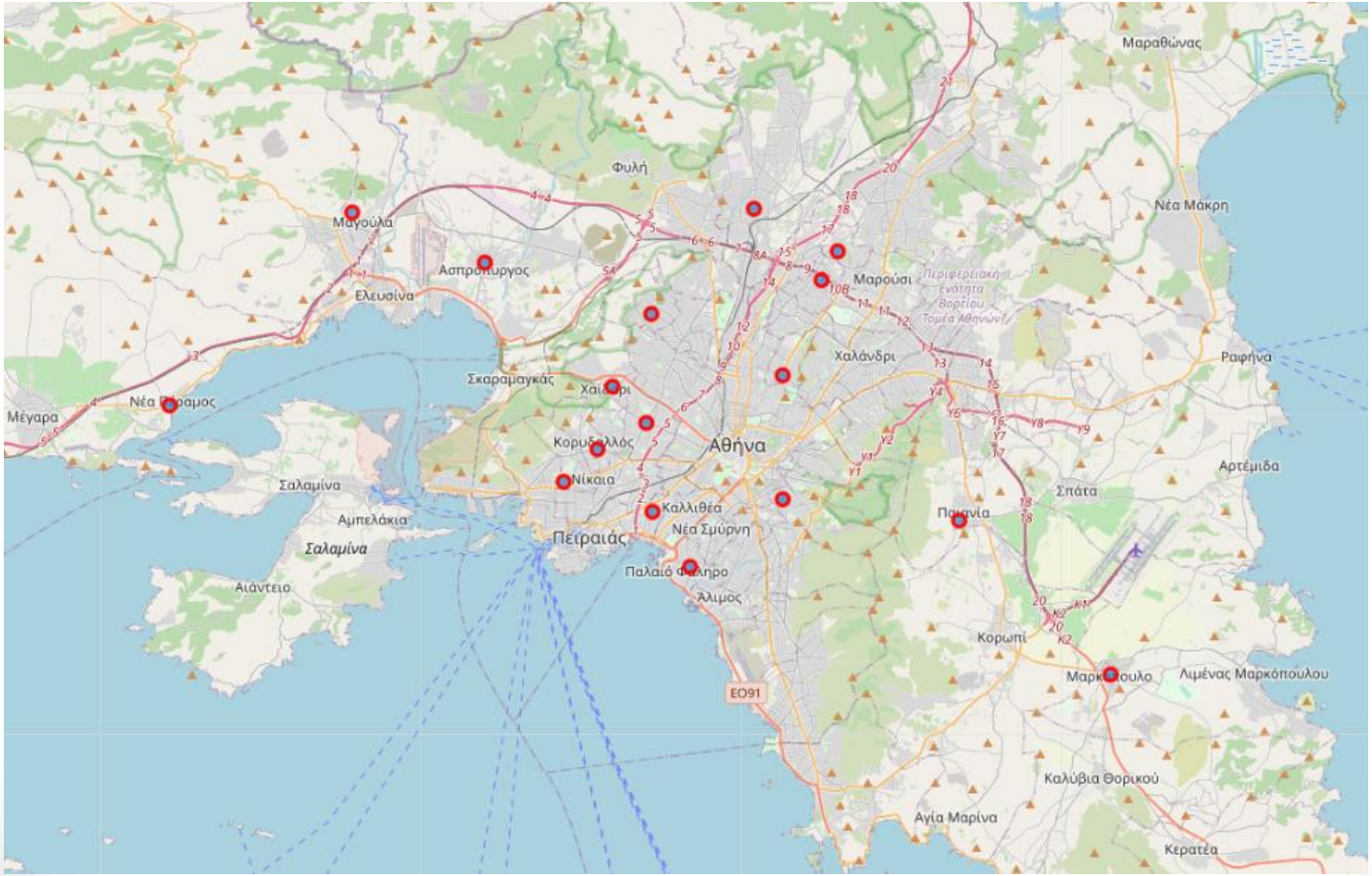
Cluster 5

Oropos

Map Visualization Cluster 1



Map Visualization Cluster 2



Conclusion and future directions

- Identify locations which restaurants is the most frequently venue in order to aid stakeholders find the best place to open a new restaurant.

- Clustering locations into 5 categories in order to know in which area a franchise will open restaurants in order to satisfy all kind of Neighborhood clusters.

- There is place for future development of this project, as we can add more variables such as economic development, demographic data and financial data about the potential customer in each Neighborhood.