Coursera Applied Data Science Capstone

Stefanos Bontas, May 2020

The objective

Find a suitable location
to establish a burger restaurant
in Athens metropolitan area, Greece

Selecting criteria

A suitable restaurant location should meet the below criteria:

- Be **close** to a metro line station, in order to be easily accessible by the mass since Athens has a huge problem with car parking
- The **total number of similar restaurants** with a radius of 1,000 meters around each metro station should be low
- The **nearest existing restaurant** to the metro line station should also have some impact on the results

An **already existing** burger restaurant can be:

- American restaurant
- Burger joint
- Fast food restaurant

Data

In order to meet the above requirements, we do need to download the following datasets from foursquare developer API:

- **Metro stations in Athens and its suburbs**. This is the venue category 4bf58dd8d48988d1fd931735 in foursquare.

- **The fast food restaurants around them** (1,000 meters radius distance should be enough). This is the venue category 4bf58dd8d48988d16e941735 in foursquare.

For obtaining the correct metro line stationsdata set, we should consider that a distance of **20 kilometers** radium around the center of Athens (Syntagma square, a point which is being used as a benchmark for all the distances count from/to the city of Athens).

Methodology

Taking as point zero the Syntagma square, we found all of the metro line stations in a radius of 20,000 meters. Then we'll find all of the relevant restaurants with radius 1,000 meters from each station and cluster them by using the **K-means** machine learning algorithm.
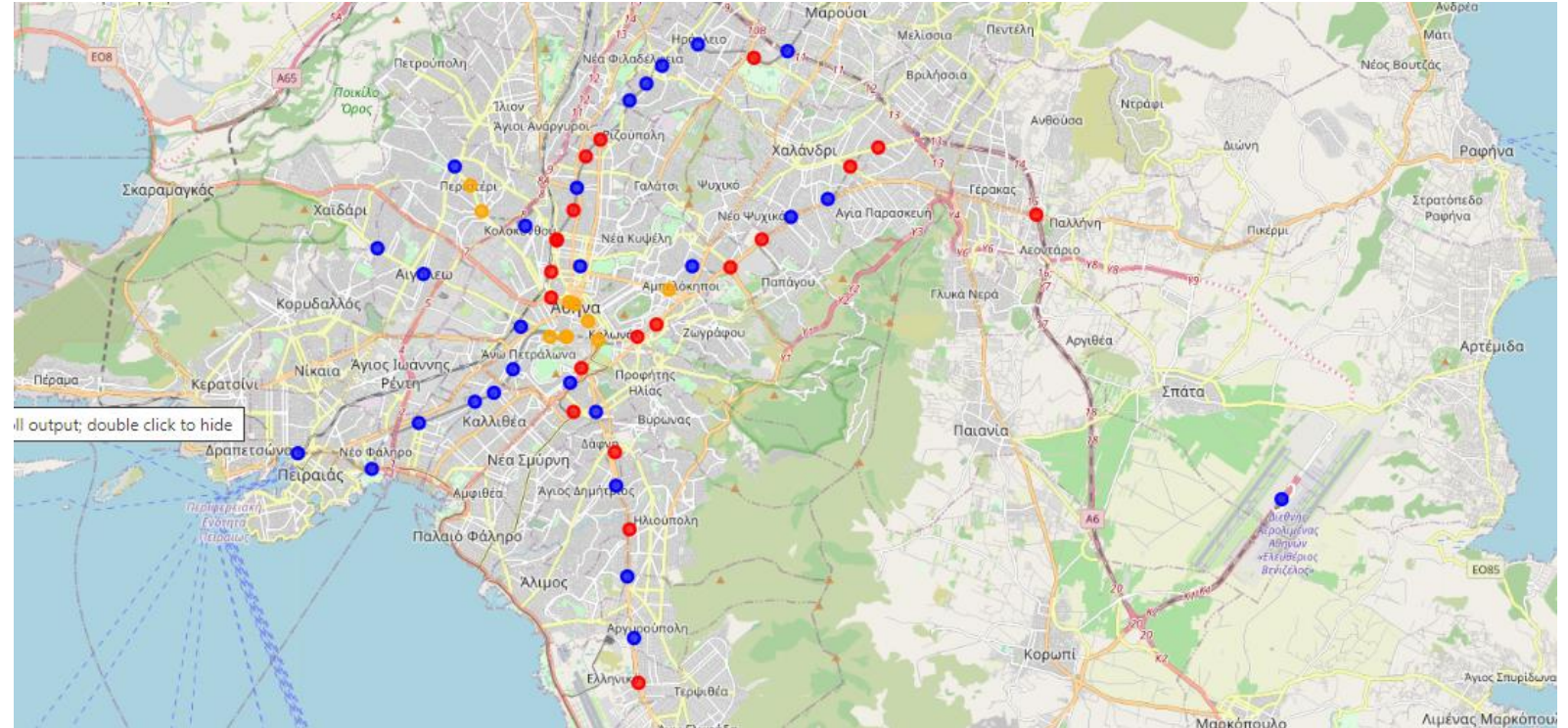
Results (1/3)

K-Means Elbow curve
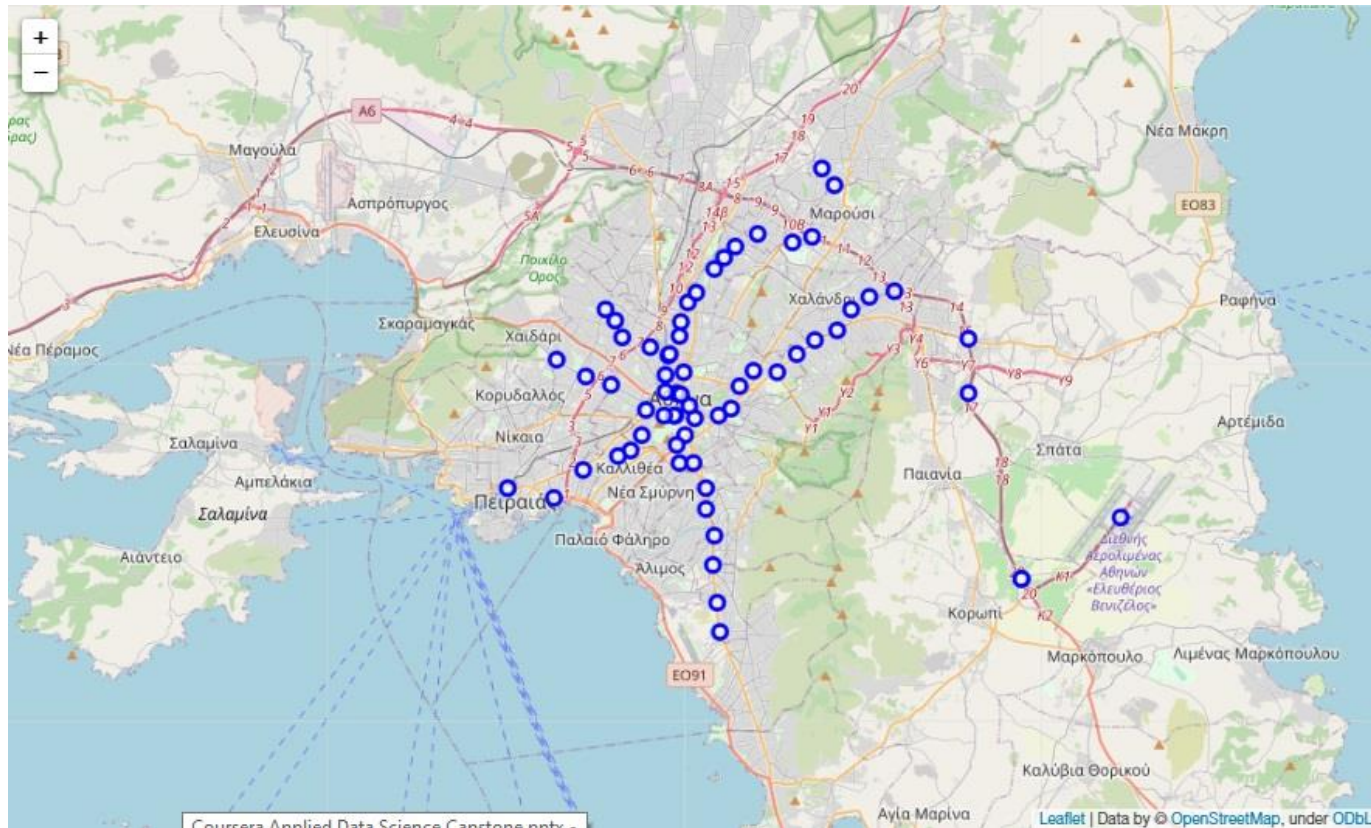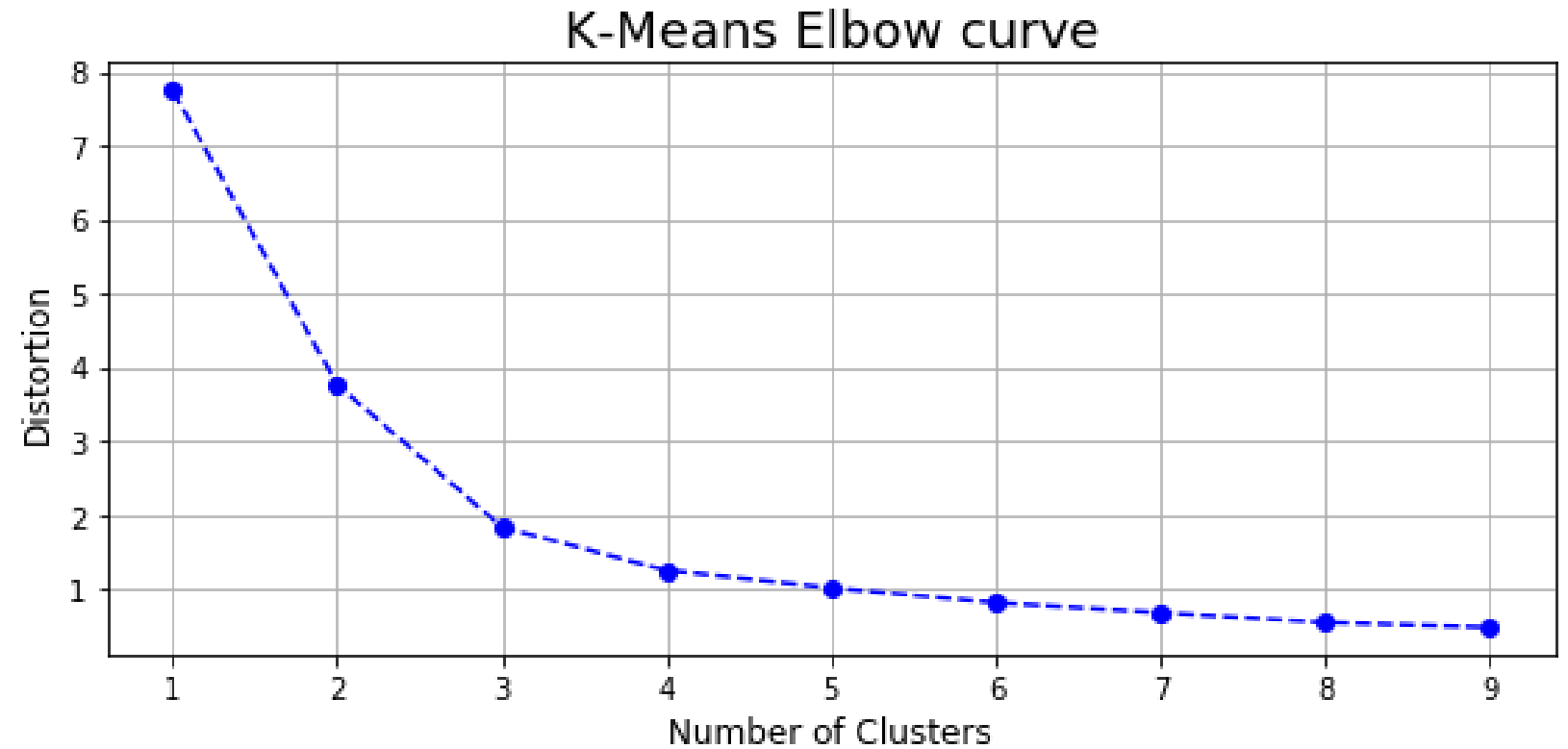
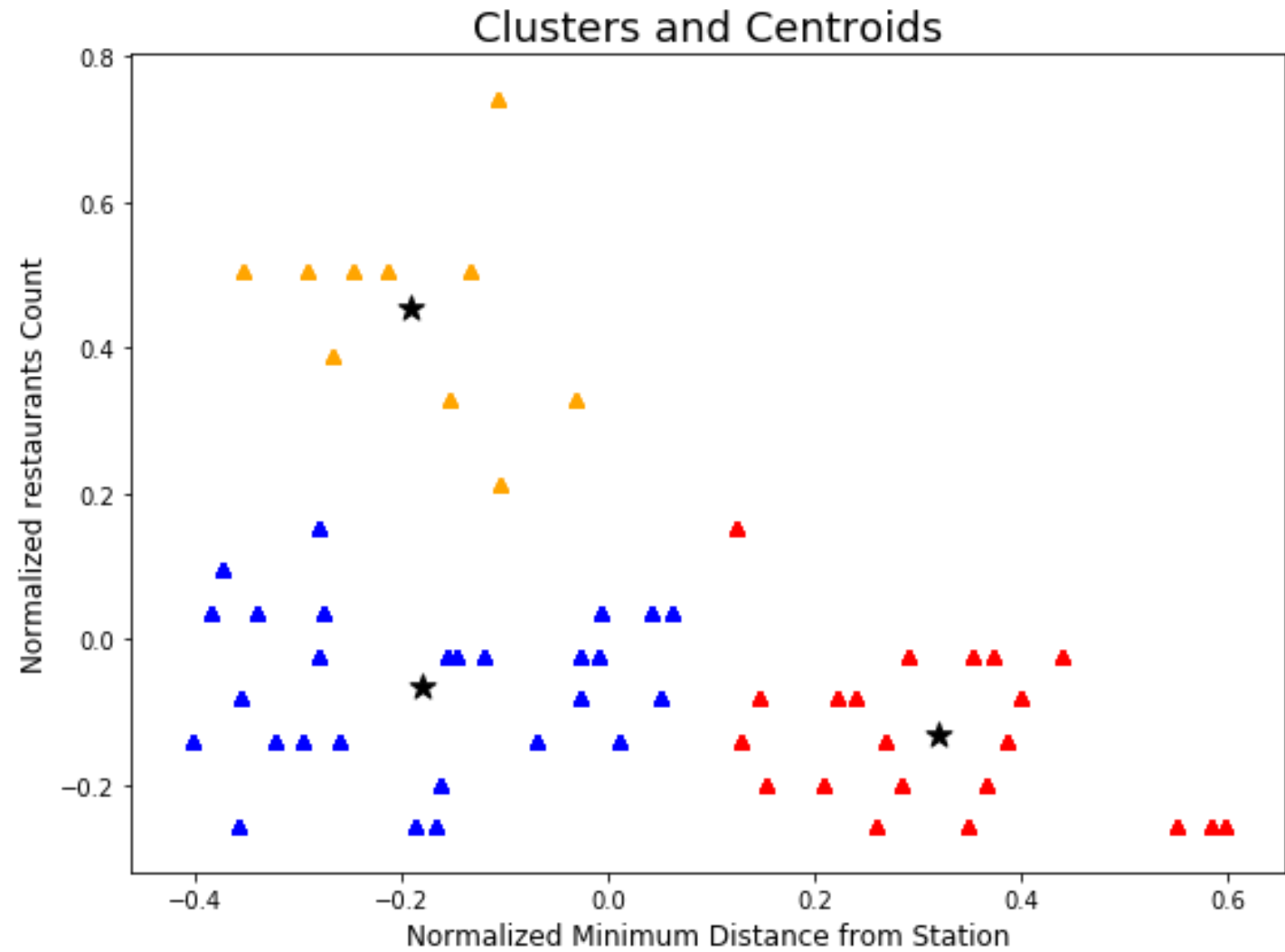We'll use **3 clusters** for K-Means clustering algorithm.

Results (2/3)

Clustered metro line map

where, **blue circles** = high potential (locations with business opportunity)
orange circles = average potential (locations to be further investigated)
red circles = low potential (locations to be avoided)

Results (3/3)

Clusters and Centroids

Conclusion

- In order the analysis to be valid, the selected location should be also checked for potential **venues not listed** in foursquare

- Since Athens metropolitan railway is expanding, **new neighborhoods** are being added as business opportunities so this analysis should be re-executed

- Some stations appear to have **no similar restaurants nearby**, which is quite worrying and further analysis is being needed.

- Other socioeconomic factors should be taken into consideration such as:
  - The **actual traffic** of each location since some stations are being served as **transportation hubs** (e.g. Syntagma square, Piraeus port, etc.) or are close to **touristic attractions** (e.g. Monastiraki & Acropolis stations)
  - Total **number of restaurants** established nearby can also be an indication if the location is good or not for dining
  - **Real estate indexes** can also distinguish an emerging location from an emerged one

Thank you!

Stefanos Bontas, May 2020