

Now that you have been equipped with the skills and the tools to use location data to explore a geographical location, over the course of two weeks, you will have the opportunity to be as creative as you want and come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of your choice or to come up with a problem that you can use the Foursquare location data to solve. If you cannot think of an idea or a problem, here are some ideas to get you started:

In Module 3, we explored New York City and the city of Toronto and segmented and clustered their neighborhoods. Both cities are very diverse and are the financial capitals of their respective countries. One interesting idea would be to compare the neighborhoods of the two cities and determine how similar or dissimilar they are. Is New York City more like Toronto or Paris or some other multicultural city? I will leave it to you to refine this idea.

In a city of your choice, if someone is looking to open a restaurant, where would you recommend that they open it? Similarly, if a contractor is trying to start their own business, where would you recommend that they setup their office?

These are just a couple of many ideas and problems that can be solved using location data in addition to other datasets. No matter what you decide to do, make sure to provide sufficient justification of why you think what you want to do or solve is important and why would a client or a group of people be interested in your project.

Review criterialess

This capstone project will be graded by your peers. This capstone project is worth 70% of your total grade. The project will be completed over the course of 2 weeks. Week 1 submissions will be worth 30% whereas week 2 submissions will be worth 40% of your total grade.

For this week, you will required to submit the following:

A description of the problem and a discussion of the background. (15 marks)

A description of the data and how it will be used to solve the problem. (15 marks)

For the second week, the final deliverables of the project will be:

A link to your Notebook on your Github repository, showing your code. (15 marks)

A full report consisting of all of the following components (15 marks):

Introduction where you discuss the business problem and who would be interested in this project.

Data where you describe the data that will be used to solve the problem and the source of the data.

Methodology section which represents the main component of the report where you discuss and describe any exploratory data analysis that you did, any inferential statistical testing that you performed, and what machine learnings were used and why.

Results section where you discuss the results.

Discussion section where you discuss any observations you noted and any recommendations you can make based on the results.

Conclusion section where you conclude the report.

3. Your choice of a presentation or blogpost. (10 marks)

Clearly define a problem or an idea of your choice, where you would need to leverage the Foursquare location data to solve or execute. Remember that data science problems always target an audience and are meant to help a group of stakeholders solve a problem, so make sure that you explicitly describe your audience and why they would care about your problem.

This submission will eventually become your Introduction/Business Problem section in your final report. So I recommend that you push the report (having your Introduction/Business Problem section only for now) to your Github repository and submit a link to it.

In this project, I will analyze several aspects of a portion of the Kadikoy district in Istanbul, Turkey. The aim is to search for the similarities and differences in the parts of the analyzed region and to cluster the similar ones together.

Burada yapılacak incelemeye göre biraz daha detay verilebilir

The region to be analyzed is shown in the image below:



Image 1: The part of Kadikoy to be analyzed

This region contains many venues such as cafes, restaurants and bars, as well as stores and also houses.

As a background information, Istanbul is the largest city in Turkey with a population of 15 million people. Kadikoy is one of the 39 districts of Istanbul. It is a big district with a population of 521 thousand residents. And Kadikoy itself has 21 neighborhoods (“mahalle” in Turkish). The region to be analyzed contains parts of Caferaga and Osmanaga neighborhoods. There are also names like Moda, Bahariye and Çarşı in this area which refers to certain areas, but they are not official names of the neighborhoods.

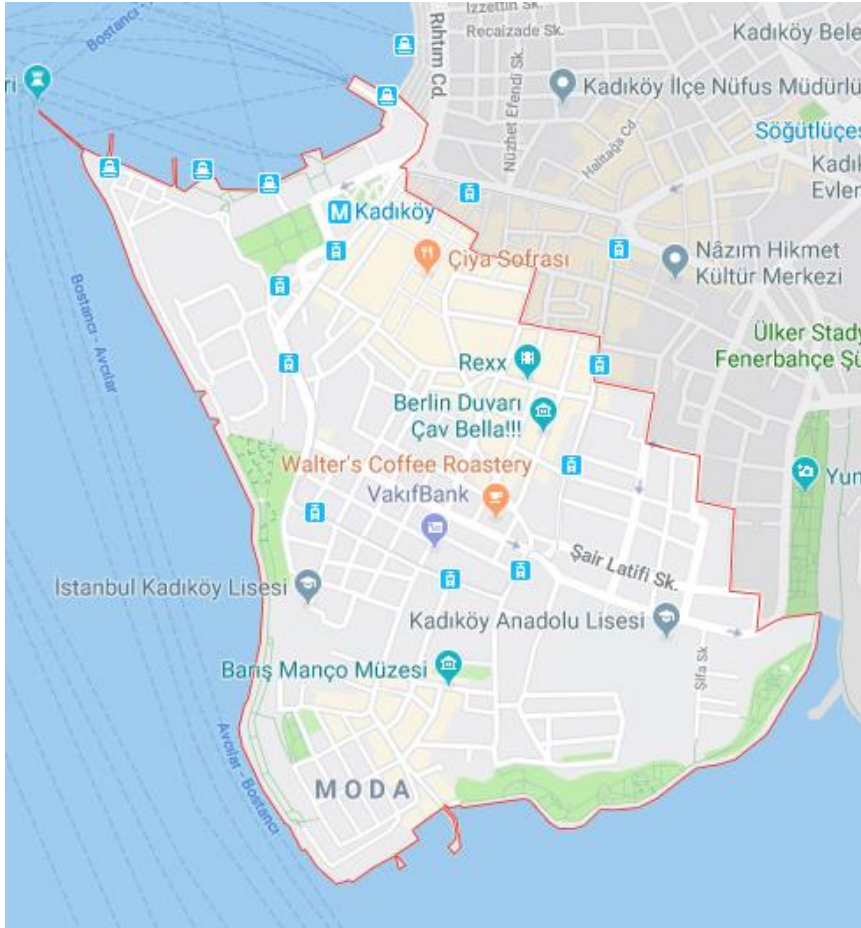


Image: Caferaga neighborhood

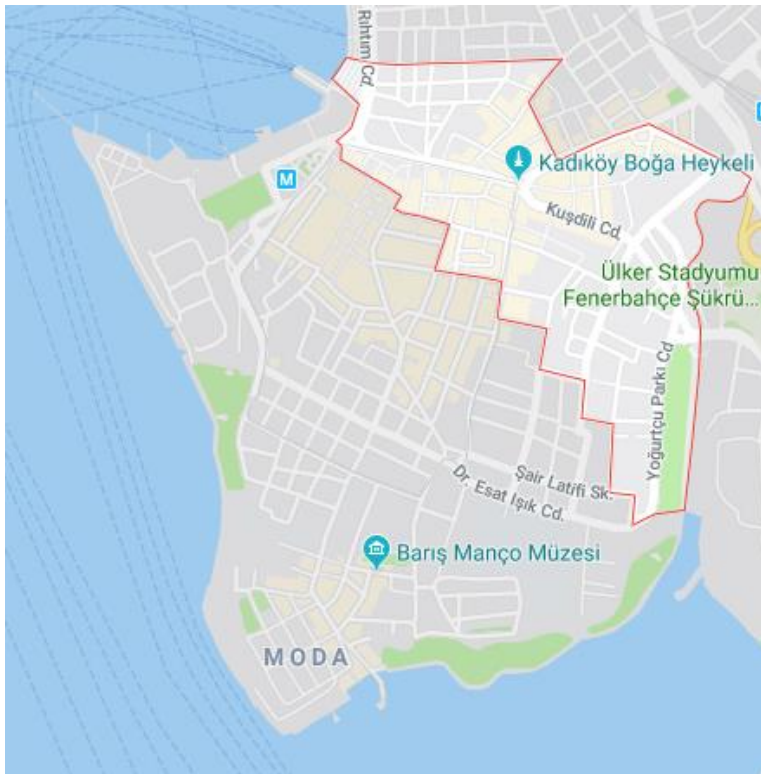
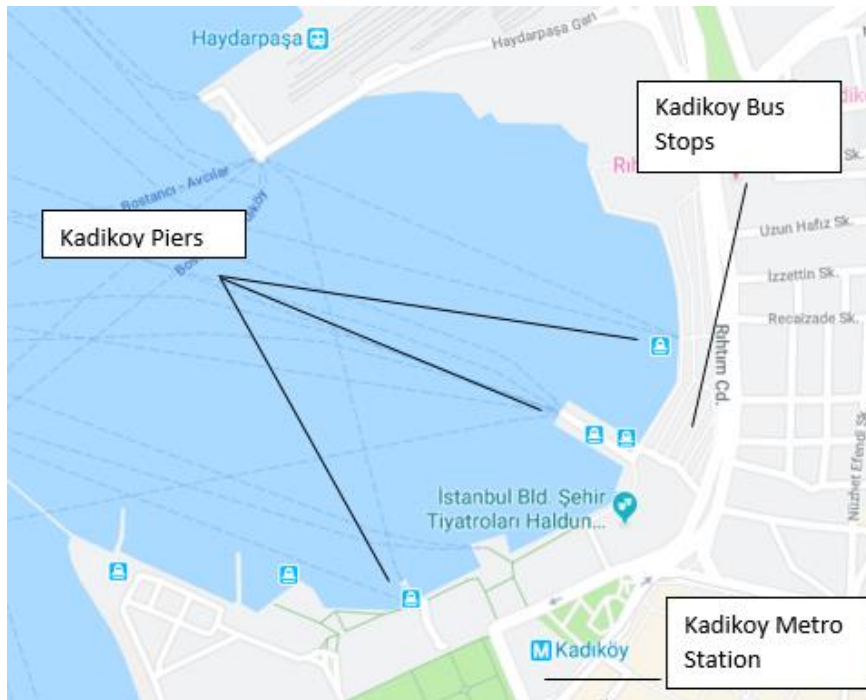
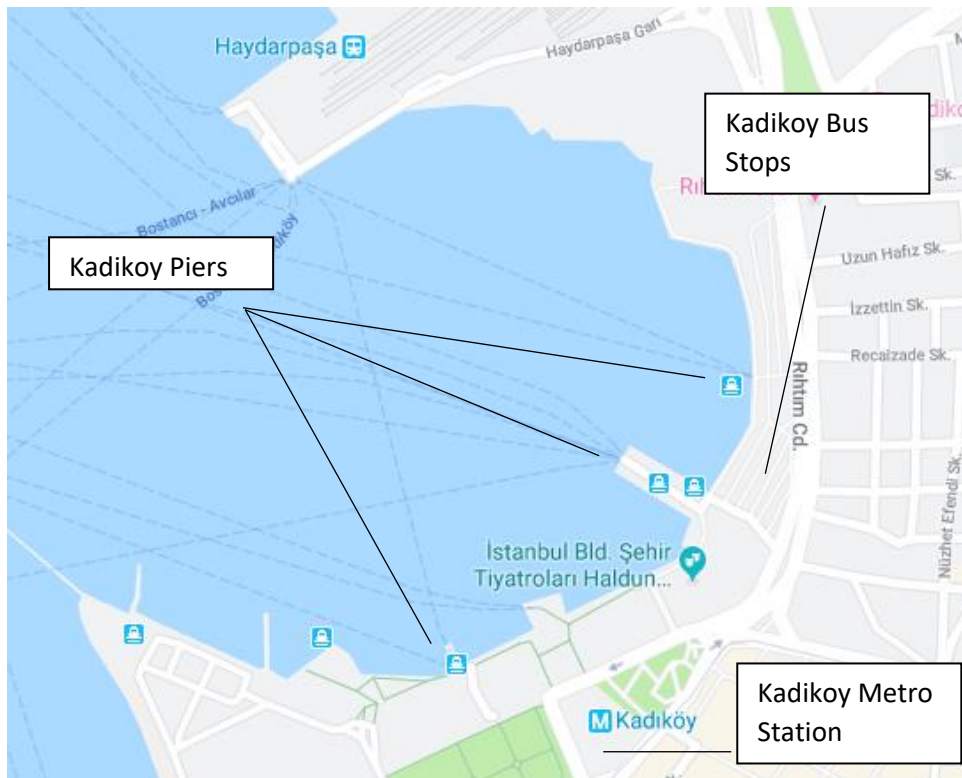


Image: Osmanaga neighborhood

Note: For the people who are not familiar with Istanbul or Kadikoy, I think it may be good to know this piece of additional information: Kadikoy may refer to (a) the district of Kadikoy as a whole (b) the area within this district, where Kadikoy Piers, Kadikoy Bus Stops and Kadikoy Metro Station or even the Kadikoy Bazaar are located. So the term “Kadikoy” changes on the context it is used. For instance, if one is talking about the election results in Kadikoy, he is referring to meaning (a), and if two friends are talking about meeting at Kadikoy, they are probably referring to meaning (b). These are also shown in the images below. The area analyzed in this project is the area where many venues such as cafes, restaurants, bars are located close to (b).





References

- (1) World Population Review
<http://worldpopulationreview.com/world-cities/istanbul-population/>
Accessed on 24.06.2019
- (2) Wikipedia / Turkcewiki
http://en.turkcewiki.org/wiki/List_of_districts_of_Istanbul
Accessed on 24.06.2019
Note: Since access to Wikipedia has been blocked by Turkish authorities since 2017 to this date, a link is given to the mirror website above
- (3) Wikipedia / Turkcewiki
<http://en.turkcewiki.org/wiki/Kad%C4%B1k%C3%B6y>
Accessed on 24.06.2019
- (4) Wikipedia / Turkcewiki (in Turkish)
<http://www.turkcewiki.org/wiki/Kad%C4%B1k%C3%B6y#Mahalleler>
Accessed on 24.06.2019

Describe the data that you will be using to solve the problem or execute your idea. Remember that you will need to use the Foursquare location data to solve the problem or execute your idea. You can absolutely use other datasets in combination with the Foursquare location data. So make sure that you provide adequate explanation and discussion, with examples, of the data that you will be using, even if it is only Foursquare location data.

This submission will eventually become your Data section in your final report. So I recommend that you push the report (having your Data section) to your Github repository and submit a link to it.