In this week, you will continue working on your capstone project. Please remember by the end of this week, you will need to submit the following:

1. 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, if any, 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.

2. A link to your Notebook on your Github repository pushed showing your code. (**15 marks**)

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

**Here are examples of previous outstanding submissions that should give you an idea of what your report would look like, what your notebook would look like in terms of clean, clear, and well-commented code, and what your presentation would look like or your blogpost would look like:**

1. **Report:**[**https://cocl.us/coursera\_capstone\_report**](https://cocl.us/coursera_capstone_report)
2. **Notebook:**[**https://cocl.us/coursera\_capstone\_notebook**](https://cocl.us/coursera_capstone_notebook)
3. **Presentation:**[**https://cocl.us/coursera\_capstone\_presentation**](https://cocl.us/coursera_capstone_presentation)
4. **Blogpost:**[**https://cocl.us/coursera\_capstone\_blogpost**](https://cocl.us/coursera_capstone_blogpost)

Ideal location for opening a restaurant in New York City

Introduction

Owning a restaurant can be a dream to most of us, but the competition is acute. There are a lot of excellent chiefs outside. The location to set up the restaurant becomes the decisive factor to win the market.

Foursquare location data is used to execute my idea. An ideal location should consider the potential passenger flow, number of competitors, and race. The data is from Manhattan, New York.

Data

New York location data from Foursquare is extracted along with Wikipedia to understand the population of different races in New York. After analyzing the ideal area, it is found that Manhattan is a good place to start a restaurant business. Since the content of Wikipedia page is crowded with many sections ,and so it is difficult to extract specific area of data. Demographics of Manhattan data is manually extracted from Wikipedia.

<https://en.wikipedia.org/wiki/Demographics_of_Manhattan>

Data Analysis

This is a map of the New York City. It is generated by folium.map function. Map

Description automatically generated

We mainly focus on the Manhattan borough.

Map

Description automatically generated

#### After that, we look into the top 5 most common venues in each neighbourhood in Manhattan.

#### 

|  |  |  |  |
| --- | --- | --- | --- |
| A screenshot of text  Description automatically generated | A screenshot of a cell phone screen with text  Description automatically generated | | A screenshot of a cell phone screen with text  Description automatically generated |
| A screenshot of text  Description automatically generated | A screenshot of text  Description automatically generated | A screenshot of text  Description automatically generated | |
| A screenshot of text  Description automatically generated | A screenshot of a cell phone screen with text  Description automatically generated | A screenshot of a cell phone screen with text  Description automatically generated | |
| A screenshot of text  Description automatically generated | A picture containing text  Description automatically generated |  | |

The figures below display the top two common venue is the entire Manhattan. From the data shown, Italian restaurant and coffee shop are easy to find in most of the neighbourhoods.

A picture containing text

Description automatically generated Text

Description automatically generated

All the neighbourhoods are grouped into five clusters by running kmeans.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Cluster 0 | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 |
| Hamilton Heights | Chinatown | Washington Heights | **Stuyvesant Town** | **Marble Hill** |
| Lower East Side | Central Harlem | Inwood |  |  |
|  | Upper East Side | Manhattanville |  |  |
|  | Yorkville | East Harlem |  |  |
|  | Lenox Hill | Morningside Heights |  |  |
|  | Roosevelt Island | Tudor City |  |  |
|  | Upper West Side |  |  |  |
|  | Lincoln Square |  |  |  |
|  | Clinton |  |  |  |
|  | Midtown |  |  |  |
|  | Murray Hill |  |  |  |
|  | Chelsea |  |  |  |
|  | Greenwich Village |  |  |  |
|  | East Village |  |  |  |
|  | Tribeca |  |  |  |
|  | Little Italy |  |  |  |
|  | Soho |  |  |  |
|  | West Village |  |  |  |
|  | Manhattan Valley |  |  |  |
|  | Gramercy |  |  |  |
|  | Battery Park City |  |  |  |
|  | Financial District |  |  |  |
|  | Carnegie Hill |  |  |  |
|  | Noho |  |  |  |
|  | Civic Center |  |  |  |
|  | Midtown South |  |  |  |
|  | Sutton Place |  |  |  |
|  | Turtle Bay |  |  |  |
|  | Flatiron |  |  |  |
|  | Hudson Yards |  |  |  |

Map

Description automatically generated

Apart from the common venue, background of the potential customers can be a key factor determining the success of the business. The race of the population is analysed. White occupies nearly 60% of the whole population, following is the African American(18.1%).

Chart, pie chart

Description automatically generated

Conclusion

The food and beverage industry is robust in Manhattan. Out of 20 the first most common venues, 13 of them are restaurant/ bar, especially the Italian restaurant and coffee shop. The market of these two cuisine is relatively saturated, so they are not suggested. For the location, Hudson Yard and Stuyvesant Town are ideal neighbourhoods since they have fewer restaurant over there compared to others. And they have much more gym/fitness centres, hotels and harbours which mean not only the locals but also the tourists would visit there. Based on the race population of Manhattan, cuisines that preferred by the white would be popular.