

THE BATTLE OF NEIGHBOURHOODS

July 2019

BUSINESS PLAN

Mr Muller wants to Open Cinema in USA and wants to find out ideal location for Investing his ideal location includes following things:

1. Population Density of the locality
2. Per Capita Income (PCI)
3. Population of each location
4. Venues in each location

DATA SECTION

Following data were used for this project:

1. https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population
2. https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income

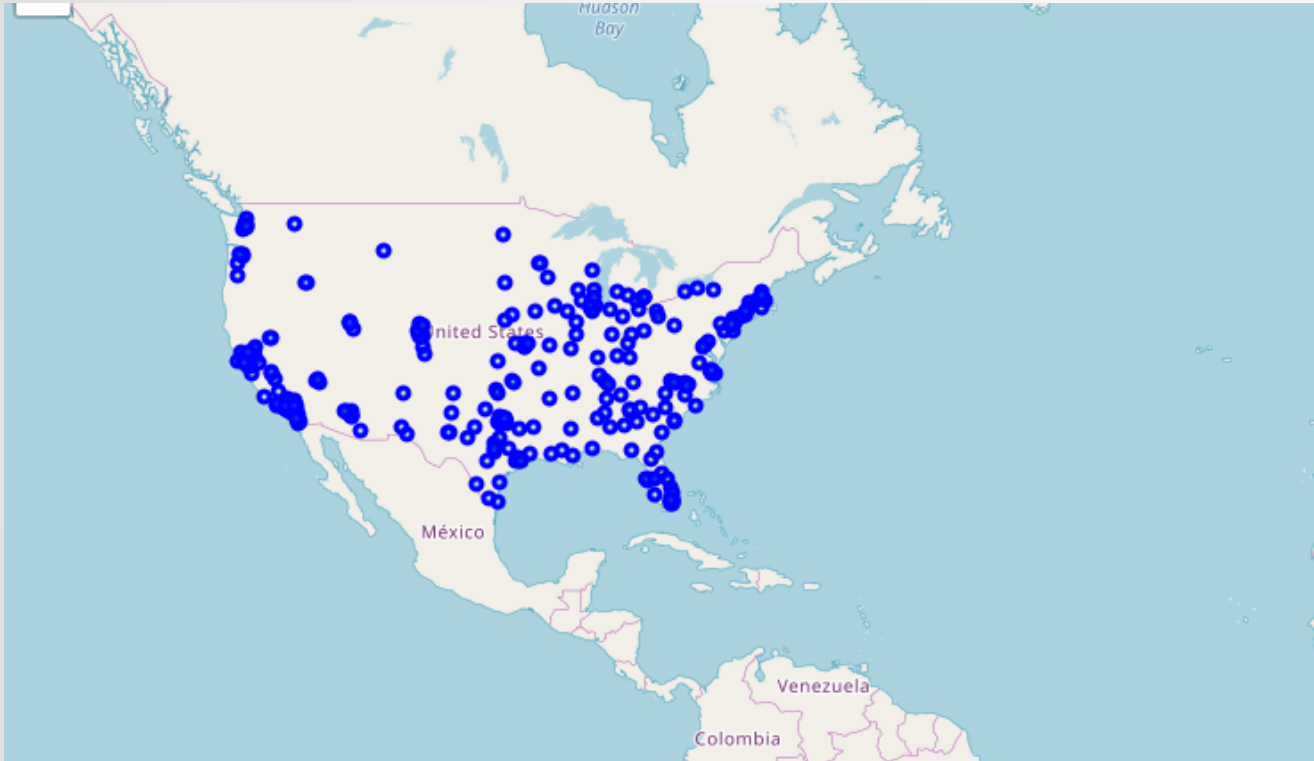
METHODOLOGY

- The Wikipedia page (https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population) was scraped using the BeautifulSoup library to build a pandas dataframe listing the cities, states, coordinates, area and population density. The dataframe was cleaned and processed appropriately.
- The Wikipedia page (https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income) was scraped using the BeautifulSoup library to build a pandas dataframe listing the cities, states and percapita income. The dataframe was cleaned and processed appropriately.
- The Foursquare API is then used to get the venues in each city of United States Based on the categories of each venue as decided by the CEO, we have assigned weights to each of them and got the city that has the maximum weight.

Once the city is finalized, we again use Four Square API to get the venues within that city and assign weights to each category. We will now use K means to cluster the venues based on the category and get the coordinates of the cluster that has maximum weight which is also our preferred location for building cinema

Plot of USA CITY

Blue dots indicates the city center in the country map of USA



Plot of usa city extracted from map

Venues in Jersey City

After analyzing several factor, we found Jersey city is the ideal location for cinema.

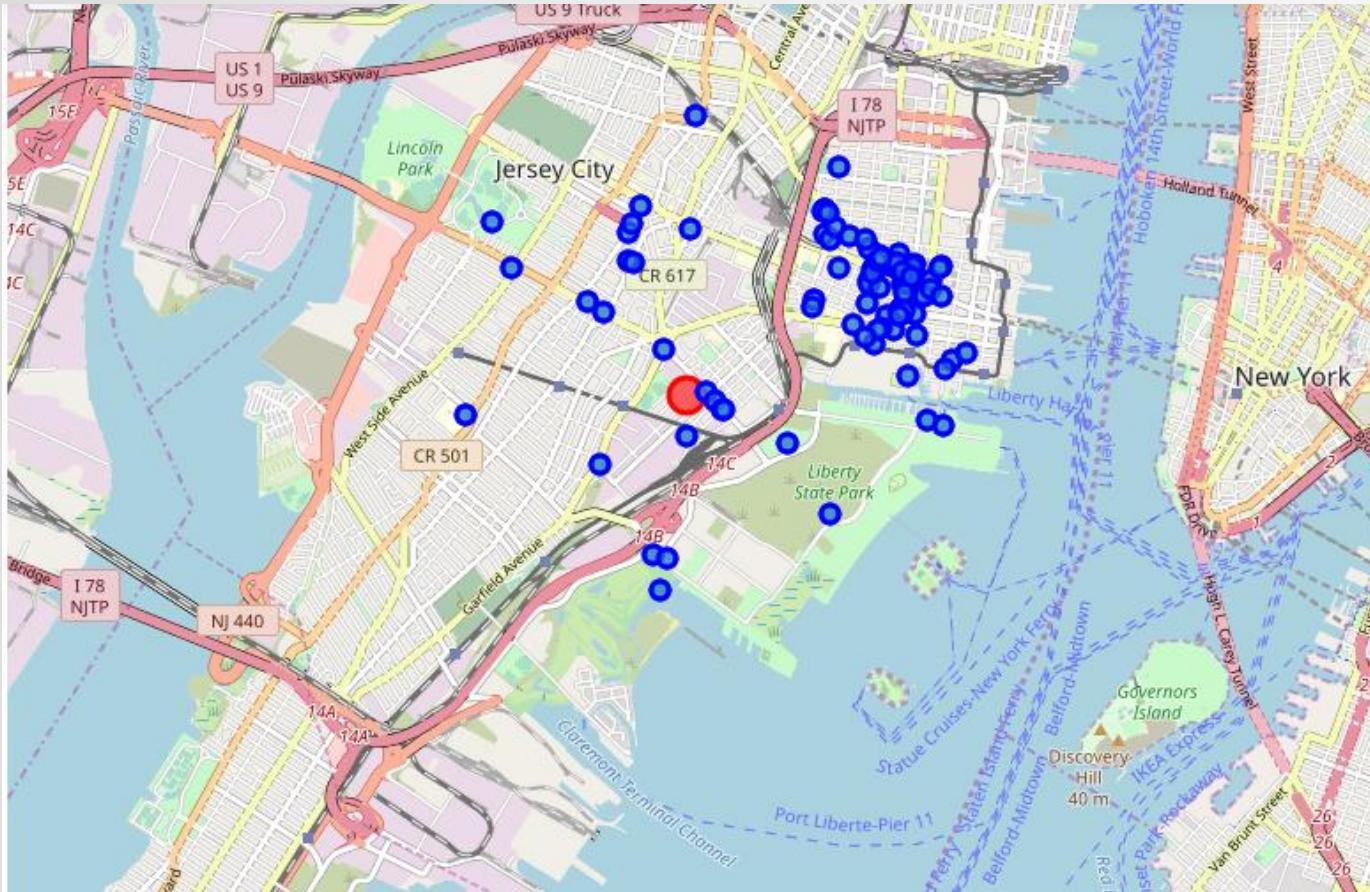


Fig: Map of venues in the Jersey City

RESULT

The green dot is the ideal place to build cinema in Jersey city.

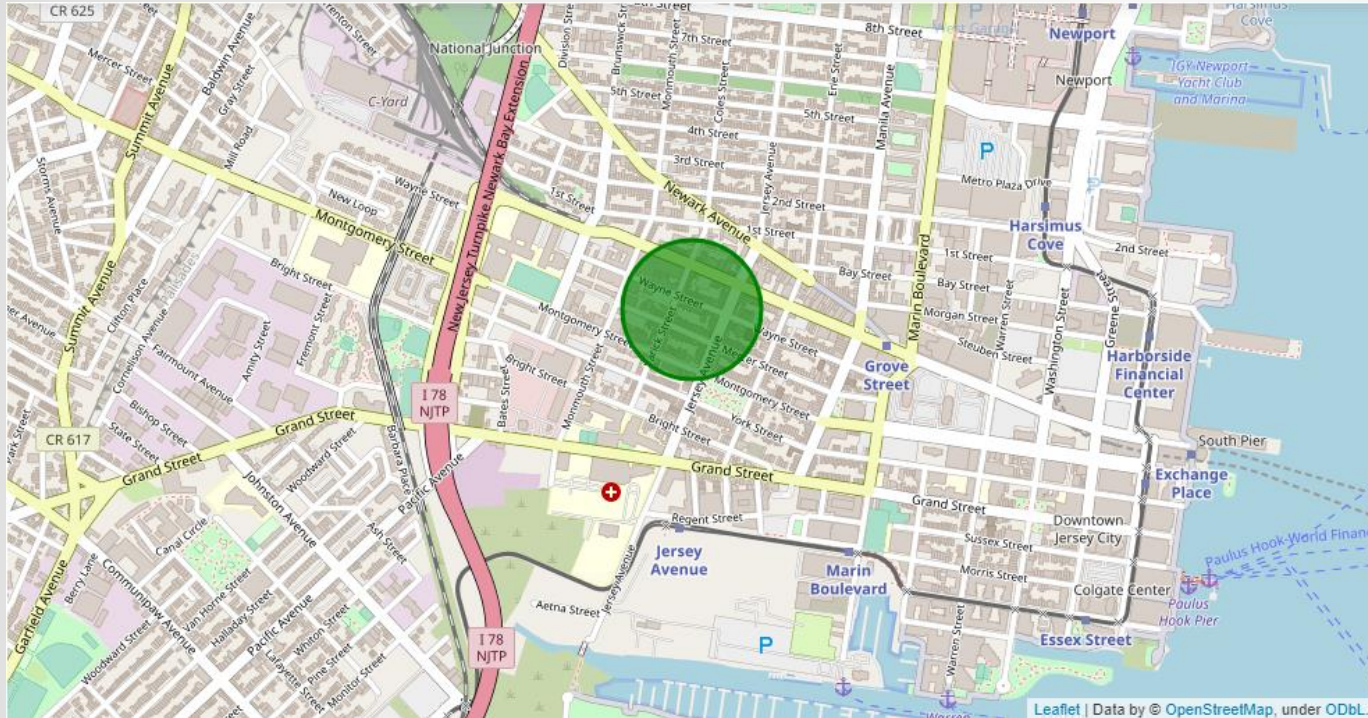


Fig: Map of ideal location in Jersey City

RECOMMENDATION/ IMPROVMENTS

- ❖ *In the Four Square API, we have queried the Venues of a locality by specifying the LIMIT and Radius of our choice. We have chosen less LIMIT as the number of API calls that can be done using a free account in Four Square are less. We can increase the limit for more accurate results.*
- ❖ *In the venue categories we are choosing only few out of 2000 that are available to give weights and identify the best cluster. Hence, assigning weights must be done relatively for each category and then considering more number of venue categories would actually yield better output.*

THANK YOU