# Determining the best location to open a departmental store - An Analysis

By Prannoy Choudhury

# BUSINESS UNDERSTANDING

Installing a departmental store in a place should be done strategically.

- Firstly, you should know why to start off with a place in particular?

- If a place is chosen, you must have a very clear understanding of who would your potential customers be?

- You need to have a clear understanding of the customers spending patterns in that locality etc.

- Last but not the least, you need to take care of the legal aspects involved in that locality

# ABOUT THE DATA

Here's some brief about the locations in particular(source: wikipedia):

As defined by the United States Census Bureau, an "incorporated place" includes a variety of designations, including city, town, village, borough, and municipality. A few exceptional census-designated places (CDPs) are also included in the Census Bureau's listing of incorporated places. Consolidated city-counties represent a distinct type of government that includes the entire population of a county, or county equivalent. Some consolidated city-counties, however, include multiple incorporated places.

About the State income levels: State income levels and income data for the United States as a whole are included for comparison. Note that county-equivalents in Louisiana are called "parishes" and in Alaska are called in "boroughs," and also that in Alaska census areas in the Unorganized Borough are county-equivalents. For states where independent cities are county-equivalents, the word "city" is included to identify the independent cities and to differentiate them from counties with identical names; the counties with the identical names have the word "county" following them. The word "county" is included in the names of counties that have names identical to the names of U.S. states or cities to differentiate them.
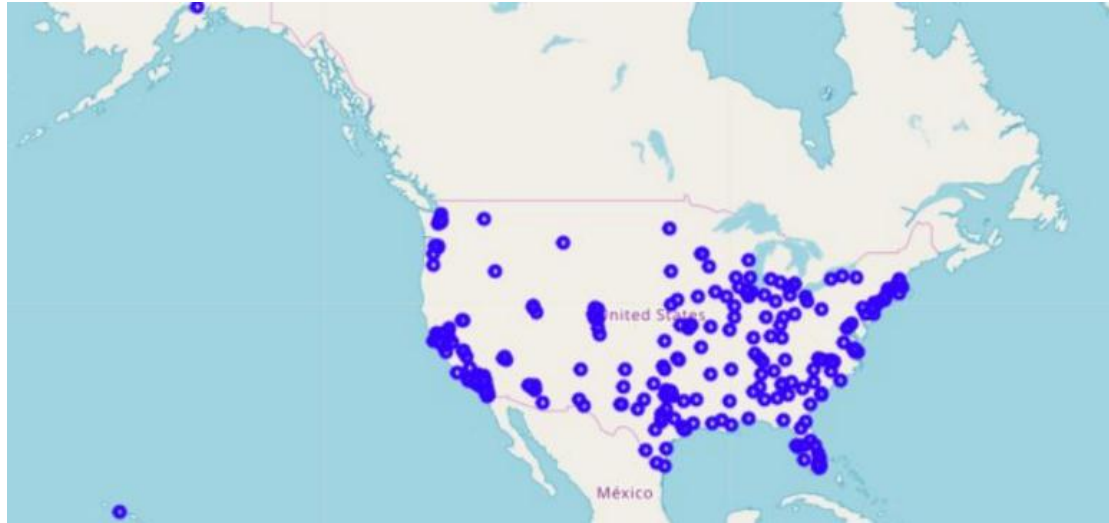
# METHODOLOGY

In order to do the analysis and suggest the best location, following are the steps we have to follow :

- 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 me, 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 to setup a gaming arcade.
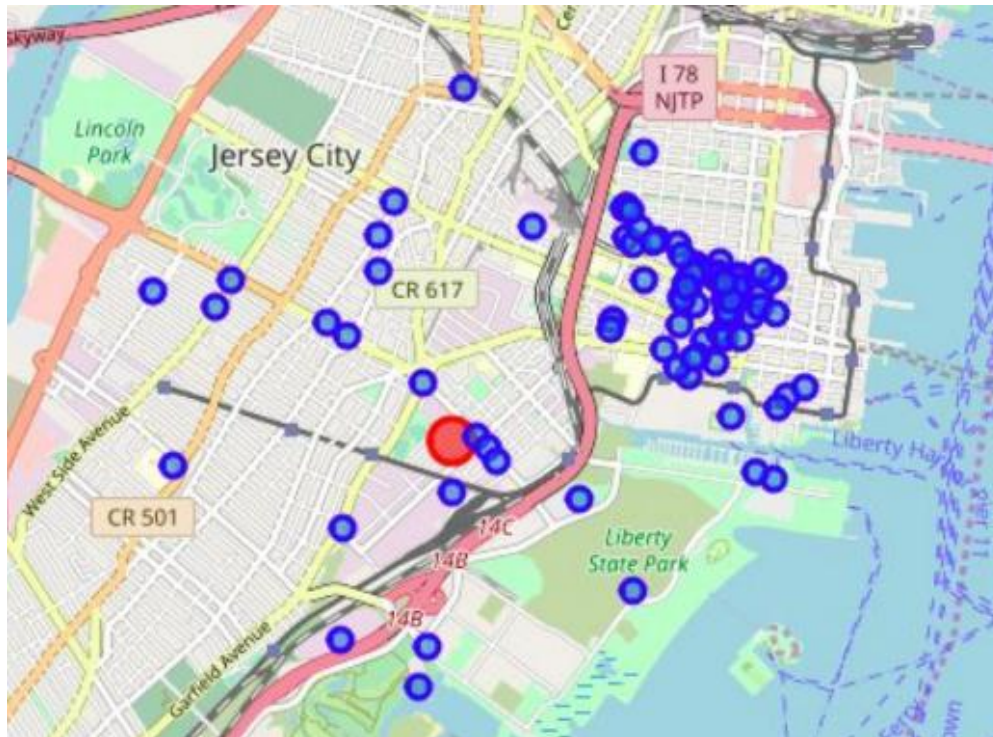
# Plots of all US cities were extracted



Blue dots represent the city centers on the map of USA

# VENUES IN JERSEY CITY



Based on constraints, we have identified Jersey city to be the best location to open a new departmental store in USA

# RESULT



The circle indicates the best place to open a departmental store in  Jersey city

# RECOMMENDATIONS/IMPROVEMENTS:

- 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.