Capstone Project - The Battle of Neighborhoods (Week 2)

-: Best location in USA to set up a gaming arcade:-

CEO of a Gaming company wants to set up a gaming arcade in USA.

Where would be the best place to install an arcade?

INTRODUCTION:

An arcade game or coin-op game is a coin-operated entertainment machine typically installed in public businesses such as restaurants, bars and amusement arcades. Most arcade games are video games, pinball machines, electro-mechanical games, redemption games or merchandisers.

Funspot Family Entertainment Center (or simply Funspot) is a video arcade which features one of the largest collections of late-1970s to mid-1980s games in the world. It is located in the village of Weirs Beach in Laconia, New Hampshire, United States.

PROBLEM AND INTEREST :

A CEO of a Gaming company wants to set up a gaming arcade in USA.

Considering all the constraints pertaining to the data and weights for each category of the venue, *Where would be the best place to install an arcade?*

Business Understanding:-

Installing a Gaming arcade 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.

CEO is interested in starting a gaming arcade in the best locality of all the cities in United states. He defines a best locality based on the following constraints :-

- Population density of a locality

- Per Capital income

- Population of each location

- Venues in each locality

- The category of the venues that he's interested in are,

- Arts and Entertainment

- Shops & Service

- College and University

- Event

- Food

- Nightlife Spot

- Outdoors & Recreation

- Professional & Other places

- Residence

- Travel & Transport

A Brief about the Data:-

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

In order to set up a gaming arcade,he is using below sources:-

- List of all the cities in United States with population density and coordinates:
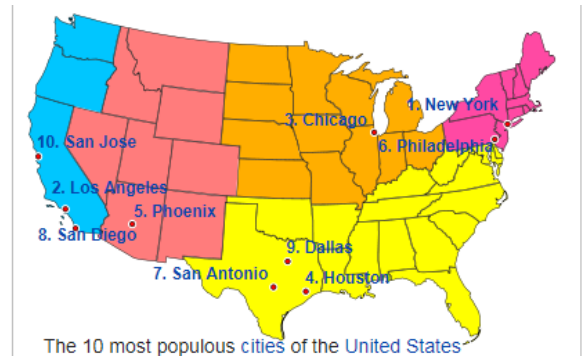 https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population

  The list is the most populous incorporated places of the United States. As defined by the United States Census Bureau, an "incorporated place" includes a variety of designations, including city, town, village, borough, and municipality.[a] A few

exceptional census-designated places (CDPs) are also included in the Census Bureau's listing of incorporated places.[b] 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. This list presents only that portion (or "balance") of such consolidated city-counties that are not a part of another incorporated place.

This list refers only to the population of individual municipalities within their defined limits; the populations of other municipalities considered suburbs of a central city are listed separately, and unincorporated areas within urban agglomerations are not included. Therefore, a different ranking is evident when considering U.S. metropolitan area populations.

The following table lists the 314 incorporated places in the United States with a population of at least 100,000 on July 1, 2018, as estimated by the United States Census Bureau. A city is displayed in **bold** if it is a state or federal capital, and in *italics* if it is the most populous city in the state. Five states—Delaware, Maine, Vermont, West Virginia and Wyoming—have no cities with populations of 100,000 or more. The table below contains the following information:
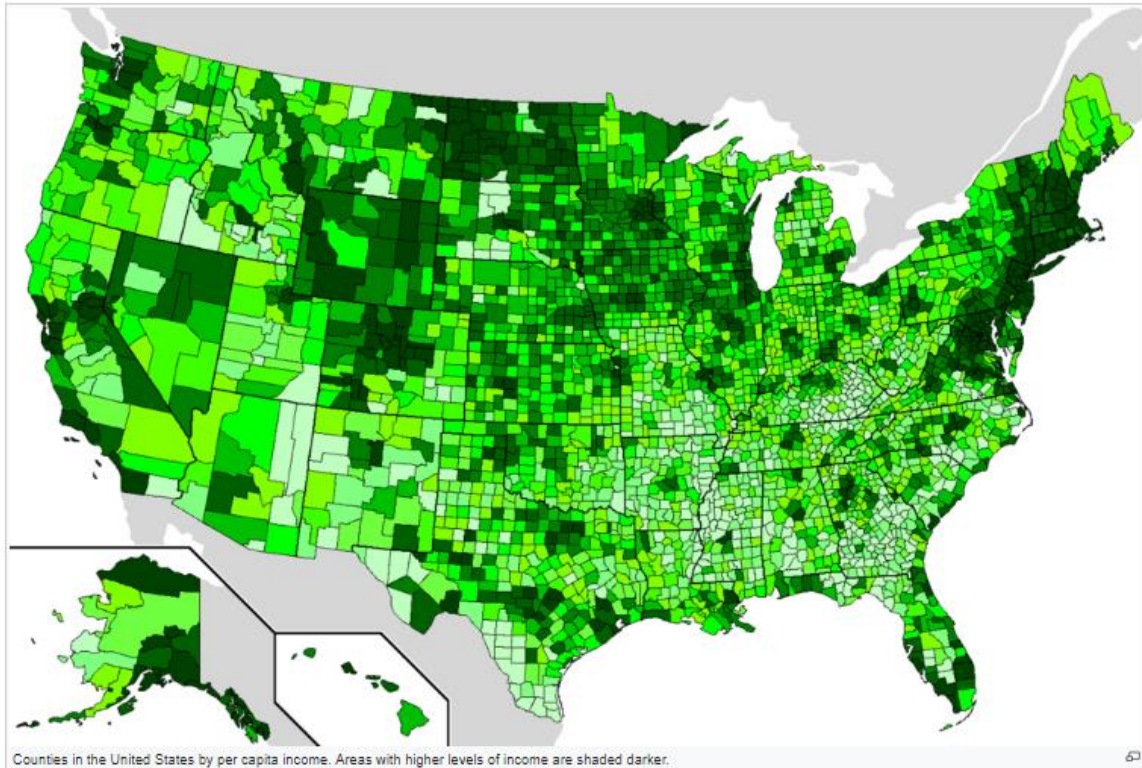
1. The city rank by population as of July 1, 2018, as estimated by the United States Census Bureau[1]
2. The city name[1]
3. The name of the state in which the city lies[1]
4. The city population as of July 1, 2018, as estimated by the United States Census Bureau[1]
5. The city population as of April 1, 2010, as enumerated by the 2010 United States Census[1]
6. The city percent population change from April 1, 2010, to July 1, 2018
7. The city land area as of January 1, 2016[2]
8. The city population density as of July 1, 2016 (residents per unit of land area)[2]
9. The city latitude and longitude coordinates[2]

The 10 most populous cities of the United States

- List of all the cities in United States with Per Capita Income :
  https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income

This is a list of United States counties by per capita income. Data for the 50 states and the District of Columbia is from the 2009-2013 American Community Survey 5-Year Estimates; data for Puerto Rico is from the 2013-2017 American Community Survey 5-Year estimates, and data for the other U.S. territories is from the 2010 U.S. Census.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.

Counties in the United States by per capita income. Areas with higher levels of income are shaded darker.

Methodology:-

- Using the above data we will first select best city to proceed with based on the values like

- Population density

- per capita income of the state

- number of venues (as we are giving weights to each venue based on its category).

Once we select a city, we then go hunting for Localities. Again, we do it using the same approach i.e. based on the scores of venues in each locality.

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 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 to setup a gaming arcade.

Based on the analysis we have done, the following is the result we have got:

Plot to show the location of final arcade that we are suggesting:

Discusions and further 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.

Conclusions:

Based on the given constraints, a gaming arcade can be installed in a place closer to the center of the circle to attract more number of diverse customers and get huge revenue.

Finally got a better place in the Jersey city

This place is between the Groove Street and the Grand Street, hence it would also have the best footfall and potential customers as well.

Resources :

List of all the cities in United States with population density and coordinates:
https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population

List of all the cities in United States with Per Capita Income :
https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income

Four Square API : https://foursquare.com