Coursera Capstone

"The Battle Of The Neighborhoods"

Which is the best venue for opening a new mall/casino in USA?

INTRODUCTION

Shopping malls and casinos are the trademarks of American culture. In daily basis we all take a trip to malls for fulfill particular needs or entertain ourselves. Also casino's are the one of the biggest part of US's entertaintmet sector.

My client ownes mall/casiono chains in Greece and he wants to open a new mall/casino in US. Fort he succes of his new venture, location is the key concept. My client has no prior information about mall / casiono sector or demographics in USA.

PROBLEM BROUGHT UP BY CLIENT

Which City and Locality is best to open a Casino/Shopping mall in USA?

ANALYTICAL APPROACH

According to the need of client first we have to find the City and then the Locality in that city for Business. To solve this particular problem and to gather data we will set few Benchmarks for the City and the locality in which opening a Casino/Shopping Mall would be Successful.

For discovering the City following benchmarks must meet
- Nightlife
- Services
- Transport
- Residence
- Universities
- Food
- Entertainment
- Travel
- Outdoor Activities

Benchmarks for the Locality in City
- Per Capita Income
- Population

- Population Density
- Venues

DATA REQUIREMENTS AND DESCRIPTION

- List of all the cities in United States with population density and coordinates
- List of all the cities in United States with Per Capita Income
- List of all venues in each city
- List of all venues in each locality in the selected city

SOURCE OF DATA

https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population

https://en.wikipedia.org/wiki/List_of_United_States_counties_by_per_capita_income

Foursquare API

DATA DESCRIPTION

By Using the Wikipedia data we will first analyse the data and then focus on certain benchmarks to discover the best City in USA to open a Casino/Shopping mall. Then by leveraging the Foursquare data we will compare and try to find the best locality. Based on Scores we will try to determine the Locality in simillar way.

About the State income levels: State income levels and income data for the United States as a whole areincluded for comparison. Note that county-equivalents in Louisiana are called "parishes" and in Alaska are called in "boroughs," and alsothat 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
•1. Using Beautiful Soup Library Wikipedia pages containing information about Cities of USA by Population and Per Capita income are scraped into Pandas Dataframe.

•2. Dataframecontained data about Cities, Coordinates, Area, Per capita Income and Population Density. ThanDataframewas Cleaned and Processed according to requirement of the problem to be solved. Proper benchmarks were set to obtain the best results.

•3. The List of Venues in a City were obtained using Foursquare API and the city with maximum weight according to the Model is selected.

•4. With the help of Unsupervised Machine Learning Algorithm (K Means Algorithm) Locality in the city is obtained where opening a Shopping Mall/Casino will be most benefitted.

RESULTS

Based on analysis we have done the results we obtain are –The plot shows the final location predicted by the Machine Learning Model –
•For opening the Shopping mall/Casino

•The Best city in USA –New Jersey

•The best locality in New Jersey –Jersey Avenue