Analysis of Land Values and Local Incomes

Bernard Beck

5.10.2021

Introduction

Problem

You work for a country club development firm and have been tasked to determine where within the United States the next destination should be placed. The next facility will include a golf course, clubhouse, tennis facilities, swimming, gym, dining, as well as other operational infrastructure. The development area will be substantial. Care should be taken to locate areas within the United States offering the most reasonable land investment costs in coordination with a local customer base with discretionary income for country club membership.

Interested Parties

Parties with vested interests in this search are numerous.

Firstly, our company stakeholders will be influenced by the financial impact of this process. The growth of our firm is determined by the popularity and prosperity of our clubs. As such, the revenues are largely determined by individual club success.

Prospective employees would of course have interest in where our next club is located. A new club means new local jobs.

Akin to employees, local individuals will also comprise the membership group.

The construction of a new club will require several subcontractors to facilitate the build. This can include local, regional, and national entities to facilitate erection of the clubhouse, golf course, tennis areas, pool and cabanas, gymnasium, dining hall, laundry services, cleaning facilities, kitchen, office structures, and much more.

Data Sources

Data collected to analyze the most primed regions of the country for our next facility include income statistics and property values. Income information was obtained from the US Census Bureau while land costs were derived from the American Enterprise Institute.

Additionally, a nationwide zip code database was also utilized to help facilitate analysis of the two data-oriented sets.

<https://data.census.gov/cedsci/table?q=median%20income&g=0100000US.860000&tid=ACSST5Y2019.S1901&hidePreview=true>

<https://www.aei.org/housing/land-price-indicators/>

<http://federalgovernmentzipcodes.us/download.html>

Data Cleaning

Both the land and income datasets were available online as CSV files. These files were downloaded into Python and imported into separate dataframes.

Within the Census income data, areas were named with an arbitrary string followed by the corresponding zip code. This string value was dropped from the name cells, revealing the individual zip codes. A superfluous header row was dropped from the set as well. Income data was cleaned until the result included the zip code column, median annual income, and mean annual income.

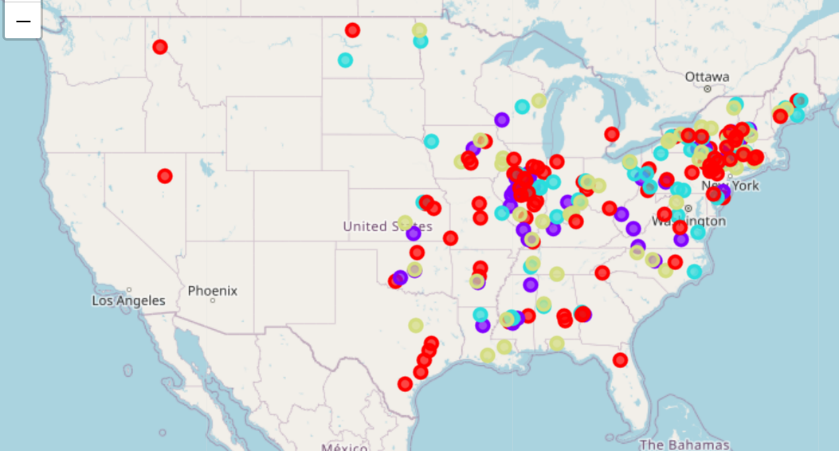
Land Data underwent a similar process. Initially, the set included land values for several years for each available zip code. This data was restricted to the most recent year, 2019. Columns provided no value were dropped. Zip codes with leading zeros lost their initial digits, which were added back utilizing zfill.

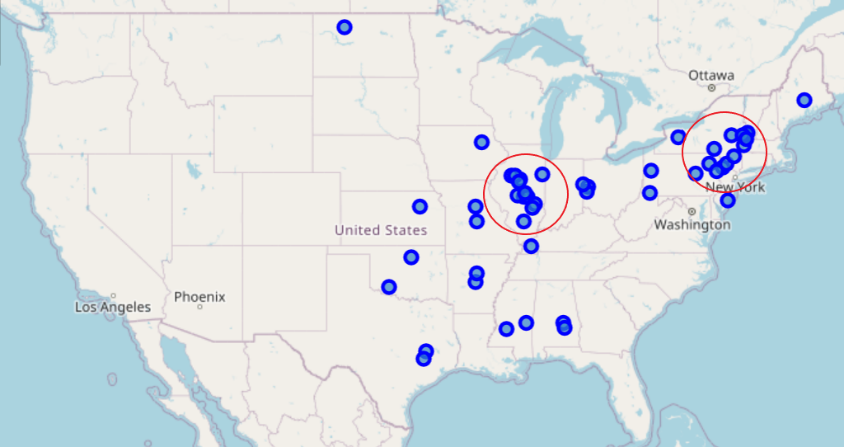
Data from both dataframes was combined upon the zip column as key. Some data within the land value dataset included for outliers and negative land values, which were removed.

Methodology

A new column was created, dividing the Average Cost per Acre by the Average Annual Income. This division gives us a relative suitability score per zip code. A filter was then created focusing upon only those zip codes whose population’s average income exceeded the median value ($78126).

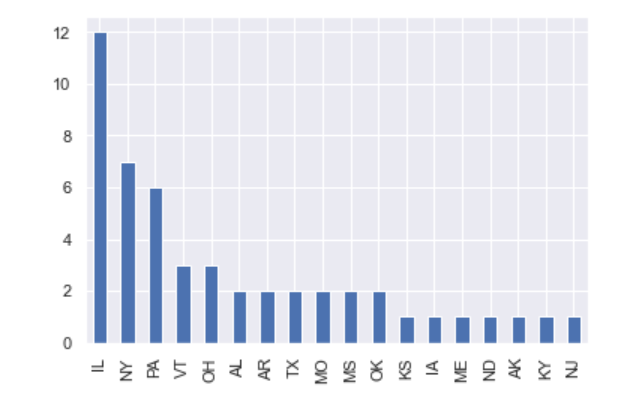
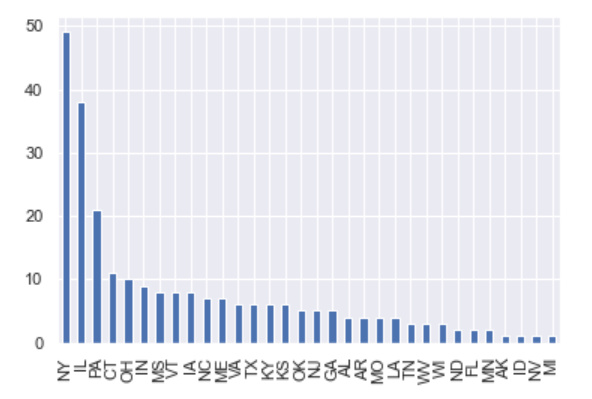
The top 50 and top 250 zip codes by suitability score were extracted to initiate our analysis. These zip codes were mapped according to latitude by longitude and correlated via state.





Results

Once the suitability score was derived, areas were mapped by strongest suitability index. These scores were also binned by state. New York and Illinois dwarfed the competing states in both number of zip codes residing in the top 50 and top 250. It was widely apparent that these two states provided the most extreme values based on our suitability index.

Further mining displayed that Illinois contained 12 (24%) of the top 50 zip codes by value. Of these values, 8 are within the Bloomington-Peoria-Springfield, Illinois region. 16% of the strongest values located in this tri-city area alone. 

Discussion & Conclusion

Once again, 8 units, or 16%, of the most valuable zip codes by suitability are included in the Bloomington-Peoria-Springfield, Illinois region. This includes the #6, #7, and #9 ranked zip codes overall.

