

The Battle of the Cities – Report

Author: Nuno Roberto

This report was prepared to complete the IBM Data Science Capstone Project requirements and offered online by Coursera. The project utilizes a similar approach to that of the Battle of the neighborhoods (NEIs) insofar as to utilize similar information to that collect in the labs and projects completed in the certification. These sources are publicly available GIS data sources, and Foursquare location based points of interest.

1. Introduction and Business Problem

I live and work in California but I “play” in Colorado, Utah, and Arizona, ie , every time I want to have some fun I drive all the way to some remote location in Colorado or Utah and go for hikes in Bryce Canyon, Arches, Zion, Capitol Reef, Grand Canyon, or one of the many awesome National Parks in any of these states.

It is always summer in California and it’s great from a weather’s point of view, if you like it but there’s no seasons. We pay high taxes in almost every aspect of our lives – sales, car, home, state, business, you name it we have it and it’s always high taxes.

When both my wife and I retire, I want to move to a state that has all the fun things we like to do, but a state that provides a much better cost of living, much better outdoor activities, low on pollution, has well defined seasons year-round, it is great for foodies, and we don’t have to drive for miles just to get a carton of milk.

Telluride is nestled between mountains and has a great yearly film festival attended by famous actors and actresses, who also own their homes outside town (NOTE: the median home price in Telluride is \$902,000, likely skewed by famous homes sold at over \$10million!!)

Telluride has winter resorts for locals and visitors alike and it’s extremely popular as a winter destination.

Telluride still has the old-west feeling, while catering to people who want to retire in a modern town with small businesses, and amenities that support a great community year round.

In the battle of the NEIs we analyzed points of interest between Toronto and NY. These are big cities, with plenty of information available on FourSquare. How will FourSquare handle information not at a neighborhood level but at a city or town level. How will it handle API calls that pull data based on my personal criteria (weather, food, supermarkets, outdoor fun, and trails)?

Once this project is completed, I expect to confirm my assertion about Telluride based on my personal criteria.

1.1. Problem Description

Identifying a town in the US to retire is a daunting task for anyone, and more so if you don’t have the right resources available to you. The retiring process is complex and it has many, many pitfalls, related to location, amenities, costs of living, medical facilities, and so on.

Availability of data and research are crucial in identifying a state and a town to retire and provide a smooth transition from active careers to a retirement phase of our lives.

The initial problem to resolve in this phase of the project is to identify the public data sources that have enough information to collate with Foursquare and analyze using one of several machine learning approaches.

Even data from universities is not bullet-proof and during my initial research I found several data points that were not available or had to be highly curated before I could use them (or not at all).

The criteria used to select the best town to retire might not be appropriate for every situation. For my use case I first had to check Foursquare information and verify if data for at least a small group of towns in Colorado was available both in Foursquare and in a GIS JSON file of Colorado towns.

The Weather data, although provided by NOAA and displayed in Wikipedia, is not used at all by some towns. This is primarily because NOAA does not have weather stations to collect data at certain locations in the US. To track the weather in small towns in the US one has to rely on aggregated weather information from larger cities (eg, Boulder, CO). The problem with this approach is that weather in mountain towns is often localized and it expresses microclimates that are not captured by weather stations.

Since weather was one of the major categories in my selection criteria to identify a town for retirement, several interesting cities could not be include in the analysis and recommendations because of a lack of sufficient weather data.

Finally, Foursquare is a site that relies on location data to show interesting points in a map, but it relies heavily on availability of wireless connectivity and precise cell phone signal and that's not always available in remote areas of Colorado. So, if I want to retire to Telluride, CO, known as a small true-american town nestled between mountain peaks, far from civilization, good luck to me finding data to analyze the assertion that Telluride is the best overall town to retire in Colorado.

1.2. Audience

Anyone who wants to retire in Colorado based on similar criteria used to analyze Colorado cities and communities.

Although I'm working on this project for my own purposes, anyone who comes from a state with high income taxes, and high property taxes, will benefit from this research.

Additionally, it will be easy to change the criteria based on other people's preferences and as long as the criteria exists in the data collect from public sources for Colorado cities.

1.3. Success Criteria

The success criteria of the project will confirm or invalidate the assertion that Telluride, CO is one of the best towns in Colorado to retire based on the preferences highlighted in the problem description. Although Telluride, CO is assumed to be the best town to retire, we need to be open to the fact that there are other more viable options based on the same criteria defined in the problem description.

2. Data

Data will be loaded from publicly available GIS data sets maintained by Stanford University, Harvard University, and Colorado State GIS Repository.

Note that the use of this research and the use of each site mentioned in this report is not a sponsorship of the website, the company or services offered there, and I did not receive any sponsorship from anyone to do this analysis.

Foursquare items of interest

We will look at cities in Colorado and compare key elements to help decide where to retire:

- Weather (NOAA / Wikipedia)
- Food (4Square)
- Trails (4Square)
- Supermarkets (4Square)
- Outdoor fun (4Square)

Websites used for initial research

My initial assumption was that Telluride would be a great place to retire and as such I started researching Telluride, CO, using the **Best Places**

This site offers simple analysis on each place population, unemployment rate, median income, median age, median home price, and comfort index. It also provides comparisons between two places on costs of living, real estate, crime, climate, schools, economy, health, religion, politics, commute, and so on.

- ASSERTION: Preferred City to Retire
 - Telluride: https://www.bestplaces.net/cost_of_living/city/colorado/telluride

I also used the site SmartAsset to find out additional information about Colorado and specifically reasons to move to Colorado (<https://smartasset.com/mortgage/15-things-to-know-before-moving-to-colorado>). Note that I did not use this site directly in my data analysis but rather to ground the research on specific cities and aspects that I wanted to analyze for each town in Colorado.

The normalization and scope of the project was defined by a bit more research using the site SmartAsset to identify the top 10 cities to retire in Colorado (<https://smartasset.com/retirement/best-places-to-retire-in-colorado>)

Although I started with 10 cities I ended up with only 7 cities since some cities did not have enough data on weather, or population demographics or failed to report on one of the categories listed in my criteria.

List of cities used to compare NEIs and that fit retirement criteria

Sorted by ascending order of overall quality (as defined in the SmartAsset site):

- 1 - Littleton: <https://www.bestplaces.net/city/colorado/littleton>
- 2 - Englewood: <https://www.bestplaces.net/city/colorado/englewood>

- 3 - Evergreen: <https://www.bestplaces.net/city/colorado/evergreen>
- 5 - Estes Park: https://www.bestplaces.net/city/colorado/estes_park
- 6 - Monument: <https://www.bestplaces.net/city/colorado/monument>
- 8 - Salida: <https://www.bestplaces.net/city/colorado/salida>
- 10 - Montrose: <https://www.bestplaces.net/city/colorado/montrose>

Available Weather data in Wikipedia:

- 0 - Telluride: https://en.wikipedia.org/wiki/Telluride,_Colorado
- 1 - Littleton: https://en.wikipedia.org/wiki/Littleton,_Colorado
- 2 - Englewood: https://en.wikipedia.org/wiki/Englewood,_Colorado
- 3 - Evergreen: https://en.wikipedia.org/wiki/Evergreen,_Colorado
- 5 - Estes Park: https://en.wikipedia.org/wiki/Estes_Park,_Colorado
- 6 - Monument: https://en.wikipedia.org/wiki/Monument,_Colorado
- 8 - Salida: https://en.wikipedia.org/wiki/Salida,_Colorado
- 10 - Montrose: https://en.wikipedia.org/wiki/Montrose,_Colorado

Each location or city listed above contains weather data from NOAA and listed on Wikipedia as below:

Climate data for Telluride 4WNW 1981-2010, extremes 1900 to present														[hide]
Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year	
Record high °F (°C)	58 (14)	65 (18)	73 (23)	78 (26)	83 (28)	91 (33)	96 (36)	91 (33)	88 (31)	83 (28)	73 (23)	60 (16)	96 (36)	
Average high °F (°C)	37.5 (3.1)	39.5 (4.2)	44.5 (6.9)	52.9 (11.6)	62.6 (17.0)	73.2 (22.9)	77.5 (25.3)	74.7 (23.7)	68.1 (20.1)	58.5 (14.7)	45.5 (7.5)	37.2 (2.9)	56.0 (13.3)	
Average low °F (°C)	4.5 (-15.3)	8.9 (-12.8)	15.7 (-9.1)	23.7 (-4.6)	30.8 (-0.7)	36.1 (2.3)	42.2 (5.7)	42.3 (5.7)	35.2 (1.8)	25.6 (-3.6)	15.5 (-9.2)	5.8 (-14.6)	23.9 (-4.5)	
Record low °F (°C)	-32 (-36)	-36 (-38)	-24 (-31)	-10 (-23)	3 (-16)	15 (-9)	26 (-3)	22 (-6)	9 (-13)	0 (-18)	-22 (-30)	-31 (-35)	-36 (-38)	
Average precipitation inches (mm)	1.49 (38)	1.58 (40)	2.02 (51)	1.94 (49)	1.81 (46)	1.12 (28)	2.50 (64)	3.02 (77)	2.68 (68)	1.92 (49)	1.87 (47)	1.46 (37)	23.41 (594)	
Average snowfall inches (cm)	27.7 (70)	26.3 (67)	34.7 (88)	20.0 (51)	8.3 (21)	0.4 (1.0)	0 (0)	0 (0)	1.0 (2.5)	8.6 (22)	24.4 (62)	25.6 (65)	177 (449.5)	
Average precipitation days (≥ 0.01 in)	10.0	10.1	11.8	10.5	8.9	6.9	13.5	16.6	12.2	9.3	8.9	10.0	128.7	
Average snowy days (≥ 0.1 in)	9.8	9.4	10.4	6.8	2.3	0.1	0	0	0.4	2.9	7.2	9.7	59	

Source: Nowdata, Telluride 4WNW from Grand Junction area

Publicly available GIS Data Sources

There are several sources of GIS information for each state in the US but I prefer to use the following data sources:

- Cities and Towns of the United States, 2014
 - <https://geo.nyu.edu/catalog/stanford-bx729wr3020>
 - Includes a downloadable JSON data source with coordinates for each city in the united states as of 2014
- Colorado Department of Local Affairs
 - <https://demography.dola.colorado.gov/gis/gis-data/#gis-data>
 - Includes different aspects of each county, borders, districts, and locations for each city. Updated often.
- US Department of Commerce – Colorado View
 - <https://www.coloradoview.org/united-states-gis/>
 - Lists several census data sources organized in different types of data files, GIS, ShapeFiles, Raster files, and others
- New York University, Spatial Data Repository

- https://geo.nyu.edu/?f%5Bdc_rights_s%5D%5B%5D=Public&f%5Bdct_spatial_sm%5D%5B%5D=Colorado
- NYU maintains a repo of spatial data aggregated from other universities and state departments. The data can be pulled in different formats, including JSON, shape, raster, or TIFF.