Battle of the Cities

Where to retire in Colorado, USA

Introduction

- I live and work in California but I "play" in Colorado, Utah, and Arizona

- It is always summer in California. 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

- The state has to 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

Introduction
Telluride, CO

- Telluride is nestled between mountains

- 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.

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

Business Problem

- 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)?

Problem Description

- Identifying a town in the US to retire is a daunting task for anyone

- Availability of data and research are crucial in identifying a state and a town to retire

- Identify the public data sources that have enough information to collate with Foursquare and analyze using k-means

- Even data from universities is not bullet-proof

- The criteria used to select the best town to retire might not be appropriate for every situation.

- The Weather data, although provided by NOAA and displayed in Wikipedia, is not used at all by some towns.

- Since weather was one of the major categories in my selection criteria, several cities could not be included in the analysis and recommendations because of a lack of sufficient weather data.

- Foursquare is a site that relies on location data to show points in a map, but it relies heavily on availability of wireless connectivity and precise cell phone signal.

- Telluride, CO, known as a small true-american town nestled between mountain peaks, far from civilization, good luck to me finding data to analyze

Selection Criteria

- 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)

Evaluation Dataset

- 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

GIS / GEO Datasets

- 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.
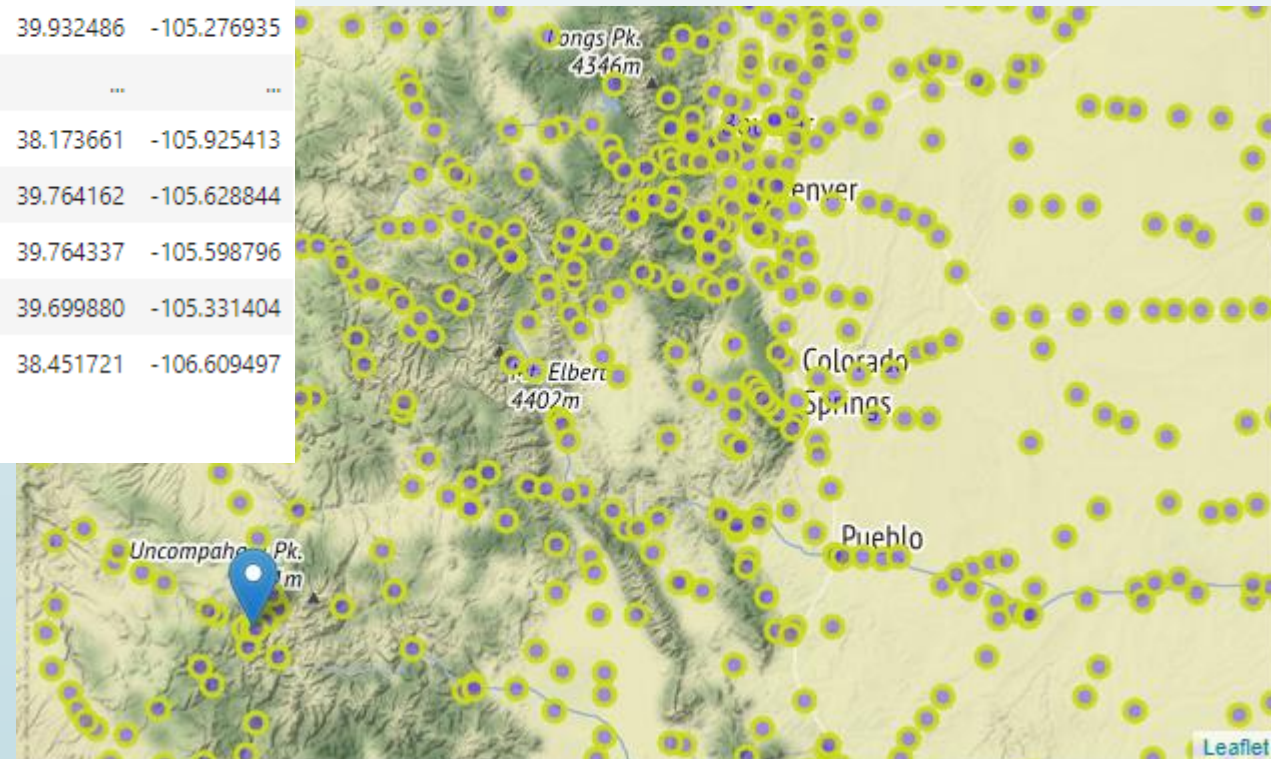
Methodology - Approach

1. Get a list of all cities in the US
2. Create a new dataframe with just the cities in colorado
3. Find Colorado on the map
4. Position Telluride and all cities in Colorado
5. Add markers for all other cities listed as best cities to retire to in Colorado
6. Sort the cities by population in ascending order
7. Add all the points of interest to the map
8. Weather Analysis
9. Start analyzing temperature data for Telluride
10. Start analyzing Precipitation data for Telluride
11. Play with Seaborn library

12. Get top 100 points of interest in Telluride within a radius of 15 miles
13. Finding points of interest based on my criteria
14. Create a function to analyze all cities of interest in Colorado
15. Analyze unique categories that can be found from all the returned venues
16. Analyze each City - ONE HOT Analysis
17. Group rows by city and take the mean of the frequency of occurrences of each category
18. Begin the CLUSTERING Process using k-means
19. Create a new dataframe that includes the cluster as well as the top 10 venues for each city
20. Visualize clusters
21. Examine Clusters

Methodology – Colorado Analysis

Methodology – Criteria Cities

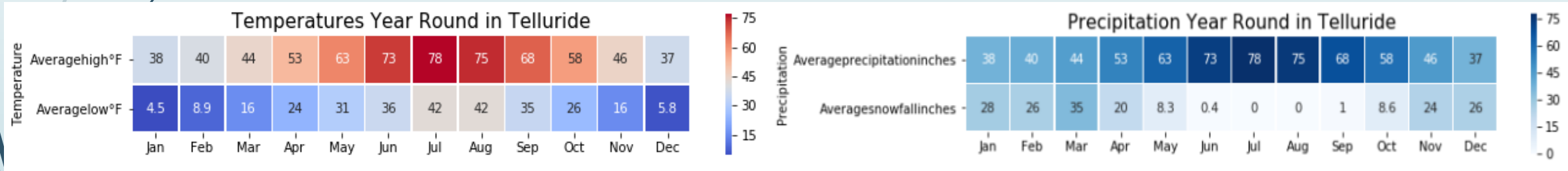|	index	City	State	Population	County	Fips	Latitude	Longitude
87	20214	Littleton	CO	41737	Arapahoe	005	39.613321	-105.016650
85	20212	Englewood	CO	30255	Arapahoe	005	39.647765	-104.987760
98	20252	Montrose	CO	19132	Montrose	085	38.478320	-107.876174
11	17539	Evergreen	CO	9038	Jefferson	059	39.633321	-105.317215
130	21383	Estes Park	CO	5858	Larimer	069	40.377206	-105.521665
267	22475	Monument	CO	5530	El Paso	041	39	
179	21474	Salida	CO	5236	Chaffee	015	38	
292	22537	Telluride	CO	2325	San Miguel	113	37	

Based on the 7 cities under review we created a table that lists out all the retirement-ready cities in Colorado, sorted by population

Methodology – Weather Analysis

	Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	Averagehigh°F	37.5	39.5	44.5	52.9	62.6	73.2	77.5	74.7	68.1	58.5	45.5	37.2
2	Averagelow°F	4.5	8.9	15.7	23.7	30.8	36.1	42.2	42.3	35.2	25.6	15.5	5.8
4	Averageprecipitationinches	1.49	1.58	2.02	1.94	1.81	1.12	2.50	3.02	2.68	1.92	1.87	1.46
5	Averagesnowfallinches	27.7	26.3	34.7	20.0	8.3	0.4	0	0	1.0	8.6	24.4	25.6

Temperatures Year Round in Telluride

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Averagehigh°F	38	40	44	53	63	73	78	75	68	58	46	37
Averagelow°F	4.5	8.9	16	24	31	36	42	42	35	26	16	5.8

Precipitation Year Round in Telluride

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Averageprecipitationinches	38	40	44	53	63	73	78	75	68	58	46	37
Averagesnowfallinches	28	26	35	20	8.3	0.4	0	0	1	8.6	24	26

- Telluride shows moderate summer temperatures between Jun and Sep
- Telluride shows heavy snow fall during winter months between Nov and Apr

Methodology – Aggregation of POIs

- Based on the criteria we created a table that lists similar points of interest (POIs) in each city, creating a new dataframe with all the results merged

[64]:

	City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Telluride	37.937494	-107.812285	Brown Dog Pizza	37.937042	-107.810671	Pizza Place
1	Telluride	37.937494	-107.812285	Taco del Gnar - Telluride	37.937216	-107.813271	Taco Place
2	Telluride	37.937494	-107.812285	The Butcher & Baker Cafe	37.936921	-107.809715	Sandwich Place
3	Telluride	37.937494	-107.812285	La Marmotte	37.935731	-107.812275	French Restaurant
4	Telluride	37.937494	-107.812285	New Sheridan Hotel & Chop House	37.937689	-107.812670	Hotel
...	...	...	...	...	...	...	...
223	Montrose	38.478320	-107.876174	Sweet Bites Bakery	38.480149	-107.875576	Bakery
224	Montrose	38.478320	-107.876174	Sams Tavern	38.479927	-107.877550	Dive Bar
225	Montrose	38.478320	-107.876174	Elevate Salon	38.480019	-107.878447	Health & Beauty Service
226	Montrose	38.478320	-107.876174	Backstreet Bagel & Deli	38.479873	-107.879339	Bagel Shop
227	Montrose	38.478320	-107.876174	Crash Burger	38.481299	-107.873507	Burger Joint

228 rows × 7 columns

Methodology – Clusters w/ Final Selection

- Telluride is great but we need to compare it to other cities in the CO.

- At this point data was already massaged, filtered, and aggregated by interests,

- We used 1-hot analysis, followed by clustering using k-means with 5 clusters

- In hindsight I probably should've used 3 or even 2 clusters to aggregate potential cities similar to Telluride in a "tighter" configuration

City	Fips	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
Telluride	113	37.937494	-107.812285	0	American Restaurant	Hotel	Bar	Pizza Place	Gift Shop
Littleton	005	39.613321	-105.016650	0	Restaurant	Gourmet Shop	Coffee Shop	Mexican Restaurant	Bar
Evergreen	059	39.633321	-105.317215	0	Bar	Restaurant	Candy Store	Breakfast Spot	Pizza Place
Estes Park	069	40.377206	-105.521665	0	Gift Shop	American Restaurant	Ice Cream Shop	Mexican Restaurant	Arts & Crafts Store

Conclusion

- Based on the results of different criteria (weather, population, restaurants, hiking trails, supermarkets, and outdoor activities) it is clear to see that Telluride is still a great town to retire

- This analysis is biased towards Telluride

- In the real world we would need to remove bias and include a much larger dataset (possibly ALL cities in Colorado) and a much larger set of features to analyze.

- It is possible that using different clusters would also yield different results and point to a different conclusion.

- The weather data and demographics for each city should be included as a relevant characteristic in the cluster formation (eg, if it snows in Telluride on average x inches, how does that compare to other cities further east or further south).