Comparison of Ski Resort Towns in the Western U.S.

And

Housing Prices

Ed Lees

January 15, 2020

Contents

Introduction / Problem to be Solved

I would like to someday purchase a vacation home in a western U.S. ski resort town with the idea of eventually retiring there. But with so many towns to potentially choose from how does one go about choosing where to buy? Certainly, one could consider various subjective factors, but one needs to seriously consider many data driven objective factors to make a wise and informed choice. If a person were to purchase a home based only on how much he enjoyed skiing a certain mountain that person might find that he had made a major mistake if he later became bored or otherwise dissatisfied with the town due to bad public services or a generally poor variety of stores, restaurants or other things to do when not skiing.

This same problem applies to not only me, but to anyone looking to move either full or part time to a ski town. In fact, a very similar data driven methodology could be applied to anyone considering moving who can choose among various comparable destinations. With the aging of the population more and more people will be ready to make such a move and could benefit from looking at data systematically before making any decisions.

Therefore, my project will be gather data from the Foursquare database as well as data on available skiing and general demographic and economic data on each of several ski towns, I will normalize the data to make it more easily comparable, weight certain variables according to my preferences and derive an objective numerical rating for each town based on the data. This analysis could be expanded to include additional variables and locations then published or used as the basis for an interactive program that would allow users to weight various factors according to their personal preferences to return suggested destinations.

Data

I will be using 3 data sets to analyze the relative attractiveness of the towns under consideration.

1. The Foursquare data to look at the number and variety of different venues in each town.
2. Quality of Life data which I will gather primarily from City-data.com, a website that contains a plethora of data on every city and town in the U.S. Additionally I will gather data on relative quality of the local schools from a USnews.com study of American high schools.
3. Statistics on local skiing mountains will be amalgamated from each mountain's own website.
4. Data on median home prices was collected from both the city-data.com website and Zillow. Data on Average Listing Price per Square Foot comes from research section of Zillow's website

Data from the various websites will be collected and then manually entered into a data frame in my code. Foursquare data will be downloaded directly from Foursquare and initially placed in its own data frame.

Data will be collected for 10 towns in the Western U.S.

1. Aspen, Colorado
2. Crested Butte, Colorado
3. Incline Village, Nevada
4. Jackson, Wyoming
5. Park City, Utah
6. Steamboat Springs, Colorado
7. Sun Valley, Idaho
8. Taos, New Mexico
9. Telluride, Colorado
10. Vail, Colorado

Foursquare Data

After downloading the information on the venues for each town the data will be added to single data frame containing all the foursquare data. After looking at the ten most frequently occurring venue types in each town I will consolidate the data into five broad categories by general venue type for easier analyses.

The five board categories that the venues will be divided into are:

1. Restaurants
2. Shopping
3. Recreation
4. Entertainment
5. Other

Quality of Life Data

The quality of Life data taken from city-data.com will come from five categories.

1. Average January Temperature (I prefer not to be too ridiculously cold!)
2. Air Quality Index
3. Average Sunny Days per Year
4. Cost of Living Index
5. Quality of Schools
6. Crime Rate

All the above statistics except for the school quality data are sourced from city-data.com. the School data comes from USnews.com/education/best-high-schools a study that attempts to rank every high school in America from 1 to 17,245. I assume that the quality of education in the primary schools is going to be roughly analogous to the high school. To calculate a quantitative value for education quality I take the ranking of the local high school according to U.S. News for each town and express that ranking as a percentile of the entire data base. For example, Aspen High School in Aspen Colorado is ranked # 1,430 in the U.S. In my data I express this as 1-(1,430/17,245) = .917.

Skiing Data

Data on available skiable terrain will be amalgamated from each mountain's website.

To evaluate the quality of the skiing experience at each mountain I will explore 6 variables.

1. Total Skiable Acres - (this includes any ski area within a 45-minute drive (according to Google) – the maximum distance I would be willing to sometimes commute for the sake of variety.)
2. Ski In/Out Acres - The number of skiable acres immediately accessible from the town.
3. Vertical – the largest vertical distance from the top to the bottom of any of the commutable skiing areas.
4. Lifts – the total number of lifts in all commutable ski areas combined.
5. Average Annual Snowfall – the more the better for skiing.
6. Lifts per Total Skiable Acres – derived from 1 and 4 above, should provide a general indication of the length of lift lines.

Data Cleaning

There is no missing data in the data sets I have chosen except some historical home price data for some towns. This missing data is not needed for my analyses as I only need the most current prices. I have included a chart of historical prices as additional interesting information that a real estate investor should be aware of, so the missing data is not in the chart, but this is of no consequence to my ratings.

Three data sets, Air Quality Index, Cost of Living Index and Crime Rate in the Quality of Life subset need to be inverted before normalizing as lower numbers are better for these three variables.

Methodology

The primary purpose of this research is to arrive at data driven score or rating for each town to determine which is the most attractive to live in. A secondary objective is to compare the ratings of each town to local home values to determine the best relative value for a real estate investment. The intention is not to derive a predictive model for real estate values based on the external factors I am examining. Furthermore, with only ten towns in the data the accuracy of a predictive model would be highly suspect. So, the objective is to create a score that will attempt to summarize in a single number the desirability of each town so they can be compared to each other. Then we can divide median housing cost by that number to determine the town with the most attractive real estate values.

For the Foursquare venue data, after amalgamating the various venue types into 5 broad categories as described in the data section above, each category will be grouped into its own data frame, then the towns will be ranked by category according to the number of venues in each town of each category. The towns will then by ranked in each category according to the number of venues of each type after normalizing each category from 0 to 1. The Restaurant and Recreation categories will be double weighted giving a total of 7 points possible from the Foursquare venue data. I am giving double weighting to restaurants because I like variety and the more restaurants you have to choose from the greater likely hood that a few will be really excellent. Choices for recreation is very important because I will want to keep up an active lifestyle 12 months out of the year, not just during ski season. Furthermore, a large number of restaurants should indicate a vibrant economy.

For the Quality of Life data, like the Foursquare data, the raw numbers in each category are normalized from 0-1. Air Quality Index, Cost of Living Index and Crime rate are indices in which lower numbers indicate a better situation, so I need to invert those statistics before normalizing them. I will double weight the crime data as that, to me, seems the most important category. I will then sum the scores for each town to arrive at a total Quality of Life score for each town. Like the Fourscore venue data, the maximum score for a town from this data set is 7.

For the skiing data I will follow the same general methodology as in the first two data sets. I will normalize each category among the different towns. The normalized value of Total Skiable Acres will be given double weight within this group giving a total of a maximum of 7 points for this data set, the same as the previous two groups.

Final Score for Each Town

Finally, the total column from each of the above 3 data frames will be placed in a final data frame and points summed to create a total score for each town to find which town is objectively most appealing based on the selected data and methodology. The maximum possible score will be 3 x 7 = 21, but I don't expect any town to come close to this. I will also then look at median home values (Also from city-data.com) to find which town represents the best real estate value.

Results

Following the above methodology, I arrived at the following results:

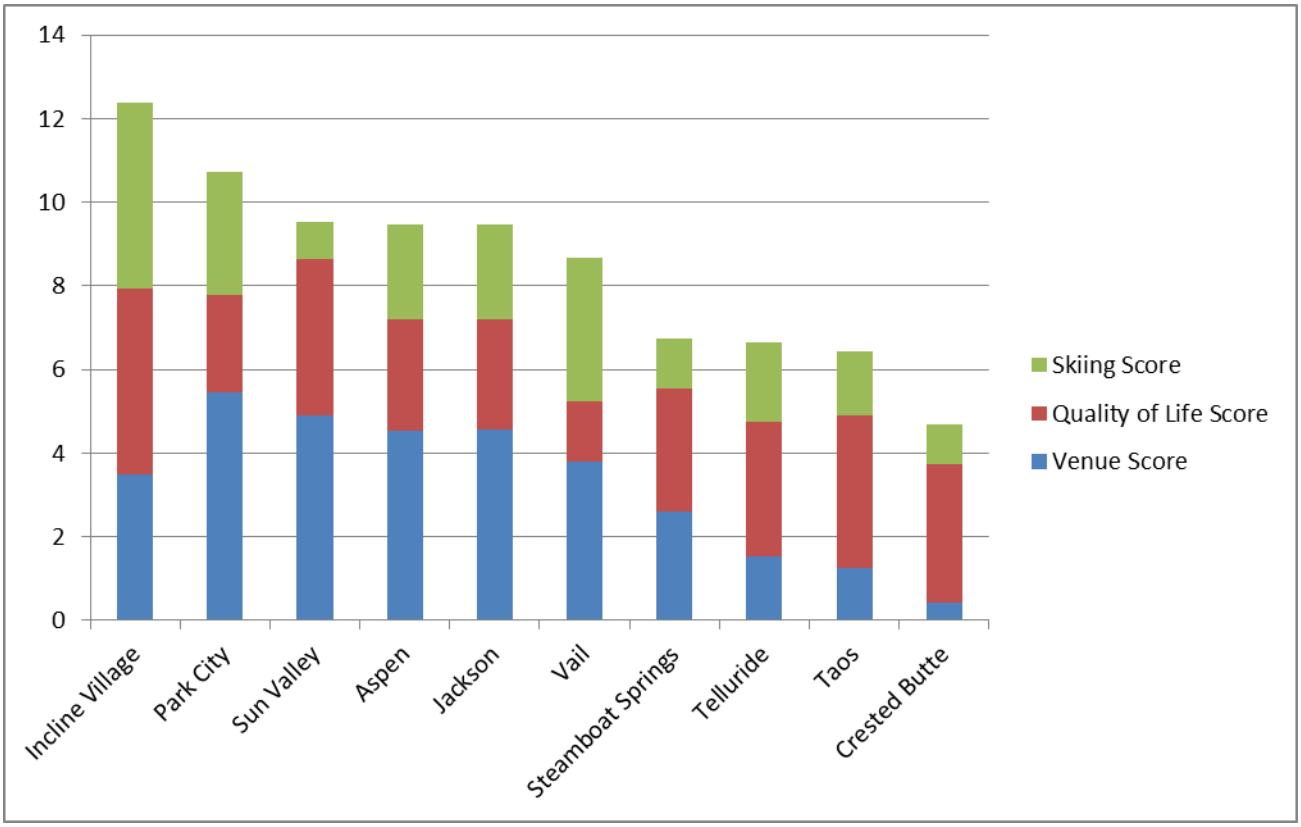
Town	Venue Score	Quality of Life Score	Skiing Score	Total Score
Incline Village	3.49244	4.43915	4.45689	12.388480
Park City	5.44118	2.3298	2.946	10.716975
Sun Valley	4.8895	3.74648	0.882779	9.518750
Aspen	4.54048	2.67294	2.25782	9.471236
Jackson	4.57087	2.61396	2.28016	9.464989
Vail	3.80462	1.42448	3.4517	8.680807
Steamboat Springs	2.60252	2.95415	1.16764	6.724310
Telluride	1.54188	3.2083	1.88559	6.635761
Taos	1.26555	3.62778	1.53168	6.425009
Crested Butte	0.428571	3.31857	0.930875	4.678017

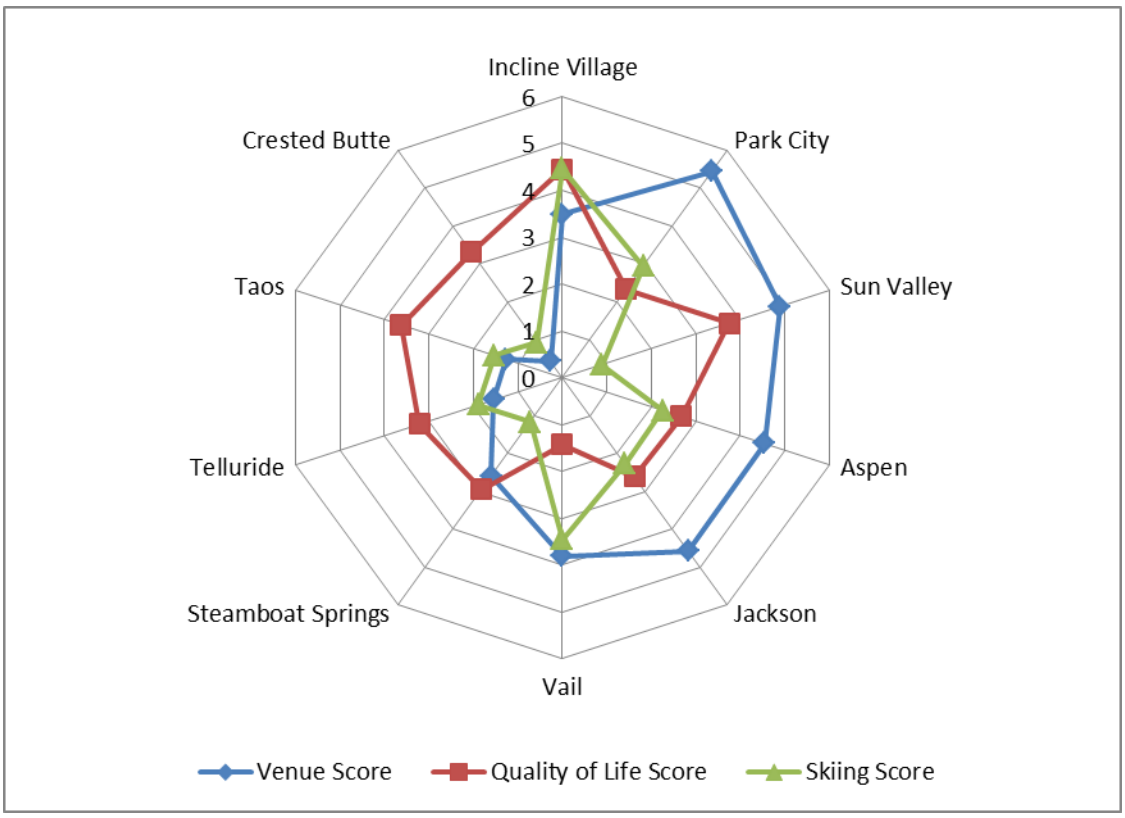
Incline Village, Nevada is rated the highest of the ten towns considered based on the highest skiing score and the highest quality of life score despite having a below average venue score. Incline's high skiing score came, despite having the least number of ski in/out acres, from having the most (by far) total skiable acres due the very many ski resorts that surround Lake Tahoe, almost all of which are within my somewhat arbitrary 45 minute commuting distance. Incline Village also came in first place in number of lifts for the same reason. Incline was also well above average in vertical drop and average annual snowfall.

Incline's top-rated quality of life score came from having the lowest crime rate and well above average scores for winter temperature and sunshine as well as school quality.

Park City finished second overall on the strength of a first-place finish in the Venues category having the most restaurants, shops and entertainment venues. Park City also scored strongly in the skiing category, finishing in third place. Unfortunately, Park City had the second worst Quality of Life score. Park City had the worst air quality and the highest cost of living index of the 10 towns. It was also well below average in sunshine, school quality and crime rate, but at least it had the warmest January temperatures.

Sun Valley came in third overall despite being in last place in the skiing category by finishing second in the venues and quality of life areas.





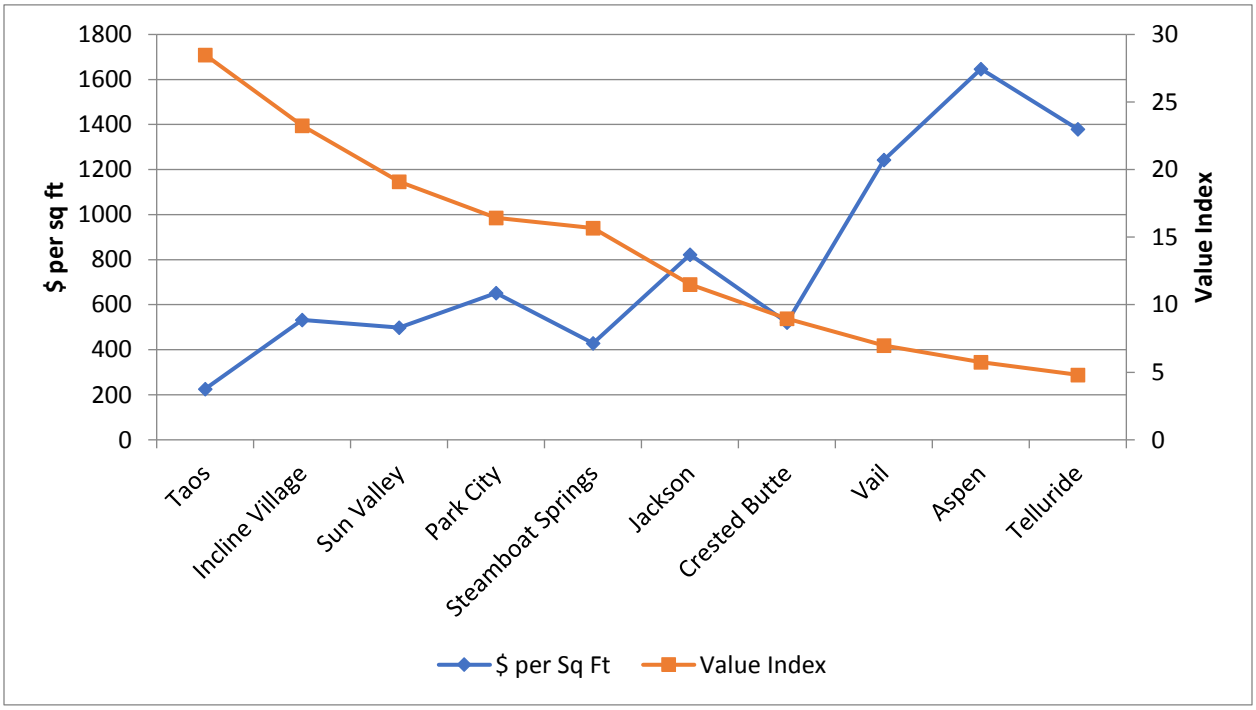
Real Estate Values

The total scores were then divided by the average listing price per square foot data as found on Zillow. Square footage pricing was used rather than actual home prices to attempt to control for variations in housing stock between towns. For example, if one market is dominated by 2-bedroom condos and another by detached 4 bedroom homes it could create a false impression of relative pricing. Pricing per square foot should give a more meaningful basis for comparison.

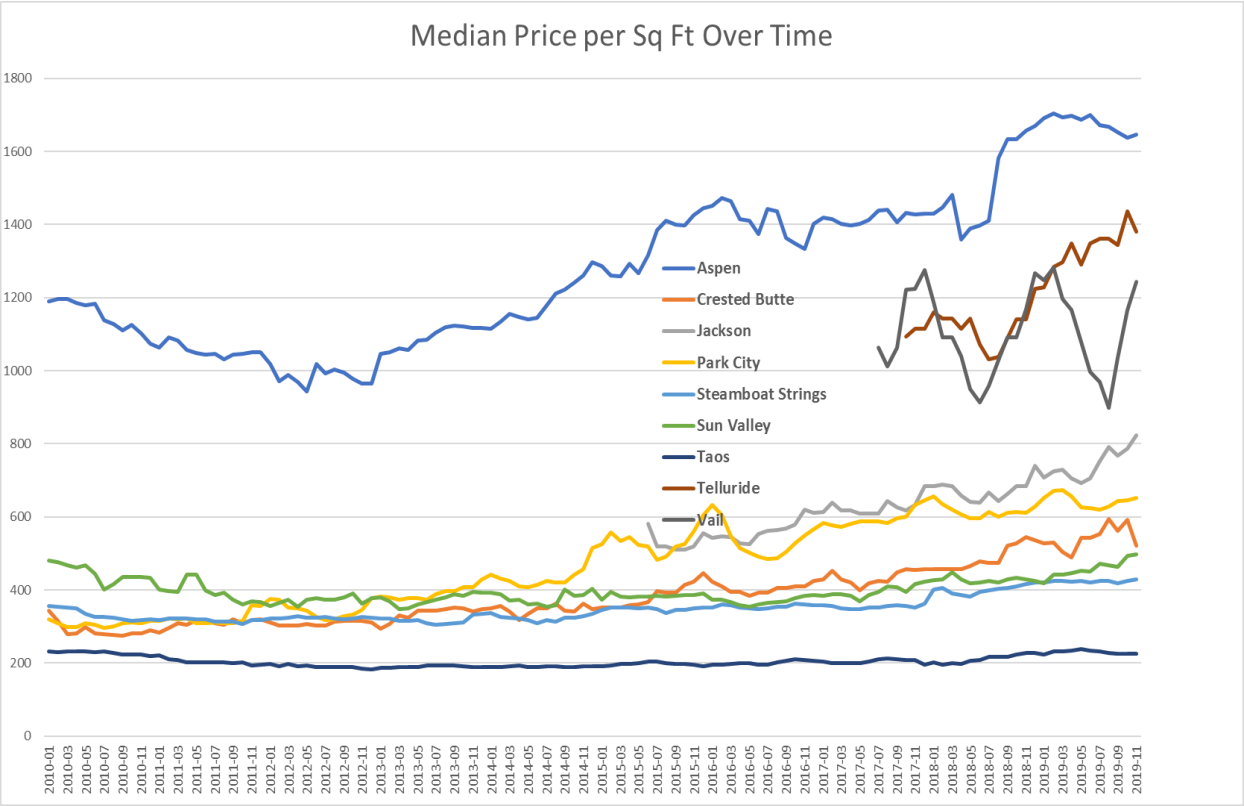
When using this method to try to establish which town gives the most value for a dollar spent on real estate, Taos, New Mexico jumps from next to last place in the raw score to first place. Real estate pricing in Taos is so very low compare to the other towns it seems to be a true outlier, so there may be more factors to consider explaining why the real estate in Taos seems so undervalued.

Incline Village fell from first place in the raw score, but only to second place in terms of real estate value. Price per square foot in Incline is about average. The most overvalued real estate values seem to be in Aspen and Telluride.

Town	Venue Score	Quality of Life Score	Skiing Score	Total Score	Median Home Price per sq ft	Value per Home $
Taos	1.26555	3.62778	1.53168	6.425009	226	28.429244
Incline Village	3.49244	4.43915	4.45689	12.388480	533	23.242927
Sun Valley	4.8895	3.74648	0.882779	9.518750	498	19.113955
Park City	5.44118	2.3298	2.946	10.716975	652	16.437079
Steamboat Springs	2.60252	2.95415	1.16764	6.724310	429	15.674383
Jackson	4.57087	2.61396	2.28016	9.464989	823	11.500594
Crested Butte	0.428571	3.31857	0.930875	4.678017	521	8.978919
Vail	3.80462	1.42448	3.4517	8.680807	1243	6.983754
Aspen	4.54048	2.67294	2.25782	9.471236	1647	5.750598
Telluride	1.54188	3.2083	1.88559	6.635761	1380	4.808523



Ten-year price trends in home values in ten ski areas are shown in the below chart.



Discussion

Obviously, there is some objectivity involved in the choice of features to use to create the ratings. Arguably, the ratings could be improved by adding additional features, but adding more features dilutes the influence of each feature as a percentage of the overall total so adding more features might not significantly change the final result.

The features used to derive the venues category seem to be fairly self-evident. Restaurants and retail shops are two obvious categories. Recreation and entertainment are also important especially if one is planning on living in the town year around rather than just during ski season. Night life could have been an additional category, but I did not include it as a separate category as I am personally at a stage of life when that is unimportant. Bars and similar venues are included in the “Other” category, so they still count towards the total.

For the most part, the features used to derive the skiing score are also common measures used to compare ski resorts. The major decision I made that other analysts might not agree with was to include any ski area within a rather arbitrary 45-minute drive of a given town in the total skiable acreage. I like variety and wouldn't want to ski the same slopes every day if I was living there the whole season. I thought that 45 minutes each way was the longest I would probably want to drive for a day trip. Another person might have a completely different opinion. Lowering the permissible commute time to 15 minutes, for example would significantly change the skiing scores. Incline Village would fall from the top of the skiing ratings to near the bottom as all but one of the 13 different ski areas surrounding Lake Tahoe other than the small area right at Incline Village are more than 15 minutes away. On the other hand, huge ski areas like Vail/Beaver Creek and Park City would advance up the rankings.

The "Quality of Life" features are not quite so obvious. Of the six features I picked two are weather related, then costs of living, air quality, school quality and crime. None of these should logically be very correlated to another. Unfortunately, the sample of only ten towns is too small to make a correlation analyses meaningful. There are conceivably a host of other factors that one could incorporate in calculating a quality of life score. There are demographic and socio-economic factors that could contribute to making one town more desirable than another. For example, population and from that one could derive many features of various types of venues per capita that one might want to examine. Population is probably not too helpful though as the number of vacationers at any time might greatly out-number the native population. It might be helpful to know the total number of hotel rooms and the average occupancy rate, but I was unable to find that data. In fact, some hotel chains do not publish occupancy rates at all. Many other features could be considered like median income, age, or educational level of residents, but none of those seem extremely important to me.

One important feature that might be good to include might be some sort of measure of the quality and completeness of the local infrastructure, but I don't know what that would be. I am sure there are other potential useful features, but I can't think of any that would be vastly more important to me or a general audience than the ones I have included. Furthermore, for the most part the venue features I have highlighted could also be considered quality of life features.

Below are the normalized ratings for the venue categories. Park City is in first place in every Category but “Other”.

	Town	Restaurant Total	Shopping Total	Rec Total	Entertainment Total	Other	Venue Score
0	Aspen	1.809524	0.35	0.714286	0.666667	1.000000	4.540476
1	Crested Butte	0.095238	0.00	0.000000	0.333333	0.000000	0.428571
2	Incline Village	0.666667	0.40	1.857143	0.333333	0.235294	3.492437
3	Jackson	1.523810	0.90	1.000000	0.500000	0.647059	4.570868
4	Park City	2.000000	1.00	1.000000	1.000000	0.441176	5.441176
5	Steamboat Springs	1.333333	0.30	0.714286	0.166667	0.088235	2.602521
6	Sun Valley	0.857143	0.65	2.000000	0.500000	0.882353	4.889496
7	Taos	0.571429	0.40	0.000000	0.000000	0.294118	1.265546
8	Telluride	0.000000	0.35	0.142857	0.666667	0.382353	1.541877
9	Vail	1.428571	0.25	0.714286	0.500000	0.911765	3.804622

This low “Other” score could be because it is the only large town not to have “Hotel” (which is an “other”) in its top 4. Surprisingly, “Hotel” or similar does not even appear in the top 10 most common venues for Park City as seen in the below table. Fully half of the towns have “Hotel” as the most common venue type. Another commonality is that only Sun Valley does not have “American Restaurant” in its top two most common venues, but it does have it in 3rd place. Other than American Restaurant the most common types of restaurants are Mexican and Italian.

	Town	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Aspen	Hotel	American Restaurant	Italian Restaurant	Ski Area	Trail	Sandwich Place	Restaurant	Park	Cocktail Bar	Resort
1	Crested Butte	American Restaurant	Coffee Shop	Breakfast Spot	Steakhouse	Pizza Place	Bar	Mexican Restaurant	Gastropub	Gas Station	French Restaurant
2	Incline Village	American Restaurant	Coffee Shop	Beach	Hotel	Tapas Restaurant	Mexican Restaurant	Deli / Bodega	Pizza Place	Golf Course	Sandwich Place
3	Jackson	Hotel	American Restaurant	BBQ Joint	Resort	Ski Lodge	Mexican Restaurant	Coffee Shop	Clothing Store	Bar	Breakfast Spot
4	Park City	Coffee Shop	American Restaurant	Ski Area	Café	Sushi Restaurant	Mexican Restaurant	Ski Lodge	Bar	Sporting Goods Shop	Pizza Place
5	Steamboat Springs	American Restaurant	Café	Italian Restaurant	Coffee Shop	Pizza Place	Candy Store	Brewery	Mexican Restaurant	Park	Restaurant
6	Sun Valley	Hotel	Ski Chairlift	American Restaurant	Sporting Goods Shop	Bar	Italian Restaurant	Clothing Store	Steakhouse	Ski Trail	Ski Lodge
7	Taos	Mexican Restaurant	American Restaurant	Coffee Shop	Brewery	Grocery Store	Hotel	Motel	Fast Food Restaurant	Café	Diner
8	Telluride	Hotel	American Restaurant	Bar	Ski Area	Pizza Place	Sporting Goods Shop	Café	Sandwich Place	Trail	Clothing Store
9	Vail	Hotel	American Restaurant	Ski Area	Resort	French Restaurant	Pizza Place	Lounge	Board Shop	Ski Chairlift	Plaza

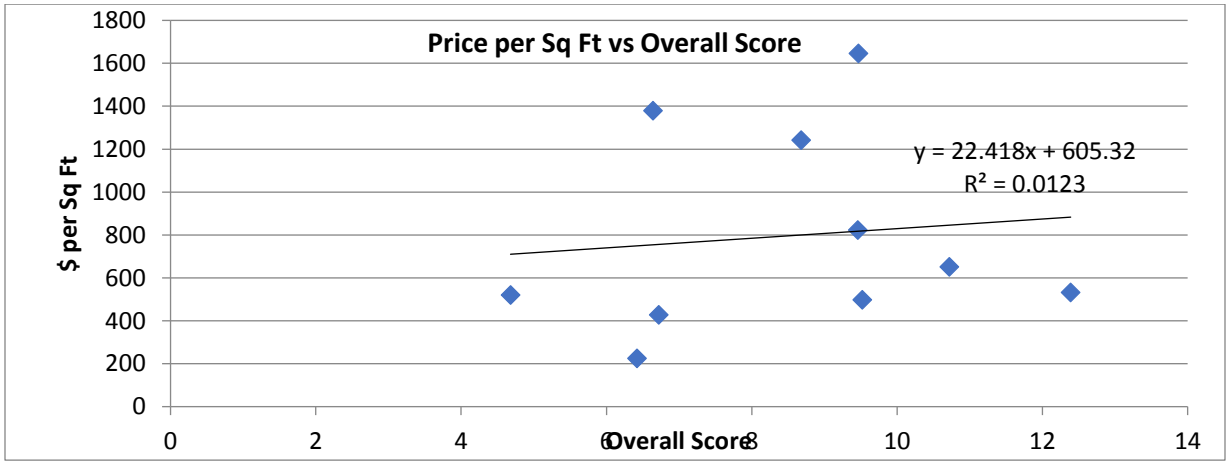
The below table show the normalized scores for the individual “Quality of Life” features. Park City and Taos are notably schizophrenic. Park City has one first place score and two last place scores in the six categories while Taos has two first places and one last place. There is a huge gap in the double weighted Crime Rate feature between the top two towns in the real estate value score. Incline Village has the lowest (best) crime rate of the ten towns while Taos has the worst crime rate. The crime rate in Taos might have something to do with its extremely low real estate prices.

	Town	Avg Jan Temp	Air Quality Index	Avg An Sunny Days	Cost of Living Index	School Quality	Crime Rate	Quality of Life Score
0	Aspen	0.533333	0.055954	0.531646	0.390977	0.835322	0.325704	2.672937
1	Crested Butte	0.000000	0.055954	0.873418	0.241486	1.000000	1.147713	3.318571
2	Incline Village	0.866667	0.026070	0.594937	0.063647	0.887828	2.000000	4.439149
3	Jackson	0.066667	0.066912	0.101266	0.342105	0.942721	1.094286	2.613956
4	Park City	1.000000	0.000000	0.316456	0.000000	0.565632	0.447713	2.329802
5	Steamboat Springs	0.400000	0.035568	0.481013	0.403635	0.880668	0.753268	2.954151
6	Sun Valley	0.066667	1.000000	0.000000	0.330309	0.570406	1.779094	3.746475
7	Taos	0.733333	0.815686	1.000000	1.000000	0.078759	0.000000	3.627779
8	Telluride	0.866667	0.035568	0.481013	0.097744	0.935561	0.791747	3.208299
9	Vail	0.266667	0.045518	0.544304	0.284705	0.000000	0.283287	1.424480

The normalized scores for the skiing related statistics appear in the table below. The selection and weighting of features favors bigness, or in Incline Village’s case proximity to numerous ski areas. Acreage correlates to the number of lifts. It would also seem that areas with the most snow seem to have grown into the biggest skiing destinations, which makes sense. So arguably there could be more variation in the features I used within the skiing category. Skiers per hour lift capacity is a metric one often sees, but not all ski areas seem to publish that statistic and it would still be correlated to the overall size of the mountain. I have tried to de-correlate that statistic by deriving a lifts per acre feature which hopefully, but not necessarily, will give an indication of the length of lift lines. A better indicator would be lifts per hotel room or acres per hotel room, but I was unable to uncover stats on number of rooms.

	Town	Total Skiable Acreage	Ski In/Out Acreage	Vertical	Lifts	Avg Snowfall	Lifts per Acre	Skiing Score
0	Aspen	0.184403	0.002709	0.986131	0.150685	0.243161	0.690734	2.257823
1	Crested Butte	0.005232	0.134236	0.000000	0.000000	0.243161	0.548246	0.930875
2	Incline Village	2.000000	0.000000	0.791241	1.000000	0.547112	0.118541	4.456895
3	Jackson	0.337117	0.277652	0.032847	0.260274	0.726444	0.645831	2.280165
4	Park City	0.773149	1.000000	0.447445	0.315068	0.410334	0.000000	2.945997
5	Steamboat Springs	0.149260	0.346275	0.251825	0.027397	0.392097	0.000783	1.167638
6	Sun Valley	0.065153	0.225583	0.324818	0.013699	0.000000	0.253527	0.882779
7	Taos	0.000000	0.096163	0.164964	0.027397	0.243161	1.000000	1.531684
8	Telluride	0.049951	0.202408	1.000000	0.013699	0.270517	0.349012	1.885585
9	Vail	0.801283	0.697366	0.288321	0.431507	1.000000	0.233227	3.451704

Finally, the measurement of relative real estate values should not be taken at face value. It is not an attempt to model or predict housing prices, only an indication of the relative "bang for the buck" one would get in each market. Clearly, the overall score is not predictive of price as shown below.



The score is obviously not capturing some major determinants of price either as absolute price or the price per square foot. There are many possible variables that could be looked at if the data could be found. Among these might be average age of buildings or quality of construction. If the homes in a wealthier area have very high-end finishes and the homes in another area have cheap finishes that would affect the price. One area might tend to have many detached houses while another is dominated by smaller condos. Large houses generally sell for more per square foot than small condos. Average lot size would also affect the price. Availability of buildable land near the mountain would also be a factor.

Aspen and to a lesser extent Telluride and Vail have a reputation as havens for the ultra-wealthy and celebrities. A high concentration of very wealthy people would tend to bid up real

estate prices in one area relative to others. It would be interesting if one were attempting to model real estate prices to see if there was a strong correlation between median income in a town and housing prices. At the other end of the spectrum, housing in Taos is relatively quite inexpensive. While, as seen on the chart on page 12, most ski town housing markets have enjoyed a strong appreciation in recent years, housing in Taos for some reason has remained flat and has not increased in value at all.

As for Incline Village on the shores of Lake Tahoe, lake front property always carries premium valuations. One would expect that this would tend to increase the median prices in that town.

Conclusion

Given my preferences, the final score does give a strong indication of where a buyer with similar preferences can get the most for his or her money. It would be interesting to build this into an interactive program that would allow an interested person to choose what features to include and to weight them according to his or her preferences. More work needs to be done on factors influencing housing prices. But for the purposes of this exercise I am only interested in which market gives the greatest value for the money.

If I was in the market today, I would want to look more deeply into why real estate in Taos is so inexpensive relative to other ski areas. Even though Taos has the best ratio of home cost to total score, Incline Village might still be a preferable place to buy based on finishing first in Total Score versus Taos at 9th (next to last). Incline Village still finished second in value due it's high overall score and close to average housing prices. Obviously, buyers need to weigh other factors besides just the numbers when making a major life decision, but the output of a structured data driven approach should also be seriously considered to avoid making a purely emotional and perhaps uninformed purchase.