**Finding the Best City to Move to in The United States**

Greg Sullivan

July 02, 2020

Introduction:

As a young professional on the job market, I have recently realized the huge discrepancy in housing prices for cities around the United States. Typically, the markets with the higher housing costs are also the markets with the higher salaries. I would like to filter through the data to find which cities I would best be able to buy a house. It is important to note that salary vs. housing cost is not the only factor considered to find the best cities to move to; happiness is probably the most important factor when looking for a place to move.

The goal of this project is to build a notebook that helps the user rank potential places to live. The intended user is anyone who is looking to move to a city within the United States.

Inherently, some very important factors will be left out of this analysis. For example, where family and friends live is not taken into account. The local school system is another major factor not considered in this analysis which many users would consider crucial.

Data:

The data used will be coming from three main sources:

**Zillow:**

* Smoothed, seasonally adjusted measure of the typical home value and market changes
* Typical value for homes within the 65th to 95th percentile range
* All single-family residences and for all homes with 1, 2, 3, 4 and 5+ bedrooms

Download can be found at https://www.zillow.com/research/data/

**U.S. Bureau of Economic Analysis:**

* Income data is counted per person over 18 years old
* Income that people get from wages, proprietors' income, dividends, interest, rents, and government benefits
* A person's income is counted in the county, metropolitan statistical area, or other area where they live, even if they work elsewhere.

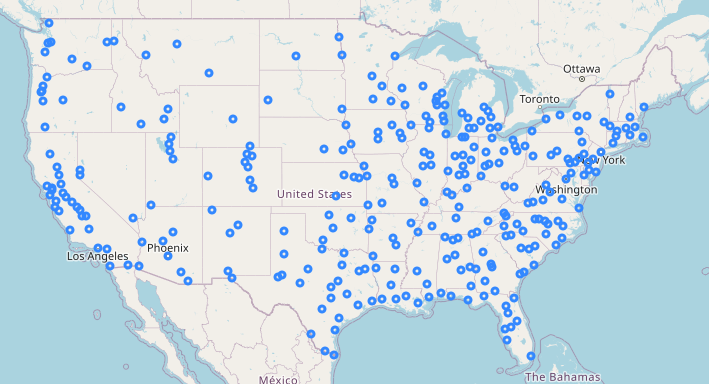
Download can be found at https://www.bea.gov/data/income-saving/personal-income-county-metro-and-other-areas

**Foursquare:**

* Data is pulled from Foursquare to better understand cities
* The number of venues in a city is taken in a 5 mile radius from city center

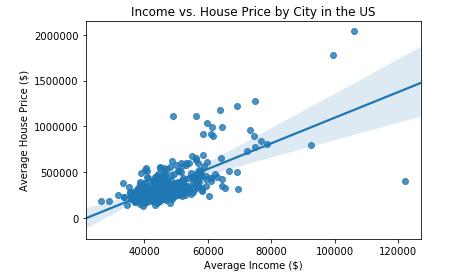
Methodology:

The data pulled from Zillow, the U.S. Bureau of Economics Analysis, and Foursquare needed to be merged to compare data. An inside merge was performed meaning that only common data was kept. This significantly reduced the data set to about 300 cities. Additionally, some data for the cities includes the metroplex while others do not. A map of the cities included in this analysis is shown below. Cities in Hawaii and Alaska with sufficient data were included in the analysis but not shown on the map below.

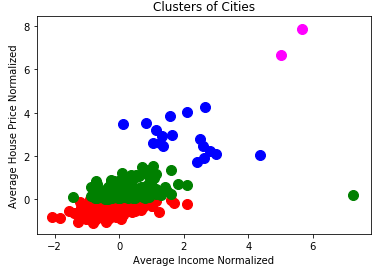


Results:

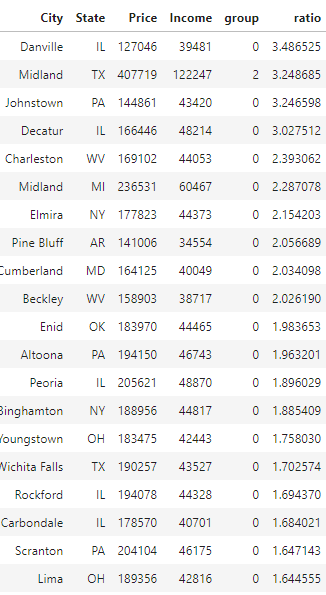
Average income vs Average House price in each city was plotted to ensure a positive correlation was found. As expected, salary and housing price increase together. There is one point well to the right of the plot below which shows a low house price with a very high salary. That point is Midland, Tx which has it’s flaws that will be discussed later.



Cities were separated into four different groups low (red), middle (green), high (blue), and luxury (magenta). The two luxury cities shown below are San Francisco, California and San Jose, California. Neither of these come as a surprise as being places with a high salary and expensive houses.



The question becomes: Does the increase in salary offset the housing cost in these cities? The answer to that question is no. A ratio between income and the home price was used to find your “purchasing power” in each city. The results are shown below.

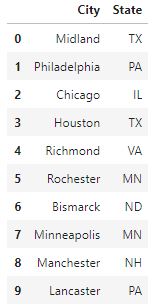


The data is heavily skewed to cities in the "low" housing cost and income ranges meaning that these cities can be considered more affordable to the average American.

The issue with this analysis is that it ignores the type of jobs available in each city.

Because I am looking for an engineering job in which I can grow, I will consider cities outside of the 'low' group. Good jobs do exist in these cities, but they would be harder to come across.

The cities with the highest purchasing power are shown below:



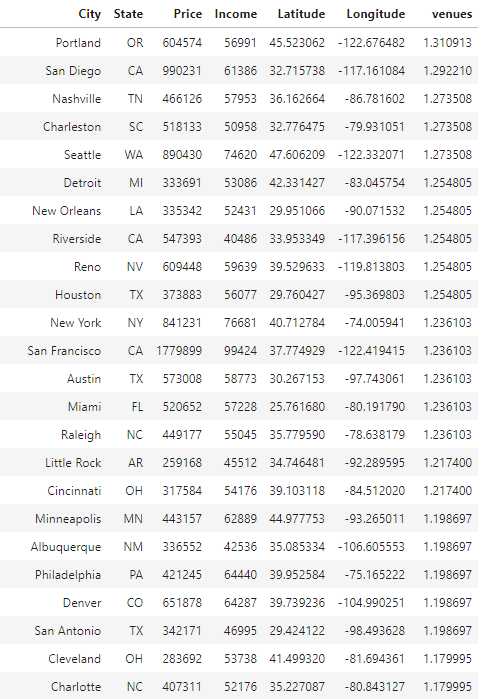
As well as the cities with the least:



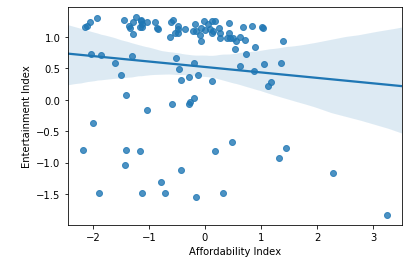
From the analysis, it appears that Midland, TX would be the best city to move to and I should avoid California. As someone who has lived in Midland, TX, I can verify that it is a great place in terms of income and affordability. However, the city was not a good fit in terms of what I was looking for.

At this point, we need to note that there is no single "perfect" city. Everyone will have their own preferences when it comes to where they live. So how can we account for this?

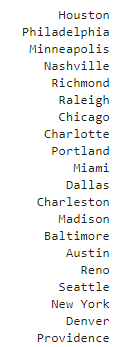
I will attempt to use the Foursquare API to find cities that have a high level of activities for residents. To this I will find the number of venues within 5 miles of downtown. Upon normalizing this, I can have an ‘Entertainment Index’ for each city. The cities raked by their entertainment index are shown below:



So now we need to analyze the Entertainment Index vs the Affordability index. A negative correlation is shown between the two. This makes sense because people sacrifice salaries and pay more for houses in more exciting locations.



Finally, the two metrics had to be multiplied to come up with an overall score for each city. Entertainment was weighted to account for 80% of the metric with affordability being worth 20%. The resulting list of best cities to live is below:



Discussion:

The methodology used in this analysis is limited and inherently flawed. The city with the lowest rating for where to live was San Luis Obispo, CA which almost anyone would move to in a heartbeat. Overall, I am pleased with the results and I believe they show a relatively good list of the top 20 places to live in the US. Additional information like crime rate, schools, and specific interests would help this analysis.

Conclusion:

Based on the results the best city to move to is **Houston, TX**. This is a reasonable result with the limited data we used. Refinement on this model can lead to better and more personalized results.