

Fastest Growing Cities in the U.S.

Tom Allen

02/07/2020

IBM Data Science Certification Capstone

1. Introduction

1.1 Background

There has never been an easier time to move and experience a different area, whether it's relocating for a different job, starting a new business or moving for fun. Technology has allowed users to explore areas much easier and quicker all while sitting in the comfort of their own home.

1.2 Problem

As cities become denser, problems start to occur including lack of infrastructure to support the masses, increase in housing costs and outdated public transportation. I wanted to see what cities are experiencing the most growth and what characteristics these areas have. For entrepreneurs, re-locating to a fast-growing city could be the kickstart they need to get their business going in the right direction.

2. Data

2.1 Data Collection

I was able to find a public [dataset](#) that comprised of the 1,000 largest cities in the U.S. This dataset included City, Rank, State, Growth from 2003 to 2013, Population and each city's geographical coordinates. The first 5 rows of this dataset are displayed in Figure 1.

Figure 1

	City	Rank	State	Growth From 2000 to 2013	Population	Coordinates
0	Marysville	552	Washington	115.7	63269	48.0517637,-122.1770818
1	Perris	466	California	98.7	72326	33.7825194,-117.2286478
2	Cleveland	48	Ohio	-18.1	390113	41.49932,-81.6943605
3	Worcester	129	Massachusetts	5.8	182544	42.2625932,-71.8022934
4	Columbia	192	South Carolina	11.7	133358	34.0007104,-81.0348144

I was also able to download a .csv file that had contained a state's region that I could eventually merge with the first dataset. This would be helpful to find what regions of the U.S. were growing the fastest. The first 5 rows of this dataset are displayed in Figure 2.

Figure 2

	State	State Code	Region	Division
0	Alaska	AK	West	Pacific
1	Alabama	AL	South	East South Central
2	Arkansas	AR	South	West South Central
3	Arizona	AZ	West	Mountain
4	California	CA	West	Pacific

2.2 Data Cleaning

The datasets were relatively clean to begin with and had no missing values. The only cleaning that was done was for formatting purposes to get the data in a format that I could easily analyze.

First, I had to separate the Coordinates column into Latitude and Longitude because it would be much easier to call a city's geographical location if they were separated columns.

Second, I merged the dataset containing the city's growth and coordinates with a dataset that contained each states region. I joined them on the key: "State".

Third, I created two different subsets of the merged data frame to get 1) data frame containing the top 10 most populous cities and 2) data frame containing the top 10 fastest growing cities. This would allow me to compare these city's locations and see if there were any patterns between them. See the top 10 fastest growing cities data frame in Figure 3.

Figure 3

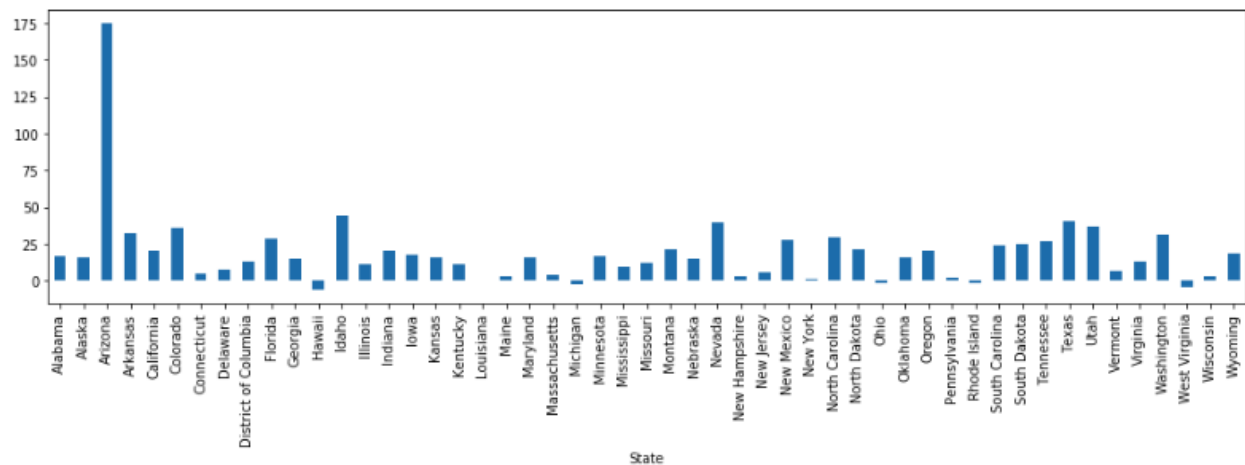
	City	Rank	State	Growth	Population	Latitude	Longitude	State Code	Division
0	Maricopa	817	Arizona	2503.4	45508	33.058106	-112.047642	AZ	Mountain
1	Buckeye	641	Arizona	480.9	56683	33.370320	-112.583777	AZ	Mountain
2	Frisco	186	Texas	287.7	136791	33.150674	-96.823612	TX	West South Central
3	Lincoln	822	California	285.2	45237	38.891565	-121.293008	CA	Pacific
4	Surprise	216	Arizona	281.9	123546	33.629234	-112.367928	AZ	Mountain
5	Goodyear	462	Arizona	271.0	72864	33.435339	-112.357657	AZ	Mountain
6	Beaumont	910	California	254.5	40481	33.929461	-116.977248	CA	Pacific
7	Plainfield	879	Illinois	203.6	41734	41.632223	-88.212031	IL	East North Central
8	Wylie	828	Texas	185.2	44575	33.015120	-96.538879	TX	West South Central
9	Lehi	673	Utah	176.3	54382	40.391617	-111.850766	UT	Mountain

3. Exploratory Data Analysis

3.1 Plotting the Fastest Growing States

I first made a quick bar chart, see Figure 4, that displayed the fastest growing states, so the viewer would not have to manually check each state's growth rate. This plot grouped by each state and aggregated the mean of each city's growth rate. The viewer could then clearly see that Arizona, Idaho, Nevada and Texas were the fastest growing states.

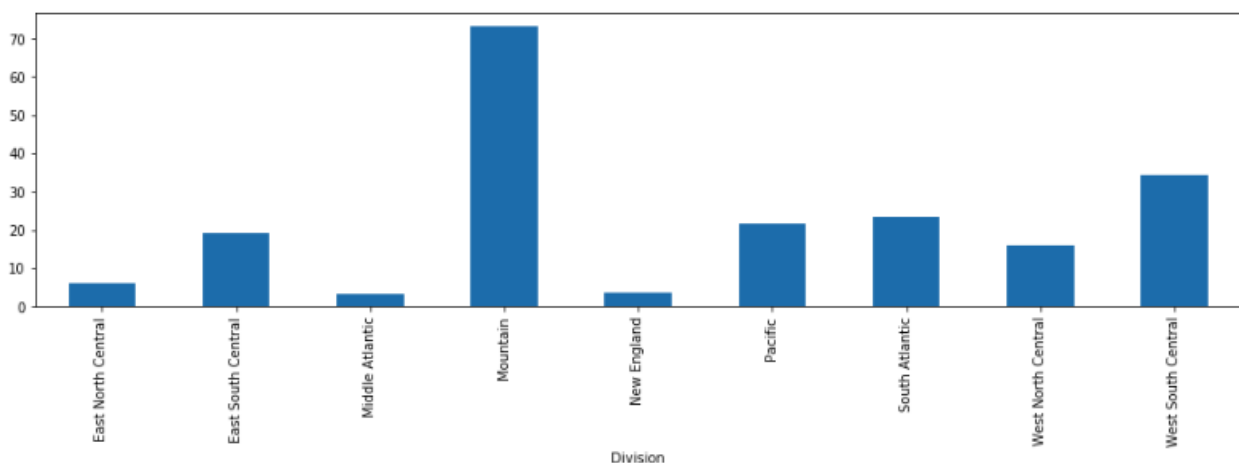
Figure 4



3.2 Plotting the Fastest Growing Regions

I made another bar chart, see Figure 5, that displayed what regions were growing the fastest. This time, I grouped by region and aggregated the mean of each city's growth rate. This chart displayed the Mountain region had the fastest average growth rate, with West South Central in second.

Figure 5

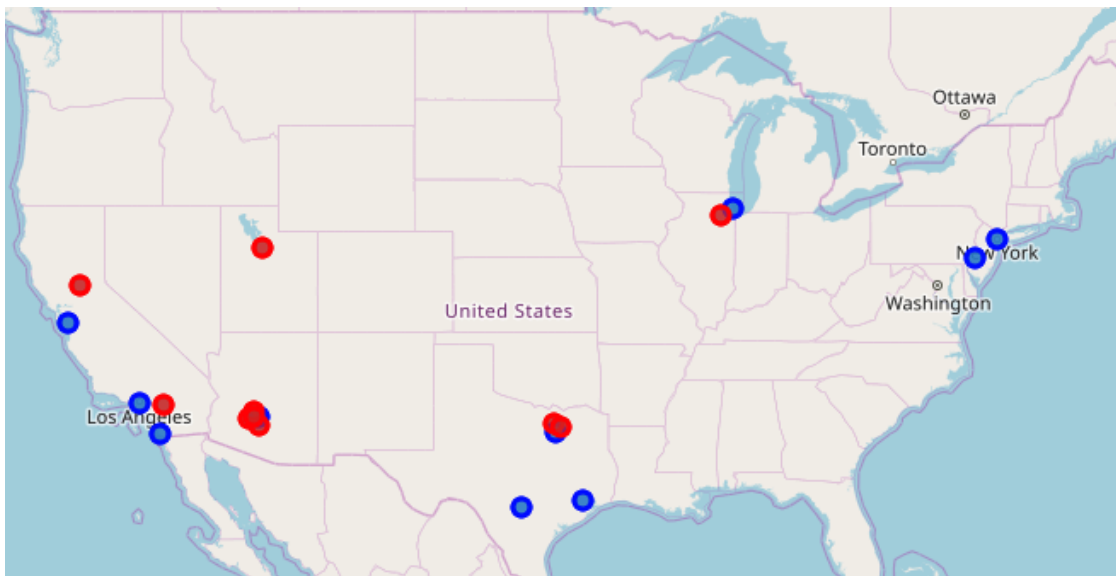


3.2 Plotting Using a Folium Map

Using a Folium Map, I could plot each city on a map to clearly see the locations or areas the fastest growing or most populous cities were located in, see Figure 6. I plotted the most populous cities with blue dots and the fastest growing cities with red dots. The map could easily display

the locations of each of these cities on a readable U.S. map. It clearly showed, most of the fastest growing cities were located in the Mountain region and California. One interesting inference from this chart, was that most of the fastest growing cities were located just outside or near major cities.

Figure 6



4. Getting City Data

4.1 Foursquare

Using the Foursquare API, I could extract data from nearby venues around these city's coordinates. I could return venue names, tips, frequency of visits, etc. I first had to set up an account in order to get client identification and client secret numbers in order for the API call to work.

4.2 Getting Nearby Venues

Using a function that used the Foursquare API, I was able to pass in all the fastest growing city's coordinates to get venues nearby. This function returned a data frame, see Figure 7, consisting of the city, location and venues that were in a 500-meter radius. For each venue, it had its official name and a category it belonged to (i.e. Domino's Pizza and Pizza Shop).

Figure 7

	City	City Latitude	City Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Maricopa	33.058106	-112.047642	Honeycutt Coffee	33.058341	-112.044423	Coffee Shop
1	Maricopa	33.058106	-112.047642	Helen's Kitchen	33.053753	-112.048868	American Restaurant
2	Maricopa	33.058106	-112.047642	Domino's Pizza	33.061536	-112.048774	Pizza Place
3	Maricopa	33.058106	-112.047642	Amtrak Maricopa Station	33.056270	-112.048117	Train Station
4	Maricopa	33.058106	-112.047642	Head Quarters	33.058179	-112.048288	American Restaurant
...
106	Lehi	40.391617	-111.850766	Thai House	40.387930	-111.853502	Thai Restaurant
107	Lehi	40.391617	-111.850766	Domino's Pizza	40.387779	-111.851021	Pizza Place
108	Lehi	40.391617	-111.850766	SHED BUILDERS	40.393970	-111.852163	Business Service
109	Lehi	40.391617	-111.850766	John Hutchings Museum of Natural History	40.388462	-111.849489	History Museum
110	Lehi	40.391617	-111.850766	Wines Park	40.394434	-111.848239	Park

4.3 Venue Category Frequency

Instead of looking at the specific venue names, grouping by the category they belonged to could tell a much better story about each city. After encoding each venue by the frequency of it being nearby to that city, I was able to get tables that displayed the most frequent categories of venues for each city. This gave me a quick idea of which types of venues were most popular in that city. Figure 8, shows the first 2 cities and the most frequent venues.

Figure 8

----Beaumont----		
	venue	freq
0	Fast Food Restaurant	0.21
1	Hotel	0.07
2	Mexican Restaurant	0.07
3	Dessert Shop	0.07
4	Coffee Shop	0.07
----Buckeye----		
	venue	freq
0	Mexican Restaurant	0.14
1	Discount Store	0.14
2	Video Store	0.07
3	Gas Station	0.07
4	Coffee Shop	0.07

5. Conclusion

Through this analysis, I was able to gather valuable information regarding city growth and what types of venues surround these areas. This could be valuable to anyone looking to move or an entrepreneur looking to start a new business.

One observation was that the Mountain region was the fastest growing area and Arizona was a big part in that growth. Are people wanting to move to a warmer climate or is businesses starting to re-locate to this area?

Another observation was that the fastest growing cities were located around the most populous cities. Personally, I expected this because after living in major city, traffic is a major problem, housing prices are increasing to low supply and more demand and the aging infrastructure is not able to keep up with the amount of people living there. It seems people are starting to move just outside of the city, to have a better cost of living or to get away from the clustering of people.

6. Recommendations

6.1 Opening a Mexican Restaurant

It is clear that in the Mountain region, Mexican Restaurants are extremely popular and there are many cities that are growing fast giving the business an opportunity to succeed.

6.2 Moving Outside of Major Cities

As noted in conclusion section, some of the fastest growing cities tend to be near major cities for reasons that may include cost of living, congestion, city taxes, etc. By getting a head start and moving to a faster growing city, a homeowner can buy for cheap, and experience growth in the property's value in order to make a solid investment.