Impact of COVID-19 on the U.S. Housing Market

Marissa Tan

Introduction

The United States has experience the steady growth in the housing market. This steady growth was interrupted by COVID-19 pandemic. This study compares the United States' housing market before and after COVID -19.

This information is important to the local government and real estate developers when planning to develop the next community.

Research Questions

1. Has there been a change in house prices since COVID-19?
2. Has there been a change in in-demand house location since COVID-19?
3. Has there been a change in buyer's house buying criteria since COVID-19?

Literature Reviews

1. Has there been a change in house prices since COVID-19?

There has been a change in house prices since COVID-19. There were construction delays in 2020 because "local business offices were closed or adopted new processes, shutdown of non-essential construction, supply chains have been disrupted, and construction labor force was at risk." (Airgood-Obrycki, 2020) Thus, there has been a short of housing inventory. "A shortage of homes for sale led to bidding wars, pushing price growth to new high." (Friedman, 2022)

2. Has there been a change in in-demand house location since COVID-19?

There has been a change in in-demand house location since COVID-19. According to Liu, "homes sales declined more in counties with higher population density and higher pre-COVID home value. The pandemic has spurred more new listings in denser neighborhoods. As a result, home inventory increased more in denser and more expensive counties. The pandemic has shifted housing toward less dense and cheaper locations." (Liu & Su, 2021, p. 12)

3. Has there been a change in buyer's house buying criteria since COVID-19?

There has been a change in buyer's house buying criteria since COVID-19. "Americans will move within 60 minutes of major cities as the historic benefits of dense urban housing and work arrangement are curb by COVID restriction." (Dotzour, 2021) Previous commuters to work "have broadly embraced work-from-home arrangements, previous home buying considerations like concerns about commutes will fall in importance." (Dotzour, 2021)

Data Sources

Data Set	Source	Description
Land Area	US Census Bureau	The file will use county land area information to compute the population densities. Link: https://www2.census.gov/library/publications/2011/compendia/usa-counties/excel/LND01.xls
Population	Economic Research Service US Department of Agriculture	This file provides population estimates for all counties in the US. Note: 2016 to 2019 estimates and 2020 actual population will be used. Links: https://www.ers.usda.gov/webdocs/DataFiles/48747/PopulationEstimates.xls https://www.ers.usda.gov/webdocs/DataFiles/48747/PopulationEstimates.csv?v=390.6
Real Estate	Realtor.com	This real estate data file contains a historical market trends and monthly statistics on active for-sale listings (including median list price, average list price, luxury list price, median days on market, average days on market, total active listings, new listings, price increases, price reductions) Link: https://econdata.s3-us-west-2.amazonaws.com/Reports/Core/RDC_Inventory_Core_Metrics_County_History.csv
Rural Area	US Census Bureau	The file will use county rural information to categorize the urban and mostly and all rural areas. Link: http://www2.census.gov/geo/docs/reference/ua/County_Rural_Lookup.xlsx

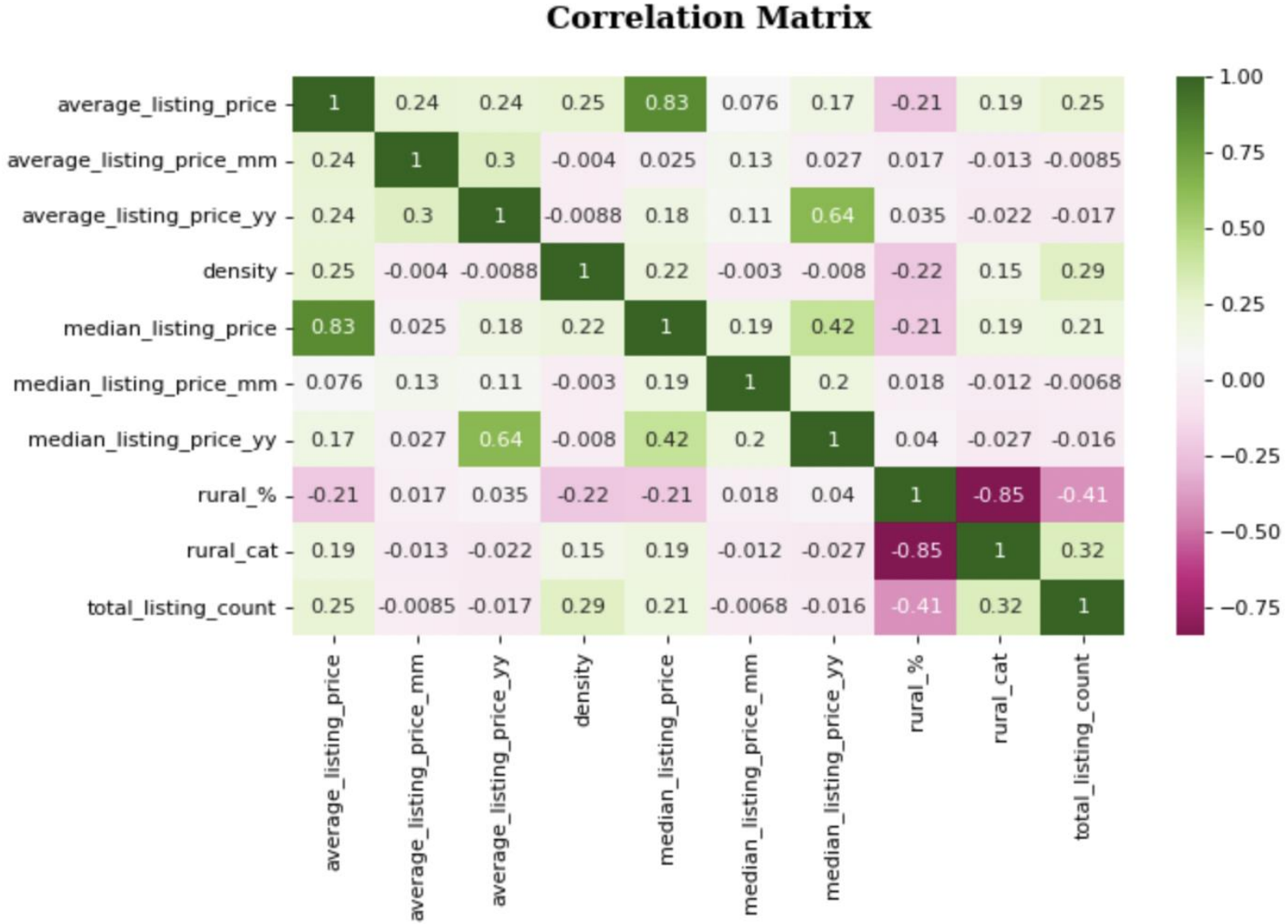
Data Variables

DATA VARIABLE	DESCRIPTION
Avg Listing Price	The average listing price within the specified geography during the specified month.
Avg Listing Price M/M	The percentage change in the average listing price from the previous month.
Avg Listing Price Y/Y	The percentage change in the average listing price from the same month in the previous year.
County	The county in the state
Date	The date of the property transaction
Density	The ratio of the population per square mile of the county (or state)
FIPS	The five digit code representing the state and the county
Median Listing Price	The median listing price within the specified geography during the specified month.
Median Listing Price M/M	The percentage change in the median listing price from the previous month.
Median Listing Price Y/Y	The percentage change in the median listing price from the same month in the previous year.
Rural Percent	The total percentage of the population that classifies almost or all rural areas and urban area
Rural Cat	The category 1 represents all or mostly rural areas; category 2 represents urban area.
State	The abbrevation of the state
Total Listing Count	The total of both active listings and pending listings within the specified geography during the specified month. This is a snapshot measure of how many total listings can be expected on any given day of the specified month.

Variables and Measurements

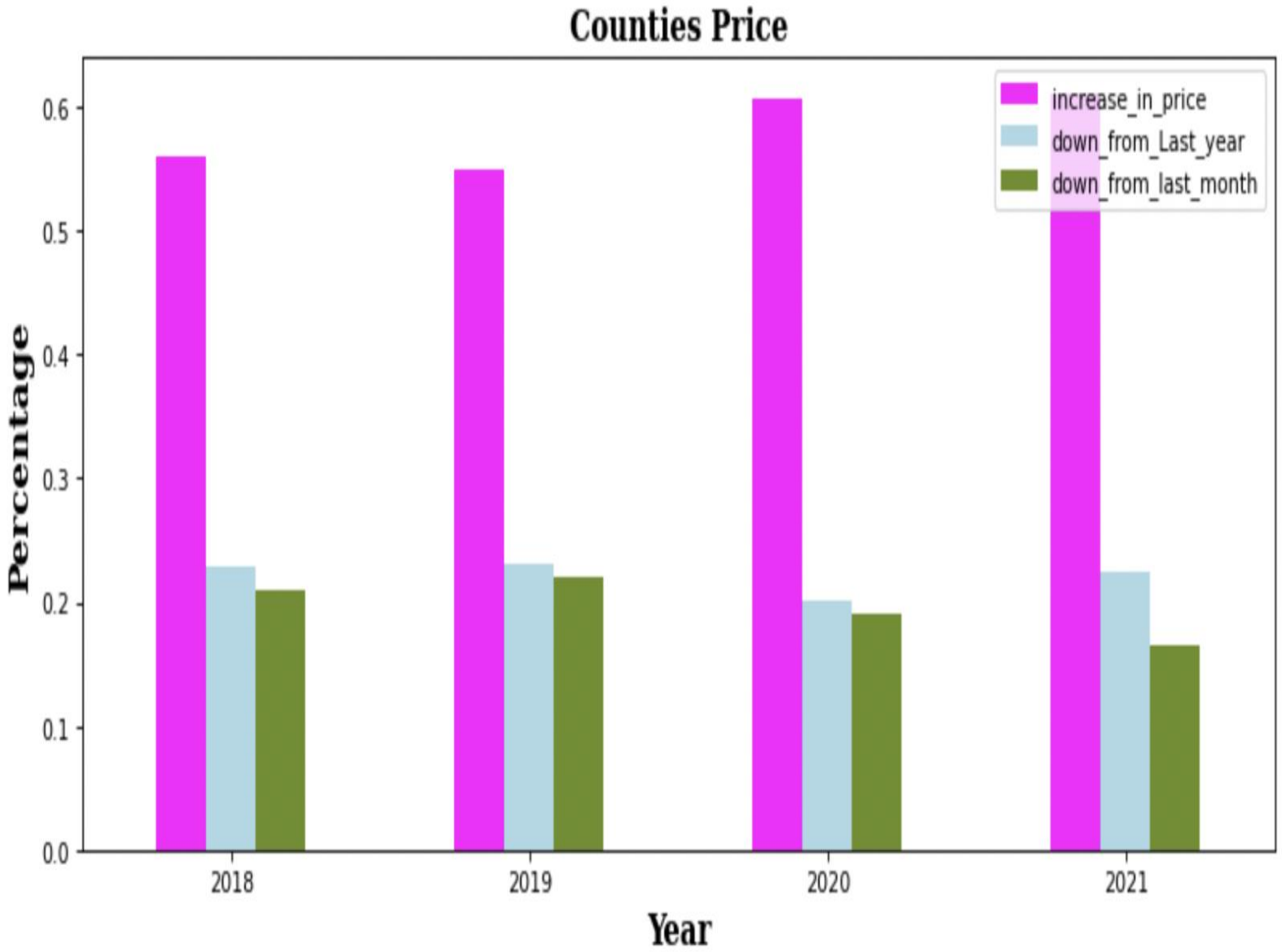
- There are 210,602 observations and 40 features in the original real estate dataset.
- Average listing price, average listing prices changes per year, density, and rural percentage for the state, individual state, and county level will be to analyzed for the period of 2019 and 2021.
- The basis of comparison is pre-COVID (2019) and post-COVID (2021)
- The unit for analysis is FIPS which represents the Federal Information Processing Standard code. FIPS is a five digit code that uniquely identifies counties in the United States.

Exploratory Data Analysis

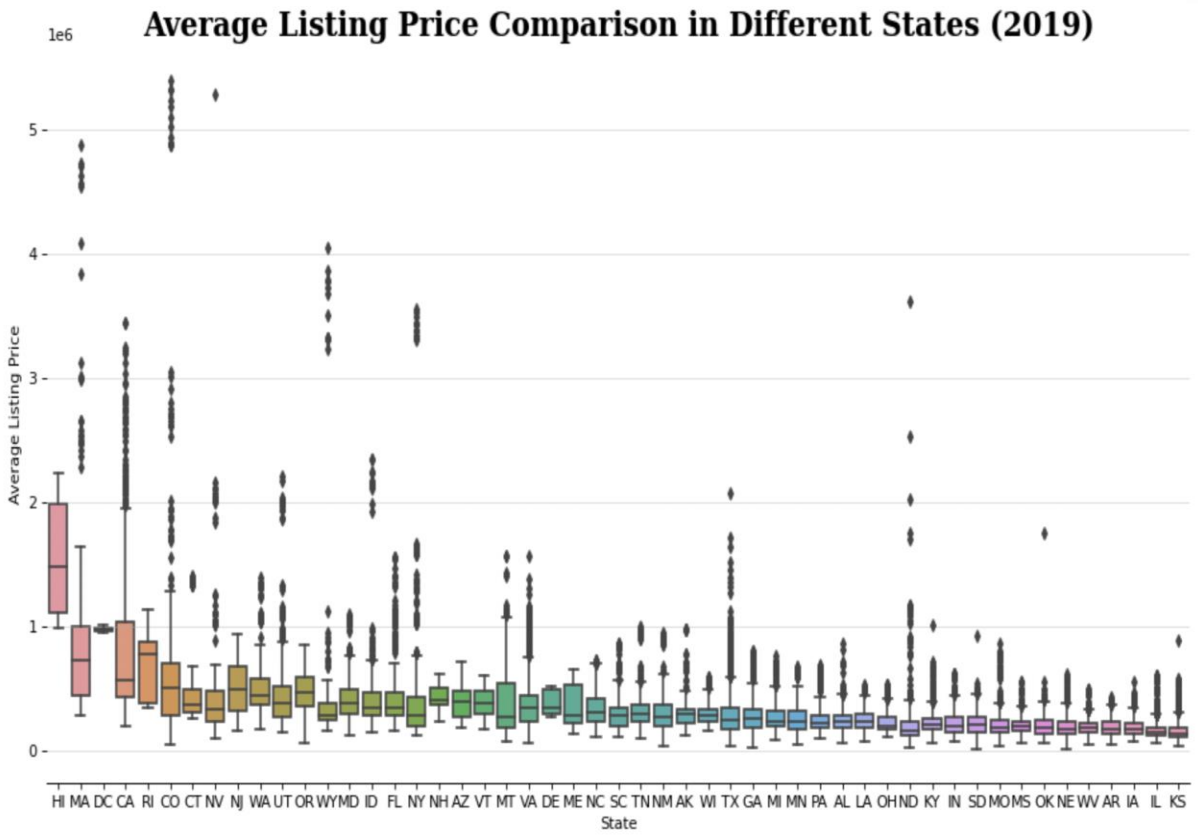


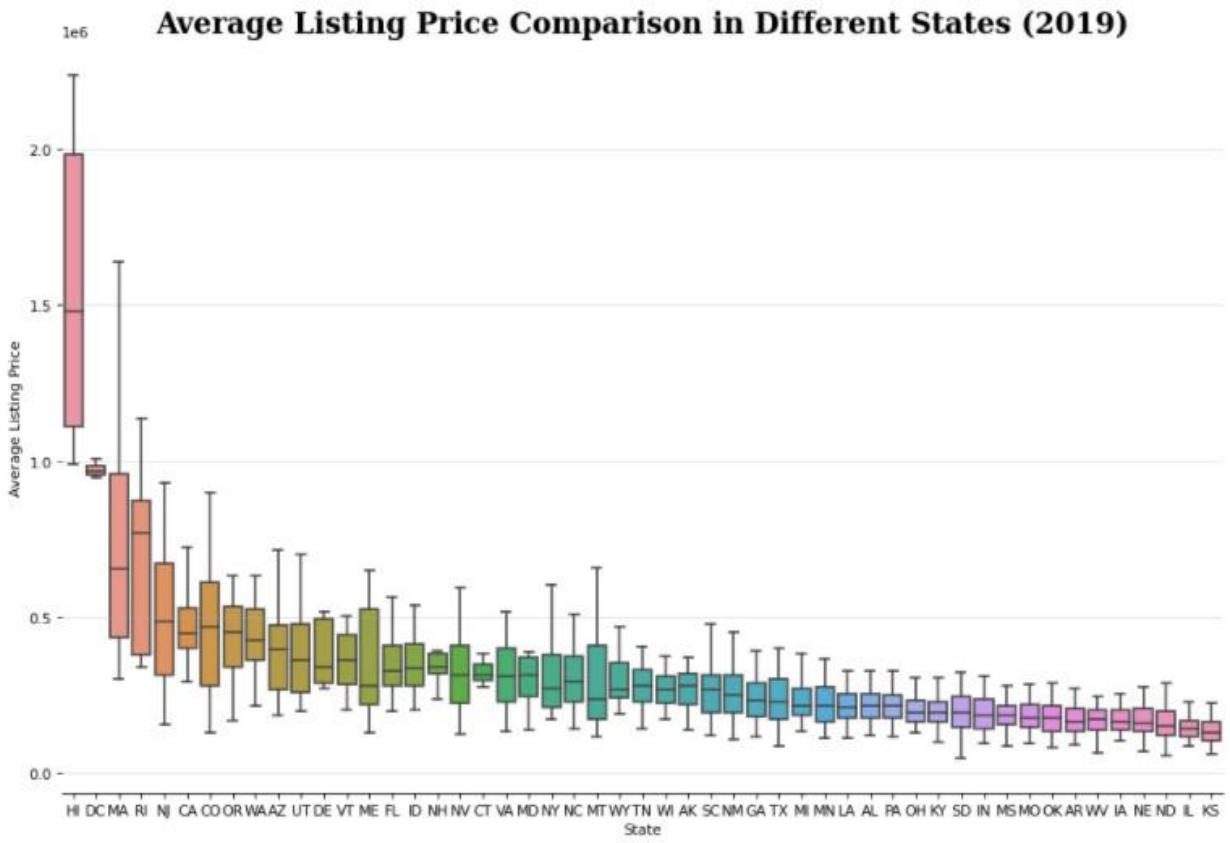
Investigative Relationship:

- Median Listing Price and Average Listing Price
- Average Listing Price y/y and Median Listing Price y/y
- Median Listing Price and Median Listing Price y/y
- Rural Category and Total Listing Count

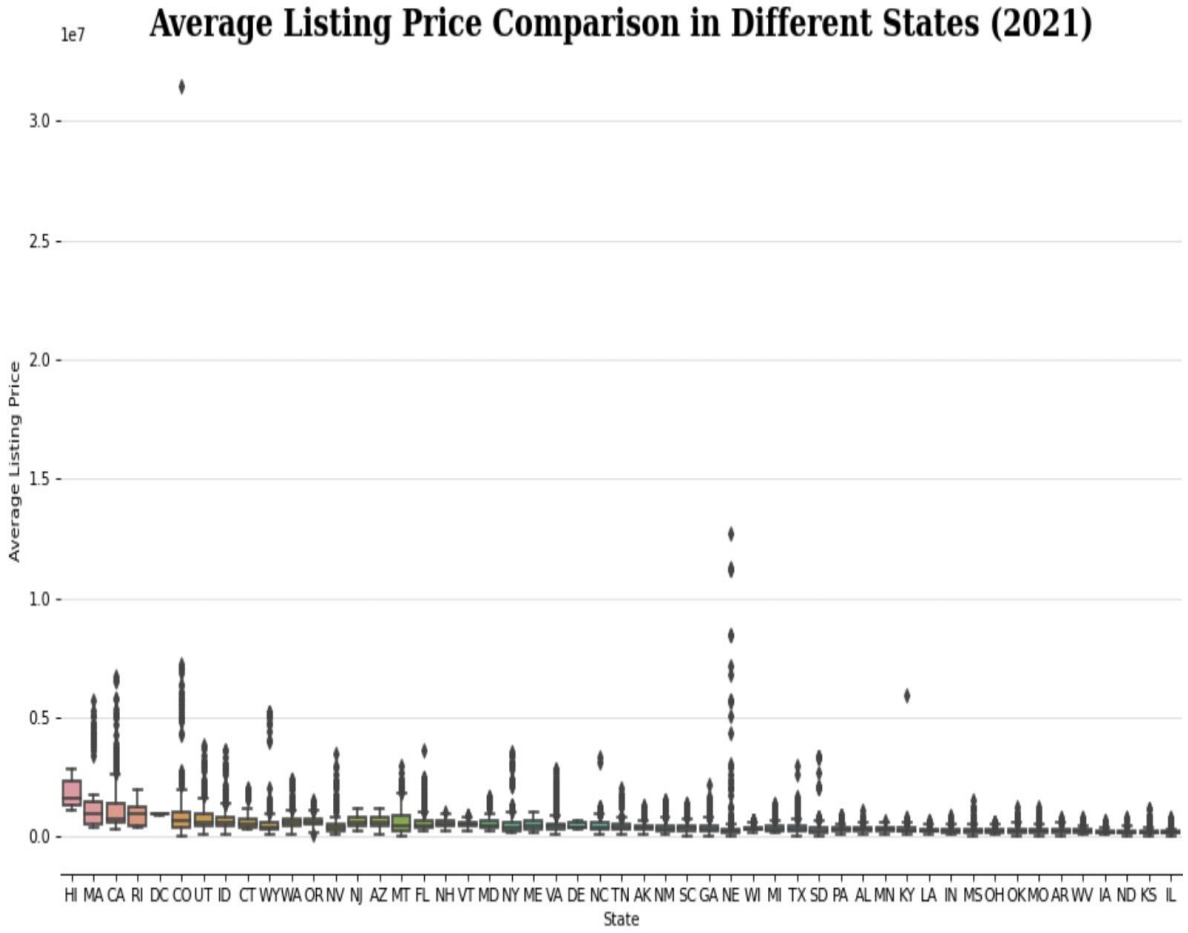


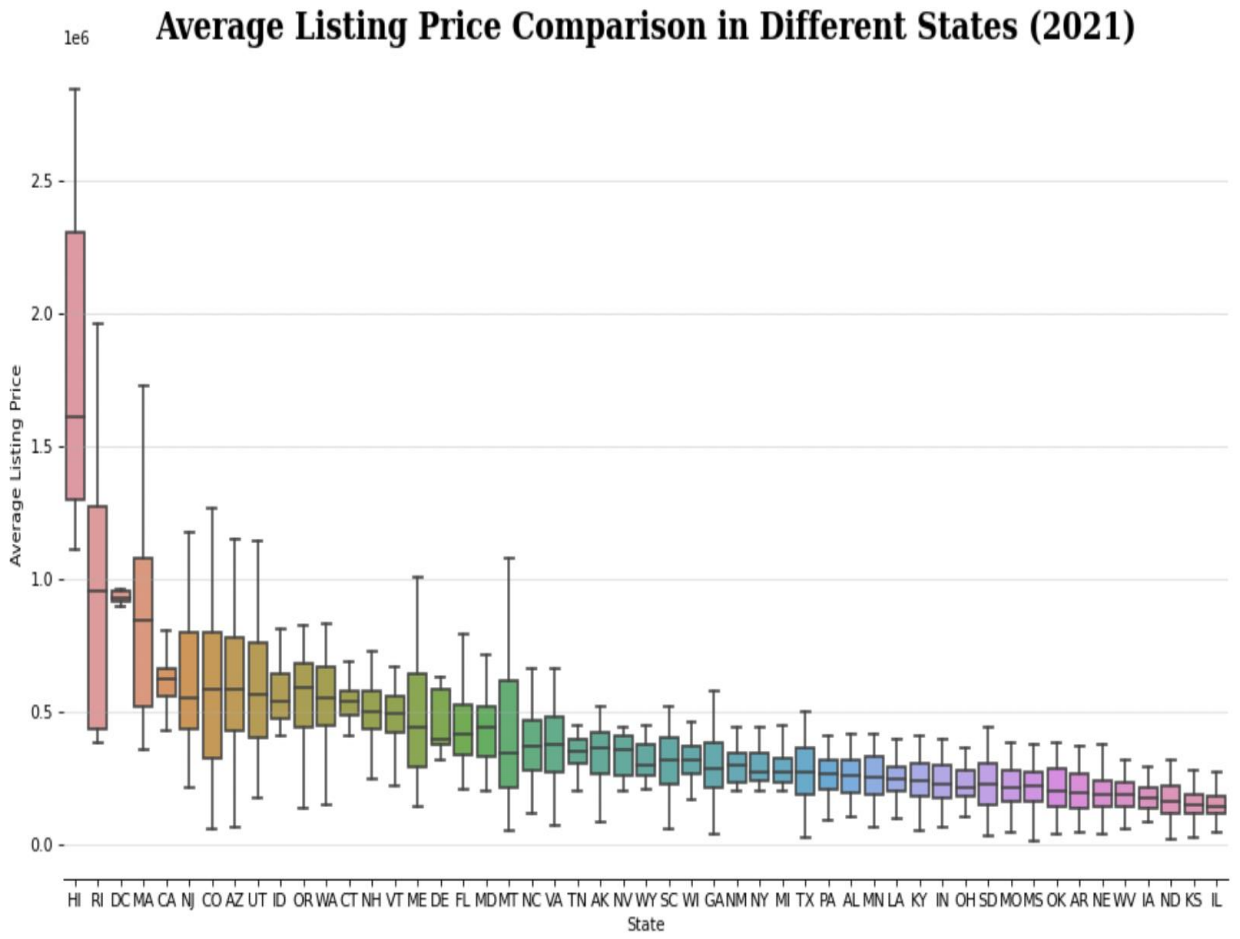
- On 2019, 54.84% of the counties in the United States experienced a price increased. 23.20 % of the counties experience down from last year and 21.96% down from last month.
- On 2021, 61.07% of the counties in the United States experienced a price increased. 22.40% of the counties experience down from last year and 16.53% down from last month.





- Hawaii, District of Columbia, and Rhode Island are some of the states that doesn't have outliers. To preserve the integrity of the dataset, outliers are removed in multiple batches.
- District of Columbia surpassed Massachusetts that has a higher average listing price mean after the outliers were cleaned.





- California and Massachusetts have outliers that were deleted in multiple batches.
- States with many outliers, likes Massachusetts, lose the counties after outliers were removed.

Analyzing Massachusetts' Counties

```
df_MA_2019['county'].unique()
```

```
array(['Barnstable County', 'Berkshire County', 'Bristol County',
       'Dukes County', 'Essex County', 'Franklin County',
       'Hampden County', 'Hampshire County', 'Middlesex County',
       'Nantucket County', 'Norfolk County', 'Plymouth County',
       'Suffolk County', 'Worcester County'], dtype=object)
```

```
df_MA_2019_cleansed['county'].unique()
```

```
array(['Barnstable County', 'Berkshire County', 'Bristol County',
       'Essex County', 'Franklin County', 'Hampden County',
       'Hampshire County', 'Middlesex County', 'Norfolk County',
       'Plymouth County', 'Suffolk County', 'Worcester County'],
      dtype=object)
```

```
df_MA_2021['county'].unique()
```

```
array(['Barnstable County', 'Berkshire County', 'Bristol County',
       'Dukes County', 'Essex County', 'Franklin County',
       'Hampden County', 'Hampshire County', 'Middlesex County',
       'Nantucket County', 'Norfolk County', 'Plymouth County',
       'Suffolk County', 'Worcester County'], dtype=object)
```

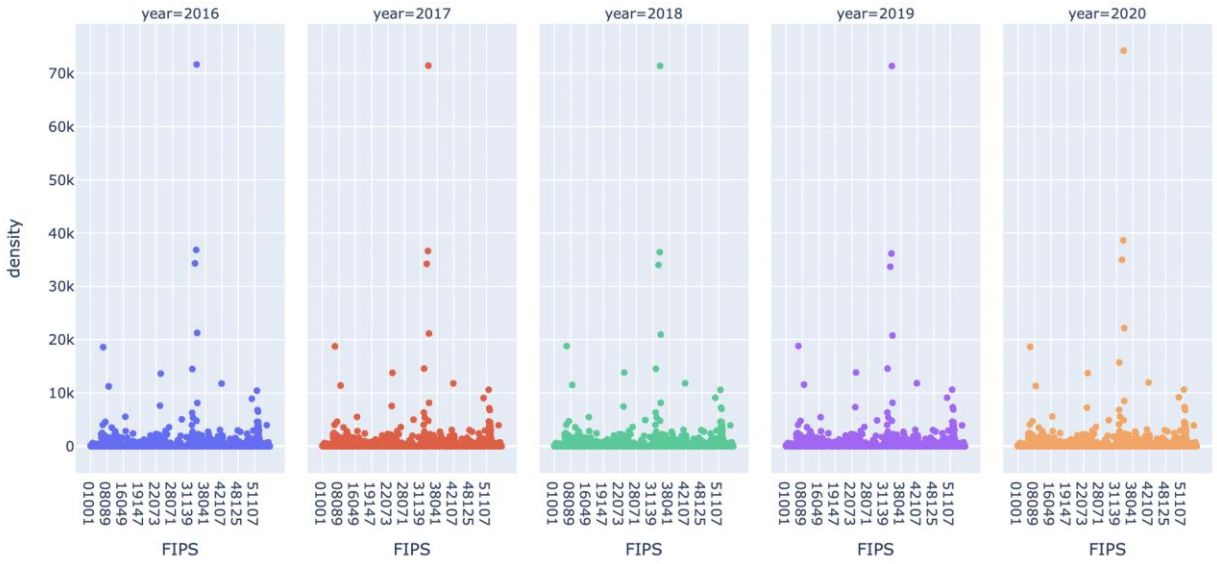
```
df_MA_2021_cleansed['county'].unique()
```

```
array(['Barnstable County', 'Berkshire County', 'Bristol County',
       'Essex County', 'Franklin County', 'Hampden County',
       'Hampshire County', 'Middlesex County', 'Norfolk County',
       'Plymouth County', 'Suffolk County', 'Worcester County'],
      dtype=object)
```

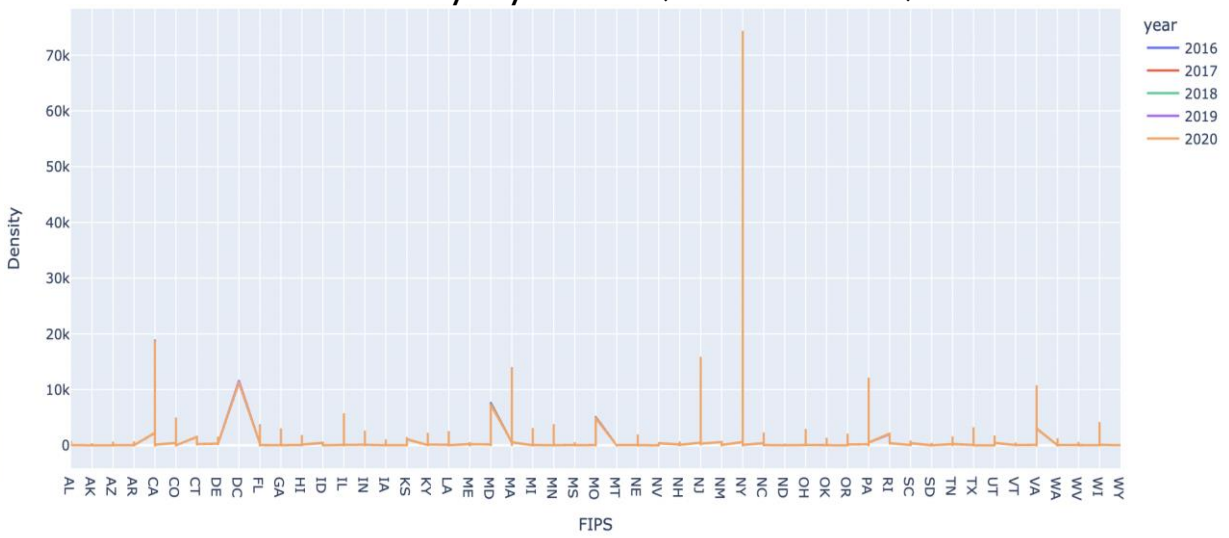
- Massachusetts has fourteen counties.
- After the outliers were removed, Massachusetts lost two counties, Duke and Nantucket, whose counties have very high average listing price.

Comparing the Density

Density by FIPS (2016 – 2020)

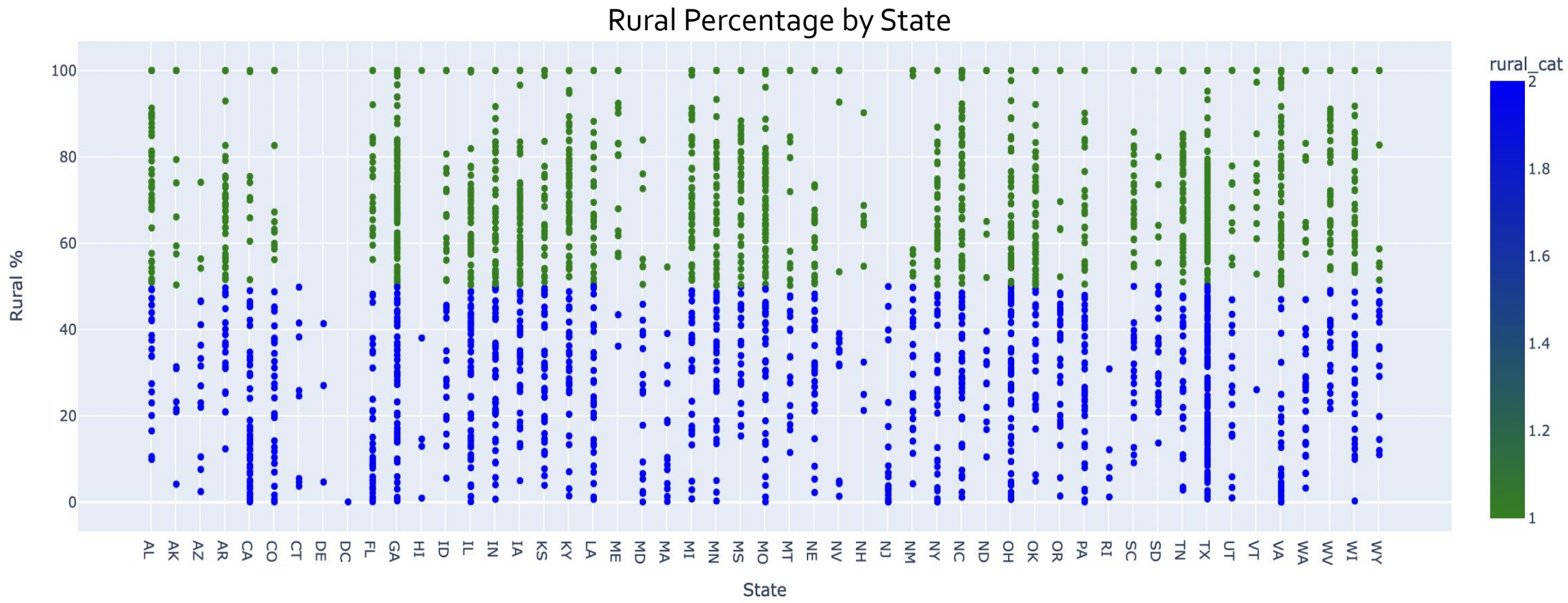


Density by State (2016 – 2020)



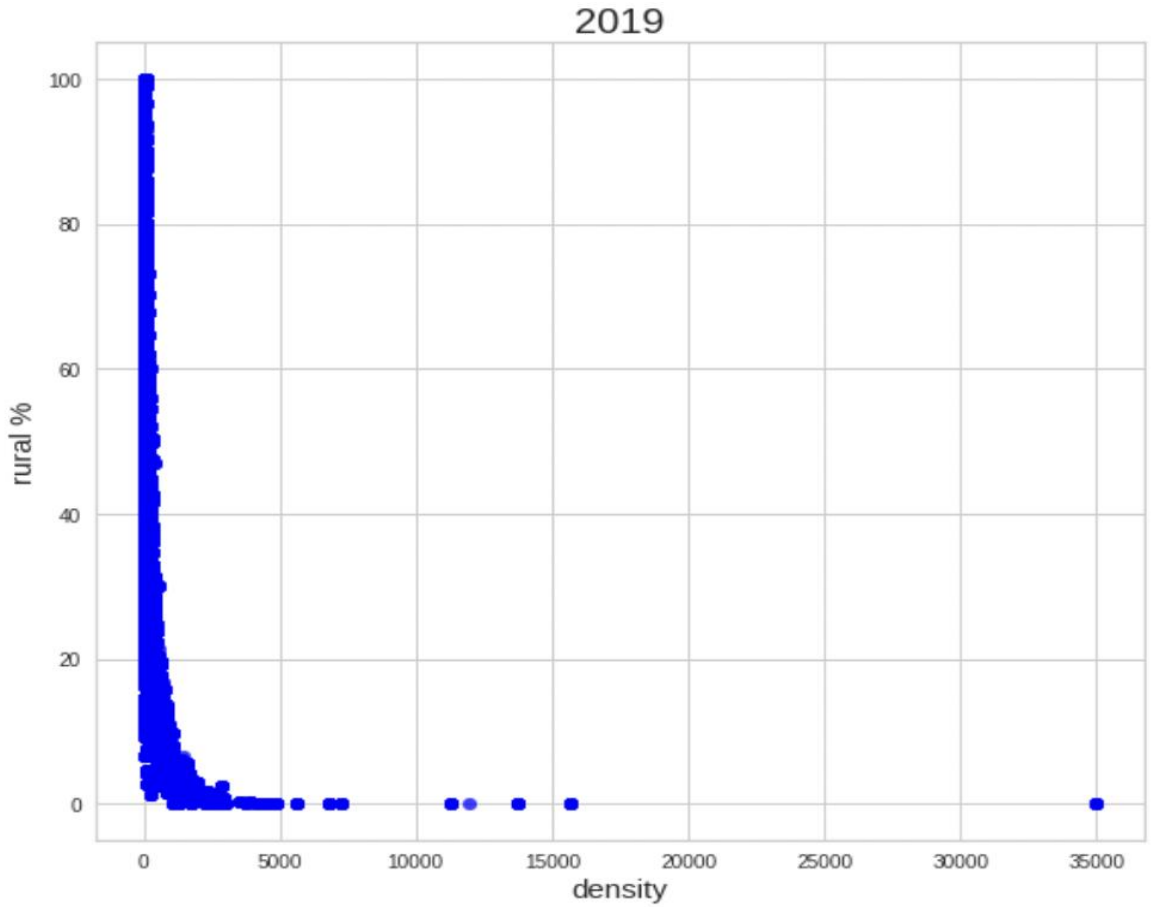
- The density for 2016 to 2020 are almost the same, so the density used is 2020 because census gathered the actual population and land area data on 2020.
- The states with very high density are New York, California, New Jersey, Massachusetts, Pennsylvania, and Virginia.

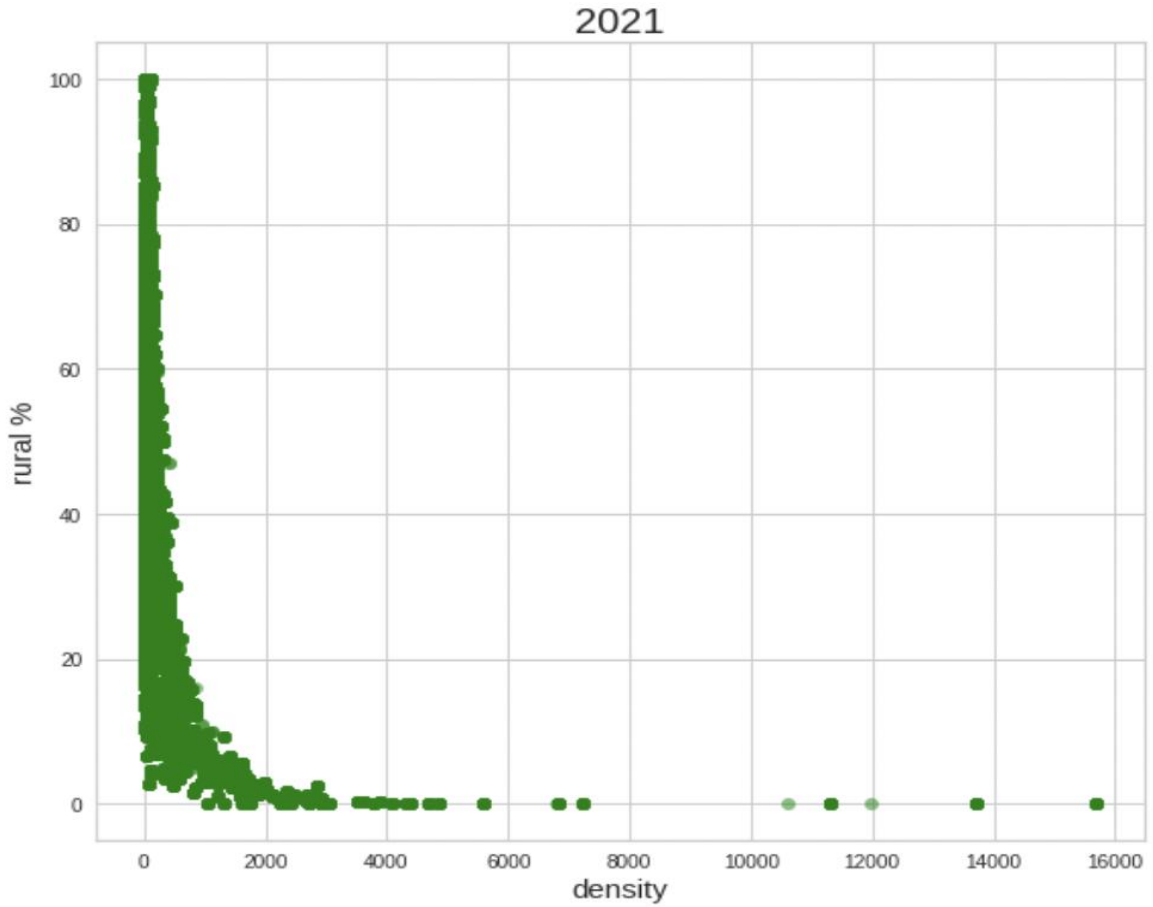
Analyzing the Rural Percentage



According to Ratcliffe, "completely rural counties have a population that is 100 percent rural. Mostly rural counties have a population that is 50.0–99.9 percent rural. Mostly urban counties have a population that is less than 50.0 percent rural."

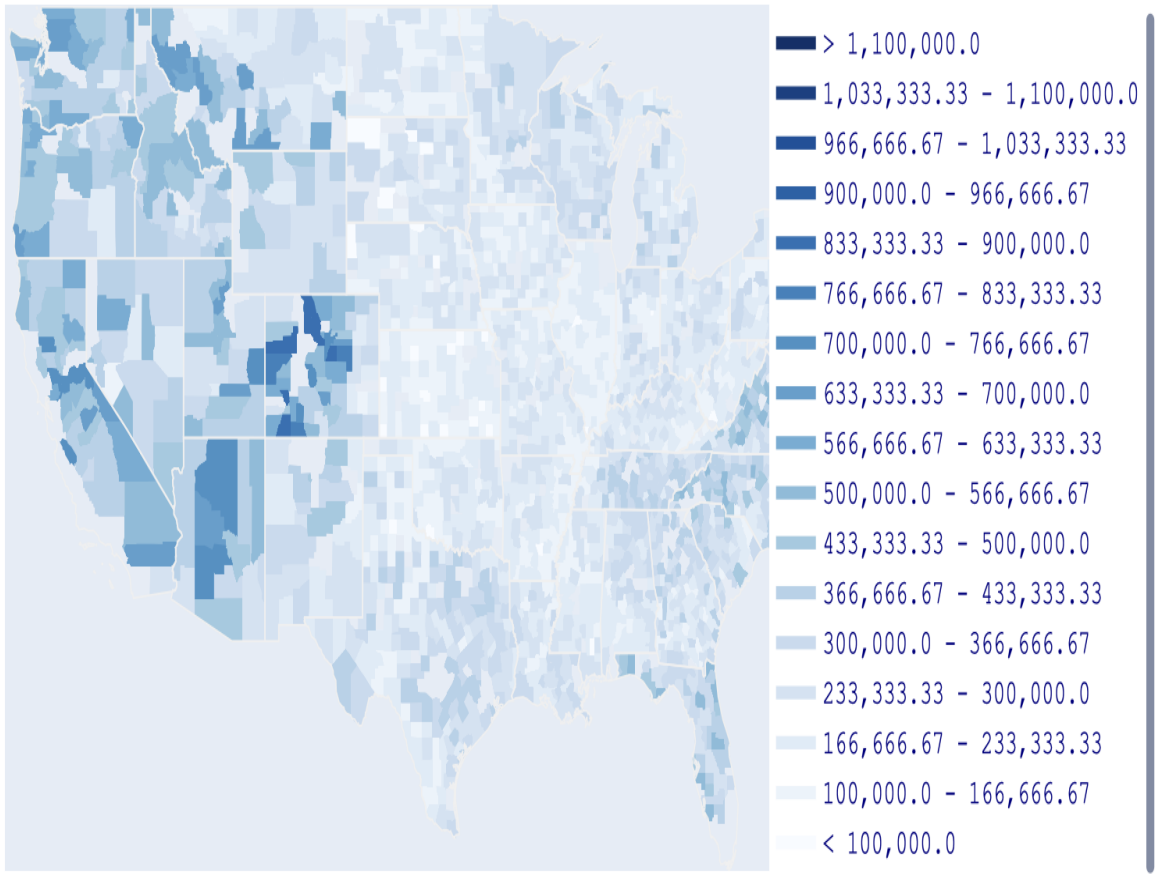
Relationship Between Rural Percentage and Density



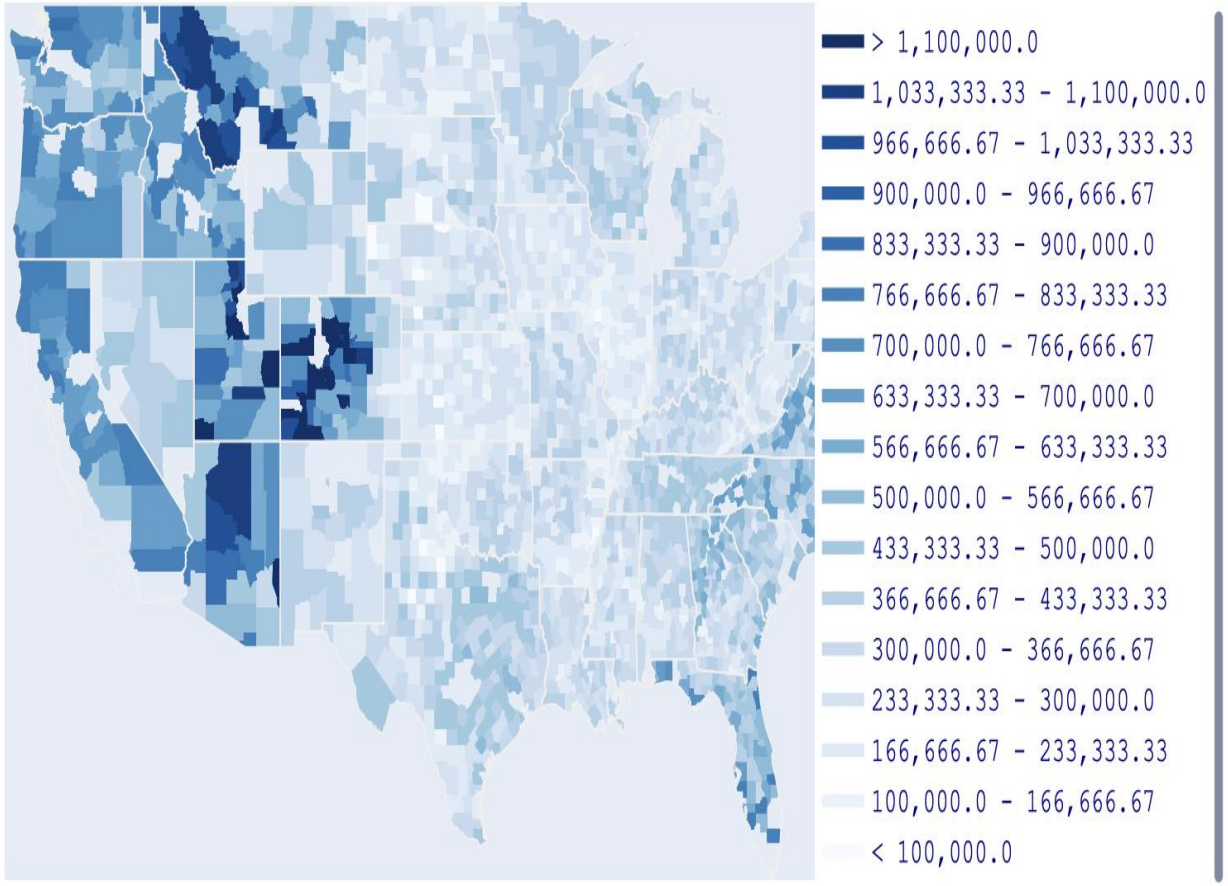


- Unevenly right skewed
- Negative correlation

2019 FIPS Average Listing Price

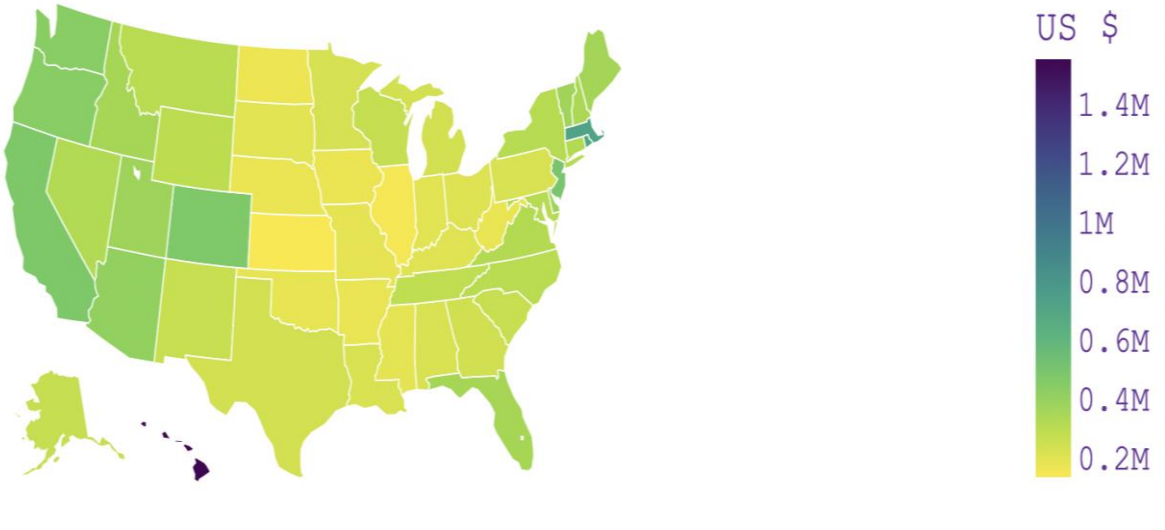


2021 FIPS Average Listing Price

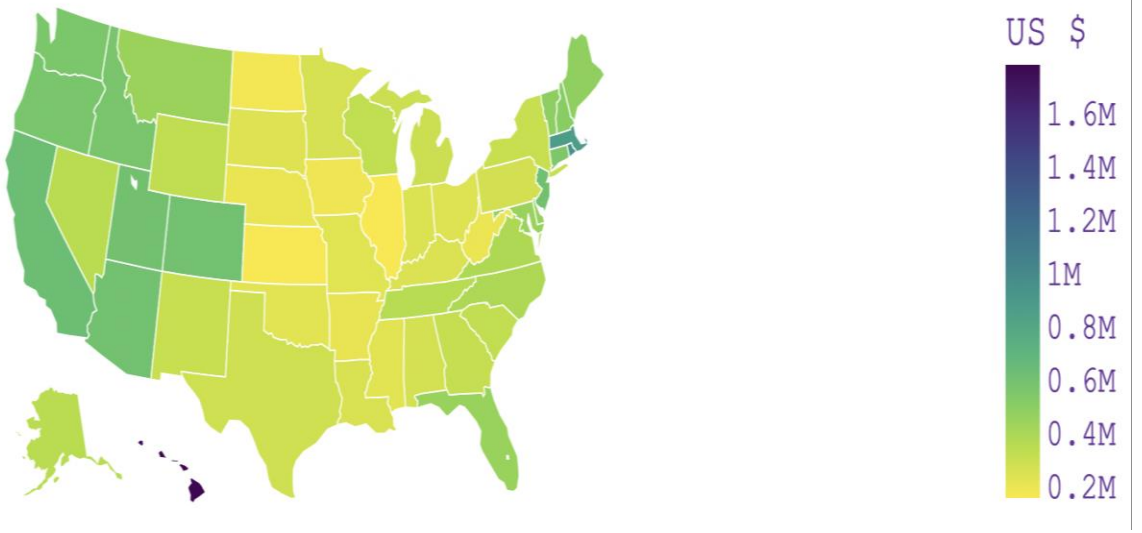


- In 2021, the color deepen in the whole United States which meant prices have gone up.
- In the Western Mountain region (Montana, Utah, Arizona, and Colorado), some of its counties' color is very distinct meaning high average listing price.

2019 US Average Listing Price per State

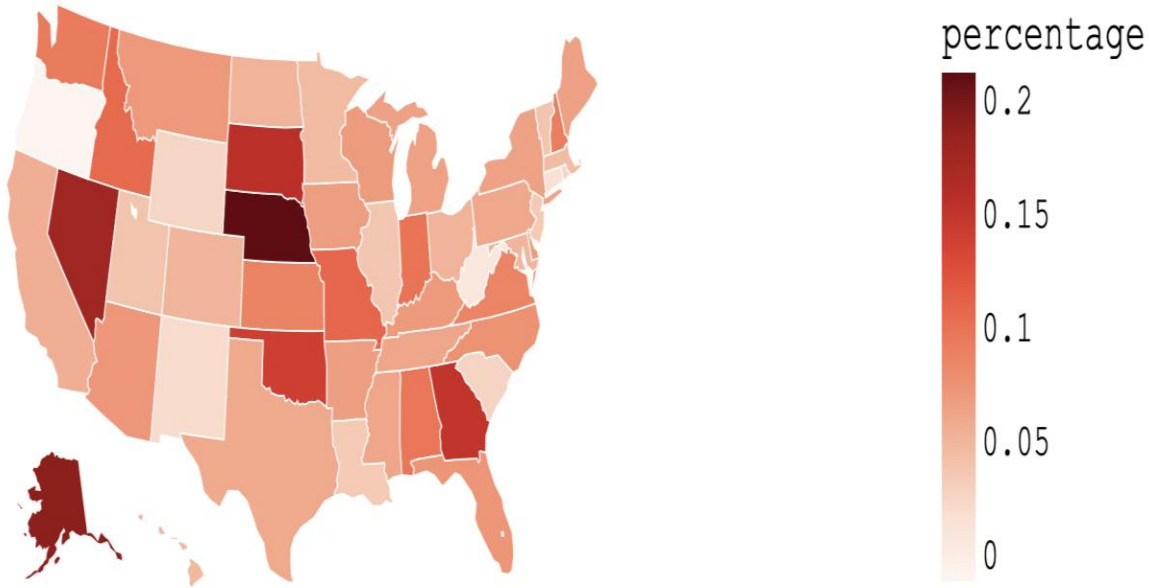


2021 US Average Listing Price per State

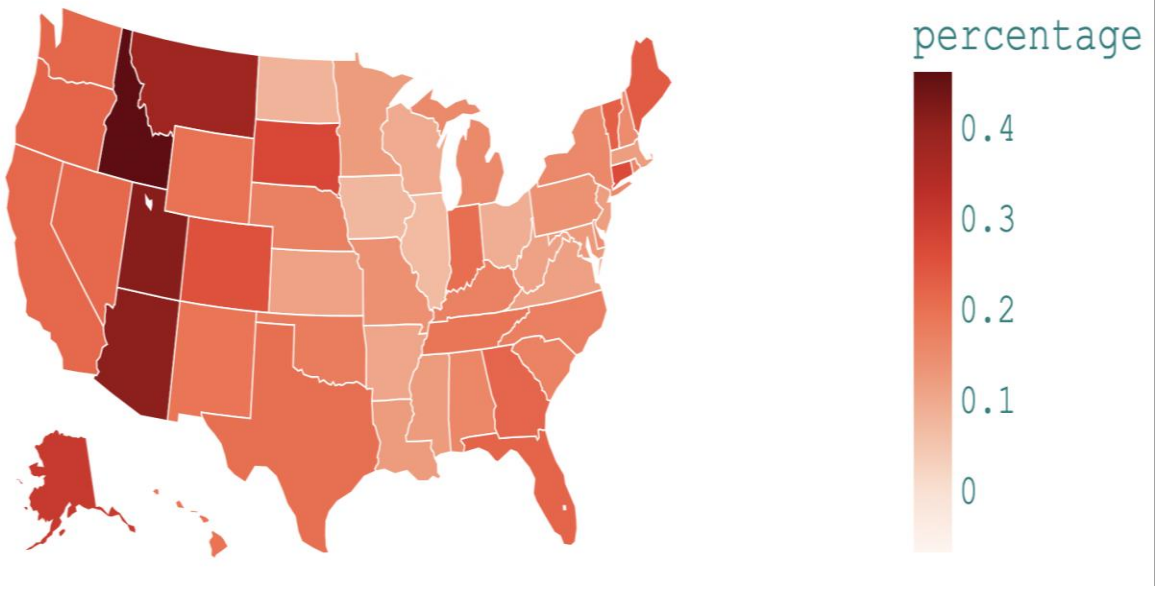


- In 2019, California, Washington, and Oregon have average listing price mean in the upper 400K's.
- Utah has a lower density (39) than Washington's (115) and Oregon's (44).
- In 2021, Utah's average listing price increased to 600K's which surpassed Washington and Oregon.
- In 2021, the color deepens more in the west coast which meant high increases.

2019 US Average Listing Price Change per State (%)

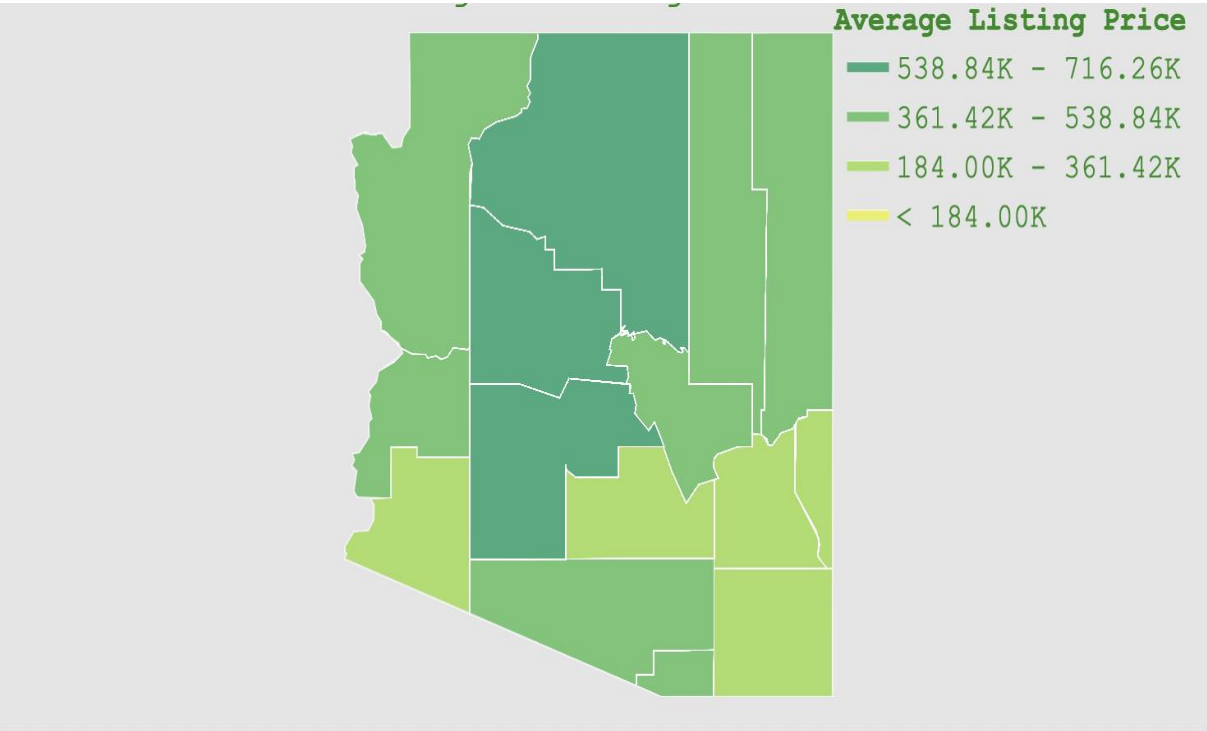


2021 US Average Listing Price Change per State (%)

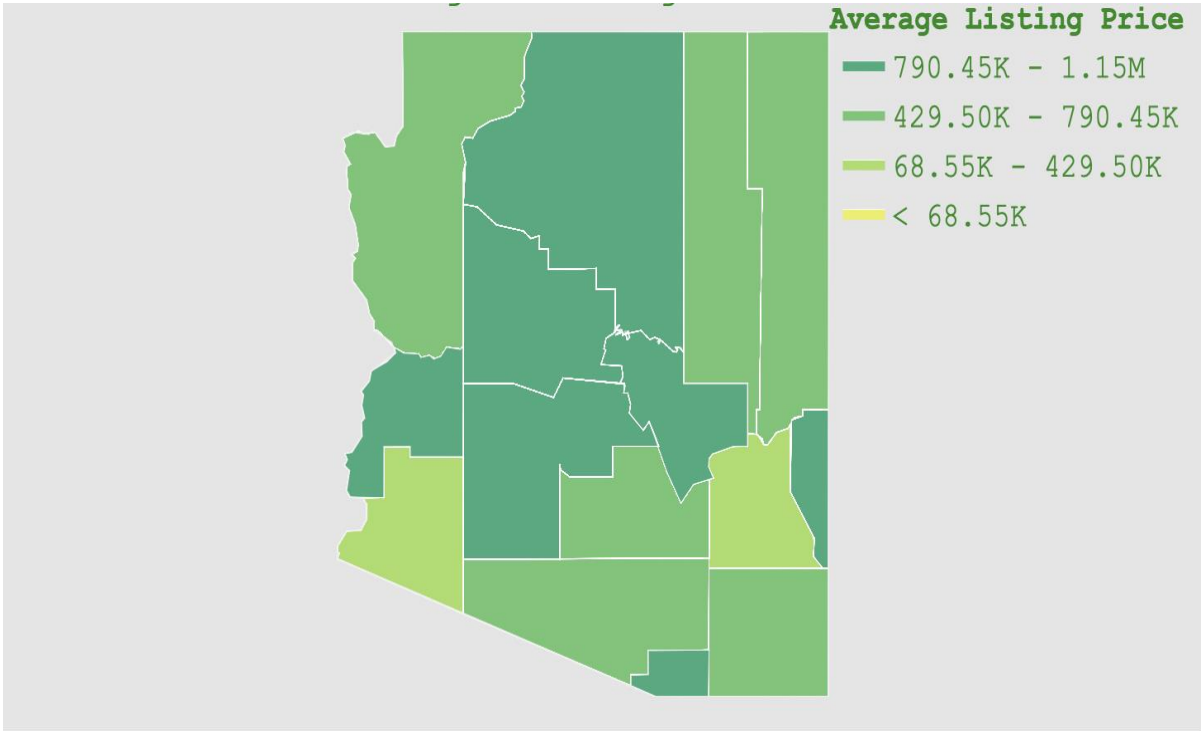


- In 2019, high average listing price mean increase was found in Alaska, Nevada, and Western North Midwest region.
- In 2021, the average listing price mean increases in West Mountain region's, such as Montana, Idaho, Utah, and Arizona, are higher.

2019 Arizona's Average Listing Price

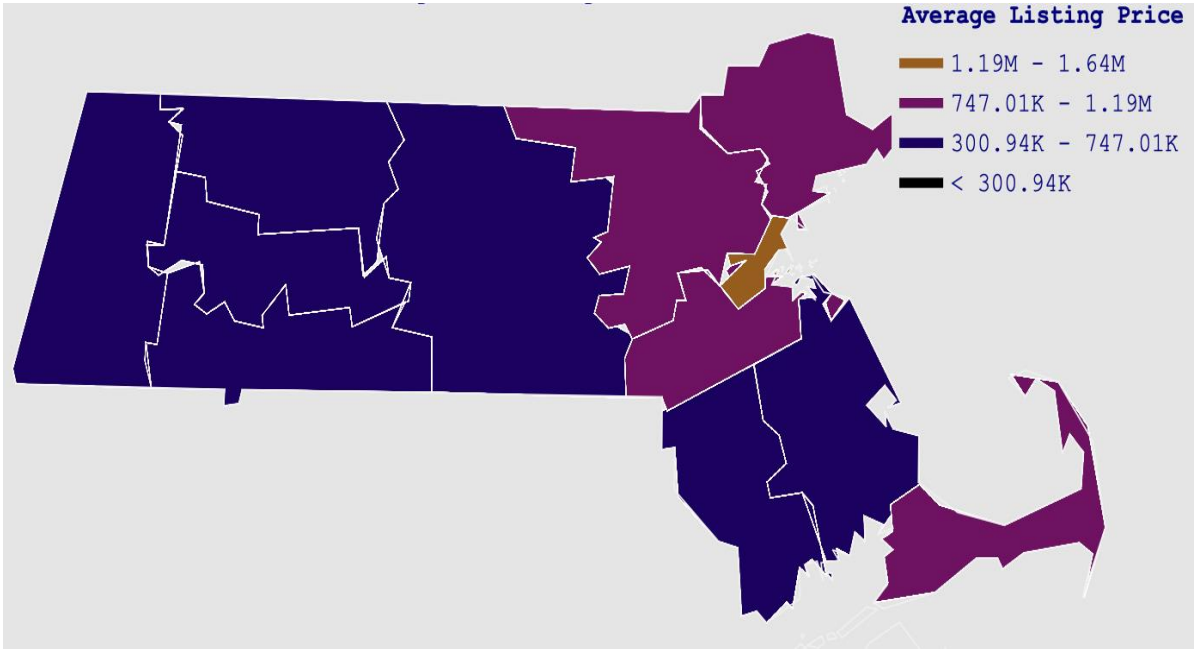


2021 Arizona's Average Listing Price

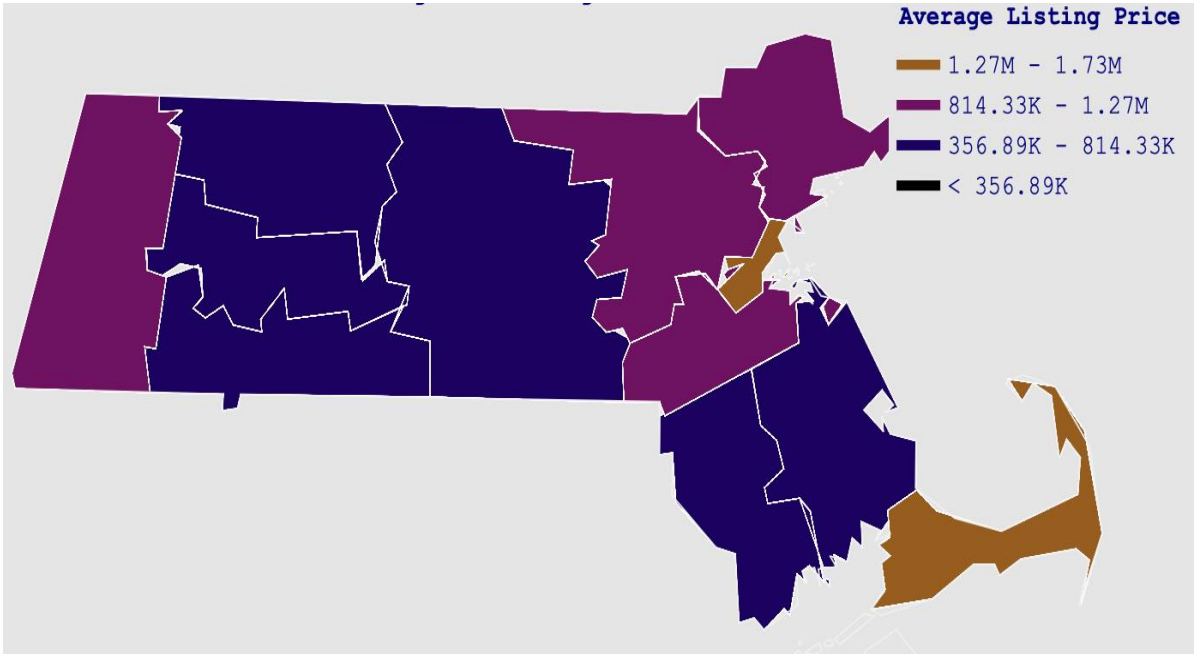


- Maricopa county with low 2.63 rural %, was displaced from 1st highest average listing price in 2019 to 3rd place in 2021 by Coconino, 31.47 rural %, and Yavapai, 33.20 rural %, counties.
- Greenlee county, an urban area with high 46.57 rural %, has a tremendous average listing price increase from 239K (in 2019) to 687K (in 2021).

2019 Massachusetts Average Listing Price



2021 Massachusetts Average Listing Price



- Berkshire county, with 31.59 rural %, surpassed the average listing price of Plymouth and Essex counties in 2021.
- Plymouth and Essex, have lower rural percentage (4.24 and 10.31 rural % respectively) than Berkshire..
- Barnstable county, 7.48 rural %, displaced Middlesex county, 3.03 rural %, from its 2nd placement to 3rd placement.

Machine Learning Model

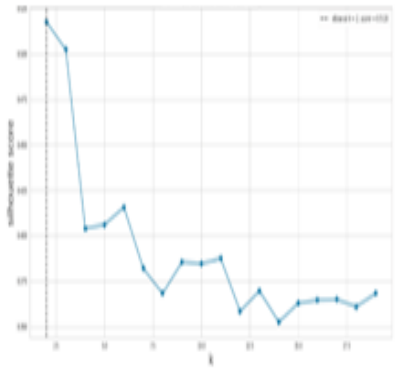
The machine learning model to be used for this project is K-means with centroid. K-means works by selecting how many clusters, k, exist in the data. Each cluster will be represented by a centroid (average) of similar points with continuous features. Principal Component Analysis (PCA), a projection technique, is used to reduce the number of dimension for a dataset.

K-means Learning Model is applied to the average listing price, rural percentage, and density on the county, state, and individual state levels.

UNITED STATES COUNTIES

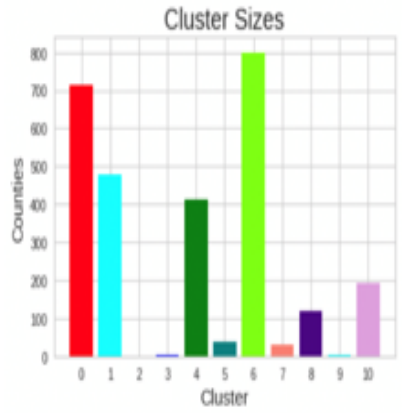
Average Listing Price, Density, Rural Percentage K-Means

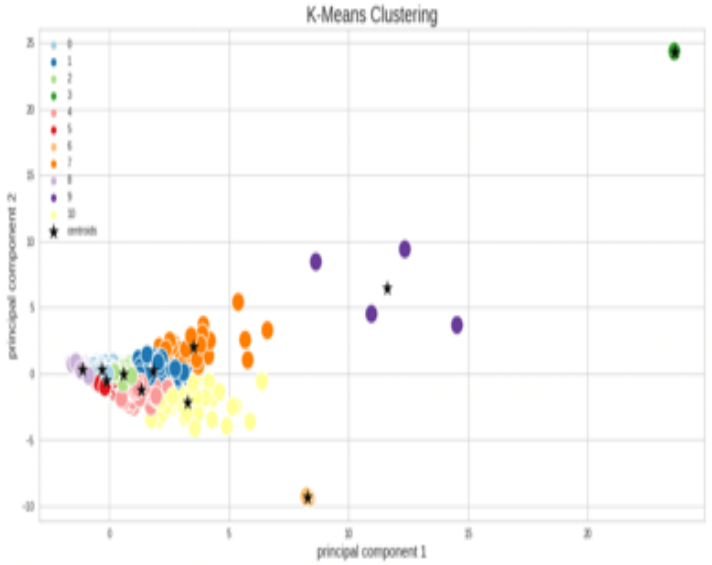
2019



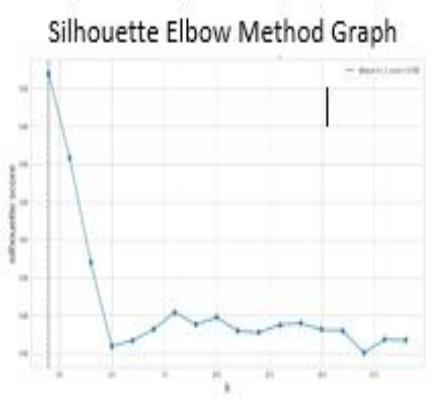
Results

- Optimal K – 11 clusters
- Inertia – 679.29
- Convergence - 29 iterations



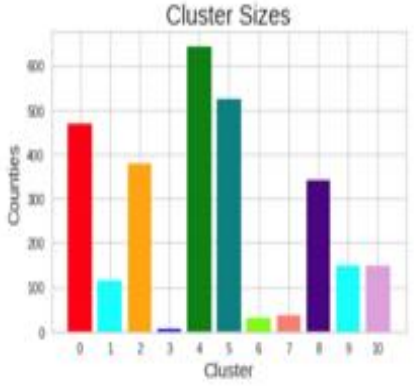


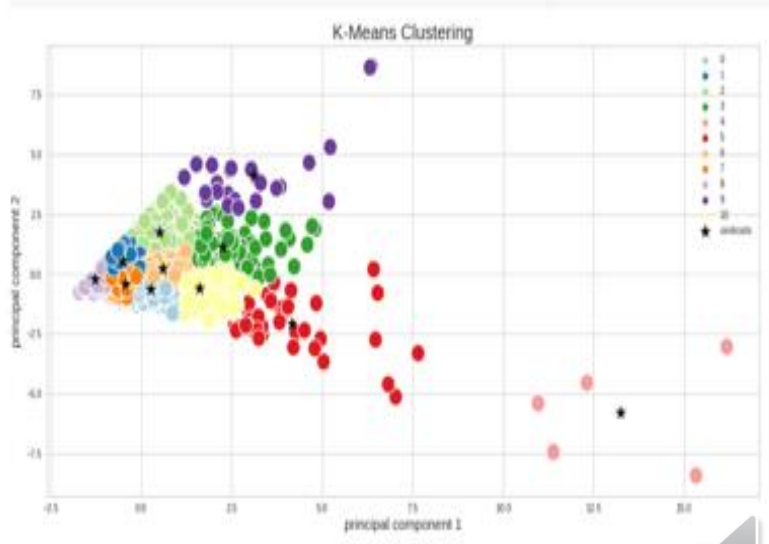
2021



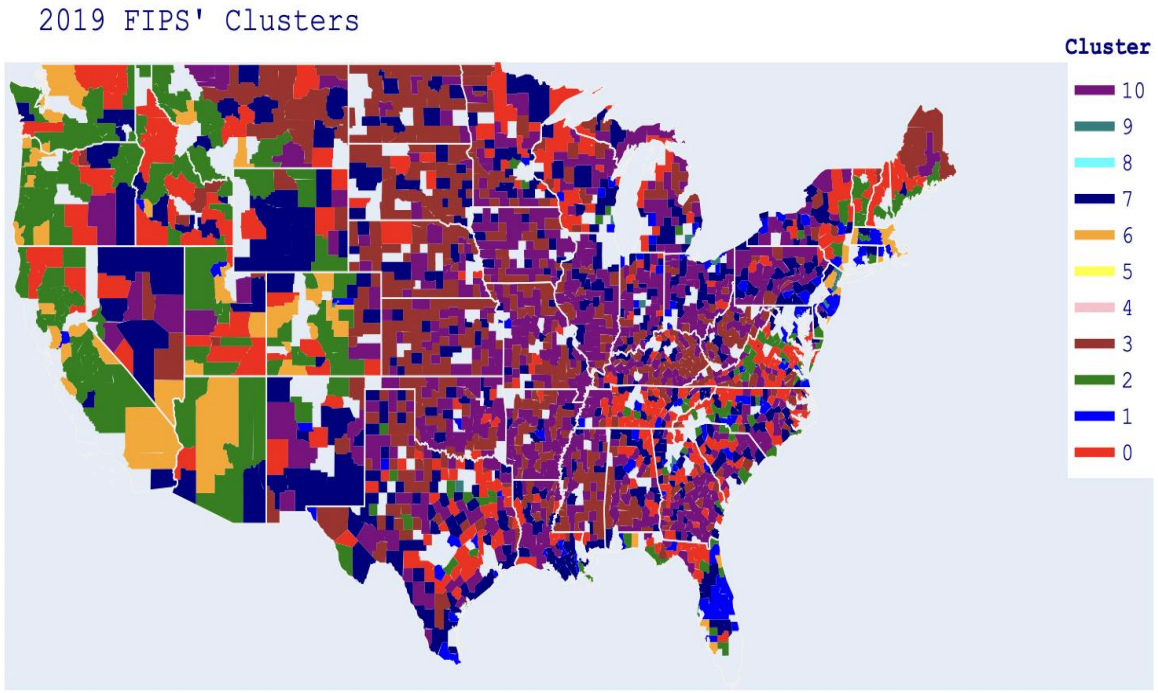
Results

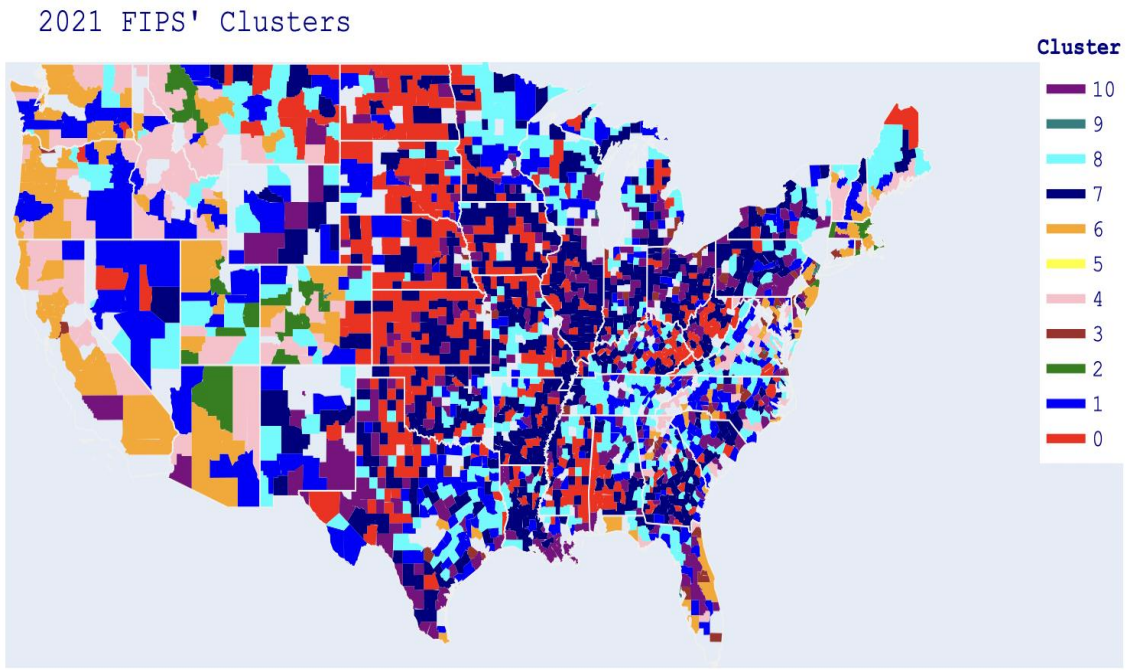
- Optimal K – 11 clusters
- Inertia- 872.44
- Convergence – 23 iterations





FIPS K-Means
Choropleth Map





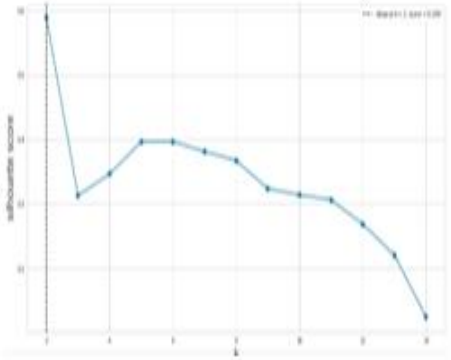
- Dominant brown cluster begins from Midwest region going towards East in 2019
- In 2021, the cluster in the Midwest region is isolated from the other regions. The formerly brown East North Central region onward to east belongs to their own cluster

ARIZONA'S COUNTIES

Average Listing Price, Density, Rural Percentage K-Means

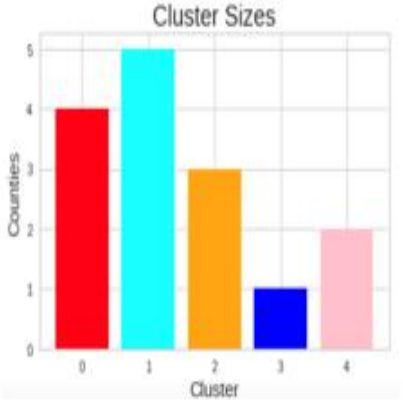
2019

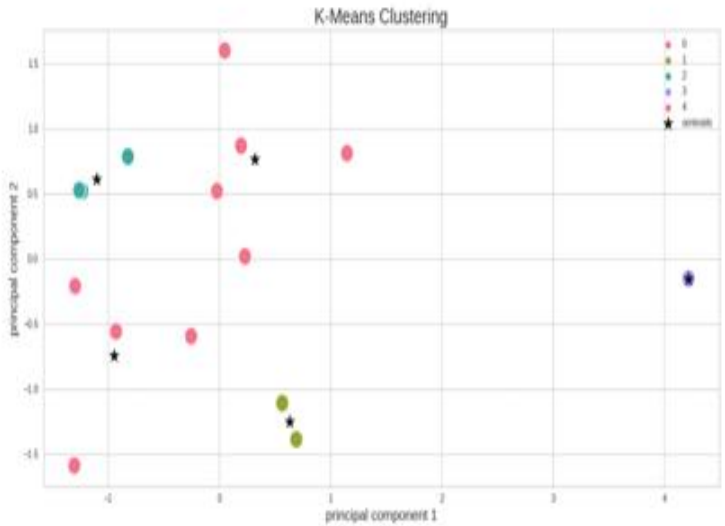
Silhouette Elbow Method Graph



Results

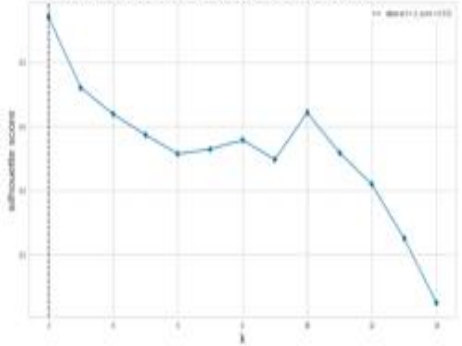
- Optimal K – 5 clusters
- Inertia – 4.2218
- Convergence - 3 iterations





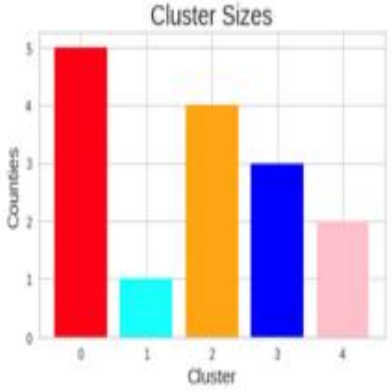
2021

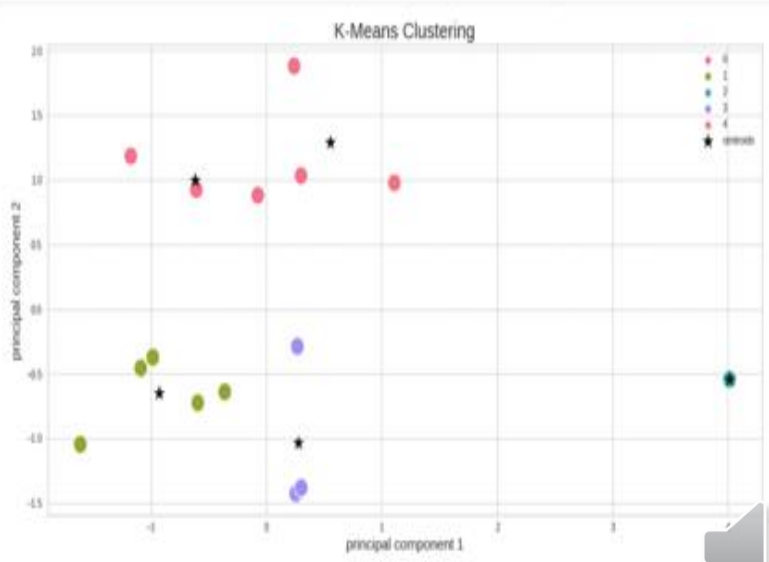
Silhouette Elbow Method Graph



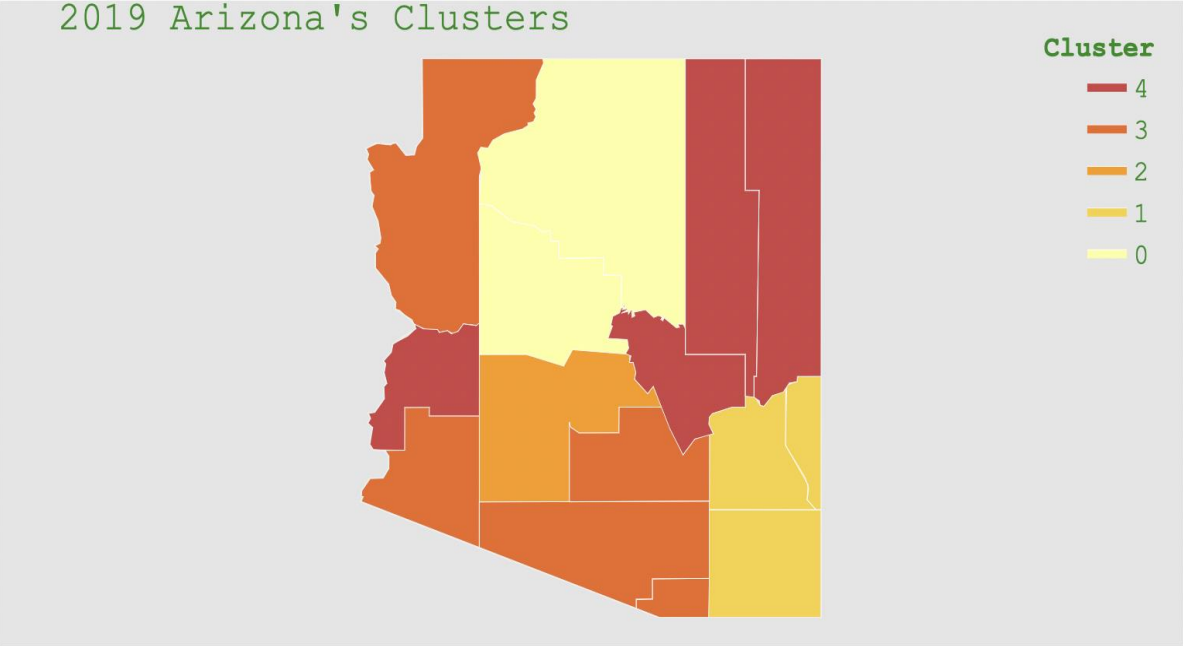
Results

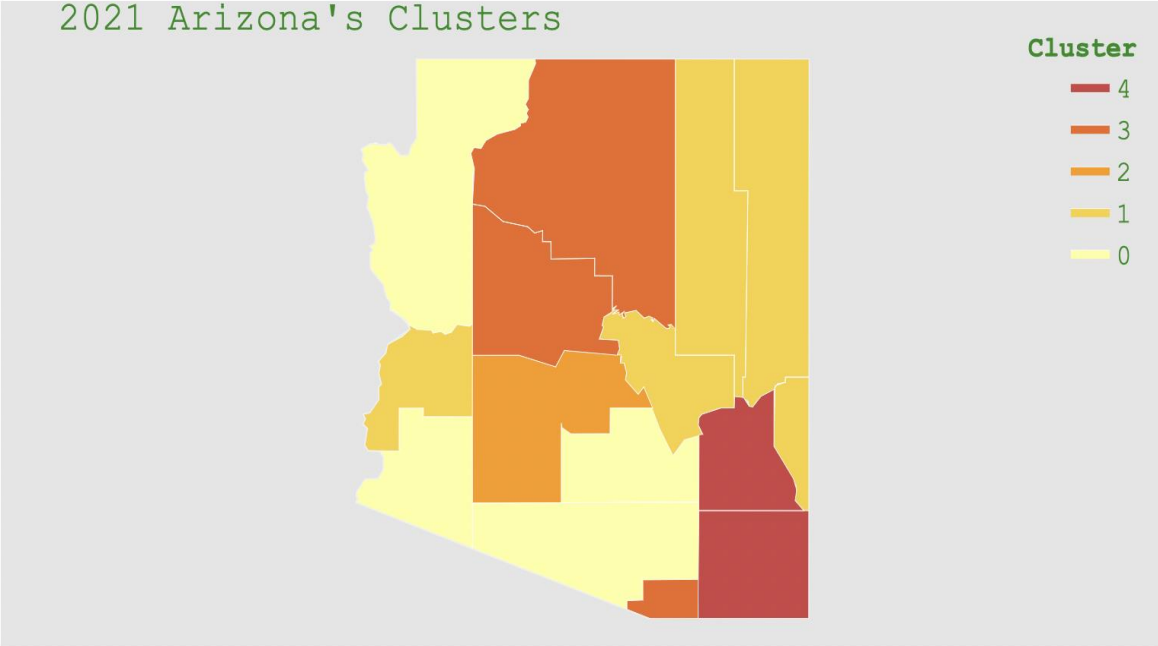
- Optimal K – 5 clusters
- Inertia- 3.6381
- Convergence – 2 iterations





Arizona K-Means
Choropleth Map



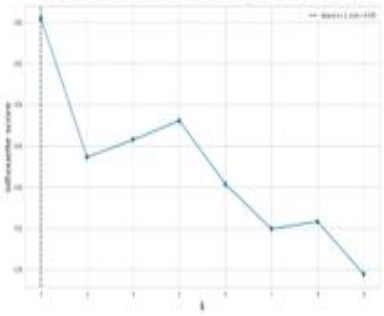


- In 2019, Greenlee and La Paz counties belong to a cluster whose average listing price is around 250K's. Greenlee is a low density urban county with 46.56 rural %. La Paz is a low density rural county with 56.32 rural %.
- In 2021, Greenlee and La Paz counties increased their average listing price to around 600K and was clustered with higher priced counties, Gila, Apache, and Navajo.

MASSACHUSETTS' COUNTIES

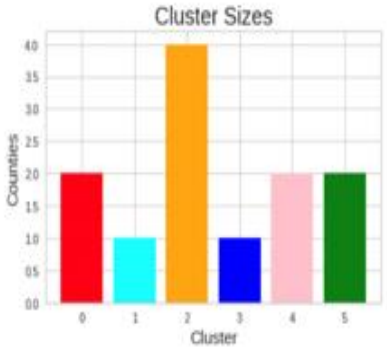
Average Listing Price, Density, Rural Percentage K-Means

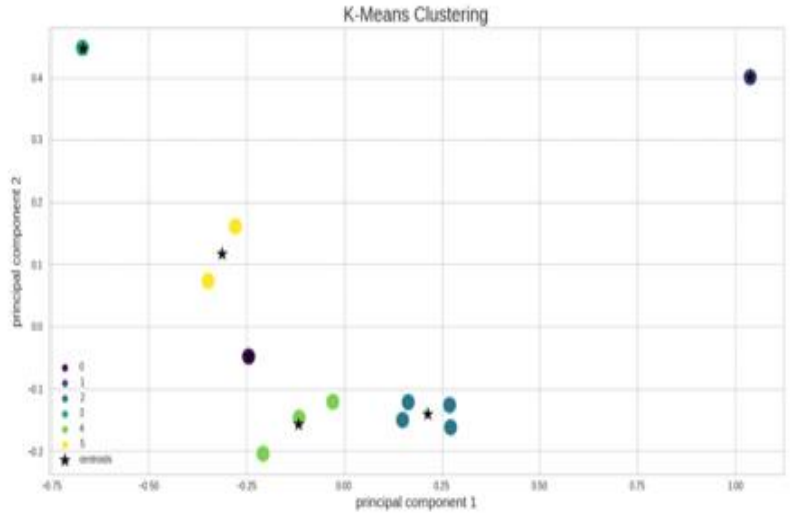
2019



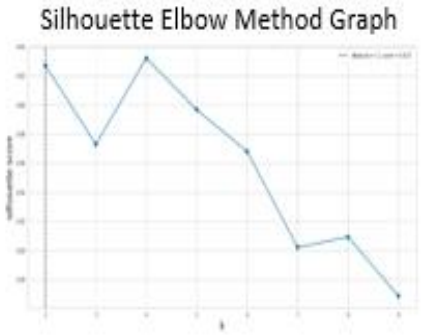
Results

- Optimal K – 6 clusters
- Inertia- 0.0375
- Convergence- 2 iterations



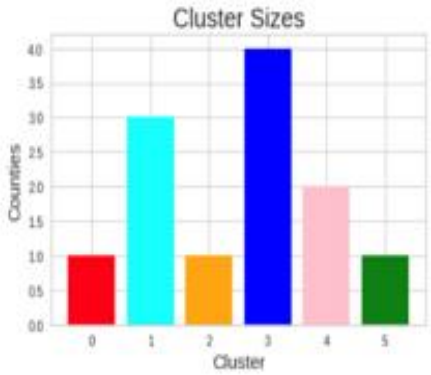


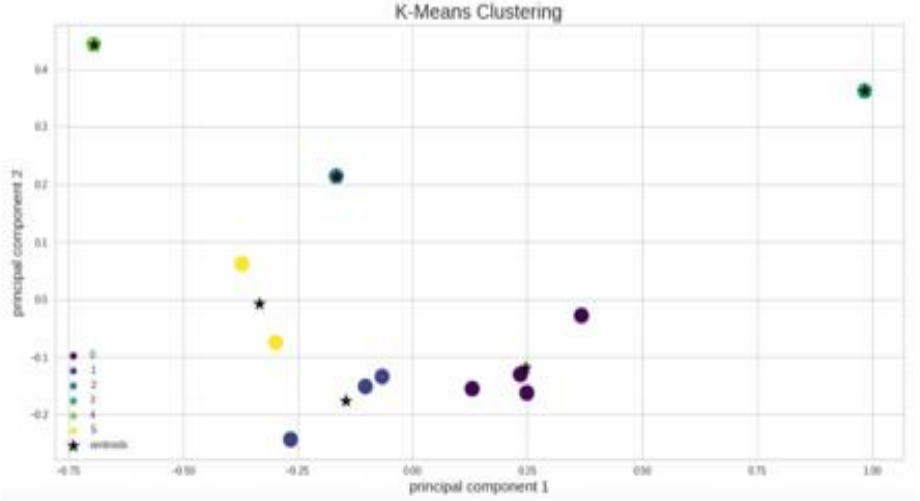
2021



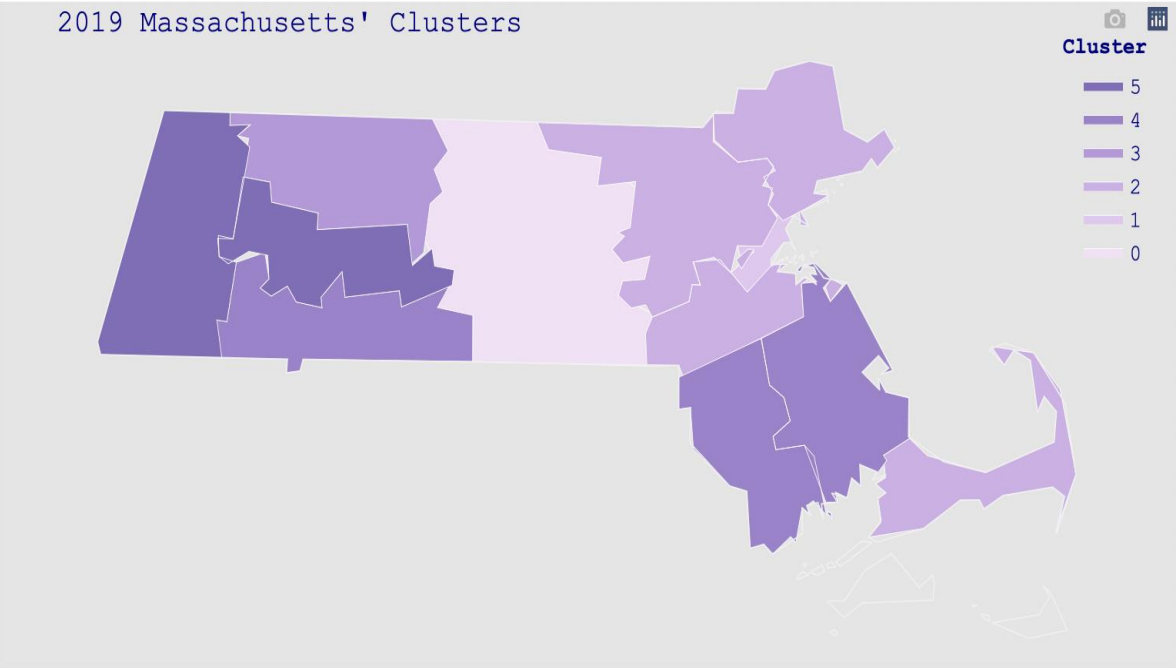
Results

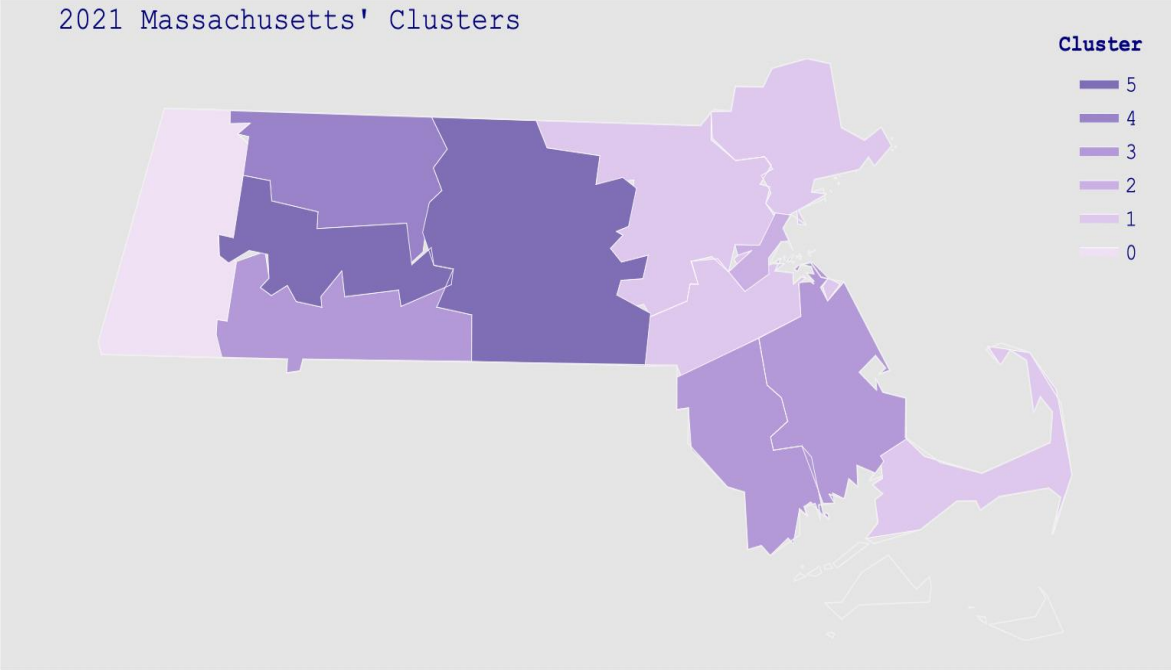
- Optimal K – 6 clusters
- Inertia- 0.0816
- Convergence – 2 iterations





Massachusetts K-Means
Choropleth Map





- In 2019, Berkshire and Hampshire counties belong to one cluster.
- Berkshire is less dense and higher rural percentage than Hampshire county
- In 2021, Berkshire is in its own cluster and its average listing price rose from 631K to 975K.

Interpretation and Implementation

1. Has there been a change in house prices since COVID-19?

Both the visualization and machine learning techniques reflect the change in house prices since COVID-19. The pre-pandemic, around 56% of the counties experience a price increase. The post-pandemic, around 61% of the counties experience price increase. The Comparison of 2019 and 2021 United States' Counties Choropleth Map shows the deepening of the shades across the United States in 2021 compared to 2019. This meant that average listing price increased. The Comparison of 2019 and 2021 United States' States Choropleth Map shows that the average listing price changed has an increased from 20% to 40%.

2. Has there been a change in in-demand house location since COVID-19?

There is a change in in-demand house location since COVID-19. There is a demand for houses in low density and/or high rural percentage. The average listing price was used to measure of the demand of house location. Demand being constant, too low of inventory will increase the property's listing price while too high of inventory will decrease the property's listing price. District of Columbia is a high urban area with zero rural percentage. Before the pandemic, District of Columbia has second in the highest average listing price in the country. After the pandemic, District of Columbia was delegated to the third highest listing price. Its average listing price dropped from 970K to 934K while many counties enjoyed increased average listing price. On the other hand, Maricopa county, an urban area with low rural percentage, has the highest listing price in Arizona in 2019. In 2021, Coconino and Yavapai counties, high rural percentage counties, have higher average listing price than Maricopa county. Another example, Berkshire is low density and high rural percentage urban area in Massachusetts. It saw its average listing price rose higher than counties, Plymouth and Essex, that are more urban and denser.

3. Has there been a change in buyers' house buying criteria since COVID-19?

There is evidence that the buyers’ house buying criteria change because the choropleth map of the K-means clustering for the average listing price, density, and rural percentage shows the partition of the clustered. In 2019, there used to be 2 dominant clusters that separate the west and east region of the US. After the pandemic, there seems to have 3 clusters. The Midwest region has its own cluster and separates the east and the west regions. However, further investigation is needed because the visually seen cluster are made up of other small clusters. (Comparison of 2019 and 2021 United States' FIPS cluster).

References

Airgood-Obrycki, W. (2020, May 7). COVID-19 Will Delay Housing Construction, but for How Long? Joint Center for Housing Studies of Harvard University. Retrieved March 6, 2022 from https://www.jchs.harvard.edu/blog/covid-19-will-delay-housing-construction-but-for-how-long

Chen, D. Y., (2018) Pandas for Everyone. Boston: Pearson.

Friedman, N. (2022, February 23). Home-Price Growth Hit Record in 2021. *Wall Street Journal - Online Edition*, N.PAG.

Liu, S., & Su, Y. (2020). The Impact of the COVID-19 Pandemic on the Demand for Density: *Evidence from the U.S. Housing Market*. 47. Retrieved Feb. 3, 2022. https://yichensu.files.wordpress.com/2021/04/liusu_covid_4_11_21.pdf

References

Polzin, P. E. (2021). The Myth of Montana's Housing Market Frenzy : Has COVID-19 Related Migration Spurred the Recent Rise in Home Prices? *Montana Business Quarterly*, *59*(2), 10–13.

Raschka, S., Mirjalili V., (2015, September). Python Machine Learning. (2nd edn.) Packt: Birmingham.

Real Estate Economist Mark Dotzour Forecasts Post-COVID Suburban Migration on HomeVestors. (2020, July 10). *Entertainment Close-up*. https://link.gale.com/apps/doc/A629042656/BIC?u=umd_umbc&sid=bookmark-BIC&xid=fb84eb82

Zhai, W., & Peng, Z.-R. (2021). Featured Graphics: Where to Buy a House in the United States amid COVID-19? Environment and Planning A, 53(1), 9–11. Retrieved Feb. 3, 2022. https://doi-org.proxy-bc.researchport.umd.edu/http://epn.sagepub.com/content/by/year

THANK YOU