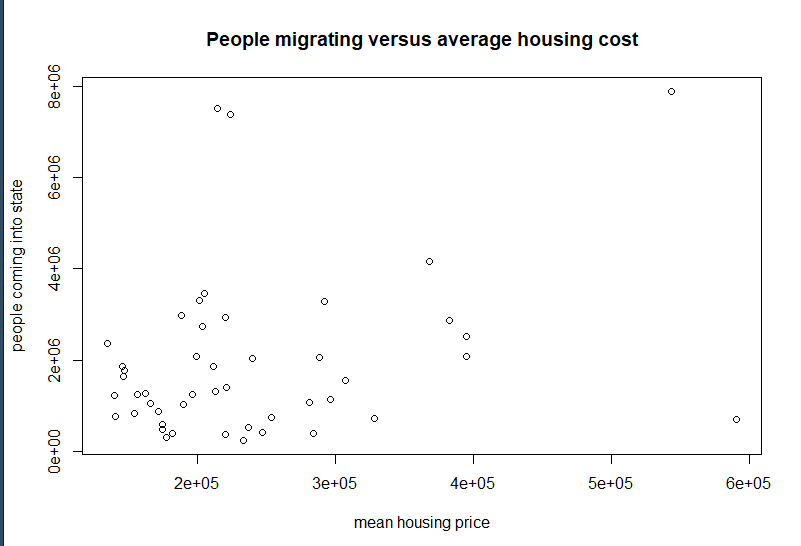
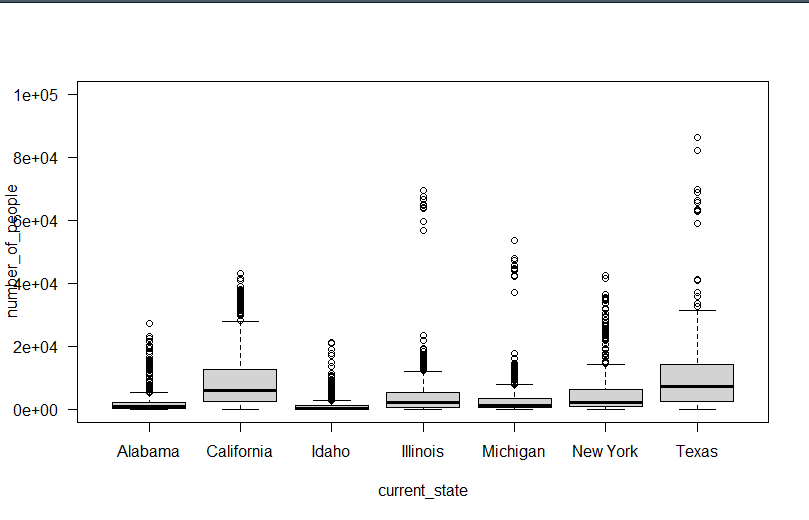
Goal: Using Rstudio to make visualization and perform research on the shared datasets. After combing through the datasets, some questions were raised about what’s happening with the housing market in the United States.

1. Is there a correlation between the mean housing price in the united states and the influx of people moving into those states?

people\_migrating\_versus\_mean\_housing\_price\_over\_all\_states\_all\_years.PNG

We can see that through the graph that people migrating into states is not solely from the housing price, or else the trend would be a reverse relationship. Upon viewing this graph, it is clear that mean housing price is not the only reason for people migration, if at all.

1. Show that California and Texas have significantly higher rates of movers than other states by comparing random states to each other. Use t-test or box-and-whisker graph to quantitatively prove this claim.

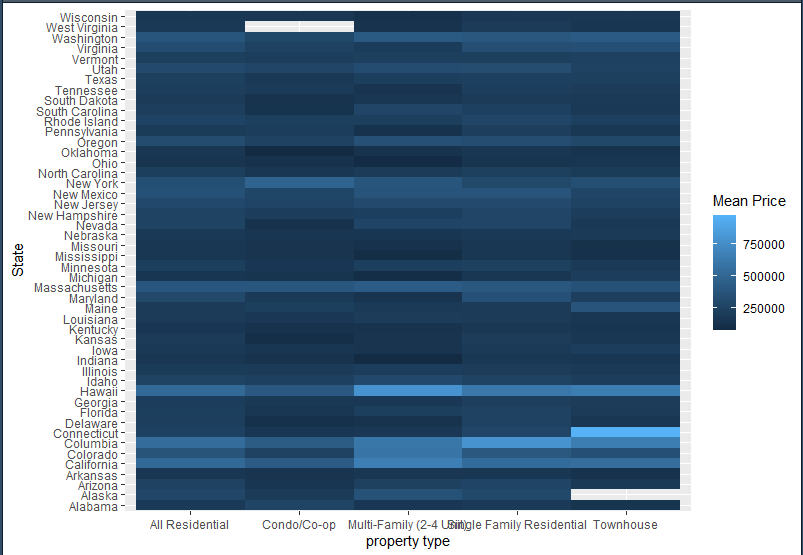


box\_whisker\_selected\_current\_state\_versus\_people\_moving\_ZOOMED.PNG

This box-and-scatter plot shows a random selection of states including Texas & California. The box and whisker for both California and Texas are significantly larger than the other states’, indicating that our states of interest have a statistically significant difference from others.

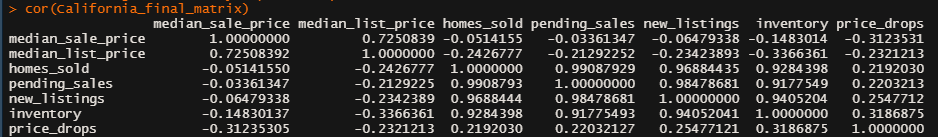
Note: California had significant outliers in the graph that were omitted from the diagram to allow proper display of the graph.

1. Do property types affect the mean housing prices in the United States?



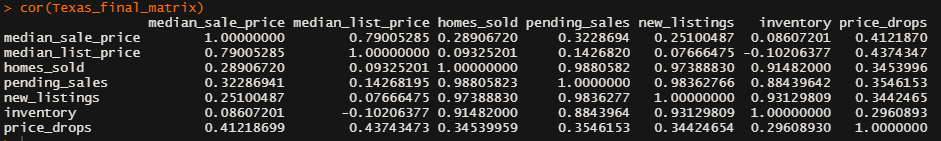
HEAT\_map\_state\_property\_type\_median\_price.PNG

The heat map presented shows the states versus property type, with the color-filled boxes representing the mean price throughout the years. One observation is that the “Single Family Residential” property type tends to have a higher mean price than its counterparts. However, not all states reflect this trend so we can say that the mean price is not solely affected by the type of property, though it could give some influence.

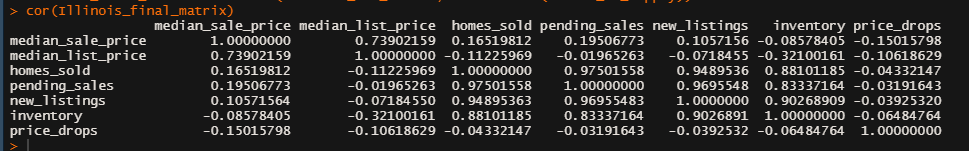


California\_cor\_matrix.PNG

* Median\_sale\_price = median\_list\_price
* Pending\_sales = homes\_sold
* Homes\_sold = new-listings = inventory
* Pending\_sales = new\_listings = inventory

Texas\_cor\_matrix.PNG

* Median\_sale\_price = median\_list\_price
* Homes\_sold = pending\_sales = new\_listings = inventory

Illinois\_cor\_matrix.PNG

* Median\_sale\_price = median\_list\_price
* Homes\_sold = pending\_sales = new\_listings = inventory
* Pending\_sales = homes\_sold =new\_listings = inventory

We can see that besides “median\_list\_price”, there are no correlations between “median\_sale\_price” among the other variables. It’s important to note that the median sale price in this dataset is not affected by the “homes\_sold”. This data suggest that if the median sale price goes up as people are moving into states & buying the available homes, the correlation would be an inverse (< -0.7) relationship of homes sold. Since this is not the case, this further indicates that there are other variables at play in the rise of mean housing prices.