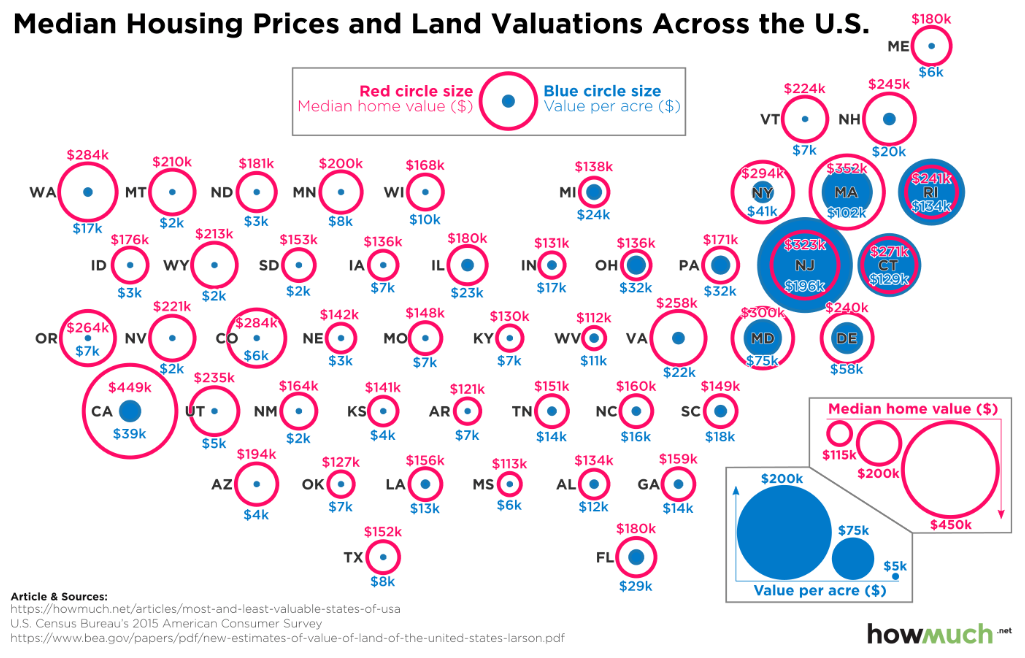
**Redesign Project - The Most (and Least) Valuable States in America**

**Introduction**

Location is everything when it comes to real estate. In an effort to display the value of each state in the United States, Raul (2017) created a map chart consisting of housing prices and land valuations. Unfortunately, the chart is ill-designed and information cannot be easily interpreted. We have redesigned the chart so that it follows good design principles, attracts readers and supports easy information extraction. The data used contains median home values, land values, total acres, and population for each state in the continental United States.

**The Bad Graph**

******

***Figure 1*. Median Housing Prices and Land Valuation Across the U.S. (Raul, 2017)**

Figure 1 is the original graph we will redesign. The intention of Raul’s graph is to illustrate home and land values across the U.S. Through these two metrics for each state, an individual can gain a basic understanding of the costs of being a homeowner in different states. Raul (2017) states, “Everyone knows location is the most important part of real estate. You can’t change where your house is (all things being equal).” In Raul’s visualization, the blue dots represent the value of an acre of land and the red circles indicate the median value of a home. The size of the circle indicates the value, with larger circles corresponding to higher values.

In addition to illustrating the cost of owning a home, Raul uses this visualization along with, presumably, previous knowledge of the area of each state to draw the conclusion that land and home values are higher in smaller states. “You can easily find cheap acres of land where they are plentiful and un-useful (sorry, Nevada), but owning property is a lot more expensive in smaller places crowded with lots of people.” (Raul, 2017) In addition, Raul concludes that home values are more expensive in densely populated areas and more established areas, primarily in the northeast.

**Disadvantages of the “Bad Graph”**

Perhaps the most significant flaw of the design is that it does not allow for comparisons of either the value per acre or median home value across states except in extreme circumstances. Except in the most extreme cases, like California and New Jersey, it is challenging for the reader to determine which state has a higher median home value or value per acre because the circles are so similar in size. While it is somewhat effective at demonstrating regional differences, it is challenging to compare any two states that are geographically separated. For example, it is apparent that both home value and value per acre tend to be lower in the midwest than on the coasts, but it is unclear how Virginia compares to Washington. Even at a regional level, interpretation of the size of circles to assess comparability is not reasonable without reviewing the actual dollar amounts shown.

In addition, the overall appearance is overcomplicated. For the entire continental United States, the graphic includes two circles representing figures on different scales, the name of each state, and the values for each data element. Given the layout, unfortunately, this is necessary. Ideally, the graphic element alone, the circles in this case, should be able to stand alone to some degree of effectiveness. This is not the case. Also, in the case of four states, blue text is used inside of a blue circle, presenting yet another challenge to the reader.

Further, the context provided is insufficient. The position of the circles does not necessarily correspond exactly to the location of each state, with Oregon and the Mid-Atlantic being the most glaringly out of place, which necessitates labeling the states. The values of the two attributes being measured are necessary due to the differences in scale and the lack of comparability. Although state abbreviations are provided, some are obscured behind the circles to which they correspond. Further, the labels are not presented consistently with respect to the corresponding circles. Most are to the left, but some are inside. Also, the scale on the graph is inconsistent between the two data points. At first glance, it would appear that the value per acre exceeds the median home value in New Jersey, for example, which is not the case. To compound this issue, it is clear that the the scale of the circle shown is not consistent. The difference in size between $115,000 and $200,000 is the same as the difference in size from $200,000-$400,000. The change in the size of the circles is four times, but the change in price is twice.

Finally, the presentation of data is not complete or accurate. The data necessary to support the conclusions that home values are positively correlated with population density and negatively correlated with size of the state is not included in the visualization. The title of the graphic cites “Housing Prices and Land Valuations”, where the data cite home values and land value per acre. In reviewing the data’s source (Larson 2015), it is evident that the data represents an estimate of the value of land, including federal lands, and value of homes as opposed to sale prices.

**Proposed Improvements**

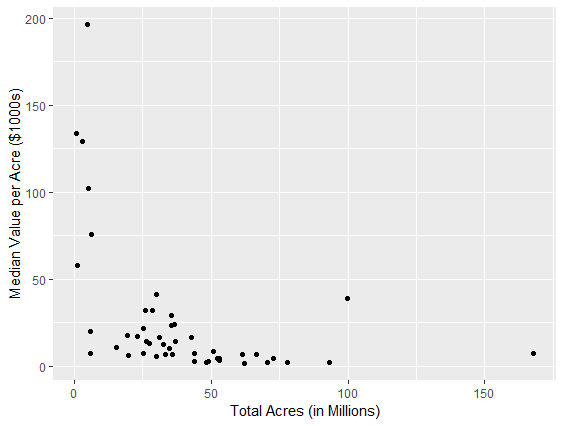
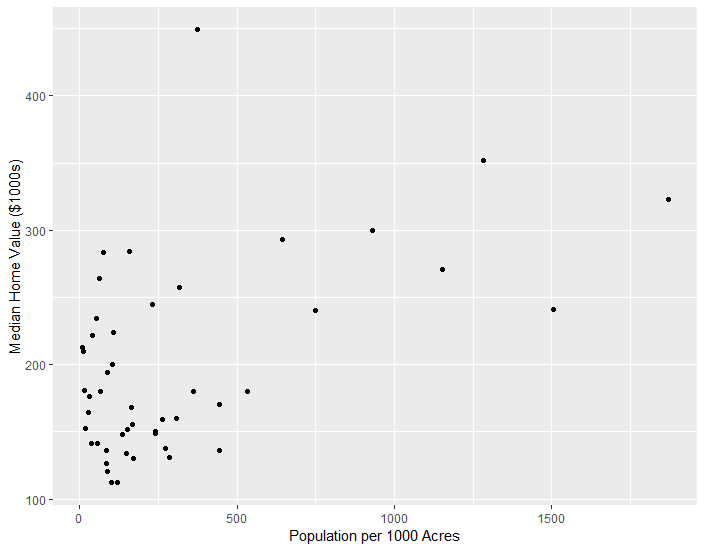
The intention of our improvements is to make it easier for the reader to draw conclusions based on the data being presented. To do so, it must be possible to draw comparisons between states on both attributes being measured. This must be possible between states and across regions, even when states are geographically separated. We can prioritize this comparison by having two separate plots side-by-side for each variable. Information should be shown using position along a scale as the encoding to improve readers’ accuracy of extraction. Each state can be displayed on a monetary scale to show median home value and land value. Within the plots, background grid lines and grid line labels will support making more accurate position comparisons. Removing the layering and showing data separately will also help the reader understand the scale of the two data elements being examined.

The information should be presented in a way that balances simplicity with the need for adequate context. The redesigned plot should sort by median home value because readers can relate to home value more than land value. Perceptual grouping should be utilized to help readers focus attention on segments of the statistics. This prevents information overload and supports comparisons on a smaller scale. For instance, by segmenting the states in groups of 5, readers can compare home values within each group. With sorting and grouping, it is much easier to discern groups such as the “top 5” or “bottom 5” states for home values.Utilizing simple symbols and the selection of high contrasting colors make the information easier to process visually. Another valuable addition is to create a linked micromap which will effectively display geographic information and will help readers see geographic trends for each variable.

**Redesigned Graph**

The following redesigns of the graph were developed using R (R Core Team, 2018) in the statistical computing environment, Rstudio (RStudio Team, 2015). We will be using the tidyverse, which includes ggplot2, and micromapST packages to create scatter plots and linked micromaps connecting data to specified geographical regions.

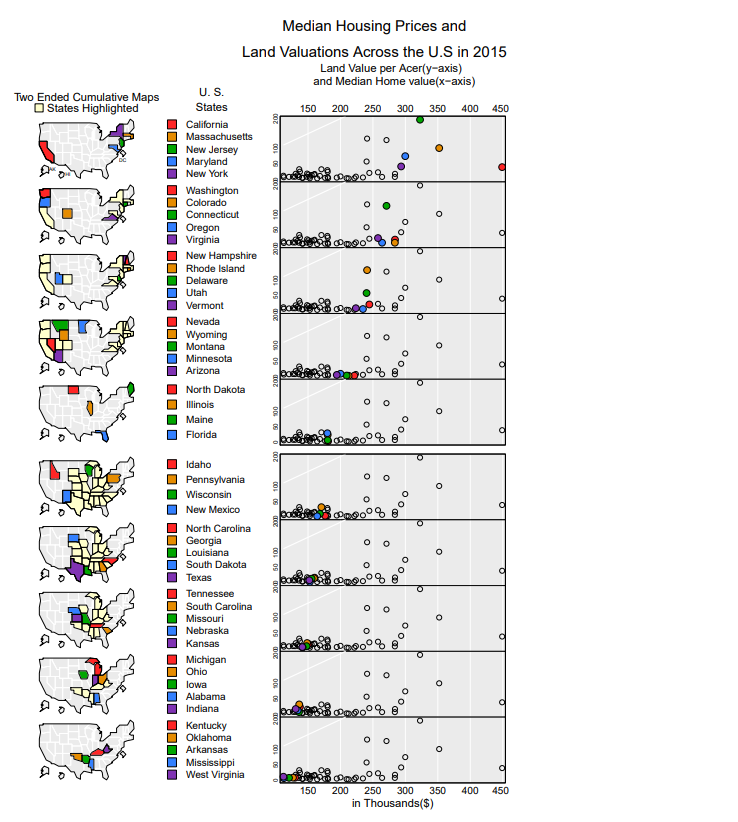
First, we wanted to address the author’s conclusions on the correlation of population density with home value and the correlation of land value with area. The scatterplots below were constructed using ggplot. Figure 2 supports the author’s assertion that land value is higher in the smallest states. Figure 3 reveals the states with the highest population density tend to have relatively high median home values, but there is no clear relationship between the two variables. These graphics, however, lack the ability to identify each individual state and discern their geographical relationship. Further, the scatter plots demonstrated the relationships in the author’s conclusion, they do not illustrate the cost of home ownership in each state.

**** 

**Figure 2: Median Value per Acre Figure 3: Median Home Value**

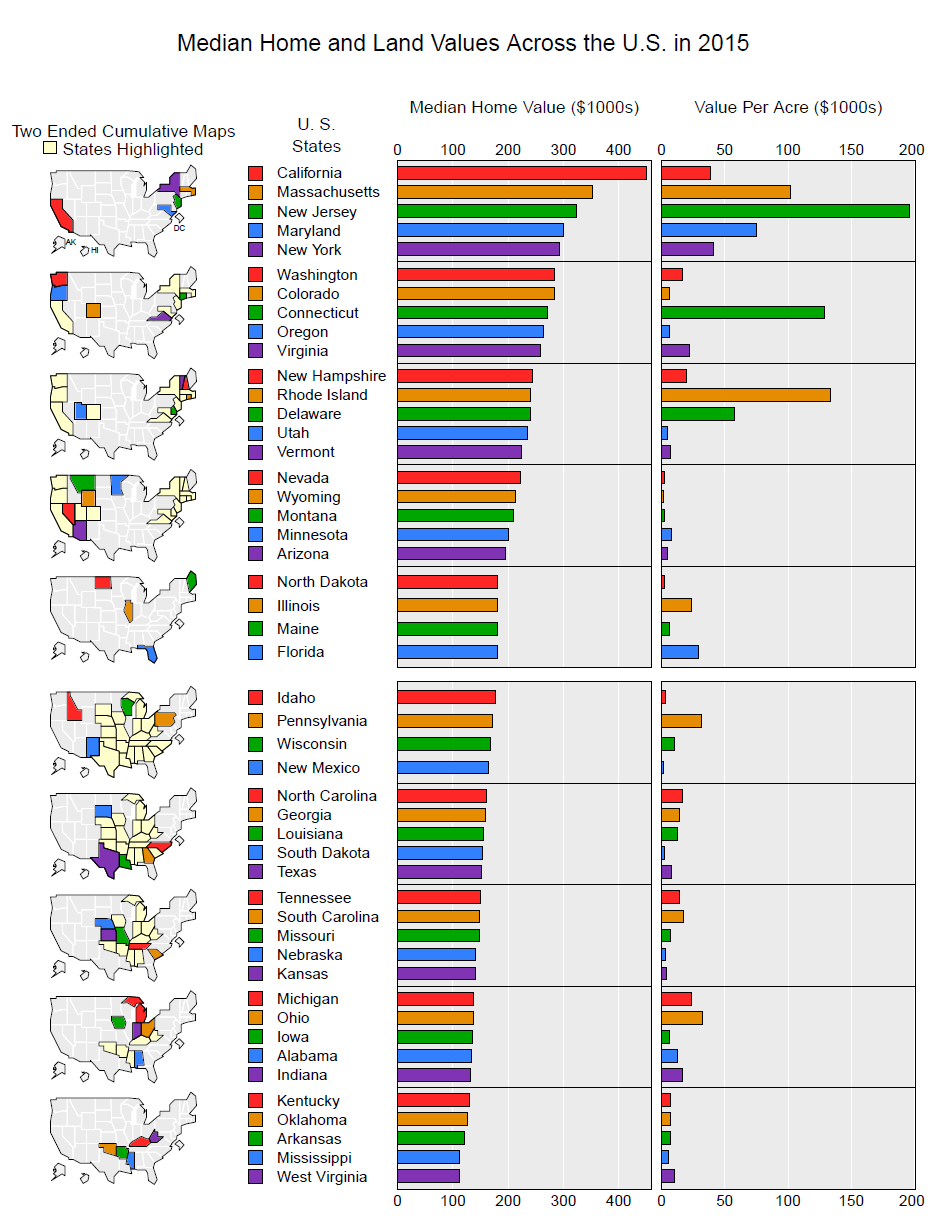
**versus Total Acres versus Population Density**

To also demonstrate the relationship between the two variables along with the visual representation of the states’ geographical locations, we next considered a scatter plot design using micromapST. While this design did provide more information about the relationship between home value and land value, we found that it was challenging to interpret as many data points were grouped closely and even overlapped. Also, the focus of the article and original graphic was on the cost of home ownership by state. Therefore, the median home value is the primary data element of concern and the land values were illustrative rather than a focus. We ultimately determined that representing the relationship between the two was less important to the objective than providing a simpler and easier to read visualization.

****

**Figure 4: MicrmapST Scatter Plot**

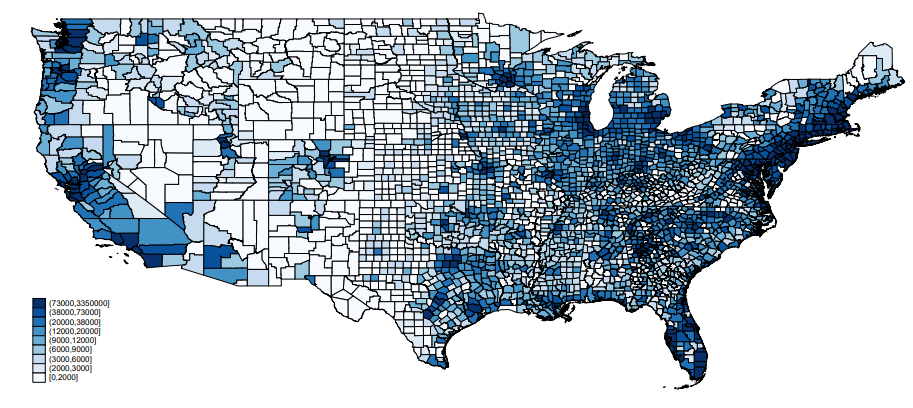
Ultimately, we settled on the side-by-side bar plots sorted by median home value. This representation showed the side-by-side home value (sorted) and value per acre. It allows for comparability between the two data elements and, like the previous plot, includes a visual representation of each state’s location. The bar plot, unlike the dot plot, made it simple to discern small differences between state values. By showing the location of states, the reader can easily identify trends based on region or geographic proximity.

****

**Figure 5**

**Next Steps**

Although the redesigned graph is substantially better than the initial chart, there are further steps that will improve the visualization. To complete the dataset, the median home and land values of Alaska, Hawaii, Washington D.C., and U.S. Territories could be obtained. For a more detailed analysis, it would also be useful to consider county-level data as home values can vary greatly in states that, like Virginia, have a single major metropolitan area confined to a small part of the overall state. Figure 6, below, was obtained from the source publication (Larson 2015) and demonstrates the utility of county-level data. It also could be useful to consider the mean home value for this circumstance. The absence data point or use of misleading data could lead the reader to make inaccurate conclusions. It would also be interesting to analyze the changes in home values over time. To keep the visualization relevant, information should be obtained for 2019. Our final, redesigned graph focuses on comparisons between different states, but the relationship between home value and other elements is unexplored.



**Figure 6: Land Price per Acre in the Lower 48 States, 2009 (Source: Larson 2015)**

**Conclusion**

Visuals following sound visualization techniques convey information in an effective and efficient manner. The original graph lacked the ability to quickly compare the median house prices and land values of two states. The design decisions for color, scaling and layout made it difficult to discern values. Our redesigned graph allows readers to easily compare median housing prices and land values and connect that information to the state locations. The most significant benefit of the micromap is the ability for readers to identify trends based on region and geographic proximity.

**Works Cited**

Daniel B. Carr and Linda Williams Pickle (2010), Visualizing Data Patterns with Micromaps

Linda Williams Pickle, James B. Pearson, Jr. and Daniel B. Carr (2014), micromapST: Exploring

and Communicating Geospatial Patterns in U. S. State Data

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for

Statistical Computing, Vienna, Austria. URL

[https://www.R-project.org/](https://www.r-project.org/)

Raul. (2017). The Most (and Least) Valuable States in America. (n.d.). Retrieved from

https://howmuch.net/articles/most-and-least-valuable-states-of-usa

RStudio Team (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL

<http://www.rstudio.com/>

Hadley Wickham (2017), Tidyverse Package, <http://tidyverse.tidyverse.org/>

Jim Pearson and Daniel Carr (2016), MicromapST Package,

<https://cran.r-project.org/src/contrib/micromapST_1.1.1.tar.gz>

William Larson (2015). New Estimates of Value of Land of the United States Retrieved from

<https://www.bea.gov/system/files/papers/WP2015-3.pdf>