***Overview***

The main goal of this project is to observe factors which would influence a house property values. Realizing there are numerous factors which contribute to the property value, we decided to narrow down on the following factors: natural disasters, property tax rates, and demand score to see if there are any trends to be observed.

***Data***

We had many different sources of data for this project. Finding an appropriate data source was more difficult as there was a lack of consistency, accessibility and dependability.

For the property value of houses, while there are many sources to choose from, we decided to keep the scope to housing listed on Zillow. The Zillow data was broken down into different sheets by the different size properties so as part of clean up, we merged together all the data. An observation of the data is that the houses are a seasonally adjusted measure of the typical home value and market changes across a given region and housing type. It reflects the typical value for homes in the 35th to 65th percentile range. This means any observable outliers or variations may have been removed.

The source of the Natural Disaster data was from Kaggle. This data source had many attributes which were not needed, so we had a clean up process for this. An observation of this was that we were looking for a total count of natural disasters recorded in each state over the last decade. While that was not a data source we found, we resorted to any natural disaster which has occured in the last decade in which the president has declared a state of emergency. Emergency declarations supplement efforts in providing emergency services, such as the protection of lives, property, public health, and safety, or to lessen or avert the threat of a catastrophe in any part of the United States. The total amount of assistance provided for in a single emergency may not exceed $5 million. This means there would have been many severe storms in which nothing would have been declared since money allocations wouldn’t have to be made for this.

The calculation of tax varies between states in the US. Each state has their own ways of calculating tax. We considered analysing the tax rate against the average market value of houses in NJ. We were able to collect the historic tax information for all counties over the decade[2010-2020]. The data comes in the downloadable excel format, from the *NJ state department of community affairs*. While analysing the data is the file, we found that the excel format is consistent through the decade. Hence the files are classified to 2 closest file formats and then required columns are extracted from the sheets in 2 different ways.

The demand score in an area will show the hotness of the real estate market in that location. Hence we took the data from *realtors’ research data* website. We are able to collect data for the period 2017-2020. Since we compared it with NJ’s property tax data, we filtered the realtors’ data for NJ. We are able to collect the hotness score, demand score, supply score and other details needed to locate the movement of the housing market. All 3 datasets[Housing value, demand index and property tax rate] are in sync to hold data for period[2017-2020] for comparison.

***Analysis***

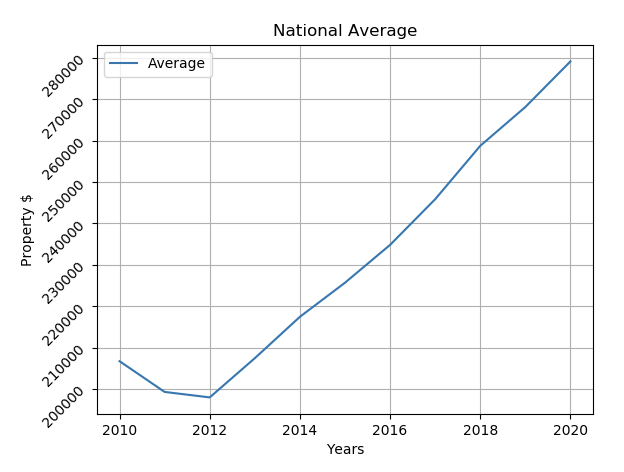
In attempts to observe our goals, our initial hypothesis is: If a state has more natural disasters in a particular year, then the average property value, property tax and demand for property in that area for that year will be lower compared to other years.

In order to get a response to this hypothesis, these were the six main questions we asked.

1. What are the average property values in the US?
2. What are the natural disaster patterns in the US?
3. How are natural disaster patterns related to US property value ?
4. What are the average property values in the US by house type?
5. What is the relationship between demand of a location and property tax?
6. What is the correlation between property tax and property value ?

***What are the average property values in the US?***

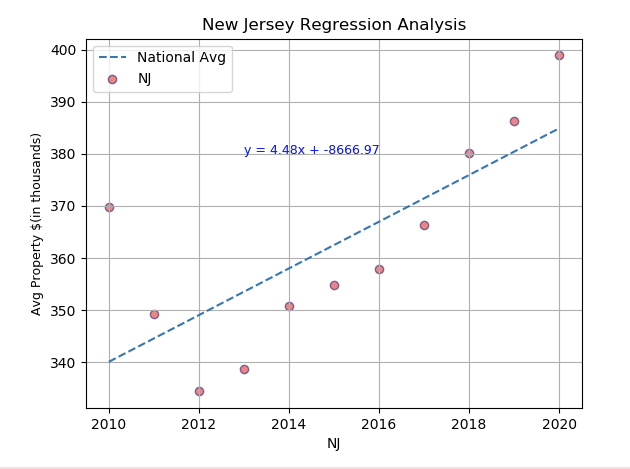
Below is the chart of the average property values of each state from 2010-2020:



Using the Zillow data, we averaged all bedroom unit residential listings over the course of 10 years (2010-2020) for all 50 states in the US.

Our initial deep dive was in the state of New Jersey. Since we had property value on New Jersey, we analyzed New Jersey by county.

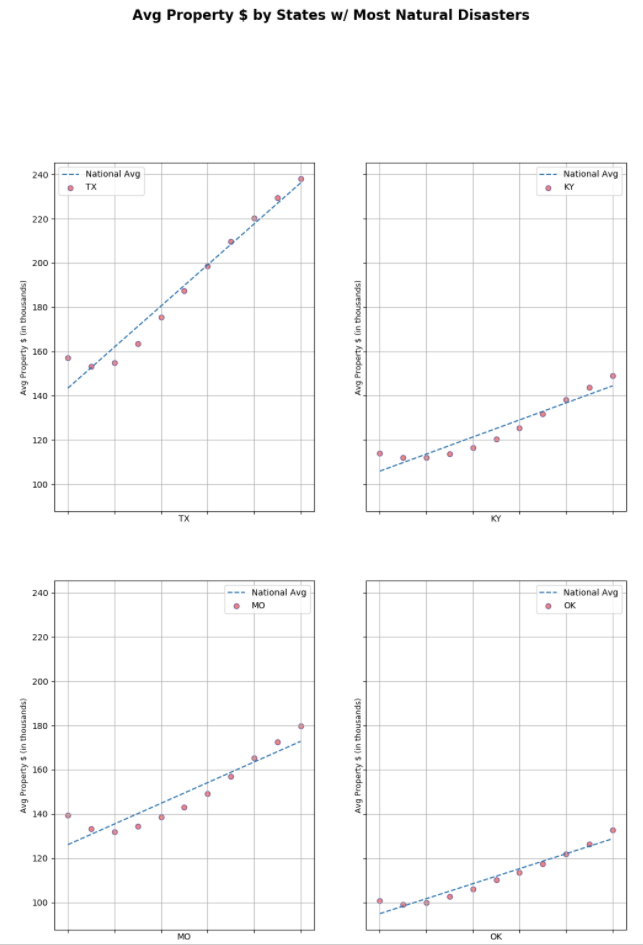
In this analysis, we took a macro point of view. We analyzed how the property value of New Jersey resendiatals units did compared to the nation.

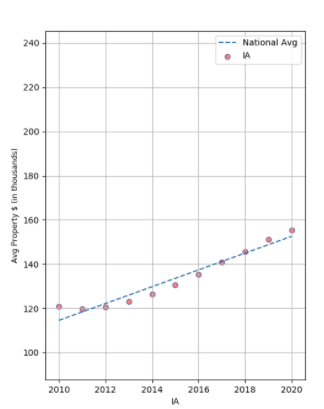


Based on the New Jersey chart, we can see in the early years of the decade, NJ property values were relatively worse compared to the rest of the country.

The regression line (dashed lines) is the national average. Also on the plot is the slope intercept for NJ Property Values.

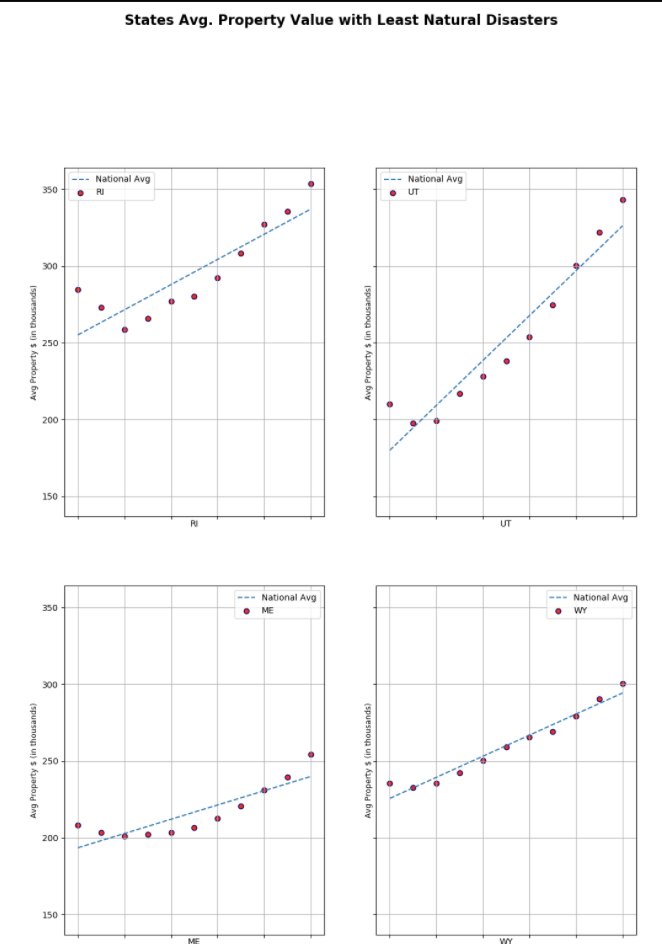
Additionally, we ran the same tests for the top 5 and bottom 5 states ranked by count of natural disasters (table list shown later below).

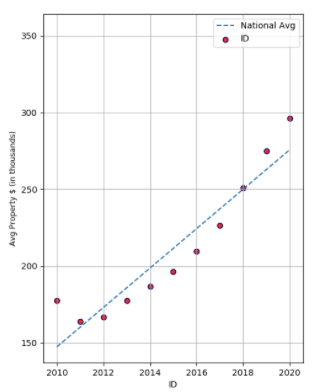




Based on the top 5 states with the highest disaster count, most of their values based on the chart stayed on par with the regression line

Bottom 5:





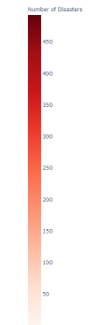
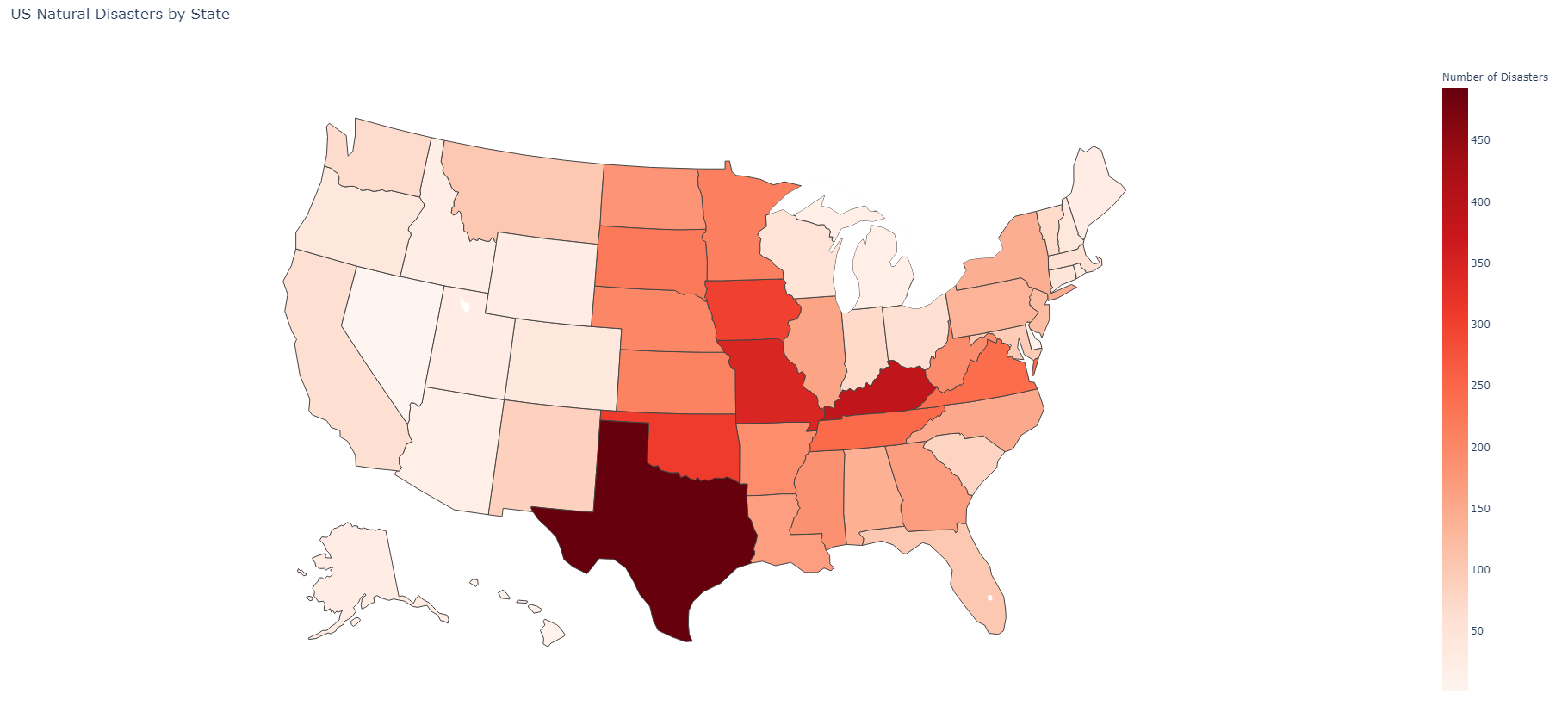
Based on the bottom 5 states by natural disaster count we can see that many of them experienced a dip in their average property values compared to the regression line during the middle part of the decade.

It is interesting to see that the states with the most natural disasters, the average property value for the respective year was relatively close to the regression line. However the states with the least amount of natural disasters, the average property value decreased.

Perhaps natural disasters do not have an implied correlation on property values as we originally thought

***What are the natural disaster patterns in the US?***

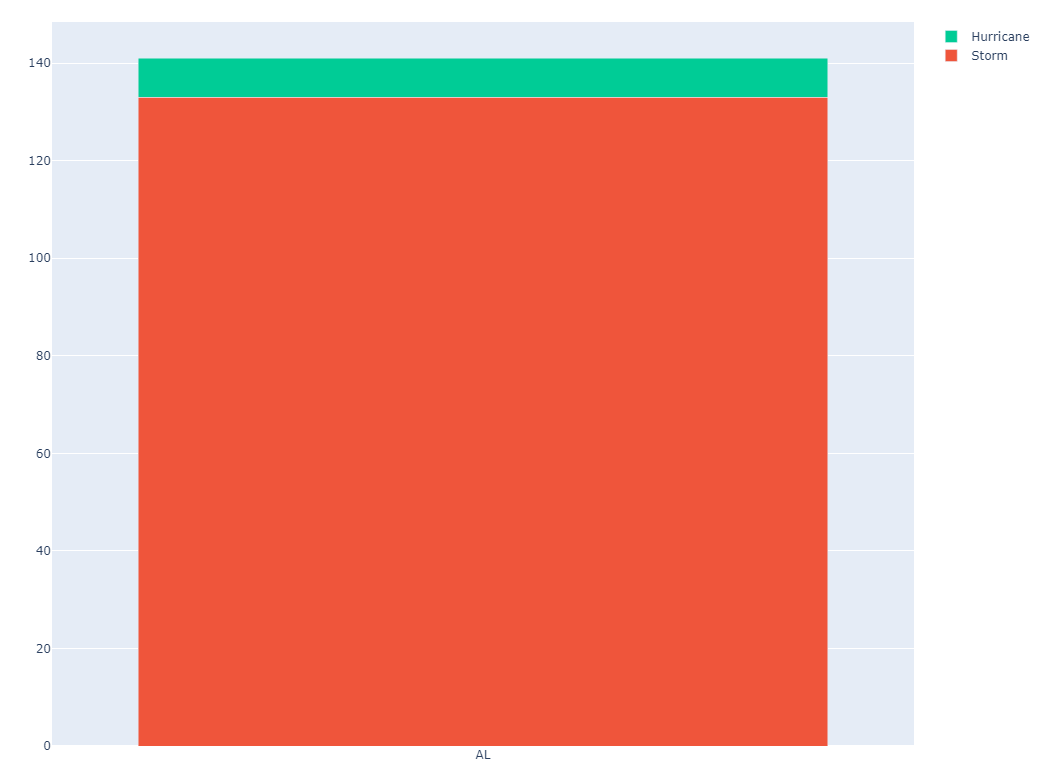
Observations made strictly based on the data is that the mid-region of the United States seem to have a lot more natural disasters then the west and east over the period of 2010-2017. When breaking down the types of disasters the most frequent in this area was storms.

US Natural Disasters By State

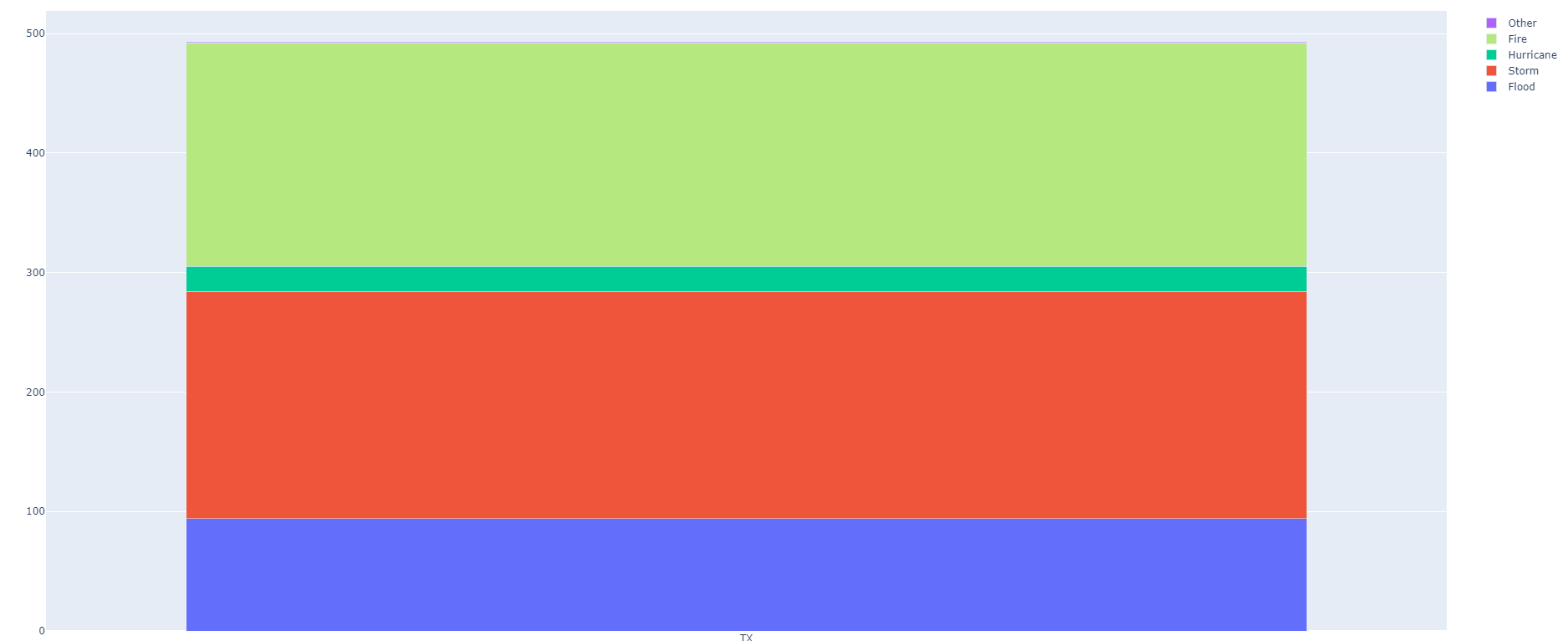
*Number of Disasters*

|  |  |  |  |
| --- | --- | --- | --- |
| TOP 5 States | | BOTTOM 5 States | |
| TX | 493 | RI | 25 |
| KY | 387 | UT | 24 |
| MO | 345 | ME | 24 |
| OK | 309 | WY | 23 |
| IA | 302 | ID | 21 |

AL



TX



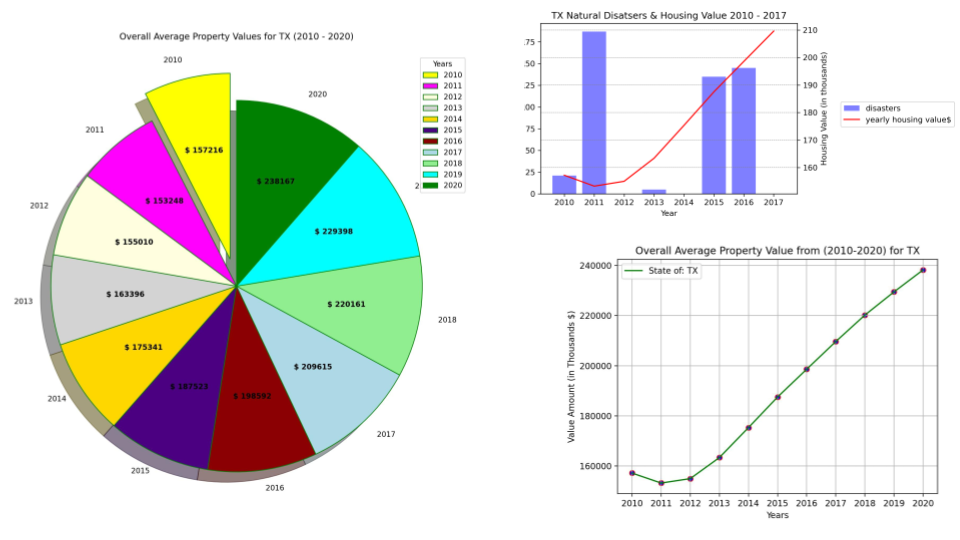
***How are natural disaster patterns related to US property value ?***

* Initially hypothesized that there would be a correlation between # of disasters and the average home value - believing that if the number of natural disasters is high then the property value will decrease
* I found that regardless of disaster pattern in most cases - states has a upward sloping trendline for their housing values from 2010 - 2017, with a dip in 2012
* After noticing this created a visualization of average housing value for each state for each year - then also created a new DF that selected the lowest value from all years and shows the state abbreviation if its lowest values was 2012 - this was the case for 28 states (over 50%)
* From this we can conclude there is no clear correlation directly between natural disasters and the property value of a house. With more time allotted we could look into the several other socioeconomic factors that may also have an effect on the value of property

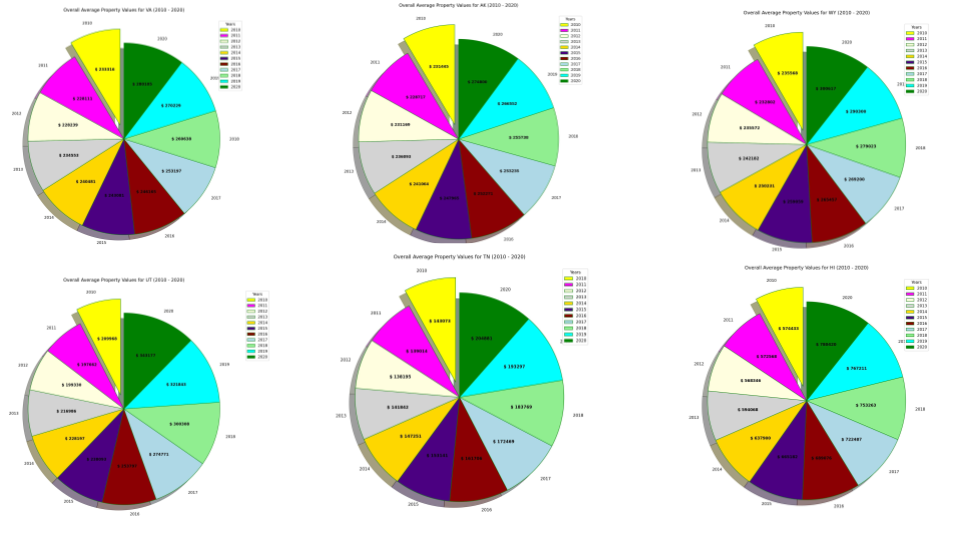
#### **Does the AVP housing units tell us anything about the effect of disasters the on market?**

Initially hypothesized that the average property values of homes/condos (regardless of size/type) would be negatively affected by natural disasters. The question then becomes, assuming that is true, do certain types of housing (ie 1BR, 2BR, 3BR, 4BR, 5BedroomRoom, Condo) suffer average property value (AVP) more, or less, in these instances? Would there be a rise in Condos and a decrease in 1BRs after a major flood for instance. I didn’t expect there to be a difference depending on housing type due to the nature of the housing market in the US and what also affects it. Housing is affected by *many* factors, from macro influences from state of the economy at large (think of the 2010 Great Recession that was brought on from the housing bubble), to the demand of housing, location, crime rates, state/local laws, property taxes policy, nearby school system, down to factors on a particular street or additions made on a home etc, the list is *vast*, so I hadn’t assumed that a natural disaster would *only* affect one type of housing option (verus another type), or if so, all housing would be affected equally. Here’s what we did::  
  
  
First, created 2 dataframes using a previously cleaned dataset by the team for other observations. This was to ensure some consistency between the project. Then I create a function that would allow inputs by the user to select the ‘State’ and ‘Room Type’ they wished to view data on, set their selection to variable that will be used as an index of sort to call future data. The one dataframe groupby “State”/user input variable to give the overall average(mean) total of that state’s housing values in a certain year between 2010-2020. The other dataframe does the same but groupby ‘State’ and ‘Room Type’. Then the program does an analysis to show 4 charts to the Output folder: a pie chart showing AVP each year in the decade, a pie chart showing that specific Room Type and its AVP in the state for the the decade, a line/scatter graph just reviewing the states AVP (to match up with other team member’s graph format), and a bar chart that shows all 6 Room types AVP for a the year’s 2010 and 2020 to show any useful change/trend.

*Observations*Reviewing pie charts and bar graph  
No notice of substantial difference of housing property values (neither overall or by housing type)...  
Reviewing data against previous disaster observations (ie Top/Bottom 10) to try something new.  
Noticing however potential correlations between the States that are the most affected by natural disaster also having low AVP overall. (Ie Over 300k VS Under 300K )



**Analysis**At a housing type level, the data results is at best inconclusive but most likely not related. All ships rise and fall in waves…. But on a national/bird’s eye view level, there may be more than meets the eye.



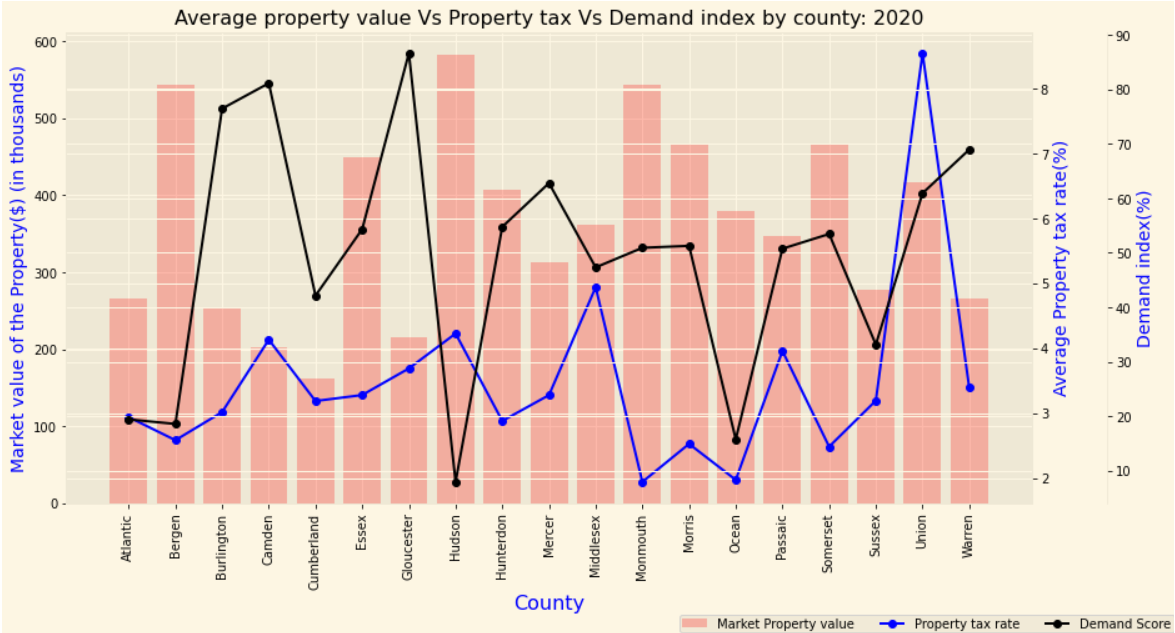
**Shortcomings**Time was a factor in really delving into the data. (ie with time constraints gathering data, coding, observations. The dataset can only be grouped and cleaned so much, but consistency between dataset (and limiting variables) is key. Lack of information on an issues like limited dataset, sporadic disasters, records kept across states data/systems etc make an already challenging question moreso, especially when many variables contribute to AVP

### **Property tax, Market value & Demand score**

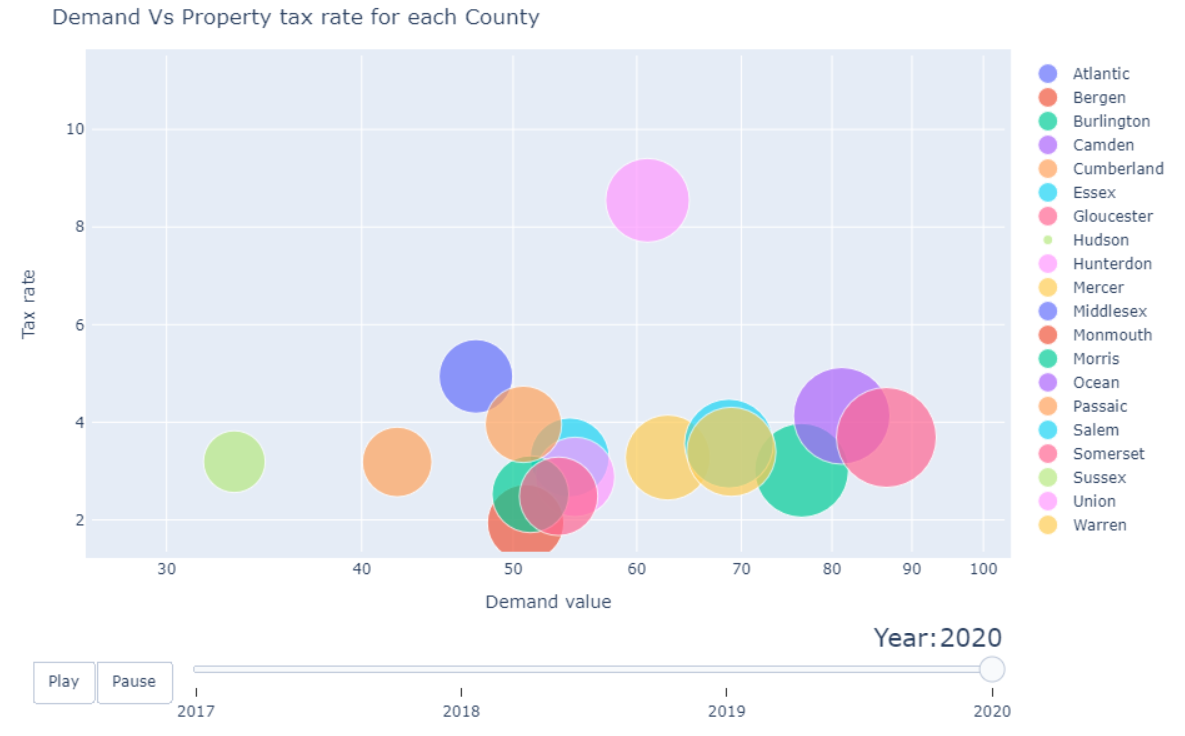
***What is the relationship between demand for a location, property tax and average market value of the house ?***

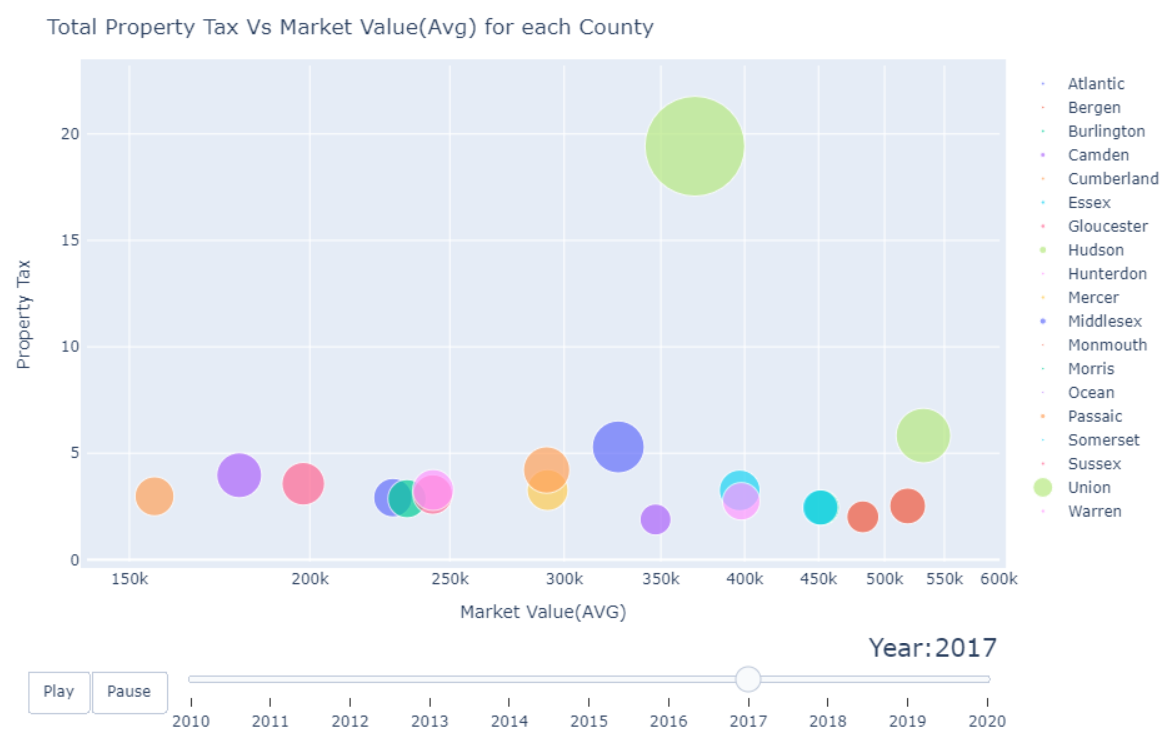
The Average market value of the house and property tax are not mutually dependent on each other. The property tax rate is calculated based on the assessed value of the property and the township’s requirement. The market value of the house need not be the same as the assessed value of the house. It can be either higher or lower than the assessed value. Since the property tax is calculated based on the assessed value there is no linear relationship between tax and market value of the house. However, the property tax also influences the trend of searching for a house [demand score].

* Gloucester county has a hotness rank: 22 (as compared to other counties in the US) & 1st place in NJ, with a demand score of 93% for year 2020.
* In Hudson county, the property tax rate is average, whereas the demand for the houses is as low as 1.5% since the market value of the home is too high.
* One of the outliers would be the Union county. The tax rate in one of the townships skews the data of the Union county to make a steep raise in tax rate as compared to the other counties.
* From the chart we could see the inverse relationship between demand score and market value of the house.
* The demand score is also influenced by the property tax. We would say it may not be one of the major factor, but property tax does have some impact on the house demand.



The bubble charts show the movement in property tax rate with respect to the market value year over year. Since we had demand scores for the period 2017-2020, other datasets [housing value & property tax] are also narrowed down to that period.





***Conclusion***

We reject our hypothesis based on the data and analysis conclusion. We noticed that property values show no direct correlation with natural disasters and property tax rates. However, we did notice that the demand score does have an inverse relationship with property values, as observed in New Jersey. Higher demand property values lean towards having a lower demand score.