Daniel Clark

Dr. Miller

CTEC 128

Storm Events – Data Science Report

A common, environmental topic that is often debated in politics, science, and many professional and non-professional settings is global warming. Effects of global warming have become more apparent as more research is being conducted. These effects include melting glaciers, plant and animal ranges shifting, and increased severity in storm events. Although all of these events are important, the main one that affects the insurance industry is the increased severity in storm events. The CEO of Bowie National LLC, an insurance company that provides a range of coverage across the United States, is concerned about the cost that these more intense storms will have on the company’s finances. The CEO wants to know more regarding trends in intensity, property damage, injuries, and other important factors. By analyzing data from the National Oceanic and Atmospheric Administration (NOAA), this report will provide the requested recommendations concerning how Bowie National LLC should adjust their consumer costs, coverage options, and other changes, in order to maximize financial efficacy.

**Preparation**

To plan the factors that the data science team would analyze, we developed a concept map that displayed the underlying factors that are effected by and affect the main issue, intensified storms. To start, global warming is a factor that affects the storms. Global warming is mainly caused by gas pollution which traps heat within the atmosphere, creating the greenhouse effect. The rising temperatures caused by the greenhouse effect, in turn, cause warming of the planet’s surface, glaciers to melt, and the sea level to rise. Global warming is hypothesized to be the main factor in intensifying storms. While this fact may be a popular belief, the data science team will use NOAA’s data to prove or deny this hypothesis. Secondly, the main population affected by storm events are homeowners and building owners. Out of pocket costs will vary depending on the insurance coverage each individual of this population has. The factors that affect the cost are the homeowner’s policy, the infrastructure and overall health of the building, and the location. The data science team will analyze the location portion of this problem to better recommend which regions within the U.S. need which policies. Lastly, a factor that affects the overall recovery from a storm event is the government. If the storm causes minimal damage, the local government will take the recovery responsibility. However, as storm events rise in scale, state, national, or even other governments may contribute towards the overall recovery. Organizations, such as the Housing and Urban Development agency (HUD), and the Federal Emergency Management Agency (FEMA), are two of the known recovery contributors of the government. Based on these concepts, the team decided to use the following research questions:

1. Is there an upward trend in the intensity of storm events?
2. Given the five states (stated later), which has the highest property damage in the time period?
3. Which of the five states has the highest amount of deaths and injuries?
4. Which are the highest costing storm events in each state?
5. Should Bowie National LLC change their policy based on the analysis?

Overall, the data team used these factors and questions to determine our approach to analyzing the data and providing recommendations.

**Getting the Data**

As previously stated, the team retrieved our data from NOAA’s database on storm events. This database contained data on every storm event from every state within the United States from 1950 to the present year. The science team, however, only used data from 2005-2018. The team decided on the year 2005 due to a significant storm event, Hurricane Katrina, occurring in that year. A decision was made to end with 2018 because it was the last year with a complete data set for the full year. The team also decided to choose a populous and large state from the five regions within the United States. This limits the processing requirement of the data and provides an accurate assessment. The five states chosen were: California representing the West, Florida representing the Southeast, Michigan representing the Midwest, New York representing the Northeast, and Texas representing the Southwest. The unfiltered data set, spanning from 2005 to 2018 and including all states, had about 50 variables and millions of observations. After filtering the data to focus on the selected states and removing unnecessary variables, the data set had 26 variables and about 121,000 observations.

In order to best analyze the data, certain techniques were employed to ease data manipulation. Overall, the team used Microsoft Excel for analysis. Able bits, an Excel add on, was used to combine the different years into one sheet. General filtering was used to pool the wanted data, while copying and pasting was used to remove gaps from between the data. Average and sum functions were used to produce the results in our recommendations and visualizations. The property damage variable also needed to be reformatted from a character representation style ($1K, $1M, $1B) to a numerical representation style ($1,000, $1,000,000, $1,000,000,000). The other variables used for analysis and recommendations were state, event type, direct and indirect injuries, direct and indirect deaths, and magnitude. Certain observations under the property damage and magnitude variables were missing, but were calculated as a void entry in analysis. Overall, our team filtered, reformatted, and sorted the data in order to provide the most effective analysis and recommendations.

**Results/Insights**

After analyzing the data, the team began to answer the research questions and form recommendations. The team created various visualizations in order to easily depict the information.

Figure 1. Shows the average magnitude value calculated for all storm events with magnitude, each year from 2005-2018, for the five focus states.

To start, the team wanted to be certain that there was reason to continue with analysis. Although the overall increase in storm intensity is widely accepted, physical evidence that proved that claim was necessary. By creating Figure 1, the team was able to show that there is definitely a rise in intensity of storm events over the years. It also shows that, while intensity was rising at a steady rate before 2016, intensity skyrocketed during 2017 and 2018.

Figure 2. Shows the amount of property damage calculated for the focus states.

\* Hurricane Michael has been removed from Florida’s value as an outlier.

For the second research question, the team calculated the property damage costs for 2018, the last year from NOAA’s database with data for the entire year. From this chart, it can be observed that Texas is higher than the other four states at $15.8 million. New York and Michigan follow closely behind at $12.3 million and $11.1 million, respectively. Florida and California are the lowest at $8.3 million and $6.6 million, respectively. However, California has recently fallen victim to an increasing number of wildfires. Therefore, the data science team concluded that either data is missing in regards to the property costs related to these fires, or the data has not yet been updated to include retroactively added costs.

Figure 3. Shows the amount of direct and indirect deaths calculated in the five focus states from 2005-2018.

The team’s third research question related to the mortality and injuries of these intensified storms for the entire observed range. By analyzing Figure 3, it is noted that, once again, Texas outclasses the other states with 633 direct deaths and 232 indirect. Although a different variable was analyzed, the data team is still wary of the fact that data from California may be inaccurate. However, due to the nature of the inaccuracy for all variables, the data team acknowledges that the values from California would only increase, rather than decrease.

Figure 4. Shows the amount of direct and indirect injuries calculated in the five focus states from 2005-2018.

Along with the mortality of storm events are the injuries. Similar to the deaths, Texas completely outclasses the other states in terms of injuries. However, unlike the deaths graph, California has more injuries than Florida, and Michigan has more than New York. Once again, the data from California may be missing observations, therefore, the value should be analyzed loosely.

Figure 5. Shows the highest costing storm events calculated by state in 2018.

Lastly, the team analyzed the amount of property damage by storm event within each of the five focus states in 2018. By doing so, we were able to conclude that wildfires, although the data may be somewhat inaccurate, were the costliest of storm events in California. Surprisingly, tornadoes were the costliest storm event in Florida, discounting Hurricane Michael. Both floods and high winds were very costly within Michigan. Flash floods and high winds again were the costliest in New York. Another revelation found within the data was hail being the costliest event in Texas with lightning coming in second.

Overall, by analyzing the data and creating visualizations, it can be seen that Texas is a very costly state in terms of insurance. The data recommends specialized infrastructure and policies implemented for Texas specifically. Another option would be to forego insurance coverage within Texas to prevent extra risks being taking by the company. In terms of California, the team recommends special measures to be taken that prevent or decrease the costs of wildfires. The costs within California may be increased, however, the team concluded that there would not be an extreme risk in generating insurance policies for the state of California. For Florida, the team recommends the company do more research to determine the risk by implementing infrastructure and insurance policies. However, if Bowie National LLC were to pursue insurance in Florida, the state appears to be a high risk factor. Therefore, the team offers the same insurance recommendations as Texas for Florida. For New York and Michigan, the team recommends that the company implement infrastructure for high winds and flooding. The team also fully recommends that the company generate insurance policies for these states. The research done by the team can be generalized to each region, but we also recommend that more research be done in order to fully understand the infrastructure and policy that would be necessary within each state.

**Discussion**

In order to recommend changes for financial efficacy within Bowie National LLC, the research team analyzed data from NOAA’s storm event database. By using a concept map and research questions, the team was able to give general recommendations for each region within the United States. From the data, the team was able to confirm a rise in the intensity of storm events. The team was able to identify Texas, representing the Southwest region, and Florida, representing the Southeast region as a high risk factor for the company. The group then recommended the type of infrastructure and policy needed for the most common storm events within that region. The team additionally identified New York and Michigan as low risk factors for the company and what infrastructure and policy would be needed to maximize effectiveness. The team was able to identify California as a medium risk factor along with the infrastructure and policy needed for the West region. However, the team acknowledges that the data regarding California was not as accurate as possible, and that any values relating to California could increase, but not decrease. Lastly, the team recommends that future research be done for each state in order to determine the best infrastructure and policy per state, rather than per region. The team also recommends that further analysis be done on California and Florida to truly determine the risk factor for insurance policy within those states and regions.

References

* NCEI. (n.d.). Storm Events Database. Retrieved from <https://www.ncdc.noaa.gov/stormevents/>