

Applied Data Science Capstone Project

Introduction

The objective of this project will be to analyze data over major hurricanes in the state of Texas. This data set was obtained from the National Weather Service (Roth, 2010). This knowledge would be of interest to local and state governments, people looking to move to the Texas Gulf Coast and any industry that is negatively affected by hurricanes (Energy, utilities, infrastructure, construction, etc..). Analyzing hurricane data is important to further learn the trends and increasing intensity of these weather events.

Data

The data that will be used consists of 31 different hurricanes that have made landfall in Texas. The dates of the data spans from 1940-2020 and includes the month of the event, latitude, longitude, wind speeds, category and casualties. This data will be used to analyze hurricane trends such as where they are most likely to make landfall and the intensities of the hurricanes.

	Year	Month	Landfall	X	Y	Winds	Category	Casualties
0	1940	August	Sabine Pass	-93.8948	29.7333	100	2	0
1	1941	September	Sabine Pass	-93.8948	29.7333	120	3	4
2	1942	September	Texas City	-94.9027	29.3838	80	1	0
3	1942	August	Galveston	-94.7977	29.3013	115	3	8
4	1943	July	Matagorda	-95.9683	28.6911	100	2	19

Methodology

The methodology I will be using is to initially explore and analyze the data to find any trends or correlation. I used the correlation and descriptive statistics to firstly analyze the data. Then I used some of the plotting functions to get a better understanding of what was related in the dataset. From there I used the folium map package to map the locations of where the hurricanes made landfall.

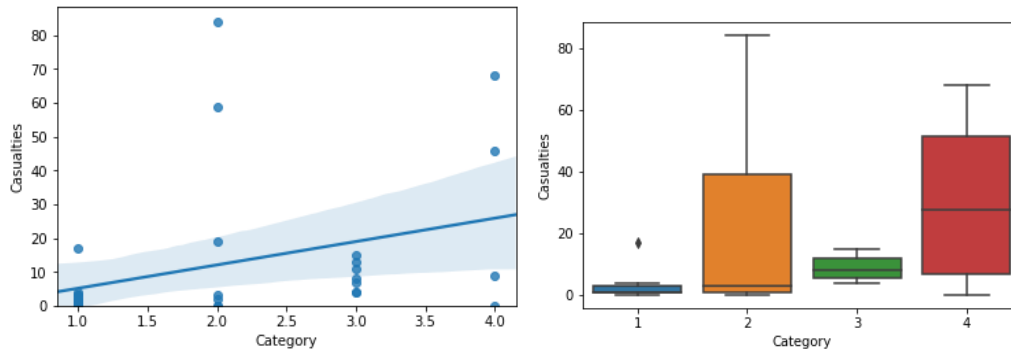


Figure 1. Regression of the category of the hurricane with the casualties and box plot of same variables.

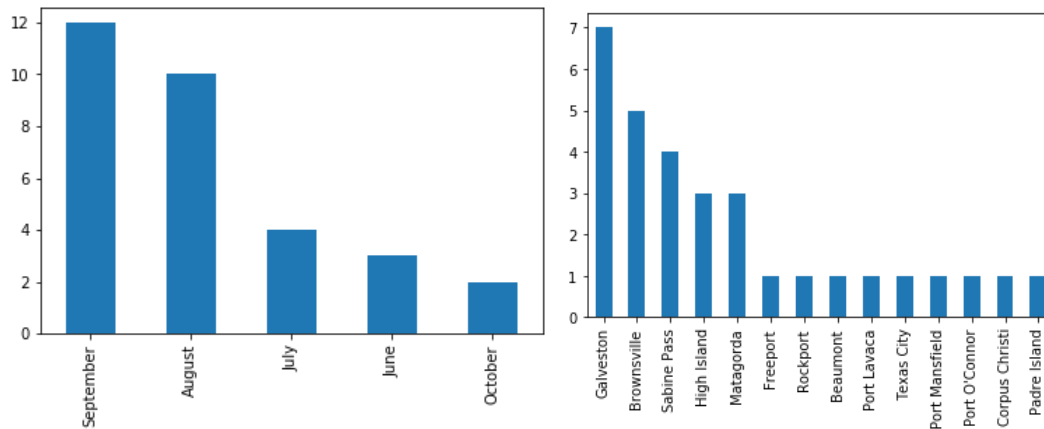


Figure 2. Bar plot of number of hurricanes per month (left) and the total number of hurricanes by the city where they made landfall (right).



Figure 3. Map of Texas Gulf Coast with locations of where the hurricanes made landfall.

Results

The results of the dataset show that there is a positive correlation between the category of the hurricane and the amount of deaths. It also shows that hurricanes are most likely to hit in the month of September and generally only between the months of June-October. Also, after mapping the locations and analyzing the data it shows that the city of Galveston has been the most frequent place that hurricanes have made landfall. The distribution of hurricanes along the Texas Gulf Coast is relatively even, they seem to make landfall all over the coast but with the majority around the Houston area.

Discussion

The observations that were noted are that hurricanes make landfall around the Houston area more often than any other region. This could be increasingly negative because of the large population of the Houston area and it has many industries that could be impacted by flooding from hurricanes. There was also a positive correlation between the strength of the hurricane and the amount of deaths. The recommendations that could be made is further forecasting and analysis of hurricanes to prevent deaths. Another recommendation that could be made is to upgrade the current infrastructure of the major cities to prevent costly damage from flooding. The hurricanes over the last 80 years have been very consistent and growing in intensity.

Conclusion

Hurricanes are very destructive and costly natural disaster that require further and frequent study. Analyzing this data is important to many industries that could be impacted by flooding and wind damage as well as state and local governments that could benefit from the results of natural disaster studies. Hurricanes are responsible for billions of dollars of damage per year and many deaths that could most likely be prevented. Further investigation of the data is needed to make better recommendation and a more comprehensive dataset would be beneficial. In the future other variables, such as water temperature, atmospheric temperature, hurricane path, evacuation routes and total damage costs, should be investigated to make more informative recommendations.

References

Roth, D. (2010, January 6). *Texas Hurricane History*. Weather.Gov.

<https://www.weather.gov/media/lch/events/txhurricanehistory.pdf>