

Applied Data Science Capstone Project

10/12/2020

Introduction

- Analyze dataset of major hurricanes in Texas using Jupyter Notebooks
- Dataset created from National Weather Service article
- The information could be used by
 - Oil/Gas Industry
 - State/Local Governments
 - Construction (Home Building)
 - Insurance Companies
 - Utilities/Infrastructure

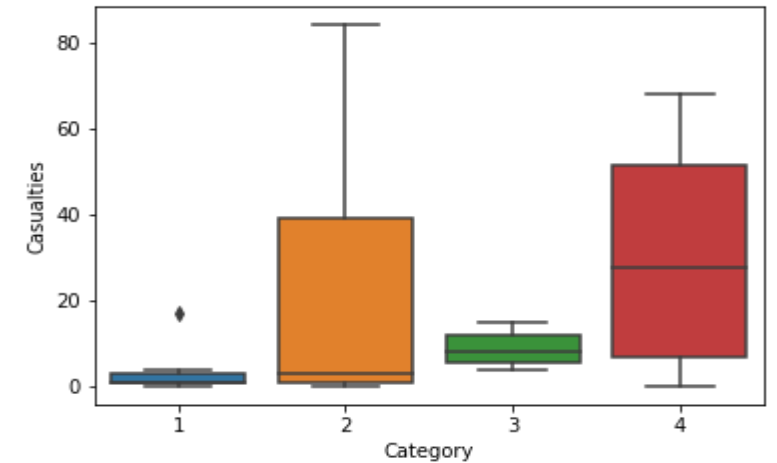
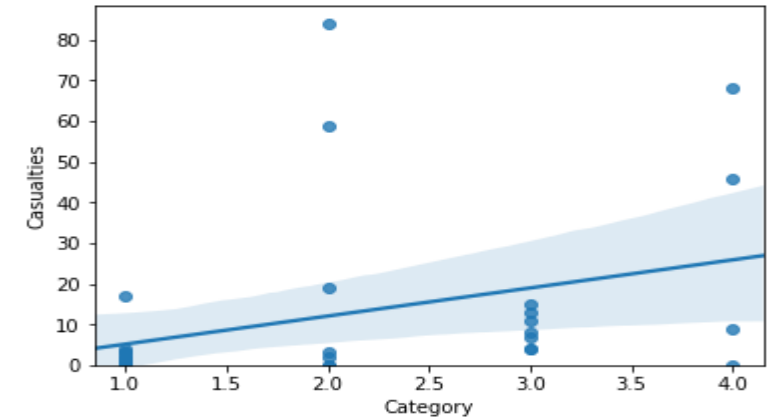
Data

	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

- Dataset created from National Weather Service article
- Includes 31 hurricanes that made landfall in Texas from 1940-2020
- Includes Year, Month, Landfall location, coordinates, wind speed, category and casualties

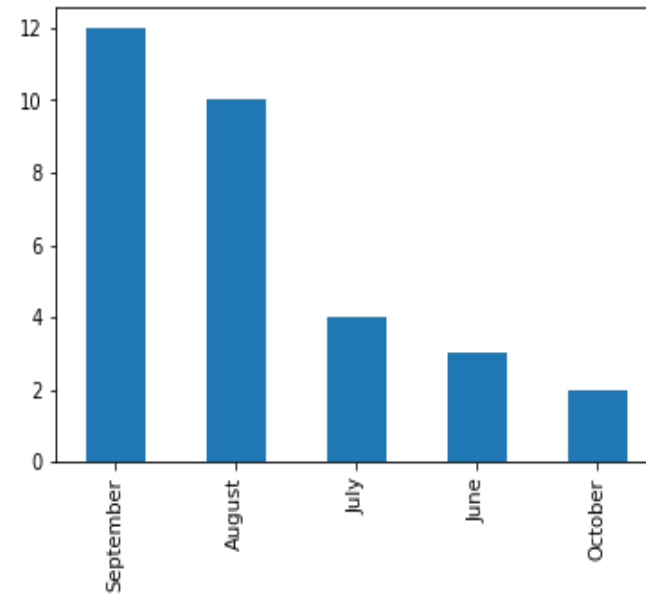
Methodology

- Explore and analyze data to find any trends or correlation
- Use regression, box plot and bar graph to analyze data
- Map location of where the hurricanes made landfall

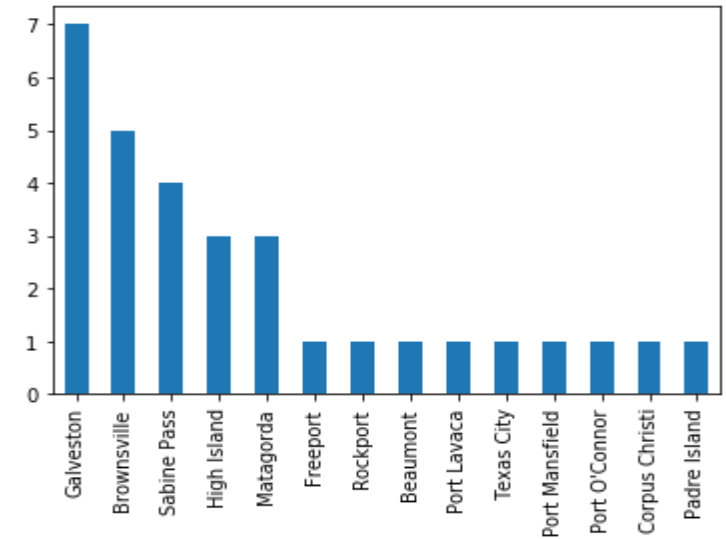


Methodology

Bar plot of most active months



Bar chart of most frequent landfall location



Map of landfall locations



Results & Discussion

Results

- Results show positive correlation between category and casualties
- Most hurricanes made landfall around Houston, with Galveston having 7 major hurricanes
- September is the most active month for hurricanes

Discussion

- Recommendations
 - Increased and upgraded flood infrastructure in and around Houston
 - Increased planning from state/local government for evacuation and preparedness
 - Find more data and variables to better understand the negative outcomes of hurricanes