Analysis of Hurricane Severity Over Time

Report

Introduction

Context

<u>What is a hurricane?</u>

A hurricane forms over tropical or subtropical waters. It is a rotating low pressure weather system with no fronts. A storm must have winds of at least 74 miles per hour to qualify as a hurricane (National Oceanic and Atmospheric Admission). Hurricane season, the time of year when hurricanes are most likely to occur, is from June through November. However, November hurricanes are less common. Hurricanes begin in the Atlantic basin, which includes the Gulf of Mexico, Atlantic Ocean, and Caribbean Sea. This means that states such as Florida, Louisiana, and Texas are often hit by hurricanes.Meteorologists measure the wind speed of hurricanes and use it to estimate the potential property damage, rating it on a scale of 1 to 5.

<u>Why are hurricanes problematic?</u>

First and foremost, hurricanes often lead to casualties. Notably, in recent years, research has found that hurricanes in the United States have killed more people than meteorologists would normally project (Borenstein, 2023). The people killed are disproportionately low-income, and in some minority or vulnerable group. Additionally, hurricanes destroy buildings and other important infrastructure (Borenstein, 2023). This takes huge amounts of money to repair. Oftentimes, local governments and individuals do not have the funds to repair damaged property, especially when hurricanes occur on a frequent (i.e. annual basis). Furthermore, hurricanes often disrupt local economies, leading to lower job growth and higher unemployment rates (Gooulborne, 2021).

<u>Motivation:</u>

Hurricanes have caused severe natural, economic and social issues to communities across tropical and subtropical areas, and it often takes a greater amount of time to rebuild the communities with human and nonhuman resources. Predicting the maximum sustained wind speed of hurricanes helps to understand a general pattern of hurricanes and should be valued in order to prevent casualties in large areas. Understanding the patterns of wind speed will also provide accountable methods of testing other related natural and/or human-induced phenomena. Developing a model of prediction will ultimately lead us to understand trends in climate change and how it might affect communities worldwide.

Topic Description:

The literature has suggested climate change as a major determinant of hurricanes in the future. Warm ocean water and more moisture in the air both contribute to the formation of hurricanes, which are both heightened by climate change (Colbert, 2022). Interestingly, models have not predicted a significant change in hurricane frequency. However, they are predicting more intense storms with stronger winds and more rainfall, both potentially devastating changes (Colbert, 2022). While the literature has observed predictions of future hurricanes, we seek to find out whether past hurricanes have gotten more severe over time. The primary goal of our study is to determine whether hurricanes become more severe over time. We utilize historical data to discern changes in storm severity that have already occurred. We hypothesize that hurricanes have increased in severity over time.

Data:

Source of Data:

The storms data set is included in the dplyr package.

When/How was it Collected:

The data was collected by the NOAA Atlantic Hurricane Center between 1975-2021. Attributes and measurements were recorded for each storm, with storms after 1979 have data recorded every six hours for the duration of the weather event.

Cases:

In the original data set, the observations (row) include measurements taken of the storms along with columns that specify the date and time of the observation.

In our filtered data set, each observation (row) represents the data collected for the name, maximum wind speed and year of a specific hurricane. Some hurricanes have multiple observations because data was taken for the storm every six hours.

Description of Variables:

The variables in our research question are year and maximum wind speed. In this case, both variables are quantitative.

Methods

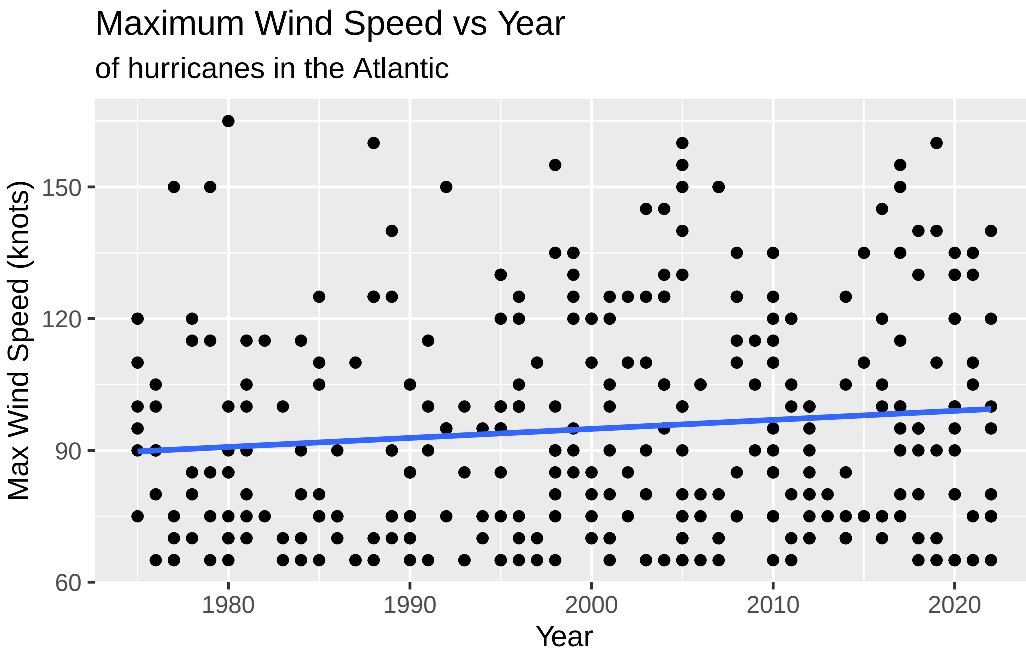
First, we looked at an overview of the storms dataset.

```
Rows: 19,537
Columns: 13
$ name                         <chr> "Amy", "Amy", "Amy", "Amy", "Amy", "Amy",~
$ year                         <dbl> 1975, 1975, 1975, 1975, 1975, 1975, 1975,~
$ month                        <dbl> 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6, 6,~
$ day                          <int> 27, 27, 27, 27, 28, 28, 28, 28, 29, 29, 2~
$ hour                         <dbl> 0, 6, 12, 18, 0, 6, 12, 18, 0, 6, 12, 18,~
$ lat                          <dbl> 27.5, 28.5, 29.5, 30.5, 31.5, 32.4, 33.3,~
$ long                         <dbl> -79.0, -79.0, -79.0, -79.0, -78.8, -78.7,~
$ status                       <fct> tropical depression, tropical depression,~
$ category                     <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ wind                         <int> 25, 25, 25, 25, 25, 25, 25, 30, 35, 40, 4~
$ pressure                     <int> 1013, 1013, 1013, 1013, 1012, 1012, 1011,~
$ tropicalstorm_force_diameter <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
$ hurricane_force_diameter     <int> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, N~
```

Then, we mutated the data set, adding a variable for maximum wind speed and selecting the variables that we planned to use in our analysis.

Here we manipulated the original storms data set to create a new data set (storms_mutated) to identify storms by their name and year (which we did via the group_by function) and add another variable that represents the maximum sustained wind speed of that storm (max_wind). We then selected our variables of interest (name, year, and max_wind) and removed the replicate rows to create a data set ready for statistical analysis.

```
`geom_smooth()` using formula = 'y ~ x'
```



Here we used a scatter plot to visualize the relationship between the variables max_wind and year. We added a line showing the linear relationship between these variables with the geom_smooth function.

According to the scatter plot, there is a weak positive relation between year and maximum sustained wind speed of hurricanes. In general, as the year increases, the maximum wind speed increases as well. A scatter plot directly shows the strength of correlation between maximum wind speed and year with a line of best fit that visualizes the trends, while each spot represents the specific maximum wind speed in a year.

```
[1] 0.1121283
```

Here we used a correlation test to assess the strength of the relationship between max_wind and year.

The correlation test shows that there is a weak positive relation between year and maximum wind speed of hurricanes. The output of 0.12 is positive, which indicates that there is a positive relation, while the fact that it is close to 0 suggests that there is a weak correlation between year and maximum wind speed.

```
parsnip model object
```

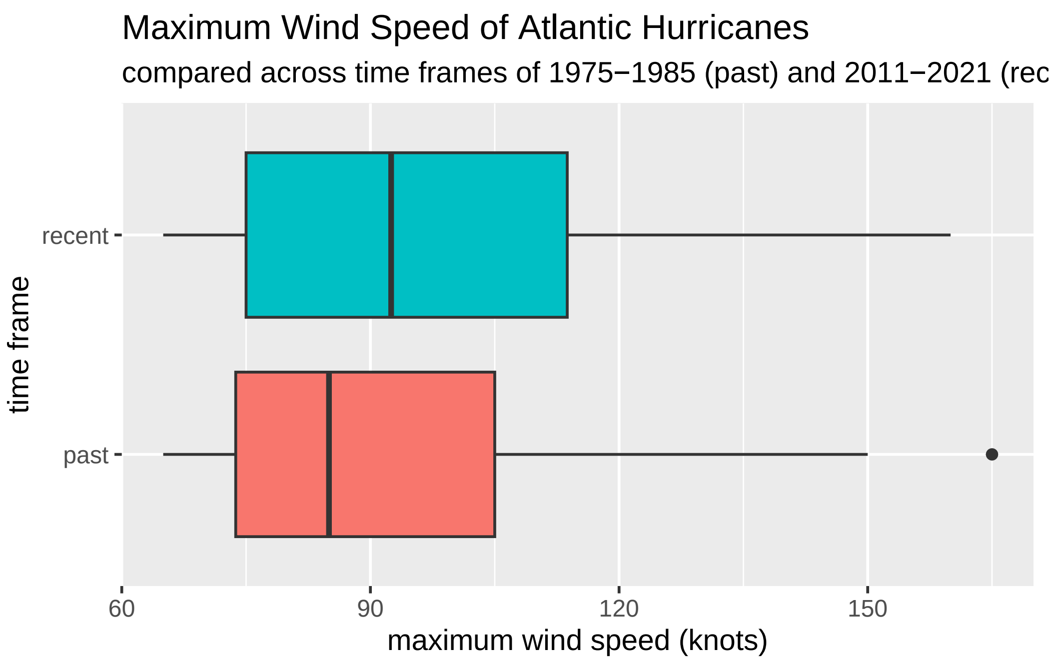
```
Call:
stats::lm(formula = max_wind ~ year, data = data)

Coefficients:
(Intercept)         year
  -316.0204       0.2055
```

Here we calculated the linear regression relationship between max_wind and year.

$$\widehat{MaxWindSpeed} = -370.7376 + 0.2329 * Year$$

After each additional year, we estimate on average an 0.2329 knot increase in the maximum wind speed of hurricanes.

We wanted to compare the first ten years of maximum wind speed of hurricanes collected in this data set to the most recent ten years. To do this, we created a boxplot to visualize the differences between these two time periods. We also calculated the mean wind speed of each time set and the difference between them. Then, we used hypothesis testing (including creating a null distribution) to cacluate the probability of observing this difference.



Here, we created a box plot showing the different ranges in maximum wind speed for "past" hurricanes (1975 - 1985) and "recent" hurricanes (2011 - 2021).

Our boxplot illustrates that the mean maximum wind speed for "recent" hurricanes is higher than that of "past" hurricanes. Additionally, "recent" hurricanes have a wider range in mean maximum wind speed than "past" hurricanes, unless one includes major outliers in range calculations (in which case "past" hurricanes would have a bigger range).

```
# A tibble: 2 x 2
  timeset  mean
  <chr>   <dbl>
1 past     90.4
2 recent   96.2
```

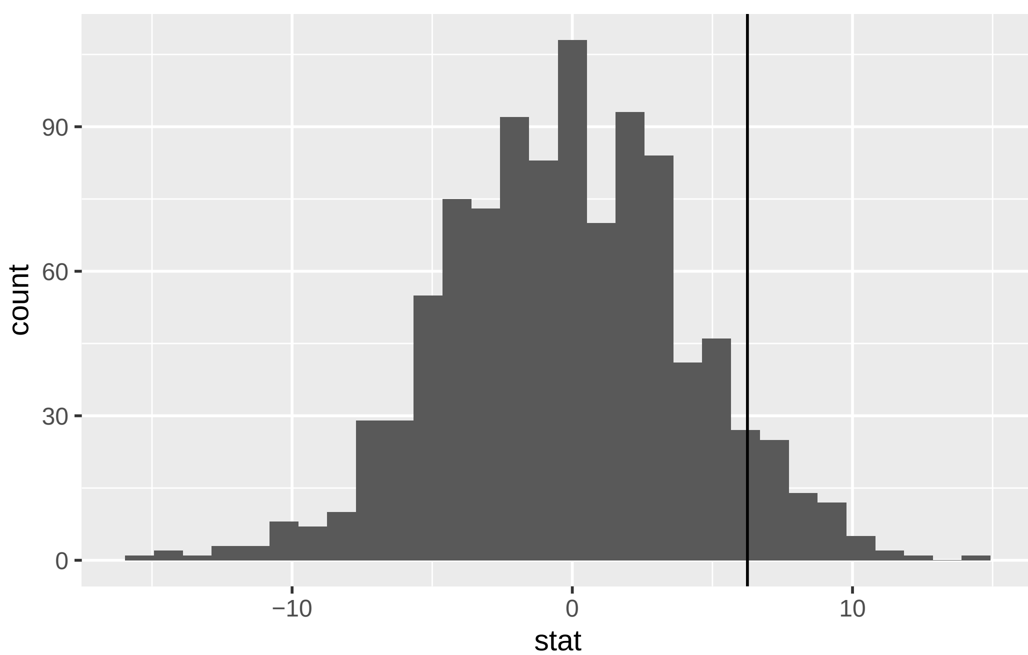
$$\bar{x}_{recent} - \bar{x}_{past} = 6.25$$

Before performing our hypothesis testing, we established our null and alternative hypotheses.

$$H_O : \mu_{recent} - \mu_{past} = 0$$

$$H_a : \mu_{recent} - \mu_{past} > 0$$

```
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```
# A tibble: 1 x 1
  p_value
    <dbl>
1    0.07
```

With a alpha level of 0.05 and a calculated p-value of 0.059, we fail to reject the null hypothesis that the true mean maximum wind speed of hurricanes from the "recent" time set is equal to that of hurricanes from the "past" time set. In conclusion, we have weak evidence to support the alternative hypothesis that the true mean maximum wind speed of hurricanes from the "recent" time is greater than that of hurricanes from the "past" time set.

Results

Scatter Plot & Correlation test:

According to the scatter plot and correlation test, there is a weak positive relation between year and maximum sustained wind speed of hurricanes. This means that while there may be some positive correlation between maximum sustained wind speed and the recentness of the year in which the hurricane occurred, it is probably not significant enough to fully support our hypothesis. Because of this, we decided to conduct further analysis before concluding whether our hypothesis was supported by the data. However, we observed that there was a larger number of more recent hurricanes that had a maximum wind speed of 120 knots or higher, which is in line with our hypothesis.

Linear Regression:

We use linear regression to measure the relation between year and the maximum wind speed for hurricanes. A slope of 0.2329 indicates that for 1 increase in year, we estimate on average an 0.2329 knot increase in the maximum wind speed of hurricanes. This linear regression model shows that the amount of increase in the maximum wind speed of hurricanes is rather small, which correspond to our conclusion that there is no true difference in means of maximum wind speed between "recent" hurricanes and "past" hurricanes, while the slope demonstrates a general trend of increase maximum wind speed in each year.

Boxplot & Hypothesis Testing:

As displayed in our boxplot, there is a higher mean maximum wind speed for "recent" hurricanes in comparison to "past" hurricanes, which is in line with our hypothesis that hurricanes have gotten more severe over time. By contrast, in performing the hypothesis testing, we failed to reject our null hypothesis that the true mean maximum wind speed of hurricanes from the "recent" time set is equal to that of hurricanes from the "past" time set. We also collected weak evidence to support our alternative hypothesis that the true mean maximum wind speed of hurricanes from the "recent" time is greater than that of hurricanes from the "past" time set. Though this does not support our initial hypothesis that hurricanes have increased in severity overtime, it does provide an example of a limitation in our analysis.

The findings from our hypothesis testing are congruent with our previous findings represented in our r-value from the correlation test and our slope from the linear regression. Though there was a slight increase in wind speed over the 1975-2021 time period, the correlation between year and wind speed is weak, as indicated by our r-value. Our slope confirms this week correlation as we estimated on average only a 0.2329 knot increase in maximum wind speed per year. Thus, it is logical that the true mean maximum wind speed of hurricanes from "recent" time would not be greater or much greater than that of the "past" time, because the change in maximum wind speeds over time is not large enough to create a significant difference between the two means.

Scope of Inference:

The limitation in our analysis is that the analyzed data only covered 46 years. In comparison to how gradually change occurs in nature, this is not a long time. To see a more significant difference, we would need to analyze the maximum wind speeds over multiple decades, potentially even one or two centuries. Even so, our overall data analysis does support our hypothesis that hurricane severity has increased over time, but further research should be done to confirm this trend over a longer time period.

Ethical Considerations

One main ethical consideration is that climate change deniers might use this to further their agenda. We found a weak positive correlation between time and wind speed. Because there was only weak evidence that the mean hurricane wind speed in recent years is greater than it was in the past and therefore that hurricane severity has been increasing, climate change deniers might use this to argue that climate change isn't real and hurricane servirity is not increasing as a result of anything other than coincidence.

Additionally, we failed to reject the null hypothesis that the true mean maximum wind speed of hurricanes from the "recent" time set is equal to that of hurricanes from the "past" time set. Because we failed to reject the null, climate change deniers might also use this to support their belief that climate change is a hoax because the increase in hurricane strength over time is often used as evidence of climate change.

Sources

[@NASA]

[@BrookingsInstitute]

[@NOAA]

[@APNEWS]