**How this all started…**

“Here we go again!”, where the words that ran through my mind when Hurricane Dorian was forecasted to come over Florida. The literal storm and stress induced by the fact that the phenomenon, which lacks malice nor mercy as it destroyed the Bahamas fueled a wide variety of storm-related debates and heated statements at my workplace. One said statement, courtesy of a good friend was that “each year it keeps getting worse”; that was met with the following question: how so? The answer being a general statement that we are seeing more storms than in past years and that they were getting stronger. My response to my co-worker: “I will get back to you.”

**Statistical/Hypothetical Questions:**

* Has the number of occurrences in terms Tropical Storms/Hurricanes/Storms in the Atlantic increased over the years?
* Are more recent Tropical Storms/Hurricanes/Storms in the Atlantic stronger (faster) then those recorded in the past?
* How does the weather from the last couple of years when storms occurred compare to more recent years?
* Is there a correlation between the strength/speed of a storm and the temperature?
* Is there a correlation between the length/duration of a storm and the temperature?

**Findings:**

* The observations presented in the “atlantic.csv”, when grouped by Event and Days show an increase in activity by day over the course of time.
* In terms of storm strength based on Maximum Winds sustained, a large number of observations fell below 50 knots (for 1-minute durations), which in the Saffir-Simpson hurricane wind scale (1) classify as Tropical Storms.
* The average land and ocean temperature during the months where storms occurred ranged between 56- and 64-degrees Fahrenheit.
* Looking at the Probability Mass Function for storm duration:
  + Between 1851 to 1933 – Storms/Hurricanes had a high probability of lasting 5 days.
  + Between 1934 to 2015 – Storms/Hurricanes had a high probability of lasting longer than 10 days or 15 days.
* Probability Mass Function for the Land-Ocean Average Temperatures:
  + Between 1851 to 1933 – there was a high probability of storm activity between 61- and 62-degrees Fahrenheit.
  + Between 1934 to 2015 – Shows a similar trend as the previous cycle, but with an increase in storm activity between 62- and 63-degrees Fahrenheit.
* The Cumulative Distribution shows that storms that last 10 days or under fall under the 80 percentiles. The average storm duration being around 11 days with a standard deviation of about 5 .58 days.
* Running a Scatter Plot and correlation analysis of:
  + Maximum Winds and Land/Ocean Average Temperatures – There appears to be a weak, negative correlation with *p < 0.05*.
  + Storm duration and Land/Ocean Average Temperatures – With the p being greater than 0.05, we can discard this association.

**The Outcome – Challenges, Learning Opportunities and other items to address**

This initial analysis has provided context that we are seeing more Tropical Storms over the course of time and that there is a correlation (albeit) weak between storm strength and temperature. However, this round of analysis did not provide any findings on whether storms where actually getting stronger nor did it explore other aspects such as movement activity.

The missing pressure values were a bit disheartening – Because of the large amount of missing/non-reported values, I had to exclude this variable from the final analysis. Had there been more data I believe it would’ve helped provide more context.

Another piece of the puzzle was the coordinate data provided in the atlantic.csv. I lacked the understanding on how to actually use this data. In theory, this data can be used to determine the actual movement ration along with the hours to determine the speed of movement. This information could provide some additional theoretical insights such as regarding a storm’s lifecycle and intensity.

As far as challenges go, I had a bit of time trying to decipher how to perform the Analytical Distribution. My attempt at a Pareto Distribution was a bit lacking so I opted to do a Normal Distribution. If anything, this an indication that I need take the time to revisit this and any other concepts that I may be struggling with.

**Conclusions**

As far as my conclusions go – I based them off on my interpretations of the data. Which could for all intents and purposes be wrong. If anything, this is good starting point to find more data and continue to refine this analysis.

**References:**

[1] Wikipedia. Saffir-Simpson scale. Article. Retrieved November 16, 2019 from <https://en.wikipedia.org/wiki/Saffir%E2%80%93Simpson_scale>