

Report

Xinran Sun

5/7/2022

Introduction

Dataset

hurrican703.csv collected the track data of 702 hurricanes in the North Atlantic area from 1950 to 2013. For all the storms, their location (longitude & latitude) and maximum wind speed were recorded every 6 hours. The data includes the following variables

1. **ID**: ID of the hurricanes
2. **Season**: In which **year** the hurricane occurred
3. **Month**: In which **month** the hurricane occurred
4. **Nature**: Nature of the hurricane
 - ET: Extra Tropical
 - DS: Disturbance
 - NR: Not Rated
 - SS: Sub Tropical
 - TS: Tropical Storm
5. **time**: dates and time of the record
6. **Latitude** and **Longitude**: The location of a hurricane check point
7. **Wind.kt** Maximum wind speed (in Knot) at each check point

From the original dataset, we built a new dataset with contains five more variables, including:

1. **Wind_prev**: wind speed at 6 hours ago
2. **Wind_prev_prev**: wind speed at 12 hours ago
3. **Lat_change**: latitude change compared to 6 hours earlier
4. **Long_change**: longitude change compared to 6 hours earlier
5. **Wind_change**: wind speed change at 6 hours earlier compared to 12 hours earlier

These variables will help us to build the model in the following part.

The *hurricaneoutcome2.csv* recorded the damages and death caused by 46 hurricanes in the U.S, and some features extracted from the hurricane records. The variables include:

1. **ID**: ID of the hurricanes
2. **Season**: In which **year** the hurricane occurred
3. **Month**: In which **month** the hurricane occurred
4. **Nature**: Nature of the hurricane
 - ET: Extra Tropical
 - DS: Disturbance
 - NR: Not Rated
 - SS: Sub Tropical
 - TS: Tropical Storm
5. **Damage**: Financial loss (in Billion U.S. dollars) caused by hurricanes
6. **Deaths**: Number of death caused by hurricanes
7. **Maxspeed**: Maximum recorded wind speed of the hurricane
8. **Meanspeed**: average wind speed of the hurricane
9. **Maxpressure**: Maximum recorded central pressure of the hurricane
10. **Meanpressure**: average central pressure of the hurricane
11. **Hours**: Duration of the hurricane in hours
12. **Total.Pop**: Total affected population
13. **Percent.Poor**: % affected population that reside in low GDP countries (i.e. GDP per Capita \leq 10,000)
14. **Percent.USA**: % affected population that reside in the United States

EDA

We use a bar plot to examine the number of hurricanes in each month. From Figure 1, we can see that September is the month with the most hurricanes, while there are no hurricanes in February and March. Hurricanes in September also have the highest average wind speed as we can see in Figure 2.

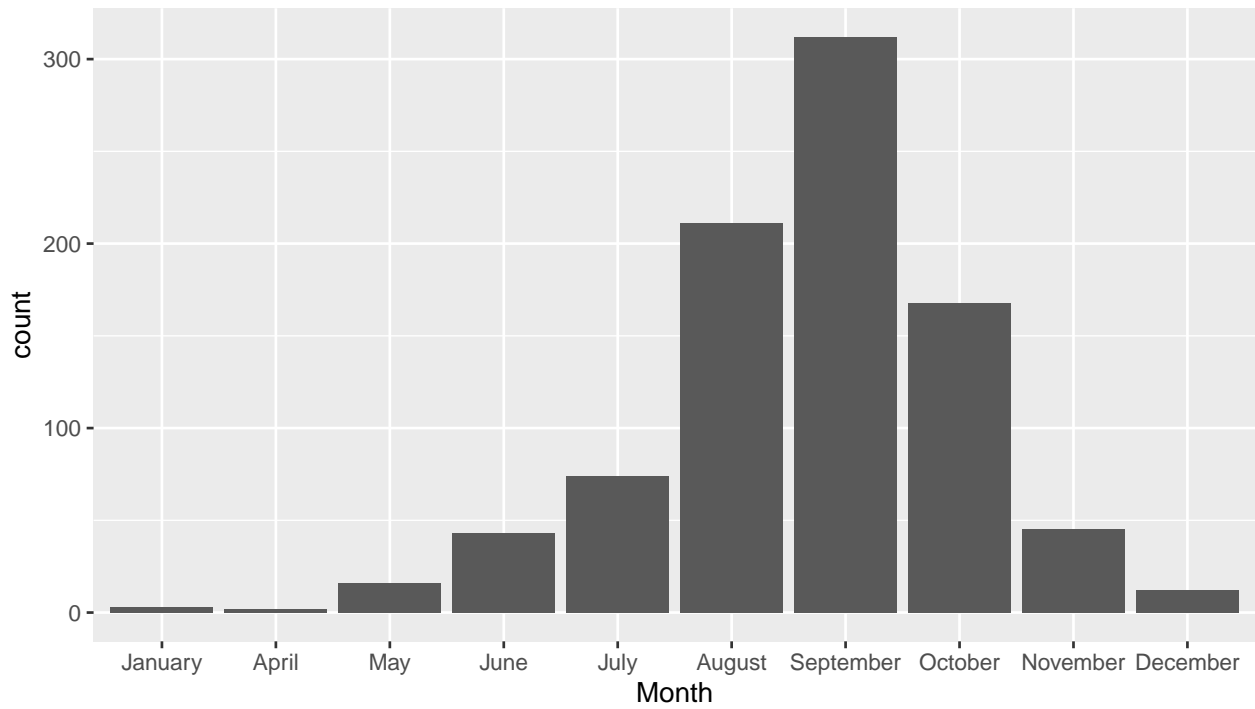


Figure 1. Count of Hurricanes in Each Month

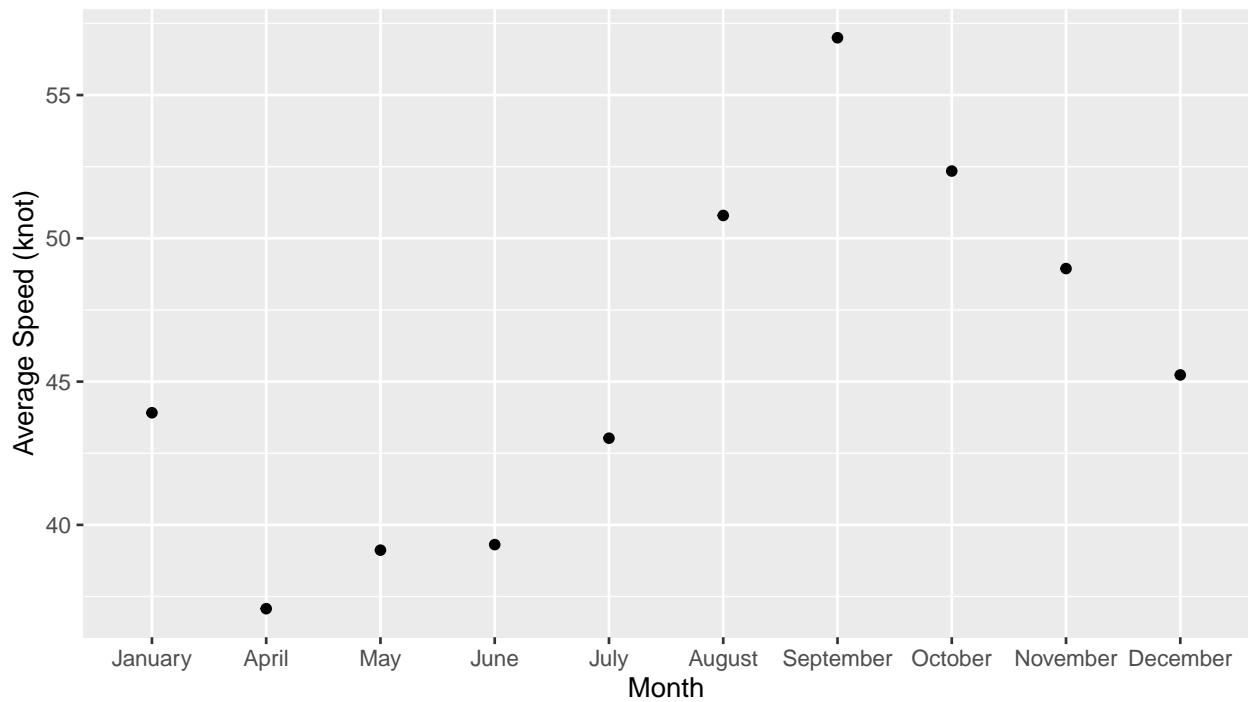


Figure 2. Average Speed (knot) of Hurricanes in Each Month

If we group the hurricanes by years, we can see in general, we have more observations in recently years compared to 50 years ago as shown in Figure 3. However, from Figure 4, the average wind speed seems to have a decreasing trend.

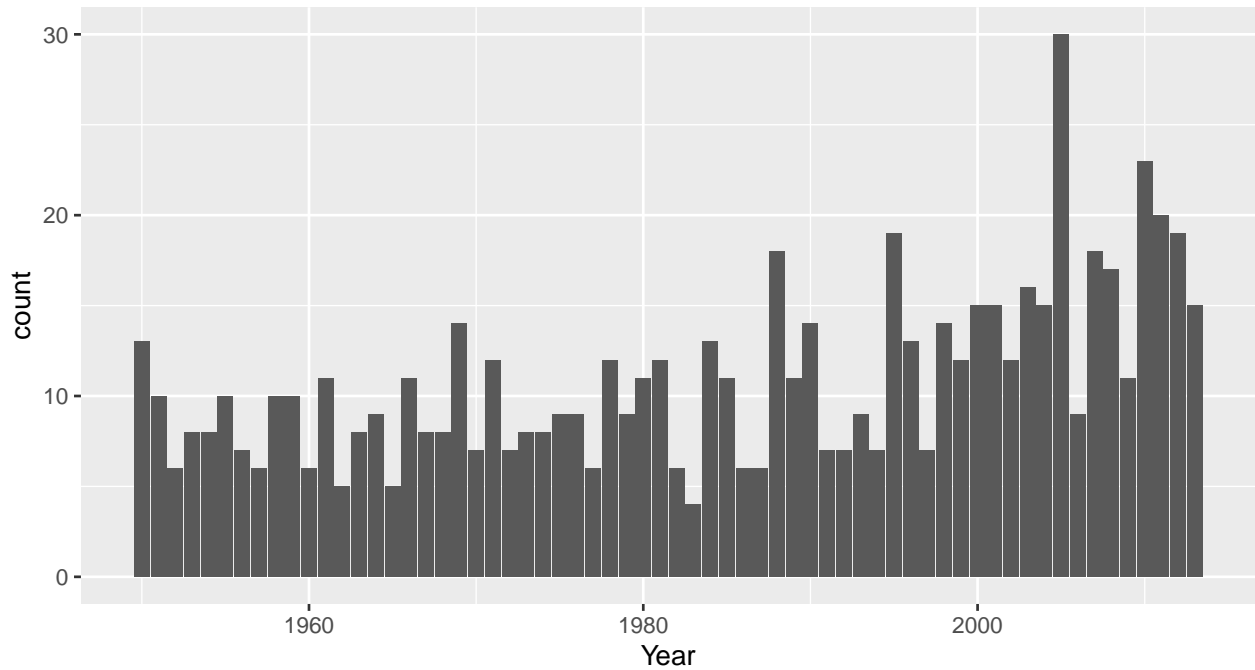


Figure 3. Count of Hurricanes in Each Year

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

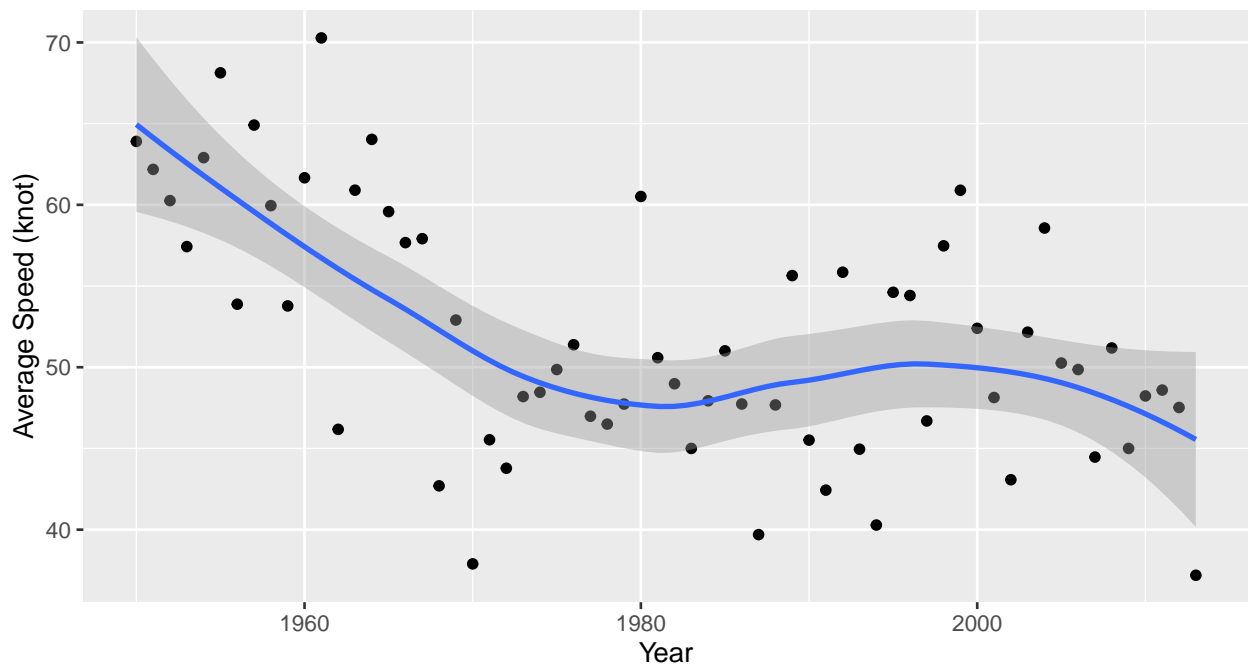


Figure 4. Average Speed (knot) of Hurricanes in Each Year

We also compare the hurricanes with different natures. In our dataset, there are 1214 different nature ratings. This number is larger than the number of hurricanes because some hurricanes are in different natures at different time. From Figure 5, we know that more than half of the natures are in Tropical Storm category.

This nature also have the highest average wind speed at about 60 knot, while the disturbance and not rated hurricanes have average wind speed at round 20 knot as Figure 6 illustrates.

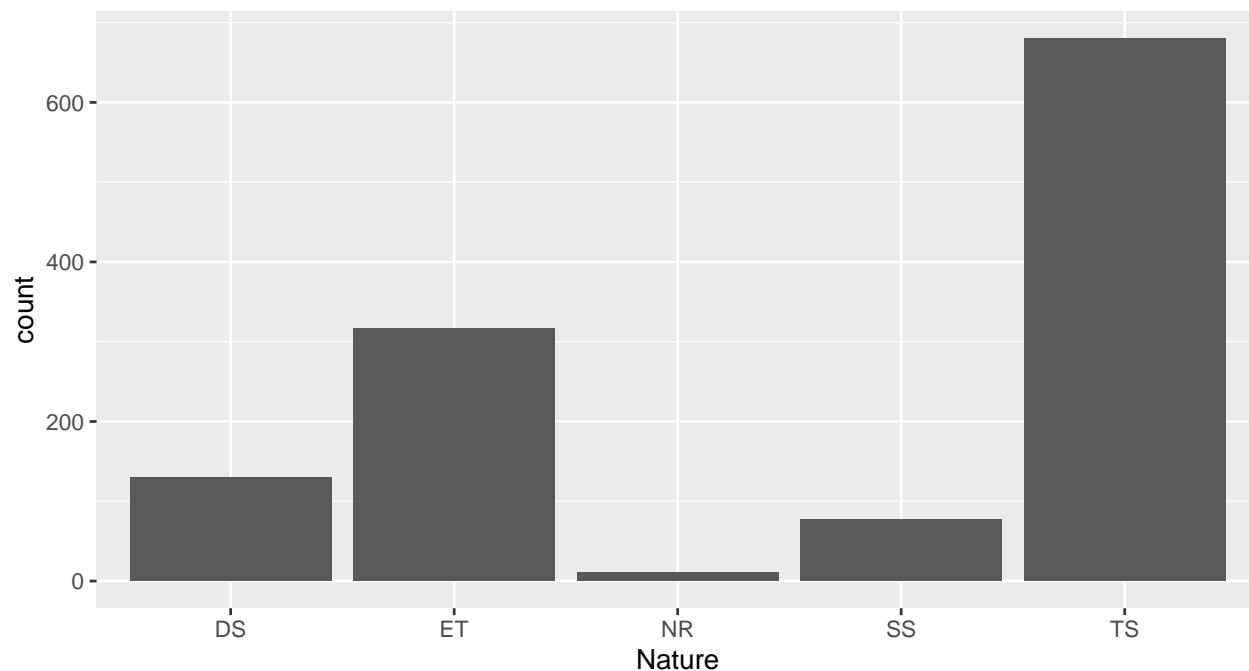


Figure 5. Count of Hurricanes in Each Nature

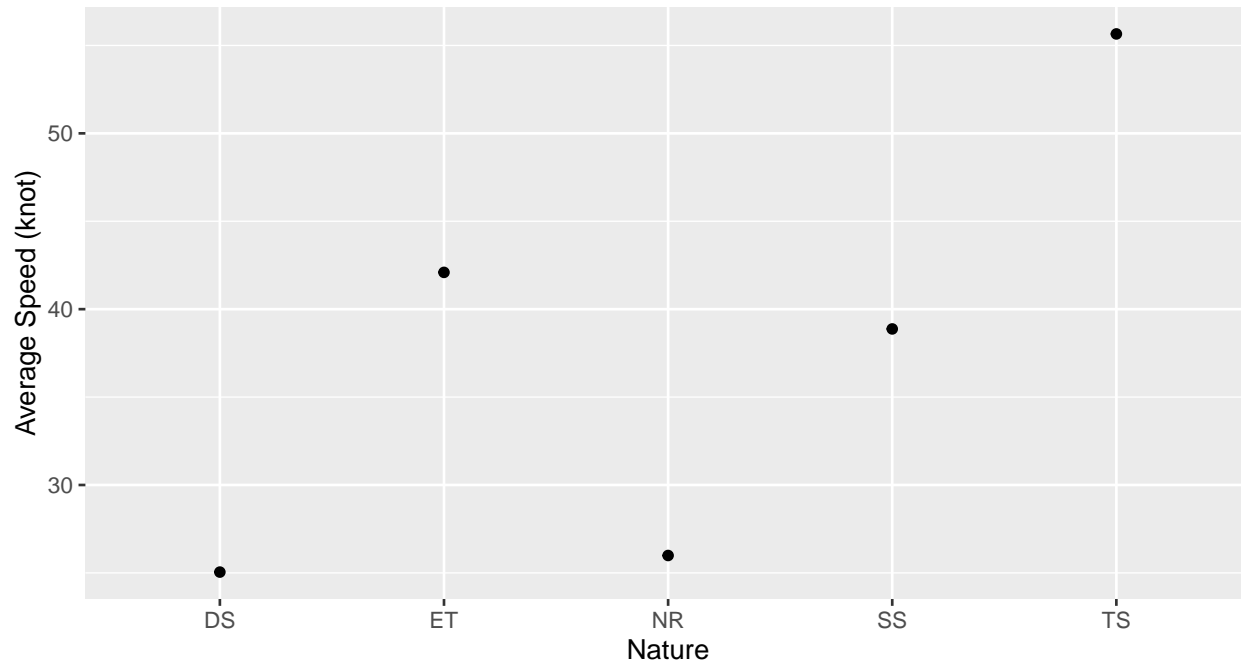


Figure 6. Average Speed (knot) of Hurricanes in Each Nature