Assignment Five

BSDS Spring 2021

(due 2/19, 11:59pm PST)

Note: You must submit this assignment as both a Jupyter Notebook file (.ipynb) and an (.html) file on Canvas. Please use markdown cells to indicate which parts of the notebook answer which questions. If a question requires a short answer rather than code, use a markdown cell.

In the next few problems you will explore the `storms` dataset which is automatically loaded into your working memory when you load the `tidyverse` library.

1. List each of the variables in the `storms` dataset and their variable type. Are there any variables which could also be given a different variable type? Explain why or why not.

2. What does each row of the dataset represent?

3. Create the following datasets.

a. All storm measurements after the year 2000 (including this year) that occur on January 6th.

b. All storm measurements with wind speed strictly higher than 30 knots or pressure strictly lower than 1000 millibars.

c. All storm measurements of category 1 or higher that did not take place in January, August, or December.

d. A dataset where the rows are ordered by wind speed (highest wind speed is first)

e. A dataset consisting of only the date, time, status, and category of the storm measurements.

4. Add the following variables to the storms dataset. Always name your new variables so that another person can understand what it might be.

a. The wind speed in miles per hour (hint: look up the conversion from knots to mph).

b. Wind speed times diameter of the area experiencing hurricane level winds.

5. Create the following datasets using the pipe. (make sure to remove missing values when taking the mean)

a. A dataset with the name of each storm and the average pressure over all its measurements, with rows sorted by average pressure (lowest first)

b. A dataset with the average tropical storm diameter for each category of storm

c. A dataset consisting of only hurricanes with the average hurricane diameter for each category.

d. A dataset with the name of each storm and the number of measurements it has

6. Which storm has the highest average wind speed? Which storm has the most measurements?

7. Exercise 5.4.1.1 (at least two different ways) using the `flights` dataset and Exercise 5.2.4.3

8. (BONUS) Pick any storm and create a map using ggplot of the path of the storm. Hint: use the latitude and longitude.