Day 3 Worksheet

Long Term Trends in Seawater Temperature

NOAA has extensive databases for long term monitoring of various environmental factors that relate to weather. One of my research objectives is to understand the long-term change in environmental factors due to climate change. One impact of climate change is that the ocean temperatures are warming but this varies a lot depending on where your site is located. Do you think the Pacific Ocean off of the California coast is getting warmer?

For this exercise I want you to analyze the dataset “mon\_seatemps.csv”, which consists of over 167,000 temperature data points. These data are downloaded from the NOAA buoy website, and are sea surface temperatures (in degrees Celsius) for a buoy in the middle of Monterey Bay. Here is the website where I downloaded the data: <https://www.ndbc.noaa.gov/station_page.php?station=46114>

I’ve compiled the spreadsheets into one dataset for sea surface temperatures for ~12 years, from September 2011 to December 2022. This dataset has 5 columns, one for the year (you should set this as a factor, ***as.factor ()***), the hour of the day, the minute (there is a datapoint every 30 minutes), temp which is the surface seawater temperature in Celsius, and date (you need to set as a Date in R, ***as.Date ()*** ). Use the following questions to practice methods to summarize a data set. Below are some commonly used R functions, but I also suggest you have your script from yesterday open for reference, for any relevant action you can copy from your old script to make a new script. Here is a brief summary of some of the R functions you might want to use;

* To understand the structure/column titles of a dataset: ***str ()***, ***head ()***
* Summarize a column*:* ***mean (), sd (), var (), min (), max (), summary (), table ()***
* Manipulate the dataset: ***subset (), aggregate ()***
* Ggplot2 functions you might need: ***ggplot(), geom\_points(), geom\_histogram (), geom\_boxplot (), geom\_col (), geom\_line ()***

Part 1: Load and Analyze the Data

Answer the following questions about the Monterey seawater temperatures (“mon\_seatemps.csv”):

1. What was the maximum seawater temperature for 2012?
2. What was the mean seawater temperature for 2019?
3. What was the minimum seawater temperature for 2020?
4. Which years have a maximum seawater temperate of more than 18 °C? Which year had the highest seawater temperature?
5. What is the minimum seawater temperature for every year? Which year had the lowest seawater temperature?
6. Graph (using ggplot) the mean seawater temperature (aggregated by year) as a bar graph, ***geom\_col()***
7. For each year graph the seawater temperature using a boxplot, ***geom\_boxplot ()***.

Part 2. Threshold Temperatures

From the scientific literature there was a marine heat wave during 2014 and 2015 that killed much of the bull kelp (*Nereocystis luetkeana*) in Northern California (Rogers-Bennett and Catton, 2019-is a scientific paper that documents this change). One species of bull kelp (*Nereocystis luetkeana*) that live in California die when sea surface temperatures get to 16-18 °C.

1. Using the entire dataset graph seawater temperature by date to make a line graph. Use ggplot to add legends to each axis. Instead of a title add a caption, start your caption with Figure 1. Then add a brief sentence about what your figure is showing.
2. Add a horizontal line to this graph at a temperature of 18 degrees C. When the temperatures go above this line the kelp described above will die.
3. How many years did the seawater temperature go above 18 degrees C?
4. Aggregate seawater temperature by date for maximum temperature and minimum temperature.
5. Make one graph with a line for each of these functions of the seawater temperature. Make each line a different color so you can see how these functions vary over the years.
6. There are two types of data that you could use to calculate a change in temperature due to climate change. Above in #3 you looked for the number of years over a threshold.
7. Or you could look for a change in the range of temperatures. Has the range (difference between the maximum and minimum temperatures) gotten larger over the years. Can you graph the range data? What type of time scale is most informative, is it by day or by year? Do these data suggest that the variation in seawater temperature is increasing?

Look at your graph and think about the following questions. Do you think that the maximum sea surface temperatures are returning to previous levels? Do you think this species of kelp will be able to survive in Monterey in the future? What evidence do you have for your conclusion? Do you think 12 years is a long enough duration to look for changes in climate?

Part 3. Bonus activity- Long Term Hawaiian Seawater Temperatures

Answer the following questions using seawater temperature taken from a NOAA buoy found on Hawai‘i. This buoy is found at Kaneohe Bay, which is a large bay located on the eastern shore of Oahu. This data set is called hi\_seatemps.csv. In this dataset date is imported in a different format. To tell R this column is a date you will need to use the ***as.POSIXct ()*** function. Additionally, there are some missing data in this dataset. These will interfere with your calculations, you can omit rows with missing values by using the fuction ***na.omit ()***

1. What was the maximum seawater temperature for 2004?
2. What was the mean seawater temperature for 2014?
3. What was the minimum seawater temperature for 2006?
4. Which years have a maximum seawater temperate of more than 30 °C?
5. What is the minimum seawater temperature for every year?
6. Graph (using ggplot) the mean seawater temperature (aggregated by year) as a bar graph, ***geom\_col()***
7. Make a line graph of the temperature by date. Use a horizontal line at 30 degree C to illustrate which years had a seawater temperature greater than 30 degrees C.

In Hawaii corals bleached in both 1996 and 2014. Do both of these years have maximum temperatures greater than 30 degrees C? Were there any other years that had high temperatures where you think corals might have bleached?

Part 4. Bonus activity- Advanced Graphing

Sometimes we want to compare sites. Can you put the data for Monterey and Hawaii on the same graph? Make sure that each line is a different color. Clearly, they are different temperatures, but can you figure out if they have different ranges? Try to plot the max and min for each site. Can you put all 4 lines on the graph. How might you change the format of the lines to clearly display each site and max vs. min. Do you see how this graph displays how wide the range is. Which site has a wider range? Do you think that the range changes over time for either/both sites?