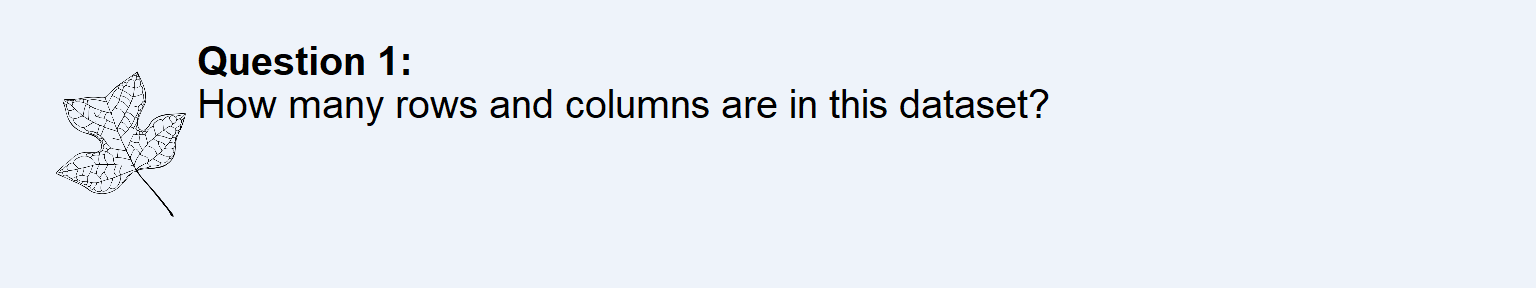
Charlie Huemmler

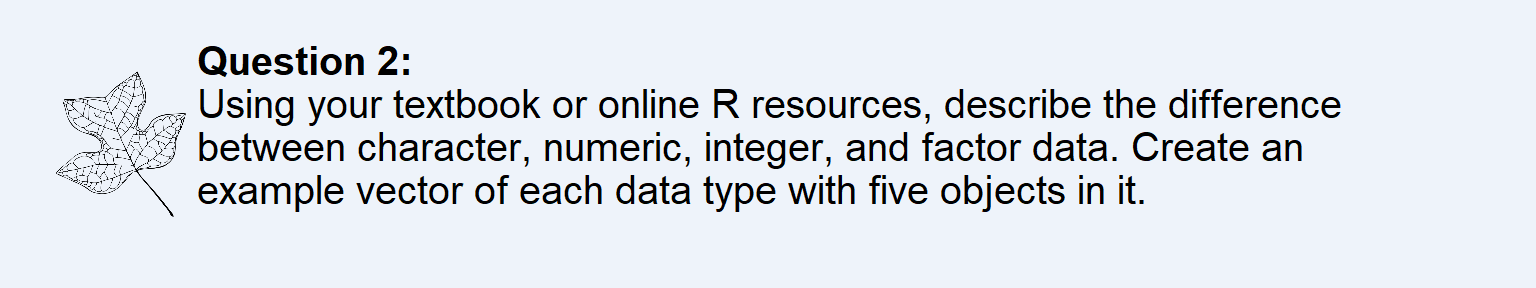
Environmental Data Science

Prof. Kropp

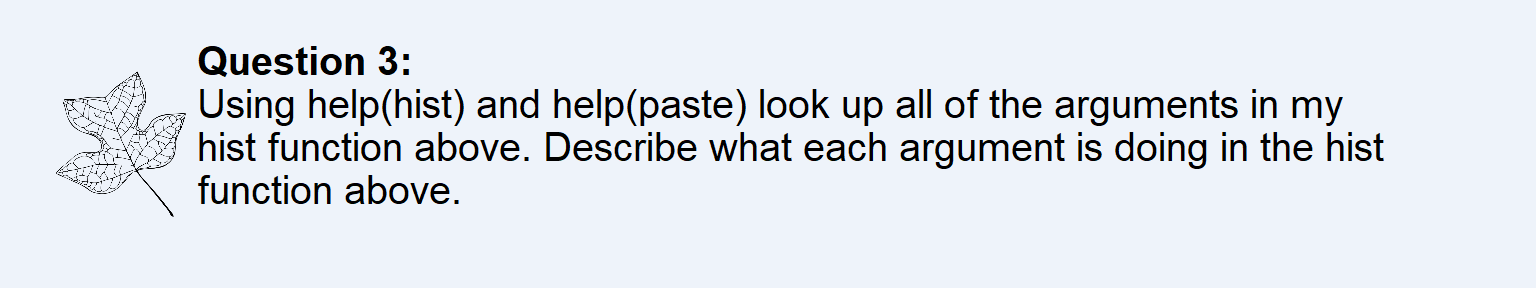
Activity 2



The NOAA data set has 157,849 rows (observations) and 9 columns (variables).

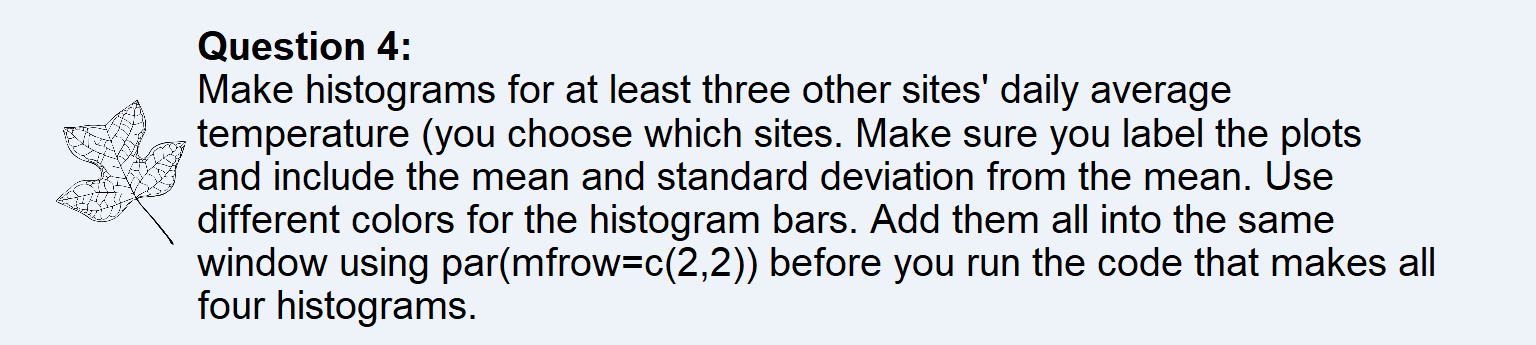


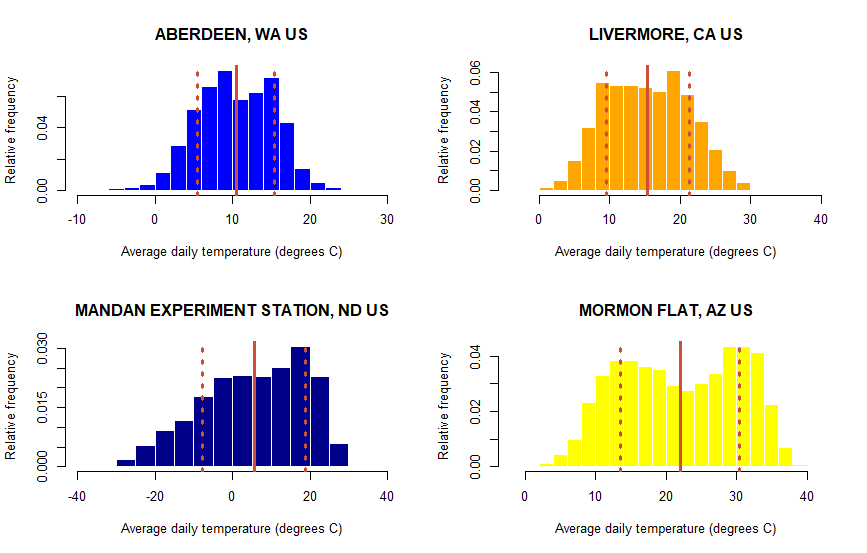
A character is a string, otherwise just words. Numeric data is numbered data, with or without a decimal point. Integers are whole number data points. Factors are character vectors with an attached integer vector for ordering.

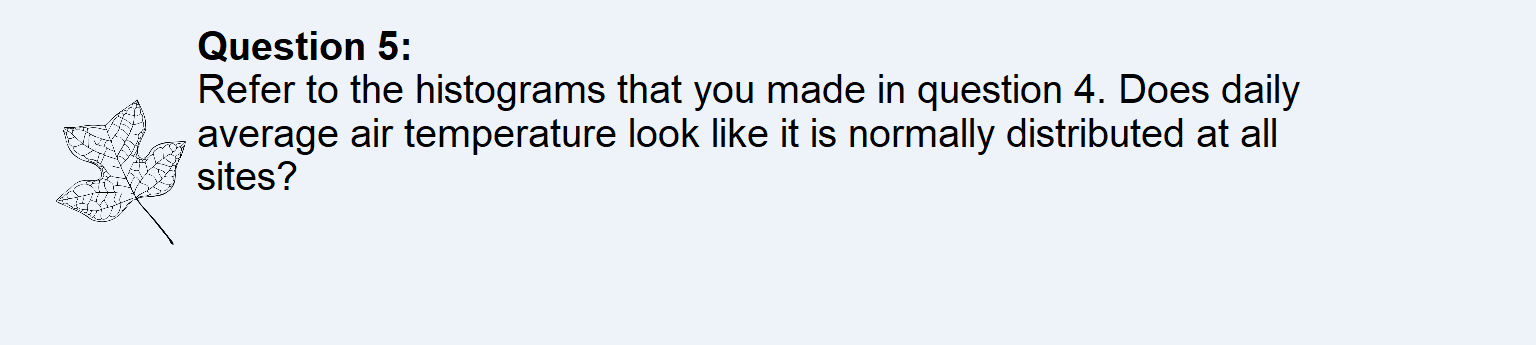


Main

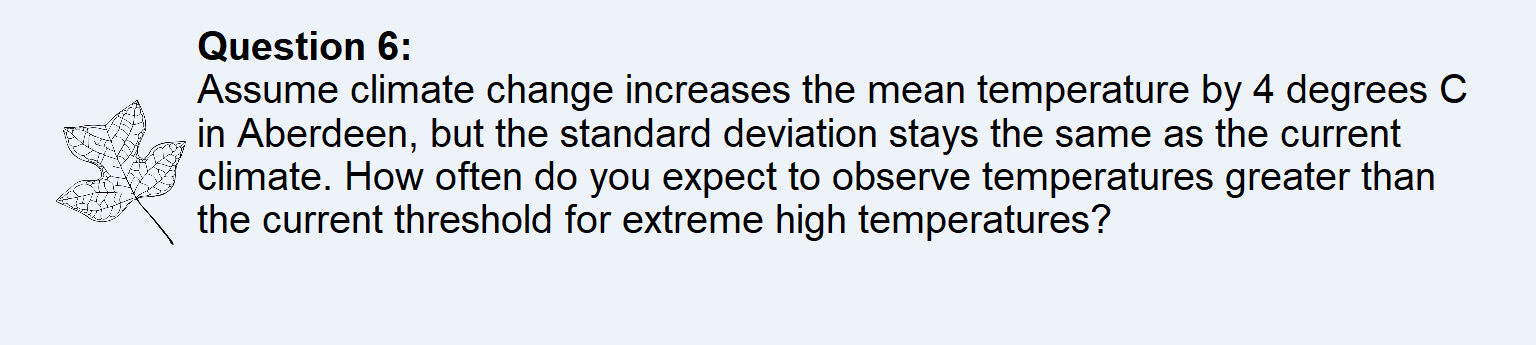
|  |  |
| --- | --- |
| Argument | What it does |
| datW$TAVE[datW$siteN == 1] | specifys the data to visualize. I.e: looking at the average temperature for site 1 |
| Freq = FALSE | tells the graph to display the density (relative frequency) instead of frequency of the data. |
| main = paste(levels(datW$NAME)[1] | Main title for graph |
| xlab = "Average daily temperature (degrees C)" | Sets X axis label |
| ylab="Relative frequency" | Sets Y axis labels |
| col="grey50"  Border = “white” | Sets the color of the bars  Sets color of border |



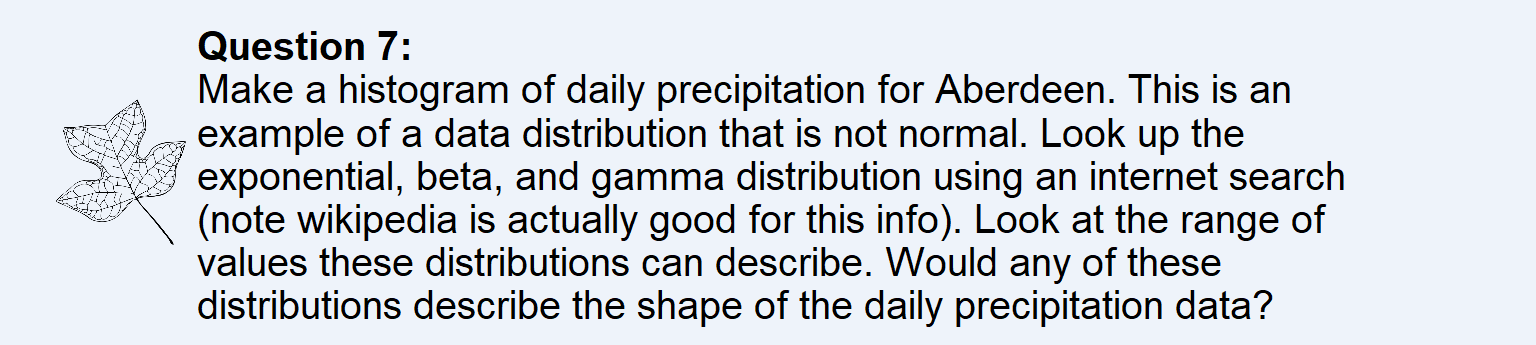


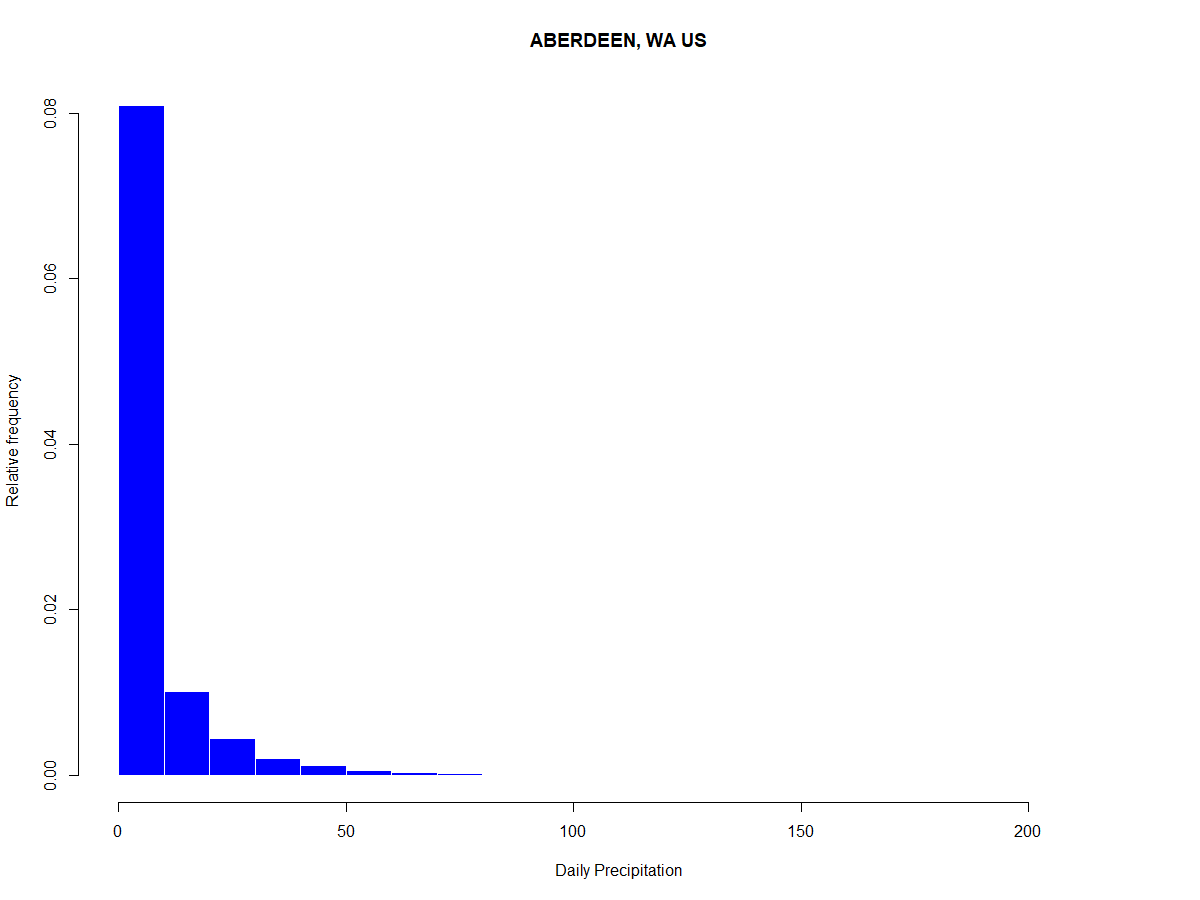


Some of the sites appear more normally distributed in air temperature than others. Livermore and Aberdeen seem fairly normal, while Mandan is left skewed and Mormon Flat is bi-modal.

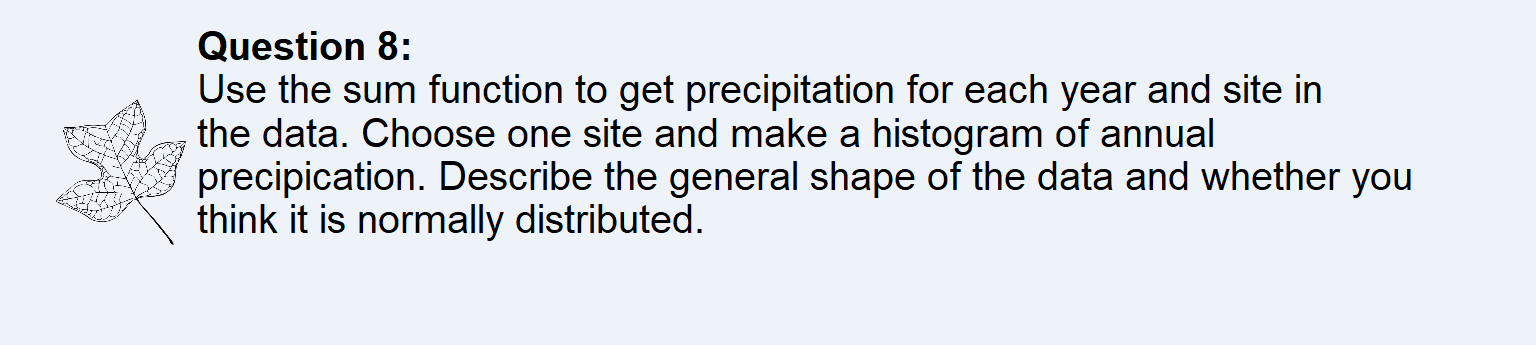


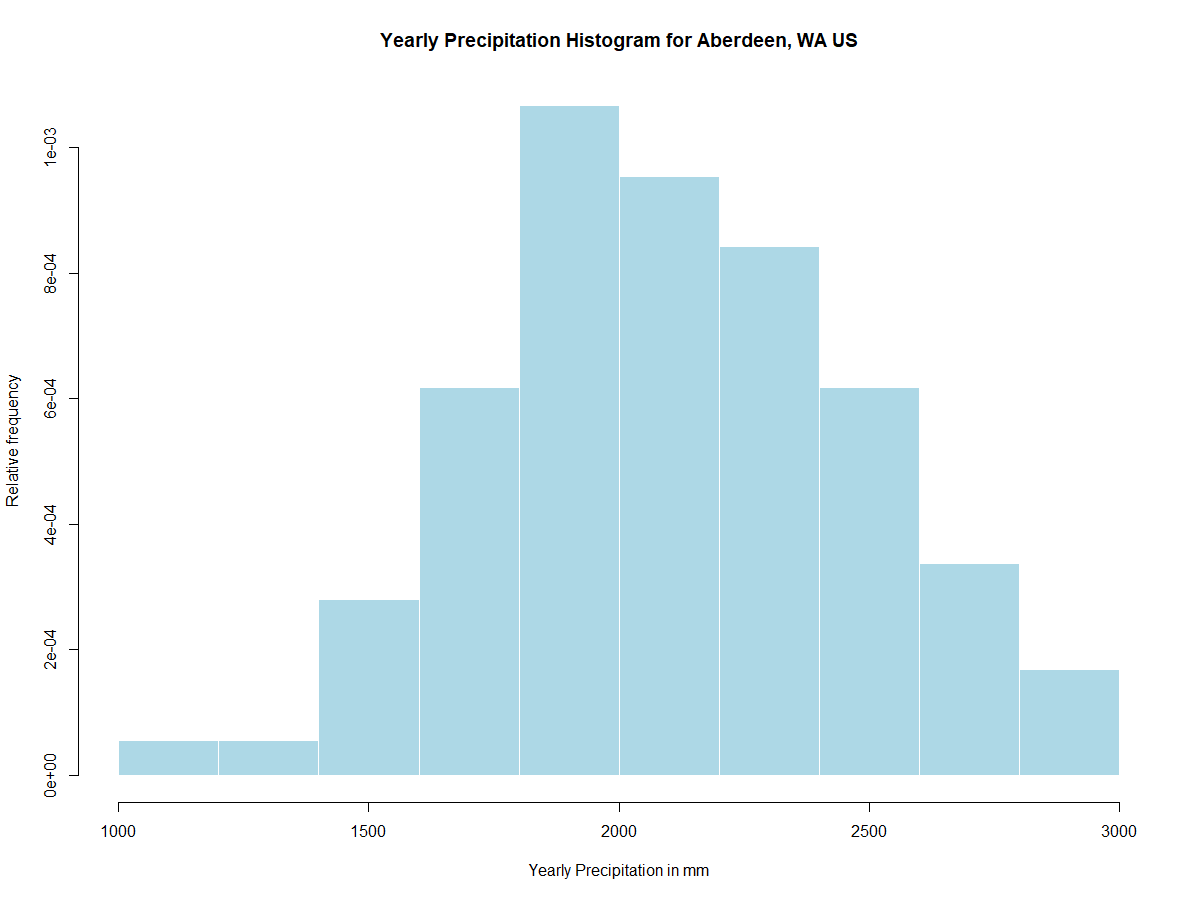
We will expect to see temperatures above the 18.51 degrees C threshold 20.31% of time, given a 4 degrees C increase in mean temperature in Aberdeen.





The probability distribution of daily precipitation appears to follow the gamma distribution with shape parameter equal to 1, which is otherwise known as the exponential distribution. Low levels of precipitation have the highest frequencies, and as the amount of precipitation increases the frequency quickly lowers.





This data is roughly normal, with a mean Yearly precipitation of approximately 1800mn. per year

