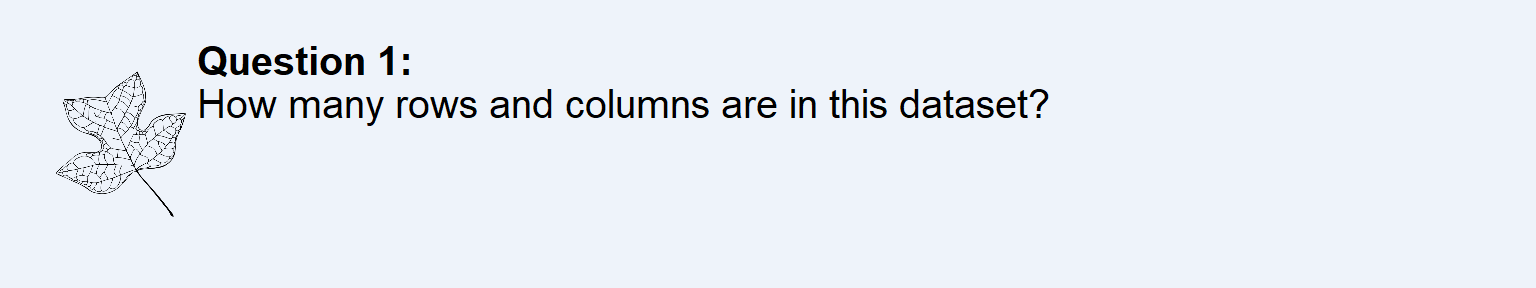
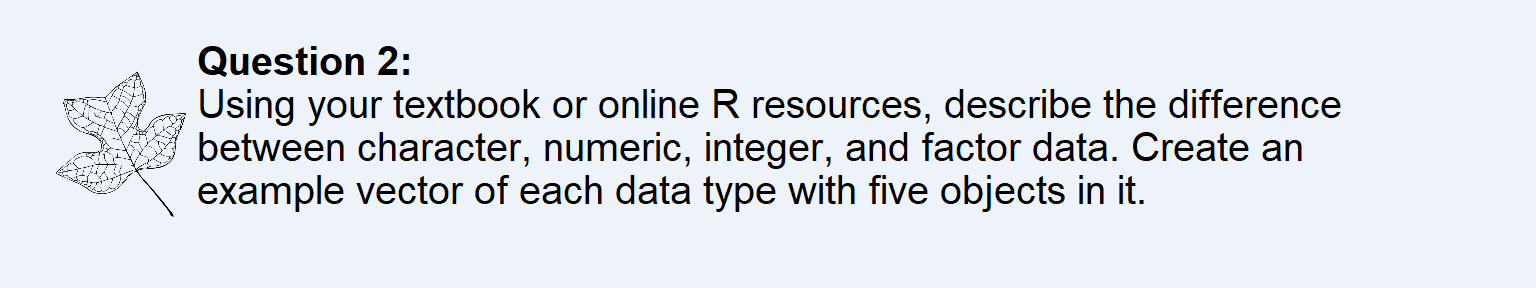
Jillian Johns

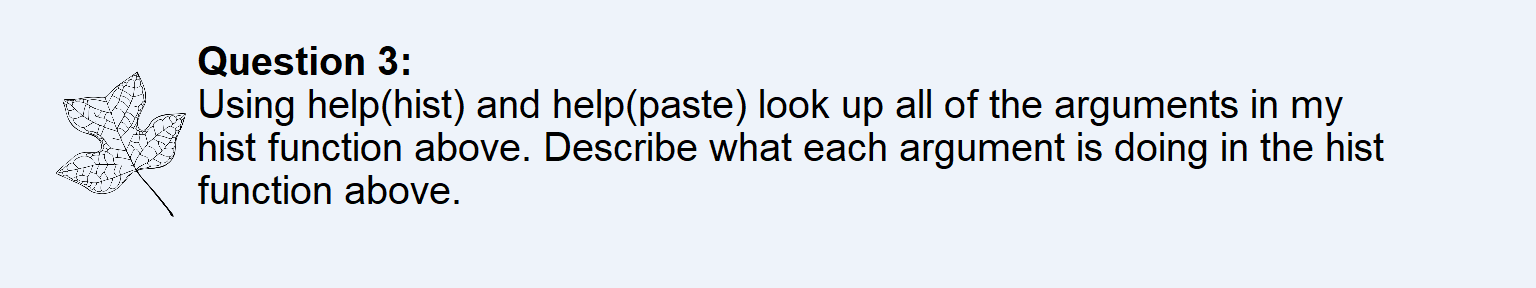
Activity #2



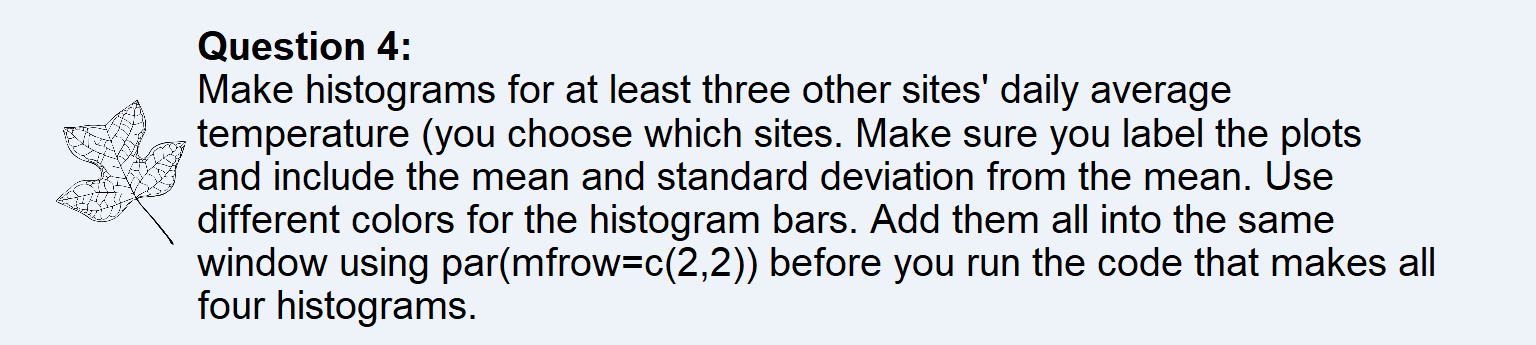
There are nine columns and 157,849 rows in the dataset.

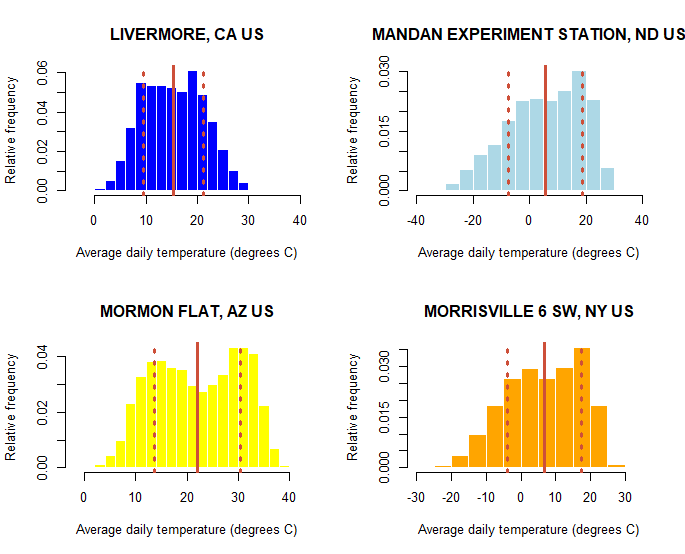


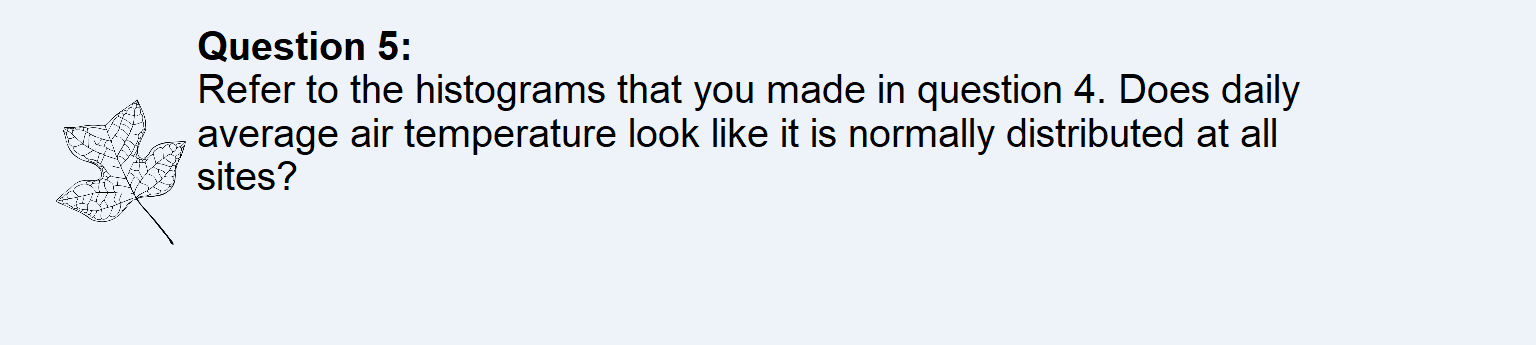
Character data is when letters are accepted as they are in a string of letters that each are unique values. Numeric data has numbers of any value, while integers are whole numbers. Factors are integer data with character values set to them.



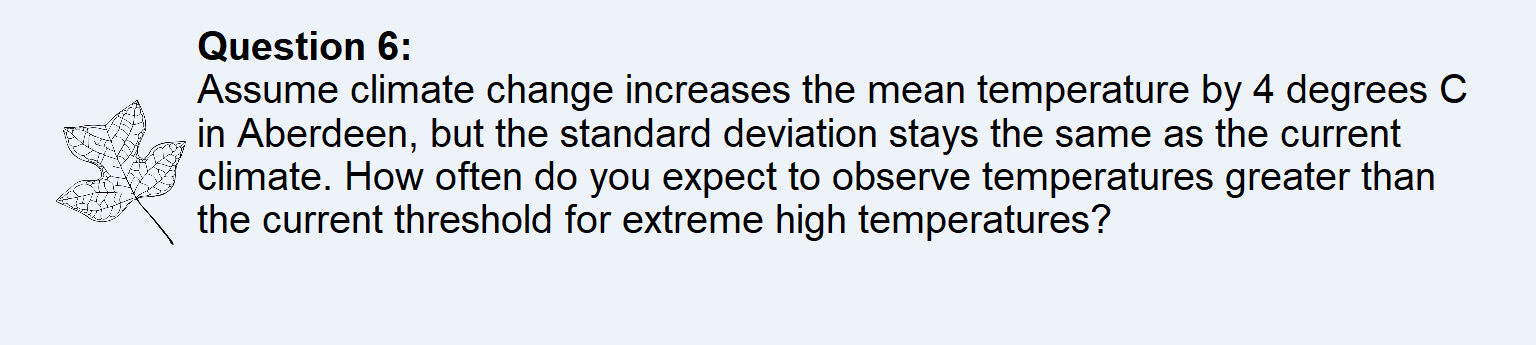
Arguments included in the histogram above are that we are ignoring NAs, going by list of the names, function is the mean and the na.rm is true, meaning that the function will skip over NA values. When freq=false, this means that the histogram goes by probability densities rather than by frequency. The xlab and ylab are the axis labels and main is the histogram label. The arguments for the border of boxes and color are white and grey.



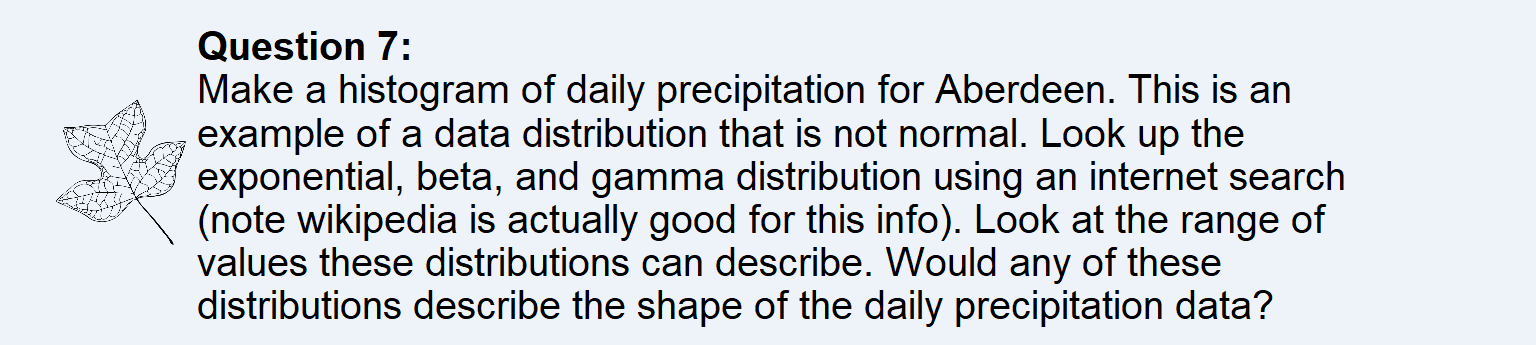




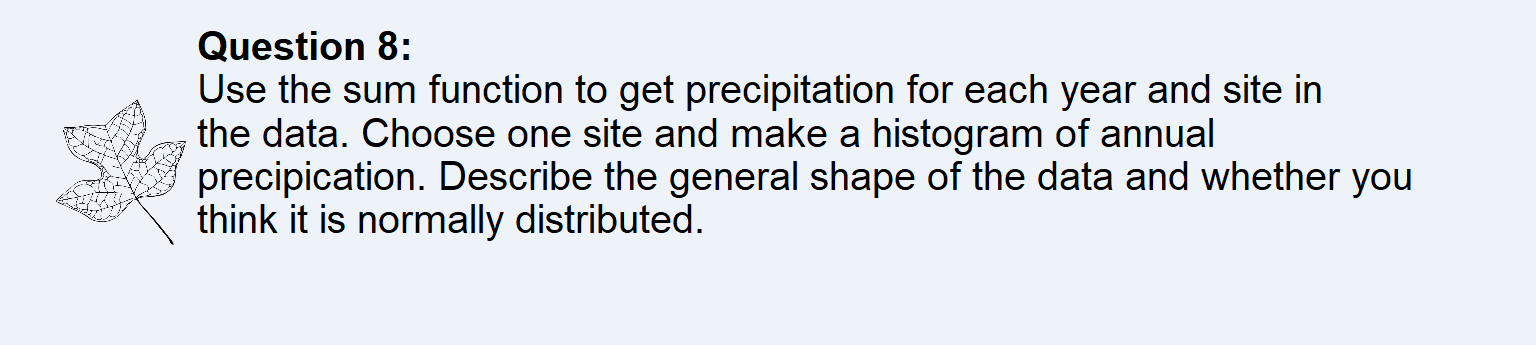
Daily average air temperatures do not look normally distributed at all sites, as at the Mandan experiment station and the morrisville 6, these both look skewed, and the Mormon Flat data looks more bimodal.



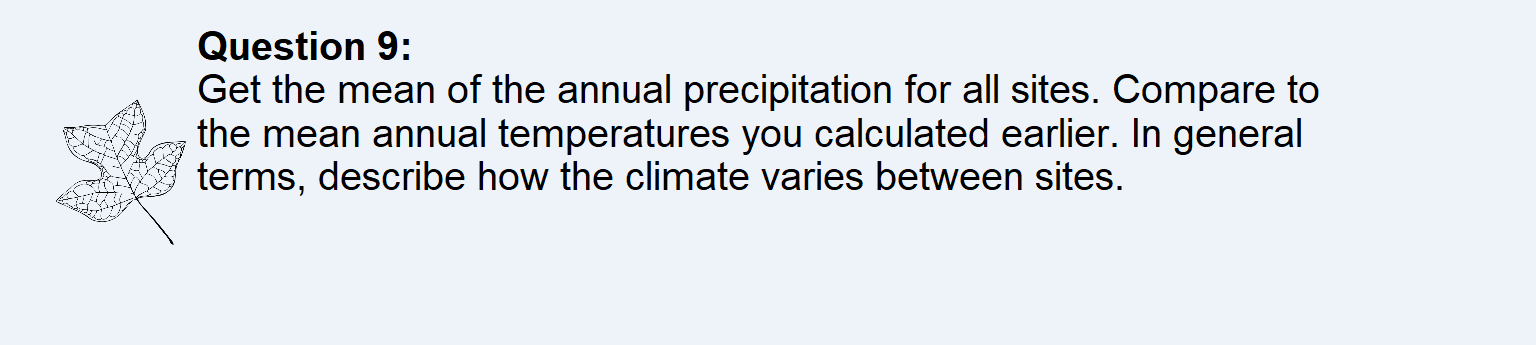
We would expect that 0.001647848% of the time, the temperatures would be greater than the current threshold for extreme high temperatures.



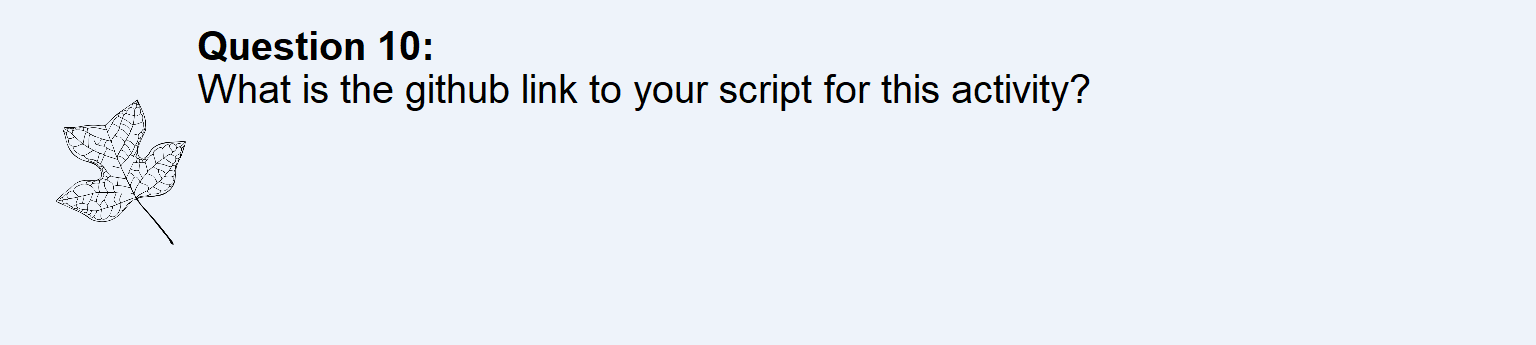
The exponential distribution seems to describe the daily precipitation data mostly, and the beta and gamma do not fit the shape. There are far more days with relatively lower precipitation rates than higher, showing a curve similar to exponential.



On R



On R



https://github.com/jillianjohns3/GEOG331.git