

Climate Change Strike

Jillian Morrison

9/18/2019

1. Use `read.table()` to read the temperature change dataset from NASA's webpage: https://data.giss.nasa.gov/gistemp/graphs/graph_data/Global_Mean_Estimates_based_on_Land_and_Ocean_Data/graph.txt

```
library(dplyr)
Temp = read.table("https://data.giss.nasa.gov/gistemp/graphs/graph_data/Global_Mean_Estimates_based_on_Land_and_Ocean_Data/graph.txt",
                  ,skip=2,header=T, sep=" ", fill=T)
#skip - skips the first 2 rows of the txt file since they are meaningless
#header uses the 3rd row(since you skipped the first 2 rows) as the column names
#Table is saved as Temp

Temp2= Temp[-c(1), ]
#Removes the first row since it is meaningless.. it has ----- .
#It then saves the new dataframe to Temp2

Temp3 = Temp%>%mutate(Year=as.numeric(as.character(Year)))
#mutate changes the column Year from Factor to numeric
#It then saves the new dataset with facor as numeric as Temp3
```

2. Try out the `round()` function in R. It takes the parameters `round(x, digits)` where `x` is the vector of numbers to be rounded and positive digits refer to rounding in decimal places and negative digits refer to rounding in whole numbers. Specifically, do:
 - a. round the Year variable from the dataset to the nearest hundredth (0.01), i.e. `digits = 2`
 - b. round the Year variable from the dataset to the nearest ten (10) , i.e. `digits= -1`
 - c. now try subtracting 4.5 from the Year and then round to the nearest ten. What is the difference between this result and the result in b?
3. Add a column called **Decade** to the dataset. **Decade** will be defined as the Year minus 4.5 and then rounded to the nearest tens. Name this new dataframe **Temp_dec**. (Hint: Use the `mutate()` function along with the `round()` function. Also use `%>%` from `{dplyr}`)
4. Summarize the temperature change (referred to as **No_Smoothing**) by decades. Count the number of observations (`n()`), find the average temperature `mean()` and the standard deviation of the temperature (`sd()`) per decade. Hint: Use `%>%` from `{dplyr}` along with `summarize()` and `group_by()`
 - a. Summarize temperature by **Decade**.
 - b. Plot a scatterplot of decade on x axis against the everage temperature on the y axis. (Hint: Use the dataset created in part a above. Be sure to add labels to your plot.)
5. Make your observations in context to the problem. Visit <https://climate.nasa.gov/vital-signs/global-temperature/> for more information on what the variables mean and for more information.