Omar

900222400

January 10, 2022

DSCI 1411

**Assignment 3**

1. **Scientific Notation:**
   1. Write the following numbers in scientific notations: Two billion and 1/2 million

**In scientific notation: 2.0005e9**

* 1. Write the following numbers without using Scientific notation: 2.143e-2 and 1.0e6

**Without scientific notation: 0.02143**

1. **The following is a line of R commands and the corresponding output for an object named “x” in an R Workspace:**

> class(x); dim(x); colMeans(x); length(x); mean(x);

[1] ”data.frame”

[1] 15 2

height weight

65.0000 136.7333

[1] 2

[1] NA

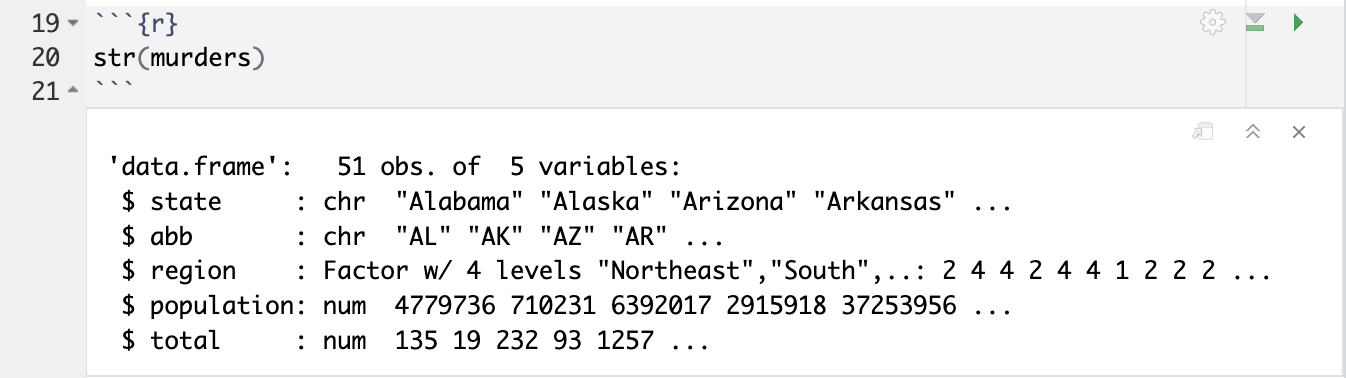
Warning message: 1 In mean.default(x) : argument is not numeric or logical: returning NA

**Using the above information, answer the following questions:**

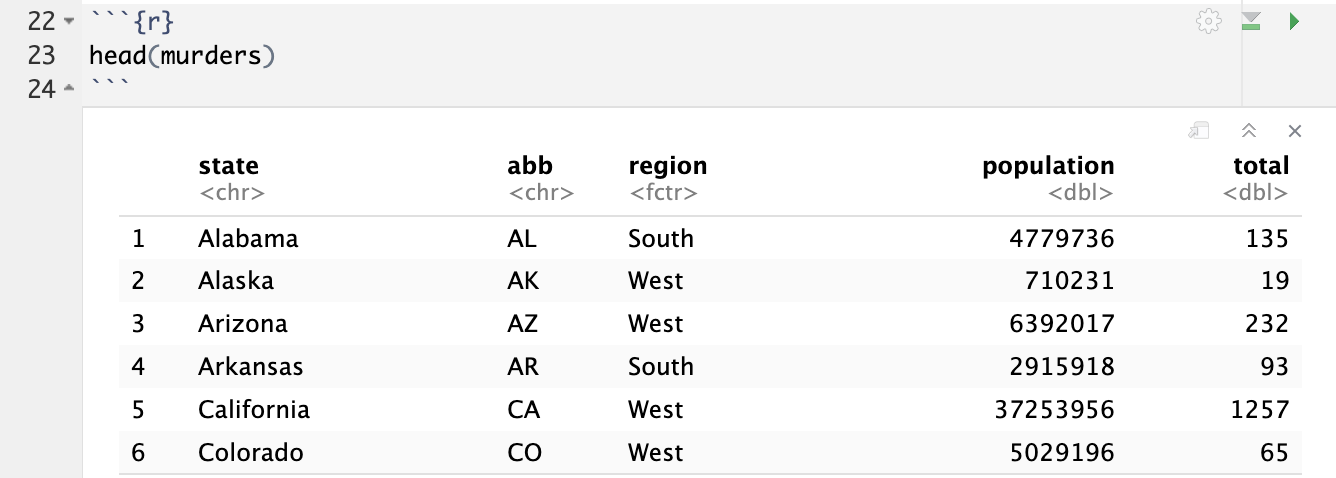
* 1. How many variables are there in object x? What are they? **In object x, there are two variables, which are height and weight.**
  2. How many observations are there in object x? **In object x, there are 15 observations.**
  3. What is the mean of x[,3]? **There is no mean value of x[,3] because there is no third column, there are only two columns. So, the output of mean(x[,3]) for this data set would be: Error in m[, 3] : subscript out of bounds**
  4. Compute the sum of each column of x. **The sum of each column of x would be equal to the mean of each column multiplied by the number of rows. Therefore, for the height column, it would be 65 times 15, which is equal to 975 and for the weight column, it would be 136.7333 times 15 which is equal to 2050.9995.**

1. **Load the US murders dataset using: install.packages(“dslabs”); library(dslabs); data(murders), then answer the following questions:**
   1. Examine the data using str(murders) and head(murders)

**str(murders): *Compactly displayed the internal structure of the data set***

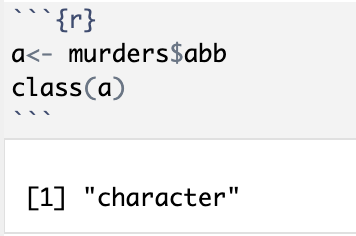


**head(murders): *Displayed the first 6 rows of the data set***



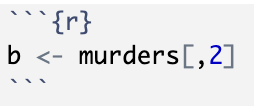
* 1. Use the accessor $ to extract the state abbreviations and assign them to an object named “a.” What is the class of this object?

**Here I assigned the extraction of State Abbreviations using the $ Accessor to an object called “a.” The class of this object is a character.**

****

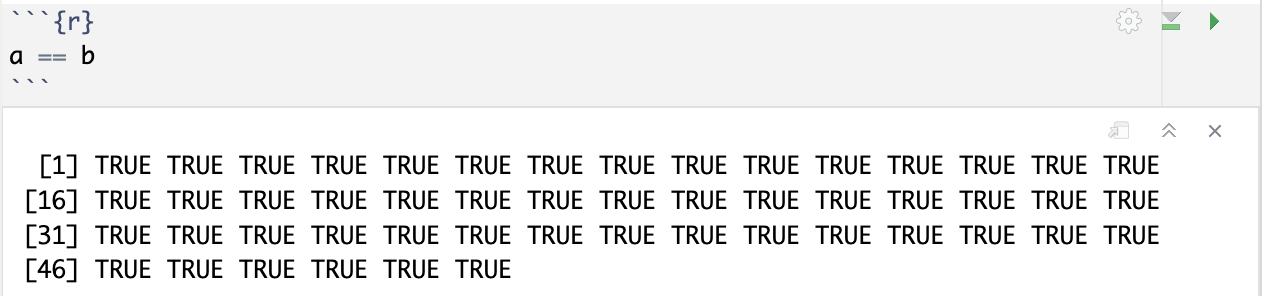
* 1. Now use the square brackets to extract the state abbreviations and assign them to an object called “b.”

**Here I assigned the extraction of State Abbreviations using the Square Brackets to an object called “b”:**

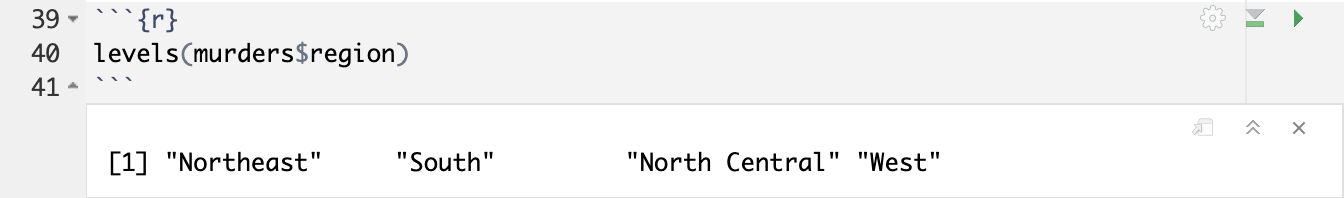


* 1. Check the equality of the two vectors using ==.

**Here I checked the equality of the two vectors using the == sign:**



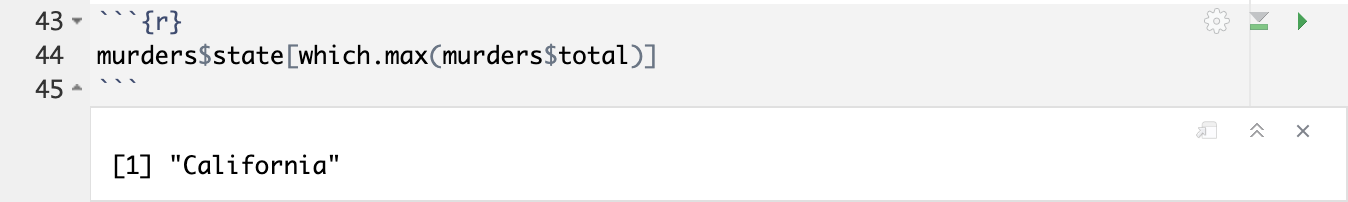
* 1. We saw that the column named region stores a factor. You can corroborate this by typing: class(murders$region). With one line of code, use the function levels and length to determine the number of regions defined by this dataset. **The number of regions defined by this dataset is 4 regions. These 4 regions are North East, North Central, West, and South.**

****

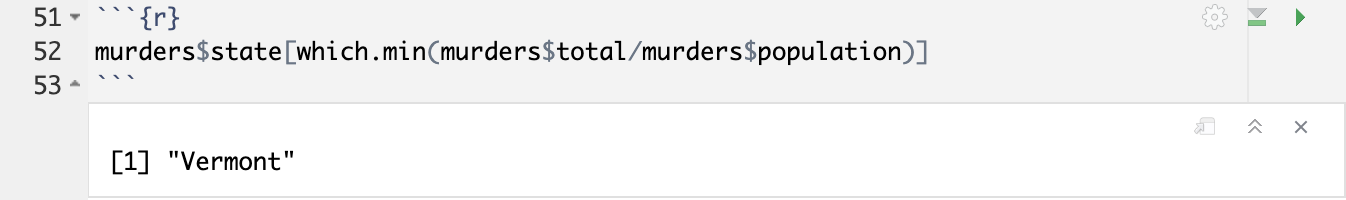
* 1. Which state has the highest population size? **The state with the highest population size is California.**

****

* 1. Which state has the highest total murders? **The state with the highest total number of murders is California.**

****

* 1. Which state has the lowest murder rate? **The state with the lowest murder rate is Vermont.**



1. **Verify this formula using sum().**

