R for Data Science Assignment 1

Faculty: Dr. Nachiket Tapas
Branch: CSE (AI)/ CSE (DS)
Submission: 6-April-2024

1. Create and store a sequence of values from 5 to -11 that progresses in steps of 0.3.
2. Overwrite the object from (1.) using the same sequence with the order reversed.
3. Repeat the vector c(-1,3,-5,7,-9) twice, with each element repeated 10 times, and store the result. Display the result sorted from largest to smallest.
4. Create and store a vector that contains, in any configuration, the following:
    i. A sequence of integers from 6 to 12 (inclusive)
    ii. A threefold repetition of the value 5.3
    iii. The number -3
    iv. A sequence of nine values starting at 102 and ending at the number that is the total length of the vector created in (3.)
5. Confirm that the length of the vector created in (4.) is 20.
6. Create and store a vector that contains the following, in this order:
    i. A sequence of length 5 from 3 to 6 (inclusive)
    ii. A twofold repetition of the vector c(2,-5.1,-33)
    iii. The value 7/42 + 2
7. Extract the first and last elements of your vector from (6.), storing them as a new object.
8. Store as a third object the values returned by omitting the first and last values of your vector from (6.).
9. Use only (7.) and (8.) to reconstruct (6.).
10. Overwrite (6.) with the same values sorted from smallest to largest.
11. Use the colon operator as an index vector to reverse the order of (10.), and confirm this is identical to using sort on (10.) with decreasing=TRUE.
12. Create a vector from (8.) that repeats the third element of (8.) three times, the sixth element four times, and the last element once.
13. Create a new vector as a copy of (10.) by assigning (10.) as is to a newly named object. Using this new copy of (10.), overwrite the first, the fifth to the seventh (inclusive), and the last element with the values 99 to 95 (inclusive), respectively.
14. Convert the vector c(2,0.5,1,2,0.5,1,2,0.5,1) to a vector of only 1s, using a vector of length 3.
15. The conversion from a temperature measurement in degrees Fahrenheit F to Celsius C is performed using the following equation:
C = 5/9 (F - 32)
Use vector-oriented behaviour in R to convert the temperatures 45, 77, 20, 19, 101, 120, and 212 in degrees Fahrenheit to degrees Celsius.
16. Use the vector c(2,4,6) and the vector c(1,2) in conjunction with rep and * to produce the vector c(2,4,6,4,8,12).

17. Overwrite the middle four elements of the resulting vector from (16.) with the two recycled values -0.1 and -100, in that order.
18. **Reversed Sequence Creation:**
Generate and store a sequence of values from 10 to -20 in steps of 0.5. How can you reverse this sequence efficiently in R without recreating it manually?
19. **Vector Repetition and Order Sorting:**
For the vector c(-2, 4, -6, 8, -10), describe how to repeat the vector itself three times and each of its elements 5 times. Following this, how would you sort the resulting vector from smallest to largest?
20. **Composite Vector Construction:**
Assemble a vector that includes:
i. A sequence of integers from 15 to 25 (inclusive)
ii. A twofold repetition of the number 4.2
iii. The number -5
iv. A sequence of twelve values starting at 200 and ending at the number which is the total length of the vector created in question 2. How would you ensure all components are accurately combined?
21. **Length Confirmation of a Complex Vector:**
After creating the complex vector described above, what function would you use to confirm its length is 25? What result from this function would indicate success?
22. **Diverse Vector Composition:**
Detail the steps to compile a vector containing, in order:
i. A sequence of length 7 from 2 to 8 (inclusive)
ii. A threefold repetition of the vector c(3,-4.2,-50)
iii. The value 14/84 + 3
How do you maintain the specified order while combining these elements?
23. **Extraction and Reconstruction:**
Extract the first and third elements from the vector created in question 5, storing them in a new object. How would you use this new object along with the original vector (minus its first and third elements) to reconstruct the original sequence?
24. **Middle Element Replacement:**
From the resulting vector in question 5, replace the middle three elements with the sequence -0.5, -200, -0.5. What strategies in R allow for such precise element replacement?
25. **Conditional Vector Transformation:**
Convert the vector c(3,1,2,3,1,2,3,1,2) into a vector of only 2s, using a vector of length 2. Which vectorized operations in R would facilitate this conversion?
26. **Temperature Conversion Vectorized:**
Convert the temperature readings 32, 68, 14, 0, 113, 104, and 198 degrees Fahrenheit to Celsius using the formula C = 5/9 * (F - 32). How can you apply this conversion in a vectorized manner in R?

27. **Vector Multiplication and Element Replacement:**
Use the vector c(3,5,7) and the vector c(2,3) with the rep function and multiplication to produce the vector c(6,15,21,9,15,21). How do you then replace the middle four elements with two alternately repeated values of -1 and -150?

*****Finished*****

R for Data Science Assignment 2

Faculty: Dr. Nachiket Tapas
Branch: CSE (AI)/ CSE (DS)
Submission: 06-February-2025

1. Construct and store a 4 x 2 matrix that's filled row-wise with the values 4.3, 3.1, 8.2, 8.2, 3.2, 0.9, 1.6, and 6.5, in that order.

2. Confirm the dimensions of the matrix from (1.) are 3 x 2 if you remove any one row.

3. Overwrite the second column of the matrix from (1.) with that same column sorted from smallest to largest.

4. What does R return if you delete the fourth row and the first column from (3.)? Use matrix to ensure the result is a single-column matrix, rather than a vector.

5. Store the bottom four elements of (3.) as a new 2 x 2 matrix.

6. Overwrite, in this order, the elements of (3.) at positions (4;2), (1;2), (4;1), and (1;1) with -1/2 of the two values on the diagonal of (e).

7. Calculate the following:

$$\frac{2}{7}\left(\begin{bmatrix}1 & 2\\ 2 & 4\\ 7 & 6\end{bmatrix} - \begin{bmatrix}10 & 20\\ 30 & 40\\ 50 & 60\end{bmatrix}\right)$$

8. Store these two matrices:

$$A = \begin{bmatrix}1\\ 2\\ 7\end{bmatrix}, and\ B = \begin{bmatrix}3\\ 4\\ 8\end{bmatrix}$$

Which of the following multiplications are possible? For those that are, compute the result.

i. $A \cdot B$
ii. $A^T \cdot B$
iii. $B^T \cdot (A \cdot A^T)$
iv. $(A \cdot A^T) \cdot B^T$
v. $[(B \cdot B^T) + (A \cdot A^T) - 100I_3]^{-1}$

9. For

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

confirm that $A^{-1} - A - I_4$ provides a 4 x 4 matrix of zeros.

10. Create and store a three-dimensional array with six layers of a 4 x 2 matrix, filled with a decreasing sequence of values between 4.8 and 0.1 of the appropriate length.

11. Extract and store as a new object the fourth- and first-row elements, in that order, of the second column only of all layers of (10.).

12. Use a fourfold repetition of the second row of the matrix formed in (11.) to fill a new array of dimensions 2 x 2 x 2 x 3.

13. Create a new array comprised of the results of deleting the sixth layer of (10.).

14. Overwrite the second and fourth row elements of the second column of layers 1, 3, and 5 of (13.) with -99.

15. Construct a 5 x 3 matrix filled column-wise with a sequence of numbers from 1 to 15. Verify the matrix is correctly populated.

16. If you add a new row to the matrix created in question (15.), confirm its new dimensions.

17. Sort the first column of the matrix from question (15.) in descending order. Keep other columns unchanged.

18. What would the result be if you remove the second row and the third column from the sorted matrix in question (17.)? Ensure the result remains a matrix.

19. From the matrix in question (17.), create a new 2 x 2 matrix using the top four elements of the last column.

20. Replace the elements at positions (2,1), (2,3), (5,1), and (5,3) in the matrix from question (17.) with the average of the four corner elements of that matrix.

21. Calculate the below operation.

$$3\left(\begin{bmatrix}2 & 5\\ 3 & 7\end{bmatrix} - \begin{bmatrix}5 & 10\\ 15 & 20\end{bmatrix}\right)$$

22. Store these two matrices:

$$A = \begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}, and\ B = \begin{bmatrix}5\\ 6\end{bmatrix}$$

which of the following multiplications are valid? For valid cases, compute the results.

i. C . D

ii. C^T . D

iii. D^T . (C . C^T)

23. For

$$B = \begin{bmatrix}1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -2\end{bmatrix}$$

confirm that B^{-1} - B – I_3 provides a 3 x 3 matrix of zeros.

24. Construct a three-dimensional array with four layers, each a 3 x 3 matrix filled with random numbers between 1 and 9. Then, extract the elements of the first row of the third column across all layers and store them as a new vector.

*****Finished*****

R for Data Science Assignment 3

Faculty: Dr. Nachiket Tapas
Branch: CSE (AI)/ CSE (DS)

1. Write an R script to create a logical vector of length 12 that alternates between TRUE and FALSE.

2. Write a R script that takes a numeric vector as input and returns a logical vector indicating which elements are negative.

3. Given two numeric vectors of equal length, write R code to check if they are element-wise equal and then output the indices where they differ using the which() function.

4. Create a 3×4 matrix from a given logical vector and display the result. Then, create another 3×4 numeric matrix and perform an element-wise comparison between them.

5. Demonstrate vector recycling by creating one vector of length 4 and another of length 10, then perform an addition operation. Print the resulting vector.

6. Write R code that calculates the sum of all TRUE values in a given logical vector (using their numeric property where TRUE = 1 and FALSE = 0).

7. Given a numeric vector, write R code to extract all elements that are greater than 10 using logical subsetting.

8. For the string "Hello World! Welcome to R programming", write R code to count the number of characters (using nchar()) and extract the substring "R programming".

9. Write R code that concatenates the strings "Data", "Science", and "R" into a single string with hyphens (-) as separators.

10. Write an R script that demonstrates the difference between cat() and paste() by printing the same set of words using both functions with a custom separator.

11. Given the string "apple, apple, and apple", write R code that uses sub() to replace only the first occurrence of "apple" with "orange", and gsub() to replace all occurrences.

12. Convert the character vector c("male", "female", "male", "female") into a factor. Display the factor and its levels.

13. Create a factor from the vector c("Jan", "Mar", "Feb", "Apr", "May", "Dec", "Nov") and order the factor to reflect the natural calendar order (January to December). Display the ordered factor.

14. Given a factor vector of genders (e.g., gender <- factor(c("male", "female", "female", "male", "female"))), write R code to extract only the entries corresponding to "female".

15. Write a function that takes a numeric vector and bins the data into three categories: "Low", "Medium", and "High". Use the cut() function and return the resulting factor.

16. Create a logical vector and demonstrate the use of the any() and all() functions. Explain, via comments, a scenario where any() returns TRUE but all() returns FALSE.

17. Given two logical vectors, write R code to perform element-wise logical AND, OR, and NOT operations. Print the results.

18. Given a numeric vector, write R code to create a new vector where all values less than the mean of the vector are replaced with NA.

19. Write an R script to compare two strings alphabetically using relational operators. Explain via comments how R determines the order (considering case sensitivity).

20. Write an R script that creates a string containing a newline, tab, and a backslash using escape sequences. Print the string to the console using cat().

*****Finished*****

R for Data Science Assignment 4

Faculty: Dr. Nachiket Tapas
Branch: CSE (AI)/ CSE (DS)

1. Create a list containing a number, a string, and a logical value.
2. How do you find the length of a list named my_list?
3. Extract the first element of a list using single brackets [ ].
4. Extract the second element from a list directly using double brackets [[ ]].
5. Name the elements of a list explicitly.
6. Access a named element of a list using the $ operator.
7. Create a simple data frame with two columns: Name and Age.
8. Access the age of the second person in a data frame.
9. Find the number of rows in a given data frame.
10. Add a new row to an existing data frame.
11. Create a nested list where one element is itself a list containing a numeric vector and a character vector.
12. Extract the second element from the nested list created above.
13. Write code to convert a character column in a data frame to a factor.
14. Create a logical subset of a data frame to include only rows where age > 20.
15. Add a new column to a data frame representing the age in months.
16. Demonstrate the difference between single bracket [ ] slicing and double bracket [[ ]] referencing with a suitable example.
17. How would you extract multiple rows and columns simultaneously from a data frame?
18. Add a new column using the cbind() function to a data frame.
19. How do you subset a data frame to exclude a particular column?
20. Extract all rows from a data frame where a character column has a specific value.
21. Write a function to merge two lists into one nested list without losing the original list structure.
22. Create a data frame that recycles shorter vectors. Demonstrate and explain the behavior.
23. Write R code to create a data frame from a list containing multiple vectors of unequal length, ensuring no data recycling occurs.
24. Create a list containing a matrix, a logical vector, and a string. Then, extract the second element of the logical vector.
25. Write R code that dynamically adds named elements to an existing list based on user input.
26. Create a function that accepts a data frame and returns a subset with only numeric columns.
27. Create a data frame and write code to reorder its columns alphabetically by column names.
28. Demonstrate how to subset a nested list to extract a deeply nested numeric value.
29. Write a function that takes a data frame as input and adds a factor-type column derived from an existing numeric column.
30. Using logical vectors, subset a data frame to extract rows based on multiple conditions across different columns.

*****Finished*****

R for Data Science Assignment 5

Faculty: Dr. Nachiket Tapas
Branch: CSE (AI)/ CSE (DS)

1. Write R code to install and load the ggplot2 package.
2. Use qplot() to create a simple scatter plot with vectors x and y.
3. Save your plot as a JPEG file using R.
4. Create a line plot connecting points using ggplot and geom_line().
5. Add titles and axis labels to your plot using ggplot2.
6. Demonstrate how to use geom_point() to modify size and shape of points.
7. Explain how to list files in the current working directory.
8. Write code to set a working directory to a specified folder.
9. Read a CSV file from your local system into an R data frame.
10. Read a CSV file from an online source into an R data frame.
11. Create a ggplot scatter plot, and differentiate points by color based on a categorical variable.
12. Demonstrate how to save a plot as a PDF file.
13. Add horizontal and vertical reference lines to a plot using ggplot2.
14. Create a ggplot plot with customized colors for different categories.
15. Write R code to add annotations (text and arrows) to a ggplot plot.
16. Create a ggplot plot with explicitly defined x and y axis limits.
17. Use logical conditions to highlight subsets of points in different colors.
18. Demonstrate adding a legend manually using ggplot2.
19. Create a plot with different line types and widths.
20. Write a code snippet using geom_segment() to add customized line segments to your plot.
21. Write a function to dynamically read and plot data from a user-selected file.
22. Demonstrate the combined usage of geom_point, geom_line, and geom_hline in a single ggplot.
23. Create a customized legend using manual scales and guides in ggplot2.
24. Write R code that handles missing values while reading external data into a data frame.
25. Construct a ggplot that categorizes points into multiple groups based on two numeric conditions.
26. Create a complex ggplot visualization that includes multiple geoms and a theme customization.
27. Write code to plot data points with condition-based shapes and colors using ggplot.
28. Develop an R function that saves plots in both JPEG and PDF formats automatically.
29. Write code to implement a custom plot theme in ggplot2, adjusting fonts, backgrounds, and grid lines.
30. Demonstrate using ifelse() in R to create a new variable and plot this conditional variable using ggplot2.

*****Finished*****