Department of Computer Science and Engineering

Osmania University(A)

B.E (CSE) – IV Semester

Subject Name: R Workshop

Prepared by

Prof. of Practice- Krishnaiah Dayyala

confidential

1. Write a R program to read the following 15 numbers from keyboard and arrange them in Ascending and descending order and print the order. Also find the mean, Standard deviation, mode, range of the numbers using the R program in built statistical functions. The numbers are given below

   100, 45, 23, 125,77,

   17, 19,88, 99, 3

   65, 15, 36, 256, 500

2. Write a program to build a matrix of Department with details of Number of students, male students, female students, average age . Use the matrix method fund the total strength of the college for CSE, ECE, EEE, Biomedical, CIVIL and Mechanical engg departs of OU College of engg. Note take eh approx. data for each department. Print each dept data and overall college data. What tie h standard deviation of the overall age of the students and mean of the students of the college

3. In TS EAMCET 2024, 150, 500 students appeared for the entrance exam. Out of them, 40% girls and 60% are boys. Out of this 70% students qualified the exam. Of the qualified, 70% are girls and 30% are boys. Among the qualified girls, 50% girls are from Telugu medium, 30% are from English medium and remaining are Urdu medium. Among the qualified boys, 70% are Telugu, 20% from English and the balance from other languages. Considering this as official data, your task is to prepare all possible visualization charts for results for newspaper printing for tomorrow. The charts can be pie, linear, bar charts and histogram.

4. Write a R program to print " R programming" 5 times using repeat loops

5. Write a program to find out the n natural numbers. The n is entered by the keyboard. Use while loop for creating this program.

6. Every Saturday, at the same time, an individual stands by the side of a road and tallies the number of cars going by within a 120-minute window. Based on previous knowledge, she believes that the mean number of cars going by during this time is exactly 107. Let X represent the appropriate Poisson random variable of the number of cars passing her position in each Saturday session.

   a. What is the probability that more than 100 cars pass her on any given Saturday?
   b. Determine the probability that no cars pass.

7. Write a program to calculate a factorial of given number. Read the number from keyboard and reject the number if it is negative, zero till the valid + number is entered. Create a function for factorial and call the function in the main program.

8. Write a program to compute the Fibonacci sequence for a given number. Take the number input from keyboard.

9. Write a program to compute the cumulative sum of the column for a given matrix

   2, 4, 3

   3, 7, 9

   1, 8, 10

10. In the vector x <- c(2,20,3,12,8,11,6), find out the odd and even numbers using the R program.
11. Using the **airquality** dataset given in R Studio, find out the mean, std deviation of temp and draw the line graph of between wind and temperature.

12. Using the data set " Iris" given in the R Studio which has 150 observations, 5 variables representing the length and width of the sepal and petal and species of 150 flowers. Length and width of the sepal and petal are numeric variables and species is a factor with levels. Find out the following using R program
    a. Min and max of sepal length
    b. Range of sepal length
    c. IQR
    d. Sd
    e. Mean
    f. Median
13. Find the following for the given matrixes

    MATRIX A:

    10 15 22

    11 23 55

    1 7 10

    MATRIX B:

    5 11 13

    2 55 8

    1 0 19

Calculate and print in the matrix format the following

C= A+B

C= A-B

C= A x B

C= A ÷ B

14. Write a program to calculate the 90th percentile of Chi-Square distribution whose degree of freedom is 5.
15. Write a program to compute the probability of turning up of a three successful " 5" while tossed the die for 7 times.
16. A car distribution company sells multiple brands of cars having different engine CC, Mileage and price.

Car Model	Engine CC	Mileage (Kmpl)	Prie (Lakhs)
Honda	1600	30	15
Toyota	1800	15	18
Kia	1346	10	14
BMW	1020	7	22
Merc	2500	5	25
Maruthi	1200	20	8
Hyundai	1400	18	16

Create a relationship model using the above data for the sales person to decide the model based on the C, KMPL and price using linear relation ship model. Note use Linear regress model. Find also Covariance and Correlation between Price CC, Correlation between mileage and CC

17. Use the data set "InsectSprays" in R studio data sets. Draw the pie chart and bar chart and histogram of the spray count for each spare type. The head information is below

Recall the built-in InsectSprays data frame, containing counts of insects on various agricultural units treated with one of six sprays.

a. Produce a histogram of the counts of insects using base R graphics.
b. Obtain the total number of insects found according to each spray. Then, use base R graphics to produce a vertical barplot and a pie chart of these totals, labeling each plot appropriately.
c. Use ggplot2 functionality to generate side-by-side boxplots of the counts of insects according to each spray type and include appropriate axis labels and a title.

```
head(InsectSprays)
  count spray
1    10     A
2     7     A
3    20     A
```

4 14 A
5 14 A
6 12 A

18. Following is the archalogical data of information for different sites. Find aou the following using R – statistical tools . the data is below

Consider the following data:

Site I	Site II	Site III	Site IV
93	85	100	96
120	45	75	58
65	80	65	95
105	28	40	90
115	75	73	65
82	70	65	80
99	65	50	85
87	55	30	95
100	50	45	82
90	40	50	
78		45	
95		55	
93			
88			
110			

These figures provide the depths (in centimeters) at which important archaeological finds were made at four sites in New Mexico (see Woosley and Mcintyre, 1996). Store these data in your R workspace, with one vector containing depth and the other vector containing the site of each observation.

a. Produce side-by-side boxplots of the depths split by group, and use additional points to mark the locations of the sample means.
b. Assuming independence, execute diagnostic checks for normality and equality of variances.
c. Perform and conclude a one-way ANOVA test for evidence of a difference between the means.

19. Create a data from named cars , contains the details of the cars purchased in year 2015 and 2020 from Jan to April. The data is shown below. Write a program to read the below csv file into data frame. Compute the 95% confidence level approximation of the variance between the mean car purchase between years 2025 and 2020

	Cars purchased	
Month	**2015**	**2020**
Jan	70	75
Feb	80	85
Mar	90	95

Apr	90	95

20. Following is the Mobile sale data for three different manufacturers. Please read this table data using R program and calculate the total mobile sales in each year and all years.

	sales data in lakhs			
	Iphone	Samsung	Vivo	Jio
2021	12	15	23	45
2022	17	14	22	56
2023	18	12	17	60
2024	22	17	18	58

What is the variance in mobile sales across all the years.