1] a) Write steps to Install R and R studio. [10M]

b) Create vectors v: elements ranging from 5 to 13, t: storing days of a week and perform following operations on vector elements. [10M]

a) Print the 2,3,and 6th element from vector t

b) Sort the elements of vector v(3,8,4,5,0,11, -9, 304)

c) Perform different operations on vectors v(3,8,4,5,0,11)

v2 (4,11,0,8,1,2): add, subtract, multiply ,division

2] a) Explain the following R objects: dataframe, List, vector [10M]

b) Create Data Frame "my_data_frame" with following information: [10M]

animal <- sheep, pig

year <- 2019:2021,

weight <- 110, 120, 140, NA, 300, 800

height <- 2.2, 2.4, 2.7, 2, 2.1, 2.3

condition <- "excellent", "good", NA, "excellent", "good", "average"

Perform following operations on Data frame.

1. Print class and structure of my_data_frame

2. Get the summary statistics for each variable of my_data_frame

3. Add the new observation: animal = "pig", year = 2018, weight = 200, height = 1.9, condition = "excellent"

4. Print the following output using subsetting : i) 110 ii) 2019 2020 2021 2019 2020 2021

3] a) Explain Descriptive statistics in detail. [10M]

b) Write R-code to perform Descriptive Data Analysis on 'iris' Dataset to compute the following:

1. Display the structure of dataset and first six observations

2. Find minimum, maximum and range of Sepal.Length

3. Find Mean, Median and Mode on Sepal.Length

4. Find First and third quartile and Interquartile range

5. Find Standard deviation and variance [10M]

4] a) How to import data from different sources (.csv file, txt file, excel file) in R? [10M]

b) Create a DataFrame "stats" with following information: [10M]

Player = ('A', 'B', 'C', 'D', 'A', 'A')

Runs = (100, 200, 408, 19, 56, 100)

Wickets = (17, 20, NA, 5, 2, 17)

Perform the following operations using the functions of 'dplyr 'package:

1. **Fetch the data of players who scored more than 100 runs**
2. **Remove duplicate rows from data frame**
3. **Arrange data based on runs low to high**
4. **Display the wickets taken by each player**
5. **Change the column name "runs" to "runs_scored" in stats data frame.**
6. **Find total no of runs scored.**

5] a) Explain different functions of dplyr package. [10M]

b) Perform data manipulation operations on "iris" dataset using 'dplyr' package: [10M]

1. Return 5 random rows from the dataset

2. Find the frequency distribution of Species in iris table

3. Select all columns from Sepal.Length to Petal.Length

4. Hide the column Sepal.Length

5. Select the first 3 rows with Species as setosa

6. Create a column "Greater.Half" which stores TRUE if given condition is true:

Sepal.Width > 0.5 * Sepal.Length