

School of Business

**Master of Business Administration MBA Dual Specialization
Semester End Examination - Jul 2024**

**Duration : 180 Minutes
Max Marks : 100**

Sem IV - MBBA6011 - Business Analytics Using R

General Instructions

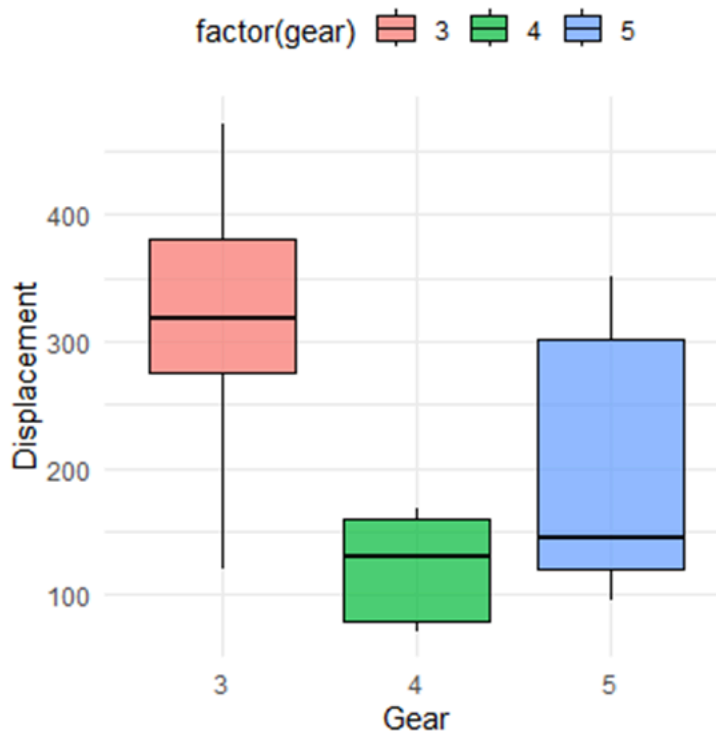
Answer to the specific question asked

Draw neat, labelled diagrams wherever necessary

Approved data hand books are allowed subject to verification by the Invigilator

- 1) Write a R program to create three vectors numeric data, character data and logical data. Display the content of the vectors. K3(6)
- 2) Explain the differences between data frame and a matrix datatypes in R language K4(8)
- 3) Analyze the application of recursive function in R programming. Write R code for user-defined function to find the factorial of a number 6. K4(8)
- 4) Explain the difference between the chi-square test for independence and the chi-square test for goodness-of-fit. K4(4)
- 5) Apply the concept of ANOVA to explain the given image, showing graphical representation of one way ANOVA result. K3(9)

One-Way ANOVA



- 6) Suppose that you are required to analyze a tabular dataset. Evaluate the Usage of array datatype with suitable example. Explain its limitations. K5(10)
- 7) In a study, you want to determine if the mean height of a sample of K5(10)

50 adults is significantly different from a known population mean height of 65 inches. Perform a one-sample t-test and state the null and alternative hypotheses.

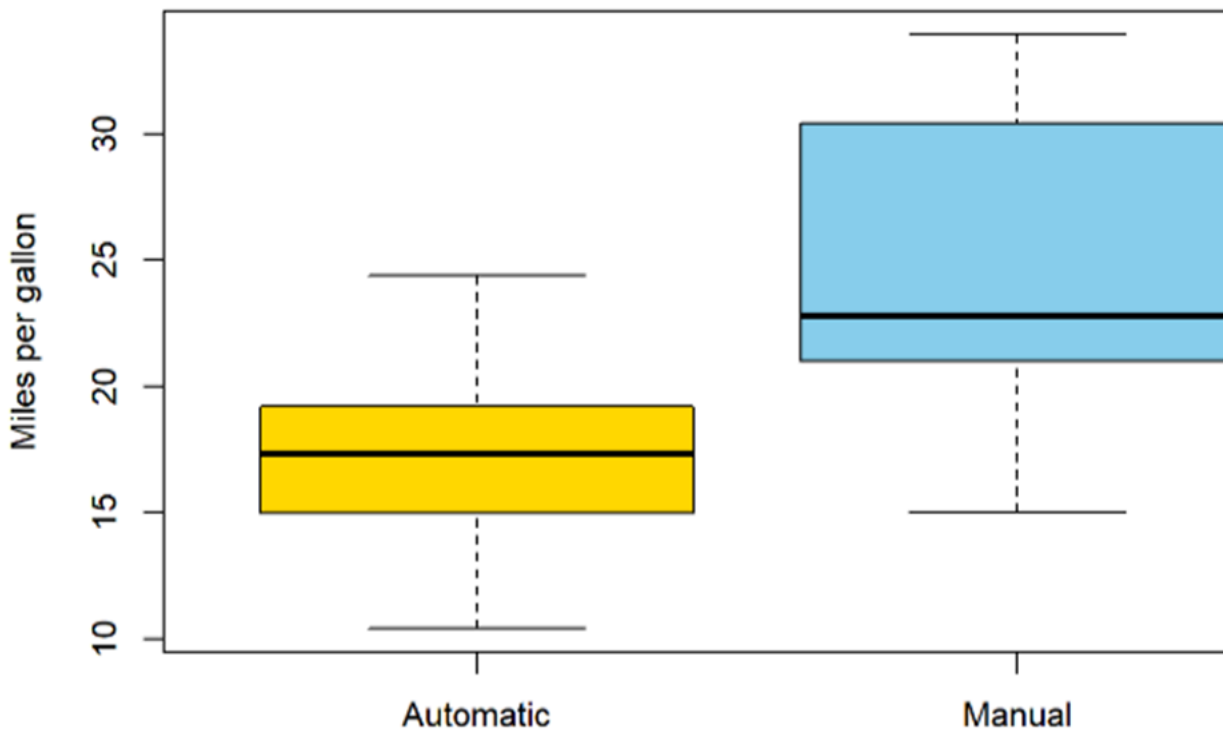
Suggest R code for it.

Interpret the output: $t = 2.9273$, $df = 49$, $p\text{-value} = 0.005173$

95 percent confidence interval: 65.18183 65.97817

- 8) Evaluate the given chart to compare fuel efficiency of two car models. Provide R code to get this image. What improvements can you suggest in the given image to make it more self-explanatory? K5(15)

Fig. 1. Fuel Efficiency



- 9) Create R program to perform the following operations and manipulations on the employee database using R: K6(18)
1. Create a dataframe using the given dataset
 2. Delete the employee named Mike Jones from the database.
 3. Sort the employee database based on Last_Name in ascending order.
 4. Update the Salary of Emily Brown to \$90,000.
 5. Calculate the total salary expenses for all employees in the database.
 6. Write code to create a cross-tabulation with column name of employee and their complete address.

Employee_ID	First_Name	Last_Name	Age	Gender	Department	Position	Salary	Start_Date	End_Date	Email
1	John	Doe	30	M	IT	Software Engineer	70000	2020-01-01	2023-12-31	john.doe@domain.com
2	Jane	Smith	25	F	HR	HR Manager	80000	2019-05-01	NA	jane.smith@domain.com
3	Mike	Jones	35	M	Finance	Financial Analyst	75000	2018-07-01	2021-12-31	mike.jones@domain.com
4	Emily	Brown	40	F	Marketing	Marketing Manager	85000	2021-03-01	NA	emily.brown@domain.com
5	Chris	Davis	28	M	Operations	Operations Manager	90000	2017-02-01	NA	chris.davis@domain.com
6	Samantha	Green	30	F	Finance	Financial Analyst	70000	2022-01-01	NA	samantha.green@domain.com
7	Alex	White	25	M	IT	Software Engineer	90000	2023-01-01	NA	alex.white@domain.com

- 10) In an experiment, you have two groups, Group A and Group B, and K6(12)

you want to compare their mean test scores to see if there is a significant difference. To perform a two-sample independent t-test, state the null and alternative hypotheses. Write R program for it. What is the expected output of the R program.