School of Business
Master of Business Administration MBA Dual Specialization
Semester End Examination - May 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) Apply the appropriate R function or syntax to write a R program to find given Number is Even or Odd. — K3 (6)

2) Analyze the various datatypes in R Language to suggest for storing name, address and mobile number of respondents for a survey-based dataset. — K4 (8)

3) Explain the usage of packages in R? How do you import it? Write R code to import package for data manipulation, excel file operations and filtering datasets. — K4 (8)

4) Analyze read.csv() function in R code: mydata <- read.csv("mydata.csv"). What are its advantages. — K4 (4)

5) Demonstrate use of switch statement to judge grades of 10th std students in exams. Write a sample R-code for given dataset and store them in a separate CSV file. — K3 (9)

Subject ID	Subject	marks	Grade
1	Maths.	86	?
2	Science	79	?
3	English	65	?
4	Social science	52	?
5	Computer	65	?

6) Compare the advantages and disadvantages of using ggplot2 for exploratory data analysis. — K5 (10)
Explain the following R code:
ggplot(data = storms75, aes(x = wind, color = name)) +
geom_density(aes(fill = name)) +
What output do you expect from it?
facet_wrap(~ name)

7) Evaluate the role of chi-square test in business decisions. Provide R program to run this test using the data:
Favorite_Color = c("Red", "Blue", "Green", "Red", "Blue", "Green"),
Gender = c("Male", "Female", "Male", "Female", "Male", "Female") **K5 (10)**

8) Suppose that you are required to analyze the given tabular dataset. Evaluate the Usage of vector datatype with suitable example. Explain its limitations. What will be the expected output of R code: x = c(1, 2, TRUE, 4+5i) print (class(x)) **K5 (15)**

Position	Salary	Start_Date	End_Date
Software Engineer	70000	2020-01-01	2023-12-31
HR Manager	80000	2019-05-01	NA
Financial Analyst	75000	2018-07-01	2021-12-31
Marketing Manager	85000	2021-03-01	NA
Operations Manager	90000	2017-02-01	2022-1-31

9) Write R program for the purpose of conducting a hypothesis test for the overall model (ANOVA) in regression analysis for the given dataset. Suggest what would be the expected output and how the results are interpretated. **K6 (18)**

sales, advertising spending
1000, 50
1200, 60
1500, 70
1800, 80
2000, 90
2200, 100
2300, 110
2500, 120
2700, 130
3000, 140

Where
"sales" represents the sales amount, which is the dependent variable.
"advertising spending" represents the amount spent on advertising, which is the independent variable.

10) A large e-commerce company is focused on optimizing its supply chain operations to improve customer satisfaction and reduce costs. One area of interest is the time it takes to fulfill orders from its warehouses. They collect data on various factors that might impact order fulfillment time, such as warehouse size, number of workers, transportation time, and order volume. The objective of this case study is to use multiple regression analysis to develop a predictive model that can estimate the time it takes to fulfill an order based on the available data with the objective of this case study is to use multiple regression analysis to develop a predictive model that can estimate the time it takes to fulfill an order based on the available data. Formulate suitable hypothesis and write R program to test and build models. What interpretations a business manager can make?

K6 (12)

Warehouse Size (sq. ft.)	Number of Workers	Transportation Time (hours)	Order Volume	Order Fulfillment Time (hours)
5000	10	2	100	4
6000	15	3	150	5
7000	12	2.5	120	4.5
5500	8	2	90	3.5
8000	20	4	200	6