**POORNIMA UNIVERSITY, JAIPUR.**

**END SEMESTER EXAMINATION, APRIL 2023**

|  |  |  |  |
| --- | --- | --- | --- |
|  | **2BC4104** | Roll No. | Total Printed Pages: 2 |
| **2BC4104** |  |
| BCA II Year IV- Semester (Main/Back) End Semester Examination, April 2023  **(AIDS)** | |
| **BASCCA4104 : R Programming** | | | |

# Time: **3** Hours. Total Marks: **60**

Min. Passing Marks: **21/24/27**

Attempt **five** questions selecting one question from each Unit. There is internal choice from Unit I to Unit V. Marks of each question or its parts are indicated against each question / parts. Draw neat sketches wherever necessary to illustrate the answer. Assume missing data suitably (if any) and clearly indicate the same in the answer.

Use of following supporting material is permitted during examination for this subject.

# **1.--------------------------Nil--------------------** **2.------------------Nil-----------------------**

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | **UNIT-I (CO1)** | **Marks** | **Bloom Level** |
| **Q.1** | **(a)** | Given a vector as (4,6,2,NA,5,NA,2,8,NA,34,NA,12)  Apply concepts of R to code for :   1. Display indexes of NA values 2. Sum of all non NA values 3. Display elements greater than 10 4. Store non NA values in new vector y.   Replace NA by mean of remaining non NA values | **(6)** | **Apply** |
|  |  |  |  |  |
|  | **(b)** | Illustrate with example: cat,paste and print function. | **(6)** | **Apply** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
| **Q.2** | **(a)** | How to perform on R console (i) Set working directory (ii) Get working directory  (iii)List files (vi) Workspace variables | **(6)** | **Apply** |
|  |  |  |  |  |
|  | **(b)** | Write a simple R code to find simple interest. The data entered must be of double type. | **(6)** | **Apply** |
|  |  |  |  |  |
|  |  | **UNIT-II (CO2)** |  |  |
|  |  |  |  |  |
| **Q.3** | **(a)** | Create a matrix of size 4x5 and populate with random elements from 10:40 without replacement. Set rownames and R1-R4 and column name as C1-C5. Display All elements having values greater than 20 but less than 30. | **(6)** | **Understand** |
|  |  |  |  |  |
|  | **(b)** | Consider a dataframe **shop** having two columns: price and type. The type has values:”food”,”toy”,”cloth”,”electronics”. Assuming random values for both the columns. Find out mean price for all the different types. | **(6)** | **Understand** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
| **Q.4** | **(a)** | Write a function to find maximum,minimum,sum and average of 3 numbers and return them as list. | **(6)** | **Apply** |
|  |  |  |  |  |
|  | **(b)** | How do you generate a covariance matrix in R? Write a function and demonstrate. | **(6)** | **Understand** |
|  |  | **UNIT-III (CO3)** |  |  |
| **Q.5** | **(a)** | How to read and write excel files in R? Write suitable code for the same. | **(6)** | **Apply** |
|  |  |  |  |  |
|  | **(b)** | Apply functions substr, regexpr ,strsplit and nchar on suitable strings. | **(6)** | **Understand** |
|  |  |  |  |  |
|  |  | **OR** |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| **Q.6** | **(a)** | Given a file :data.txt with contents:  12122023919878797879  05122023919878767879  11122023919878795879  Write code to create dataframe with columns day,month,year,countrycode,mobile no. By reading the fixed format data. | **(6)** | **Understand** |
|  |  |  |  |  |
|  | **(b)** | Write function to check a file with given suffix in R? | **(6)** | **Understand** |
|  |  | **UNIT-IV (CO4)** |  |  |
| **Q.7** | **(a)** | How do you find mode of a given vector? Mention at least two methods | **(6)** | **Understand** |
|  |  |  |  |  |
|  | **(b)** | A cosmetics company fills its best-selling 8-ounce jars of facial cream by an automatic dispensing machine. The machine is set to dispense a mean of 8.1 ounces per jar. Uncontrollable factors in the process can shift the mean away from 8.1 and cause either underfill or overfill, both of which are undesirable. In such a case the dispensing machine is stopped and recalibrated. Regardless of the mean amount dispensed, the standard deviation of the amount dispensed always has value 0.22 ounce. A quality control engineer routinely selects 30 jars from the assembly line to check the amounts filled. On one occasion, the sample  Mean is ǜ = 8.2 ounces and the sample standard deviation is s = 0.25 ounce. Determine if there is sufficient evidence in the sample to indicate, at the 1% level of significance, that the machine should be recalibrated. (Z0.005=2.576) | **(6)** | **Understand** |
|  |  | **OR** |  |  |
| **Q.8** | **(a)** | Before an increasing in exercise duty on tea, 800 persons out of a sample of 1000 persons were found to be tea drinkers. After an increasing in duty, 800 people were tea drinkers in a sample of 1200 people. State null and alternate hypothesis. Using SE of a proportion, state whether there is a significant decrease in consumption of tea after the increase in the excise duty (alpha=0.05) | **(6)** | **Understand** |
|  |  |  |  |  |
|  | **(b)** | Describe the concept of cross tabulation? How do you recreate original data from contingency table? | **(6)** | **Understand** |
|  |  | **UNIT V (CO5)** |  |  |
| **Q.9** | **(a)** | A new approach to prenatal care is proposed for pregnant women living in a rural community. The new program involves in-home visits during the course of pregnancy in addition to the usual or regularly scheduled visits. A pilot randomized trial with 15 pregnant women is designed to evaluate whether women who participate in the program deliver healthier babies than women receiving usual care. The outcome is the APGAR score measured 5 minutes after birth. Recall that APGAR scores range from 0 to 10 with scores of 7 or higher considered normal (healthy), 4-6 low and 0-3 critically low. The data are shown below. | **(6)** | **Understand** |
|  |  | Is there statistical evidence of a difference in APGAR scores in women receiving the new and enhanced versus usual prenatal care?The critical value for this test with n1=8, n2=7 and α =0.05 is 10 |  |  |
|  | **(b)** | Acme Toy Company prints baseball cards. The company claims that 30% of the cards are rookies, 60% veterans but not All-Stars, and 10% are veteran All-Stars. Suppose a random sample of 100 cards has 45 rookies, 50 veterans, and 5 All-Stars. Is this consistent with Acme's claim? Use a 0.05 level of  significance.(ChiSquare at 0.05 for df=2 is 5.991 | **(6)** | **Understand** |
|  |  | **OR** |  |  |
| **Q.10** |  | The data represents sale of cold drinks : 3,4,1,1,3,4,3,3,1,3,2,1,2,1,2,3,2,3,1,1,1,1,1,4,3,2  (1-Limca,2-Pepsi 3-Mirinda 4-Fanta).. Represent data using bar chart with Cold drink name in the center. | **(6)** | **Remember** |
|  |  |  |  |  |
|  | **(b)** | How two way ANOVA is different than ONE Way ANOVA? Explain using example. | **(6)** | **Remember** |