UPES

UNIVERSITY OF PETROLEUM AND ENERGY STUDIES

End Semester Examination – December, 2018

Program/course: MBA (BA)   **Semester – III**
Subject: Programming for Analytics   **Max. Marks : 100**
Code   : DSBA 8004   **Duration   : 3 Hrs**
No. of page/s: 6

Section A

1. Select most appropriate answer.   (co_1)   **(2x10)**

    I.   In the well-known package ggplot2 of R "gg" stands for

        (a) Grammar of Geometry
        (b) Grammar of Graphics
        (c) Grammar of Geom
        (d) None of these

    II.   How to represent the represent impossible values in R?

        (a) NA
        (b) NaN
        (c) NoN
        (d) none of these

    III.   The R language is a dialect of which of the following programming languages?

        (a) S
        (b) SAS
        (c) C
        (d) Matlab

IV.   What function is used to test the missing observation in data frame
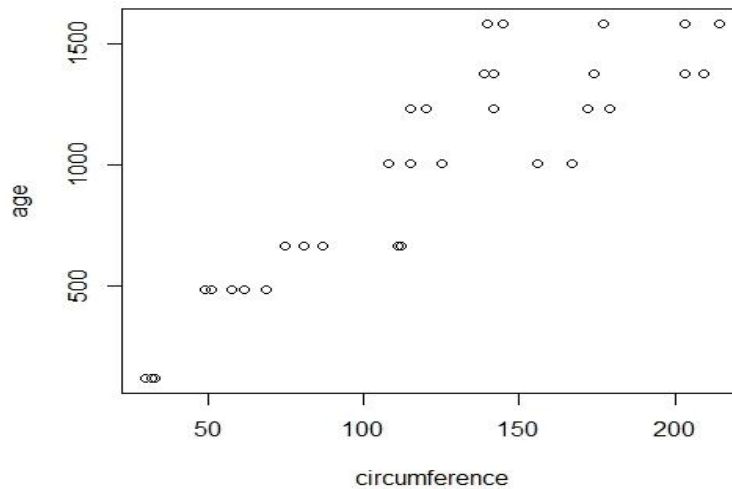
(a) Missing()
(b) NA.miss()
(c) na()
(d) is.na

V.   How many states are present in the data frame named state.name already in build in R

(a) 52
(b) 25
(c) 50
(d) 10

VI.   Which command is used to know the structure of the data frame in R

(a) seq()
(b) str()
(c) structure()
(d) none of these

VII.   What can be the value of r (correlation coefficient) for this given graph

circumference

(a) 0.5 to 0.7
(b) -0.5 to – 0.7
(c) 0.7 to 0.9
(d) -0.7 to 0.9

VIII.   Which function is generally used after the installation of a package in R

(a) library ()
(b) library.dyam()
(c) library.install()
(d) none of these

IX.   For Preparing the word cloud in R, which package is required before "wordcloud"

(a) tm ()
(b) lm()
(c) gm()
(d) none of these

X.   Which function helps you perform sorting in R language?

(a) Order ()
(b) Short()
(c) Merge()
(d) None of these

Section B

Attempt any four questions.   **(5x4)**

2. Suppose traffic accidents occur at an intersection with a mean rate of 3.7 per year. Simulate the annual number of accidents for a 10-year period, assuming a Poisson model?   (co_2)
3. What is the difference between data frame and a matrix in R?   (co_2)
4. Toss a fair coin 100 times (using R). How many heads do you have? (write R Codes)   (co_3)
5. How can you find the mean and variance for the data sets containing missing value in R? (write the R code)   (co_2)
6. Write R codes using rep() and seq() as required to create the following vectors (co_3)
   0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4 and
   1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5

Section-C

Answer any four questions.   **(10 x 4)**

7. The following are a sample of observations on incoming solar radiation at a greenhouse:
    11.1, 10.6, 6.3, 8.8, 10.7, 11.2, 8.9, 12.2

    (a) Assign the data to an object called solar.radiation.   (co_2)
    (b) Find the mean, median, range, and variance of the radiation observations.   (co_2)
    (c) Add 10 to each observation of solar.radiation, and assign the result to sr10. Find the
       mean, median, range, and variance of sr10. Which statistics change, and by how much?

8. The factorial $n!$ counts how many ways n different objects could be ordered. It is defined
    as $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n.$ Write R codes to calculate n factorial using for() loop and
    find the value of 10! using the same code.   (co_2)

9. Run the following codes and write the interpretation line by line including the nature of
    data set used.   (co_2)

    (a) y=rnorm(1000)
       f=factor(rep(1:10,100))
       boxplot(y ~ f, main="Boxplot of normal random data with model notation")
       hist(y)

    (b) x <- rnorm(100)
       hist(x,probability=T)
       x <- x[(0 < x) & (x < 3)]
       hist(x, probability=TRUE)

10. Write R codes to generate hypothetical random data using normal Distribution of Size 10
    with mean 3 and standard deviation 1. Select 10 samples of size 5 form the generated data
    set without replacement and find their mean and standard deviation. What conclusion you
    can draw based on this exercise.   (co_2)

11. Why this following commands / functions used in R ?   (co_3)

       a. head(newdata2)
       b. xtabs(~admit+rank,data=b)
       c. sample.split()
       d. class(AirPassengers)

      e.   abline(98.0054, 0.9528)

12. There are some errors / missing statements in the following program, find those errors and correct the following program, run on R save the output.   (co3)

```
install.packages("car")
install.packages("corrplot")
install.packages("visreg")
install.packages("rgl")
install.packages("knitr")
install.packages("scatterplot3d")

library(Car)
library(corrplot)
library(visreg)
library(rgl)
library(knitr)

Prestige
head(Prestige)
str(Prestige)
summary(Prestige)
newPdata=Prestige[,c(1:4)]
summary(newPdata)
newPdata
plot(newPdata,col="blue")

education.c= scale(newPdata$education, center=TRUE, scale=FALSE)
prestige.c= scale(newPdata$prestige, center=TRUE, scale=FALSE)
women.c= scale(newPdata$women, center=TRUE, scale=FALSE)
new.c.vars= cbind(education.c, prestige.c, women.c)
newPdata1 = cbind(newPdata, new.c.vars)
names(newPdata1)[5:7] = c("education.c", "prestige.c", "women.c" )
summary(newPdata1)

mod1 = lm(income ~ education.c+ prestige.c+ women.c, data=newPdata)
mod1
summary(mod1)
```

Section-D

Answer the questions. (co3) **(20)**

16. Are You Going To Hate Your New Job?

Getting a new job can be an exciting and energizing event in your life. But what if you discover after a short time on the job that you hate your job? Is there any way to determine ahead of time whether you will love or hate your job? Sue Shellenbarger of The Wall *Street Journal* discuss some of the things to look for when interviewing for a position that may provide clues as to whether you will be happy on that job.

Among other things, work culture vary from hip, freewheeling start-ups to old school, organizations place pressure on workers to feel tense and to work long hours while others place more emphasis on creativity and the bottom line. Shellenbarger suggest that job interviewees pay close attention to how they are treated in an interview. Are they just another cog in the wheel or are they valued as an individual? Is a work life balance apparent within the company? Ask what a typical workday is like at that firm. Inquire about the values that undergird the management by asking questions such as "What is your proudest accomplishment". Ask about flexible schedules and how job training is managed. For example, does a worker have to go to job training on their own time?

A "Work Trends" survey undertaking by the John J. Heldrich Center for Workforce Development at Rutgers University and the Center for survey Research and Analysis at the University for Connecticut posted several questions to employees in a survey to ascertain their job satisfaction. Some other things included in these questions for relationship with your supervisor, overall quality of the work environment, total hours worked each week and opportunities for advancement at the job.

Suppose another research are gathered survey data from 19 employees on these questions and also ask the employees to read their job satisfaction on a scale from 0 to 100 (with 100 being perfectly satisfied). Suppose the following data represent the results of this survey. Assume that relationship which supervisor is rated on a scale from 100 to 50 (0 represent poor relationship and 50 represent an excellent relationship), oral quality of the work environment is rated on a scale from 0 to 100 (0 represents poor work environment and 100 represents an excellent work environment), and opportunities for advancement is rated on a scale from 0 to 50 (0 represents no opportunities and 50 represent excellent opportunities).

Job Satisfaction	Relationship with supervisor	Overall quality of work environment	Total hours worked per week	Opportunities for advancement
55	27	65	50	42
20	12	13	60	28
85	40	79	45	7
65	35	53	65	48
45	29	43	40	32
70	42	62	50	41
35	22	18	75	18
60	34	75	40	32
95	50	84	45	48
65	33	68	60	11
85	40	72	55	33
10	5	10	50	21
75	37	64	45	52
80	42	82	40	46
50	31	46	60	48
90	47	95	55	30
75	36	82	70	39
45	20	42	40	22
65	32	73	55	12

Managerial and Statistical Questions

1. Several variables are presented that maybe related to job satisfaction. Which variables are stronger predictors of job satisfaction? Might other variables not mentioned here be related to job satisfaction?

2. Is it possible to develop a mathematical model to predict job satisfaction using the data given? If so, how strong is the model? With for independent variables, will we need to develop four different simple regression models and compare their results?