

Name:
Enrolment No:



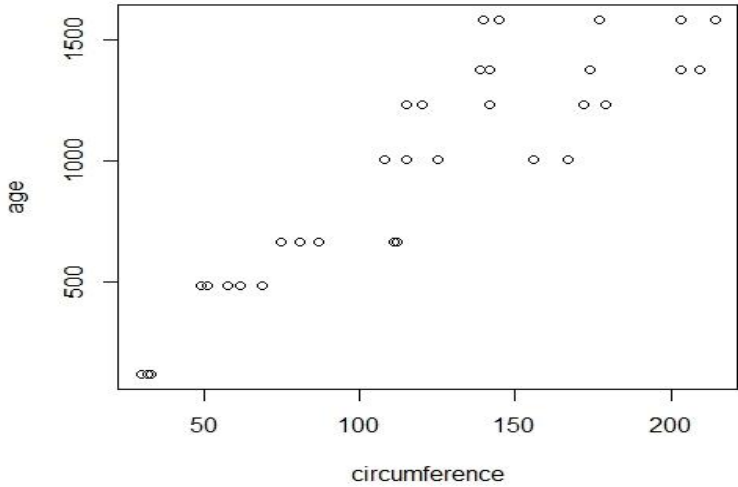
UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
End Semester Examination, December 2019-Set-I

Course: Programming for Analytics
Program: MBA BA
Course code: DSBA 8004
Instructions:

Semester: III
Time: 03 Hours
Max. Marks: 100

SECTION A

		Marks	CO
Q 1	Select the most appropriate answer.	2 X 10=20	CO ₁
I.	How to represent the impossible values in R? (a) NA (b) NaN (c) NoN (d) none of these		
II.	Which function is generally used after the installation of a package in R (a) library () (b) library.dynam() (c) library.install() (d) none of these		
III.	Which function helps you perform sorting in R language? (a) Order () (b) Short() (c) Merge() (d) None of these		
IV.	airquality\$Ozone command used for (a) access the Ozone variable of airquality data (b) access the airquality variable of Ozone data (c) for multiplying both the variables (d) none of these		
V.	To find cumulative probability of Poisson distribution we use (e) dpois (f) rpois (g) ppois (h) none of these		

VI.	<p>The R language is a dialect of which of the following programming languages?</p> <ul style="list-style-type: none"> (a) S (b) SAS (c) C (d) Matlab 		
VII.	<p>What function is used to test the missing observation in data frame</p> <ul style="list-style-type: none"> (a) Missing() (b) NA.miss() (c) na() (d) is.na 		
VIII.	<p>Which command is used to know the structure of the data frame in R</p> <ul style="list-style-type: none"> (a) seq() (b) str() (c) structure() (d) none of these 		
IX.	<p>What can be the value of r (correlation coefficient) for this given graph</p>  <p>The scatter plot shows a positive linear relationship between circumference (x-axis) and age (y-axis). The x-axis ranges from 50 to 200, and the y-axis ranges from 500 to 1500. The data points are scattered but show a clear upward trend.</p> <ul style="list-style-type: none"> (a) 0.5 to 0.7 (b) -0.5 to -0.7 (c) 0.7 to 0.9 (d) -0.7 to 0.9 		
X.	<p>R files has an extension _____</p>		

- (a) .R
- (b) .S
- (c) .Rp
- (d) .c

SECTION B

Q	Attempt any four questions	5 X 4=20	
2.	How can you find the mean and variance for the data sets containing missing value in R? (write the R code)		CO₂
3.	Simulate the process of tossing a coin 1000 times and find the probability of getting Head.		CO₂
4.	What is the difference between data frame and a matrix in R?		CO₂
5.	Generate the 1000 random normal population, take a sample of size 100 and test whether the sample has been taken from same population or not.		CO₂
6.	Explain the method of setting the directory in R.		CO₁

SECTION-C

Q	Attempt any four questions :	10 X 4 =40																											
7.	<p>The Airline cost data (given in table 1), shown in figure 1. Develop the simple regression equation to predict the cost of flying using number of passengers. Also Estimate the cost if number of passenger increase to 120.</p> <p align="center">Table: 1</p> <table border="1"> <thead> <tr> <th>Number of Passengers(X)</th> <th>Cost (\$1000):y</th> </tr> </thead> <tbody> <tr><td>61</td><td>4.28</td></tr> <tr><td>63</td><td>4.08</td></tr> <tr><td>67</td><td>4.42</td></tr> <tr><td>69</td><td>4.17</td></tr> <tr><td>70</td><td>4.48</td></tr> <tr><td>74</td><td>4.3</td></tr> <tr><td>76</td><td>4.82</td></tr> <tr><td>81</td><td>4.7</td></tr> <tr><td>86</td><td>5.11</td></tr> <tr><td>91</td><td>5.13</td></tr> <tr><td>95</td><td>5.64</td></tr> <tr><td>97</td><td>5.56</td></tr> </tbody> </table>	Number of Passengers(X)	Cost (\$1000):y	61	4.28	63	4.08	67	4.42	69	4.17	70	4.48	74	4.3	76	4.82	81	4.7	86	5.11	91	5.13	95	5.64	97	5.56		CO₂
Number of Passengers(X)	Cost (\$1000):y																												
61	4.28																												
63	4.08																												
67	4.42																												
69	4.17																												
70	4.48																												
74	4.3																												
76	4.82																												
81	4.7																												
86	5.11																												
91	5.13																												
95	5.64																												
97	5.56																												
8.	Write an R program to Generate 1000 random numbers and calculate their mean and variance. Take 5 samples of size 100 and calculate their mean and variance. What can you conclude by this exercise.		CO₃																										
9.	<p>Why this following commands / functions used in R ?</p> <ul style="list-style-type: none"> a. head(newdata2) b. xtabs(~admit+rank,data=b) c. sample.split() 																												

	<p>d. <code>class(AirPassengers)</code> e. <code>abline(98.0054, 0.9528)</code></p>		CO₂
10.	<p>Run the following codes and write the interpretation line by line including the nature of data set used.</p> <p>(a) <code>y=rnorm(1000)</code> <code>f=factor(rep(1:10,100))</code> <code>boxplot(y ~ f, main="Boxplot of normal random data with model notation")</code> <code>hist(y)</code></p> <p>(b) <code>x <- rnorm(100)</code> <code>hist(x,probability=T)</code> <code>x <- x[(0 < x) & (x < 3)]</code> <code>hist(x, probability=TRUE)</code></p>		CO₂
11.	<p>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</p> <p>(a) Assign the data to an object called <code>solar.radiation</code>. (b) Find the mean, median, range, and variance of the radiation observations. (c) Add 10 to each observation of <code>solar.radiation</code>, and assign the result to <code>sr10</code>.</p> <p>Find the mean, median, range, and variance of <code>sr10</code>. Which statistics change, and by how much?</p>		CO₁
12	<p>The factorial $n!$ counts how many ways n different objects could be ordered. It is defined as $n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot (n - 1) \cdot n$. Write R codes to calculate n factorial using <code>for()</code> loop and find the value of $50!$ using the same code.</p>		CO₃
SECTION-D			
Q	Answer the Question	20 X 1=20	
13.	<p>Are You Going To Hate Your New Job?</p> <p>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 <i>The Wall Street Journal</i> 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.</p> <p>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</p>		CO₃

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 undertaken by the John J. Heldrich Center for Workforce Development at Rutgers University and the Center for survey Research and Analysis at the University of 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

	<p>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?</p> <p>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?</p>		
--	---	--	--