

USN

--	--	--	--	--	--	--	--	--	--

18IM652

## Sixth Semester B.E. Degree Examination, July/August 2022

### Data Analytics for Engineers

Time: 3 hrs.

Max. Marks: 100

**Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.  
2. Use of Statistical Tables is permitted.**

#### Module-1

- 1 a. Discuss the different methods of Data input in R. (08 Marks)
- b. Explain the advantages of R language as a statistical software. (06 Marks)
- c. Write a brief note on Data Accessing in R. (06 Marks)

**OR**

- 2 a. Explain briefly the following with respect to R : i) Plot function ii) Data frame function  
iii) Matrix function iv) Class function. (08 Marks)
- b. Explain briefly the method of Saving , Storing and Retrieving work in R (06 Marks)
- c. Explain some of the built in functions in R. (06 Marks)

#### Module-2

- 3 a. For the following data, draw i) Histogram ii) Frequency polygon. (10 Marks)

Class	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
Frequency	14	26	37	44	56	44	38	26	15

Write the R commands to obtain Histogram and Frequency Polygon for the above data.

- b. Calculate the Mean, Median and Mode for the following data :

Class	66-67	67-68	68-69	69-70	70-71	71-72
Frequency	15	24	40	20	14	11

Write the R commands to compute Mean and Median for the above data.

(10 Marks)

**OR**

- 4 a. Construct Histogram and Frequency Polygon for the data given in table below. Write the R commands to draw histogram and frequency polygon on the same graph for the following table : (10 Marks)

Marks Scored	20-29	30-39	40-49	50-59	60-69	70-79	80-89
No. of Students	50	64	58	42	38	25	13

- b. Find Mean and Median and then write R commands to compute Mean and Median for the following frequency distribution : (10 Marks)

Height in cm	No. of students
145-150	4
150-155	6
155-160	28
160-165	58
165-170	64
170-175	30
175-180	5
180-185	5

**Module-3**

- 5 a. Discuss the properties of Probability. (06 Marks)
- b. A Committee of University teachers consists of 3 Professors, 5 Associate Professors and 2 Assistant Professors. A sub committee of 6 is selected by drawing names out of a hat. What is the probability that the sub committee is composed of 2 Professors, 3 Associate Professor and 1 Assistant Professor? (08 Marks)
- c. If the number of defective components produced by a machine per day follows Poisson distribution, with parameters 2. What is the probability that, out of total production of day, there are i) Exactly 2 defective components ii) Atleast 1 defective component. (06 Marks)

**OR**

- 6 a. Explain the characteristics of Normal distribution. (06 Marks)
- b. A random sample of 20 is selected from a lot containing 200 articles, 12 of which are defective. Find the probability that the sample will contain : i) Less than 2 defects ii) Exactly 2 defects iii) Utmost 3 defects iv) Atleast 1 defect. (08 Marks)
- c. If the weight of 300 students are normally distributed with a mean of 68 kg and variance of 9kg. How many students have this weight at i) greater than 75 kgs ii) less than 64 kgs iii) between 64 & 72 kgs. (06 Marks)

**Module-4**

- 7 a. Discuss the Null hypothesis and Alternate hypothesis. (06 Marks)
- b. A die is rolled 100 times with the following distribution :

Number	1	2	3	4	5	6
Observed frequency	17	14	20	17	17	15

At 0.01 level of significance, determine whether the die is true (or uniform). Using Chi – Square for goodness of fit test. (06 Marks)

- c. An IQ test was administered to 5 persons before and after they were trained. Results are given below in table :

Candidates	I	II	III	IV	V
IQ before training	110	120	123	132	125
IQ after training	120	118	125	136	121

Test whether there is any change in IQ after training programme. (08 Marks)

**OR**

- 8 a. Explain the Type I errors and Type II errors in Testing of Hypothesis. (06 Marks)
- b. The number of scooter accidents per month in a certain city were as follows : 12, 8, 20, 2, 14, 10, 15, 6, 9, 4. Check whether these frequencies are in agreement with that of accident conditions same during 10 month period by applying Chi – square test of goodness of fit. (06 Marks)
- c. The sales data of an item in 6 shops before and after a special promotional campaign are as follows :

Shops	A	B	C	D	E	F
Before Campaign	53	28	31	48	50	42
After Campaign	58	29	30	55	56	45

Can the Campaign be judged to be a success? Test at 5% level of significance using paired t – test. (08 Marks)

**Module-5**

- 9 a. Explain briefly the significance of Regression analysis in statistics. (06 Marks)  
 b. Obtain the two regression equations for the following data :

Purchases	62	72	98	76	81	56	76	92	88	49
Sales	112	124	131	117	132	96	120	136	97	85

- i) Estimate the likely sales when the purchases are equal to 100.  
 ii) Write the R commands to obtain the two regression equations. (14 Marks)

**OR**

- 10 a. Discuss the various types of correlation. (06 Marks)  
 b. Following table gives the Ages and Blood Pressure of ten women :

Age X years	56	42	36	47	49	42	60	72	63	53
Blood pressure Y	147	125	118	128	145	140	155	160	149	143

- i) Find the correlation coefficient between X and Y.  
 ii) Determine the Regression equation of Y on X.  
 iii) Write the R commands to find Correlation coefficient and Regression equation of Y on X. (14 Marks)

\* \* \* \* \*