Total No. of Questions: 6 *Total No. of Printed Pages:2*

Enrollment No.....

Faculty of Science
End Sem Examination Dec-2023
BC3EA02 Exploratory Data Analysis

Programme: B.Sc. Branch/Specialisation: Computer Science

Duration: 3 Hrs. **Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

Q.1 i. Data Analysis is defined by the statistician- **1**
(a) William S.
(b) Hans Peter Luhn
(c) Gregory Piatetsky-Shapiro
(d) John Tukey

ii. In EDA, what is the term for a graphical representation that displays the distribution of a continuous variable- **1**
(a) Pie Chart (b) Histogram (c) Bar Chart (d) Scatter Plot

iii. What is the primary purpose of Exploratory Data Analysis (EDA) in data science? **1**
(a) To make data more complicated
(b) To store data
(c) To discover patterns and insights in data
(d) To manipulate data

iv. How many main statistical methodologies are used in data analysis? **1**
(a) 2 (b) 3 (c) 4 (d) 5

v. This metric indicates how dispersed a range of numbers is- **1**
(a) Skewness (b) Median (c) Standard Deviation (d) Mode

vi. Which of the following gave rise to need of graphs in data analysis? **1**
(a) Data visualization (b) Communicating results
(c) Decision making (d) All of these

vii. Most often, EDA relies on _____. **1**
(a) Visual techniques (b) Assumptions
(c) Fixed Models (d) Testing for statistical significance

viii. Which of the following is characteristic of exploratory graph? **1**
(a) Made slowly
(b) Axes are not cleaned up
(c) Color is used for personal information
(d) All of the mentioned

ix. Data analysis is a process of- **1**
(a) Inspecting data (b) Cleaning data
(c) Transforming data (d) All of these

x. _____ is used in ANOVA to compare the variance between multiple groups to the variance within those groups. **1**
(a) F-test (b) Z-test (c) T-test (d) X^2-test

Q.2 i. Write the difference between classic data analysis and exploratory data analysis. **2**

ii. What is autocorrelation graph? Explain in brief. **3**

iii. Write any two techniques for testing assumptions. **5**

OR iv. Distinguish between data and information. **5**

Q.3 i. What is Bartlett's test? Explain in brief. **2**

ii. Explain about the benefits of data visualization. **8**

OR iii. What are the best settings for factors? **8**

Q.4 i. What is ANOVA? Explain in brief. **3**

ii. Explain location and scale parameters along with estimation of parameters. **7**

OR iii. Discuss all four cases of autocorrelation plot. **7**

Q.5 i. What is autocorrelation graph? Explain in brief. **4**

ii. Explain in brief random distribution. **6**

OR iii. Explain probability distribution in detail. **6**

Q.6 Attempt any two:

i. Explain sinusoidal correlation. **5**

ii. What is summary analysis? Explain. **5**

iii. Explain hypothesis. **5**

******

Scheme of Marking

Faculty of Science

End Sem Examination Dec-2023

Exploratory Data Analysis-BC3EA02(T)

Programme: B.Sc. Branch/Specialisation:

Note: The Paper Setter should provide the answer wise splitting of the marks in the scheme below.

Q.1	i)	John Tukey	**1**
	ii)	Histogram	**1**
	iii)	To discover patterns and insights in data	**1**
	iv)	2	**1**
	v)	Standard Deviation	**1**
	vi)	All of the mentioned	**1**
	vii)	Visual techniques	**1**
	viii)	Color is used for personal information	**1**
	ix)	All of the above	**1**
	x)	F-test	**1**
Q.2	i.	Write the Difference between classic data analysis and exploratory data analysis 4 Difference **2 marks**	**2**
	ii.	What is autocorrelation graph, explain in brief? **3 marks**	**3**
	iii.	Write any two techniques for testing assumptions. Two techniques **5 marks**	**5**
OR	iv.	Distinguish the difference between data and Information? 10 Difference **5 marks**	**5**
Q.3	i.	What is Bartlett's test? Explain in brief. **2 marks**	**2**
	ii.	Explain about the benefits of data visualization? **8 marks**	**8**
OR	iii.	What are the best settings for factors? **8 marks**	**8**
Q.4	i.	What is ANOVA? Explain in brief. **3 marks**	**3**
	ii.	Explain location and scale parameters along with estimation of parameters. **7 marks**	**7**
OR	iii.	Discuss all four cases of Autocorrelation plot. **7 marks**	**7**
Q.5	i.	What is autocorrelation graph, explain in brief? **4 marks**	**4**
	ii.	Explain in Brief Random distribution. **6 marks**	**6**
OR	iii.	Explain probability distribution in detail. **6 marks**	**6**
Q.6		Attempt any two:	
	i.	Explain sinusoidal correlation. **5 marks**	**5**
	ii.	What is summary analysis explain. **5 marks**	**5**
	iii.	Explain Hypothesis. **5 marks**	**5**

******