7-4-25

Paper / Subject Code: 53542 / Exploratory Data Analysis TYDS - Lieb

(2½ Hours) **[Total Marks: 75]**

N. B.: (1) **All** questions are **compulsory**.
(2) Make **suitable assumptions** wherever necessary and **state the assumptions** made.
(3) Answers to the **same question** must be **written together**.
(4) Numbers to the **right** indicate **marks**.
(5) Draw **neat labeled diagrams** wherever **necessary**.
(6) Use of **Non-programmable** calculators is **allowed**.

1. **Attempt *any three* of the following:** **15**

a. Distinguish between structured and unstructured data.
b. What role does data play in Exploratory Data Analysis?
c. With the help of an example, explain the importance of data at ordinal and ratio level.
d. State the differences between population and sample. What are the entities that help gather information about population and sample?
e. Compare inferential and descriptive statistics.
f. What is meant by frequency distribution of data?

2. **Attempt *any three* of the following:** **15**

a. EDA is considered a crucial step in data analysis. Explain this statement.
b. Compare and contrast the Classical, Bayesian and Exploratory approaches of data analysis.
c. Explain the use of a lag plot.
d. Discuss the normal probability plot.
e. Elaborate the role of graphics in EDA.
f. Why is EDA considered important?

3. **Attempt *any three* of the following:** **15**

a. Explain in brief the data discovery process.
b. Discuss methods to overcome the following challenges to data accessibility
   i. Data Processing Operational Challenges
   ii. Compliance Challenges with Sensitive Data Access
c. Define data consistency. How can it be measured?
d. Discuss the ways to mitigate data pollution.
e. State the causes of duplicates and ways to detect them.
f. What is meant by noisy data?

4. **Attempt *any three* of the following:** **15**

a List the techniques used in analysis of univariate data.
b How does a bar and pie chart help in analysis of univariate data?
c Explain percentile in brief.
d Discuss positive skewness in univariate data analysis.
e What are measures of concentration?
f Explain the types of kurtosis.

5. **Attempt *any three* of the following:** **15**

a. Explain the chi square calculation for bivariate data analysis.
b. Discuss the significance of Bravais-Pearson Correlation Coefficient.
c. What is the Phi Coefficient?
d. Explain monotonic relationship.
e. Illustrate the significance of scatter plot.
f. How can one measure the association between nominal and metric variables?