

BACHELOR OF TECHNOLOGY (CBCS) (2021-COURSE)
B. Tech. Sem - III COMPUTER SCIENCE & BUSINESS SYSTEM :
WINTER : 2023
SUBJECT : COMPUTATIONAL STATISTICS

Day : Monday

Date : 11/12/2023

Time : 10:00 AM-01:00 PM

Max. Marks : 60

W-24145-2023

N.B.:

- 1) All questions are **COMPULSORY**.
- 2) Figures to the right indicate **FULL** marks.

Q.1 Explain in detail how the process of data wrangling works. (10)

OR

Explain four data types in Python i) List ii) Tuple iii) Set iv) Dictionary. (10)
Also write important functions of these data types along with examples.

Q.2 What is the significance of Groupby function? Find out the statistical (10)
measures such as mean, variance, minimum and maximum value on the
sample data. Also split the data using Groupby function on any column.

OR

What are the crosstab and pivot tables used for in python? Use an example (10)
to show the working of both.

Q.3 Explain Multiple Linear Regression Model. Also explain the assumptions (10)
such as linearity, collinearity, outliers, auto correlation and
homoscedasticity.

OR

Explain Univariate, Bivariate and Multivariate exploratory analysis. (10)
Explain random variable, its mean and variance.

Q.4 Explain the curse of dimensionality and how can it be resolved. (10)

OR

Design a classifier model using Fisher Linear Discriminant Function. (10)

Q.5 Design the model for Factor Analysis. List the assumptions of Factor (10)
Analysis.

OR

How to decide the total number of principle components to be retained using (10)
3 different methods.

Q.6 What are different types of clustering algorithms? Explain important types (10)
of agglomerative clustering.

OR

Write and explain with example K-means algorithm. Also write its (10)
advantages and disadvantages.

* * * * *