

P P SAVANI UNIVERSITY

Seventh Semester of B. Tech. Examination

December 2022

SECE4513 Data Analysis

01.12.2022, Thursday

Time: 10:00 a.m. To 12:30 p.m.

Maximum Marks: 60

Instructions:

1. The question paper comprises of two sections.
2. Section I and II must be attempted in separate answer sheets.
3. Make suitable assumptions and draw neat figures wherever required.
4. Use of scientific calculator is allowed.

SECTION - I

- | Q - 1 | MCQ (Any Five) | [05] | CO | BTL |
|-------|---|------|----|-----|
| (i) | Data Analytics uses _____ to get insights from data.
A. Statistical figures
B. Numerical aspects
C. Statistical methods
D. None of the mentioned above | | 1 | 2 |
| (ii) | Amongst which of the following is / are the branch of statistics which deals with the development of statistical methods is classified as _____.
A. Industry statistics
B. Economic statistics
C. Applied statistics
D. None of the mentioned above | | 1 | 2 |
| (iii) | _____ are used when we want to visually examine the relationship between two quantitative variables.
A. Bar graph
B. Scatterplot
C. Line graph
D. Pie chart | | 1 | 6 |
| (iv) | A graph that uses vertical bars to represent data is called a _____.
A. Bar graph
B. Line graph
C. Scatterplot
D. All of the mentioned above | | 2 | 4 |

(v) Data Analysis is a process of, 2 2

A. Inspecting data
 B. Data Cleaning
 C. Transforming of data
 D. All of the mentioned above

(vi) What is a hypothesis? 1 3

A. A statement that the researcher wants to test through the data collected in a study
 B. A research question the results will answer
 C. A theory that underpins the study
 D. A statistical method for calculating the extent to which the results could have happened by chance

(vii) For each value of the _____, the distribution of the dependent variable must be normal. 3 5

A. Independent variable
 B. Dependent variable
 C. Intermediate variable
 D. None of the mentioned above

Q - 2 (a) Explain Data Analytic Life cycle. [05] 2 4

Q - 2 (b) What is classification? List the different classifiers [05] 3 2

OR

Q - 2 (a) What are the various steps involved in Data Analytics ? [05] 1 4

Q - 2 (b) What is Big Data ? Explain in Detail. [05] 2 2

Q - 3 (a) Which are the technical tools that you have used for analysis and presentation purposes? [05] 2 6

Q - 3 (b) What are the best methods for Data Cleaning ? [05] 3 6

OR

Q - 3 (a) What is the significance of Exploratory Data Analysis (EDA)? [05] 2 3

Q - 3 (b) Describe univariate, bivariate, and multivariate analysis. [05] 3 2

Q - 4	Attempt any one.	[05]		
(i)	Explain descriptive, predictive, and prescriptive analytics.		2	2
OR				
(ii)	Explain the Difference Between Data Mining and Data Profiling.		2	5
SECTION - II				
Q - 1	Short Question (Any Five)	[05]		
(i)	What is Mapping ?		1	1
(ii)	What is NumPy?		1	1
(iii)	Why is seaborn used for Data Visualization ?		3	4
(iv)	Define Inferential Statistics.		2	3
(v)	What is Normal Distribution ?		3	4
(vi)	What is Overfitting ?		3	1
(vii)	What is Underfitting ?		2	1
Q - 2 (a)	What are different types of Hypothesis Testing ? Explain.	[05]	2	1
Q - 2 (b)	What are the Different ways to create a Dataframe in Pandas ?	[05]	1	4
OR				
Q - 2 (a)	How will you print 4 random numbers between 1 and 15 using NumPy ?	[05]	2	6
Q - 2 (b)	How to Handle the Data with lots of NaN and Missing Value ?	[05]	2	3
Q - 3 (a)	Explain Outliers in Detail.	[05]	1	3
Q - 3 (b)	Write a Short Note on Time Series Analysis.	[05]	2	2
OR				
Q - 3 (a)	Differentiate Between Variance and Co- Variance.	[05]	2	2
Q - 3 (b)	What is Data Validation ? Explain in Detail.	[05]	3	
Q - 4	Attempt any one.	[05]		
(i)	When and Why Histogram and Bar Graph Chart is used ?		2	2
(ii)	Explain Data Types in Details.		2	1

CO : Course Outcome Number

BTL : Bloom's Taxonomy Level

Level of Bloom's Revised Taxonomy in Assessment

1: Remember	2: Understand	3: Apply
4: Analyze	5: Evaluate	6: Create