CBCS SCHEME

USN | | | | | | | | | | |

BDS613B

Sixth Semester B.E./B.Tech. Degree Examination, June/July 2025
Exploratory Data Analysis

Time: 3 hrs. Max. Marks: 100

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M : Marks , L: Bloom's level , C: Course outcomes.

		Module – 1	M	L	C
Q.1	a.	Develop a python code to demonstrate the concept of the following basic numpy operations on one dimensional and 2 dimensional data: (i) Indexing (ii) Slicing (iii) Joining	10	L3	CO1
	b.	List and explain the various numpy functions used for sorting of arrays.	10	L2	CO1
		OR			
Q.2	a.	Explain how Numpy handles structured data with examples.	10	L2	CO1
	b.	With python program explain the following basic array manipulations: (i) Splitting (ii) Reshapping (iii) Attributes of the array	10	L3	CO1
		Module – 2			
Q.3	a.	Develop python code for constructing the Pandas Data Frame objects from the following: (i) Single series object (ii) List of dicts (iii) Dictionary of series objects (iv) Two-dimensional Numpy array	10	L2	CO1
	b.	List the methods used to operate on Null values and Also demonstrate the use of these methods to handle the null values present in the data with examples.	10	L2	CO2
		OR			
Q.4	a.	Develop python code for constructing the Pandas series objects from list and dictionary.	10	L3	CO2
	b.	Explain hierarchical indexing in Pandas. Also write a python program to create a multi indexed Data frame and perform data aggregation.	10	L3	CO2
		Module – 3			
Q.5	a.	Develop python code to create a vectorized string and do the following: (i) Apply pandas string methods on string (ii) Apply pandas string methods along with regular expression on string	10	L3	CO3
	b.	Explain High performance Pandas eval( ) and query( ) functions with examples.	10	L2	CO3
		OR			
Q.6	a.	Develop python program to demonstrate the arithmetic, comparison and bitwise operation using pandas.eval( ) function.	10	L3	CO3
	b.	How does pandas handle time-series data? Provide examples.	10	L2	CO3

VTUCircle

		Module – 4			
Q.7	**a.**	Develop python code for simple line plot with specific line color, line style, Axes limits and Labeling.	**10**	**L3**	**CO4**
	b.	Explain the general tips for creating virtualizations with matplotlib.	**10**	**L2**	**CO4**
		OR			
Q.8	**a.**	With python program explain simple scatter plot using plot( ) and scatter( ) function of matplotlib.	**10**	**L3**	**CO4**
	b.	With example illustrate the creation of histogram, KDE and pair plot using Seaborn.	**10**	**L2**	**CO4**
		Module – 5			
Q.9	**a.**	Define machine learning and explain the categories of machine learning with examples.	**10**	**L2**	**CO5**
	b.	Write a python program to demonstrate the use of Scikit-Learn for simple linear regression.	**10**	**L3**	**CO4**
		OR			
Q.10	**a.**	How is model validation performed in Scikit-Learn?	**05**	**L2**	**CO5**
	b.	Explain the process of splitting datasets into training and testing sets with example.	**05**	**L2**	**CO5**
	c.	Discuss the steps involved in building a machine learning pipeline in Scikit-Learn.	**10**	**L2**	**CO5**

* * * * *

VtuCircle