

GUJARAT TECHNOLOGICAL UNIVERSITY
BE- SEMESTER-V (NEW) EXAMINATION – WINTER 2020

Subject Code:3150713**Date:22/01/2021****Subject Name:Python for Data Science****Time:10:30 AM TO 12:30 PM****Total Marks: 56****Instructions:**

1. Attempt any FOUR questions out of EIGHT questions.
2. Make suitable assumptions wherever necessary.
3. Figures to the right indicate full marks.

		MARKS
Q.1	(a) Discuss the role of indentation in python.	03
	(b) Explain range() function with suitable examples.	04
	(c) Write a python program to find the factorial of a given number using recursion.	07
Q.2	(a) Explain sampling in terms of data science?	03
	(b) List and explain different coding styles supported by python.	04
	(c) Discuss why python is a first choice for data scientists?	07
Q.3	(a) Explain TF-IDF transformations.	03
	(b) Explain categorical variables in detail.	04
	(c) Write a python program to read the data from XML file using pandas library.	07
Q.4	(a) Describe date time transformation using datetime module.	03
	(b) Explain a bag of words model in detail.	04
	(c) Explain imputation in detail with example.	07
Q.5	(a) List the features of matplotlib.	03
	(b) Write a python program to read data from a text file using pandas library.	04
	(c) Explain time series plot with appropriate examples.	07
Q.6	(a) List the type of plots that can be drawn using matplotlib.	03
	(b) Write a python program to read data from CSV files using pandas.	04
	(c) Explain pie chart plot with appropriate examples.	07
Q.7	(a) List and explain interfaces of SciKit-learn.	03
	(b) List the multiprocessing tasks that can be done using SciKit-learn?	04
	(c) Define the classification problem. How can it be solved using	07

SciKit-learn?

- Q.8** (a) Define EDA. List the tasks need to be carried out in EDA? **03**
- (b) How hash functions can be useful to solve data science problems? **04**
- (c) Define the regression problem. How can it be solved using SciKit-learn? **07**
