Enrolment No./Seat No_____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE- SEMESTER– V EXAMINATION – SUMMER 2026

Subject Code:3150713 Date:21-05-2026

Subject Name:Python for Data Science

Time:02:30 PM TO 05:00 PM Total Marks:70

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q. 1	**(a)**	Discuss the role of indentation in python	**3**
	(b)	What is the role of Python in Data science?	**4**
	(c)	Write a python program to read data from CSV files using pandas and Explain any two operations on data.	**7**
Q. 2	**(a)**	Explain Slicing rows and columns with example	**3**
	(b)	List and explain any three Magic functions	**4**
	(c)	What is the use of hash function in EDA? Express various hashing trick along with example	**7**
		OR	
	(c)	What do you mean by Exploratory Data Analysis (EDA)? How t-test is useful for EDA?	**7**
Q. 3	**(a)**	Explain join and merge functions in pandas with Example	**3**
	(b)	Write a program to perform the following operation about Function 1. Function with multiple return values 2. Default arguments	**4**
	(c)	What kind of data is analyzed with Bag of word model? Explain it with example.	**7**
		OR	
Q. 3	**(a)**	Explain Z-score standardization	**3**
	(b)	Write python program to demonstrate Tuple, list and Dictionary	**4**
	(c)	Define the term Data wrangling. Explain the steps needed to perform data wrangling	**7**
Q. 4	**(a)**	Explain hist() and bar()function with code	**3**
	(b)	Write a python program to perform the following operation on Array (Create a single Program to demonstrate all) a. Create b. Insert Element c. Print Array d. Slice Array	**4**
	(c)	Write a python program to demonstrate the concept of skewness and kurtosis	**7**
		OR	
Q. 4	**(a)**	Explain any three functions from Scikit learn	**3**
	(b)	Define the classification problem. How can it be solved using SciKit-learn?	**4**
	(c)	List various types of graph/chart available in the pyplot of matplotlib library for data visualization. Explain any two of them in brief	**7**
Q. 5	**(a)**	Write python program to find area of circle (user defined function)	**3**
	(b)	Define the regression problem. How can it be solved using SciKit-learn?	**4**
	(c)	Explain Labels, Annotation and Legends in MatPlotLib	**7**
		OR	
Q. 5	**(a)**	Differentiate Supervised and Unsupervised learning	**3**
	(b)	Write a brief note on NetworkX library	**4**
	(c)	What are the different ways to remove duplicate values from dataset? Explain with Example	**7**

****

Enrolment No./Seat No_____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–V EXAMINATION – SUMMER 2025

Subject Code:3150713 Date:13-05-2025

Subject Name:Python for Data Science

Time:02:30 PM TO 05:00 PM Total Marks:70

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q.1	**(a)**	Explain range() function of python with suitable example.	**03**
	(b)	Write a short note on List.	**04**
	(c)	Write a python program to display the Fibonacci series. Number of terms will be inputted by user.	**07**
Q.2	**(a)**	Why data visualization is important in Data Science?	**03**
	(b)	List and Explain advantages & disadvantages of Python in context of Data Science.	**04**
	(c)	Explain different stages of the data science pipeline.	**07**
		OR	
	(c)	Explain HTML parsing using Beautiful soup with suitable program.	**07**
Q.3	**(a)**	Differentiate NumPy and pandas.	**03**
	(b)	Write a program in Python to read .CSV file named “PDS. csv”. Display the contents of Column “Grade”. File contents are as follows, Roll No Grade 101 AA 102 BB 103 AB	**04**
	(c)	Explain use of categorical variable with the help of suitable program.	**07**
		OR	
Q.3	**(a)**	Write a program to read and display a text file named “PDS.txt”.	**03**
	(b)	Explain how to deal with the missing data.	**04**
	(c)	Explain time transformation with suitable program.	**07**
Q.4	**(a)**	Explain use of Labels, Annotations and Legends with suitable example.	**03**
	(b)	Explain pie chart with suitable example.	**04**
	(c)	Explain Bag of words model.	**07**
		OR	
Q.4	**(a)**	Explain histograms with suitable example.	**03**
	(b)	Explain bar chart with suitable example.	**04**
	(c)	Explain TF-IDF transformations.	**07**
Q.5	**(a)**	Explain of covariance and correlation in EDA.	**03**
	(b)	Explain skewness and kurtosis.	**04**

	(c)	What is the use of hashing trick and hash function in Scikit-learn? Explain in detail with suitable example.	**07**

OR

Q.5	**(a)**	What is the use of %timeit and %%timeit magic functions?	**03**
	(b)	Explain Z-score standardization.	**04**
	(c)	Explain EDA approach in detail.	**07**

****

Enrolment No./Seat No_____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–V (NEW) EXAMINATION – SUMMER 2024

Subject Code:3150713 Date:16-05-2024

Subject Name:Python for Data Science

Time:02:30 PM TO 05:00 PM Total Marks:70

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q.1	**(a)**	Briefly discuss major coding styles supported by python programming language.	**03**
	(b)	What will be the output of the below python code? import numpy as np a = np.array([[1, 2.5, 3],[4, 5, 6.5]]) print(a.shape) print(a.ndim) print(a.itemsize) print(a.nbytes)	**04**
	(c)	Write a python program to take two strings and position value (starting from zero) as input from the user. Insert the second string in the first string after the given position. (Example: String-1 = "Gujarat University", String-2="Technological ", position = 7, Output="Gujarat Technological University")	**07**
Q.2	**(a)**	Discuss characteristics of Set and Dictionary data structures in python.	**03**
	(b)	Write a short note on TF-IDF transformations.	**04**
	(c)	Explain different stages of the data science pipeline.	**07**
		OR	
	(c)	Discuss major libraries available with python language with its application in the field of Data Science.	**07**
Q.3	**(a)**	Explain different data structures available in pandas library.	**03**
	(b)	Below is the content of a csv file named "data.csv". Name,Salary abc,10000 xyz,20000 pqr,40000 Write a python program to read data from given csv file and print each pair of names and salaries in the format "Name -> Salary". (Example: abc -> 10000 and so on)	**04**
	(c)	Why it is important to find missing values from a dataset? Discuss different approaches to handle missing values with appropriate examples.	**07**
		OR	
Q.3	**(a)**	Write a short note on categorical variables.	**03**
	(b)	Consider below dataframe in python	**04**

A B C

0	7	1	4
1	2	5	8
2	3	6	9

i. Write a code to sort this dataframe by column A.
ii. Write a code to sort this dataframe by row 0 (first row).

(c) Explain the process of Word Tokenization, Stemming and Stop word removal using nltk library with suitable example code. **07**

Q.4 **(a)** Discuss how to create pie chart using matplotlib library with suitable example. **03**

(b) Demonstrate how to remove a row and a column from a dataframe with appropriate example code. **04**

(c) Data of marks for five students in three subjects is as given below. **07**

Maths = [100, 82, 76, 51, 88]
Physics = [92, 89, 81, 83, 80]
Chemistry = [71, 90, 75, 88, 45]

Write a python program to create boxplot from the above data. What type of important observation about data can be derived from the boxplot?

OR

Q.4 **(a)** Explain how to add markers and change line color and line style in a line graph using matplotlib library. **03**

(b) An adjacency matrix for a graph with for nodes named 0, 1, 2, 3 is given below. **04**

	0	1	2	3
0	0	1	1	0
1	0	0	0	0
2	0	1	0	1
3	1	0	0	0

Write a python program to draw the undirected graph from above matrix using networkx library.

(c) Below is the data about average temperature (in °C) at a place for 15 days. **07**

temp = [30, 34, 35, 32, 38, 26, 29, 45, 42, 32, 40, 33, 36, 34, 36]

Write a python program to create histogram with 4 bins from the above data. What type of important observation about data can be derived from the histogram?

Q.5 **(a)** What is the use of %timeit and %%timeit magic functions? **03**

(b) Discuss importance of covariance and correlation in EDA. **04**

(c) What is regression and classification problems? Explain any one with suitable example. **07**

OR

Q.5 **(a)** Explain different interfaces from Scikit-learn library. **03**

(b) Explain the importance of chi-square test in EDA. **04**

(c) Write a note on different descriptive statistics measures for numeric data. **07**

****

Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE – SEMESTER- V EXAMINATION-SUMMER 2023

Subject Code: 3150713 Date: 23/06/2023

Subject Name: Python for Data Science

Time: 02:30 PM TO 05:00 PM Total Marks: 70

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q.1	**(a)**	Write a single line code to get the value of "type" from the given dictionary in such a way that it does not produce any error or exception even if any key from the dictionary is misspelled. e.g. batters is misspelled as bateers. Still, your code must traverse the dictionary and fetch the value "Regular" of the key "type".	**03**

```
{
   "batters": {
      "batter": [
         {
            "batter": [
               {
                  "batter": [{
                        "type": "Regular"
                  }]
               }]
         }]
   }
}
```

	(b)	What is chi-square test? why it is necessary in data analysis?	**04**
	(c)	Explain following string functions with suitable example. len, count, title, lower, upper, find, rfine, replace	**07**
Q.2	**(a)**	List and explain the reasons which make python programming popular in Data Science.	**03**
	(b)	Differentiate: Dictionary and List	**04**
	(c)	What do you mean by Exploratory Data Analysis? List and explain the task which needs to be performed in EDA.	**07**
		OR	
	(c)	Define Standardization. Explain Z-score standardization with suitable example.	**07**
Q.3	**(a)**	Depict steps to create a scatter plot with example.	**03**
	(b)	Establish relationship between AI, data science and big data.	**04**
	(c)	Provide explanations on the importance of Graphs in Data Science.	**07**
		OR	
Q.3	**(a)**	Define correlation and explain its importance in Data Science.	**03**
	(b)	Provide duties performed by a Data Scientist with suitable example	**04**
	(c)	Elaborate Graphs along with its types.	**07**
Q.4	**(a)**	Differentiate: Bar graph vs. Histogram	**03**
	(b)	Explain Training and Testing with suitable example.	**04**
	(c)	Elaborate a bag of word concept in detail.	**07**

OR

Q.4	**(a)**	Why data visualization is important in Data Science?	**03**
	(b)	Explain HTML parsing using Beautiful soup.	**04**
	(c)	Explain stemming in detail with relatable example.	**07**
Q.5	**(a)**	Provide your views on Data wrangling with suitable example.	**03**
	(b)	Define covariance and explain its importance with appropriate example.	**04**
	(c)	Elaborate XPath in detail with relatable example.	**07**

OR

Q.5	**(a)**	Explain Hashing Tricks and its importance with suitable example.	**03**
	(b)	Explain importance of Legends, Labels and Annotations in Graphs.	**04**
	(c)	Describe sampling along with its types in detail with suitable example.	**07**

****

Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–V(NEW) EXAMINATION – SUMMER 2022

Subject Code:3150713 Date:02/06/2022

Subject Name:Python for Data Science

Time:02:30 PM TO 05:00 PM Total Marks: 70

Instructions:

1. **Attempt all questions.**
2. **Make suitable assumptions wherever necessary.**
3. **Figures to the right indicate full marks.**
4. **Simple and non-programmable scientific calculators are allowed.**

			MARKS
Q.1	**(a)**	List Advantages of Python.	**03**
	(b)	Differentiate Numpy and Pandas.	**04**
	(c)	Explain Exploratory Data Analysis (EDA).	**07**
Q.2	**(a)**	Explain String Slicing in python with Example.	**03**
	(b)	List and Explain different programming styles in python.	**04**
	(c)	Write a program to check whether the given number is prime or not.	**07**
		OR	
	(c)	Write a program to print Fibonacci series up to number given by user.	**07**
Q.3	**(a)**	Differentiate rand and randn function in Numpy.	**03**
	(b)	Explain DataFrame in Pandas with example.	**04**
	(c)	Write a program to print following patterns.	**07**

1)

```
*
* *
* * *
* * * *
```

2)

```
$ $ $ $
$ $ $
$ $
$
```

3)

```
# # # # #
 # # #
  #
 # # #
# # # # #
```

OR

Q.3	**(a)**	Explain Groupby function in pandas with example.	**03**
	(b)	Explain how to deal with missing data in Pandas.	**04**
	(c)	Explain Web Scrapping with Example using Beautiful Soup library.	**07**
Q.4	**(a)**	Explain Bag of Word model.	**03**
	(b)	Differentiate join and merge functions in pandas.	**04**
	(c)	Write a program which takes 2 digits, X,Y as input and generates a 2-dimensional array of size X * Y. The element value in the i-th row and j-th column of the array should be i*j.	**07**

OR

Q.4	**(a)**	Explain Hashing Trick in python with example.	**03**
	(b)	Write a brief note on NetworkX library.	**04**
	(c)	List and Explain different graphs in MatPlotLib.	**07**
Q.5	**(a)**	Explain Labels, Annotation and Legends in MatPlotLib.	**03**
	(b)	Differentiate Supervised and Unsupervised learning.	**04**
	(c)	Explain Regression with example.	**07**

OR

Q.5	**(a)**	Write a program to print Current date and time.	**03**
	(b)	Write a program to interchange the List elements on two positions entered by a user	**04**
	(c)	Explain Classification with example.	**07**

****