Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE- SEMESTER–V EXAMINATION – WINTER 2025

Subject Code:3150713 Date:17-11-2025

Subject Name:Python for Data Science

Time:10:30 AM TO 01:00 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)**	What are the advantages of python as a preferred programming language for data science?	**03**
	(b)	Write a program that takes a list of words as input and returns the longest word in the list.	**04**
	(c)	In EDA, why is it essential to analyze both covariance and correlation when exploring relationships between variables? Provide examples of scenarios where covariance might indicate a strong relationship while correlation might suggest otherwise.	**07**
Q.2	**(a)**	Describe the difference between a list and a tuple in Python (Provide at-least three points).	**03**
	(b)	Implement a function that takes a string as input and returns the count of each character in the string as a dictionary.	**04**
	(c)	Explain Data Science Pipeline in detail.	**07**
		OR	
	(c)	Write a note on python ecosystem for data science.	**07**
Q.3	**(a)**	What are the advantages of Jupyter Notebook?	**03**
	(b)	Explain how one-hot encoding is performed on a given text snippet, such as "The quick brown fox jumps over the lazy dog." What specific steps are involved in transforming this text into one-hot encoded vectors?	**04**
	(c)	Describe the process of parsing an HTML document using the Beautiful Soup library in Python.	**07**
		OR	
Q.3	**(a)**	Discuss the advantages of pandas library in python.	**03**
	(b)	Explain the process of calculating the TF-IDF score for a term in a document with proper example.	**04**
	(c)	Discuss different options to deal with missing data as part of data cleaning process.	**07**
Q.4	**(a)**	Discuss importance of pie chart and box plot in data visualization.	**03**
	(b)	What are magic functions? Briefly explain any two magic functions.	**04**
	(c)	Write code to create an undirected graph of six nodes using networkx and matplotlib.	**07**
		OR	
Q.4	**(a)**	How can we add labels, annotations and legends in our graphs using matplotlib?	**03**
	(b)	Given the sample CSV file "trigonometry_data.csv" containing trigonometry data (Angle (degrees), Sine, Cosine, Tangent), how would	**04**

you use Pandas to read the file in DataFrame and print the angle in degrees and its corresponding cosine value from the DataFrame?
Content of "trigonometry_data.csv" file is as below.

```
Angle (degrees), Sine, Cosine, Tangent
0,0.0,1.0,0.0
30,0.5,0.866,0.577
45,0.707,0.707,1.0
60,0.866,0.5,1.732
90,1.0,0.0,undefined
```

(c) What is the importance of histogram in data visualization? Write code to create histogram using matplotlib. **07**

Q.5 **(a)** What role does the chi-square test play in data analysis? **03**

(b) Briefly discuss skewness and kurtosis. **04**

(c) Describe various methods for slicing rows and columns in a Pandas DataFrame, supported by relevant examples. **07**

OR

Q.5 **(a)** What is z-score standardization and why it is useful in EDA? **03**

(b) Explain working of %timeit and %memit magic functions. **04**

(c) Elaborate on the concepts of sorting data row-wise, column-wise, and shuffling data within a DataFrame by providing appropriate examples. **07**

****

Enrolment No./Seat No_____

GUJARAT TECHNOLOGICAL UNIVERSITY

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

Subject Code:3150713 Date:25-11-2024

Subject Name:Python for Data Science

Time:10:30 AM TO 01:00 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 with suitable examples	**3**
	(b)	Differentiate List and Tuple in Python	**4**
	(c)	Write a python program to read data from CSV files using pandas and Explain Slicing operation using appropriate examples.	**7**
Q. 2	**(a)**	What do you mean by slicing operation in string of python? Write an example of slicing to fetch first name and last name from full name of person and display it	**3**
	(b)	Explain data science pipeline in details	**4**
	(c)	Define term n-gram. Explain the TF-IDF techniques	**7**
		OR	
	(c)	What is the use of following operations on Panda's Data Frames? Explain with a small example of each. shape 2. tail() 3. describe()	**7**
Q. 3	**(a)**	Explain Groupby function 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)**	Why we need to perform Z-score standardization in EDA? Justify it with example	**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)**	Define covariance and correlation	**3**
	(b)	Define stemming. Explain the concept of stemming with example	**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)	Explain pie chart plot with appropriate examples	**7**
Q. 5	**(a)**	Explain Labels, Annotation and Legends in MatPlotLib	**3**
	(b)	Explain Regression with example	**4**

	(c)	Explain Hashing Trick in python with example	**7**
		OR	
Q. 5	**(a)**	Differentiate Supervised and Unsupervised learning	**3**
	(b)	Write a brief note on NetworkX library	**4**
	(c)	What do you mean by missing values? Explain the different ways to handle the missing value with example	**7**

****

Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

BE - SEMESTER–V (NEW) EXAMINATION – WINTER 2023

Subject Code:3150713 Date:05-12-2023

Subject Name: Python for Data Science

Time:10:30 AM TO 01:00 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)**	Why python is very popular language in 21st Century?	**03**
	(b)	Write any four-python command that demonstrate Python as calculator.	**04**
	(c)	Write a python program to display the Fibonacci sequence up to nth term.	**07**
Q.2	**(a)**	Explain with example use of range () function in python.	**03**
	(b)	Explain the phases of Data Science Pipeline.	**04**
	(c)	List the different data structure used in python. Explain any three with example of each.	**07**
		OR	
	(c)	Write a Python program to input the string from user and count number of Vowel in given string.	**07**
Q.3	**(a)**	What is the use of magic function in python? Explain with the help of example.	**03**
	(b)	Explain the use of following libraries of python in data science. i. Matplotlib ii. Scikit-learn	**04**
	(c)	Write a python code to read the local image, obtain statistics about that image, and display the image onscreen.	**07**
		OR	
Q.3	**(a)**	Write a python code to print current date and time.	**03**
	(b)	Write the difference between numpy and pandas.	**04**
	(c)	Describe the three levels of flat-file dataset. Explain any two with the help of example.	**07**
Q.4	**(a)**	List the different line styles used in charts.	**03**
	(b)	Explain with the example use of axes () and grid () method in plots.	**04**
	(c)	Write a python code to draw two lines on plot. Both lines should be depicted with different colors.	**07**
		OR	
Q.4	**(a)**	What is the use of %matplotlib inline command?	**03**
	(b)	What is the use of labels and legends in plot? Explain with example.	**04**

	(c)	Write a python code to draw a bar chart utilizing at least three properties of it.	**07**
Q.5	**(a)**	Explain Classification and clustering class of Scikit-learn.	**03**
	(b)	Illustrate the concept of regression class of Scikit-learn with the help of small example.	**04**
	(c)	What is the use of hashing trick and hash function in Scikit-learn? Explain in detail with example.	**07**

OR

Q.5	**(a)**	What do you mean by categorical data? Explain with small example.	**03**
	(b)	Explain the use of skew () and kurtosis () function.	**04**
	(c)	Explain EDA approach in detail.	**07**

****

Seat No.: _____ Enrolment No._____

GUJARAT TECHNOLOGICAL UNIVERSITY

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

Subject Code:3150713 Date:04-01-2023

Subject Name:Python for Data Science

Time:10:30 AM TO 01:00 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.**

Q.1	**(a)**	What is the role of python in data science?	**03**
	(b)	Explain the input function of python that demonstrates type casting.	**04**
	(c)	Explain following data structures of python with suitable example. 1. String 2. List 3. Tuple 4. Dictionary	**07**
Q.2	**(a)**	Differentiate: C and Python.	**03**
	(b)	How to format Date and Time in python. Explain it with example.	**04**
	(c)	Give comparison between Numpy and Pandas.	**07**
		OR	
	(c)	Write a python code to read data from text file.	**07**
Q.3	**(a)**	Explain %matplotlib magic function.	**03**
	(b)	Explain stemming and stop words removal operation in python.	**04**
	(c)	Write a python program to implement Fibonacci sequence for given input.	**07**
		OR	
Q.3	**(a)**	What are the magic functions in Jupyter? Explain with example.	**03**
	(b)	With example explain the concept of bags of words model.	**04**
	(c)	Write a python program that finds the factorial of a natural number n.	**07**
Q.4	**(a)**	Explain labels, annotations and legends.	**03**
	(b)	Explain with example how to parse XML and HTML.	**04**
	(c)	Write a python code that demonstrate hashing trick.	**07**
		OR	
Q.4	**(a)**	Explain any three functions from Scikit learn.	**03**
	(b)	Explain how to create data science pipeline.	**04**
	(c)	Write a python program to demonstrate the concept of skewness and kurtosis.	**07**
Q.5	**(a)**	Explain EDA in detail.	**03**
	(b)	Write a python code to access data from web.	**04**
	(c)	Write a small code to perform following operations on data: Slicing, Dicing, Concatenation, Transformation.	**07**
		OR	
Q.5	**(a)**	Explain Z-score standardization.	**03**
	(b)	How to Obtain online graphics and multimedia. Explain with example.	**04**
	(c)	Write a code to draw pie chart using python's library.	**07**

****