GUJARAT TECHNOLOGICAL UNIVERSITY

BE 5 SEMESTER – OLD PAPER – S22 TO S25 – QUESTION BANK

Subject Name & Code: Python for Data Science - 3150713

Chapter 1: Overview of Python and Data Structures

Repeated Questions:

1. Python Features and Basics (Theory):

- **Q.** Why Python is popular / List advantages of Python. **(SUMMER 2022, 3 Marks; WINTER 2023, 3 Marks)**
- **Q.** List and explain different programming/coding styles supported by Python. **(SUMMER 2022, 4 Marks; SUMMER 2024, 3 Marks)**
- **Q.** What is the use of magic functions in Python/Jupyter? Explain with example. **(WINTER 2022, 3 Marks OR; WINTER 2023, 4 Marks)**

2. Data Structures and Strings (Theory):

- **Q.** Explain/List different data structures used in Python and explain any three/four with suitable examples. **(WINTER 2022, 7 Marks; WINTER 2023, 7 Marks)**
- **Q.** Explain String Slicing in Python with example. **(SUMMER 2022, 3 Marks; WINTER 2024, 3 Marks)**

3. Built-in Functions:

- **Q.** Explain range() function of Python with suitable example. **(WINTER 2023, 3 Marks; WINTER 2024, 3 Marks; SUMMER 2025, 3 Marks)**
- **Q.** Write a Python program/code to print current date and time. **(SUMMER 2022, 3 Marks OR; WINTER 2023, 3 Marks OR)**

4. Programs:

- **Q.** Write a Python program to display the Fibonacci sequence / Fibonacci series. **(SUMMER 2022, 7 Marks OR; WINTER 2022, 7 Marks; WINTER 2023, 7 Marks; SUMMER 2025, 7 Marks)**
- **Q.** Write a Python code/program to read data from a text file. **(WINTER 2022, 7 Marks OR; SUMMER 2025, 3 Marks OR)**

Other Important Questions:

1. Python Fundamentals (Theory):

- **Q.** Differentiate: C and Python. **(WINTER 2022, 3 Marks)**
- **Q.** Explain the input function of Python that demonstrates type casting. **(WINTER 2022, 4 Marks)**
- **Q.** How to format Date and Time in Python? Explain with example. **(WINTER 2022, 4 Marks)**
- **Q.** Write any four Python commands that demonstrate Python as calculator. **(WINTER 2023, 4 Marks)**

2. Data Structures (Theory):

- **Q.** Discuss characteristics of Set and Dictionary data structures in Python. **(SUMMER 2024, 3 Marks)**
- **Q.** Write a short note on List. **(SUMMER 2025, 4 Marks)**
- **Q.** Differentiate List and Tuple in Python. **(WINTER 2024, 4 Marks)**

3. Functions:

- **Q.** Write a program to perform the following operations about Function: 1. Function with multiple return values 2. Default arguments. **(WINTER 2024, 4 Marks)**

4. Programs: Strings and Dictionaries:

- **Q.** 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. Dictionary: {"batters": {"batter": [{"batter": [{"batter": [{"type": "Regular"}]}]}]}} **(SUMMER 2023, 3 Marks)**
- **Q.** 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. **(SUMMER 2024, 7 Marks)**
- **Q.** Write a Python program to input the string from user and count number of vowels in given string. **(WINTER 2023, 3 Marks)**

5. Programs: Numbers and Patterns (Numerical):

- **Q.** Write a Python program to find the factorial of a natural number n. **(WINTER 2022, 7 Marks OR)**
- **Q.** Write a program to check whether the given number is prime or not. **(SUMMER 2022, 7 Marks)**
- **Q.** Write a program to print the following patterns. (1) * / * * / * * * / * * * *; (2) $ $ $ / $ $ / $; (3) # # # # / # # # / # # # / # # # / # # # **(SUMMER 2022, 7 Marks)**
- **Q.** 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 should be i*j. **(SUMMER 2022, 7 Marks)**

6. Programs: Lists, Tuples and Dictionaries:

- **Q.** Write a program to interchange the List elements on two positions entered by a user. **(SUMMER 2022, 4 Marks OR)**
- **Q.** Write a Python program to demonstrate Tuple, List and Dictionary. **(WINTER 2024, 4 Marks OR)**

Chapter 2: Data Science and Python

Repeated Questions:

1. Data Science Pipeline and Role of Python:

- **Q.** Explain different stages/phases of the Data Science Pipeline / Explain data science pipeline in detail / How to create data science pipeline. **(WINTER 2022, 4 Marks OR; WINTER 2023, 4 Marks; SUMMER 2024, 7 Marks; WINTER 2024, 4 Marks; SUMMER 2025, 7 Marks)**
- **Q.** Explain the role/popularity/advantages of Python in Data Science. **(WINTER 2022, 3 Marks; SUMMER 2023, 3 Marks; SUMMER 2025, 4 Marks)**
- **Q.** Discuss major libraries available with Python language and their application in Data Science / Explain use of Matplotlib and Scikit-learn in Data Science. **(WINTER 2023, 4 Marks; SUMMER 2024, 7 Marks OR)**

2. NumPy and Pandas:

- **Q.** Differentiate/Compare NumPy and Pandas. **(SUMMER 2022, 4 Marks; WINTER 2022, 7 Marks; WINTER 2023, 4 Marks OR; SUMMER 2025, 3 Marks)**
- **Q.** Write a Python program to read data from a CSV file using pandas. [Diagram required – refer to the original question paper] **(SUMMER 2024, 4 Marks; WINTER 2024, 7 Marks; SUMMER 2025, 7 Marks)**
- **Q.** Explain Groupby function in pandas with example. **(SUMMER 2022, 3 Marks OR; WINTER 2024, 3 Marks)**

Other Important Questions:

1. NumPy:

- **Q.** 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) **(SUMMER 2024, 4 Marks)**
- **Q.** Differentiate rand and randn function in NumPy. **(SUMMER 2022, 3 Marks)**

2. Pandas DataFrames:

- **Q.** Explain different data structures available in pandas library. **(SUMMER 2024, 3 Marks)**
- **Q.** Explain DataFrame in Pandas with example. **(SUMMER 2022, 4 Marks)**
- **Q.** Consider the given dataframe. Write code to sort this dataframe by column A and by row 0 (first row). [Diagram required – refer to the original question paper] **(SUMMER 2024, 4 Marks OR)**
- **Q.** Demonstrate how to remove a row and a column from a dataframe with appropriate example code. **(SUMMER 2024, 4 Marks)**
- **Q.** Differentiate join and merge functions in pandas. **(SUMMER 2022, 4 Marks)**
- **Q.** What is the use of following operations on Pandas DataFrames? Explain with a small example of each: 1. shape 2. tail() 3. describe(). **(WINTER 2024, 7 Marks OR)**

3. Data Sources and File Formats:

- **Q.** Write a Python code to read the local image, obtain statistics about that image, and display the image onscreen. **(WINTER 2023, 7 Marks)**
- **Q.** Describe the three levels of flat-file dataset. Explain any two with the help of example. **(WINTER 2023, 7 Marks OR)**

Chapter 3: Visualizing the Data

Repeated Questions:

1. Importance of Visualization and Magic Functions:

- **Q.** Why data visualization is important in Data Science? **(SUMMER 2023, 3 Marks; SUMMER 2025, 3 Marks)**
- **Q.** Explain %matplotlib magic function / use of %matplotlib inline command. **(WINTER 2022, 3 Marks; WINTER 2023, 3 Marks OR)**

2. Chart Types:

- **Q.** Explain/create pie chart using matplotlib library with suitable example. **(WINTER 2022, 7 Marks OR; SUMMER 2024, 3 Marks; WINTER 2024, 7 Marks OR; SUMMER 2025, 4 Marks)**
- **Q.** Explain/create histogram with suitable example / create histogram with 4 bins from given data. [Diagram required – refer to the original question paper] **(SUMMER 2024, 7 Marks OR; SUMMER 2025, 3 Marks OR)**
- **Q.** Explain/draw bar chart with suitable example/utilizing at least three properties. **(WINTER 2023, 7 Marks OR; SUMMER 2025, 4 Marks OR)**

3. Customising Plots:

- **Q.** Explain importance of Labels, Annotations and Legends in Graphs/Matplotlib with example. **(SUMMER 2022, 3 Marks; WINTER 2022, 3 Marks; SUMMER 2023, 4 Marks OR; WINTER 2023, 4 Marks OR; WINTER 2024, 3 Marks; SUMMER 2025, 3 Marks)**
- **Q.** Explain line styles / markers / line color / draw lines with different colors in charts. **(WINTER 2023, 3 Marks; WINTER 2023, 7 Marks; SUMMER 2024, 3 Marks OR)**

4. NetworkX:

- **Q.** Write a brief note on NetworkX library / Draw graph using NetworkX / Draw undirected graph from adjacency matrix. [Diagram required – refer to the original question paper] **(SUMMER 2022, 4 Marks OR; SUMMER 2024, 4 Marks OR; WINTER 2024, 4 Marks OR)**

Other Important Questions:

1. Plot Types:

- **Q.** Depict steps to create a scatter plot with example. **(SUMMER 2023, 3 Marks)**
- **Q.** Write a Python program to create boxplot from given marks data. What important observation about data can be derived from the boxplot? [Diagram required – refer to the original question paper] **(SUMMER 2024, 7 Marks)**
- **Q.** List and explain different graphs in Matplotlib. **(SUMMER 2022, 7 Marks OR)**

2. Plot Customisation:

- **Q.** Explain with example use of axes() and grid() method in plots. **(WINTER 2023, 4 Marks)**

Chapter 4: Data Wrangling

Repeated Questions:

1. Data Wrangling and Missing Values:

- **Q.** Explain Data wrangling / Define Data wrangling and explain steps needed to perform data wrangling. **(SUMMER 2023, 3 Marks; WINTER 2024, 7 Marks OR)**
- **Q.** Why is it important to find/handle missing values? Discuss different approaches to handle missing values with examples. **(SUMMER 2022, 4 Marks OR; SUMMER 2024, 7 Marks; WINTER 2024, 7 Marks OR; SUMMER 2025, 4 Marks OR)**

2. Web Scraping and Parsing:

- **Q.** Explain HTML/XML parsing / Web Scraping using Beautiful Soup with example/program. **(SUMMER 2022, 7 Marks OR; WINTER 2022, 4 Marks; SUMMER 2023, 4 Marks; SUMMER 2025, 7 Marks OR)**

3. Text Processing:

- **Q.** Explain stemming / stemming and stop words removal / Word Tokenization, Stemming and Stop word removal using nltk. **(WINTER 2022, 4 Marks; SUMMER 2023, 7 Marks; SUMMER 2024, 7 Marks OR; WINTER 2024, 4 Marks)**
- **Q.** Explain TF-IDF transformations / Define n-gram and explain TF-IDF techniques. **(SUMMER 2024, 4 Marks; WINTER 2024, 7 Marks; SUMMER 2025, 7 Marks OR)**
- **Q.** Explain Bag of Words model / What kind of data is analyzed with Bag of Words model? **(SUMMER 2022, 3 Marks; WINTER 2022, 4 Marks OR; WINTER 2024, 7 Marks; SUMMER 2025, 7 Marks)**
- **Q.** Explain Hashing Trick / hashing trick and hash function in Scikit-learn with example. **(SUMMER 2022, 3 Marks OR; WINTER 2022, 7 Marks; SUMMER 2023, 3 Marks OR; WINTER 2023, 7 Marks; WINTER 2024, 7 Marks; SUMMER 2025, 7 Marks)**

4. Scikit-learn and Machine Learning:

- **Q.** Explain any three functions/interfaces from Scikit-learn library. **(WINTER 2022, 3 Marks OR; SUMMER 2024, 3 Marks OR; WINTER 2024, 3 Marks OR)**
- **Q.** Explain Regression and Classification problems / Regression class / Classification class of Scikit-learn with example. **(SUMMER 2022, 7 Marks; SUMMER 2022, 7 Marks OR; WINTER 2023, 3 Marks; WINTER 2023, 4 Marks; SUMMER 2024, 7 Marks; WINTER 2024, 4 Marks OR; WINTER 2024, 4 Marks)**
- **Q.** Differentiate Supervised and Unsupervised learning. **(SUMMER 2022, 4 Marks; WINTER 2024, 3 Marks OR)**

5. Magic Functions:

- **Q.** What is the use of %timeit and %%timeit magic functions? **(SUMMER 2024, 3 Marks; SUMMER 2025, 3 Marks OR)**

Other Important Questions:

1. Web Data Access:

- **Q.** Elaborate XPath in detail with relatable example. **(SUMMER 2023, 7 Marks)**
- **Q.** Write a Python code to access data from web. **(WINTER 2022, 4 Marks)**
- **Q.** How to obtain online graphics and multimedia? Explain with example. **(WINTER 2022, 4 Marks OR)**

2. Data Manipulation Programs:

- **Q.** Write a small code to perform following operations on data: Slicing, Dicing, Concatenation, Transformation. **(WINTER 2022, 7 Marks)**

- **Q.** Explain time transformation with suitable program. **(SUMMER 2025, 7 Marks OR)**

Chapter 5: Exploring Data Analysis (EDA)

Repeated Questions:

1. EDA Concepts:

- **Q.** Explain EDA / EDA approach in detail. **(SUMMER 2022, 7 Marks; WINTER 2022, 3 Marks; WINTER 2023, 7 Marks OR; SUMMER 2025, 7 Marks OR)**
- **Q.** Define/Explain covariance and correlation and its importance in EDA with example. **(SUMMER 2023, 4 Marks; SUMMER 2024, 4 Marks; WINTER 2024, 3 Marks; SUMMER 2025, 3 Marks)**

2. Statistical Measures:

- **Q.** Explain skewness and kurtosis / use of skew() and kurtosis() function / program to demonstrate skewness and kurtosis. **(WINTER 2022, 7 Marks OR; WINTER 2023, 4 Marks OR; WINTER 2024, 7 Marks; SUMMER 2025, 4 Marks)**
- **Q.** Explain Z-score standardization / Why we need to perform Z-score standardization in EDA? **(WINTER 2022, 3 Marks OR; WINTER 2024, 3 Marks OR; SUMMER 2025, 4 Marks OR)**

3. Categorical Data:

- **Q.** What do you mean by categorical data/variables? Explain with example/program. **(WINTER 2023, 3 Marks OR; SUMMER 2024, 3 Marks OR; SUMMER 2025, 7 Marks)**

Other Important Questions:

1. Sampling and Statistics:

- **Q.** Describe sampling along with its types in detail with suitable example. **(SUMMER 2023, 7 Marks OR)**
- **Q.** Explain the importance of chi-square test in EDA. **(SUMMER 2024, 4 Marks OR)**
- **Q.** Write a note on different descriptive statistics measures for numeric data. **(SUMMER 2024, 7 Marks OR)**

****