Question Bank - PDS

3 Marks:

1. Discuss the role of indentation in python. *[WINTER 2020]*
2. Explain sampling in terms of data science? *[WINTER 2020]*
3. Explain TF-IDF transformations. *[WINTER 2020]*
4. Describe date time transformation using datetime module. *[WINTER 2020]*
5. List the features of matplotlib. *[WINTER 2020]*
6. List the type of plots that can be drawn using matplotlib. *[WINTER 2020]*
7. List and explain interfaces of SciKit-learn. *[WINTER 2020]*
8. Define EDA. List the tasks need to be carried out in EDA? *[WINTER 2020]*
9. What is the role of Python in Data science? *[WINTER 2021]*
10. What is HTML parsing? *[WINTER 2021]*
11. List and explain any three Magic function. *[WINTER 2021]*
12. What is Categorical Variables? Explain it with example. *[WINTER 2021]*
13. Explain hist() function with code. *[WINTER 2021]*
14. Explain bar() function with code. *[WINTER 2021]*
15. What is Scikit-learn? *[WINTER 2021]*
16. Define covariance and correlation. *[WINTER 2021]*
17. Differentiate the list and dictionary data types of python by their characteristics along with example in brief. *[SUMMER 2021]*
18. What is the core competencies needed to become a data scientist? Explain in brief. *[SUMMER 2021]*
19. Compare the numpy and pandas on the basis of their characteristics and usage. *[SUMMER 2021]*
20. How XPath is useful for analysis of html data? Explain in brief. *[SUMMER 2021]*
21. List various types of graph/chart available in the pyplot of matplotlib library for data visualization. Explain any two of them in brief. *[SUMMER 2021]*
22. Compare bar graph, box-plot and histogram with respect to their applicability in data visualization. *[SUMMER 2021]*
23. Define the term Data wrangling. Explain the steps needed to perform data wrangling. *[SUMMER 2021]*
24. What do you mean by Exploratory Data Analysis (EDA)? How t-test is useful for EDA? *[SUMMER 2021]*

4 Marks:

1. Explain range() function with suitable examples. *[WINTER 2020]*
2. List and explain different coding styles supported by python. *[WINTER 2020]*
3. Explain categorical variables in detail. *[WINTER 2020]*
4. Explain a bag of words model in detail. *[WINTER 2020]*
5. Write a python program to read data from a text file using pandas. *[WINTER 2020]*
6. Write a python program to read data from CSV files using pandas. *[WINTER 2020]*
7. List the multiprocessing tasksthat can be done using SciKit-learn? *[WINTER 2020]*
8. How hash functions can be useful to solve data science problems? *[WINTER 2020]*
9. Differentiate List and Tuple in Python. *[WINTER 2021]*
10. Write a python code to find factorial of number using function. *[WINTER 2021]*
11. Explain Slicing rows and columns with example. *[WINTER 2021]*
12. How to read data from relational database? Briefly explain it. *[WINTER 2021]*
13. Write a program using Numpy to count number of "C" element wise in a given array. *[WINTER 2021]*
14. What are the different ways to remove duplicate values from dataset? *[WINTER 2021]*
15. Explain Box plot with example. *[WINTER 2021]*
16. Explain scatterplots with example. *[WINTER 2021]*
17. 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. *[SUMMER 2021]*
18. Compare and summarize four different coding styles supported by Python language. *[SUMMER 2021]*
19. For what purpose sampling is used. Demonstrate random sampling with example. *[SUMMER 2021]*
20. Define term n-gram. Explain the TF-IDF techniques. *[SUMMER 2021]*
21. What kind data is analyzed with Bag of word model? Explain it with example. *[SUMMER 2021]*
22. Define stemming. Explain the concept of stemming with example. *[SUMMER 2021]*
23. Why we need to perform Z-score standardization in EDA? Justify it with example. *[SUMMER 2021]*
24. What do you mean by covariance? What is the importance of covariance in data analysis? Explain it with example. *[SUMMER 2021]*

7 Marks:

1. Write a python program to find the factorial of a given number using recursion. *[WINTER 2020]*
2. Discuss why python is a first choice for data scientists? *[WINTER 2020]*
3. Write a python program to read the data from XML file using pandas library. *[WINTER 2020]*
4. Explain imputation in detail with example. *[WINTER 2020]*
5. Explain time series plot with appropriate examples. *[WINTER 2020]*
6. Explain pie chart plot with appropriate examples. *[WINTER 2020]*

7. Define the classification problem. How can it be solved using SciKit-learn? *[WINTER 2020]*
8. Define the regression problem. How can it be solved using SciKit-learn? *[WINTER 2020]*
9. Explain data science pipeline in details. *[WINTER 2021]*
10. Explain Dictionary in Python with example. *[WINTER 2021]*
11. Is String a mutable data type? Also explain the string operations length, indexing and slicing in detail with an appropriate example. *[WINTER 2021]*
12. What do you mean by missing values? Explain the different ways to handle the missing value with example. *[WINTER 2021]*
13. What is the use of following operations on Panda's Data Frames? Explain with a small example of each. 1. shape 2. tail() 3. describe() *[WINTER 2021]*
14. What do you understand by Data visualization? Discuss some Python's data visualization techniques. *[WINTER 2021]*
15. Write a simple python program that draws a line graph where x = [1,2,3,4] and y = [1,4,9,16] and gives both axis label as "X-axis"and "Y-axis". *[WINTER 2021]*
16. Write a Python programming to create a pie chart with a title of the popularity of programming Languages. *[WINTER 2021]*
    Sample data:
    Programming languages: Java, Python, PHP, JavaScript, C#, C++
    Popularity: 22.2, 17.6, 8.8, 8, 7.7, 6.7
17. What is Data Wrangling process? Define data exploratory data analysis? Why EDA is required in data analysis? *[WINTER 2021]*
18. Which are the basic activities we performed as a part of data science pipeline? Summarize and explain in brief. *[SUMMER 2021]*
19. Summarize the characteristics of NumPy, Pandas, Scikit-Learn and matplotlib libraries along with their usage in brief. *[SUMMER 2021]*
20. What do you mean by prototyping? List the phases of prototyping and experimentation process and explain in brief. *[SUMMER 2021]*
21. What is the need of streaming the data? Explain data uploading and streaming data with example. *[SUMMER 2021]*
22. List the techniques to handle missing data. Explain various techniques with example.*[SUMMER 2021]*
23. What do you mean by time series data? How can we plot it? Explain it with example to plot trend over time. *[SUMMER 2021]*
24. What is the use of scatter-plot in data visualization? Can we draw trendline in scatter-plot? Explain it with example. *[SUMMER 2021]*
25. What is the use of hash function in EDA? Express various hashing trick along with example. *[SUMMER 2021]*
26. List different way for defining descriptive statistics for Numeric Data. Explain them in brief. *[SUMMER 2021]*