What are some reasons why the average number of comments per user would be decreasing and what metrics would you look into ?

Questions :

(i) What will be your clarifying questions specific to the case ?

(ii) How will you make assumptions using the case question ?

(iii) Make a hypothesis about the data.

Discuss your data analysis process.

H Roll No. .....

MB–404(BA3)

M. B. A. (FOURTH SEMESTER)

END SEMESTER

EXAMINATION, May, 2023

DATA SCIENCE WITH PYTHON

Time : Three Hours

Maximum Marks : 100

Note : (i) This question paper contains two Sections—Section A and Section B.

(ii) Both sections are compulsory.

(iii) Answer any *two* sub-questions among (a), (b) and (c) in each main question of Section A. Each question carries 10 marks.

(iv) Section B consisting of case study is compulsory. Section B is of 20 marks.

P. T. O.

Section—A

1. (a) Define Data Science. Discuss the role of Python in Data Science. (CO1)

   (b) Assess the basic skills, roles and responsibilities of a data scientist. (CO4)

   (c) Outline various applications of Data Science and the role of big data in Data Science. (CO3)

2. (a) Recommend various plots for visualization according to their data types. (CO4)

   (b) Describe the hypothesis test and also discuss the level of confidence and significance. (CO1)

   (c) Demonstrate descriptive, diagnostic, predictive and prescriptive analysis with an example. (CO2)

3. (a) Explain various features of the Anaconda and JupyterNotebook for analysis. (CO1)

   (b) How to construct clusters using cluster analysis ? Discuss hierarchical and k-means clustering. (CO5)

   (c) Discuss the concept of multiple regression analysis. Also, explain the concept of multicollinearity. (CO1)

4. (a) Compare various data structures supported by Python. (CO3)

   (b) What is the use of data visualization ? discuss different plots with their advantages. (CO2)

   (c) Assess how multiple scientific domain scan be managed using Scipy package in Python. (CO5)

Section—B

5. **Case Study :** (20 Marks) (CO5)

Let's say you work for a social media company that has just done a launch in a new city. Looking at weekly metrics, you see a slow decrease in the average number of comments per user from January to March in this city. The company has been consistently growing new users in the city from January to March.