Graphic Era HILL UNIVERSITY
Established by an Act of the State Legislature of Uttarakhand (Adhiniyam Sankhya 12 of 2011)
University under section 2(f) of UGC Act, 1956

End Term (Even) Semester Examination May-June 2025

Roll no.....

Name of the Program and semester: II MBA
Name of the Course: ***FUNDAMENTALS OF BUSINESS ANALYTICS***
Course Code: MBA 205(BA1)
Time: 3 hour Maximum Marks: 100

Note:

(i) This question paper contains two Sections-Section A and B
(ii) Both Sections are compulsory
(iii) Answer any two sub questions from a, b & c in each main question of Section A. Each sub question carries 10 marks.
(iv) Section B, consisting of a case study, is compulsory. It is of 20 Marks.

Section A

Q1. (2X10=20 Marks)

a. Discuss the data collection methods and data formats used in Business Analytics. How does data preprocessing improve analysis? CO1, CO2

b. Describe the scope of Business Analytics in today's digital economy and its impact on decision-making. CO3

c. Explain Descriptive, Diagnostic, Predictive and Prescriptive analytics with example. CO1, CO2

Q2. (2X10=20 Marks)

a. Discuss the steps involved in Exploratory Data Analysis (EDA) for a dataset. How do descriptive statistics and visualization techniques complement each other? CO2, CO4

b. Name three Python libraries used for data visualization and their primary functions. CO5

c. What are outliers in a dataset? How can they be detected using visualization? CO3, CO4

Q3. (2X10=20 Marks)

a. Define probability and explain its role in business decision-making with an example. Differentiate between discrete and continuous probability distributions with one example each. CO4

b. Distinguish between ANOVA and MANOVA with an example. CO3

c. Define simple linear regression and write its mathematical equation. List and explain four key assumptions of linear regression. CO2

Q4. (2X10=20 Marks)

a. Distinguish between Supervised and Unsupervised learning. Explain its classification with example. CO4

b. What is distributed computing? How does it address Big Data scalability? CO5

c. Compare **Hadoop** and **Spark** in terms of: CO4, CO5
i) Processing speed
ii) Ease of use
iii) Fault tolerance
iv) Use case suitability

Graphic Era HILL UNIVERSITY
Established by an Act of the State Legislature of Uttarakhand (Adhiniyam Sankhya 12 of 2011)
University under section 2(f) of UGC Act, 1956

End Term (Even) Semester Examination May-June 2025

Section B

Q5. Case Study (20 Marks)

An online retailer wants to segment customers based on purchasing behavior for targeted marketing. The dataset includes:

Age, Gender, Income (numerical & categorical features)

Purchase Frequency, Average Order Value (continuous features)

Customer ID	Age	Gender	Income ($)	Purchase Frequency (per month)	Average Order Value ($)
1	25	Male	35,000	2	50
2	32	Female	42,000	4	75
3	45	Male	60,000	1	120
4	28	Female	38,000	3	60
5	50	Male	55,000	2	90
6	35	Female	48,000	5	45
7	22	Male	28,000	4	55
8	40	Female	65,000	1	150
9	30	Male	45,000	3	70
10	27	Female	33,000	2	65

Answer the following questions based on the case study:

End Term (Even) Semester Examination May-June 2025

a. Should the retailer use supervised or unsupervised learning? Justify your answer. CO5

b. Which algorithm would you recommend for customer segmentation? Explain its working. CO4

c. How would you pre-process the data before applying the algorithm? CO5

d. The retailer later wants to predict if a customer will make a purchase in the next month. Which supervised learning approach would you use? CO5