18026

BRAINWARE UNIVERSITY

Term End Examination 2025-2026
Programme – MBA-2024
Course Name – Fundamentals of Data Science
Course Code - MBA30202E
(Semester III)

Full Marks : 60 **Time : 2:30 Hours**

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A

(Multiple Choice Type Question) 1 x 15=15

1. *Choose the correct alternative from the following :*

(i) Select the primary goal of data science.
a) Create infographics
b) Extract knowledge from data
c) Build websites
d) Send emails

(ii) Identify the key feature of big data.
a) Small datasets
b) Low velocity
c) High variety, volume, and velocity
d) Manual data entry

(iii) Cite a key responsibility of a data scientist.
a) Filing reports
b) Cleaning the office
c) Extracting insights from data
d) Installing hardware

(iv) Select a common property of big data.
a) Simplicity
b) Structured only
c) Scalability
d) Static data

(v) Identify the term "datafication".
a) Turning ideas into graphs
b) Converting real-world processes into data
c) Encrypting files
d) Data backup process

(vi) Cite a primary skill for a data scientist.
a) Public speaking
b) SQL and Python
c) Bricklaying
d) HTML design

(vii) Select the first step in data munging.
a) Visualizing data
b) Collecting raw data
c) Presenting findings
d) Publishing data

(viii) Identify a purpose of data cleaning.
a) Adding errors
b) Formatting data for presentations
c) Removing inconsistencies and errors
d) Encrypting documents

(ix) Cite a programming language widely used in data science.
a) JavaScript
b) C#

c) Python d) CSS

(x) Select the property of high "velocity" in big data.

a) The data changes slowly b) Data is static
c) Data is generated in real time d) Data is archived

(xi) Select the correct interpretation of a residual plot with a pattern.

a) Good fit b) Heteroscedasticity or model misfit
c) Random variation d) High R^2

(xii) Identify the appropriate plot to evaluate prediction errors in regression.

a) Pie chart b) Residual plot
c) Heat map d) Line graph

(xiii) Select the regression type that includes interactions between variables.

a) Simple regression b) Linear regression
c) Interaction regression d) Multiple regression

(xiv) Identify which value should be close to zero in a good residual plot.

a) R^2 b) Mean of residuals
c) Slope d) Standard error

(xv) Choose the programming language most associated with data science.

a) Java b) Python
c) HTML d) CSS

Group-B

(Short Answer Type Questions) 3 x 5=15

2. Examine the concept of Data Science and its significance in modern decision-making. (3)
3. Enumerate the key characteristics of Big Data and their implications. (3)
4. Evaluate the concept of Datafication and its role in digital transformation. (3)
5. Evaluate the benefits of visualizing regression results in model interpretation. (3)
6. Evaluate the importance of the coefficient of determination (R^2) in regression analysis. (3)

OR

Evaluate the implications of multicollinearity in multiple regression models. (3)

Group-C

(Long Answer Type Questions) 5 x 6=30

7. Define the concept of data science and big data, and explain their interrelationship in modern business ecosystems. (5)
8. Examine the role of a data scientist in the context of datafication and digital transformation. (5)
9. Explain the importance of data structures in data science and EDA. (5)
10. Establish the role of scatterplots in evaluating variable relationships during EDA. (5)
11. Explain how data cleaning aligns with the objectives of EDA. (5)
12. Examine the role of summary statistics in supporting EDA. (5)

OR

Examine the strengths and weaknesses of SPSS for EDA. (5)

****