

INDIAN STATISTICAL INSTITUTE

MS in QMBA

TEST ON BUSINESS ANALYTICS I

Date: 11 September 2026 Time: 2 hours Maximum Marks: 50

Answer as many questions as you can. The maximum you can score is 50

1. Answer whether the following statements are true or false.
 - a. Identifying and suggesting additional products a customer may be willing to buy, based on their current and past browsing (surfing) data, is an example of a supervised learning problem.
 - b. Hypothesis testing can be applied to evaluate whether a particular web page design results in a higher conversion rate compared to an alternative design.
 - c. The `describe()` function is a part of the `dplyr` package in *R*.
 - d. The `quantile()` function in Python can be used to compute the median of a dataset.
 - e. The expression `(heat_treatment_data == 0).sum()` in Python will return the count of zeros in each column of the `heat_treatment_data` DataFrame.

[5]

2.
 - a. Explain business analytics and its role in modern organizations? Discuss the key reasons behind the growing popularity of business analytics? Briefly describe the major types of analytical solutions?
 - b. Differentiate between supervised learning and unsupervised learning with suitable examples?
 - c. Explain the difference between a regression problem and a classification problem in supervised learning. Illustrate with one business-related example for each?

[10]

3.
 - a. Describe the syntax and structure of a `for` loop in *R*, highlighting its key components. Write an *R* program using a `for` loop to display all numbers between 200 and 500 that are divisible by 21.
 - b. Describe the syntax and structure of functions in Python, including the use of parameters and return statements. Write a Python function that prints all numbers between 500 and 1000 which are divisible by 19.

[10]

- 4.
- Explain the data visualization techniques commonly used to study the relationship between a response variable (y) and an explanatory variable (x)
 - Describe the scenarios in which each visualization tool is most appropriate
 - Discuss how to interpret these visualizations to assess the strength, direction, and nature of the relationship between x and y

[10]

5. Explain the different methods available in Python for handling missing values in datasets. For each method, describe how it addresses missing data.

[10]

- 6.
- State and prove the Weak Law of Large Numbers ($WLLN$)? Provide one application of the $WLLN$ in the context of data preprocessing?
 - Define the Min–Max transformation technique and present its general formula. Show mathematically that Min–Max transformation is an affine (linear) transformation, i.e., it can be expressed in the form $y = \alpha x + \beta$

[10]