

Faculty of Engineering

End Semester Examination May 2025

RA3EL38 Industrial Data Analysis & Decision Making

Programme	:	B.Tech.	Branch/Specialisation	:	RA
Duration	:	3 hours	Maximum Marks	:	60

Note: All questions are compulsory. Internal choices, if any, are indicated. Assume suitable data if necessary.
Notations and symbols have their usual meaning.

Section 1 (Answer all question(s))

Marks CO BL

- Q1.** Which of the following statements correctly explains the difference between a population and a sample in statistics? 1 1 1
- A sample includes all data points, while a population consists of only a few.
 - A population refers to a group of individuals or data points, while a sample is a smaller group selected from it.
 - A sample is always larger than a population in size.
 - A population is just a random selection from a sample.
- Q2.** Which of the following most accurately defines "data" in the context of industrial data analytics? 1 1 1
- Information is processed and analyzed using machine learning techniques.
 - Unprocessed numbers were collected for the purpose of analysis.
 - A sample of information obtained from population-based research.
- Q3.** What is involved in the process of "web scraping"? 1 2 1
- Collecting data manually from websites.
 - Extracting data from government-provided databases.
 - Distributing surveys through email to collect data.
 - Using automated scripts to gather data from websites.
- Q4.** Which of the following data collection methods is most effective for obtaining quick responses from a large audience? 1 2 1
- Observational study
 - Focus group discussions
 - Online survey
 - In-person interviews
- Q5.** A bar chart is most effective for: 1 3 2
- Comparing the distribution of continuous data.
 - Showing the relationship between two variables.
 - Displaying the frequency of categories or groups.
 - Representing the proportion of a whole in different categories.
- Q6.** Which measure of central tendency is the middle value when the data is ordered from lowest to highest? 1 3 1
- Mode
 - Mean
 - Range
 - Median
- Q7.** When should a t-test be used instead of a z-test? 1 4 2
- When the data is categorical in nature.
 - When the sample size is large ($n > 30$).
 - When the population variance is known.
 - When the sample size is small ($n < 30$) and the population variance is unknown.

- Q8.** Which of the following is important when performing linear regression? 1 4 2
- The independent variables should be unrelated.
 - The data should have no outliers.
 - The residuals should be perfectly correlated.
 - The dependent variable should be categorical.
- Q9.** Which of the following is the main purpose of a decision tree? 1 5 1
- To predict future values based on historical data.
 - To display the relationship between two variables.
 - To group data points into categories.
 - To make decisions based on multiple conditions.
- Q10.** ISM helps in: 1 5 1
- Predicting future market trends.
 - Identifying and analyzing the relationships between different factors.
 - Collecting data from surveys.
 - Classifying data into categories.

Section 2 (Answer all question(s))

Marks CO BL

- Q11.** Explain the difference between "qualitative" and "quantitative" data in the context of industrial data analytics. 2 1 2

Rubric	Marks
Explain the "qualitative" data in the context of industrial data analytics.	1
Explain the "quantitative" data in the context of industrial data analytics.	1

- Q12.** What are the three main measures of central tendency in descriptive statistics? Explain how each one is calculated. 3 1 2

Rubric	Marks
What are the three main measures of central tendency in descriptive statistics?	1
Explain how each one is calculated.	2

- Q13. (a)** How can industrial data analytics improve decision-making in manufacturing? Explain with a few examples. 5 1 3

Rubric	Marks
How can industrial data analytics improve decision-making in manufacturing?	3.5
Explain with a few examples	1.5

(OR)

- (b)** A factory produces 500 items, and the weights of items (in kg) are measured as a sample. The following data is recorded for 8 items:
5, 7, 9, 11, 13, 15, 17, 19.
Calculate the range and standard deviation for this sample.

Rubric	Marks
Calculate the range for this sample.	1
Calculate the standard deviation for this sample.	4

Section 3 (Answer all question(s))

Marks CO BL

Q14. What is the distinction between structured and unstructured data?

2 2 2

Rubric	Marks
structured data	1
unstructured data	1

Q15. (a) Briefly explain data collection methods in industrial data analytics, such as observation, surveys, IoT sensors, and web scraping. Also, highlighting their pros and cons. 8 2 2

Rubric	Marks
explain observation	2
explain surveys	2
explain IoT sensors	2
explain web scraping	2

(OR)

(b) Briefly explain the process of creating a questionnaire for industrial data collection, focusing on key factors for effective surveys in manufacturing.

Rubric	Marks
Briefly explain the process of creating a questionnaire for industrial data collection.	4
focusing on key factors for effective surveys in manufacturing.	4

Section 4 (Answer all question(s))

Q16. What is Kurtosis? How does it influence the interpretation of a dataset?

Marks CO BL
3 3 2

Rubric	Marks
What is kurtosis?	1.5
How does it influence the interpretation of a dataset?	1.5

Q17. (a) The following data represents the number of vehicles passing through a toll booth each hour for 12 hours: 7 3 3

12,15,10,17,14,18,19,21,20,17,14,16
12, 15, 10, 17, 14, 18, 19, 21, 20, 17, 14, 16

- (i) Construct a frequency distribution with class intervals of width 5.
- (ii) Draw the cumulative frequency curve (Ogive) for this data.

Rubric	Marks
(a) Construct a frequency distribution with class intervals of width 5.	4
(b) Draw the cumulative frequency curve (Ogive) for this data.	3

(OR)

(b) A factory produces 400 units of product B each day. The following data shows the number of defective units produced on 8 randomly selected days: 3, 5, 7, 4, 6, 8, 3, 5.
Calculate the mean, variance, and coefficient of variation for the number of defective units.

Rubric	Marks
Calculate the mean	2
Calculate variance	2
Calculate coefficient of variation	3

Section 5 (Answer all question(s))

Q18. What is the purpose of a one-way ANOVA? How do you interpret its results?

Marks CO BL
4 4 2

Rubric	Marks
What is the purpose of a one-way ANOVA?	2
How do you interpret its results?	2

Q19. (a) Compare linear regression and decision trees as predictive modeling techniques, emphasizing their strengths, limitations, and appropriate use cases. 6 4 4

Rubric	Marks
linear regression as predictive modeling techniques	3
decision trees as predictive modeling techniques	3

(OR)

(b) Describe the concept of Confirmatory Factor Analysis (CFA). How does it differ from Exploratory Factor Analysis (EFA)? What are its key advantages in model validation?

Rubric	Marks
Describe the concept of Confirmatory Factor Analysis (CFA)	2
How does it differ from Exploratory Factor Analysis (EFA)	2
What are its key advantages in model validation?	2

Section 6 (Answer any 2 question(s))

Q20. How does the Analytic Hierarchy Process (AHP) help in prioritizing decisions with multiple conflicting objectives? 5 5 3

Marks CO BL
5 5 3

Rubric	Marks
Explanation	5

Q21. Explain the role of Multiple Criteria Decision Making (MCDM) in statistical analysis. How it differs from MODM? **5 5 3**

Rubric	Marks
Explain the role of Multiple Criteria Decision Making (MCDM) in statistical analysis	3
how it differs from MODM.	2

Q22. Describe the methodology of Interpretive Structural Modeling (ISM) and its role in structuring complex decision problems. **5 5 3**

Rubric	Marks
Description	3
role	2
