

TILAK MAHARASHTRA VIDYAPEETH, PUNE
MASTER OF SCIENCE (DATA SCIENCE) (768)
EXAMINATION : JANUARY - 2026
SEMESTER - I
Sub : Research Methodology (76825 -107)

Date : 10/01/2026**Total Marks : 60****Time: 10.00 am To 12.30 pm**

Instruction:

1. All questions are compulsory unless and otherwise stated.
2. Bold figures to the right of every question are the maximum marks for that question.
3. Candidates are advised to attempt questions in order.
4. Answers written illegibly are likely to be marked zero.
5. Use of scientific calculators, Log tables, Mollier Charts is allowed.
6. Draw neat and labelled diagram wherever necessary.

Q. 1. Answer the following in 2-3 lines. (Any 5) (10)

1. Define data preprocessing.
2. What is mean imputation?
3. What are outliers?
4. Define qualitative research methods.
5. What is the objective of content analysis?
6. Define thematic analysis.
7. Why is dataset preparation important?

Q. 2. Answer the following in short. (Any 4) (20)

1. Explain data cleaning and preparation for analysis.
2. Describe thematic analysis with steps.
3. Explain machine learning algorithms used in research.
4. Explain model evaluation and validation techniques.
5. Explain GDPR and HIPAA in data science research.
6. Explain formulation of research questions and hypotheses.

Q. 3. Answer the following in detail. (Any 3) (30)

1. One-Way ANOVA (Numerical)

Technique A	Technique B	Technique C
30	25	28
32	27	29
31	26	30

Test whether the mean processing times differ.

2. Two-Way ANOVA (Numerical)

Dataset Size	SVM	Decision Tree
Small	82	78
Large	88	85

Apply Two-Way ANOVA and interpret results.

3. Principal Component Analysis (Numerical)

Observation	X ₁	X ₂
1	1	2
2	2	4
3	3	6

Apply PCA and find the first principal component.

4. Factor Analysis (Numerical)

	X ₁	X ₂
X ₁	1.0	0.9
X ₂	0.9	1.0

Perform Factor Analysis assuming one factor.

5. Descriptive Statistics

Given the data:

5, 7, 9, 7, 10, 12

Calculate mean, median, mode, variance and standard deviation.
