Roll No. ..... Total Pages: 03

8253-A

M.Sc. (IT) III[rd] SEMESTER EXAMINATION, 2019

INTRODUCTION TO DATA SCIENCE

Paper - III

Time: Three Hours

Maximum Marks: 80

PART – A (खण्ड – अ) *[Marks: 20]*

Answer all questions (50 words each).

All questions carry equal marks.

सभी प्रश्न अनिवार्य हैं। प्रत्येक प्रश्न का उत्तर 50 शब्दों से अधिक न हो।

सभी प्रश्नों के अंक समान हैं।

PART – B (खण्ड – ब) *[Marks: 40]*

Answer five questions (250 words each).

Selecting one from each unit. All questions carry equal marks.

*प्रत्येक इकाई से **एक–एक** प्रश्न चुनते हुए, कुल **पाँच** प्रश्न कीजिए।*

प्रत्येक प्रश्न का उत्तर 250 शब्दों से अधिक न हो।

सभी प्रश्नों के अंक समान हैं।

PART – C (खण्ड – स) *[Marks: 20]*

Answer any two questions (300 words each).

All questions carry equal marks.

*कोई **दो प्रश्न** कीजिए। प्रत्येक प्रश्न का उत्तर 300 शब्दों से अधिक न हो।*

सभी प्रश्नों के अंक समान हैं।

PART – A

Q.1 (i) What is relational algebra?

(ii) What are parallel databases?

(iii) Define Analytics.

(iv) What is Machine Learning?

(v) What is visualization?

(vi) What is the importance of communicating results in data science?

(vii) Define Page Rank.

(viii) What is semantic web?

(ix) What is R?

(x) What is Dirty Data?

PART – B

UNIT –I

Q.2 What is In-database analytics? What is it used for?

Q.3 What is NoSQL? How is it different from SQL?

UNIT –II

Q.4 List and explain various techniques used in statistical modelling giving suitable examples.

Q.5 What is supervised learning? Explain the terminology commonly used in supervised learning.

UNIT –III

Q.6 What is visual data analytics? Explain its significance and use.

Q.7 What is the importance of provenance during data science project lifecycle? Explain.

UNIT –IV

Q.8 What is Graph Analytics? Explain its various components along with their use.

Q.9 What is community detection in graph analytics? Explain its methods and significance.

UNIT –V

Q.10 What is exploratory data analysis? Explain how is it implemented in R?

Q.11 What is a hypothesis? List and explain various methods of hypothesis testing.

PART – C

Q.12 Explain the following giving suitable examples:

(i) MapReduce

(ii) Key-Value Store

Q.13 List and explain various Optimization Methods for Supervised Machine Learning giving suitable examples.

Q.14 Write short notes on:

(a) Ethics

(b) Governance

Q.15 What are recursive queries? Explain giving suitable examples and use.

Q.16 Write short notes on:

(a) ANOVA

(b) Type I and Type II Errors
