

**MASTER OF COMPUTER APPLICATIONS DEGREE EXAMINATION,  
SEPTEMBER/OCTOBER - 2024**

**FOURTH SEMESTER**

**MCA 402A - ESSENTIALS OF DATA SCIENCE**

*(Under C.B.C.S New Regulations w.e.f. 2020-2021)*

*(Common paper to University and All Affiliated Colleges)*

**Time : 3 Hours**

**Max. Marks : 70**

**PART - A**

**(Compulsory)**

**Answer any Five of the following questions. Each question carries 4 marks.**

**(5×4=20)**

1. a) What are components of data science?
- b) Explain about the role of Domain Expertise.
- c) Mention different types of EDA.
- d) What are main features of R Language?
- e) Write about Challenges of Data Science Technology.
- f) Explain about Filters in Feature Selection.
- g) Define Curse of Dimensionality.
- h) What are the principles of Data Visualization?
- i) Define the security accepts of Data Science.
- j) Write about graphs in Data Analysis.

## PART - B

Answer Five questions. Choosing One Full question from each unit. Each question carries 10 marks.  
(5×10=50)

### UNIT - I

2. Explain about Statistical Hypothesis testing in detail.

(OR)

3. What is sampling? Explain about different types of sampling methods in statistics?

### UNIT - II

4. Define EDA. Explain about steps involved in Exploratory Data Analysis.

(OR)

5. Write about various Summary Statistics of EDA.

### UNIT - III

6. Describe about methods of Feature Extraction in detail.

(OR)

7. Explain Decision Tree Algorithm with example.

### UNIT - IV

8. Write about 3 different ways of Plotting available in Data Visualization Projects.

(OR)

9. Create your own visualization of the complex data set including all visualization methods.

### UNIT - V

10. What are the Ethical issues in data science explain in detail?

(OR)

11. Explain in detail about Next Generation Data Scientists.
-

12-00-2-06

MASTER OF COMPUTER APPLICATIONS DEGREE EXAMINATION — APRIL/MAY 2018

SECOND SEMESTER

MCA 203 — DATA SCIENCE ESSENTIALS

(Under C.B.C.S. Revised New Regulations w.e.f. 2016-2017)

(Common paper to university and all affiliated colleges)

Time : 2 hours

Max. Marks : 50

**PART – A (Compulsory)**

Answer any FIVE of the following questions. Each question carries 4 marks.

(Marks :  $5 \times 4 = 20$ )

1. (a) What is data science? Explain.
- (b) What is Statistical modeling? Explain.
- (c) What is the role of Data Scientist's data analysis Process.
- (d) What is data analysis?
- (e) Feature Selection algorithms? Explain.
- (f) Write a short note on Random forest.
- (g) What are the basic principles of Data Visualization?
- (h) What are the Examples of inspiring projects in data visualization?
- (i) Discuss the ethical issues on data.
- (j) What are the different policies on privacy issues on data.

**PART – B**

Answer ONE full question from each Unit. Each question carries 6 marks.

(Marks :  $5 \times 6 = 30$ )

**UNIT – I**

2. Explain the different Statistical Inference - Populations and samples.

Or

3. Give the introduction to the R language.

[P.T.O.]

## UNIT - II

4. With neat diagram explain Data Science Process.

Or

5. What is EDA? Explain Philosophy of Exploratory Data Analysis.

## UNIT - III

6. What is Feature Generation? Explain the different methods of feature generation.

Or

7. What is Decision Trees? Explain the general model of the decision tree.

## UNIT - IV

8. What is Data Visualization? What are the tools for data visualization?

Or

9. Create your own visualization of a complex dataset. Explain the process of visualization.

## UNIT - V

10. Discussions on privacy, security and ethics of data science.

Or

11. What are the principles of the Next-generation data scientists?
-

MASTER OF COMPUTER APPLICATIONS DEGREE EXAMINATION —  
NOVEMBER 2020

SECOND SEMESTER

**MCA 205 — DATA SCIENCE ESSENTIALS**

*(Under C. B. C. S. Revised New Regulations w.e.f. 2016-2017)*

*(Common paper to University and all Affiliated Colleges)*

*(Regular/Supplementary)*

Time : 3 hours

Max. Marks : 50

**PART - A**

**(Compulsory)**

Answer any FIVE of the following questions. Each question carries 4 marks.

(Marks :  $5 \times 4 = 20$ )

1. (a) Write down the relationship between BigData and Data Science.
- (b) Give the process of Statistical Inference.
- (c) What is EDA? Explain.
- (d) How patterns are identified from the given data?
- (e) What is meant by feature generation?
- (f) Briefly write about filters.
- (g) Define Data Visualization 3 with example.
- (h) List down the tools of Data Visualization 3.
- (i) Specify the measure related to privacy.
- (j) Identify the role of data scientists.

**PART - B**

Answer FIVE questions, choosing ONE question from each Unit.

Each question carries 6 marks.

(Marks :  $5 \times 6 = 30$ )

**UNIT - I**

2. Explain how sampling can be made from population.

Or

3. Specify the algorithm that is used for fitting a model.

**UNIT - II**

4. Describe the Data Science process in detail.

Or

5. Explain the role of Data Scientist in Data Science process.

**UNIT - III**

6. Explain step wise regression algorithm used for wrapper.

Or

7. Briefly explain about random forests

**UNIT - IV**

8. Explain the basic principles of Data Visualization 3.

Or

9. Give an example for Data Visualization. Explain in detail.

**UNIT - V**

10. Elucidate the methods to handle ethical issues.

Or

11. How do next generation data scientists act as problem solvers?
-