

CBSE | DEPARTMENT OF SKILL EDUCATION

DATA SCIENCE (SUBJECT CODE - 844)

Blue-print for Sample Question Paper for Class XII (Session 2022-2023)

Max. Time: 2 Hours

Max. Marks: 50

PART A - EMPLOYABILITY SKILLS (10 MARKS):

UNIT NO.	NAME OF THE UNIT	OBJECTIVE TYPE QUESTIONS	SHORT ANSWER TYPE QUESTIONS	TOTAL QUESTIONS
		1 MARK EACH	2 MARKS EACH	
1	Communication Skills - IV	-	-	-
2	Self-Management Skills - IV	2	2	4
3	ICT Skills - IV	2	1	3
4	Entrepreneurial Skills - IV	2	2	4
5	Green Skills - IV	-	-	-
TOTAL QUESTIONS		6	5	11
NO. OF QUESTIONS TO BE ANSWERED		Any 4	Any 3	07
TOTAL MARKS		1 x 4 = 4	2 x 3 = 6	10 MARKS

PART B - SUBJECT SPECIFIC SKILLS (40 MARKS):

UNIT NO.	NAME OF THE UNIT	OBJECTIVE TYPE QUESTIONS	SHORT ANSWER TYPE QUESTIONS	DESCRIPTIVE/ LONG ANS. TYPE QUESTIONS	TOTAL QUESTIONS
		1 MARK EACH	2 MARKS EACH	4 MARKS EACH	
1	Data Governance	4	1	1	6
2	Exploratory Data Analysis	4	1	1	6
3	Classification Algorithms I	4	1	1	6
4	Classification Algorithms II	4	1	1	6
5	Regression Algorithms I	4	1	-	5
6	Regression Algorithms II	3	1	-	4
7	Unsupervised Learning	1	-	1	2
TOTAL QUESTIONS		24	6	5	35
NO. OF QUESTIONS TO BE ANSWERED		20	Any 4	Any 3	27
TOTAL MARKS		1 x 20 = 20	2 x 4 = 8	4 x 3 = 12	40 MARKS

CBSE | DEPARTMENT OF SKILL EDUCATION

DATA SCIENCE (SUBJECT CODE -844)

Sample Question Paper for Class XII (Session 2022-2023)

Max. Time: 2 Hours

Max. Marks: 50

General Instructions:

1. Please read the instructions carefully.
2. This Question Paper consists of **21 questions** in two sections: Section A & Section B.
3. Section A has Objective type questions whereas Section B contains Subjective type questions.
4. **Out of the given (5 + 16 =) 21 questions, a candidate has to answer (5 + 10 =) 15 questions in the allotted (maximum) time of 2 hours.**
5. All questions of a particular section must be attempted in the correct order.
6. **SECTION A - OBJECTIVE TYPE QUESTIONS (24 MARKS):**
 - i. This section has 05 questions.
 - ii. Marks allotted are mentioned against each question/part.
 - iii. There is no negative marking.
 - iv. Do as per the instructions given.
7. **SECTION B – SUBJECTIVE TYPE QUESTIONS (26 MARKS):**
 - i. This section has 16 questions.
 - ii. A candidate has to do 10 questions.
 - iii. Do as per the instructions given.
 - iv. Marks allotted are mentioned against each question/part.

iii.	Which of the following is NOT a part of a Decision Tree? i. Decision Node ii. Leaf Node iii. Branch iv. Result fruit	1
iv.	Cross-validation helps us do the following. i. Removes bias ii. Helps to test tune parameters iii. Both of the mentioned	1
v.	Regression can be performed for which kind of variables? i. Continuous ii. Discrete iii. Categorical iv. All of the above	1
vi.	Which of the following statement is false? i. Unsupervised learning is done on unlabeled data sets. ii. Unsupervised learning needs human intervention. iii. Unsupervised learning is a continuous process. iv. None of the mentioned	1

Q. 3	Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)	
i.	What are some important ethical guidelines for data? i. Keeping data secure. ii. Making models impartial. iii. Being open and accountable. iv. All of the above.	1
ii.	You need to perform a univariate analysis. Which graph can you use? i. Scatter plot ii. Histogram iii. Contour plot iv. Both scatter plot and histogram can be used	1
iii.	A decision tree can be further divided into further sub-trees. i. True ii. False	1
iv.	K-NN works well with imbalanced data sets. i. True ii. False	1
v.	Which of the following statement is false? i. Regression can be done for only two variables. ii. Regression is a supervised learning technique. iii. Regression can be done for categorical variables. iv. Regression helps us predict the value of the dependent variable.	1
vi.	Non-linear regression is _____ than linear regression. i. more flexible ii. less flexible iii. equally flexible	1
Q. 4	Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)	
i.	Some organizations store more personal information than required. This is in accordance with Data Ethics. i. True ii. False	1

ii.	You need to check the relationship between the two variables. Which graph would you use? i. Histogram ii. Pair plot iii. Box plot iv. Any of the mentioned graphs can be used	1
iii.	Which of the following statement is false? i. Decision Trees are intuitive to understand. ii. Decision Trees can be used for classification and regression. iii. Random Forests algorithm uses Decision Trees. iv. Decision Trees can only handle numerical variables.	1
iv.	Which of the following statement is false? i. K-Nearest Neighbors uses proximity for prediction. ii. K-Nearest Neighbors takes a majority vote on K data points. iii. K-Nearest Neighbors can be used for clustering. iv. K-Nearest Neighbors is especially used in data mining and pattern recognition	1
v.	The value of RSME is usually greater than the value of MAE. i. True ii. False	1
vi.	The equation for Non-Linear Regression follows the equation of a curve. i. True ii. False	1

Q. 5	Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)	
i.	Which of the following is not a focus area under data governance? i. Data integrity ii. Data availability iii. Data traceability iv. Data consistency	1
ii.	Which of the following is a data cleaning step? i. Adding duplicates ii. Ignoring outliers iii. Removing duplicates iv. Adding features	1
iii.	Choose the correct statement i. Regression trees are used when the dependent variable is continuous. Classification trees are used when the dependent variable is categorical ii. Classification trees are used when the dependent variable is continuous. Regression trees are used when the dependent variable is categorical iii. Both classification and regression trees are used when the dependent variable is continuous iv. Both classification and regression trees are used when the dependent variable is categorical	1
iv.	What is the K-Nearest Neighbors algorithm used for? i. Regression ii. Classification iii. Both Regression and Classification iv. Neither Regression nor Classification	1
v.	MEA considers the direction of the residuals i. True ii. False	1

vi.	Non-Linear Regression can be performed for which kind of variables i. Continuous ii. Discrete iii. Categorical iv. All of the above	1
------------	---	----------

SECTION B: SUBJECTIVE TYPE QUESTIONS

Answer any 3 out of the given 5 questions on Employability Skills (2 x 3 = 6 marks)

Answer each question in 20 – 30 words.

Q. 6	Explain the concept of Interpersonal skills. Give an example displaying these skills.	2																																																
Q. 7	Write the steps for inserting different shapes in your presentation.	2																																																
Q. 8	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 10%;">S.No.</th> <th style="width: 15%;">A</th> <th style="width: 15%;">B</th> <th style="width: 15%;">C</th> <th style="width: 15%;">D</th> <th style="width: 15%;">E</th> <th style="width: 15%;">F</th> </tr> <tr> <th>Name</th> <th>Emp. ID</th> <th>Grade Pay</th> <th>Basic Pay</th> <th>TA/DA</th> <th>Total Salary</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">1</td> <td>Ravi</td> <td>ab1</td> <td style="text-align: center;">3000</td> <td style="text-align: center;">2000</td> <td style="text-align: center;">1000</td> <td></td> </tr> <tr> <td style="text-align: center;">2</td> <td>Reena</td> <td>ab2</td> <td style="text-align: center;">3500</td> <td style="text-align: center;">2200</td> <td style="text-align: center;">800</td> <td></td> </tr> <tr> <td style="text-align: center;">3</td> <td>Vikas</td> <td>ab3</td> <td style="text-align: center;">4000</td> <td style="text-align: center;">3000</td> <td style="text-align: center;">1200</td> <td></td> </tr> <tr> <td style="text-align: center;">4</td> <td>Harsh</td> <td>ab5</td> <td style="text-align: center;">3000</td> <td style="text-align: center;">3300</td> <td style="text-align: center;">800</td> <td></td> </tr> <tr> <td style="text-align: center;">5</td> <td>Geeta</td> <td>ab5</td> <td style="text-align: center;">2800</td> <td style="text-align: center;">5000</td> <td style="text-align: center;">1500</td> <td></td> </tr> </tbody> </table>	S.No.	A	B	C	D	E	F	Name	Emp. ID	Grade Pay	Basic Pay	TA/DA	Total Salary	1	Ravi	ab1	3000	2000	1000		2	Reena	ab2	3500	2200	800		3	Vikas	ab3	4000	3000	1200		4	Harsh	ab5	3000	3300	800		5	Geeta	ab5	2800	5000	1500		2
	S.No.	A	B	C	D	E	F																																											
	Name	Emp. ID	Grade Pay	Basic Pay	TA/DA	Total Salary																																												
	1	Ravi	ab1	3000	2000	1000																																												
	2	Reena	ab2	3500	2200	800																																												
	3	Vikas	ab3	4000	3000	1200																																												
	4	Harsh	ab5	3000	3300	800																																												
5	Geeta	ab5	2800	5000	1500																																													
Answer the following questions using above table:- a) Write formula for finding the total salary of Vikas. b) Which function shall be used to arrange total salary in ascending order.																																																		
Q. 9	How do 'Openness' & Consciousness factor of 'FF Model' contribute in an individual's personality?	2																																																
Q. 10	Q5.Match the situations in the columns to the meaning of the word: (Any two)	2																																																
	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 70%;">Situations</th> <th style="width: 30%;">Word</th> </tr> </thead> <tbody> <tr> <td>a) Shobit wants to start more batches in his competitive coaching classes but is finding it difficult to hire teachers</td> <td style="text-align: center;">Venture</td> </tr> <tr> <td>b) Karan has a business of selling Momos. He wants to open a restaurant selling all types of snacks</td> <td style="text-align: center;">Barriers</td> </tr> <tr> <td>c) Aman has started an electronics store, where he sells LED,AC,Refrigerator etc</td> <td style="text-align: center;">Risk</td> </tr> </tbody> </table>	Situations	Word	a) Shobit wants to start more batches in his competitive coaching classes but is finding it difficult to hire teachers	Venture	b) Karan has a business of selling Momos. He wants to open a restaurant selling all types of snacks	Barriers	c) Aman has started an electronics store, where he sells LED,AC,Refrigerator etc	Risk																																									
	Situations	Word																																																
	a) Shobit wants to start more batches in his competitive coaching classes but is finding it difficult to hire teachers	Venture																																																
b) Karan has a business of selling Momos. He wants to open a restaurant selling all types of snacks	Barriers																																																	
c) Aman has started an electronics store, where he sells LED,AC,Refrigerator etc	Risk																																																	

Answer any 4 out of the given 6 questions in 20 – 30 words each (2 x 4 = 8 marks)

Q. 11	What are some of the aspects covered by data governance?	2
Q. 12	What are the ways to handle missing data?	2
Q. 13	What are Decision Trees?	2
Q. 14	Mention two important characteristics of K-NN.	2
Q. 15	Define Mean Absolute Error and Root Mean Square Deviation	2
Q. 16	What is Multiple Linear Regression?	2

Answer any 3 out of the given 5 questions in 50– 80 words each (4 x 3 = 12 marks)

Q. 17	Write a short note on the California Consumer Privacy Act.	4
Q. 18	What are some of the differences between univariate and multivariate analysis? Give some examples.	4
Q. 19	Write the steps to create a Decision Tree.	4
Q. 20	What is cross-validation? Mention the steps involved in cross-validation.	4
Q. 21	Mention some common applications of Unsupervised Learning.	4