CBSE | DEPARTMENT OF SKILL EDUCATION

DATA SCIENCE (SUBJECT CODE - 844)

BLUE-PRINT FOR SAMPLE QUESTION PAPER – 1 (PRACTICE PAPER-1)

CLASS XII (SESSION 2025-2026)

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	1	1	2
2	Self-Management Skills - IV	2	1	3
3	ICT Skills - IV	1	1	2
4	Entrepreneurial Skills - IV	1	1	2
5	Green Skills - IV	1	1	2
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	3	1	1	5
3	Classification Algorithms – I	3	1	1	5
4	Classification Algorithms – II	4	1	-	5
5	Regression Algorithms – I	4	1	-	5
6	Regression Algorithms – II	2	-	1	3
7	Unsupervised Learning	4	1	1	6
TOTAL QUESTIONS		**24**	**6**	**5**	**35**
NO. OF QUESTIONS TO BE ANSWERED		**Any 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 - 1 (PRACTICE PAPER-1)
CLASS XII (SESSION 2025-2026)

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.

SECTION A: OBJECTIVE TYPE QUESTIONS

<table>
<tr><td>Q. 1</td><td colspan="2">Answer any 4 out of the given 6 questions on Employability Skills (1 x 4 = 4 marks)</td></tr>
<tr><td>i.</td><td>An indirect object answers the questions, such as

a. What?
b. Why?
c. To whom?
d. Which?</td><td>1</td></tr>
<tr><td>ii.</td><td>“I enjoy singing and practice it as it gives me a lot of pleasure. It also works as a stress buster.” The sentence given above is an example of

a. Intrinsic motivation
b. Extrinsic motivation
c. External motivation
d. Lack of motivation</td><td>1</td></tr>
<tr><td>iii.</td><td>Statement1: Personality traits are defined as relatively lasting patterns of thoughts, feelings and behaviours.
Statement2: They distinguish individuals from one another.

a. Both Statement1 and Statement2 are correct
b. Both Statement1 and Statement2 are incorrect
c. Statement1 is correct but Statement2 is incorrect
d. Statement2 is correct but Statement1 is incorrect</td><td>1</td></tr>
<tr><td>iv.</td><td>Match the following:
<table><tr><td>A. To underline text</td><td>1. Ctrl + B</td></tr><tr><td>B. To make text bold</td><td>2. Ctrl + I</td></tr><tr><td>C. To make text italic or slanting</td><td>3. Ctrl + U</td></tr></table>
a. A - 2; B - 3; C - 1
b. A - 3; B - 1; C - 2
c. A - 1; B - 3; C - 2
d. A - 1; B - 2; C - 3</td><td>1</td></tr>
<tr><td>v.</td><td>After 20 years of being a successful entrepreneur in Lucknow, Ravi decided to move back to his village. He opened a clothing store but people in his village said that the style did not match what they wanted. They wanted more variety and brighter colours. So, he sold off the old clothes to his friend in a city and bought good quality clothes from a local seller.
Which attitude can you witness in Ravi?
a. Interpersonal skills
b. Perseverance
c. Competitiveness
d. Barrier</td><td>1</td></tr>
</table>

vi.	Electronic waste, as known as e-waste, is generated when any electronic or electrical equipment becomes unfit for the intended use or if it has crossed its expiry date. Due to rapid technological advancements and the production of newer electronic equipment, the old ones get easily replaced with new models. It has particularly led to an exponential increase in e-waste in India. Which of the following is the correct way to handle e-waste? a. Sell the e-waste to a local scrap dealer b. Throw the e-waste in the dustbin c. Dispose-off the e-waste with the help of a certified partner d. Dump the e-waste in the local landfill	**1**

Q. 2	**Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)**	
i.	Assertion (A): Data Governance is a data management concept. Reason (R): It ensures that high data quality exists throughout the complete lifecycle of the data along with effective controls on how and with whom data is shared. a. Both A and R are correct and R is the correct explanation of A b. Both A and R are correct but R is not the correct explanation of A c. A is correct and R is incorrect d. R is correct and A is incorrect	**1**
ii.	An observation that lies in an abnormal distance from other values in a random sample and needs to be removed is an _____ a. Observation b. Missing data c. Outlier d. Duplicate data	**1**
iii.	Identify the cardiologist who developed a decision tree to determine if a person was likely to have a heart attack in 1970. a. Lee Goldman b. John Mc Carthy c. Claude Shannon d. Marvin Minsky	**1**
iv.	Name the technique in which we reserve a particular portion of a dataset on which we do not train the model.	**1**
v.	A large RSME value shows that a. the model is a better bit to the data b. the model is accurate c. the model is well trained d. the model is not a good fit and might need to be retrained.	**1**
vi.	Apps like amazon, flipkart predict what products a customer is likely to purchase. This is an example of a. Targeted advertisements b. Voice assistance c. Recommendation systems d. Anomaly detection	**1**

Q. 3	**Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)**	
i.	The moral principles that govern the behavior or actions of an individual or a group is known as _____ a. Access b. Privacy c. Ethics d. Bias	**1**
ii.	Statement1: Bivariate analysis is a method to perform exploratory data analysis. Statement2: This is performed on one feature variable in the raw dataset by preparing visualizations along with summary statistics. a. Both Statement1 and Statement2 are correct b. Both Statement1 and Statement2 are incorrect c. Statement1 is correct but Statement2 is incorrect d. Statement2 is correct but Statement1 is incorrect	**1**
iii.	Decision Trees and other tree-based learning algorithms are considered to be one of the best and most used _____ a. Reinforcement learning model b. Unsupervised learning model c. Supervised learning model d. Dynamic learning model	**1**
iv.	Which of the following is correct about 'K' in K-Nearest Neighbour algorithm? a. The value of 'K' is fixed b. Value of K is set for optimum accuracy c. K is always an even number d. The algorithm takes a minority vote	**1**
v.	How are RSME and MAE values related? a. The value of RSME is usually greater than the value of MAE b. The value of RSME is usually lower than the value of MAE c. The value of RSME is usually equal to the value of MAE d. The value of RSME is not related to the value of MAE	**1**
vi.	_____refers to the process of dividing entire raw data into several groups such that the data points in one group are similar to other data points in the same group but different from those in other groups. a. Clustering b. Classification c. Regression d. Dimensionality reduction	**1**

Q. 4	**Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)**	
i.	Which of the following statement is incorrect? a. GDPR was passed in the year 2018 to protect European Union consumer data. b. PDP bill was passed in India the year 2019 c. CCPA was passed in California in the year 2018 d. HIPAA was passed in India to protect healthcare information from fraud and theft.	**1**
ii.	_____is done with the help of summary statistics and graphical representations.	**1**
iii.	K-NN algorithm is used to solve the problems related to ____ i) Clustering ii) Classification iii) Regression a. i and ii b. i and iii c. ii and iii d. ii only	**1**
iv.	The equation for a simple linear regression is of the form **Y = mX+ C** Here X is the a. Slope b. Intercept c. Dependent variable d. Independent variable	**1**
v.	Statement1: The graph of non-linear regression follows the equation of line. Statement2: The graph of linear regression follows the equation of curve. a. Both Statement1 and Statement2 are correct b. Both Statement1 and Statement2 are incorrect c. Statement1 is correct but Statement2 is incorrect d. Statement2 is correct but Statement1 is incorrect	**1**
vi.	Identify the term that is not related to clustering: a. Partitioning b. Density-Based c. Multi-class d. Distribution Model-Based	**1**

<table>
<tr><td>Q. 5</td><td>Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)</td><td></td></tr>
<tr><td>i.</td><td>Statement1: Ethical guidelines must be well-defined in any organization dealing with lots of data.
Statement2: Transparency is an important ethical guideline.
a. Both Statement1 and Statement2 are correct
b. Both Statement1 and Statement2 are incorrect
c. Statement1 is correct but Statement2 is incorrect
d. Statement2 is correct but Statement1 is incorrect</td><td>1</td></tr>
<tr><td>ii.</td><td>Identify the correct match
<table><tr><td>1. Decision tree</td><td>a) Holds a class label</td></tr><tr><td>2. Branch node</td><td>b) Tree like structure</td></tr><tr><td>3. Leaf node</td><td>c) Outcome of the test</td></tr></table>a. 1 – c; 2 – a; 3 – b
b. 1 – c; 2 – b; 3 – a
c. 1 – b; 2 – a; 3 – c
d. 1 – b; 2 – c ; 3 – a</td><td>1</td></tr>
<tr><td>iii.</td><td>Name any one distance technique used by k-NN algorithm.</td><td>1</td></tr>
<tr><td>iv.</td><td>Renu has enrolled for a weight loss program. She keeps a track of her weight on her mobile app. The app predicts her expected weight after a period of three months using linear regression. Which type of data is Renu tracking?
a. Discrete
b. Continuous
c. Qualitative
d. Categorical</td><td>1</td></tr>
<tr><td>v.</td><td>How many independent variables and dependent variables are used in multiple linear regression?
a. Zero, zero
b. One, many
c. Many, many
d. Many, one</td><td>1</td></tr>
<tr><td>vi.</td><td>Which of the following statement is false?
a. Unsupervised learning is done on unlabeled data sets.
b. Unsupervised learning needs human intervention.
c. Supervised learning is done on labeled data sets.
d. Unsupervised learning is a continuous process.</td><td>1</td></tr>
</table>

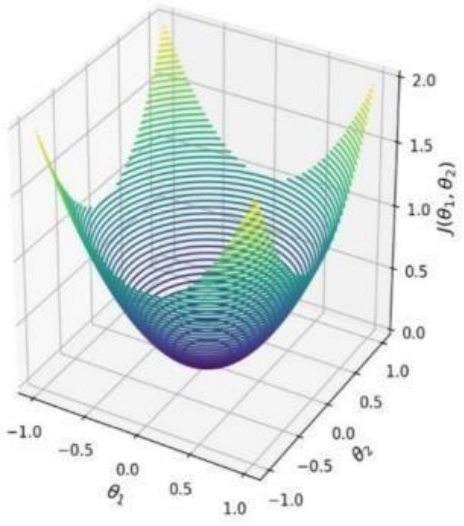
SECTION B: SUBJECTIVE TYPE QUESTIONS

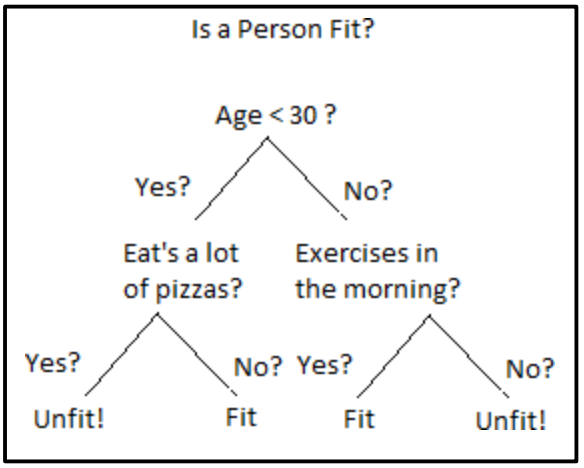
Answer any 3 out of the given 5 questions on Employability Skills (2 x 3 = 6 marks)

Answer each question in 20 – 30 words.

Q. 6	What does 'P' stand for in the acronym 'RESPECT' used to ensure active listening? Explain.	**2**
Q. 7	Shikha is an elderly woman. She stays with her family. She has a habit of washing her hands at least 20 times a day. Even after washing her hands, she feels they are not clean, and continues rubbing or washing them. She neither talks to her grandchildren, nor does she participate in any family activity. Suggest her one ways in which she can overcome her personality disorder.	**2**
Q. 8	Differentiate between the terms worksheet and workbook.	**2**
Q. 9	Name two government schemes that help small-scale entrepreneurs gain capital.	**2**
Q. 10	The greening of economy presents a major opportunity to start new businesses, develop new markets and lower energy costs. Write any two benefits of green jobs.	**2**

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

Q. 11	An online shopping portal sells only clothes. The portal tracks the browsing habits of the customers which includes the kind of clothes they buy, frequency of buying clothes, name, gender, mobile number, etc., This portal sends promotional offers to all their customers on their mobile numbers as a message. This is a concern relating to data privacy. What do you mean by the term data privacy?	**2**
Q. 12	How is the given figure related to multivariate analysis? Explain. 	**2**
Q. 13	"K-NN follows the principle of lazy learning. Explain this statement.	**2**

Q. 14	What can you relate the given picture to? Give two advantages of the concept.  	**2**
Q. 15	Give examples of any two situations in which we use linear regression.	**2**
Q. 16	How do unsupervised learning algorithms work?	**2**

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

Q. 17	Sneha was making a project related to data science. She collected data from reliable and authentic sources. But, she wanted to know about governing the data. Explain her the term data governance. Suggest her three points to keep in mind while dealing about data.	**4**
Q. 18	Data cleaning is a fundamental aspect of data science. If you have a dataset that has been cleaned well, even simple algorithms can learn and give impressive insights from the data. Explain any four steps of data cleaning.	**4**
Q. 19	A Decision tree is a flowchart like tree structure. With a suitable example, explain the steps to create and draw a decision tree.	**4**
Q. 20	Assessment of triglyceride and cholesterol in overweight people is based on effects of age, weight, and height. a. Which type of linear regression can you relate this to? b. Define the type of linear regression. c. Write the formula and explain the terms used.	**4**
Q. 21	Let's say you have multiple documents and you need to cluster similar documents together. Clustering helps us group these documents such that similar documents are in the same clusters. Explain the working of k-means clustering.	**4**