# CBSE | DEPARTMENT OF SKILL EDUCATION

## DATA SCIENCE (SUBJECT CODE - 844)

### MARKING SCHEME FOR 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.** There is no negative marking.
   - **iii.** Do as per the instructions given.
   - **iv.** Marks allotted are mentioned against each question/part.
7. **SECTION B – SUBJECTIVE TYPE QUESTIONS (26 MARKS):**
   - **i.** This section contains 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

| Q. No. | QUESTION | Source Material (NCERT/PSSCIVE/ CBSE Study Material) | Unit/ Chap. No. | Page no. of source material | Marks |
|---|---|---|---|---|---|
| **Q.1** | **Answer any 4 out of the given 6 questions on Employability Skills (1 x 4 = 4 marks)** | | | | |
| **i.** | a. Only statement 1 | NCERT | 1 | 3 | **1** |
| **ii.** | d. Result-Oriented Goal | NCERT | 2 | 30 | **1** |
| **iii.** | a. Schizoid personality | NCERT | 2 | 35 | **1** |
| **iv.** | a. A  2; B  3; C 1 | NCERT | 3 | 42, 76 | **1** |
| **v.** | c. Standard of Excellence | NCERT | 4 | 86 | **1** |
| **vi.** | c. Green Collar | NCERT | 5 | 112 | **1** |

| | | | | | |
|---|---|---|---|---|---|
| **Q.2** | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | | | | |
| **i.** | c. Statement1 is correct, but Statement2 is incorrect | NCERT | 1 | 3 | **1** |
| **ii.** | b. Exploratory Data Analysis | NCERT | 2 | 7 | **1** |
| **iii.** | c. Random Forest | NCERT | 3 | 15 | **1** |
| **iv.** | False | NCERT | 4 | 23 | **1** |
| **v.** | b. Regression | NCERT | 5 | 33 | **1** |
| **vi.** | c. Anomalies | NCERT | 7 | 47 | **1** |
| **Q.3** | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | | | | |
| **i.** | b. Create machine learning models that are biased and robust | NCERT | 1 | 2 | **1** |
| **ii.** | b. Univariate Analysis | NCERT | 2 | 8 | **1** |
| **iii.** | c. Class Label | NCERT | 3 | 15 | **1** |
| **iv.** | d. No Training Step | NCERT | 4 | 24 | **1** |
| **v.** | b. Y = mx + b | NCERT | 5 | 34 | **1** |
| **vi.** | a. Clustering | NCERT | 7 | 45 | **1** |
| **Q.4** | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | | | | |
| **i.** | a. Health Insurance Portability and Accountability Act | NCERT | 1 | 3 | **1** |
| **ii.** | c. Both graphical and statistical methods | NCERT | 2 | 8 | **1** |
| **iii.** | False | NCERT | 3 | 15 | **1** |
| **iv.** | c. Mean Absolute Error | NCERT | 5 | 34 | **1** |
| **v.** | d. Non-Linear Functions | NCERT | 6 | 41 | **1** |
| **vi.** | c. Supervised Learning | NCERT | 7 | 46 | **1** |
| **Q.5** | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | | | | |
| **i.** | a. Both Assertion (A) and Reason (R) are true, and Reason (R) is the correct explanation of Assertion (A) | NCERT | 1 | 4 | **1** |
| **ii.** | a. 4, 3 | NCERT | 3 | 15 | **1** |
| **iii.** | d. 10 | NCERT | 4 | 25 | **1** |
| **iv.** | b. Nonlinear Regression | NCERT | 6 | 41 | **1** |
| **v.** | b. small | NCERT | 5 | 35 | **1** |

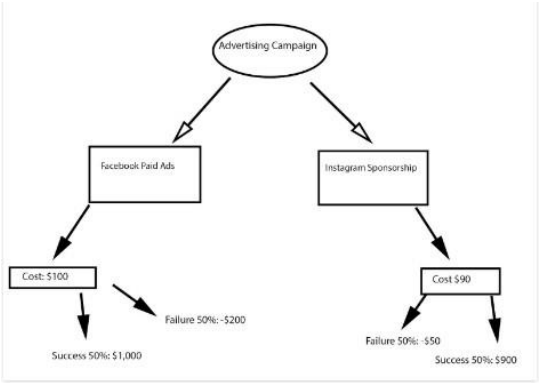
| | | | | | |
|---|---|---|---|---|---|
| **vi.** | Expectation Maximization | NCERT | 7 | 46 | **1** |

## SECTION B: SUBJECTIVE TYPE QUESTIONS

| Q. No. | QUESTION | Source Material (NCERT/PSSCIVE/ CBSE Study Material) | Unit/ Chap. No. | Page no. of source material | Marks |
|---|---|---|---|---|---|
| **Answer any 3 out of the given 5 questions on Employability Skills in 20 – 30 words each (2 x 3 = 6 marks)** | | | | | |
| **Q.6** | The five stages of active listening are as follows.<br>1. Receiving: It involves listening attentively.<br>2. Understanding: It is an informed agreement about something or someone.<br>3. Remembering: It refers to the retrieval or recall of some information from the past.<br>4. Evaluating: It is about judging the value, quantity, importance, and amount of something or someone.<br>5. Responding: It is about saying or doing something as a response to something that has been said or done<br>Receiving and Understanding - ½ mark<br>Remembering, Evaluating, and Responding - 1½ mark | NCERT | 1 | 5 | **2** |
| **Q.7** | a. Openness ½ mark<br>b. Agreeableness ½ mark<br>c. Neuroticism ½ mark<br>d. Consciousness ½ mark | NCERT | 2 | 33, 34 | **2** |
| **Q.8** | a. =SUM(B2:E2) or =B2+C2+D2+E2<br>b. Active Cell<br>c. Range of Cells<br>d. Integer and Text<br>(½ mark for each) | NCERT | 3 | 41, 45, 50 | **2** |
| **Q.9** | They rely on capital that comes via angel investors or venture capital firms, while small business operations rely on loans and grants. | NCERT | 4 | 89 | **2** |

| | | | | | |
|---|---|---|---|---|---|
| | A lot of Indian startups, too, have received funding from global investors<br>( 1 mark each) | | | | |

| | | | | | |
|---|---|---|---|---|---|
| **Q.10** | Identify the name: National Action Plan on Climate Change (NAPCC) (1 mark)<br>The NAPCC comprises eight missions in specific areas such as solar energy, enhanced energy efficiency, habitat, water, sustaining Himalayan ecosystems, forestry, agriculture, and strategic knowledge for climate change (Any 4 areas) which address issues relating to the mitigation of greenhouse gases and adaptation to the adverse impacts of climate change on the environment, forests, habitat, water resources, and agriculture.<br>( 1 mark) | NCERT | 5 | 114 | **2** |
| **Answer any 4 out of the given 6 questions in 20 – 30 words each (2 x 4 = 8 marks)** | | | | | |
| **Q.11** | The legislation implemented is the General Data Protection Regulation.<br>The European Union made the General Data Protection Regulation effective on May 25th, 2018, to protect European Union consumer data. All the reformed laws are made to help consumers gain a high level of control over their data.<br>(1 mark for identifying the legislation)<br>(1 mark for an explanation - ½ mark for each point) | NCERT | 1 | 3 | **2** |
| **Q.12** | 1. Remove duplicate observations<br>2. Remove irrelevant observations<br>3. Remove unwanted outliers<br>4. Fix data type issues<br>5. Handle missing data<br>(Any 4 points - 1/2 mark for each point) | NCERT | 2 | 10 | **2** |
| **Q.13** | Cross-validation refers to a technique in which we reserve a particular portion of a dataset on which we do not train the model. After the training is over, we test the resulting model on this portion of the data before finalizing it.<br>K-fold Cross Validation<br>(1 mark for each point) | NCERT | 4 | 24 | **2** |

| | | | | | |
|---|---|---|---|---|---|
| **Q.14** | <br>1 mark for decision nodes 1 mark for the branches | NCERT | 3 | 15 | **2** |
| **Q.15** | Linear regression is known as "linear" because the relation of the dependent to the independent variables is a linear function of some parameters.<br>The equation for a simple linear regression is of the form y = mx + b<br>(1 mark per point) | NCERT | 5 | 33,34 | **2** |
| **Q.16** | Some online news websites use unsupervised learning techniques to categorize articles from various online news outlets to put them under various sections like sports, entertainment, international news, etc.<br>(2 marks for any relevant explanation) | NCERT | 7 | 45 | **2** |
| **Answer any 3 out of the given 5 questions in 50– 80 words each (4 x 3 = 12 marks)** | | | | | |
| **Q.17** | Data privacy is the right of any individual to have control over how his or her personal information is collected and used.<br><br>The following points are to be kept in mind while dealing with data.<br>● Keep the data secure.<br>● Create machine learning models that are impartial and robust<br>● Be as open and accountable as possible<br>● Use technologies and data architecture that have the minimum intrusion necessary.<br>(1 mark for defining the term data privacy; any 3 benefits – 3 marks) | NCERT | 1 | 2 | **4** |

| | | | | | |
|---|---|---|---|---|---|
| **Q.18** | Exploratory Data Analysis is the process of carrying out an initial analysis of the available data to find out more about the data.<br>Univariate analysis can be considered the easiest form of data analysis, where we only analyze one variable from the entire dataset.<br>Univariate analysis techniques involve both statistical and graphical methods.<br><br>Multivariate analysis is Bivariate analysis refers to the analysis of two or more variables in the dataset.<br>Bivariate analysis is usually done by using graphical methods like scatter plots, line charts, pair plots, and more complex forms of statistical analysis techniques.<br><br>Definition of EDA - 2 marks<br>Differences - 2 marks (½ mark for each valid point) | NCERT | 2 | 7,8,9 | **4** |
| **Q.19** | **Advantages:**<br>i) Their graphical representation is very intuitive, and users can easily relate to their hypothesis.<br>ii) Another major advantage of the decision tree is that it can handle both numerical and categorical variables.<br>iii) They are also not influenced much by outliers and missing values to a fair degree.<br>Difference between Classification and Regression<br>i) Regression trees are used when the dependent variable is continuous. Classification trees are used when the dependent variable is categorical.<br>ii) In the case of a regression tree, the value of the terminal nodes after training is the mean of the observations. Thus, predictions on unseen data are made using the mean.<br>In the case of a classification tree, the value or class of the terminal nodes after training is the mode of the observations. | NCERT | 3 | 16 | **4** |

| | | | | | |
|---|---|---|---|---|---|
| | Thus, predictions on unseen data are made using the mode.<br><br>(Any 2 advantages stated above, 1 mark per point; Two differences: 1 mark for each difference) | | | | |
| **20.** | a: Linear Regression 1 mark b: Positive 1 mark<br>c: Y = mx + b - ½ mark<br>Y is the Dependent Variable,<br>X is an Independent Variable, b is the intercept and<br>m is slope<br>Labeling the variables - ½ mark<br>d. The basic objective of linear regression is to try to reduce the vertical distance between the line and the data points to make it a minimum - ½ mark<br>"fitting the line to the data - ½ mark | NCERT | 5 | 34 | **4** |
| **21.** | Some of the main clustering methods are - Partitioning Clustering<br>Density-Based Clustering<br>Distribution Model-Based Clustering<br>Hierarchical Clustering<br>(Any 1 method - ½ mark)<br><br>The k-means clustering algorithm works is as follows:<br>1. First, specify the number of clusters K depending on the input.<br>2. Initialize the centroids by shuffling the data points and then selecting<br>K data points for the centroids randomly.<br>3. Do iterations till the point that there is no change to the centroids such that the assignment of data points to clusters aren't changing.<br>4. Compute the sum of the squared distance between data points and all centroids.<br>5. Classify or mark each data point to the cluster it is closest to.<br>(Each point - ½ mark each - 2½ marks) | NCERT | 7 | 45, 46 | **4** |