# CBSE | DEPARTMENT OF SKILL EDUCATION
## DATA SCIENCE (SUBJECT CODE - 419)

## BLUE PRINT FOR SAMPLE QUESTION PAPER - 1 (PRACTICE PAPER-1)
### CLASS X (SESSION 2024-2025)

**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 –II | 1 | 1 | 2 |
| 2 | Self-Management Skills –II | 2 | 1 | 3 |
| 3 | ICT Skills – II | 1 | 1 | 2 |
| 4 | Entrepreneurial Skills –II | 1 | 1 | 2 |
| 5 | Green Skills -II | 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 | Use of statistics in Data Science | 4 | 1 | 1 | 6 |
| 2 | Distributions in Data Science | 4 | 1 | 1 | 6 |
| 3 | Identifying Patterns | 4 | 1 | 1 | 6 |
| 4 | Data Merging | 7 | 2 | 1 | 10 |
| 5 | Ethics in Data Science | 5 | 1 | 1 | 7 |
| | **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 - 419)

### SAMPLE QUESTION PAPER - 1 (PRACTICE PAPER-1)
### CLASS X (SESSION 2024-2025)

**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

| Q. 1 | **Answer any 4 out of the given 6 questions on Employability Skills (1 x 4 = 4 marks)** | |
|---|---|---|
| i. | How much % of the communication that we do in our day-to-day communication is non-verbal?<br>    a) 55%<br>    b) 93%<br>    c) 7%<br>    d) 38% | **1** |
| ii. | 'Prepare yourself for new changes, so that you can transition seamlessly'.<br>Which term can you relate the given sentence to?<br>    a) Self-awareness<br>    b) Responsibility<br>    c) Adaptability<br>    d) Time-management | **1** |
| iii. | _____is the ability to plan and control how you spend the hours of your day well and do all that you want to do. | **1** |
| iv. | Which key opens the start menu?<br>    a) Ctrl key<br>    b) Alt key<br>    c) Shift key<br>    d) Windows key | **1** |
| v. | Ravi's customer comes to his store and starts shouting at him. He does not get angry. He listens to what his customer is saying. He is_____.<br>    a) hardworking<br>    b) confident<br>    c) patient<br>    d) trying new ideas | **1** |
| vi. | Statement1: Economic development is using up resources of the world.<br>Statement2: This would have serious environmental problems, much worse than those that we are facing at present.<br>    a) Both Statement1 and Statement2 are correct<br>    b) Both Statement1 and Statement2 are incorrect<br>    c) Statement1 is correct but Statement2 is incorrect<br>    d) Statement2 is correct but Statement1 is incorrect | **1** |

| Q. 2 | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | |
|---|---|---|
| i. | Which of the following term means 'an average value of a data set'?<br>    a) Mean<br>    b) Median<br>    c) Mode<br>    d) Standard deviation | **1** |

| | | | |
|---|---|---|---|
| ii. | At the animal shelter, after counting the cats, they are weighed. The counts are _____ values while their weights are _____ values.<br>    a) continuous, discrete<br>    b) continuous, qualitative<br>    c) discrete, continuous<br>    d) qualitative, continuous | | 1 |
| iii. | $\mu$ and $\sigma$ denotes ___ and ___ respectively.<br>    a) Mean and standard deviation<br>    b) Mean and Sample size<br>    c) Standard deviation and mean<br>    d) Standard deviation and sample size | | 1 |
| iv. | _____ the process of combining two or more data sets into a single data frame. | | 1 |
| v. | _____ aims at creating methods, set of responsibilities and processes to standardize, integrate, protect and store data.<br>    a) National Curriculum Framework<br>    b) Data Governance Framework<br>    c) Current Curriculum Framework<br>    d) Business Process Framework | | 1 |
| vi. | Confidential data can be stored in which of the following format?<br>    a) Digital Data<br>    b) Physical Copies<br>    c) Both a) and b)<br>    d) Neither a) nor b) | | 1 |

| | | |
|---|---|---|
| Q. 3 | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | |
| i. | If we consider all the flowers in a garden as one set, then what do we call a rose in the garden?<br>    a) Set<br>    b) Subset<br>    c) Superset<br>    d) Powerset | 1 |
| ii. | Which term in data science refers to the method that shows the probable values for a variable and how often they occur? | 1 |
| iii. | Choose the correct set of matches.<br><br>| (i) Selection Bias | (A) Focusing on successful examples and ignoring the failure. |<br>|---|---|<br>| (ii)Linearity Bias | (B) Recommendation Systems |<br>| (iii)Survivor Bias | (C) Common at the data labeling stage of any object |<br>| (iv) Recall Bias | (D) Cognitive Bias |<br><br>    a) (i)-(A), (ii)-(B),(iii)-(C), (iv)-(D)<br>    b) (i)-(B), (ii)-(D),(iii)-(C), (iv)-(A)<br>    c) (i)-(B), (ii)-(D),(iii)-(A), (iv)-(C)<br>    d) (i)-(B), (ii)-(A),(iii)-(D), (iv)-(C) | 1 |

| iv. | Which of the following formula is used to calculate the z-score?<br>    a) $(x-\mu)/\sigma$<br>    b) $(\mu-x)/\sigma$<br>    c) $(\sigma-\mu)/x$<br>    d) $(\mu-\sigma)/x$ | 1 |
|---|---|---|
| v. | i.  The z-score is always negative if the value of z-score lies above the mean.<br>ii.  The z-score is positive if its value is below the mean.<br>iii. The z-score is always positive if the value of z-score lies above the mean.<br>iv. The z-score is negative if its value is below the mean.<br>Identify the correct statements.<br>    a)  i and ii<br>    b)  i and iv<br>    c)  ii and iii<br>    d)  iii and iv | 1 |
| vi. | Which of the following is not an appropriate way of discarding confidential data?<br>    a)  Shredding the data<br>    b)  Cutting the files which contain confidential data<br>    c)  Burning the confidential data<br>    d)  Crumbling the papers which contain confidential data and throwing it in the dustbin | 1 |

| Q. 4 | Answer any 5 out of the given 6 questions (1 x 5 = 5 marks) | |
|---|---|---|
| i. | Statement1: Mean Absolute Deviation (MAD) is the average of how far away all values in a data set are from the mean.<br>Statement2: The Standard Deviation is the measure of how spread out the numbers are around the mean.<br>    a)  Both Statement1 and Statement2 are correct<br>    b)  Both Statement1 and Statement2 are incorrect<br>    c)  Statement1 is correct but Statement2 is incorrect<br>    d)  Statement2 is correct but Statement1 is incorrect | 1 |
| ii. | What is the probability that in a pack of 52 cards, the card taken is 5?<br>    a) 4/52<br>    b)  1/26<br>    c)  8/52<br>    d) 1/169 | 1 |
| iii. | Bias data can have the following properties:<br>    i)   Partial<br>    ii)  Prejudiced<br>    iii)  preferential<br>    iv)  selective<br>    a)  i, iii and iv<br>    b)  ii, iii, and iv<br>    c)  i, ii and iii<br>    d)  i, ii and iv | 1 |

| iv. | Quartiles of dataset partitions the data into_____equal parts.<br>a) Three<br>b) Four<br>c) Two<br>d) Six | 1 |
|---|---|---|
| v. | What is the value of mean and standard deviation in a standard normal distribution?<br>a) 1 and 0<br>b) 0.5 and 0.5<br>c) 0 and 1<br>d) 1 and 1 | 1 |
| vi. | If you soft delete a file<br>a) It gets deleted forever<br>b) One can easily restore it<br>c) The hard disk gets formatted<br>d) A duplicate copy gets generated | 1 |

| Q. 5 | **Answer any 5 out of the given 6 questions (1 x 5 = 5 marks)** | |
|---|---|---|
| i. | The scores of five students in a subject are 72, 82, 94, 85, 74. Find the median. | 1 |
| ii. | The collection of one or more outcomes from an experiment is called_____<br>a) Probability<br>b) Distribution<br>c) Event<br>d) Random Experiment | 1 |
| iii. | Statement1: Probability is a very essential tool in statistics.<br>Statement2: There is no relation between probability and how we make predictions in statistics.<br>a) Both Statement1 and Statement2 are correct<br>b) Both Statement1 and Statement2 are incorrect<br>c) Statement1 is correct but Statement2 is incorrect<br>d) Statement2 is correct but Statement1 is incorrect | 1 |
| iv. | What measures of position divides the distribution into 10 equal parts?<br>a) Quartiles<br>b) Deciles<br>c) Percentiles<br>d) Range | 1 |
| v. | Statement A: A z-score describes the position of a fair score in terms of its distance from the mean when measured in standard deviation units.<br>Statement B: The z-score is positive if the value lies above the mean and negative if it lies below the mean.<br>a) (A)is false but (B) is true<br>b) (A)is true but (B) is false<br>c) Both (A) and (B) are true<br>d) Both (A) and (B) are false | 1 |

| **vi.** | Assertion (A) : While storing the data in your device, you can encrypt the it.<br>Reason (R) : In the case of a data leak, hackers are not able to read your data.<br>    a) Both A and R are correct and R is the correct explanation of A<br>    b) Both A and R are correct but R is not the correct explanation of A<br>    c) A is correct and R is incorrect<br>    d) R is correct and A is incorrect | **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** | Feedback is the final component and one of the most important factors in the process of communication. Give two reasons to justify why feedback is important. | **2** |
|---|---|---|
| **Q. 7** | In SMART goals, what does 'R' stand for? Explain. | **2** |
| **Q. 8** | What will happen if you overcharge the battery of a laptop? | **2** |
| **Q. 9** | Differentiate between wage employment and self-employment. | **2** |
| **Q. 10** | Explain the term: Organic farming | **2** |

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

| **Q. 11** | A survey of 20 people (10 kids and 10 adults) was taken on the type of food they like to snack. The following responses were recorded:<br>    ● 7 kids liked pizza while the other kids preferred burger<br>    ● 5 adults liked burger while the other adults preferred pizza<br>Based on the information given above, build a two-way frequency table. | **2** |
|---|---|---|
| **Q. 12** | Explain uniform distribution with the help of an example. | **2** |
| **Q. 13** | What is survivorship bias? Give an example. | **2** |
| **Q. 14** | Do you think it is helpful to standardize the values of a normal distribution by converting them into z-score? Explain | **2** |
| **Q. 15** | Consider the data set 27,22,44,35,14,34,32<br>Find the percentile for the element 22 | **2** |
| **Q. 16** | Confidential data is maintained in the form of digital as well as physical copies. Suggest two ways in which we can discard the physical copies of confidential data. | **2** |

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

| Q. 17 | Ashish downloaded a dataset from Kaggle and now wants to use parts of it for data analysis. Explain him the term 'subset' and the different ways of subsetting the data. | 4 |
|---|---|---|
| Q. 18 | Manya, a student of class 10 was learning the topic "Statistical Problem Solving Process". She did not understand the concept and had trouble understanding it. Help her by explaining the process with a good example. | 4 |
| Q. 19 | "Given a significantly large sample size from a population with finite variance, the mean of all samples from same set of population will be roughly equal to the mean of the population". What is this statement about? Explain it with the help of an example. | 4 |
| Q. 20 | We can perform data merging by implementing data joins on the databases in frame. How many types of joins are there? Explain. | 4 |
| Q. 21 | Data analytics raises many ethical issues, especially when anyone starts making money from their data externally for the purposes different from the ones for which the data was initially collected. While there are no specific rules for ethical guidelines around data analysis, there are a few principles to agree upon. Explain any four of such principles. | 4 |