

General Instructions

- (i) This question paper contains 65 questions. All questions are compulsory.
- (ii) It comprises 35 single-correct multiple-choice questions, 12 multiple-correct questions and 18 numerical / integer-type questions.
- (iii) Attempt every question; detailed solutions are provided in the companion solutions booklet.
- (iv) For multiple-correct questions, all correct options must be selected to earn full marks.
- (v) For numerical questions, report the answer rounded exactly as asked.

1.

Verbosity : Brevity :: Insolence : _____

Choose the word that best fills the blank.

- (A) Innocence
- (B) Respect
- (C) Solace
- (D) Wealth

...

2.

The product of the digits of a three-digit number is 70. The sum of the digits of this three-digit number is _____

- (A) 12
- (B) 14
- (C) 16
- (D) 18

...

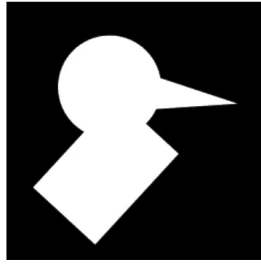
3.

The four pieces of a puzzle are shown in the figure below.



Which one of the figures labelled as P, Q, R, and S can be constructed by using each of the four pieces only once without overlaps?

(A)



P

Figure P

(B)



Q

Figure Q

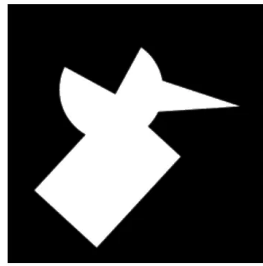
(C)



R

Figure R

(D)



S

Figure S

...

4.

Consider two distinct positive real numbers m, n , with $m > n$.

Let $x = n^{\log_{10}(m)}$ and $y = m^{\log_{10}(n)}$. The relation between x and y is _____.

- (A) $x > y$
- (B) $x < y$
- (C) $x = y$
- (D) $x = \log_{10}(y)$

...

5.

'If his latest movie had been a commercial success, the actor would have made enough money to sponsor his next movie.'

Based only on the above sentence, which one of the following statements is true?

- (A) The actor will certainly sponsor his next movie.
- (B) His latest movie was a commercial success.
- (C) The actor made enough money from his latest movie.
- (D) His latest movie was not commercially successful.

...

6.

'My friend and I parted ___ the door ___ the cabin that I had rented ___ the night.'

Choose the option with the correct sequence of words to fill the blanks.

- (A) at; of; for
- (B) for; at; of
- (C) of; for; in
- (D) in; of; for

• • •

7.

Five integers are picked from 0 to 20, with possible repetitions, such that their mean is 12, median is 18, and they have a single mode of 20.

Ignoring permutations, the number of ways to pick these five integers is _____

- (A) 0
- (B) 1
- (C) 2
- (D) 3

• • •

8.

Rishi and Swathi are students of Class 5. Pavan and Tanvi are students of Class 4. Rishi and Pavan are boys. Swathi and Tanvi are girls. The four students played a total of three games of chess. The games were played one after another. A player who lost a game did not participate in any more games. It was observed that:

1. the first game was the only game where two students of the same class played against each other,
2. the students of Class 5 won more games than the students of Class 4, and
3. the boys won two games and the girls won one game.

The student who did not lose any game is _____.

- (A) Pavan
- (B) Rishi
- (C) Swathi
- (D) Tanvi

• • •

9. P, Q, R, S, X , and Y are distinct single-digit whole numbers taking values from 0 to 9.

PQ is a two-digit number with Q in the units place and P in the tens place. Similarly, RS is a two-digit number.

It is known that PQ and RS are consecutive numbers and $(PQ)^2 + (RS)^2 = XYP$, with XYP being a three-digit number.

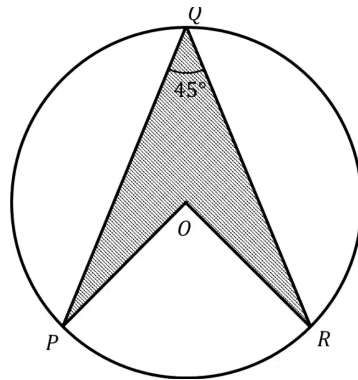
The value of Y is _____.

- (A) 4
- (B) 5
- (C) 6
- (D) 7

• • •

10. In the given figure, P , Q , and R are three points on a circle of radius 10 cm with O as its center, $\overline{PQ} = \overline{RQ}$, and $\angle PQR = 45^\circ$. The figure is representative.

The area of the shaded region $PQRO$ is _____ cm^2 .



- (A) 50
- (B) $25\sqrt{2}$
- (C) $50\sqrt{2}$
- (D) 100

...

11. For a classification problem, Principal Component Analysis (PCA) has been used to reduce the dimensionality of a feature space from 100 to 10.

Which of the following options is true about the angle θ between the first and the tenth principal components?

- (A) $\theta = 0^\circ$
- (B) $\theta = 90^\circ$
- (C) $90^\circ < \theta \leq 180^\circ$
- (D) $0 < \theta < 90^\circ$

...

12. Consider that you are training a classifier for a 10-class classification problem. Each input is represented as a 512-dimensional vector. There are 1000 samples, out of which the first 100 will be used for testing.

Let Leave-One-Out-Cross-Validation (LOOCV) be used for selection of the classifier model before testing.

Which of the following options is the correct number of validation splits that will be generated?

- (A) 10
- (B) 512
- (C) 900
- (D) 1000

...

13. Which of the following algorithms is NOT an example of uninformed search?

- (A) Breadth First Search
- (B) Depth First Search
- (C) A* Search
- (D) Depth-limited Search

• • •

14. Which of the following statements is NOT true? (The names of the predicates are intuitive.)

- (A) $\forall x \forall y \text{ Classmate}(x, y) \Rightarrow \text{Classmate}(y, x)$
- (B) $\forall x \text{ Likes}(x, \text{Icecream}) \Rightarrow \neg \exists x \neg \text{Likes}(x, \text{Icecream})$
- (C) "Each king is a person" is equivalent to $\forall x \text{ IsKing}(x) \wedge \text{IsPerson}(x)$
- (D) "All humans are mortal" is equivalent to $\forall x \text{ IsHuman}(x) \Rightarrow \text{IsMortal}(x)$

• • •

15. Consider that the quick sort algorithm is used to sort an array of n distinct randomly ordered elements. In every call, the pivot is chosen as the first element of the current subarray.

Let $T(n)$ denote the expected time to sort the array. Assume that the time to partition is linear in the size of the current subarray.

Which of the following recurrence relations correctly represents $T(n)$ in this scenario?

- (A) $T(n) = T(1) + T(n - 1) + O(n)$
- (B) $T(n) = T\left(\frac{n}{4}\right) + T\left(\frac{3n}{4}\right) + O(n)$
- (C) $T(n) = 2T\left(\frac{n}{2}\right) + O(n)$
- (D) $T(n) = \frac{1}{n} \sum_{k=0}^{n-1} [T(k) + T(n - k - 1)] + O(n)$

• • •

16. Consider the given Python program.

```
def append_to_lst(val, lst=[]):
    lst.append(val)
    return lst

print(append_to_lst(1))
print(append_to_lst(2))
print(append_to_lst(3, []))
```

Which of the following is the correct output of this program?

- (A) [1]
[2]
[3]
(B) [1]
[1, 2]
[3]
(C) [1]
[2]
[1, 2, 3]
(D) [1]
[1, 2]
[1, 3]

...

17. Let $R(A, B, C, D, E)$ be a relational schema with functional dependency set $F = \{A \rightarrow BC, CD \rightarrow E, E \rightarrow A\}$. Which of the following statements is correct?

- (A) AD , ED and CD are the only candidate keys of R .
(B) AD and ED are the only candidate keys of R .
(C) A , E and CD are the only candidate keys of R .
(D) A and CD are the only candidate keys of R .

...

18. Consider that the visualization of a 3-dimensional data cube is showing Sales Quantity for each combination of the attributes Product Type, Month and Country.

From this, if we want to further visualize the Sales Quantity for each combination of Product Type, Month and State, which of the following OLAP operations should be performed?

- (A) Slicing
(B) Dicing
(C) Roll-up
(D) Drill-down

...

19. Let M be a randomly chosen non-empty subset of $S = \{1, 2, 3, \dots, 2026\}$.

Which of the following is the probability that the product of all the elements of M is even?

- (A) $\frac{2^{1013}(2^{1013} - 1)}{2^{2026}}$
(B) $\frac{2^{1013}}{2^{2026}}$
(C) $\frac{2^{1013}(2^{1013} - 1)}{2^{2026} - 1}$
(D) $\frac{1}{2^{2026} - 1}$

...

20. Suppose that a computer program provides a non-negative and integer-valued random solution to the equation $n_1 + n_2 + n_3 + n_4 = 20$.

Which of the following is the probability that all of n_1, n_2, n_3, n_4 in the provided solution are positive?

- (A) $\binom{19}{3} / \binom{23}{3}$
- (B) $\binom{20}{4} / \binom{24}{4}$
- (C) $\binom{20}{3} / \binom{23}{3}$
- (D) $\binom{19}{4} / \binom{24}{4}$

...

21. Let $M = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ be a 2×2 matrix, where $\theta = \frac{2\pi}{5}$, and $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Which of the following options is equal to M^{2026} ?

- (A) M^2
- (B) M
- (C) M^{-1}
- (D) I_2

...

22. Consider a set $S_1 = \{x = (x_1, x_2, x_3)^T \in \mathbb{R}^3 \mid x^T x \leq 16\}$. Let S_2 be another set which is a subspace of $\mathbb{R}^3$ with dimension two.

Which of the following gives the area of $S_1 \cap S_2$?

- (A) 16π
- (B) 4π
- (C) $4\pi^2$
- (D) $16\pi^2$

...

23. In the following table, the Task column lists a few tasks related to machine learning. The Algorithm column lists a few algorithms.

Each entry "t" from the Task column is to be matched with an appropriate entry "a" from the Algorithm column such that the task "t" can be solved using the algorithm "a". Denote such a match as t:a.

Task	Algorithm
T1 - Clustering	A1 - Markov Chain Monte Carlo
T2 - Classification	A2 - K-Medoid
T3 - Sampling	A3 - Linear Discriminant Analysis
T4 - Feature Extraction	A4 - Naive Bayes

Which of the following options is/are the correct matching(s)?

- (A) T1:A4, T2:A3, T3:A1, T4:A2
- (B) T1:A2, T2:A4, T3:A1, T4:A3
- (C) T1:A3, T2:A4, T3:A1, T4:A2
- (D) T1:A4, T2:A2, T3:A1, T4:A3

• • •

24. Sentence X is said to entail Sentence Y if whenever X is **TRUE**, Y also must hold **TRUE**. Which of the following statements is/are correct if X entails Y ?

- (A) $X \Rightarrow Y$
- (B) $X \wedge \neg Y$ is **FALSE**
- (C) if X then Y
- (D) if Y then X

• • •

25. You are given the following Pre-order and In-order traversals of a Binary Tree T with nodes E, F, G, P, Q, R, S.

Pre-order: P Q S E R F G
In-order: S Q E P F R G

Which of the following statements is/are true about the Binary Tree T ?

- (A) Node P is the root of T
- (B) The Post-order traversal of T is: S E Q F G R P
- (C) Node Q has only one child
- (D) The left subtree of node R contains the node G

• • •

26. Consider two relations r and s defined on the relational schemas $R(A, B)$ and $S(E, C)$, respectively. A is the primary key of R and E is a foreign key of S referencing A in R .

Which of the following operations will NEVER violate the foreign key constraint?

- (A) Inserting records into relation r
- (B) Deleting records from relation s
- (C) Deleting records from relation r
- (D) Inserting records into relation s

• • •

27. Let $f(x) = x^3 - 3x^2 + 2$ be a function defined on $(-1, 3]$.

Which of the following statements is/are correct?

- (A) $f(x)$ has exactly two roots in $[-0.9, 0]$.
- (B) $f(x)$ has a minimum at 2 only.
- (C) $f(x)$ has maximum at 0 only.
- (D) $f(x)$ has a root at 1.

• • •

28. Suppose a random variable Z follows $Normal(\mu = 0, \sigma^2 = 1)$ distribution with probability density function $g(z)$ and cumulative distribution function $G(z)$. Another random variable Y follows t_1 distribution with probability density function $h(y)$ and cumulative distribution function $H(y)$. Let c be the positive real number for which $g(c) = h(c)$.

Which of the following statements is/are correct?

- (A) $G(0) = H(0)$
- (B) $G(c) < H(c)$
- (C) $G(-c) < H(-c)$
- (D) $g(0) = h(0)$

• • •

29. Consider that for a supervised learning task, the objective function being minimized is $f_w(x) = wx$, where $x \in \mathbb{R}$ is the input and $w \in \mathbb{R}$ is the parameter. Stochastic Gradient Descent with a learning rate of 0.10 is used for parameter updates.

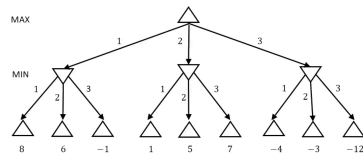
Suppose that at the end of iteration i , the value of w becomes 10.00.

Let $x = 10.00$ be the input for iteration $(i + 1)$.

The value of w at the end of iteration $(i + 1)$ is _____. (Rounded off to two decimal places)

• • •

30. Consider the game tree for a two-player turn-taking minimax game as shown in the figure. The value of a terminal node represents the utility of the game state if the game ends there. The numbers written next to the edges denote the strategies.



There are two players MAX and MIN. At any particular state of the game, MAX prefers to move to a state of maximum value. On the other hand, MIN prefers to move to a state of minimum value.

Suppose MAX starts the game at the root and has three strategies: 1, 2 and 3. Next, MIN plays and also has three strategies: 1, 2 and 3. The game ends there. Both players always take optimal strategies throughout the game.

At the root, the best strategy for MAX is _____. (Answer in integer)

...

31. Let A be a sorted array containing 1000 distinct integers. You perform a recursive binary search on A to find an element y . Suppose each comparison checks whether the middle element computed during the current recursive step is equal to, less than, or greater than y .

The maximum number of comparisons that may have to be performed if y is not an element of A is _____. (Answer in integer)

...

32. In a relational database, a B+ Tree Index is to be constructed for a relation on a key field. In a B+ Tree, a Node Pointer points to a sub-tree and a Data Record Pointer points to a block of database records.

Let, Node size = 4096 bytes, Node Pointer size = 10 bytes, Search Key Field size = 11 bytes and Data Record Pointer size = 12 bytes.

The maximum number of Node Pointers that can be present in a non-leaf node of the B+ Tree is _____. (Answer in integer)

...

33. The number of bijections $f(\cdot)$ from the set $S = \{1, 2, 3, 4\}$ to itself such that $f(f(n)) = n$, for all $n \in S$, is _____. (Answer in integer)

...

34. Let X be an exponentially distributed random variable with mean $\lambda(> 0)$. If $P(X > 5) = 0.35$, then the conditional probability $P(X > 10 \mid X > 5)$ is _____. (Rounded off to two decimal places)

• • •

35. The value of $\sum_{i=0}^{\infty} \sum_{j=1}^{\infty} 2^{-i} 3^{-j}$ is _____. (Answer in integer)

• • •

36. Let four points in three-dimensional space be:

$P_1 : [2, 3, -1]$, $P_2 : [3, 1, 1]$, $P_3 : [5, -2, 3]$ and $P_4 : [3, 3, 3]$.

Hierarchical Agglomerative Clustering is used to cluster the above points. If Manhattan Distance is used as the distance metric during clustering, which of the following options indicates the two points that will be merged first?

- (A) P_1, P_2
- (B) P_2, P_3
- (C) P_3, P_4
- (D) P_2, P_4

• • •

37. Which of the following statements is true for Ridge Regression?

- (A) The regularizer in the objective function of Ridge Regression is used to guard against scenarios where the model works well for the test data, but poorly for the training data.
- (B) The regularizer of Ridge Regression uses L_1 norm.
- (C) Ridge Regression aims to reduce the number of parameters that have negative values.
- (D) The regularizer of Ridge Regression may increase the bias of the model, but it helps in reducing the variance in predictions.

• • •

38. Assume that a Creative (C) person will Succeed (S) if the person is also Disciplined (D), but will not succeed otherwise. Now, consider the following statements:

- (i) $C \wedge S \Leftrightarrow D$
- (ii) $C \Rightarrow (S \Leftrightarrow D)$
- (iii) $C \Leftrightarrow ((D \Rightarrow S) \vee \neg S)$

Which of the following options is correct?

- (A) Both (i) and (ii) are TRUE
- (B) Only (ii) is TRUE
- (C) Both (ii) and (iii) are TRUE
- (D) Only (iii) is TRUE

• • •

39. A recursive function in Python is given.

```
def mystery(n):  
    if n <= 0:  
        return 1  
    else:  
        return mystery(n-1) + mystery(n-2)
```

Now, consider the following function call:

mystery(4)

Assume that a typical runtime stack is used to manage function calls. Each function call is pushed onto the stack and removed only after it finishes execution.

Which of the following options denotes the total number of function calls (i.e., the total number of stack activations), including the initial call, to compute mystery(4)?

- (A) 5
- (B) 9
- (C) 15
- (D) 17

• • •

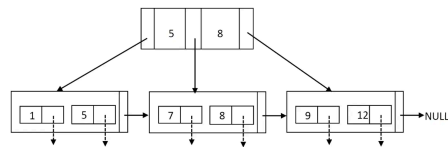
40. Consider a directed graph $G = (V, E)$, where V is the finite set of vertices and E is the set of directed edges between the vertices. G may contain cycles but there is no self-loop. Further, G may not be strongly connected. Let G^R be the graph obtained by reversing the directions of all the edges in G without changing the set of vertices. Assume that Breadth First Search (BFS) or Depth First Search (DFS) from any given vertex v of a graph visits only the reachable vertices from v in that graph.

Which of the following statements must always be true, regardless of the structure of G ?

- (A) If u is a reachable vertex in the BFS of G^R from v , then u is also a reachable vertex in the DFS of G from v .
- (B) In G^R , the BFS traversal from v will visit exactly the same set of vertices as the DFS from v in G .
- (C) The order of vertices visited in the BFS of G^R from v is the reverse of the order of vertices visited in the DFS of G from v .
- (D) If u is a reachable vertex in the DFS of G from v , then v is also a reachable vertex in the BFS of G^R from u .

• • •

41. Consider a B+ Tree where the maximum number of key values in each leaf node is 2 and the maximum number of pointers in each non-leaf node is 3. Let the content of the B+ Tree be as shown in the figure.



Which of the following options denotes the key value(s) stored in the root node after inserting a key value 3 in the given B+ Tree?

- (A) 5
- (B) 8
- (C) 3 and 5
- (D) 3, 5 and 8

...

42. Consider the given relations **X**, **Y** and **Z**. The relation **X** has three columns P, Q and R. The relation **Y** has three columns P, Q and S. The relation **Z** has two columns P and T.

Relation X:

P	Q	R
P1	Q1	R1
P2	Q2	R2
P3	Q3	R2

Relation Y:

P	Q	S
P1	Q1	2
P1	Q2	5
P2	Q1	6
P3	Q3	1

Relation Z:

P	T
P1	T1
P3	T2
P4	T3
P4	NULL

Consider the relational algebra expression

$$\Pi_{P,R,S} \left[\left(\sigma_{(Q=Q3 \vee R=R2)} [X \bowtie Y] \right) \bowtie \left(\sigma_{(S>1)} [Y \bowtie Z] \right) \right]$$

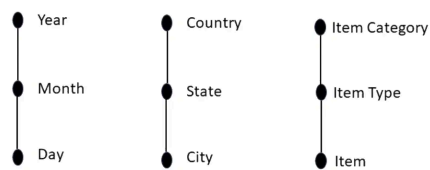
where $\bowtie$ denotes natural join operation.

Which of the following options is the correct output for the given expression?

- (A) Two rows (P1, R1, 2) and (P1, R1, 5)
- (B) Three rows (P1, R1, 2), (P1, R1, 5) and (P2, R2, 6)
- (C) One row (P1, R1, 2)
- (D) Zero rows

...

43. Consider the concept hierarchies as shown in the figure.



Which of the following options denotes the total number of possible data cuboids from these concept hierarchies?

- (A) 4^3
- (B) 2^3
- (C) $2!$
- (D) $4!$

...

44. Let X and Y be two independent random variables. X follows *Bernoulli*($p = 0.3$) distribution and Y follows *Normal*($\mu = 0, \sigma^2 = 100$) distribution.

Which of the following options is the variance of $(2X - 1)Y$?

- (A) 100
- (B) 90
- (C) 49
- (D) 21

...

45. Let

$$L = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{e^{-n} n^k}{k!}$$

Which of the following is the value of L ?

- (A) 0.5
- (B) 1.0
- (C) 0
- (D) e^{-1}

...

46. Let $\gamma_1, \gamma_2, \gamma_3$ be the eigenvalues of the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos t & \sin t \\ 0 & -\sin t & \cos t \end{bmatrix}$$

where $t \in [-\pi, \pi]$ is in radians.

Which one of the following options lists all the possible values of t satisfying $\gamma_1 + \gamma_2 + \gamma_3 = 1 + \sqrt{2}$?

- (A) $\{\pi/3, -\pi/4\}$
- (B) $\{\pi/4, -\pi/3\}$
- (C) $\{\pi/4, -\pi/4\}$
- (D) $\{\pi/3, -\pi/3\}$

• • •

47. Consider that 20 stories of Author X and 10 stories of Author Y were kept together without mentioning the names of the authors. A classifier was then asked to predict the author (X or Y) of each of these stories. Let, out of X's stories, 6 were classified as that of Y. On the other hand, out of Y's stories, 2 were classified as that of X.

Considering X and Y as two classes, which of the following statements is/are true?

- (A) Accuracy of the classifier is 11/15.
- (B) Precision of Class X is higher than the Precision of Class Y.
- (C) Recall of Class X is higher than the Recall of Class Y.
- (D) Accuracy of the classifier is 14/15.

• • •

48. Let $P(x)$ be a predicate.

Which of the following statements is/are NOT valid in first-order logic?

- (A) $\forall x P(x) \Rightarrow \exists x P(x)$
- (B) $\exists x P(x) \Rightarrow \forall x P(x)$
- (C) $\exists x P(x) \Leftrightarrow \forall x P(x)$
- (D) $\forall x P(x) \Rightarrow \exists x \neg P(x)$

• • •

49. Consider the problem of sorting the given array in ascending order:

$P = [1, 2, 3, 5, 4]$

Consider two sorting algorithms Bubble Sort (BS) and Insertion Sort (IS).

Let N_1 be the total number of comparisons done by BS on the elements of P and N_2 be the total number of comparisons done by IS on the elements of P .

Which of the following options is/are correct?

- (A) $N_1 = 10, N_2 = 4$
- (B) $N_1 > N_2$
- (C) IS on P will perform only one swap
- (D) Both BS and IS on P will make at least one unnecessary comparison (i.e., comparing elements that are already in correct order)

• • •

50. Consider the given Python program.

```
def outer():
    x = []
    def inner(val):
        x.append(val)
        return x
    return inner

f1 = outer()
f2 = outer()
print(f1(10))    # Line P
print(f1(20))    # Line Q
print(f2(30))    # Line R
print(f1(40))    # Line S
```

Which of the following options is/are correct?

- (A) f_1 and f_2 share the same list x
- (B) Output of Line Q is $[10, 20]$
- (C) Output of Line R is $[10, 20, 30]$
- (D) Output of Line S is $[10, 20, 40]$

• • •

51. Consider a table Employee(EmpID, TeamID), where the column EmpID (ID of an employee) is the primary key. The column TeamID denotes the team ID of the team of which the employee is a member. TeamID is a NOT NULL column.

We want to display the size of the team (denoted as TeamSize) in which each employee is a member by using SQL. As an example, the desired output for the given Employee table is also shown in tabular form.

Which of the following is/are correct?

Employee table:

EmpID	TeamID
-------	--------

1	8
2	8
3	8
4	7
5	7
6	9

Output table:

EmpID	TeamSize
-------	----------

1	3
2	3
3	3
4	2
5	2
6	1

(A)

```
SELECT E.EmpID, B.TeamSize
FROM Employee AS E, (SELECT TeamID, COUNT(TeamID) AS TeamSize FROM Employee GROUP BY TeamID) AS B
WHERE E.TeamID = B.TeamID
```

(B)

```
SELECT A.EmpID, COUNT(B.TeamID) AS TeamSize
FROM Employee AS A, Employee AS B
WHERE A.TeamID = B.TeamID AND A.EmpID = B.EmpID
GROUP BY A.EmpID
```

(C)

```
SELECT B.EmpID, B.TeamSize
FROM (SELECT EmpID, COUNT(TeamID) AS TeamSize
FROM Employee GROUP BY EmpID) AS B
```

(D)

```
SELECT A.EmpID, B.TeamSize
FROM Employee AS A, (SELECT COUNT(TeamID) AS TeamSize FROM Employee GROUP BY TeamID) AS B
WHERE A.TeamID = B.TeamID
```

• • •

52. Let $M = (I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T)$ be a matrix, where $\mathbf{1} = (1, 1, 1, \dots, 1)^T \in \mathbb{R}^n$ and I_n is the identity matrix of order n .

Which of the following options is/are correct?

- (A) $M^T = M$
- (B) $M^2 = I_n$
- (C) $\text{Trace}(M) = n$
- (D) M is a projection matrix

• • •

53. Let $X_1, X_2, \dots, X_n$ be n independent random variables. Each of the random variables follows $\text{Normal}(\mu = 0, \sigma^2 = 1)$ distribution. Define $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

Which of the following statements is/are correct?

- (A) $\sum_{i=1}^n X_i^2$ follows Chi-square distribution with n degrees of freedom.
- (B) $\sum_{i=1}^n (X_i - \bar{X})^2$ follows Chi-square distribution with $(n - 1)$ degrees of freedom.
- (C) $X_1^2 + X_n^2$ follows exponential distribution with mean 2.
- (D) $(\sqrt{n}\bar{X})^2$ follows Chi-square distribution with 2 degrees of freedom.

• • •

54. Let X be a discrete valued random variable with cumulative distribution function $F(x)$.

Which of the following statements is/are correct?

- (A) $F(x)$ is always a positive function.
- (B) $F(x)$ is a non-decreasing function.
- (C) $F(x)$ has jump discontinuity.
- (D) $F(x)$ is a left continuous function.

• • •

55. Consider that Linear Ridge Regression is being used to learn a prediction function $y_{pred} = w^T x$, where $w, x \in \mathbb{R}^2$ and Mean Absolute Error (MAE) is used to measure the prediction error. A weight of 0.20 is associated with the regularizer.

At an intermediate step of the training process, assume that the parameter $w = [-3.00, 4.00]^T$. In the next step, for the input $x = [1.00, 2.00]^T$, the predicted value of y is noted. Let the relation between $x = [x_1, x_2]^T$ and the true value of y be $y_{true} = x_1 + x_2$.

The value of the overall regularized loss function for this instance is _____. (Rounded off to two decimal places)

• • •

56. Consider a fully-connected feed-forward multi-layer perceptron. It has 30 neurons in the input layer, followed by two hidden layers and an output layer. The first hidden layer has 4 neurons and the second hidden layer has 3 neurons. The output layer has only one neuron. Assume that no bias parameters are used.

The number of learnable parameters in the multi-layer perceptron is _____. (Answer in integer)

• • •

57. A clinic specializes in testing for a disease D . The result of the test can be either positive or negative.

A study revealed that if a person suffers from the disease D , the test result in that clinic comes out positive 80% of the time, and negative 20% of the time. If a person is not suffering from the disease D , the test comes out positive 10% of the time and negative 90% of the time. It is also known that among the general population, the disease D occurs in 30% of the individuals.

If a person tests positive for D in that clinic, the probability that he/she actually suffers from the disease D is _____. (Rounded off to two decimal places)

• • •

58. Consider the given Python program.

```
def fun(L, i=0):
    if i >= len(L)-1:
        return 0
    if L[i] > L[i+1]:
        L[i+1], L[i] = L[i], L[i+1]
        return 1+fun(L, i+1)
    else:
        return fun(L, i+1)

data = [5, 3, 4, 1, 2]
count = 0
for _ in range(len(data)):
    count += fun(data)
print(count)
```

The output of the program is _____. (Answer in integer)

...

59. Let there be two relations X and Y as shown. X has three columns P , Q and R . Y has two columns P and S .

Relation X:

P	Q	R
P1	Q1	R1
P2	Q2	R2
P3	Q3	R2

Relation Y:

P	S
P1	10
P1	15
P2	20
P3	1

Consider that the following tuple relational calculus expression is evaluated.

$$\{t \mid t \in X \wedge \exists z \in X (t[P] = z[P]) \wedge \exists m \in Y (m[P] = t[P] \wedge m[S] > 1)\}$$

The number of tuples that will be returned is _____. (Answer in integer)

...

60. Let Account be a relation as shown.

Table: Account

AccNo Balance

A1	5000
A2	5000
A3	10000
A4	15000
A5	18000

Consider the given SQL query.

```
SELECT AccNo FROM Account AS A
WHERE (SELECT COUNT(*) FROM Account AS B
WHERE A.Balance < B.Balance) >= (SELECT COUNT(*)
FROM Account AS C WHERE A.Balance > C.Balance)
```

The number of rows returned by the SQL query is _____. (Answer in integer)

...

61. Consider an ER model with the entities E1(A11, A12, A13) and E2(A21, A22, A23), where A11, A12, A13 are the attributes of E1, and A21, A22, A23 are the attributes of E2. Let A22 be a multi-valued attribute. A11 and A21 are the primary keys of E1 and E2, respectively.

Let R12 be a many-to-many relationship between E1 and E2. Participation of both E1 and E2 in R12 is total.

The minimum number of relations required to convert the ER model into relational model (assuming there is no other functional dependency) where each relation is in third normal form (3NF) is _____. (Answer in integer)

...

62. For a given data set $\{x_1, x_2, \dots, x_n\}$, where $n = 100$, it is known that

$$\frac{1}{2000} \sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2 = 99$$

Let us denote $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$

The value of $\frac{1}{99} \sum_{i=1}^n (x_i - \bar{x})^2$ is _____. (Answer in integer)

...

63. Let X be a random variable that follows $Uniform(-1, 1)$ distribution. The conditional distribution of the random variable Y given $X = x$ is the $Uniform(x^2 - 0.1, x^2 + 0.1)$ distribution.

The value of $correlation(X, Y)$ is _____. (Answer in integer)

• • •

64. Let $A_{5 \times 5}$ be a matrix such that each of its elements follows $Bernoulli(p = 0.50)$ distribution independently.

The probability that the row-sum of the second row and the column-sum of the third column are both equal to 3 is _____. (Rounded off to two decimal places)

• • •

65. Let $A = (I_n - \frac{1}{n}\mathbf{1}\mathbf{1}^T)$ be a matrix, where $\mathbf{1} = (1, 1, 1, \dots, 1)^T \in \mathbb{R}^n$ and I_n is the identity matrix of order n .

The value of $\max_S x^T A x$, where $S = \{x \in \mathbb{R}^n \mid x^T x = 1\}$, is _____. (Answer in integer)