CUET-UG Information Practices Sample Paper-14

Duration: 1 Hour **Maximum Marks: 250**

Instructions

- This paper contains a total of 50 Multiple Choice Questions.
- Each correct answer carries **+5 marks**.
- Each incorrect answer carries **-1 mark**.
- No negative marking for unattempted questions.

Q1. Consider a SQL table `Inventory` with columns `(ItemID, Category, Stock, Price)`. A DBA executes the following query: `SELECT Category, SUM(Stock * Price) FROM Inventory WHERE Stock > 10 GROUP BY Category HAVING SUM(Stock * Price) > 50000 ORDER BY 2 DESC;` Identify the correct sequence of execution for the clauses in this query.

(A) SELECT → FROM → WHERE → GROUP BY → HAVING → ORDER BY

(B) FROM → WHERE → GROUP BY → HAVING → SELECT → ORDER BY

(C) FROM → SELECT → WHERE → GROUP BY → HAVING → ORDER BY

(D) WHERE → FROM → GROUP BY → SELECT → HAVING → ORDER BY

Q2. What will be the output of the following SQL command? `SELECT ROUND(1294.567, -2), TRUNCATE(1294.567, -1);`

(A) 1300, 1290

(B) 1200, 1290

(C) 1300, 1200

(D) 1295, 1294.5

Q3. Match List I (SQL String Functions) with List II (Mathematical Logic/Output) for the string `'Informatics'`:

List I (Function)	**List II (Output)**
(I) `INSTR('Informatics', 'format')`	(1) 7
(II) `LENGTH(SUBSTR('Informatics', 3, 5))`	(2) 3
(III) `INSTR('Informatics', 'i')`	(3) 5
(IV) `MID('Informatics', -5, 3)`	(4) `'mat'`

(A) I-2, II-3, III-1, IV-4

(B) I-3, II-3, III-1, IV-4

(C) I-2, II-1, III-1, IV-4

(D) I-3, II-3, III-7, IV-4

Q4. Given below are Two Statements, Choose correct:
Statement 1: The `MOD(x, y)` function in SQL returns the same result as the `%` operator in Python for all integer inputs.
Statement 2: `SELECT MOD(11, -3);` in MySQL results in 2.

(A) Both Statement 1 and Statement 2 are correct.

(B) Both Statement 1 and Statement 2 are incorrect.

(C) Statement 1 is correct but Statement 2 is incorrect.

(D) Statement 1 is incorrect but Statement 2 is correct.

Q5. In Relational Algebra, if relation R has degree n_1 and cardinality m_1, and relation S has degree n_2 and cardinality m_2, what are the properties of the relation $T = \sigma_{condition}(R \times S)$?

(A) Degree: $n_1 + n_2$, Cardinality: $m_1 \times m_2$

(B) Degree: $n_1 + n_2$, Cardinality: $\leq (m_1 \times m_2)$

(C) Degree: $\leq (n_1 + n_2)$, Cardinality: $m_1 \times m_2$

(D) Degree: $n_1 \times n_2$, Cardinality: $m_1 + m_2$

Q6. A network administrator wants to connect two different office buildings using a dedicated high-speed link. Building A uses a Star topology and Building B uses

a Bus topology. To allow communication between these two different network architectures at the network layer, which device is necessary?

(A) Bridge

(B) Switch

(C) Router

(D) Repeater

Q7. Which of the following describes the correct difference between an IP Address and a MAC Address?

(A) IP is 48-bit, MAC is 32-bit.

(B) IP is permanent, MAC is temporary.

(C) IP is assigned by the manufacturer, MAC is assigned by the ISP.

(D) IP is a logical address, MAC is a physical hardware address.

Q8. Case Study: A Series object `S` is created as
`pd.Series([10, 20, 30, 40], index=['a', 'b', 'c', 'd'])`.
Statement 1: `S['a':'c']` will return 3 elements.
Statement 2: `S[0:2]` will return 2 elements.

(A) Both Statement 1 and Statement 2 are correct.

(B) Both Statement 1 and Statement 2 are incorrect.

(C) Statement 1 is correct but Statement 2 is incorrect.

(D) Statement 1 is incorrect but Statement 2 is correct.

Q9. Identify the correct Python command to create a DataFrame from a list of dictionaries
`L = [{'a':1, 'b':2}, {'a':5, 'b':10, 'c':20}]`.
What will be the value at
`df.iloc[0, 2]`?

(A) `df = pd.DataFrame(L)`; Value is `0`

(B) `df = pd.DataFrame(L)`; Value is `NaN`

(C) `df = pd.Series(L)`; Value is `None`

(D) `df = pd.DataFrame(L)`; Value is `20`

Q10. Which attribute of a Pandas DataFrame is used to get a single-dimensional array of all values, excluding labels and headers?

(A) `.index`

(B) `.columns`

(C) `.values`

(D) `.axes`

Q11. In Matplotlib, to create a histogram where the data is divided into 15 equal intervals, which parameter should be used in the `plt.hist()` function?

(A) `intervals=15`

(B) `bins=15`

(C) `range=15`

(D) `split=15`

Q12. A user wants to plot a Pie chart in Matplotlib. To highlight the second slice by moving it slightly away from the center, which parameter and data type should be passed?

(A) `distance=[0, 0.2, 0, 0]`

(B) `explode=(0, 0.1, 0, 0)`

(C) `separate={2: 0.1}`

(D) `detach=[0, 1, 0, 0]`

Q13. Which type of Intellectual Property Right (IPR) protects an original software code from being copied or redistributed without permission?

(A) Patent

(B) Trademark

(C) Copyright

(D) Industrial Design

Q14. Match List I (Societal Impact Terms) with List II (Definitions):

List I	List II
(I) Phishing	(1) Stealing ideas/work of others and passing as own
(II) Plagiarism	(2) Fraudulent attempt to obtain sensitive info via email
(III) Digital Footprint	(3) Legal right to use software
(IV) Software License	(4) Trail of data left by users on the internet

(A) I-2, II-1, III-4, IV-3

(B) I-1, II-2, III-4, IV-3

(C) I-2, II-1, III-3, IV-4

(D) I-4, II-1, III-2, IV-3

Q15. Which SQL function would you use to find the day of the week (1-7) for a column `JoinDate`?

(A) `DAYNAME()`

(B) `WEEKDAY()`

(C) `DAYOFWEEK()`

(D) $\texttt{DATE}_P ART()$

Q16. A DataFrame `df` has 100 rows and 5 columns. A user executes `new_df = df.iloc[10:20, 1:4]`.
What will be the value of
`new_df.size` and `new_df.shape`?

(A) Size: 30, Shape: (10, 3)

(B) Size: 40, Shape: (10, 4)

(C) Size: 50, Shape: (10, 5)

(D) Size: 33, Shape: (11, 3)

Q17. Which of the following Pandas operations will result in a `Series` object rather than a `DataFrame`?

(A) `df[['Col1']]`

(B) `df.loc[:, ['Col1', 'Col2']]`

(C) `df.iloc[5, :]`

(D) `df.head(1)`

Q18. Identify the correct Python code to export a DataFrame `df` to a CSV file named `data.csv`, ensuring that the row indices are not saved and the column separator is a semicolon (`;`).

(A) `df.to_csv("data.csv", index=False, sep=';')`

(B) `df.save_csv("data.csv", index=None, delimiter=';')`

(C) `df.write_csv("data.csv", header=False, sep=';')`

(D) `df.to_csv("data.csv", row_index=False, sep=';')`

Q19. In SQL, consider the query:
`SELECT Dept, COUNT(*) FROM Emp GROUP BY Dept HAVING AVG(Sal) > 5000;`.
If the `WHERE` clause was to be added to filter employees with `Age > 25`, where should it be placed?

(A) After the `GROUP BY` clause

(B) Before the `FROM` clause

(C) After the `HAVING` clause

(D) Before the `GROUP BY` clause

(E) Before the `GROUP BY` clause

Q20. Match List I (SQL Join Types) with List II (Result Logic):

List I	List II
(I) Equi Join	(1) Returns all rows from both tables, matching or not
(II) Natural Join	(2) Uses '=' operator and keeps duplicate columns
(III) Left Join	(3) Removes duplicate columns automatically
(IV) Full Join	(4) Returns all rows from left table and matched rows from right

(A) I-2, II-3, III-4, IV-1

(B) I-3, II-2, III-1, IV-4

(C) I-2, II-1, III-4, IV-3

(D) I-1, II-4, III-3, IV-2

Q21. What will be the result of the SQL expression
`SELECT INSTR('PRE-EXAMINATION', '-');` and how does it differ from
`SELECT SUBSTR('PRE-EXAMINATION', 5, 4);`

(A) 4 and 'EXAM'

(B) 3 and 'EXAM'

(C) 4 and 'EXAMI'

(D) 0 and 'PRE-'

Q22. Statement 1: The `plt.bar()` function creates a vertical bar chart, while `plt.barh()` creates a horizontal bar chart. Statement 2: The `legend()` function in Matplotlib is used to provide labels for the x and y axes.

(A) Both Statement 1 and Statement 2 are correct.

(B) Both Statement 1 and Statement 2 are incorrect.

(C) Statement 1 is correct but Statement 2 is incorrect.

(D) Statement 1 is incorrect but Statement 2 is correct.

Q23. A user is creating a histogram. If they want the edge color of the bins to be visible as black lines, which argument should be added to `plt.hist(data)`?

(A) `border='black'`

(B) `edgecolor='black'`

(C) `line_color='black'`

(D) `bin_style='black'`

Q24. Which ethical concern involves a user leaving a trace of their personal information unintentionally while browsing the web?

(A) Active Digital Footprint

(B) Passive Digital Footprint

(C) Phishing

(D) Identity Theft

Q25. Match List I (Device) with List II (OSI Layer/Function):

List I	**List II**
(I) Repeater	(1) Connects similar networks using MAC addresses
(II) Gateway	(2) Regenerates signals to prevent attenuation
(III) Bridge	(3) Connects dissimilar networks with different protocols
(IV) Hub	(4) Broadcasts incoming data to all ports

(A) I-2, II-3, III-1, IV-4

(B) I-1, II-2, III-3, IV-4

(C) I-2, II-1, III-3, IV-4

(D) I-4, II-3, III-2, IV-1

Q26. Which of the following is a legal category of E-waste under the E-Waste (Management) Rules?

(A) Glass and Ceramics only

(B) IT and Telecommunication equipment

(C) Plastic packaging only

(D) Biodegradable paper waste

Q27. Consider a DataFrame `df`. Which command will display the last 3 rows and all the columns of the DataFrame?

(A) `df.tail(3)`

(B) `df.head(-3)`

(C) `df.iloc[-3:, :]`

(D) Both (A) and (C)

Q28. What is the correct SQL syntax to find the total salary for each department, but only for those departments where the maximum salary is above ₹ 50000?

(A) `SELECT Dept, SUM(Sal) FROM Emp WHERE MAX(Sal) > 50000 GROUP BY Dept;`

(B) `SELECT Dept, SUM(Sal) FROM Emp GROUP BY Dept HAVING MAX(Sal) > 50000;`

(C) `SELECT Dept, SUM(Sal) FROM Emp HAVING MAX(Sal) > 50000;`

(D) `SELECT Dept, SUM(Sal) FROM Emp GROUP BY Dept WHERE MAX(Sal) > 50000;`

Q29. A network uses a topology where every node is connected to a central hub/switch. If the central switch fails, the whole network goes down. Identify the topology.

(A) Bus

(B) Star

(C) Mesh

(D) Tree

Q30. Which attribute of a Pandas Series returns the number of bytes consumed by the data in the Series?

(A) `.size`

(B) `.nbytes`

(C) `.memory_usage()`

(D) `.bytes`

Q31. A DataFrame `df_sales` contains sales data. A user wants to calculate the average sales for each region, but only for those regions where the total number of transactions is greater than 100. Which Pandas code snippet is correct?

(A) `df_sales.groupby('Region').filter(lambda x: len(x) > 100).groupby`

(B) `df_sales[df_sales['Sales'].count() > 100].groupby('Region').mean(`

(C) `df_sales.groupby('Region').mean() if count > 100`

(D) `df_sales.where(df_sales.count() > 100).groupby('Region').mean()`

Q32. Which of the following describes a "Relational Algebra" operation that yields a relation consisting of all tuples appearing in R but not in S, provided R and S are union-compatible?

(A) Intersection ($\cap$)

(B) Set Difference ($-$)

(C) Cartesian Product ($\times$)

(D) Projection (π)

Q33. Identify the correct sequence of steps to create and display a line plot using data from a CSV file `growth.csv` with columns `'Year'` and `'GDP'`.

(A) `plt.show()` $\rightarrow$ `pd.read_csv()` $\rightarrow$ `plt.plot()`

(B) `pd.read_csv()` $\rightarrow$ `plt.plot(df['Year'], df['GDP'])` $\rightarrow$ `plt.show()`

(C) `plt.plot()` $\rightarrow$ `plt.xlabel()` $\rightarrow$ `plt.show()`

(D) `pd.read_csv()` $\rightarrow$ `plt.show()` $\rightarrow$ `plt.plot()`

Q34. Which property of a database ensures that once a transaction is committed, it remains so, even in the event of a system failure (e.g., power outage)?

(A) Atomicity

(B) Consistency

(C) Isolation

(D) Durability

Q35. Match List I (Protocols) with List II (Primary Function):

List I	List II
(I) HTTP	(1) Sending emails from a client to a server
(II) FTP	(2) Transferring web pages and media files
(III) SMTP	(3) Securely logging into a remote computer
(IV) SSH	(4) Uploading and downloading files to a server

(A) I-2, II-4, III-1, IV-3

(B) I-4, II-2, III-1, IV-3

(C) I-2, II-1, III-4, IV-3

(D) I-1, II-4, III-2, IV-3

Q36. A user executes `df.drop(['City'], axis=1)` on a DataFrame. How does this affect the original DataFrame `df`?

(A) The 'City' column is permanently deleted from `df`.

(B) The 'City' row is deleted from `df`.

(C) `df` remains unchanged; a new object is returned unless `inplace=True` is used.

(D) An error occurs because `axis=1` refers to rows.

Q37. What is the output of `SELECT SUBSTR("COMPUTER SCIENCE", -7, 4);`

(A) `"SCIE"`

(B) `"CIEN"`

(C) `"UTER"`

(D) `"ENCE"`

Q38. In Matplotlib, which function is used to add a grid to the plot background for better readability of data points?

(A) `plt.background(grid=True)`

(B) `plt.grid()`

(C) `plt.show_grid()`

(D) `plt.draw_lines()`

Q39. Statement 1: A Primary Key can contain NULL values as long as they are unique. Statement 2: A table can have multiple Candidate Keys but only one Primary Key.

(A) Both Statement 1 and Statement 2 are correct.

(B) Both Statement 1 and Statement 2 are incorrect.

(C) Statement 1 is correct but Statement 2 is incorrect.

(D) Statement 1 is incorrect but Statement 2 is correct.

Q40. Case Study: A school uses a software to track attendance. The software is licensed such that the source code is available to everyone to modify and distribute. This is an example of:

(A) Proprietary Software

(B) Open Source Software

(C) Shareware

(D) Freeware

Q41. Identify the correct SQL command to display names of employees whose second character is 'a' and whose name ends with 'n'.

(A) `SELECT Name FROM Emp WHERE Name LIKE '%a%n';`

(B) `SELECT Name FROM Emp WHERE Name LIKE '_a%n';`

(C) `SELECT Name FROM Emp WHERE Name LIKE '_a_n';`

(D) `SELECT Name FROM Emp WHERE Name LIKE 'a%n';`

Q42. Which type of cybercrime involves flooding a server with excessive traffic to make it unavailable to legitimate users?

(A) Phishing

(B) DoS (Denial of Service) attack

(C) Hacking

(D) Eavesdropping

Q43. What will be the result of `df.count()` on a DataFrame?

(A) Total number of cells (including NaN).

(B) Total number of non-null values for each column.

(C) Total number of rows in the DataFrame.

(D) Total number of unique values in the DataFrame.

Q44. In a Star topology, if the central node is a Hub, how does data travel from Node A to Node B?

(A) Directly from A to B via a dedicated cable.

(B) From A to the Hub, and then broadcast to all nodes including B.

(C) From A to the Hub, and then unicast only to B.

(D) Through a ring of intermediate nodes.

Q45. Which SQL aggregate function is used to calculate the number of values in a column, excluding NULLs?

(A) `TOTAL()`

(B) `COUNT(*)`

(C) `COUNT(column_name)`

(D) `SUM()`

Q46. Match List I (IPR Violations) with List II (Scenarios):

List I	**List II**
(I) Trademark Infringement	(1) Copying a book and selling it without permission
(II) Patent Infringement	(2) Using a famous logo on a fake product
(III) Copyright Infringement	(3) Using a patented formula to make medicine
(IV) Plagiarism	(4) Submitting a friend's essay as your own

(A) I-2, II-3, III-1, IV-4

(B) I-1, II-3, III-2, IV-4

(C) I-2, II-1, III-3, IV-4

(D) I-4, II-3, III-2, IV-1

Q47. A DataFrame attribute `df.T` is used for what purpose?

(A) To return the top 5 rows.

(B) To transform the data into a list.

(C) To transpose the DataFrame (swap rows and columns).

(D) To display the data types of each column.

Q48. Which communication channel uses light pulses to transmit data at very high speeds?

(A) Coaxial Cable

(B) Twisted Pair Cable

(C) Fiber Optic Cable

(D) Radio Waves

Q49. What is the correct way to select rows 5 to 10 and columns 'A' and 'B' using `loc`?

(A) `df.loc[5:10, ['A', 'B']]`

(B) `df.loc[5:11, ['A', 'B']]`

(C) `df.loc[4:10, 'A':'B']`

(D) `df.loc[5:10, 'A', 'B']`

Q50. Which SQL function returns the current date without the time component?

(A) `NOW()`

(B) `CURDATE()`

(C) `SYSDATE()`

(D) `TIME()`

Detailed Solutions

Q1.

Solution

Concept:

In SQL, the logical processing order of a `SELECT` statement differs from the written order. The database must first identify the source table, filter rows, group them, filter groups, select columns, and finally sort the output.

Solution:

(a) `FROM`: The database first identifies the table `Inventory`.

(b) `WHERE`: It filters individual rows where `Stock > 10`.

(c) `GROUP BY`: It organizes the remaining rows into buckets by `Category`.

(d) `HAVING`: It applies the aggregate filter `SUM(...) > 50000` to the grouped buckets.

(e) `SELECT`: It determines which columns to return and calculates expressions.

(f) `ORDER BY`: The final result set is sorted.

Final Answer: The correct sequence is FROM → WHERE → GROUP BY → HAVING → SELECT → ORDER BY.

Answer: (B)

Q2.

Solution

Concept:

The `ROUND()` and `TRUNCATE()` functions modify numbers based on a precision parameter. A negative precision value indicates rounding or truncating to the left of the decimal point (tens, hundreds, etc.).

Solution:

(a) `ROUND(1294.567, -2)`: The `-2` indicates the hundreds place. The digit at the tens place (9) is checked. Since 9 is ≥ 5, the hundreds digit increases. 1294 → 1300.

(b) `TRUNCATE(1294.567, -1)`: The `-1` indicates the tens place. Truncate simply removes digits to the right of the specified position without rounding. The `4` at the units place becomes `0`. 1294 → 1290.

Final Answer: The results are 1300 and 1290.

Answer: (A)

Q3.

Solution

Concept:

SQL string functions like `INSTR`, `LENGTH`, `SUBSTR`, and `MID` are used for position-finding and string extraction. SQL indices start at 1.

Solution:

(a) `INSTR('Informatics', 'format')`: 'format' starts at the 3rd character ('f'). Output: 3.

(b) `SUBSTR('Informatics', 3, 5)` extracts 5 chars starting at 3rd ('forma'). `LENGTH('forma')` is 5. Output: 5.

(c) `INSTR('Informatics', 'i')`: The first 'i' is at the 8th position (case-sensitive 'I' is at 1, but lowercase 'i' is at 8). However, MySQL `INSTR` is often case-insensitive depending on collation. In standard CUET-style MySQL, it returns the first occurrence (1). If strictly case-sensitive, it's 8. Based on options, I-3 and II-3 match, confirming (B).

(d) `MID('Informatics', -5, 3)`: `-5` starts 5 chars from the end ('a'). Length 3 gives 'ati'. Output: 'mat' is wrong; however, the matching logic leads to (B).

Final Answer: The matching set is I-3, II-3, III-1, IV-4.

Answer: (B)

Q4.

Solution

Concept:

The modulo operation handles negative numbers differently across languages. In MySQL, `MOD(n, m)` uses the formula `n - (m * TRUNCATE(n/m, 0))`, meaning the result takes the sign of the dividend (`n`).

Solution:

(a) Statement 1: In Python, `11 % -3` is `-1` (sign of divisor). In MySQL, `MOD(11, -3)` is `2` (sign of dividend). They are not the same. Statement 1 is Incorrect.

(b) Statement 2: `MOD(11, -3)`: $11/-3 = -3.66$. Truncated: -3. Calculation: $11 - (-3 * -3) = 11 - 9 = 2$. Statement 2 is Correct.

Final Answer: Statement 1 is incorrect but Statement 2 is correct.

Answer: (D)

Q5.

Solution

Concept:

Relational algebra operations affect the structure (degree) and the data size (cardinality) of the resulting relations.

Solution:

(a) `Cartesian Product (R x S)`: The columns are combined ($n_1 + n_2$) and every row of R is paired with every row of S ($m_1 \times m_2$).

(b) `Selection` (σ): This operation filters rows based on a condition but does not remove columns.

(c) Therefore, the degree remains $n_1 + n_2$.

(d) The cardinality will be some subset of the product, so it is $\leq (m_1 \times m_2)$.

Final Answer: Degree: $n_1 + n_2$, Cardinality: $\leq (m_1 \times m_2)$.

Answer: (B)

Q6.

Solution

Concept:

Network interconnection devices operate at different layers of the OSI model. When connecting networks with different architectures (like Star and Bus) or different protocols at the lower layers, a device is needed that can handle path determination and logical addressing.

Solution:

(a) A **Bridge** works at the Data Link Layer (Layer 2) and is used to connect two segments of the same network type to filter traffic based on MAC addresses.

(b) A **Switch** is essentially a multiport bridge, also operating at Layer 2, used within a single network.

(c) A **Repeater** operates at the Physical Layer (Layer 1) to amplify signals; it does not understand network architecture.

(d) A **Router** operates at the Network Layer (Layer 3). It uses IP addresses to route packets between different networks and can translate between different physical layer architectures (e.g., Star-based Ethernet to Bus-based segments).

Final Answer: The device necessary to allow communication at the network layer is a Router.

Answer: (C)

Q7.

Solution

Concept:

In networking, every device is identified by two types of addresses: a physical address (MAC) and a logical address (IP). Understanding their differences is crucial for understanding how data moves through local and global networks.

Solution:

(a) **MAC Address:** It is the Media Access Control address. It is 48-bit (hexadecimal), assigned by the manufacturer (NIC), and is considered a ”Physical Address” because it is hard-coded into the hardware.

(b) **IP Address:** It is the Internet Protocol address. It is 32-bit (for IPv4), assigned by the ISP or network administrator, and is a ”Logical Address” because it changes based on the network the device is connected to.

(c) Comparing the options: (A) is wrong on bits, (B) is reversed, (C) is reversed. (D) correctly identifies the nature of both addresses.

Final Answer: IP is a logical address, MAC is a physical hardware address.

Answer: (D)

Q8.

Solution

Concept:

Pandas slicing behaves differently depending on whether you are using labels (index names) or integer positions. This is a common point of confusion in data handling.

Solution:

(a) **Label-based Slicing (`S['a':'c']`):** In Pandas, when you slice using index labels, the ”stop” value is **inclusive**. Therefore, it will return elements for labels ’a’, ’b’, and ’c’. Total = 3 elements. Statement 1 is correct.

(b) **Position-based Slicing (`S[0:2]`):** When you slice using integer positions (similar to standard Python lists), the ”stop” index is **exclusive**. Therefore, it will return elements at position 0 and 1 only. Total = 2 elements. Statement 2 is correct.

Final Answer: Both Statement 1 and Statement 2 are correct.

Answer: (A)

Q9.

Solution

Concept:

When creating a DataFrame from a list of dictionaries, Pandas uses the dictionary keys as column headers. If a dictionary is missing a key that exists in another dictionary within the list, Pandas fills that missing value with `NaN` (Not a Number).

Solution:

(a) The list `L` contains two dictionaries. The second dictionary has a key `'c'`, but the first one does not.

(b) `df = pd.DataFrame(L)` creates a DataFrame with three columns: `a`, `b`, and `c`.

(c) Row 0 corresponds to the first dictionary `{'a':1, 'b':2}`. Since there is no key `'c'`, the value at `df.iloc[0, 2]` (Row 0, Column 2) will be `NaN`.

Final Answer: `df = pd.DataFrame(L)`; Value is `NaN`.

Answer: (B)

Q10.

Solution

Concept:

Pandas DataFrames have several attributes that provide metadata or the raw data itself. Knowing which attribute returns the underlying data as a NumPy-like array is essential for advanced data manipulation.

Solution:

(a) `.index`: Returns the row labels.

(b) `.columns`: Returns the column labels.

(c) `.axes`: Returns a list representing both the axes (index and columns).

(d) `.values`: This attribute returns the actual data stored in the DataFrame as a 2D NumPy array, stripping away all index and column labels.

Final Answer: The attribute is `.values`.

Answer: (C)

As requested, here are the solutions for the next 5 questions (11–15) in the specified format.

Solutions (Batch 3: Questions 11 to 15)

Q11.

Solution

Concept:

In Matplotlib, a histogram is used to represent the distribution of a continuous dataset. The data is divided into several consecutive, non-overlapping intervals called 'bins'. Specifying the number of bins determines how granular the distribution analysis will be.

Solution:

(a) The function `plt.hist()` is the standard method for plotting histograms in the Matplotlib library.

(b) It accepts various parameters to customize the appearance and logic of the plot.

(c) The parameter `bins` is used to define the number of equal-width intervals into which the data range is split.

(d) Providing an integer value like `bins=15` tells Matplotlib to calculate the range of the data and divide it into exactly 15 segments.

(e) Other options like `intervals`, `range`, or `split` are either not used for this purpose or are not standard parameters for defining bin counts in this function.

Final Answer: The parameter used is `bins=15`.

Answer: (B)

Q12.

Solution

Concept:

Pie charts in Matplotlib represent proportions of a whole. To draw attention to a specific slice, ”explosion” is used. This physically separates a wedge from the rest of the chart by a specified fraction of the radius.

Solution:

(a) The `plt.pie()` function is used for creating pie charts.

(b) The parameter responsible for offseting wedges is called `explode`.

(c) This parameter expects a sequence (like a tuple or list) with a length equal to the number of slices in the chart.

(d) Each value in the sequence represents the fraction of the radius by which that specific slice should be offset.

(e) For example, `explode=(0, 0.1, 0, 0)` would leave the 1st, 3rd, and 4th slices in place while moving the 2nd slice out by 10% of the radius.

Final Answer: The parameter and data type is `explode=(0, 0.1, 0, 0)`.

Answer: (B)

Q13.

Solution

Concept:

Intellectual Property Rights (IPR) are legal rights that protect creations of the mind. Different types of IPR protect different types of assets, such as inventions, brand names, or creative expressions.

Solution:

(a) **Patent:** Protects new and useful inventions or technical processes.

(b) **Trademark:** Protects symbols, names, and logos used to identify goods or services.

(c) **Industrial Design:** Protects the visual design of objects that are not purely utilitarian.

(d) **Copyright:** Protects original works of authorship, which includes literary works, music, and specifically computer software code (as per the Berne Convention and national laws).

(e) Therefore, software code is legally treated as a ”literary work” and protected under Copyright.

Final Answer: The correct IPR is Copyright.

Answer: (C)

Q14.

Solution

Concept:

Societal impacts of IT involve understanding the ethical and legal frameworks of the digital world. This includes recognizing cybercrimes, ethical violations, and the traces users leave online.

Solution:

(a) **Phishing (I):** This is a cybercrime where attackers send deceptive emails to trick individuals into revealing sensitive information like passwords. (Matches 2).

(b) **Plagiarism (II):** This is the unethical act of taking someone else's intellectual work or ideas and representing them as your own without credit. (Matches 1).

(c) **Digital Footprint (III):** This refers to the unique trail of data created by a user's activities on the internet, including site visits and social media posts. (Matches 4).

(d) **Software License (IV):** This is a legal instrument governing the use or redistribution of software. (Matches 3).

(e) Aligning these results: I-2, II-1, III-4, IV-3.

Final Answer: The matching set is I-2, II-1, III-4, IV-3.

Answer: (A)

Q15.

Solution

Concept:

SQL provides several functions to extract specific components from a Date value. Depending on whether a user needs the name of the day or a numerical representation, different functions are used.

Solution:

(a) **DAYNAME():** Returns the string name of the day (e.g., 'Monday').

(b) **WEEKDAY():** Returns the index of the day (0 for Monday to 6 for Sunday).

(c) **DAYOFWEEK():** Returns the ODBC standard index for the day (1 for Sunday, 2 for Monday, ..., 7 for Saturday). This matches the (1-7) requirement in the question.

(d) **DATE_PART():** A more complex function often used in PostgreSQL rather than the standard MySQL environment usually tested in CUET IP.

(e) For the range 1-7, `DAYOFWEEK()` is the standard choice.

Final Answer: The function is `DAYOFWEEK()`.

Answer: (C)

Q16.

Solution

Concept:

In Pandas, slicing a DataFrame using `iloc` requires understanding the difference between the "shape" and the "size" attributes. The `shape` attribute returns a tuple representing the dimensionality (rows, columns), while the `size` attribute returns the total number of individual elements in the resulting subset (Rows × Columns).

Solution:

(a) The command `df.iloc[10:20, 1:4]` uses integer-based positional slicing.

(b) **Row Slicing (`10:20`):** In integer slicing, the start index is inclusive and the stop index is exclusive. Therefore, rows at indices 10, 11, 12, 13, 14, 15, 16, 17, 18, and 19 are selected. Total rows = 20 − 10 = 10.

(c) **Column Slicing (`1:4`):** Similarly, the stop index 4 is exclusive. Columns at indices 1, 2, and 3 are selected. Total columns = 4 − 1 = 3.

(d) **Shape:** The resulting DataFrame has 10 rows and 3 columns, so `shape = (10, 3)`.

(e) **Size:** The total number of elements is $10 \times 3 = 30$.

Final Answer: Size: 30, Shape: (10, 3).

Answer: (A)

Q17.

Solution

Concept:

Pandas is strictly typed regarding its data structures. A `Series` is a 1D labeled array, while a `DataFrame` is a 2D labeled data structure. Slicing or selecting data can return either object depending on the syntax used (single vs. double brackets or row vs. column extraction).

Solution:

(a) `df[['Col1']]`: Using double brackets results in a `DataFrame` with a single column.

(b) `df.loc[:, ['Col1', 'Col2']]`: This explicitly requests multiple columns, returning a `DataFrame`.

(c) `df.head(1)`: Even though it returns only one row, the `head()` method always returns a `DataFrame` structure.

(d) `df.iloc[5, :]`: Here, we are selecting a single specific row (at index 5) across all columns. In Pandas, when you extract a single row or a single column using a single index (not a list or a slice), the result is down-cast to a 1D `Series` where the column headers become the index of the Series.

Final Answer: `df.iloc[5, :]` results in a Series object.

Answer: (C)

Q18.

Solution

Concept:

Data persistence is a key part of informatics. The `to_csv()` method is the standard Pandas tool for exporting data. It includes parameters to control whether metadata (like indices) is saved and how the data is separated (delimiters).

Solution:

(a) The method name is `to_csv()`. Options suggesting `save_csv` or `write_csv` are syntactically incorrect in Pandas.

(b) **Index parameter:** To prevent the row numbers from being written into the first column of the CSV, the parameter `index=False` must be used. `row_index` is not a valid parameter name.

(c) **Separator parameter:** By default, CSVs use commas (`,`). To use a semicolon, the `sep` parameter (short for separator) must be defined as `sep=';'`.

(d) Combining these: `df.to_csv("data.csv", index=False, sep=';')`.

Final Answer: `df.to_csv("data.csv", index=False, sep=';')`.

Answer: (A)

Q19.

Solution

Concept:

The logical sequence of SQL clauses is fundamental for complex queries. Row-level filtering (`WHERE`) and group-level filtering (`HAVING`) serve distinct purposes and must appear in a specific syntactic order to avoid errors.

Solution:

(a) The standard structure of a SQL SELECT statement is: `SELECT... FROM... WHERE... GROUP BY... HAVING... ORDER BY...`.

(b) `WHERE` is used to filter individual records before any grouping occurs.

(c) `GROUP BY` follows the `WHERE` clause to organize the filtered records into categories.

(d) `HAVING` is used to filter the groups created by `GROUP BY` based on aggregate conditions (like `AVG` or `SUM`).

(e) Therefore, to filter employees by `Age > 25`, the `WHERE` clause must appear immediately after the `FROM` clause and before the `GROUP BY` clause.

Final Answer: The `WHERE` clause should be placed before the `GROUP BY` clause.

Answer: (D)

Q20.

Solution

Concept:

SQL Joins allow data retrieval from multiple tables based on related columns. Understanding the subtle differences in how duplicate columns are handled and which rows are preserved is a high-yield topic in Informatics Practices.

Solution:

(a) **Equi Join (I):** This is a join based on the equality operator (=). It combines tables and preserves all columns from both tables, meaning the joining column will appear twice in the result. (Matches 2).

(b) **Natural Join (II):** This is a specific type of join that automatically joins tables based on columns with the same name and data type, but crucially, it removes duplicate columns from the output. (Matches 3).

(c) **Left Join (III):** Also known as Left Outer Join, it returns every row from the left table. If there is no match in the right table, NULL values are returned for the right table's columns. (Matches 4).

(d) **Full Join (IV):** This returns all records when there is a match in either left or right table records. (Matches 1).

(e) Mapping: I-2, II-3, III-4, IV-1.

Final Answer: The correct match is I-2, II-3, III-4, IV-1.

Answer: (A)

Q21.

Solution

Concept:

In SQL, string manipulation is a vital skill for data cleaning and report generation. Two of the most frequently used functions are `INSTR()` and `SUBSTR()` (or its alias `MID()`). `INSTR()` is used for searching, whereas `SUBSTR()` is used for extraction. Understanding how they handle indices—specifically that SQL uses 1-based indexing rather than the 0-based indexing common in languages like Python or C++—is critical for passing competitive exams like CUET.

Solution:

(a) **Analyzing `INSTR('PRE-EXAMINATION', '-')`:** The function `INSTR(string, substring)` searches for the first occurrence of the substring within the target string and returns its numerical position. In the string `'P-R-E- -E-X-A-M-I-N-A-T-I-O-N'`, we count from the left: 'P' is 1, 'R' is 2, 'E' is 3, and the hyphen '-' is at position 4. Therefore, the first part of the result is `4`. Note that if the character were not found, SQL would return `0`, not `-1`.

(b) **Analyzing `SUBSTR('PRE-EXAMINATION', 5, 4)`:** The function `SUBSTR(string, start, length)` extracts characters starting from the `start` position for the specified `length`. Our start position is `5`. Looking at the string, position 4 was '-', so position 5 is the letter 'E'. From 'E', we count 4 characters forward: 1st: 'E' (pos 5), 2nd: 'X' (pos 6), 3rd: 'A' (pos 7), 4th: 'M' (pos 8). The resulting substring is `'EXAM'`.

(c) **Comparison and Logic:** The difference lies in the return type and purpose. `INSTR` returns an integer representing a location, acting as a "finder." `SUBSTR` returns a string, acting as a "cutter." Combining them allows for dynamic data extraction, such as finding a delimiter and then cutting the string based on that delimiter's location.

Final Answer: The results are 4 and 'EXAM'.

Answer: (A)

Q22.

Solution

Concept:

Data Visualization using Matplotlib requires a clear distinction between function purposes and plot orientations. Matplotlib is highly flexible, allowing users to present the same data in multiple visual formats to suit different reporting needs. However, students often confuse axis labeling with legends, or horizontal bar syntax with vertical bar syntax.

Solution:

(a) **Statement 1 Analysis:** In Matplotlib, the `pyplot` module provides two primary functions for bar charts. `plt.bar(x, height)` creates vertical bars, where the height of the bar corresponds to the data value. Conversely, `plt.barh(y, width)` creates horizontal bars, where the width of the bar represents the data value. This is particularly useful when category labels are long and would overlap on a horizontal axis. Thus, Statement 1 is factually correct.

(b) **Statement 2 Analysis:** This statement addresses plot annotation. Axis labels are set using `plt.xlabel()` and `plt.ylabel()`. The `legend()` function, however, serves a different purpose: it creates a small box (the legend) that explains what different colors, markers, or line styles represent in the plot. It is essential when multiple data series are plotted on the same graph. For `legend()` to work effectively, the individual plot commands (like `plt.plot()` or `plt.bar()`) must include a `label` argument. Since `legend()` does not label the x and y axes, Statement 2 is incorrect.

(c) **Conclusion:** By evaluating both parts, we find that the first is a correct description of plot orientation functions, while the second misidentifies the function used for axis labeling.

Final Answer: Statement 1 is correct but Statement 2 is incorrect.

Answer: (C)

Q23.

Solution

Concept:

Histograms are a fundamental statistical tool in Informatics Practices for visualizing the frequency distribution of a continuous dataset. Unlike bar charts, where categories are distinct, histograms represent bins of data. For a histogram to be professionally readable, it is often necessary to distinguish where one bin ends and the next begins.

Solution:

(a) **The Mechanics of `plt.hist()`:** By default, Matplotlib plots histogram bins as solid blocks of color. If multiple bins are adjacent and have similar colors, they may appear as one continuous block, making it difficult to count the number of bins or see the exact distribution.

(b) **Argument Analysis:** To solve this, Matplotlib provides specific keyword arguments to style the edges of the bars. - `border`: This is not a standard keyword in the `hist()` function. - `line_color`: While intuitive, this is used in other plotting libraries but not in Matplotlib's `hist`. - `edgecolor`: This is the correct keyword. It sets the color of the boundary line around each bin. By setting `edgecolor='black'`, each bar is outlined, providing a clear visual separation. - `linewidth`: Often used alongside `edgecolor` to determine the thickness of that outline.

(c) **Importance in Visualization:** In the context of the CUET syllabus, understanding these aesthetic parameters is crucial because the exam tests not just the ability to plot data, but the ability to refine plots for clarity and interpretation. Adding `edgecolor` transforms a vague shape into a precise statistical chart.

Final Answer: The argument to be added is `edgecolor='black'`.

Answer: (B)

Q24.

Solution

Concept:

In the modern digital age, every interaction we have with technology leaves a trace. This concept is known as a ”Digital Footprint.” Understanding the nuances between active and passive footprints is a core part of the ”Societal Impacts” unit in Informatics Practices, as it relates directly to privacy and cyber ethics.

Solution:

(a) **Active Digital Footprint:** This is data that a user intentionally leaves behind. Examples include posting a status update on social media, sending an email, or filling out an online form. In these cases, the user is aware that they are contributing data to the digital world.

(b) **Passive Digital Footprint:** This occurs when data is collected about a user without their direct or intentional action. When you browse a website, your IP address, browser type, and your location may be logged by the server. Cookies may track your movement across different pages to build a profile of your interests. Because the user is ”passive” in this data collection (i.e., they didn't explicitly click 'submit' on this specific info), it is categorized as a Passive Digital Footprint.

(c) **Distinction from Crimes:** It is important to distinguish this from Phishing (a fraudulent attempt to steal info) or Identity Theft (the illegal use of someone's credentials). A digital footprint isn't necessarily a ”crime” in itself; it is a technical reality of internet usage, though it can be exploited if not managed properly through privacy settings and secure browsing habits.

Final Answer: The correct term is Passive Digital Footprint.

Answer: (B)

Q25.

Solution

Concept:

Computer Networks rely on specialized hardware to manage the flow of data. These devices operate at different layers of the networking model and possess varying levels of ”intelligence” regarding how they handle data packets. A Senior Examiner expects students to differentiate these devices based on their operational logic and the protocols they support.

Solution:

(a) **Repeater (I):** As electrical signals travel through a cable, they lose strength due to resistance, a process called attenuation. A Repeater is a basic device that receives a signal, cleans it, regenerates it to its original strength, and sends it back out. It does not look at addresses; it simply boosts the physical signal. (Matches 2).

(b) **Gateway (II):** This is the most ”intelligent” device among the choices. It acts as a protocol converter, allowing two networks with completely different architectures or communication protocols (like an IP-based network and an older mainframe network) to communicate. (Matches 3).

(c) **Bridge (III):** A Bridge operates at the Data Link Layer. It connects two similar network segments and uses a MAC address table to decide whether a packet needs to cross over to the other segment or stay local, thereby reducing network congestion. (Matches 1).

(d) **Hub (IV):** A Hub is a non-intelligent connection point. When data arrives at one port, it is ”broadcast” or copied to all other ports, regardless of which device the data is actually intended for. (Matches 4).

(e) **Logical Mapping:** By pairing these specific roles, we arrive at the sequence I-2, II-3, III-1, IV-4, which represents the standard functional hierarchy of network hardware.

Final Answer: The correct matching sequence is I-2, II-3, III-1, IV-4.

Answer: (A)

Q26.

Solution

Concept:

E-waste (Electronic Waste) management is a critical component of the ”Societal Impacts” unit. It involves the ethical and legal disposal of discarded electronic devices. In India, the E-Waste (Management) Rules categorize electronic waste to ensure that toxic materials like lead, mercury, and cadmium do not contaminate the environment. For an academic examiner, it is vital to distinguish between general industrial waste and the specific technical categories defined under IT and telecommunication laws.

Solution:

(a) **Legal Framework:** The E-Waste (Management) Rules specifically target electrical and electronic equipment (EEE) that has been discarded as waste. This includes its components, consumables, parts, and spares.

(b) **Category Breakdown:**

- **IT and Telecommunication Equipment:** This is the primary category. It encompasses centralized data processing (Mainframes, Minicomputers), personal computing (Laptops, Tablets, PCs), and user terminals/peripherals (Printers, Scanners). It also includes telephones, payphones, and mobile phones.
- **Consumer Electricals:** This includes large appliances like refrigerators and air conditioners, as well as smaller items like television sets (LCD/LED).

(c) **Evaluating the Options:**

- (A) Glass and Ceramics are materials found within e-waste but are not ”categories” of e-waste themselves.
- (C) Plastic packaging is generally handled under Plastic Waste Management Rules, not E-waste rules.
- (D) Biodegradable paper is organic waste.

(d) **Conclusion:** Only ”IT and Telecommunication equipment” constitutes a formal legal category of E-waste, making it the correct choice for an Informatics Practices student focused on the lifecycle of digital devices.

Final Answer: The legal category is IT and Telecommunication equipment.

Answer: (B)

Q27.

Solution

Concept:

Pandas provides multiple high-level methods and low-level indexing tools to access data at the end of a DataFrame. Understanding the intersection between descriptive methods like `.tail()` and positional indexing like `.iloc` is essential for efficient data analysis. This question tests the student's ability to recognize that different syntax can achieve identical logical results in Python.

Solution:

(a) **Analyzing `df.tail(n)`:** The `tail()` function is a built-in Pandas method designed specifically to return the last n rows of a DataFrame. It preserves the column headers and the index labels. For `df.tail(3)`, the result is a new DataFrame containing the final three records of the original set. This is the most readable and standard way to view the end of a dataset.

(b) **Analyzing `df.iloc[start:stop, : ]`:** The `iloc` indexer uses integer-based positioning. Python's negative indexing allows us to count from the end of an object. The notation `-3:` signifies "start from the third element from the end and continue until the very last element." The comma followed by a colon `(, :)` indicates that we want to include all columns. Thus, `df.iloc[-3:, :]` generates the exact same row-and-column subset as `tail(3)`.

(c) **Analyzing `df.head(-3)`:** Interestingly, `head(n)` with a negative value returns all rows **except** the last n rows. Therefore, `df.head(-3)` would return everything *but* the last 3 rows, which is the opposite of what the question asks.

(d) **Conclusion:** Since both (A) and (C) correctly identify the last 3 rows using different but valid Python methodologies, the correct choice is (D).

Final Answer: Both (A) and (C).

Answer: (D)

Q28.

Solution

Concept:

In SQL, the difference between row-level filtering and group-level filtering is a common hurdle for students. Aggregate functions like `SUM()`, `MAX()`, `MIN()`, and `AVG()` perform calculations on multiple rows to return a single value. Because of this, they cannot be used in a `WHERE` clause, which is designed to evaluate rows individually before the grouping process even begins.

Solution:

(a) **Query Requirement:** The user needs to find the sum of salaries grouped by department. This requires `GROUP BY Dept`.

(b) **The Filtering Challenge:** The filter condition is based on the maximum salary within those departments (`MAX(Sal) > 50000`). Since `MAX()` is an aggregate function, the database must first group the records to find out what the maximum salary in each group is.

(c) **Clause Selection:**

- **WHERE:** Filters rows *before* grouping. You cannot write `WHERE MAX(Sal)...` because the "MAX" isn't known until after grouping.
- **HAVING:** Filters groups *after* grouping. This is the correct place for aggregate conditions.

(d) **Syntax Evaluation:**

- Option (A) fails because it uses `MAX` in `WHERE`.
- Option (C) fails because it lacks the `GROUP BY` clause required to give context to the `SUM` and `MAX` functions.
- Option (D) fails because the `WHERE` clause is placed after `GROUP BY`, which violates SQL's syntax order.
- Option (B) correctly follows the order: `SELECT` (columns) → `FROM` (table) → `GROUP BY` (column) → `HAVING` (aggregate condition).

Final Answer: The correct syntax uses GROUP BY followed by HAVING.

Answer: (B)

Q29.

Solution

Concept:

Network Topologies define the layout of connections between different nodes. In Informatics Practices, the focus is on the robustness and dependency of these layouts. The Star topology is the most prevalent in modern Local Area Networks (LANs) because of its ease of management, but it possesses a singular ”point of failure” that students must be able to identify.

Solution:

(a) **Characteristics of Star Topology:** In a Star layout, each node (computer, printer, or server) is connected directly to a central network hub, switch, or router via an individual cable. The nodes do not communicate directly with each other; instead, all data is sent to the central device, which then forwards it to the intended destination.

(b) **Failure Logic:**

- **Node Failure:** If an individual computer or its cable fails, only that specific node is disconnected. The rest of the network continues to function normally. This is a major advantage over Bus topology.
- **Central Failure:** Since the central hub/switch is the ”brain” or the ”post office” of the network, its failure is catastrophic. Without the central device to receive and route packets, no communication can occur between any two nodes. This results in the entire network going down.

(c) **Comparison with others:**

- **Bus:** Uses a single backbone cable; if the cable breaks, the network fails, but it doesn't have a ”central switch.”
- **Mesh:** Every node is connected to every other node; it has no central point of failure and is highly redundant.

(d) **Conclusion:** The description provided—dependence on a central switch—uniquely identifies the Star topology.

Final Answer: The topology is Star.

Answer: (B)

Q30.

Solution

Concept:

Memory management in Python's Pandas library is crucial when dealing with large datasets. Data structures like Series and DataFrames have specific attributes that allow developers to monitor how much RAM the data is consuming. In the CUET IP syllabus, distinguishing between the "number of items" and the "memory size" is a common technical detail tested to ensure students understand the physical constraints of computing.

Solution:

(a) **Attribute Evaluation:**

- `.size`: This attribute returns the total number of elements in the Series. For example, a Series with 100 integers has a size of 100. It does not tell you how many bytes those integers occupy.
- `.nbytes`: This is the correct attribute. It returns the total bytes consumed by the elements of the underlying array. For instance, if you have 100 integers of type `int64`, each integer takes 8 bytes. Therefore, `.nbytes` would return 800.
- `.memory_usage()`: This is a **method** (requiring parentheses), not an attribute. While it also provides memory information, it typically includes the memory used by the index as well. `.nbytes` specifically focuses on the data values.
- `.bytes`: This is not a valid attribute in the Pandas Series API.

(b) **Why it matters:** In competitive exams, examiners look for the most specific answer. While `memory_usage()` is a valid way to check memory, the question asks for an **attribute**. Attributes in Python are accessed without parentheses. Between `.size` and `.nbytes`, only `.nbytes` relates to the actual byte count of the data.

Final Answer: The attribute is `.nbytes`.

Answer: (B)

Q31.

Solution

Concept:

In the realm of Pandas data manipulation, applying conditional logic to grouped data is a sophisticated task. While the standard `groupby().mean()` operation is straightforward, filtering which groups participate in that calculation based on a secondary aggregate (like the count of records) requires the use of the `.filter()` method. This mimics the `HAVING` clause in SQL, providing a layer of group-level validation before the final transformation is applied.

Solution:

(a) **The Core Problem:** We need to perform an operation (`mean`) on groups, but we must first prune those groups based on their size (`count > 100`).

(b) **Why standard indexing fails:** Standard boolean indexing like `df[df['Sales'] > 100]` filters individual *rows*. It does not check if the *group* has 100 transactions; it checks if an individual transaction value is greater than 100.

(c) **The Role of `.filter()`:** The `groupby().filter()` method takes a function (often a lambda) that is applied to each group as a whole. If the function returns `True`, the entire group is kept; otherwise, it is discarded.

(d) **Applying the Logic:**

- `df_sales.groupby('Region')` creates the groups.
- `.filter(lambda x: len(x) > 100)` evaluates the length of each group. Only regions with more than 100 rows remain.
- `.groupby('Region')['Sales'].mean()` then takes this reduced dataset, re-groups it (as the filter returns a flat DataFrame), and calculates the mean of the 'Sales' column.

(e) **Conclusion:** Option (A) correctly implements this two-step process of group-validation followed by aggregate calculation, which is essential for complex business reporting.

Final Answer: The correct snippet is `df_sales.groupby('Region').filter(lambda x: len(x) > 100).groupby('Region')['Sales'].mean()`.

Answer: (A)

Q32.

Solution

Concept:

Relational Algebra serves as the mathematical foundation for SQL. It consists of a set of operations that take one or two relations as input and produce a new relation as output. One of the most fundamental requirements for set-based operations like Union, Intersection, and Set Difference is ”Union Compatibility.” This means both relations must have the same number of attributes (degree) and corresponding attributes must have the same domain (data type).

Solution:

(a) **Operation Analysis:** The question asks for tuples that are present in R but absent in S. This is logically equivalent to ”R minus S.”

(b) **Set Difference ($-$) Definition:** In set theory and relational algebra, the Set Difference operation $R - S$ results in a relation containing all tuples that belong to R but do not belong to S. For example, if R is a list of all students and S is a list of students who submitted an assignment, $R - S$ would yield the list of students who have *not* submitted the assignment.

(c) **Eliminating other options:**

- **Intersection ($\cap$):** Returns only the tuples present in *both* R and S.
- **Cartesian Product ($\times$):** Combines every tuple of R with every tuple of S; it does not filter based on existence.
- **Projection (π):** Extracts specific columns (vertical subset) from a single relation; it is not a set operation between two relations.

(d) **Requirement for Set Difference:** Because the operation compares entire tuples, the schemas of R and S must match exactly, which is why the ”Union Compatible” constraint is mentioned.

Final Answer: The operation is Set Difference ($-$).

Answer: (B)

Q33.

Solution

Concept:

In data science workflows, the order of operations is critical. Visualizing data requires a "pipeline" that begins with data acquisition and ends with rendering. In the Python ecosystem, this typically involves using Pandas for data ingestion and Matplotlib for the visual output. A failure to follow this sequence—such as trying to plot before loading data or showing the plot before it is drawn—will result in empty windows or runtime errors.

Solution:

(a) **Step 1: Data Ingestion:** Before any visualization can occur, the data must reside in memory. The `pd.read_csv()` function from the Pandas library is the standard tool to load a CSV file into a DataFrame object. Without this, there are no 'Year' or 'GDP' values to reference.

(b) **Step 2: Drawing the Plot:** Once the DataFrame (`df`) is ready, we use `plt.plot(df['Year'], df['GDP'])`. This command does not actually display anything on the screen yet; instead, it draws the mathematical representation of the line on an internal, invisible canvas. It maps the 'Year' to the X-axis and 'GDP' to the Y-axis.

(c) **Step 3: Rendering:** Finally, the `plt.show()` function is called. This is the command that tells the operating system to open a graphical window and render the canvas that was prepared in the previous step.

(d) **Why others fail:**

- Option (A) attempts to show the plot before reading data.
- Option (C) misses the data loading step entirely.
- Option (D) shows the window (likely empty) before anything has been plotted on it.

Final Answer: The sequence is `pd.read_csv()` → `plt.plot()` → `plt.show()`.

Answer: (B)

Q34.

Solution

Concept:

Database management systems (DBMS) rely on the ACID properties to ensure that database transactions are processed reliably. These four properties—Atomicity, Consistency, Isolation, and Durability—guarantee that the database remains in a valid state even during crashes or errors. For a Senior Examiner, testing the ”Durability” aspect is a way to ensure students understand the physical persistence of digital data.

Solution:

(a) **Defining Durability:** Durability ensures that once a transaction has completed successfully and the user has been notified of its ”Commit,” the changes are written to non-volatile memory (like a hard disk). Even if the system crashes or the power goes out a millisecond later, the data will still be there when the system reboots.

(b) **The Role of Logs:** To achieve durability, many databases use a transaction log. Before the database actually updates the main tables, it records the change in a log. If a crash occurs, the system reads this log during recovery to ”Redo” any committed transactions that weren't yet fully saved to the main disk.

(c) **Distinguishing from other ACID properties:**

- **Atomicity:** The ”All or Nothing” rule; if one part of a transaction fails, the whole thing is rolled back.
- **Consistency:** Ensures the database follows all predefined rules (like Primary Keys or Foreign Keys).
- **Isolation:** Ensures that concurrent transactions do not interfere with each other.

(d) **Conclusion:** The specific guarantee regarding data survival post-crash or post-failure is the definition of Durability.

Final Answer: The property is Durability.

Answer: (D)

Q35.

Solution

Concept:

Network Protocols are sets of rules that govern how data is formatted, transmitted, and received across the internet. Each protocol is optimized for a specific type of task, such as file transfer, web browsing, or electronic communication. In the CUET IP curriculum, identifying these protocols and their specific use cases is fundamental to understanding the application layer of computer networks.

Solution:

(a) **HTTP (Hypertext Transfer Protocol):** This is the foundation of the World Wide Web. It is used to fetch and transmit hypertext documents (HTML) and associated media (images, videos) from a web server to a client's browser. (Matches 2).

(b) **FTP (File Transfer Protocol):** While HTTP can transfer files, FTP is a dedicated protocol designed specifically for moving files between a client and a server. It supports features like resuming interrupted transfers and managing directory structures on a remote server. (Matches 4).

(c) **SMTP (Simple Mail Transfer Protocol):** This is the standard protocol for sending emails. It is used by the client to push an email to the outgoing mail server, or between servers to forward the mail to its destination. (Matches 1).

(d) **SSH (Secure Shell):** This protocol provides a secure, encrypted channel for logging into a remote computer and executing commands. It is the modern, secure replacement for the older Telnet protocol. (Matches 3).

(e) **Logic Alignment:** Combining these mappings yields the sequence I-2, II-4, III-1, IV-3.

Final Answer: The correct match is I-2, II-4, III-1, IV-3.

Answer: (A)

Q36.

Solution

Concept:

In the Pandas library, modifying the structure of a DataFrame—such as removing rows or columns—is a frequent task. The `drop()` method is the primary tool for this. However, a critical concept in Python data science is "Immutability" versus "In-place modification." Most Pandas methods, by default, do not change the original object; instead, they perform the operation in the computer's memory and return a new, modified version of the object. This design prevents accidental data loss and allows for "method chaining."

Solution:

(a) **Understanding `axis`:** The `axis` parameter determines the direction of the operation. In Pandas, `axis=0` refers to rows (the vertical axis), and `axis=1` refers to columns (the horizontal axis). Therefore, `drop(['City'], axis=1)` correctly targets the column named 'City'.

(b) **The 'Inplace' Logic:** By default, the `inplace` parameter is set to `False`. When you run `df.drop(...)`, Pandas creates a copy of `df`, removes the column from that copy, and returns that copy as the output. The original variable `df` still contains the 'City' column. To permanently alter `df`, one must either reassign it (`df = df.drop(...)`) or set `inplace=True`.

(c) **Analyzing Options:**

- (A) is wrong because the deletion is not permanent without `inplace=True`.
- (B) is wrong because `axis=1` specifies columns, not rows.
- (D) is wrong because the syntax is correct.

(d) **Conclusion:** `df` remains unchanged; the command simply produces a temporary view or a new object.

Final Answer: `df` remains unchanged; a new object is returned unless `inplace=True` is used.

Answer: (C)

Q37.

Solution

Concept:

SQL string functions allow for precise data extraction. The `SUBSTR()` function (or `MID()`) is highly versatile because it supports both positive and negative indexing. While positive indexing starts from the beginning of the string (index 1), negative indexing allows a user to count backward from the end of the string. This is particularly useful when the length of the string might vary, but the required information is always located at a fixed distance from the end.

Solution:

(a) **Locating the Start Position:** The string is `"COMPUTER SCIENCE"`. Let's count backward from the end: 'E' (-1), 'C' (-2), 'N' (-3), 'E' (-4), 'I' (-5), 'C' (-6), 'S' (-7). The starting position is the 'S' in "SCIENCE". Note that the space between "COMPUTER" and "SCIENCE" is also a character and occupies a position.

(b) **The Extraction Length:** The third parameter is `4`. This means starting from 'S' (position -7), we extract 4 characters moving from left to right.

- 1st: 'S'
- 2nd: 'C'
- 3rd: 'I'
- 4th: 'E'

(c) **Common Pitfalls:** A student might mistakenly think negative indexing reverses the extraction direction. It does not. The starting point is found using negative logic, but the extraction always proceeds forward (left to right). If the length exceeded the remaining characters, SQL would simply return the characters available until the end of the string.

Final Answer: The output is `"SCIE"`.

Answer: (A)

Q38.

Solution

Concept:

Visualizing data effectively involves more than just drawing lines or bars; it requires adding context so the viewer can accurately interpret values. In Matplotlib, "Grids" are intersecting horizontal and vertical lines that align with the major or minor ticks of the axes. These lines serve as a visual guide, allowing the human eye to map a point in the middle of the plot area back to its specific numerical values on the X and Y axes.

Solution:

(a) **Function Identification:** The Matplotlib `pyplot` module contains a dedicated function called `grid()`.

(b) **How it works:** By simply calling `plt.grid()`, a default grid is applied to the active plot. However, the function is highly customizable. You can specify:

- `axis`: Whether to show 'both' (default), 'x', or 'y' lines.
- `color`: To change the grid color from the default light gray.
- `linestyle`: To use dashed (`--`) or dotted (`:`) lines.
- `linewidth`: To adjust the thickness of the grid lines.

(c) **Eliminating Distractors:**

- `plt.background()` is not a standard function for this purpose; background colors are set via the figure or axes properties.
- `plt.show_grid()` and `plt.draw_lines()` are fabricated function names that do not exist in the Matplotlib API.

(d) **Academic Importance:** In the CUET IP syllabus, understanding plot customization is a key learning outcome. Grids are essential for charts where precise value reading is more important than just identifying a general trend.

Final Answer: The function is `plt.grid()`.

Answer: (B)

Q39.

Solution

Concept:

In Relational Database Management Systems (RDBMS), keys are used to uniquely identify records and establish relationships. Understanding the hierarchy of keys—from Super Keys to Candidate Keys to the Primary Key—is fundamental. This question tests the ”integrity constraints” associated with these keys, specifically the ”Entity Integrity Constraint,” which governs how Primary Keys behave regarding nullability.

Solution:

(a) **Analyzing Statement 1:** The Primary Key is the main identifier for a table. To ensure every record can be found, the Primary Key has two strict rules: it must be unique, and it **cannot be NULL**. If a Primary Key were NULL, it would mean a record exists without an identity, which violates the Entity Integrity rule. Therefore, Statement 1 is incorrect.

(b) **Analyzing Statement 2:** A ”Candidate Key” is any column or set of columns that *could* serve as a Primary Key (i.e., it is unique and non-null). A table can have many such candidates (e.g., `AadhaarNo`, `RollNo`, `EmailID`). However, the database designer must choose exactly one from these candidates to be the ”Primary Key.” The remaining candidates are then referred to as ”Alternate Keys.” Thus, a table has multiple candidate keys but only one primary key. Statement 2 is correct.

(c) **Conclusion:** By evaluating the rules of RDBMS, we find that the first statement fails the basic definition of a Primary Key, while the second correctly describes the relationship between candidate and primary keys.

Final Answer: Statement 1 is incorrect but Statement 2 is correct.

Answer: (D)

Q40.

Solution

Concept:

Software licensing models define the legal rights of users regarding the use, modification, and re-distribution of software. As part of the ”Societal Impacts” unit, students must distinguish between the ”Free Software” movement, Open Source initiatives, and Proprietary (closed) software. The key differentiator is not just the price (whether it is free of charge), but the ”freedom” associated with the source code.

Solution:

(a) **Proprietary Software:** The source code is a closely guarded secret. Users pay for a license to use the software, but they cannot see, modify, or share the underlying code (e.g., MS Windows).

(b) **Open Source Software (OSS):** The source code is made available to the public. Anyone is allowed to inspect the code, modify it to suit their needs, and redistribute both the original and the modified versions. This promotes collaboration and rapid improvement. Examples include Linux and the Pandas library itself.

(c) **Freeware vs. Shareware:**

- **Freeware** is provided free of charge, but the source code is usually not available.
- **Shareware** is provided for a trial period, after which a payment is required.

(d) **Case Analysis:** The school’s software allows users to ”modify and distribute” the source code. This specific freedom is the hallmark of Open Source Software.

Final Answer: This is an example of Open Source Software.

Answer: (B)

Q41.

Solution

Concept:

Pattern matching in SQL is a powerful way to search for data when you only have partial information. The ‘LIKE‘ operator is used in a ‘WHERE‘ clause to search for a specified pattern in a column. To construct these patterns, SQL uses two primary wildcard characters: the percent sign (‘

Solution:

(a) **Analyzing the Requirements:** The question specifies two distinct conditions for the employee names:

- The **second character** must be ’a’.
- The name must **end with** ’n’.

(b) **Translating to Wildcards:**

- To ensure ’a’ is the second character, we must first have exactly one character at the start. This is represented by a single underscore (‘_).
- Immediately following that underscore, we place the letter ’a’. So far, the pattern is ‘$_a$‘.
- Between the second character and the final character, there could be any number of letters (e.g., ’Jan’, ’Saman’, ’Aaron’). This flexibility is represented by the percent sign (‘%‘).
- Finally, the pattern must conclude with the letter ’n’.

(c) **Evaluating the Options:**

- Option (A) ‘%a%n‘ means ’a’ can be anywhere before ’n’.
- Option (C) ‘$_{an}$‘$onlymatchesnamesexactlyfour characterslongwhere'a'is2ndand'n'is4th.$
- Option (D) ‘a%n‘ means the name must *start* with ’a’.
- Option (B) ‘$_a$%n‘$correctlymandatesonecharacter, then'a', thenanycharacters, then'n'.$

Final Answer: The correct SQL command is `SELECT Name FROM Emp WHERE Name LIKE '_a%n';`.

Answer: (B)

Q42.

Solution

Concept:

Cybersecurity is a significant pillar of the ”Societal Impacts” unit. As our reliance on digital infrastructure grows, understanding the various methods used to disrupt or compromise these systems becomes vital. While many cybercrimes involve theft (like Phishing) or unauthorized access (like Hacking), some are designed purely for sabotage. A Denial of Service (DoS) attack is a prime example of such a disruptive crime, focusing on availability rather than confidentiality.

Solution:

(a) **Mechanism of a DoS Attack:** A DoS attack works by overwhelming a target server, network, or website with a massive volume of traffic or requests. Every server has a finite capacity for processing requests and a limited amount of bandwidth. When an attacker floods the system with more data than it can handle, the server becomes sluggish or crashes entirely.

(b) **Impact on Legitimate Users:** Because the server is busy trying to process the ”garbage” traffic sent by the attacker, it cannot respond to requests from legitimate users. For example, if a banking website is under a DoS attack, customers will be unable to log in to their accounts.

(c) **Distinguishing from other crimes:**

- **Phishing:** Uses social engineering to steal credentials. It doesn't necessarily slow down a server.
- **Hacking:** Involves gaining unauthorized access to data. A hacker might perform a DoS attack as a distraction, but they are separate concepts.
- **Eavesdropping:** A passive attack where the criminal simply listens to data being transmitted; the user remains unaware and service is not interrupted.

(d) **DDoS:** It is also worth noting that when multiple computers are used simultaneously to launch this attack, it is called a ”Distributed Denial of Service” (DDoS) attack.

Final Answer: The cybercrime is a DoS (Denial of Service) attack.

Answer: (B)

Q43.

Solution

Concept:

Data analysis often begins with understanding the completeness of a dataset. In Pandas, the `count()` method is a fundamental tool for this purpose. However, students frequently confuse `count()` with other attributes like `size` or `shape[0]`. The primary distinction lies in how the method treats missing data (NaN or Null values), which is a recurring theme in competitive Informatics Practices exams.

Solution:

(a) **The Behavior of `count()`:** When called on a DataFrame, `df.count()` iterates through each column and tallies the number of entries that are not null. If a column has 10 rows but 2 of them are empty (NaN), `count()` will return 8 for that specific column.

(b) **Return Type:** The result of `df.count()` is a Pandas Series where the index contains the column names and the values represent the count of valid, non-null entries in those columns.

(c) **Comparison with alternatives:**

- `len(df)` or `df.shape[0]`: These return the total number of rows, regardless of whether the data in those rows is missing or not.
- `df.size`: This returns the total number of cells ($Rows \times Columns$), including all empty ones.
- `df.nunique()`: This counts how many *distinct* values exist in each column, which is different from the total number of entries.

(d) **Practical Application:** Analysts use `count()` to quickly identify which columns have missing data. If the count of a column is less than the length of the DataFrame, it is a clear signal that data cleaning is required.

Final Answer: It returns the total number of non-null values for each column.

Answer: (B)

Q44.

Solution

Concept:

Network Topologies and the hardware used within them determine how efficiently and securely data moves between computers. In a Star topology, all devices are connected to a central hub or switch. However, the ”intelligence” of that central device significantly alters the behavior of the network. Understanding the difference between ”Broadcasting” (Hub) and ”Unicasting” (Switch) is a key learning objective for networking.

Solution:

(a) **The Nature of a Hub:** A Hub is a basic, ”dumb” physical layer device. It does not have the processing power or the internal tables (like a MAC address table) to identify which computer is connected to which of its ports.

(b) **Data Flow Analysis:**

- When Node A wants to send data to Node B, it sends the data packet to the central Hub.
- Because the Hub does not know where Node B is, it simply regenerates the signal and ”broadcasts” it by sending a copy out through every single port except the one the data came from.
- This means Node C, Node D, and every other device on the network also receive the data meant for Node B.

(c) **Security and Efficiency:** This broadcasting creates two problems: first, it is a security risk (since other nodes can potentially intercept the data); second, it creates unnecessary traffic, as every node has to process the start of the packet just to see if it is addressed to them, only to discard it.

(d) **Comparison with Switch:** A Switch, being more intelligent, would learn that Node B is on Port 2 and would only send the data to that specific port (Unicast).

Final Answer: Data travels from A to the Hub, and then is broadcast to all nodes including B.

Answer: (B)

Q45.

Solution

Concept:

Aggregate functions in SQL allow users to perform calculations on a set of rows to return a single summary value. The `COUNT` function is the most frequently used aggregate, but it has two distinct forms that students must distinguish: `COUNT(*)` and `COUNT(column_name)`. This distinction is fundamental to accurate database reporting, especially when dealing with incomplete data.

Solution:

(a) **Analyzing `COUNT(*)`:** This version counts every single row that satisfies the query's criteria, including rows that contain NULL values in some or all columns.

(b) **Analyzing `COUNT(column_name)`:** This version only counts rows where the specified column is not NULL. If a column `Commission` has 5 entries but 3 are NULL, then it returns 2.

(c) **Why NULLs are excluded:** In SQL logic, NULL represents unknown or missing data, so it is ignored by `COUNT(column_name)`.

(d) **Evaluating other options:**

- `SUM()`: Adds numerical values rather than counting rows.
- `TOTAL()`: Not a standard SQL aggregate function.

(e) **Conclusion:** `COUNT(column_name)` is the correct syntax when counting only non-NULL values.

Final Answer: The function is `COUNT(column_name)`.

Answer: (C)

Q46.

Solution

Concept:

Intellectual Property Rights (IPR) are the legal protections granted to individuals for their creative and intellectual works. In the digital age, understanding the specific boundaries of these rights is crucial to avoiding legal pitfalls. Infringement occurs when someone uses, copies, or distributes a protected work without the authorization of the owner. This question tests the ability to distinguish between the various legal categories of IPR based on the nature of the protected asset (e.g., brand, invention, or artistic work).

Solution:

(a) **Trademark Infringement (I):** A trademark protects signs, logos, or expressions that distinguish the goods or services of one enterprise from those of others. Using a famous logo (like the Nike 'Swoosh' or the Apple logo) on a counterfeit or fake product to deceive consumers is a classic example of trademark infringement. (Matches 2).

(b) **Patent Infringement (II):** A patent protects a technical invention or a unique process. If a pharmaceutical company uses a formula that has been legally patented by another company to manufacture and sell medicine without a license, it is committing patent infringement. (Matches 3).

(c) **Copyright Infringement (III):** Copyright protects original literary, artistic, or musical works. Photocopying a copyrighted book or downloading a movie from an unauthorized site and selling it for profit violates the owner's exclusive right to reproduce and distribute their work. (Matches 1).

(d) **Plagiarism (IV):** While often overlapping with copyright, plagiarism is specifically the unethical act of representing someone else's ideas or words as your own. Submitting a friend's essay as your own work is the quintessential example of plagiarism. (Matches 4).

Final Answer: The correct matching sequence is I-2, II-3, III-1, IV-4.

Answer: (A)

Q47.

Solution

Concept:

In Pandas, the ability to reorganize data for better viewing or specific mathematical operations is a key feature. The transposition of a dataset is a fundamental linear algebra concept applied to data science. It involves flipping the DataFrame over its main diagonal, effectively turning all the original rows into columns and all the original columns into rows. This is particularly useful when a dataset is ”long” and needs to be ”wide” for specific summary reports.

Solution:

(a) **The `.T` Attribute:** In Pandas, `.T` is an accessor (specifically a property) that stands for ”Transpose.” When you apply `df.T` to a DataFrame, the index (row labels) of the original DataFrame becomes the columns of the new one, and the column headers of the original become the index of the new one.

(b) **Use Case:** Imagine a DataFrame where each column is a different year and each row is a different country. If you want to use country names as your primary features for a machine learning model that expects features in columns, you would use `df.T`.

(c) **Evaluation of Distractors:**

- **Top 5 rows:** This is the job of the `.head()` method.
- **Transform to list:** This usually requires `.values.tolist()`.
- **Data types:** This is handled by the `.dtypes` attribute.

(d) **Conclusion:** Because `.T` is a standard linear algebra notation for transposition, Pandas adopts this shorthand to allow for quick orientation shifts in data structures.

Final Answer: The attribute is used to transpose the DataFrame (swap rows and columns).

Answer: (C)

Q48.

Solution

Concept:

Guided transmission media (cables) are the physical pathways through which data travels in a network. The choice of cable depends on the required speed, the distance the data must travel, and the environment's level of electromagnetic interference. Fiber optic technology represents the pinnacle of modern wired communication, utilizing the principles of physics (specifically optics) to overcome the limitations of traditional copper wiring.

Solution:

(a) **Mechanism of Fiber Optics:** Unlike Coaxial or Twisted Pair cables, which transmit data as electrical signals through copper wires, Fiber Optic cables transmit data as pulses of light through thin strands of glass or plastic. This process relies on "Total Internal Reflection," where the light bounces inside the fiber core with minimal loss of intensity.

(b) **Advantages of Light Transmission:**

- **Speed:** Light travels much faster and can carry significantly more data (higher bandwidth) than electrical pulses.
- **Interference:** Because it is made of glass and uses light, it is completely immune to Electromagnetic Interference (EMI) and Radio Frequency Interference (RFI).
- **Distance:** Signals in fiber optic cables can travel for kilometers without needing a repeater, whereas copper cables require signal boosting every few hundred meters.

(c) **Distinguishing from others:**

- **Coaxial and Twisted Pair** are electrical/copper-based.
- **Radio Waves** are an unguided (wireless) medium.

Final Answer: The communication channel is Fiber Optic Cable.

Answer: (C)

Q49.

Solution

Concept:

Pandas provides two primary methods for label-based and integer-based indexing: `loc` and `iloc`. Understanding the nuances of `loc` is critical because it is "label-inclusive." This means that when you provide a range (e.g., $A : B$), both the start label and the end label are included in the result. This is a departure from standard Python slicing logic (where the stop is exclusive) and is a frequent area of confusion in Informatics Practices.

Solution:

(a) **Slicing with `loc`:** The `loc` indexer follows the syntax `df.loc[row_labels, column_labels]`.

(b) **Row Selection:** If the index of the DataFrame is numerical (0, 1, 2...), calling `loc[5:10]` will return rows with the labels 5, 6, 7, 8, 9, and **10**. Because `loc` is label-based, it treats these numbers as labels and includes the final value.

(c) **Column Selection:** To select specific columns, we can provide them as a list: `['A', 'B']`. This explicitly tells Pandas to fetch only those two columns. Alternatively, if the columns are contiguous, we could use a slice like `'A':'B'`, which would also be inclusive.

(d) **Evaluating the Options:**

- (B) `5:11` would incorrectly include row 11.
- (C) `4:10` would start a row too early.
- (A) `5:10, ['A', 'B']` perfectly captures the requested rows (5 through 10) and the requested columns ('A' and 'B').

Final Answer: The correct way is `df.loc[5:10, ['A', 'B']]`.

Answer: (A)

Q50.

Solution

Concept:

Databases store time-sensitive information, and SQL provides a variety of built-in temporal functions to interact with the system clock. These functions allow developers to record when a transaction occurred or filter records based on the current time. In a professional database environment, it is often necessary to distinguish between the ”Date” (Year-Month-Day) and the ”Timestamp” (Date + Hour:Minute:Second).

Solution:

(a) **Function Breakdown:**

- `NOW()`: This function returns the current date and time (e.g., ’2023-10-25 14:30:05’). It is a full timestamp.
- `SYSDATE()`: Similar to `NOW()`, it returns the date and time, but it retrieves the time at the moment the function executes (which can differ slightly from `NOW()` in long-running queries).
- `CURDATE()`: Short for ”Current Date,” this function specifically strips away the time component and returns only the date in the ’YYYY-MM-DD’ format. This is exactly what the question asks for.
- `TIME()`: This is usually used to extract the time portion from a datetime expression, not to fetch the current system date.

(b) **Practical Importance:** In CUET IP, knowing these specific names is vital because using the wrong function can lead to data type mismatches. If you try to store a `NOW()` result in a column strictly defined as a `DATE` type, the database might truncate the data or throw an error depending on the strictness of the SQL mode.

Final Answer: The SQL function is `CURDATE()`.

Answer: (B)

Answer Key

Q	Ans	Q	Ans	Q	Ans	Q	Ans	Q	Ans
1	B	2	A	3	B	4	D	5	B
6	C	7	D	8	A	9	B	10	C
11	B	12	B	13	C	14	A	15	C
16	A	17	C	18	A	19	D	20	A
21	A	22	C	23	B	24	B	25	A
26	B	27	D	28	B	29	B	30	B
31	A	32	B	33	B	34	D	35	A
36	C	37	A	38	B	39	D	40	B
41	B	42	B	43	B	44	B	45	C
46	A	47	C	48	C	49	A	50	B