

# Python Institute PCAD-31-02

Python Institute Certified Associate Data Analyst with  
Python (PCAD-31-02)

For More Information – Visit link below:

<https://www.examsempire.com/>

**Product Version**

1. Up to Date products, reliable and verified.
2. Questions and Answers in PDF Format.



<https://examsempire.com/>

Visit us at: <https://www.examsempire.com/pcad-31-02>

# Latest Version: 6.0

## Question: 1

What is a major challenge in aggregating data from multiple sources?  
Response:

- A. Excessive disk space usage
- B. Duplicate records and format inconsistencies
- C. Increased algorithmic complexity
- D. Overfitting in predictive models

**Answer: B**

## Question: 2

What may occur if a model is evaluated using the same data it was trained on?

- A. Overgeneralization
- B. Data leakage
- C. Overfitting
- D. Underfitting

**Answer: C**

## Question: 3

What is the primary goal of pre-processing data before performing analysis?

- A. To minimize the number of data points
- B. To obscure sensitive information
- C. To ensure the data is clean, consistent, and suitable for modeling
- D. To compress datasets for storage efficiency

**Answer: C**

## Question: 4

Which method allows you to detect and remove rows with duplicate values across all columns in a Pandas DataFrame?

- A. DataFrame.unique()
- B. DataFrame.remove\_duplicates()
- C. DataFrame.drop\_duplicates()
- D. DataFrame.nunique()

**Answer: B**

### Question: 5

Which operations are recommended when organizing messy tabular data in Pandas for further transformation and statistical modeling?  
(choose two)

- A. Resetting the index after filtering
- B. Dropping column headers
- C. Ensuring consistent data types
- D. Converting the DataFrame to a Series

**Answer: A,C**

### Question: 6

Which conditions typically necessitate data normalization or scaling before analysis?  
(choose two)

- A. Features with different units and scales
- B. Presence of duplicate data
- C. Distance-based modeling algorithms
- D. Boolean feature encoding

**Answer: A,C**

### Question: 7

Which SQL commands can be used to modify the contents of a table's data directly?  
(Choose all that apply)

- A. UPDATE
- B. DELETE

- C. SELECT
- D. INSERT

**Answer: A,D**

### Question: 8

Which of the following best describes the purpose of the `plt.subplot()` function in Matplotlib?

- A. It overlays multiple plots in the same figure without axes
- B. It creates a 3D surface plot using pandas
- C. It configures multiple plots in a grid layout within a single figure
- D. It links Matplotlib with Seaborn's grid styling

**Answer: C**

### Question: 9

Which Python constructs are useful for comparing and managing object identity in a data pipeline?  
(Choose all that apply)

- A. `is` operator for memory-level comparison
- B. `id()` function to retrieve unique object identifiers
- C. `str()` function to convert data formats
- D. `==` operator for checking value equivalence

**Answer: A,B**

### Question: 10

What is a key assumption of linear regression that distinguishes it from logistic regression?

- A. Output values must be binary
- B. Predictors must be independent
- C. The relationship between predictors and the outcome is linear
- D. The target variable is unordered categorical

**Answer: C**

**Thank You for Trying Our Product**  
**Special 16 USD Discount Coupon: NSZUBG3X**  
**Email: [support@examsempire.com](mailto:support@examsempire.com)**

**Check our Customer Testimonials and ratings  
available on every product page.**

**Visit our website.**

**<https://examsempire.com/>**