

Mock_Test_I - Shashank Katti

MCQs

1. Which of the following libraries in Python is commonly used for numerical computing and working with arrays?
A) Pandas
B) Seaborn
C) NumPy
D) Matplotlib
2. Which of the following methods is used to remove missing data from a Pandas DataFrame?
A) dropna()
B) fillna()
C) replace()
D) isnull()
3. In a Seaborn pairplot(), what does the diagonal plot show by default?
A) Scatter plot
B) KDE (Kernel Density Estimate)
C) Box plot
D) Histogram
4. What is the output of the following code?

```
X = 12
if (X > 10 & X < 15):
    print('YES')
else:
    print('No')
```

- YES
- NO
- Error
- No Output

5. Complete the piece of code to print a maximum of three variables, a,b and c.

```
a=10
b=16
c=20
if(**missing part1**):
    print("a")
elif(**missing part 2**):
    print("b")
else:
    print("c")
```

- Missing Part 1: $a > b, c$
Missing Part 2: $b > a, c$
- Missing Part 1: $a > b$ and $a > c$
Missing Part 2: $b > a$ and $b > c$
- Missing Part 1: $c < a$ and $c < b$
Missing Part 2: $a < b$
- Missing Part 1: $(b < a, a > c)$
Missing Part 2: $(b > a, c < b)$

Subjective Questions

1. Define the term 'Data Wrangling in Data Analytics?'

Data Wrangling: The process of cleaning, transforming, and preparing raw data for analysis, involving handling missing values, outliers, data normalization, and data transformation.

2. How do you treat outliers in a dataset?

- Identify outliers using statistical methods (e.g., Z-score, IQR) or visualization.
- Handle outliers based on context:
- Remove or replace with mean/median/mode.
- Transform data (e.g., log scaling).
- Use robust algorithms.

3. How would you handle missing data in a dataset?

- Identify missing values.
- Handle missing data based on context:
- Delete rows/columns.
- Replace with mean/median/mode.
- Impute using regression or machine learning models

4. What are the most commonly used Python libraries for data analysis and their functions?

- NumPy: numerical computing.
- Pandas: data manipulation and analysis.
- Matplotlib/Seaborn: data visualization.
- Scikit-learn: machine learning.

5. How do you manipulate and aggregate data using pandas in Python?

- DataFrames for structured data.
- Use merge/join for combining data.
- groupby() for aggregation.
- pivot_table() for data summarization.

Coding Questions

1. Write a code to check if the string that the user enters starts with a vowel or not. Print capital YES or NO. For example, if input = 'analytics' then, your output should print 'YES'.

Sample Input:

Alpha, analytics

Sample Output:

YES

Sample Input:

Time, Sample

Sample Output:

NO

2. Given a single positive integer n, create a NumPy array of size (n x n) with alternating ones and zeros to form a checkerboard pattern.

Sample Input:

6

Sample Output:

```
[[1 0 1 0 1 0]
 [0 1 0 1 0 1]
 [1 0 1 0 1 0]
 [0 1 0 1 0 1]
 [1 0 1 0 1 0]
 [0 1 0 1 0 1]]
```