INFOLABZ IT SERVICES PVT. LTD.

Internship : Data Science

EXPLORE AND APPLY.

NUMPY:-

Numpy Image

1. Generate a script where the user can input the directory of an image, convert the image into grayscale, and save it into another folder.

Conditional Filtering & Boolean Indexing

1. Replace all negative values in an array with 0.
2. Count how many elements in an array are greater than a given number 25.
3. Extract all odd numbers from an array.

Sales Data Analysis using NumPy

Objective:

Analyze and visualize sales data for a company's products across four quarters using NumPy arrays. You will compute key metrics like total sales, average sales per product, and identify the best-performing product.

Dataset:

Here's a sample dataset with columns representing products and sales across four quarters.

```
Product    Q1 Sales    Q2 Sales    Q3 Sales    Q4
Sales
```

A	12000	15000	17000	20000
B	18000	22000	24000	28000
C	10000	12000	14000	16000
D	9000	11000	13000	15000

Steps:

1. **Create a NumPy Array for Sales Data: Convert the sales data into a 2D NumPy array.**
2. **Compute Total Sales for Each Product: Calculate the total sales for each product by summing across the quarters.**
3. **Compute Average Quarterly Sales for Each Product: Find the average sales per quarter for each product.**
4. **Find the Best-Performing Product: Identify the product with the highest total sales.**
5. **Visualize the Data (optional): If you'd like, use `matplotlib` to create a bar chart for the total sales and average quarterly sales of each product.**

PANDAS [SERIES]

Package installation : - Pip install pandas

Basic Operations

1. Create a Pandas Series from a Python list [10, 20, 30, 40, 50] and print its values.
2. Create a Series from a dictionary: {'a': 100, 'b': 200, 'c': 300}.
3. Convert a NumPy array [1.5, 2.5, 3.5, 4.5] into a Pandas Series.

Indexing & Selection

4. Access the third element of a Series.

5. Retrieve all elements greater than 20 from a Series.
6. Select elements at index positions [0, 2, 4] from a given Series.

Mathematical Operations

7. Given a Series s = pd.Series([5, 10, 15, 20]), multiply each element by 2.
8. Compute the mean, median, of a numerical Series.

Filtering & Transformation

10. Given a Series s = pd.Series([1, 2, 3, 4, 5, 6]), replace all even numbers with 0.
11. Find the count of missing (NaN) values in a Series.
12. Replace all NaN values in a Series with the mean of the non-null values.

String Operations

13. Create a Series of string values [Laptop, keyboard', Mouse'] and convert all elements to uppercase.
14. Count the occurrences of the letter 'a' in each string in a Series.

Sorting & Ranking

15. Sort a Series in ascending and descending order.
16. Rank the elements of a numerical Series.

GroupBy & Value Counts

19. Count the occurrence of each unique element in a Series.
20. Group a Series by values and calculate the sum for each unique value.