



S. B. JAIN INSTITUTE OF TECHNOLOGY, MANAGEMENT & RESEARCH, NAGPUR.

Practical No. 0 (Pre-Lab)

Aim: Understand and apply basic programming concept of Python.

Name of Student: ShrutiKA Pradeep Bagdi

Roll No.: CS22130

Semester/Year: IV/VII

Academic Session: 2024-2025

Date of Performance: _____

Date of Submission: _____

AIM: Understand and apply basic programming concept of Python

OBJECTIVE/EXPECTED LEARNING OUTCOME:

The objectives and expected learning outcome of this practical are:

- To understand basic libraries like pandas and numpy
- Able to perform basic operations using python

HARDWARE AND SOFTWARE REQUIREMENTS:

Hardware Requirement:

Software Requirement:

THEORY:

Pandas

Pandas is a powerful and popular Python library for data manipulation and analysis. It provides data structures and functions that make working with structured data (like spreadsheets or SQL tables) more efficient and convenient. Pandas is commonly used for tasks such as data cleaning, exploration, transformation, and analysis.

1. **DataFrame:** A DataFrame is a two-dimensional table-like data structure that consists of rows and columns. It's similar to a spreadsheet or a SQL table. Each column in a DataFrame can have a different data type.
2. **Series:** A Series is a one-dimensional labeled array that can hold any data type. It's like a single column of a DataFrame.
3. **Index:** An Index is a labeled array that identifies the rows or elements in a DataFrame or Series. It provides a way to uniquely label data points.
4. **Reading and Writing Data:** Pandas supports reading data from various file formats such as CSV, Excel, SQL databases, and more. Similarly, you can write data to these formats.

5. **Data Cleaning and Transformation:** Pandas provides functions for handling missing data, removing duplicates, and reshaping data. You can also perform various transformations like sorting, grouping, and pivoting.
6. **Data Selection and Filtering:** You can select specific rows and columns using labels or indexes. Pandas also allows you to filter data based on conditions.
7. **Operations on Data:** You can perform mathematical and statistical operations on data in a DataFrame or Series. Pandas provides flexible ways to apply functions to data.

Numpy

NumPy (Numerical Python) is a fundamental Python library for numerical computations, particularly when dealing with large arrays and matrices of numerical data. It provides support for multi-dimensional arrays, along with a collection of mathematical functions to operate on these arrays efficiently. NumPy is the foundation for many scientific and data analysis libraries in Python.

1. **Arrays:** The core data structure in NumPy is the **ndarray** (n-dimensional array), which is a homogeneous collection of elements of the same data type. NumPy arrays can have any number of dimensions.
2. **Array Creation:** NumPy provides functions to create arrays from lists, tuples, or other array-like objects. You can create arrays with specific data types and initial values.
3. **Array Operations:** NumPy allows you to perform element-wise mathematical operations, broadcasting (performing operations on arrays of different shapes), and aggregation functions (sum, mean, max, etc.).
4. **Indexing and Slicing:** You can index and slice NumPy arrays similar to Python lists. Slicing works in multiple dimensions.
5. **Array Shape and Reshaping:** You can query the shape of an array and reshape it using functions like **reshape()**, **resize()**, and **transpose()**.
6. **Array Concatenation and Splitting:** NumPy provides functions to concatenate and split arrays along specified dimensions.

7. **Universal Functions (ufuncs):** These are functions that operate element-wise on arrays. They provide high-performance mathematical operations.
8. **Array Broadcasting:** Broadcasting allows you to perform operations on arrays with different shapes, making it easier to work with arrays of different sizes.
9. **Linear Algebra:** NumPy provides various linear algebra operations such as matrix multiplication, eigenvalue decomposition, and more.

CODE:

```
import numpy as np
import pandas as pd
import nltk
from nltk.tokenize import word_tokenize
from nltk.corpus import stopwords
print("NumPy Array Example:")
arr = np.array([1, 2, 3, 4, 5])
print("Original Array:", arr)
print("Array + 5:", arr + 5)
print("Array squared:", arr ** 2)
print("Array mean:", np.mean(arr))
print("\n-----\n")
print("Pandas DataFrame Example:")
data = {
    "Name": ["Alice", "Bob", "Charlie"],
    "Age": [25, 30, 35],
    "Score": [85, 90, 95]
}
df = pd.DataFrame(data)
print("Original DataFrame:")
print(df)
print("\nFilter rows where Age > 28:")
print(df[df["Age"] > 28])
```

```

print("\n-----\n")
print("NLP Example with NLTK:")
nltk.download('punkt')
nltk.download('stopwords')
text = "Natural Language Processing with Python is fun and powerful."
print("Original Text:", text)
tokens = word_tokenize(text)
print("Tokens:", tokens)

```

OUTPUT (SCREENSHOT):

```

→ NumPy Array Example:
Original Array: [1 2 3 4 5]
Array + 5: [ 6  7  8  9 10]
Array squared: [ 1  4  9 16 25]
Array mean: 3.0

-----
Pandas DataFrame Example:
Original DataFrame:
   Name  Age  Score
0  Alice    25     85
1    Bob    30     90
2 Charlie    35     95

Filter rows where Age > 28:
   Name  Age  Score
1    Bob    30     90
2 Charlie    35     95

-----
NLP Example with NLTK:
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to /root/nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
[nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data]  Unzipping tokenizers/punkt_tab.zip.
Original Text: Natural Language Processing with Python is fun and powerful.
Tokens: ['Natural', 'Language', 'Processing', 'with', 'Python', 'is', 'fun', 'and', 'powerful', '.']

```

Select a Language which you know better

English

Select a word from the below dropdown and do a morphological analysis on that word

watching

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	watching	
ROOT	watch	✓
CATEGORY	verb	✓
GENDER	male	✓
NUMBER	singular	✓
PERSON	second	✓
CASE	na	✓
TENSE	present-continuous	✓

Check

Right answer!!!



Word Analysis

Select a Language which you know better

English

Select a word from the below dropdown and do a morphological analysis on that word

train

Select the Correct morphological analysis for the above word using dropboxes (NOTE : na = not applicable)

WORD	train	
ROOT	train	✓
CATEGORY	noun	✓
GENDER	female	✓
NUMBER	singular	✓
PERSON	first	✓
CASE	oblique	✗
TENSE	past-continuous	✗

Wrong answer!!!

Check

Get Answers

CONCLUSION:

DISCUSSION AND VIVA VOCE:

- Explain pandas and its functions
- Explain Numpy and its functions
- What are the applications of pandas and numpy?

REFERENCE:

- www.w3schools.com
- www.tutorialsmade.com
- <https://www.javatpoint.com/>