

Python Data Handling with Pandas and NumPy

Internship Assignment Report

Assignment Title: Python Data Handling with Pandas and NumPy

Submitted by: Sai Manikanta

Internship Domain: Machine Learning

Year: 2025

*This report contains code snippets and explanations for basic data handling
operations using Pandas and NumPy in Python.*

Python Data Handling with Pandas and NumPy

Python Data Handling with Pandas and NumPy

1. Creating and Manipulating DataFrames

```
import pandas as pd

data = {
    'Name': ['Alice', 'Bob', 'Charlie', 'David', 'Eva'],
    'Age': [24, 27, 22, 32, 29],
    'Department': ['HR', 'Finance', 'HR', 'IT', 'Finance'],
    'Salary': [40000, 52000, 41000, 58000, 60000]
}

df = pd.DataFrame(data)
print(df)
```

2. Filtering Data

```
# Filter employees with Salary > 50000
high_salary = df[df['Salary'] > 50000]
print(high_salary)
```

3. Grouping Data

```
# Group by Department and calculate average salary
avg_salary = df.groupby('Department')['Salary'].mean()
print(avg_salary)
```

4. NumPy Basics

```
import numpy as np

# Create NumPy array
arr = np.array([10, 20, 30, 40, 50])

# Perform basic operations
print("Mean:", np.mean(arr))
print("Max:", np.max(arr))
print("Min:", np.min(arr))
```