Training Day-11 Report:

What is Pandas?

Pandas is a Python library used for working with data sets.It has functions for analyzing, cleaning, exploring, and manipulating data.The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis" and was created by Wes McKinney in 2008.

Why Use Pandas?

Pandas allows us to analyze big data and make conclusions based on statistical theories.Pandas can clean messy data sets, and make them readable and relevant.Relevant data is very important in data science.

What Can Pandas Do?

Pandas gives you answers about the data. Like:

1. Is there a correlation between two or more columns?

2. What is average value?
3. Max value?
4. Min value?

Pandas Series

A Pandas Series is like a column in a table.It is a one-dimensional array holding data of any type.

Example: Create a simple Pandas Series from a list:

```
import pandas as pd
a = [1, 7, 2]
myvar = pd.Series(a)
print(myvar)
```

Pandas DataFrames

```
import pandas as pd

data = {
  "calories": [420, 380, 390],
  "duration": [50, 40, 45]
}
df = pd.DataFrame(data)
print(df)
```

Pandas Read CSV:

A simple way to store big data sets is to use CSV files (comma separated files).CSV files contains plain text and is a well know format that can be read by everyone including Pandas.In our examples we will be using a CSV file called 'data.csv'.

```
import pandas as pd
df = pd.read_csv('data.csv')
print(df.to_string())
```

Pandas Read JSON: g data sets are often stored, or extracted as JSON.JSON is plain text, but has the format of an object, and is well known in the world of programming, including Pandas.In our examples we will be using a JSON file called 'data.json'.

Example

```
import pandas as pd
df = pd.read_json('data.json')
print(df.to_string())
```