Networking Academy

Python DTS PRoA 2022

Library Session 3: Pandas

Muhammad Ogin Hasanuddin

KK Teknik Komputer
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung

Introduction

- Pandas is an open-source library for data analysis and manipulation. It is a go-to toolkit for data scientists.
- Pandas integrates seamlessly with other Python libraries such as NumPy and Matplotlib for numeric processing and visualizations.
- When using Pandas, we will primarily interact with DataFrames and Series.
- Pandas “panel data”

Importing Pandas

- In order to use Pandas, you must import it. This is as simple as:

```
import pandas
```

- However, you'll rarely see Pandas imported this way. By convention programmers rename Pandas to pd. This isn't a requirement, but it is a pattern that you'll see repeated often.
- To import Pandas in the conventional manner run the code block below.

```
import pandas as pd
```

- After importing Pandas as pd we can use pandas by calling methods provided by pd.
- For example we can print pandas version by run code below

```
pd.__version__
```

Pandas Series

A Series represents a sequential list of data. It is a foundational building block of the powerful DataFrame.

Creating a Series

- We create a new Series object as we would any Python object:

```
s = pd.Series()
```

- This creates a new, empty Series object, which isn't very interesting. You can create a series object with data by passing it a list or tuple:

```
temperatures = [55, 63, 72, 65, 63, 75, 67,
59, 82, 54]

series = pd.Series(temperatures)

print(type(series))
print(series)
```

- Here we created a new pandas.core.series.Series object with ten values presumably representing some temperature measurement.

Pandas DataFrame

If you picture Series as a list of data, you can think of DataFrame as a table of data. A DataFrame consists of one or more Series presented in a tabular format. Each Series in the DataFrame is a column.

Creating a DataFrame

- We can create an empty DataFrame using the DataFrame class in Pandas:

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

Terimakasih!