

Welcome to **INTERNSHIP STUDIO**

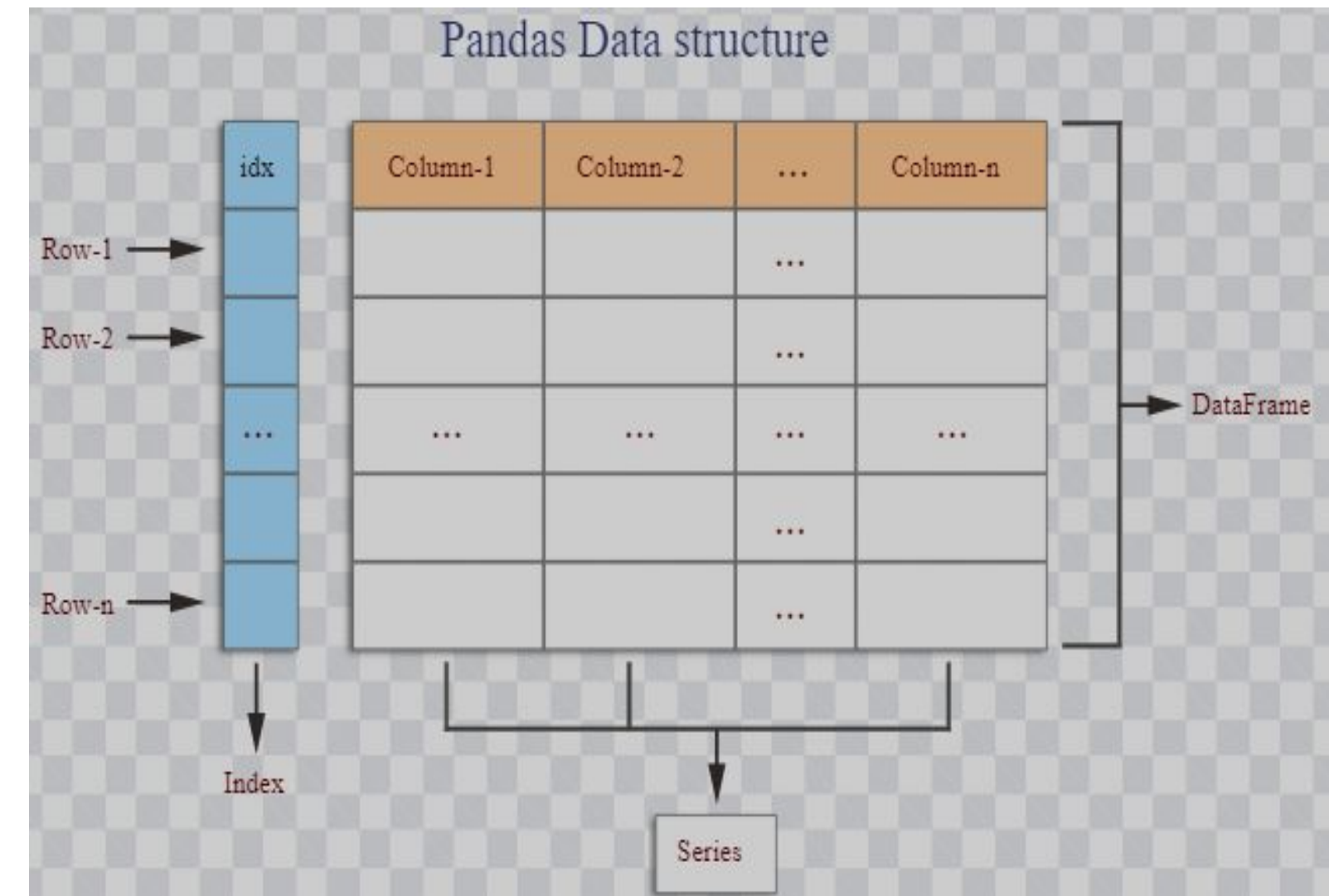
Module 04 | Lesson 03

Pandas

Data Structures in Pandas

Pandas Data Structures Overview

- Pandas provides two main data structures: Series and DataFrame.
- Series: a one-dimensional labeled array capable of holding any data type.
- DataFrame: a two-dimensional labeled data structure with columns of potentially different data types.



Creating a Series

- Use the `pd.Series()` function to create a Series object.
- Syntax: `pd.Series(data, index)`

Manipulating a Series

- **Accessing elements**: `series[index]`
- **Slicing**: `series[start:end]`
- **Arithmetic operations**: `series1 + series2`
- **Applying functions**: `series.apply(function)`

Creating a DataFrame

- Use the `pd.DataFrame()` function to create a DataFrame object.
- Syntax: `pd.DataFrame(data, index, columns)`

Manipulating a DataFrame

- Accessing columns: `df['column_name']`
- Accessing rows: `df.loc[index]` or `df.iloc[row_number]`
- Adding a new column: `df['new_column'] = values`
- Filtering rows: `df[df['column'] > value]`
- Applying functions: `df.apply(function, axis=1)`

SUMMARY

You got

this

- Pandas provides powerful data structures: Series and DataFrame.
- Series is a one-dimensional labeled array.
- DataFrame is a two-dimensional labeled data structure.
- Use `pd.Series()` and `pd.DataFrame()` to create these structures.
- Manipulate the data using indexing, slicing, arithmetic operations, and functions.

Next

session

Some more live coding on
Jupyter notebook