**Lab – 24**

**Student Name: Navneet P**

**Student ID: AF0411619**

***Topic-Pandas Library***

* What is Pandas?

Pandas is an open-source data manipulation and analysis library for the Python programming language that is made mainly for working with relational or labeled data both easily and intuitively. It provides various data structures and operations for manipulating numerical data and time series and for working with structured data, making it one of the most popular and widely used libraries for data analysis and manipulation in the Python ecosystem.

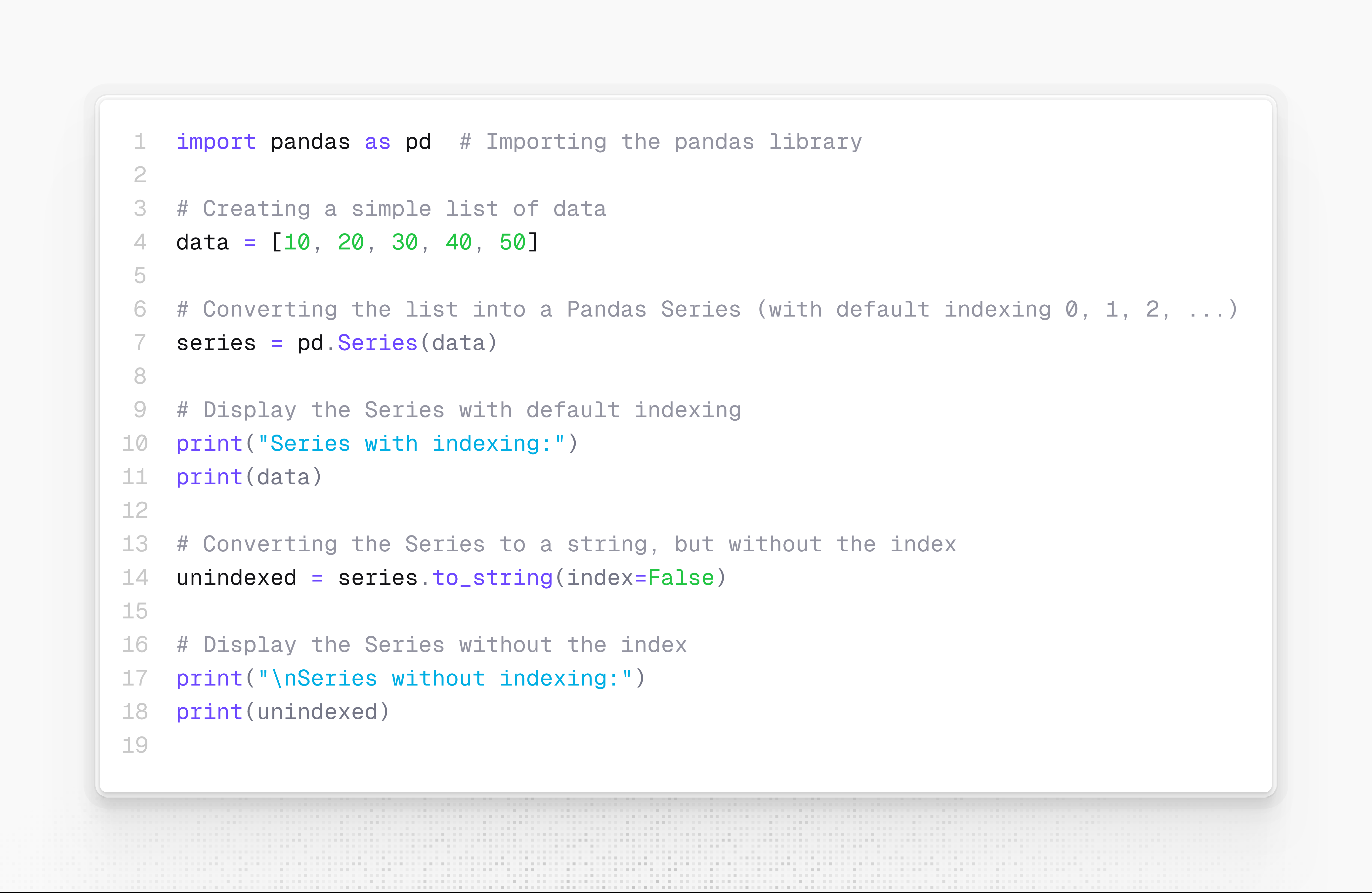
* Installation
  + Pip install pandas
* Data Cleaning
* Pandas includes functions for handling missing data, removing duplicates, and transforming data to make it suitable for analysis.
* Dataframes
  + The DataFrame is one of the central data structures in Pandas. It is a two-dimensional table with rows and columns, similar to a spreadsheet or a SQL table.
* Series
  + A Series is a one-dimensional array-like object in Pandas. It can be thought of as a single column of data within a DataFrame, with an associated index
* Grouping
  + Pandas allows you to group data by one or more columns and perform aggregation operations on the groups.
* Why Pandas?
  + Fast and efficient for manipulating data. Data from different file objects can be easily loaded. Flexible reshaping, time-series functionality.

Functions used in this assignment:-

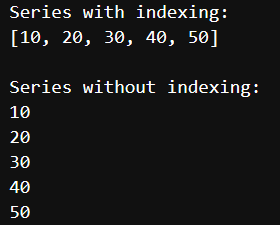
1. **Pandas.Series** – A one-dimensional labeled array capable of holding any data type, similar to a column in a DataFrame.
2. **Pandas.DataFrame** – A two-dimensional, size-mutable, and potentially heterogeneous tabular data structure with labeled axes (rows and columns).
3. **to\_string()** – Converts a DataFrame or Series to a string representation for more readable console output.
4. **pandas.read\_excel()** – Reads an Excel file into a DataFrame, allowing you to load data from spreadsheets.
5. **head()** – Returns the first few rows of a DataFrame (default is 5 rows).
6. **tail()** – Returns the last few rows of a DataFrame (default is 5 rows).
7. **loc[]** – Accesses a group of rows and columns by labels or a boolean array in a DataFrame.

Q1. Create Pandas series [10,20,30,40,50] without index

Solution:



Output:

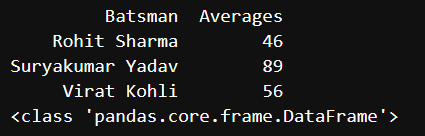


Q2. Create Pandas DataFrame example.

Solution:



Output:

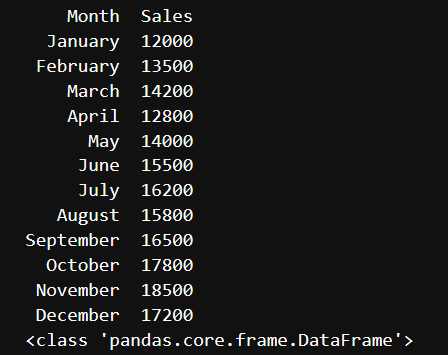


Q3. Example: Monthly Sales Data Imagine you are a sales manager for a retail company, and you want to analyze the monthly sales performance of a particular product in a given year. You have recorded the monthly sales figures for that product, and you want to represent this data using a Pandas Series.

Solution:



Output:



Example: Read an excel file and demonstrate the use of loc slicing, head and tail methods.

Solution:



Output:

