Pandas

- 1. Pandas is a Python library for data manipulation and analysis.
- 2. It provides two main data structures, Series (1-dimensional) and DataFrame (2-dimensional).
- 3. Pandas can read data from various sources such as CSV, Excel, SQL databases, and web scraping.
- 4. Pandas also supports data visualization through integration with Matplotlib.
- 5. It has built-in methods for handling missing data, time series data, and text data.

Series

- 1. A pandas Series is a one-dimensional labeled array that can hold any data type (integers, floats, strings, etc.).
- 2. It is similar to a Python list or NumPy array, but with added functionality for indexing and labeling.
- 3. Series can be created from various data sources such as lists, NumPy arrays, dictionaries.
- 4. The index of a Series is a sequence of labels that can be customized.
- 5. Series can be sliced and indexed by both position and label.
- 6. operations can be performed on the entire Series at once without the need for loops.
- 7. Pandas offers many built-in methods for data manipulation and analysis on Series, including arithmetic operations, statistical analysis, and data alignment.
- 8. Series can be used as input to many pandas functions, including the creation of DataFrames.

DataFrame

- 1. A pandas DataFrame is a two-dimensional labeled data structure that can hold any data type (integers, floats, strings, etc.).
- 2. It is similar to a spreadsheet or SQL table, but with added functionality for indexing and labeling.
- 3. DataFrames can be created from various data sources such as lists, dictionaries, NumPy arrays, and CSV/Excel files.

- 4. DataFrames have both row and column labels, which can be customized.
- 5. DataFrames can be sliced and indexed by both position and label, on both rows and columns.
- 6. Pandas offers many built-in methods for data manipulation and analysis on DataFrames.
- 7. DataFrames support vectorized operations, meaning that operations can be performed on entire columns or rows at once without the need for loops.
- 8. Pandas also offers many functions for handling missing data, reshaping data, and data cleaning.

REFs:

Pandas Tutorial: DataFrames in Python | DataCamp Pandas_Cheat_Sheet.pdf (pydata.org)