

Pandas Series

Pandas Series is a one-dimensional labeled array capable of holding any data type. Pandas Series is built on top of NumPy array objects.

How Pandas Series is different from 1-D Numpy Array

1. Pandas Series can hold a variety of data types whereas Numpy supports only numerical data type
2. Pandas Series supports index labels.

Pandas DataFrame

Pandas DataFrame is a two dimensional labeled data structure. It consists of rows and columns.

Each column in Pandas DataFrame is a Pandas Series.

`df.head()`

Returns first 5 rows of dataframe (by default).

`df.tail()`

Returns the last 5 rows of the dataframe (by default).

`df.info()`

It prints the concise summary of the dataframe. This method prints information of the dataframe like column names, its datatypes, nonnull values, and memory usage

`df.dtypes()`

Returns a series with the datatypes of each column in the dataframe.

`df.shape`

Return the number of rows and columns of the dataframe.

`df.values`

Return the NumPy representation of the DataFrame.

`df.to_numpy()` → This also returns the NumPy representation of the dataframe.

`df.columns`

Return the column labels of the dataframe

`df.describe()`

Generates descriptive statistics. It describes the summary of all numerical columns in the dataframe.

`df.describe(include="all")`

It describes the summary of all columns in the dataframe.

`df.col_name.unique()`

Returns the unique values in the column as a NumPy array.

`df.col_name.value_counts()`

Return a Series containing counts of unique values.

```
df.col_name.astype()
```

Converting datatype of a particular column.

```
df.sort_values(by="Col_1")
```

It will sort the dataframe by column "Col_1"