PANDAS.

- It is a python library used for working with data sets.

- It have functions for analyzing, cleaning, exploring and manipulating data.

has a reference

Pandas ⟶ "Panel Data" & " Python Data Analysis.

* It gives answers about the data. like.

    * Max value ?

    * Min Value ?

    * Avg value ?

    * Is there a correlation b/w 2 or more columns ?

→ mean () : Calculate the mean of a Series or Data Frame.

    std () : Calculate the standard deviation of a series a DataFrame.

    describe () : Calculate summary statistics for a series a Data Frame.

    plot() : Plot a series a Data Frame.

Functions

1. **Read Data**

   Import pandas as pd.

   df = pd.read_csv("heart_failure.csv")

2. **Head & Tail**

   df.head() & df.tail.

3. **Shape, Size and Info**

   df.shape — total no. of rows & then columns.

   df.size() — returns the number of rows times number of columns in the data frame.

   df.info() — Information such as rows from Range Index, Data columns and then data type of each column.

* **Describe**

   df.describe() → it will give you count, mean, standard deviation & also 5 number summary.

* **Nunique**

   df.nunique().

   It will give all the unique values

a variable contains:

* Value counts.

    df.anaemia.value_counts().

Pandas dtype	Python type	Usage
Object	Str or mixed	Txt or mixed numeric eq non-numeric values.
int64	int	integer numbers
float64	float	Floating point numbers
bool	bool	True / False values.