

19/10/23

## Pandas

Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring and manipulating data.

→ Import pandas as pd

Once you import it you can take your data analysis to whole new level.

→ `df = pd.read_csv("data file name.csv")`

Used to read the csv file imported.

→ `df.head`

Find the headings from the csv file

→ `df.tail`

Find the last 5 heading

→ `df['Name'].dtype`

`df.dtypes`

Used to find the data types in it

→ `df.columns`

Used to find the number of columns and heading

→ `df.shape`

→ `df['name'].mean()`

find the mean value

→ `df.groupby('name')[''] - mean()`

→ df.describe().

It will give you count, mean standard deviation and also 5 number summary

→ df.numunique().

It will give all the unique values a variable counting

→ df.anaemia.value\_counts().

Panda dtype - Pythontype - Usage

Object

- str - mixed  
- Int - a mixed number & non-numeric values

int 64

- int

- Integer numbers

Float 64

- float

- Floating point numbers

bool

- bool

- True / False values