Pandas 1

Wednesday, 29 October 2025 9:03 PM

{ np.array( ) } Notes

* python * } 2x

Extra

{ ① axis → 1 — Row
0 — col.

np.sum(axis) —

★ ★ ★ Pandas ★ ★ ★ ★

↓

Most powerful tool

ML → learn patterns

↓

EDA → DAV

Num Pands Mat./seabo

3

Numerical

{Numerical votes Rating cost } ↑

text { Zomato City Address Name, cuisine, reviews

↳ Banglore

MUMBAI

AMRAVATI

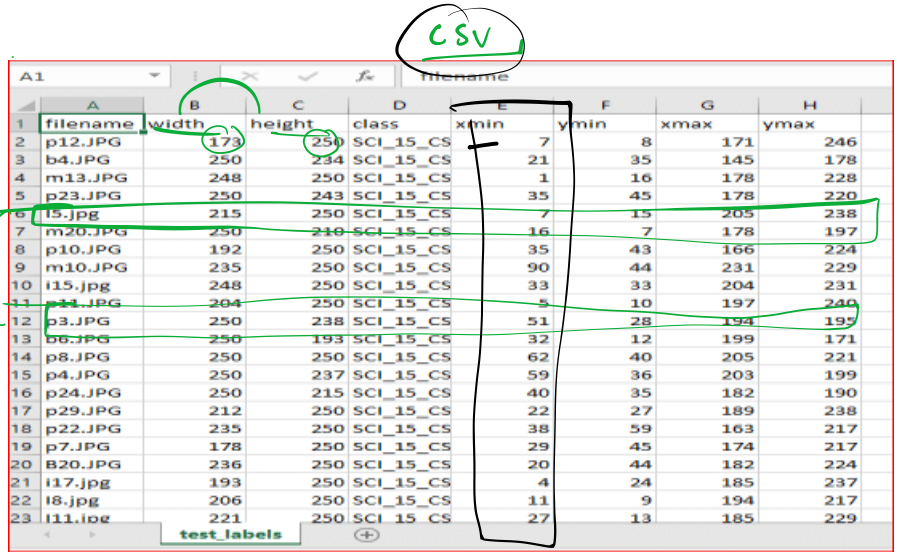
Numpy

Pandas

SQL }

[ ] [ ]

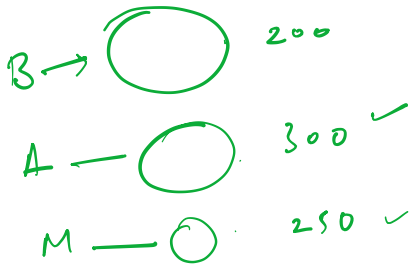
csv excel json

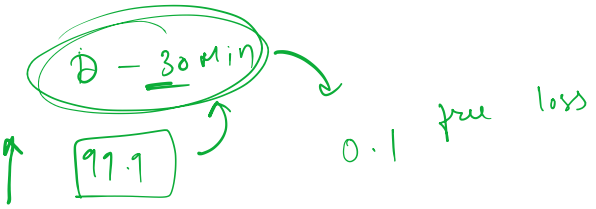


	A	B	C	D	E	F	G	H
1	filename	width	height	class	xmin	ymin	xmax	ymax
2	p12.JPG	173	250	SCI_15_CS	7	8	171	246
3	b4.JPG	250	234	SCI_15_CS	21	35	145	178
4	m13.JPG	248	250	SCI_15_CS	1	16	178	228
5	p23.JPG	250	243	SCI_15_CS	35	45	178	220
6	I5.jpg	215	250	SCI_15_CS	7	15	205	238
7	m20.JPG	250	210	SCI_15_CS	16	7	178	197
8	p10.JPG	192	250	SCI_15_CS	35	43	166	224
9	m10.JPG	235	250	SCI_15_CS	90	44	231	229
10	I15.jpg	248	250	SCI_15_CS	33	33	204	231
11	p11.JPG	204	250	SCI_15_CS	5	10	197	240
12	p3.JPG	250	238	SCI_15_CS	51	28	194	195
13	b6.JPG	250	193	SCI_15_CS	32	12	199	171
14	p8.JPG	250	250	SCI_15_CS	62	40	205	221
15	p4.JPG	250	237	SCI_15_CS	59	36	203	199
16	p24.JPG	250	215	SCI_15_CS	40	35	182	190
17	p29.JPG	212	250	SCI_15_CS	22	27	189	238
18	p22.JPG	235	250	SCI_15_CS	38	59	163	217
19	p7.JPG	178	250	SCI_15_CS	29	45	174	217
20	B20.JPG	236	250	SCI_15_CS	20	44	182	224
21	I17.jpg	193	250	SCI_15_CS	4	24	185	237
22	I8.jpg	206	250	SCI_15_CS	11	9	194	217
23	I11.jpg	221	250	SCI_15_CS	27	13	185	229

test_labels

Bag
A
B
A
B
M





Agenda

Pandas

Col / 1D — Series

R, C — DF

① Accessing one column — dot: df.colname; []: df['col']

② index ★★★★ pd.Series([ ], index)

③ head (top), tail (last), describe (stat), shape, info — dtype, non null

. df.rename(col = {old: new})

- Add a column
- drop a column

} inplace ★★★ Permanent

df[new] = ⇐

df.drop axis = 0 K
1 C

{ ① Null values }

None
null
empty

Age

27
24

Gender Diab

Gender	Diab
M	1
F	0
M	1
M	0
M	0
F	1
M	1
M	1
F	1
F	1

Analysis X

M – 1

F – 0

✓ unique → 10

✓ value count (B) → ✓ 100 ★★★★

(A) – 20

★★★★ LOC and ILOC ★★★

import numpy as np

import pandas as pd

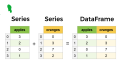
Dataframe → Table (2D)

Series (1D)

C

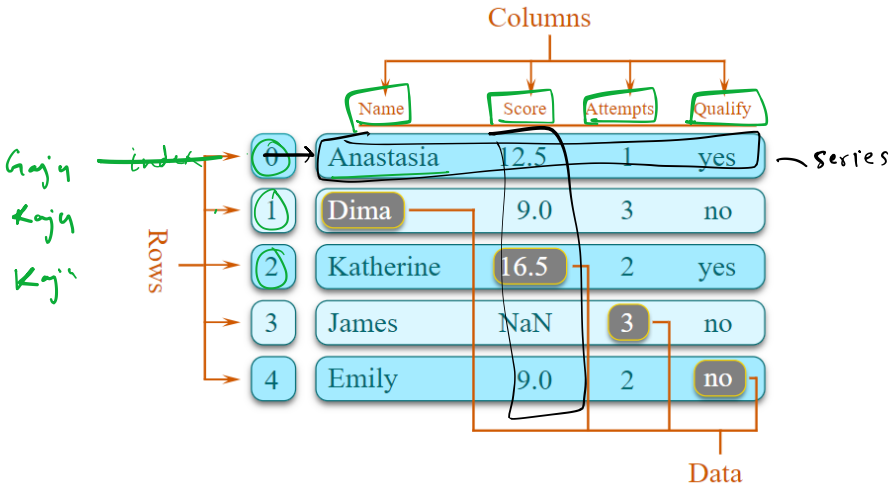
R

1 single column



Series Series DataFrame

Series		Series		DataFrame	
	apples		oranges		apples	oranges
0	3	0	0	0	3	0
1	2	+ 1	3	= 1	2	3
2	0	2	7	2	0	7
3	1	3	2	3	1	2



DF

Python + Pandas

1 hr.

Series → 1 Column

1 Column - Series

1 ≥ DF

focus X Explaining

Slicing

CSV

excel

df [" name of col "]

df . name_of_column

Space

* Multiple column. [list of col]

Rename a column.

df.rename → It is creating a new df → changes

↓ old → remains same.

inplace = True → do not create new df modify the old.

[[[[

2D

output → operation — is not permanent

no output → operation → permanent oper

0 1 2 3 4

A K A S H

N → empty

error → "N"

df["co"].replace("N", np.nan) → empty.

pd.to_numeric( ) , errors = coerce

4.1
7.2
4.3
4.1
{A N}
4.2

A

np.nan B

C

K

AK

Numpy

axis = 0 → G

K, C

x

Pandas

axis = 0 → K

= 1 → C

1 →

- handling Missing values
- Loc and iLoc
- duplicate
- Sort

Rating

3 → Awesome

5 —

4
3
2
1 } Menti

Metrics — Placed

① quiz

What → sol

100 — 15-20

9545126535