

Target \Rightarrow Product house prices

Fitting / Training \rightarrow predict

CSV

Table \rightarrow dataframe

excel / csv

20% \rightarrow \downarrow

20% test

Connect w. db

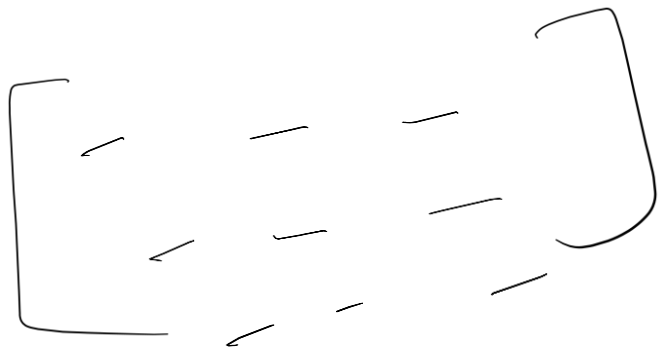
70% \rightarrow

30% \rightarrow

pre-processing

Modelling

Numpy



~~Numpy~~ -
Numerical python

np.array([list])

→ [-, -, -, -]
→ 1D

Scalar :-

1×1 matrix

$$a = 5 \quad (\text{scalar})$$

$$b = 10 \times 3 \Rightarrow 30 \quad (\text{scalar})$$

$$c = 10^5 = 100000 \quad (\text{scalar})$$

Vector:

$[5, 7, 8, 10, 15]$

\hookrightarrow 1D

$(\underbrace{1}_n, \underbrace{5}_n)$

Numpy

Series of continuous element

Matrix: 2D \Rightarrow array



Matrix

$n \times n$
matrix

Col 1	Col 2	Col 3	Col 4
5	1	2	8
9	11	12	13
16	7	4	5

row \times cols

(3×4)

\rightarrow shape of 2D
Numpy array

Matrix Multiplication: (Dot product)

np. Dot

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$\begin{bmatrix} 7 & 10 \\ 15 & 20 \\ 11 & 13 \end{bmatrix}$$

$$\begin{matrix} \textcircled{A} & \textcircled{3} & \times & \textcircled{3} \\ \downarrow & \downarrow & & \downarrow \\ m & n & & n \\ \text{(row)} & \text{(col)} & & \end{matrix}$$

$$\begin{matrix} \textcircled{3} & \times & \textcircled{2} & \textcircled{B} \\ \downarrow & & \downarrow & \\ \text{(row)} & & \text{(col)} & \end{matrix}$$

Col₁ Col₂

$$\begin{bmatrix} 7+30+33 & 10+40+39 \\ \hline \end{bmatrix}$$

shape (3, 2)

$$\begin{bmatrix} 1 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix} + \begin{matrix} 1 \\ 5 \end{matrix} = \begin{pmatrix} 1+5 & 4+5 & 5+5 \\ 6+5 & 7+5 & 8+5 \end{pmatrix}$$

2x3

1 \Rightarrow $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ Append

2 \Rightarrow

col₁, col₂, col₃

$$axis = 0$$

$$axis = 1$$

$$\begin{bmatrix} \textcircled{5} & \textcircled{6} & \textcircled{7} \\ 8 & 9 & 10 \\ 11 & 12 & 13 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\left[\begin{array}{l} \textcircled{5} + 0 + 0 \\ 0 + \textcircled{6} + 0 \\ 0 + 0 + \textcircled{7} \end{array} \right]$$

↳ transpose :

→
→
→

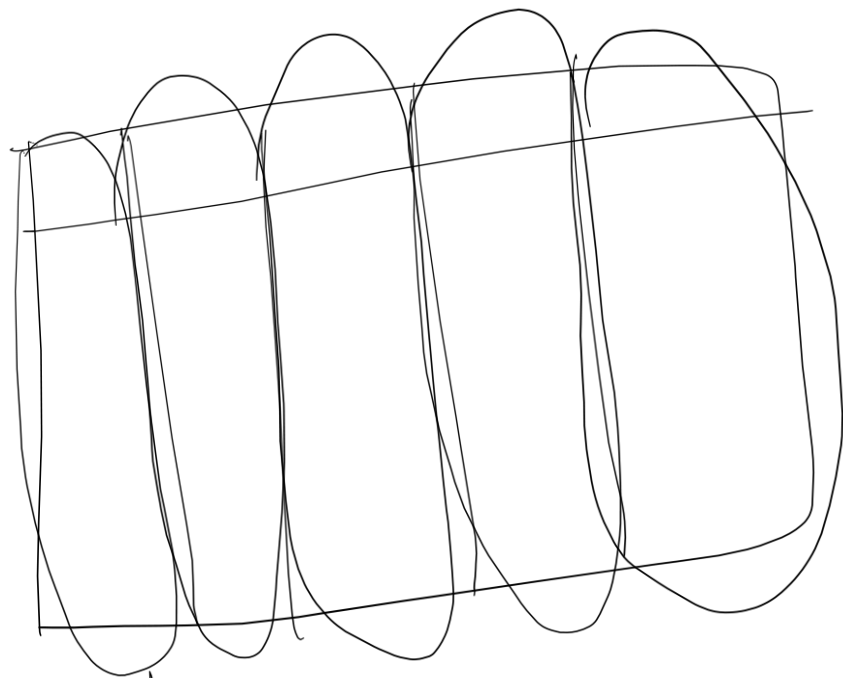
$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

3 × 2

• T

$$\Rightarrow \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ \textcircled{0} & \textcircled{2} & \textcircled{3} \end{bmatrix}$$

2 × 3

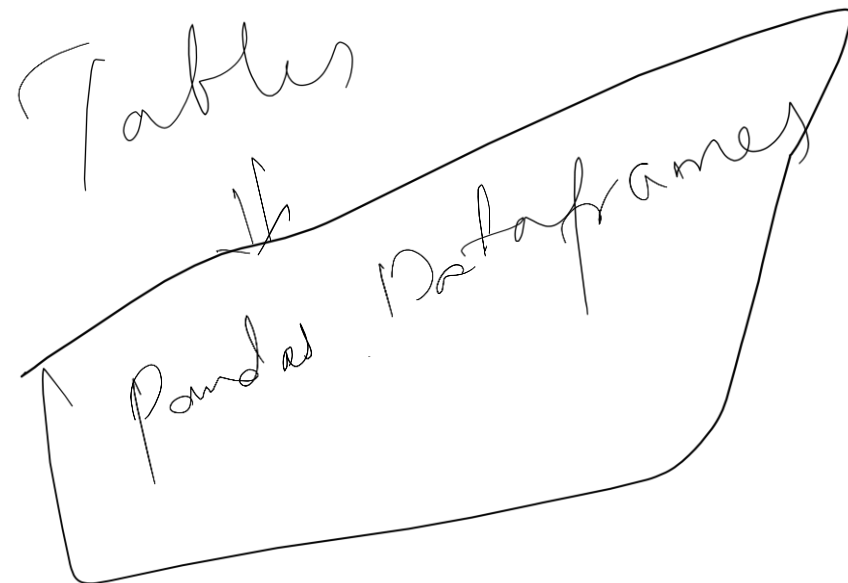


$Col_1 \rightarrow$ Series 1

$Col_2 \rightarrow$ " 2

$Col_n \rightarrow$ " n

multiple series



loc → Names of 'index' / 'col'

i loc → position / index number
of index / col

|

df

	col 1	col 2	col 3
a' ← 0	1	A	SS
b' ← 1 ✓	2	B	RR
c' ← 2 ✓	3	C	VV
d' ← 3 ✓	4	D	LL

df.iloc[1, ~~2~~] = B

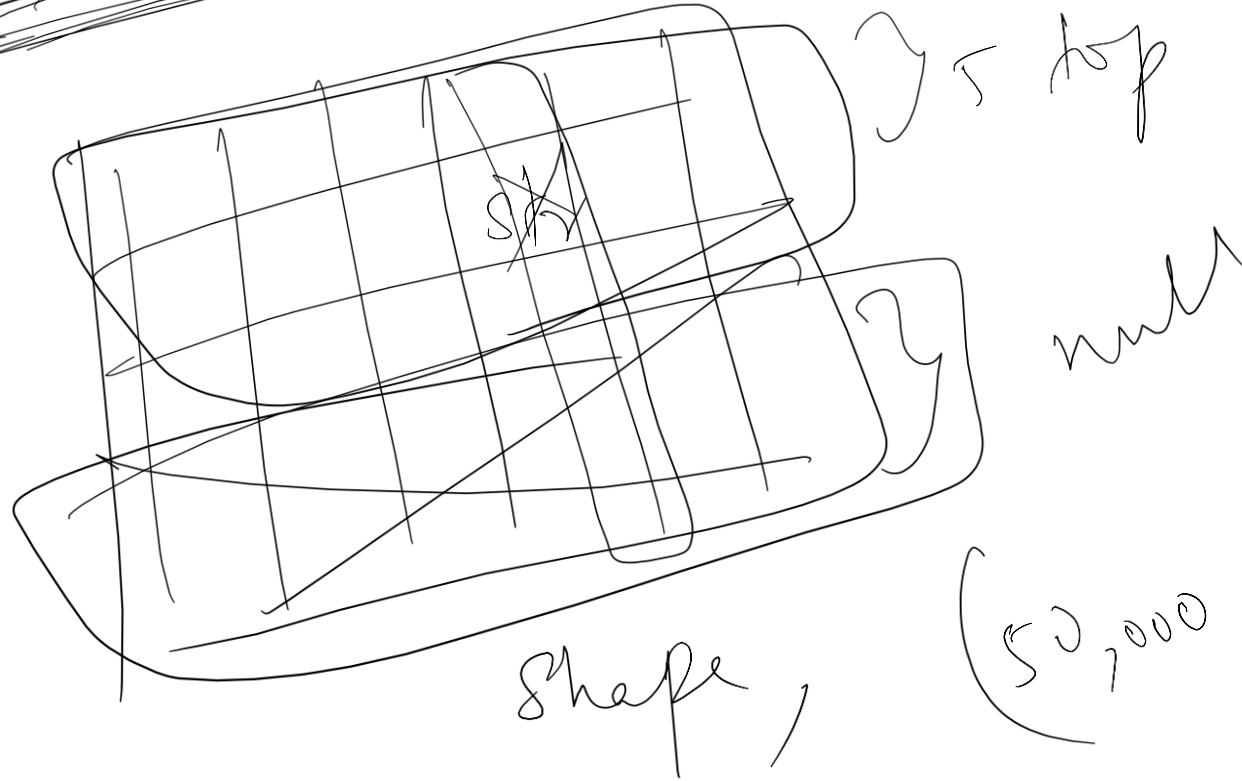
df.loc[:, col 3]

✓ df.iloc[1] ⇒ 2, B, RR

df.loc['c'] ⇒ 3, C, VV

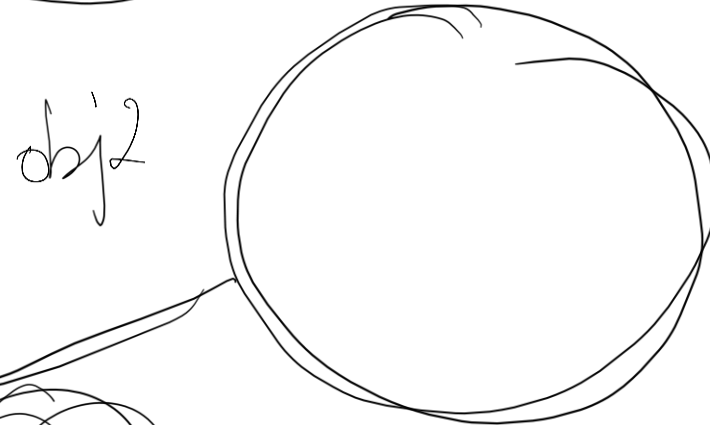
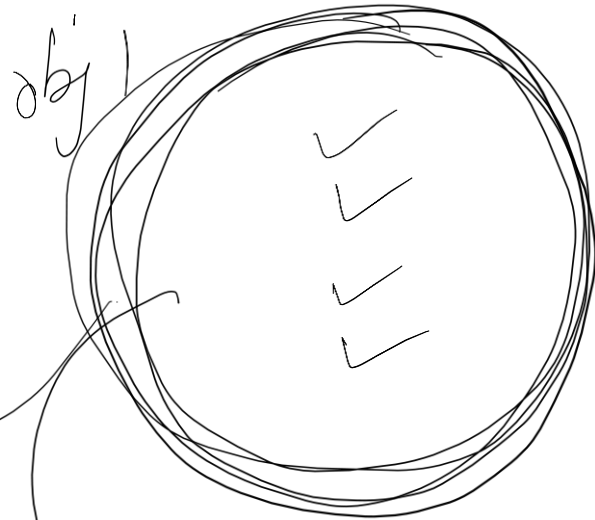
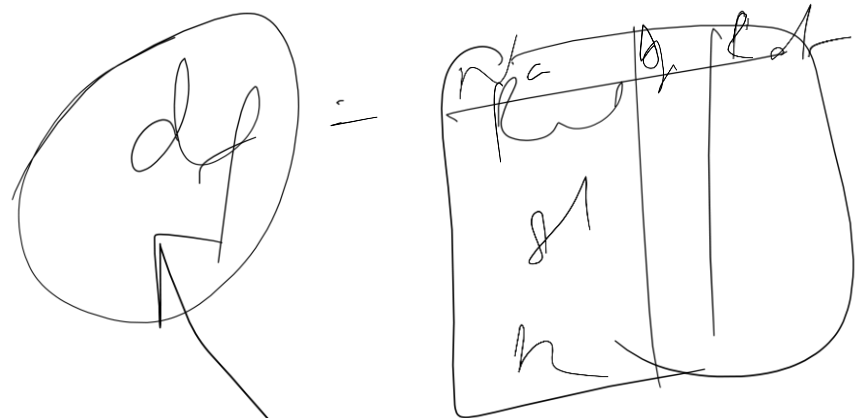
SS, RR, VV, ...

EDA → Exploratory Data Analysis



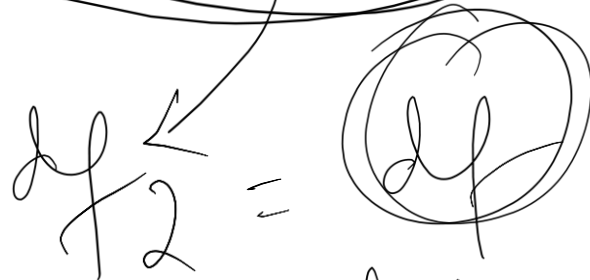
(img = 10;

(50,000 / 20) drop



id(df) → 123

id(df1) → 456



id(df2) ⇒ 123

