

Agenda

- \* Restructuring Data
- \* Dealing with Missing Values
- \* String Method in Pandas
- \* Handling Datetime
- \* Writing to a File.

world bank CO<sub>2</sub>

easy - medium  
=

Nan

1/2 → 1 hr

Assignments

Goal →

Data FAI

# Grouping

Group based

Aggr.

$d_1$

60m  
~~60~~

50m  
~~50~~

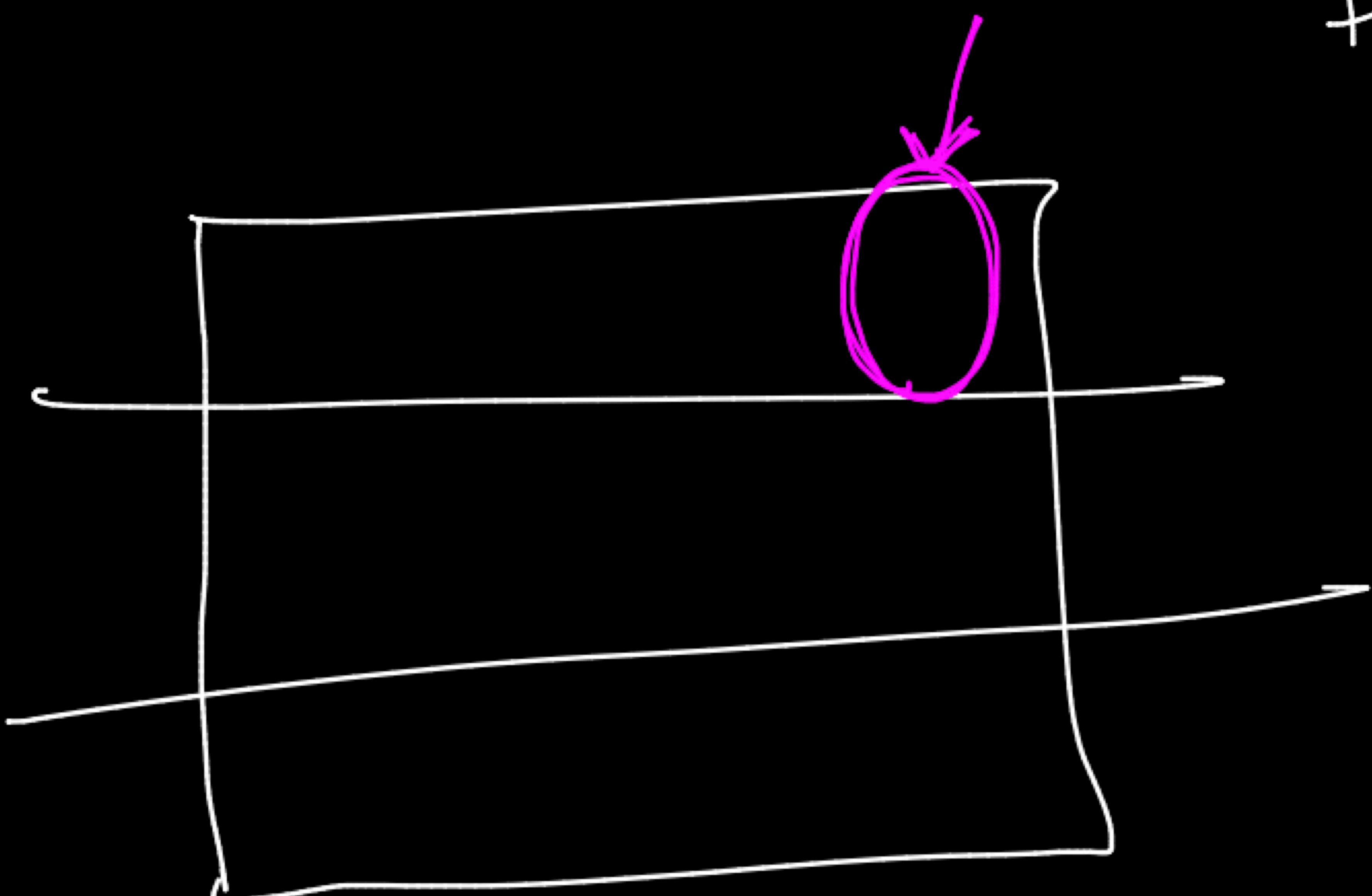
filter

Apply

$>=0$

reject

True  
False

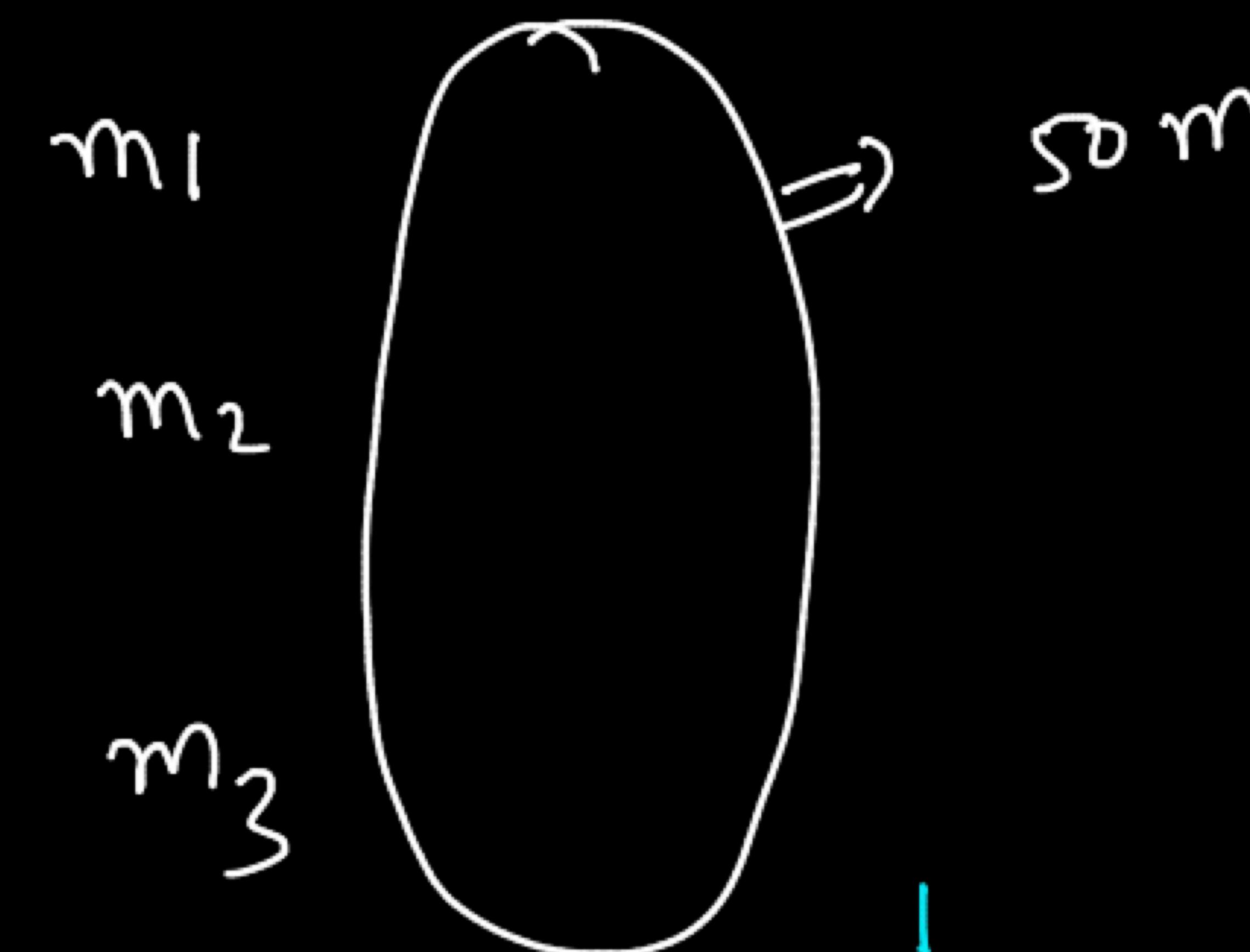


$budget - avg >= 0$

~~$budget - avg >= 0$~~

$60 - 50 >= 0$

d<sub>1</sub> 60  
~~34~~

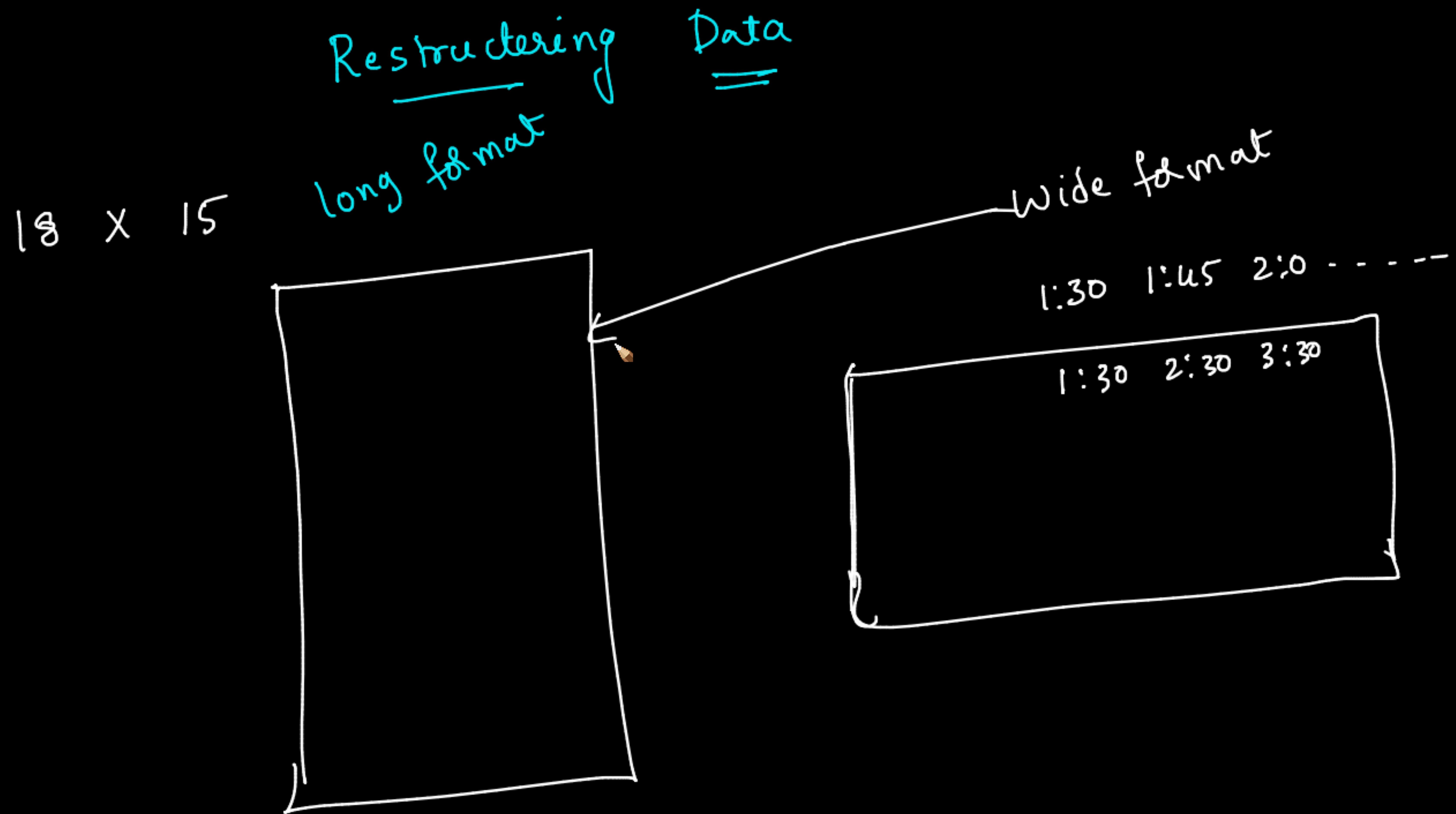


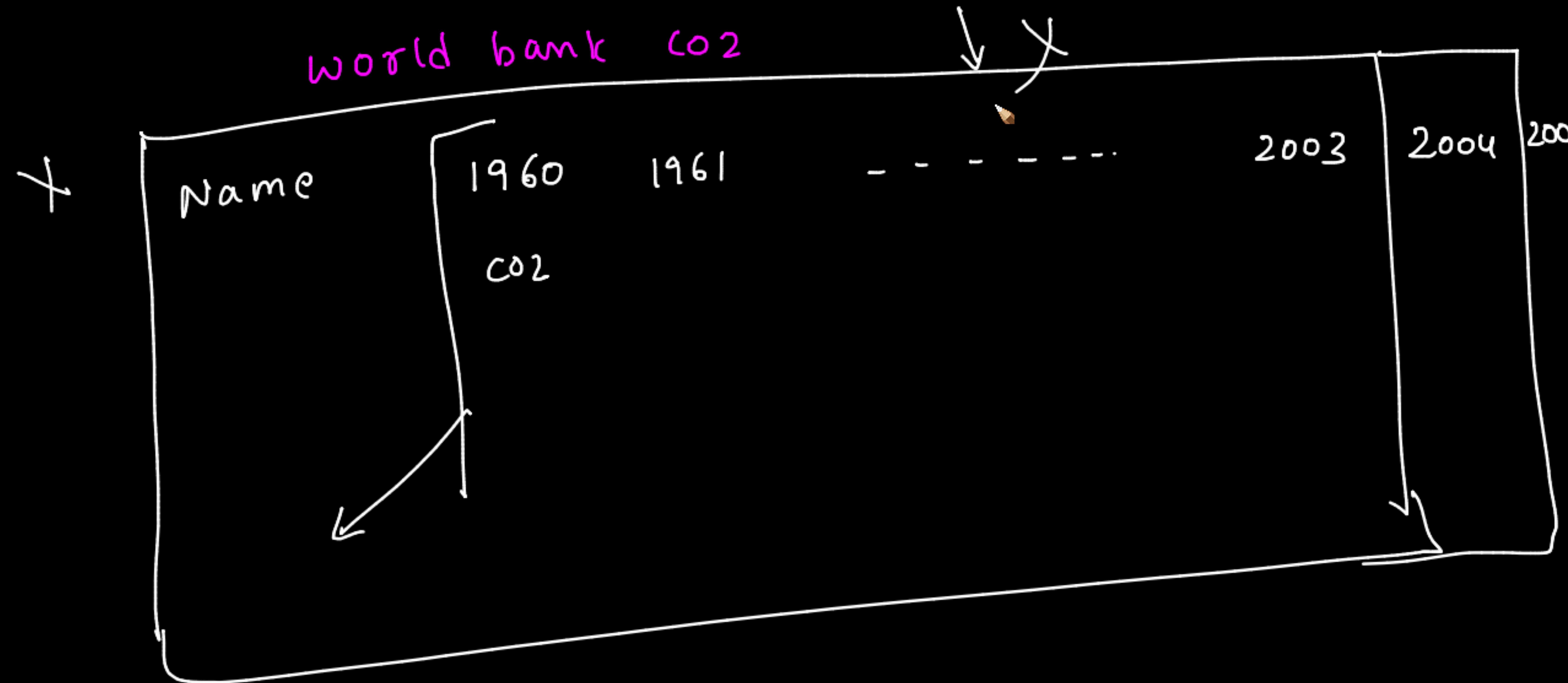
A hand-drawn diagram illustrating a geological fault. A horizontal line representing the ground surface is intersected by a diagonal line representing a fault. The upper block above the fault is labeled "Fault". A vertical line extends downwards from the fault line, with the text "not risky" written next to it. A circle on the ground surface contains the text "50m". To the right of the circle, the text " $\gamma \geq 0$ " is written, with a small downward-pointing arrow below it.

```
lst = [o
      ('title', 'count'),
      ('year', 'min'),
      ('year', 'max')]
```

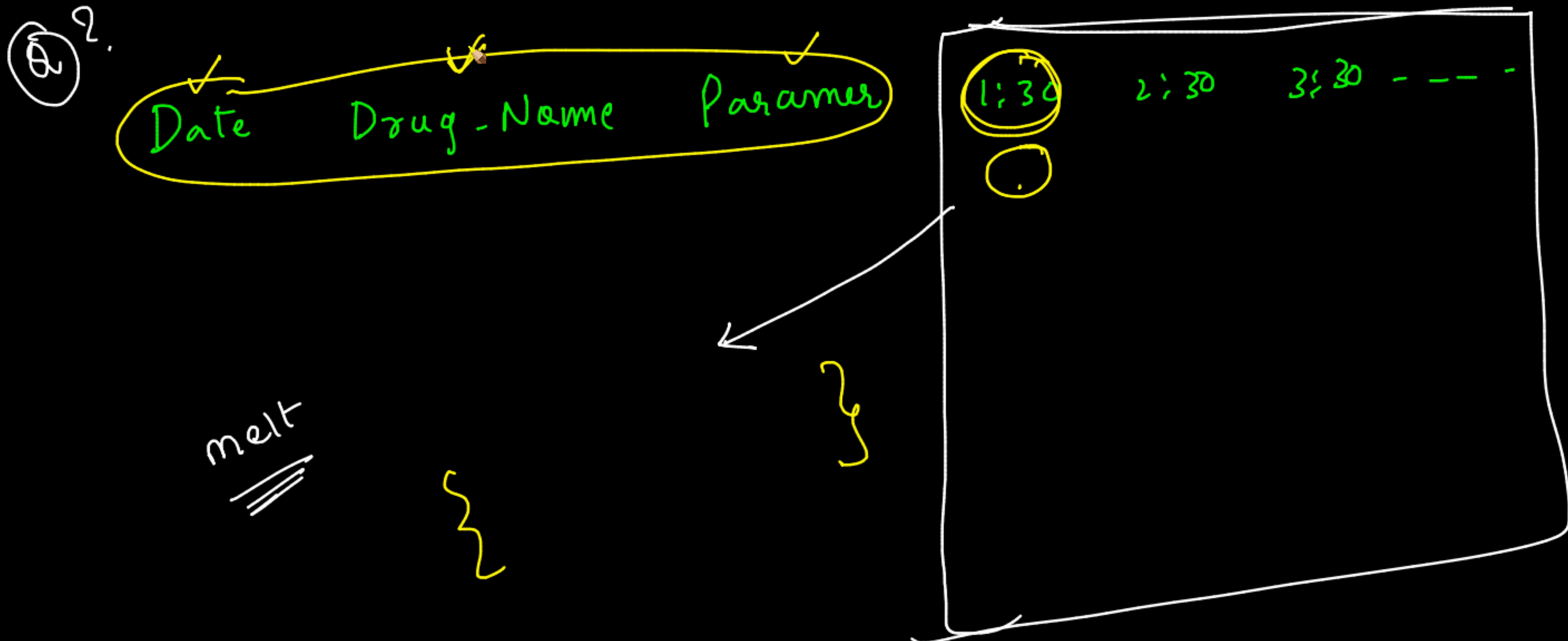
```
for i in list:  
    print(i)  
  
title, count = input("Enter title and count: ").split()  
print(f"\nTitle: {title}\nCount: {count}\n")  
  
list = [int(x) for x in range(1, int(count)+1)]  
print(f"\nList: {list}\n")  
  
min = min(list)  
max = max(list)  
avg = sum(list)/len(list)  
  
print(f"\nMin: {min}\nMax: {max}\nAvg: {avg}\n")
```

True





Name	Year	CO2 Emission
	1960	
	1961	



$$(18 \times 15) \\ (3 \quad 12) \\ 2$$

melt( )

$$18 \times 12 \Rightarrow ( \underline{\underline{216}}, 5 )$$

long → wide  
≡

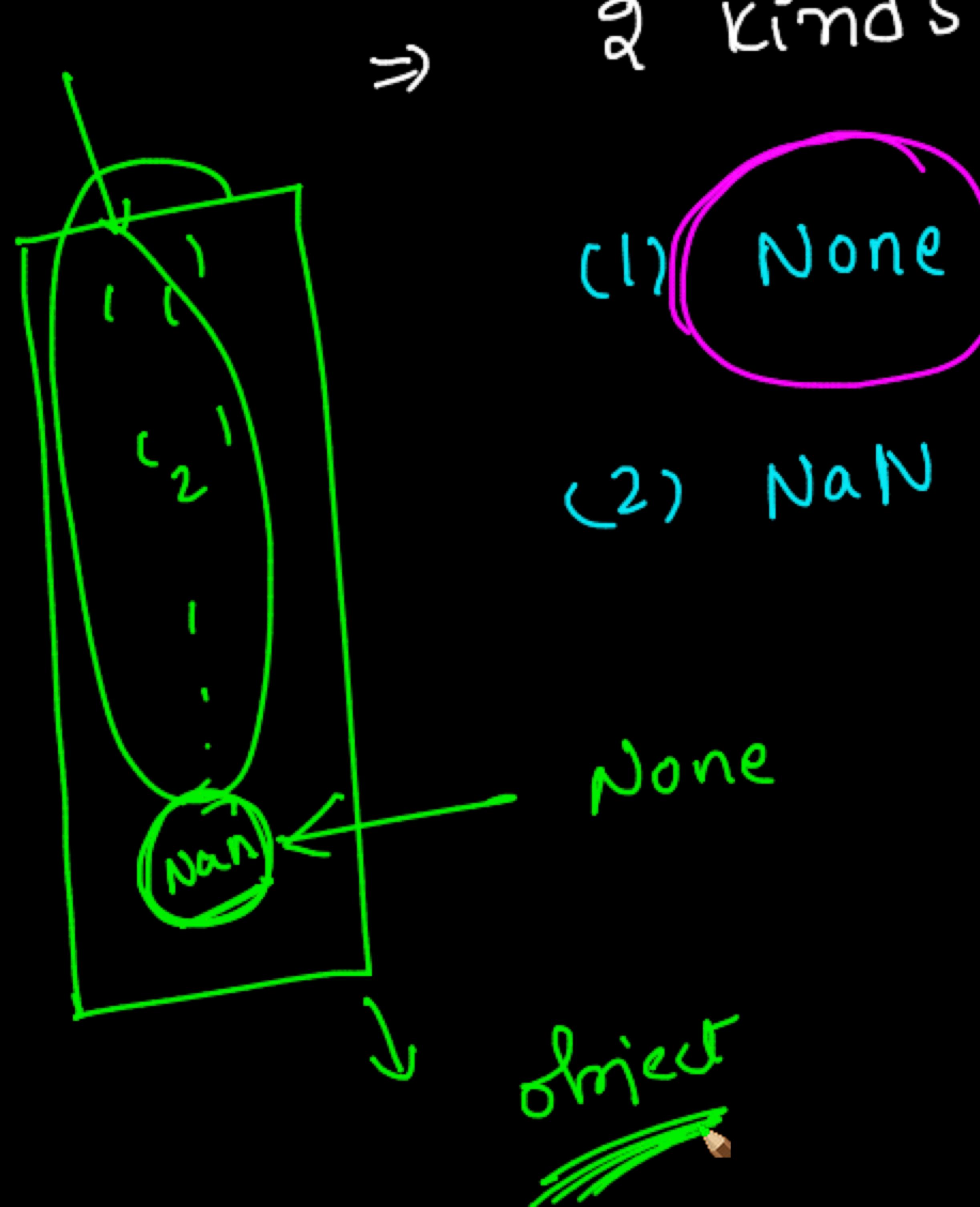
data - melt

Pivot()  
≡

216 ⇒ 108  
2

NaN → Missing values  
empty cell / no data

int, float

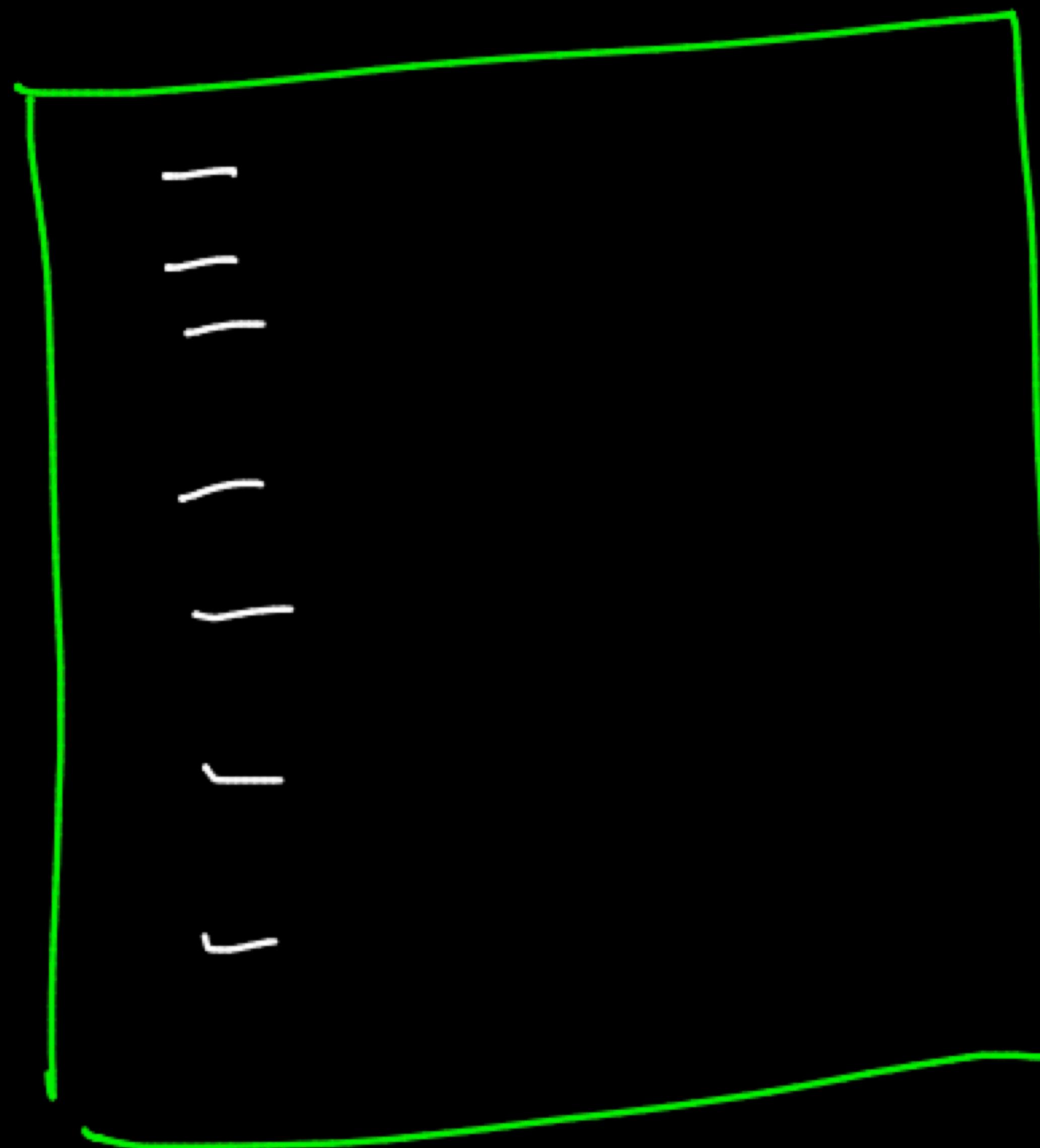


missing values

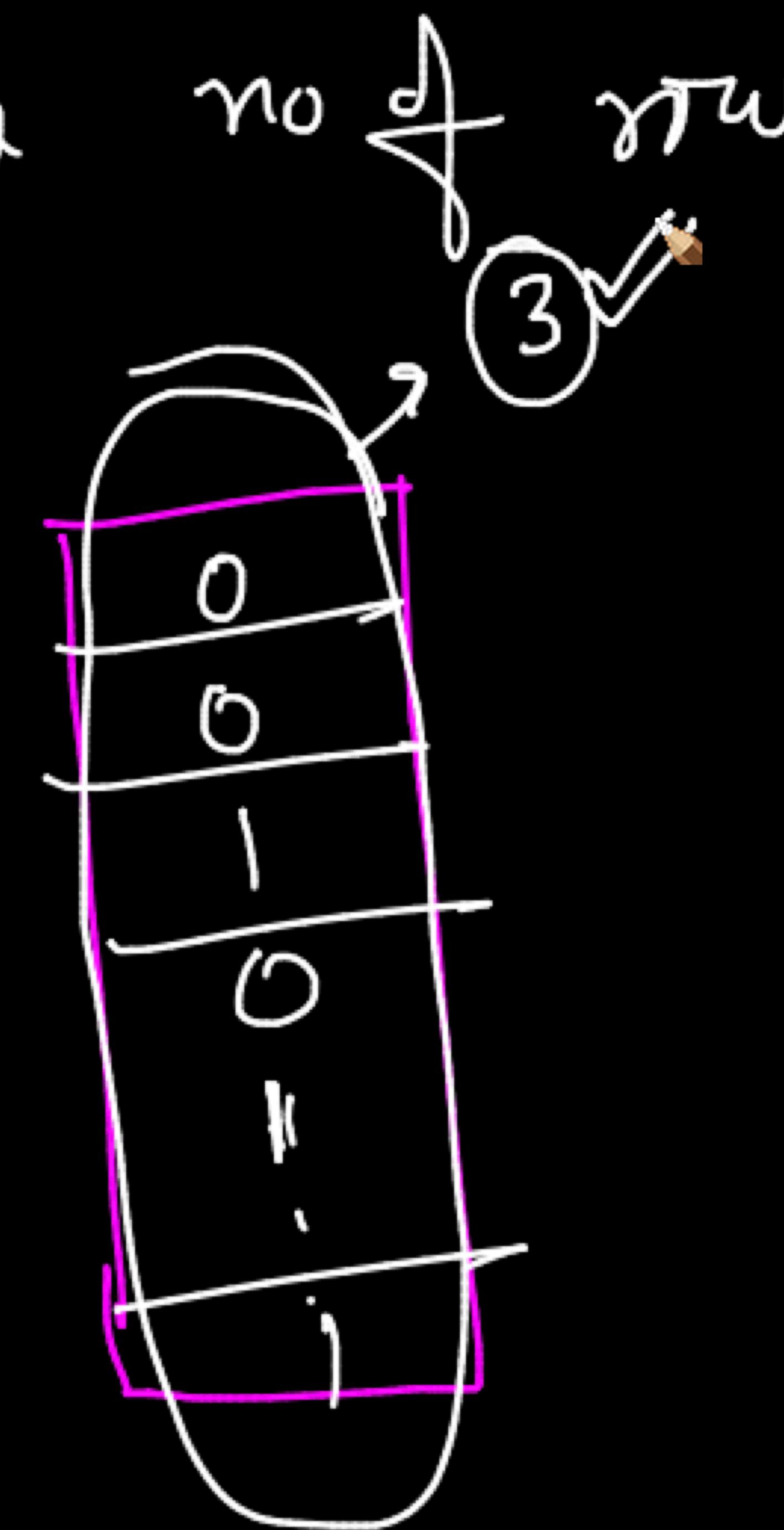
- ✓ None:
- ✓ NaN:

object: missing values in a column with non-number entries

columns with number entries



total      no ✓  
~~d~~      yrs



(1)

remove those rows / columns

(2)

just fill missing values.

NaN

any()

[ F, F, T, F, F, F ]

all()

