

21-06-2025

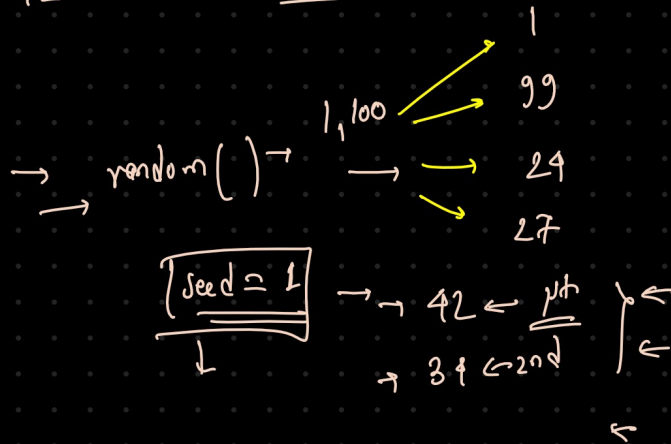
Agenda -

→ Data Visualization

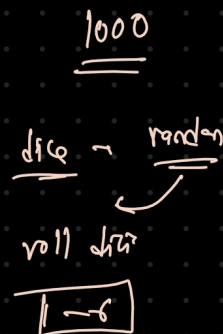
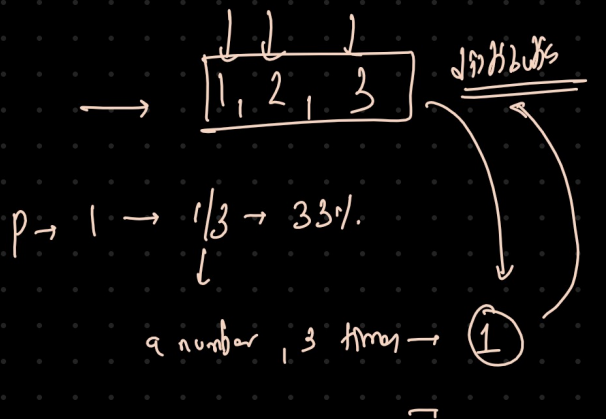
random\_state = 1, 42

seed = 42

random



→ seed 2)



product

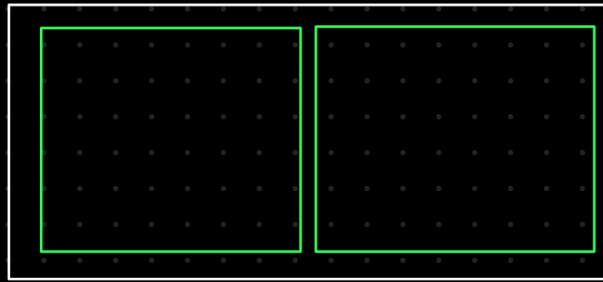
Quantity ↑ Revenue ↑ → X

→ 1\$, 1M

→ 3000

2 - plot ->

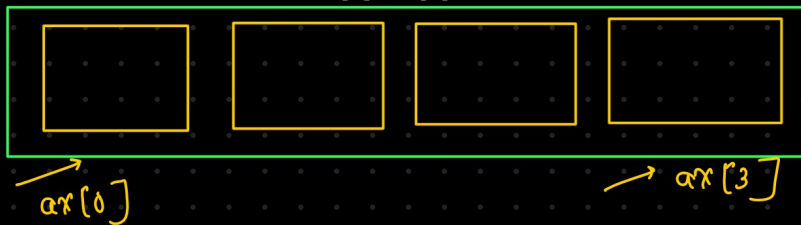
(1) row=1 col=2



(2) row=2 col=1



nrows=1, ncols=4  
ax[0]...ax[3]



- plt.figure → creates a canvas
- sns.histplot() → histogram, important active matplotlib fig
- plt,

matplotlib

numpy

Seaborn

Pandas

KDE  $\rightarrow$  kernel Density Estimation

1 — 100

number count

1  $\rightarrow$  30

2  $\rightarrow$  20

3  $\rightarrow$  5

4  $\rightarrow$  1

5  $\rightarrow$  1

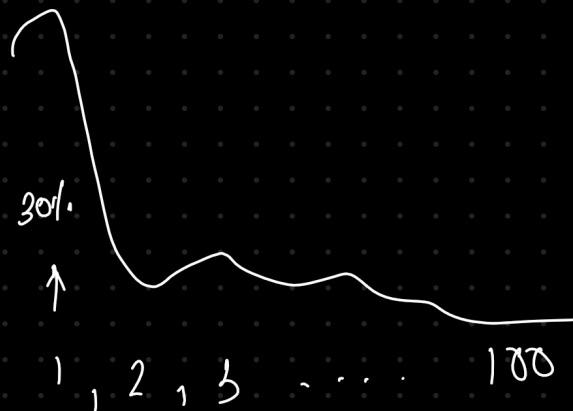
7  $\rightarrow$  1

8  $\rightarrow$  0

99  $\rightarrow$  1

100  $\rightarrow$  1

adds upto  
100



random  $\rightarrow$  stand

probability of getting Quantile(x)

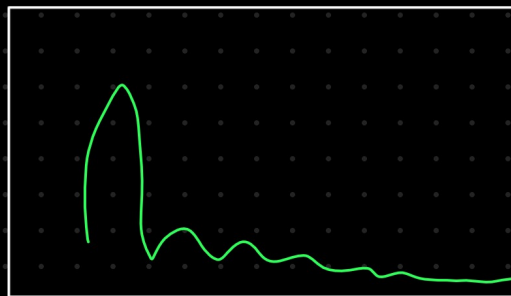
$x=0 \rightarrow p1$

$x=1 \rightarrow p2$


$x=2 \rightarrow p3$

$\vdots$

$x=100 \rightarrow p101$




histogram →  $X, y$   
 $\downarrow \quad \downarrow$   
 $C \quad \text{revenue}$

R   
 $C$

count plot →  $X$   
 $\downarrow$   
 $C \rightarrow \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array} \rightarrow f \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \end{array}$   
 $C$

Hospital  
 Analyg  
 shid

restaurant →  $\begin{array}{c} 10 \\ 1 \\ 40 \end{array}$



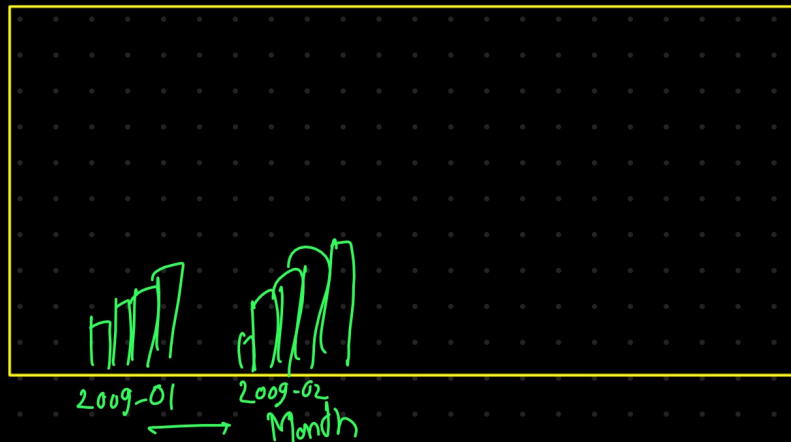
← count plot.

$\begin{array}{ccc} \boxed{C} & \boxed{M} & \boxed{R} \rightarrow \text{Sun} \\ \hline \text{A} & \text{Jan} & 20 \\ \hline \end{array}$

$\begin{array}{ccc} B & Feb & 10 \\ \hline \text{A} & \text{Jan} & 20 \\ \hline \end{array}$

$\boxed{A \quad \text{Jan} \quad 50}$

revenue

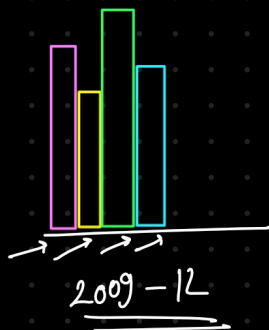


Tracking apps (health apps)

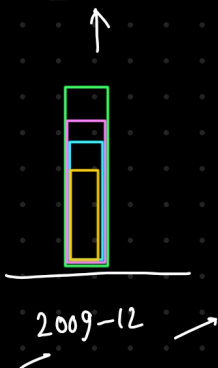
battery health

stop

nnnnn → nnn  
86n ←



→ grouped bar plot



→ stacked bar plot

Month	Country	Revenue
2020-01	UK	10000
2020-01	Germany	8000
2020-02	UK	11000
2020-02	France	9000

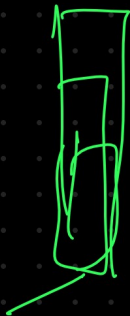
↓  
index → column → value

`pivot(index = 'month', columns = 'country', values = 'Revenue')`

Month	France	Germany	UK
2020-01	NAN	8000	10000
2020-02	9000	NAN	11000
2020-03	7000	NAN	NAN

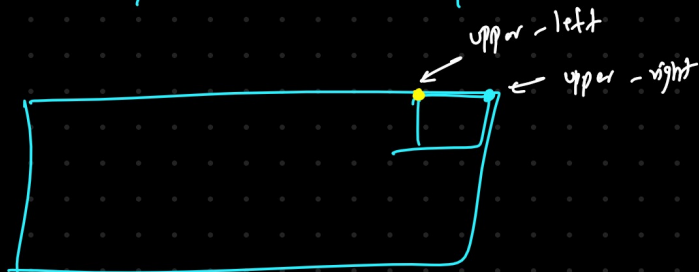
→  
`fillna(0)`

Month	France	Germany	UK
2020-01	0	→ 8000	→ 10000
2020-02	9000	0	11000
2020-03	7000	0	0



relative to axes

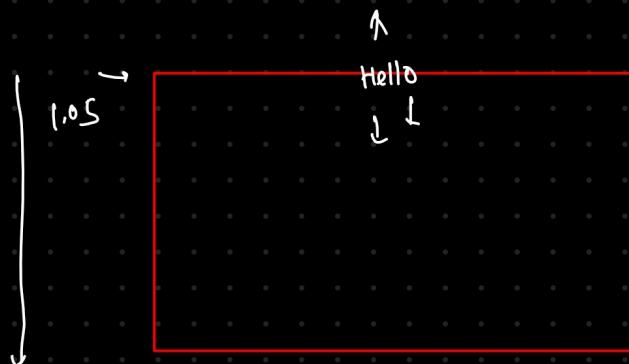
$x = 1.05 \rightarrow$  Just a bit right to the plot boundary  
 $y = 1 \rightarrow$  Align to the top of the axes



Quantity	Price	Revenue
(1)	(2)	(3)

$\rightarrow$

(1) - (2)	(2) - (3)	<del>(3) - (3)</del>
(1) - (3)	(2) - (1)	(3) - (2)
<del>(1) - (1)</del>	<del>(2) - (2)</del>	(3) - (1)



Height weight

150  $\rightarrow$   
 180  $\rightarrow$   
 190  $\rightarrow$   
 210  $\rightarrow$   
 250  $\rightarrow$



Height  $\rightarrow$  weight

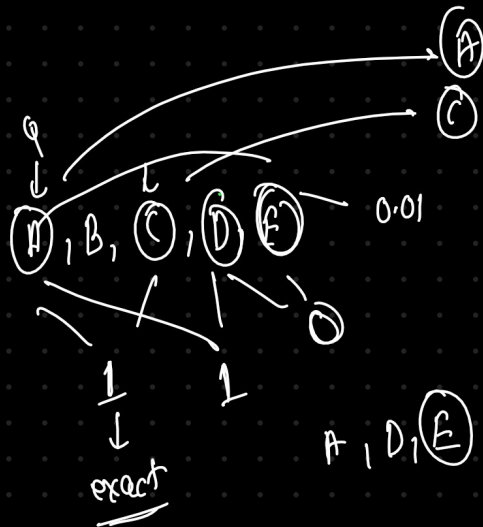
$\rightarrow$  +ve

Car value	<span style="border: 1px solid black; padding: 2px;">-ve</span>
Total - years	
<span style="border: 1px solid black; padding: 2px;">15L</span>	<span style="border: 1px solid black; padding: 2px;">5</span>
<span style="border: 1px solid black; padding: 2px;">13L</span>	10
<span style="border: 1px solid black; padding: 2px;">12L</span>	15



Height	Salary
120	200k
140	400k
160	100k
180	80k
200	10k

Height  $\Rightarrow 0 \Rightarrow$  Salary



ML

derive complex formulas