


Data Visualisation - 3

| Start at 9:03

Agenda

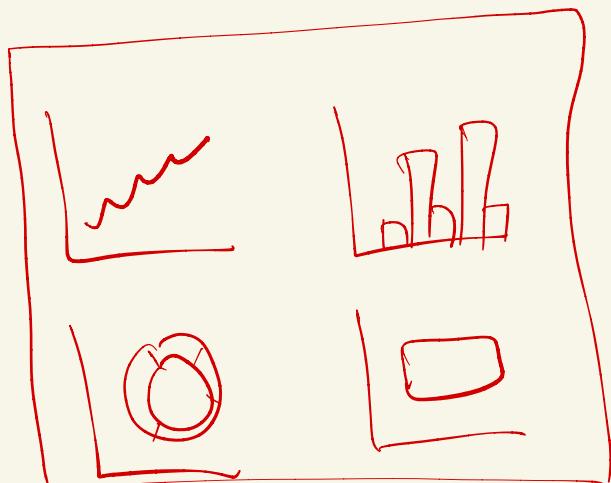
- 1) Revision (Bivariate)
- 2) Subplots
- 3) Multivariate
- 4) Advance charts
 - Joint plot
 - Pair plot
 - Heatmap

⇒ Bivariate Analysis

Review

- N - N
 - Line
 - Scatter
- C - C
 - Dodged bar
 - Stacked bar
- C - N
 - Boxplot
 - Bar chart

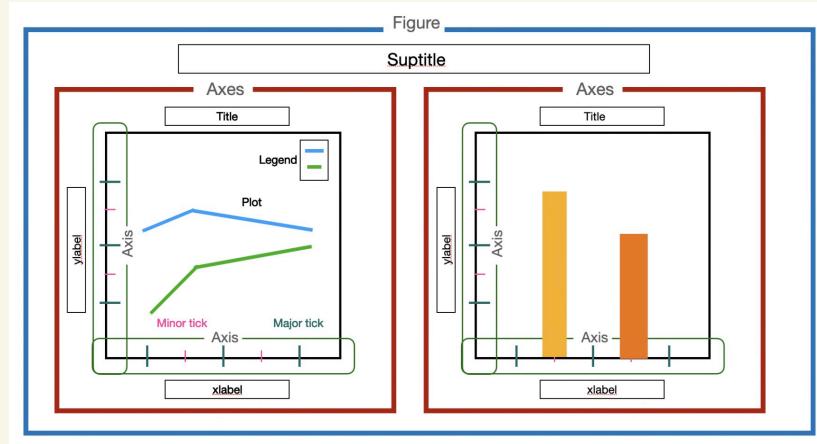
⇒



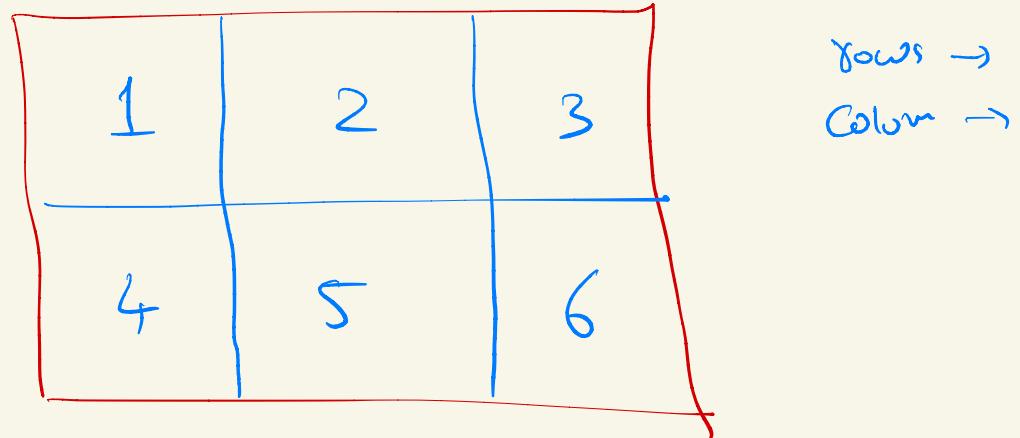
most insightful DataViz

more than 1 chart

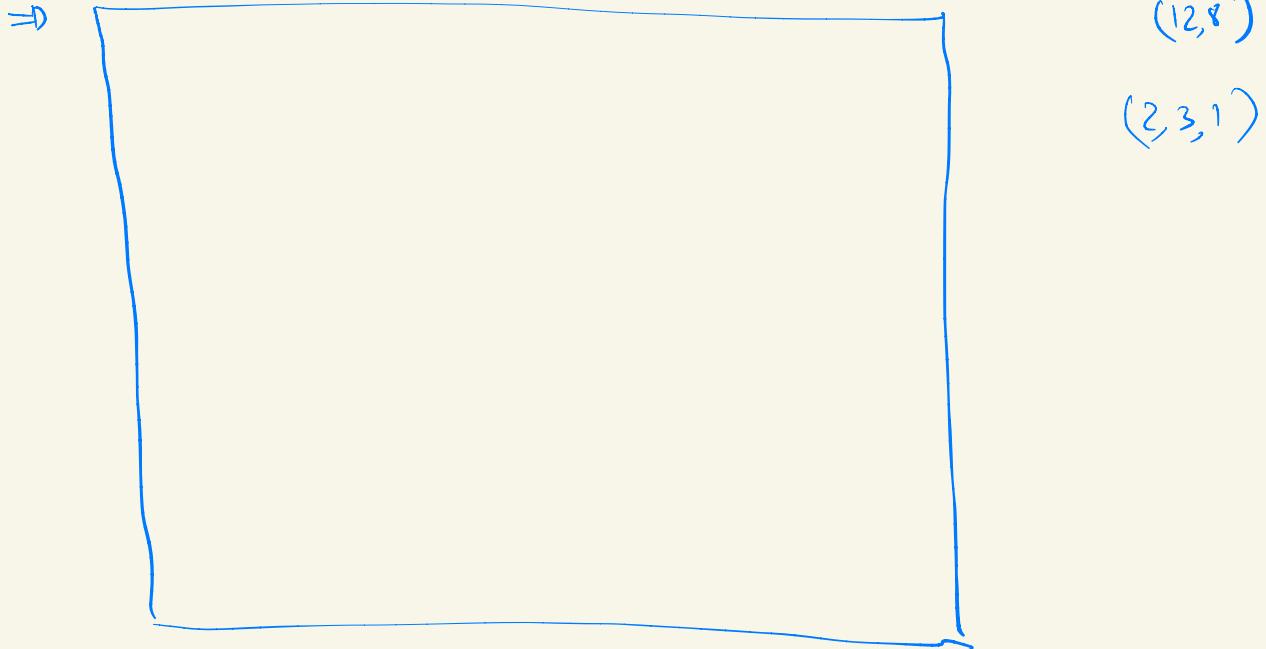
⇒



⇒



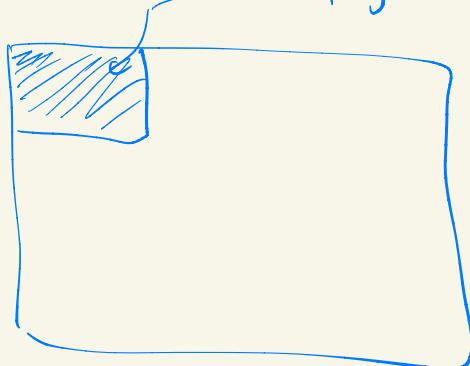
⇒ 6 Sub plot → → ↓



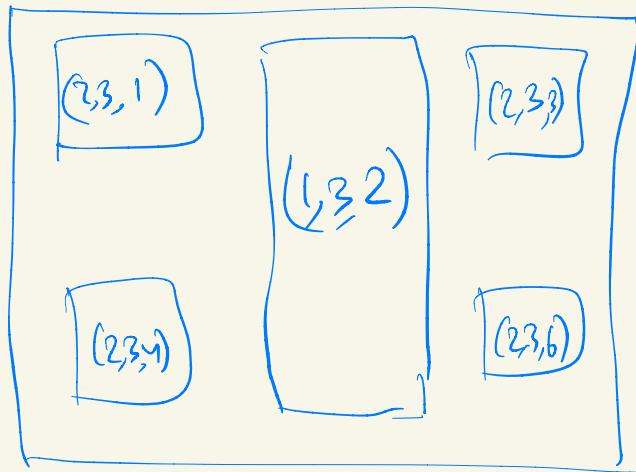
⇒ First Create an array reference for all subplots
 ↓
 populate it

- Create a figure → `fig`
- Subplot → all upcoming plots should be placed in a designated area

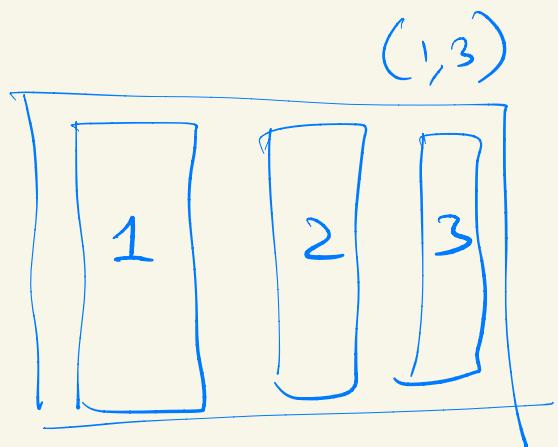
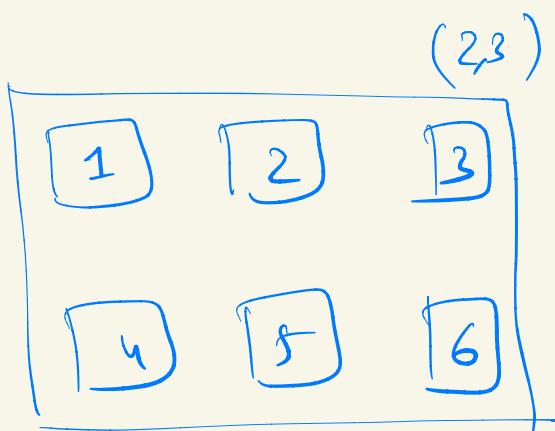
→ `Subplot(2,3,1)` → display the graph



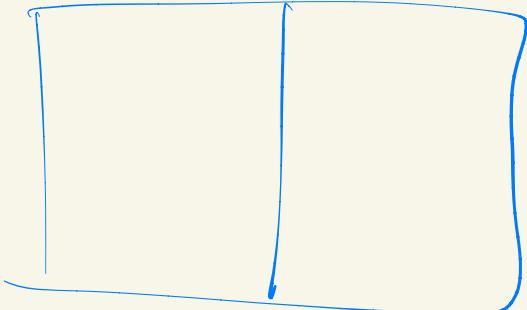
\Rightarrow



\Leftarrow a figure like this will look nice



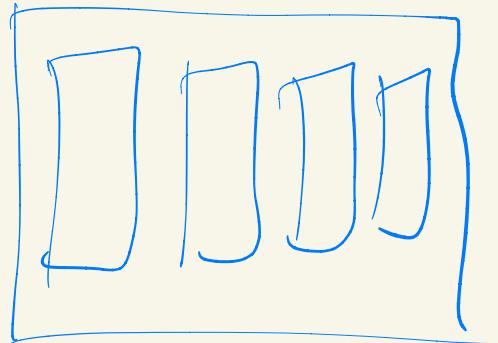
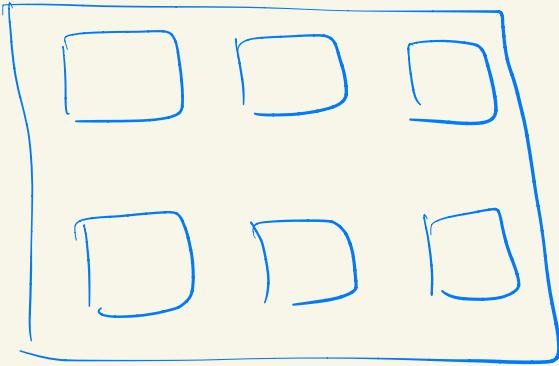
\Rightarrow



(1,2,3)

$\Rightarrow (2, 3,)$

$(1, 4, \sim)$



$\Rightarrow \underline{\text{Multivariate}} \rightarrow$

$\rightarrow \underline{3 \text{ variables}} \rightarrow$

$\rightarrow C - C - C \rightarrow \times \times \times$

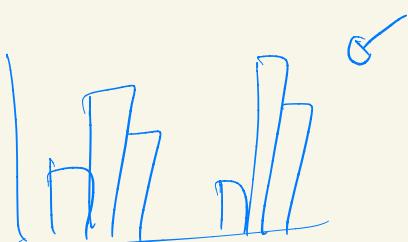
$\rightarrow N - N - N$

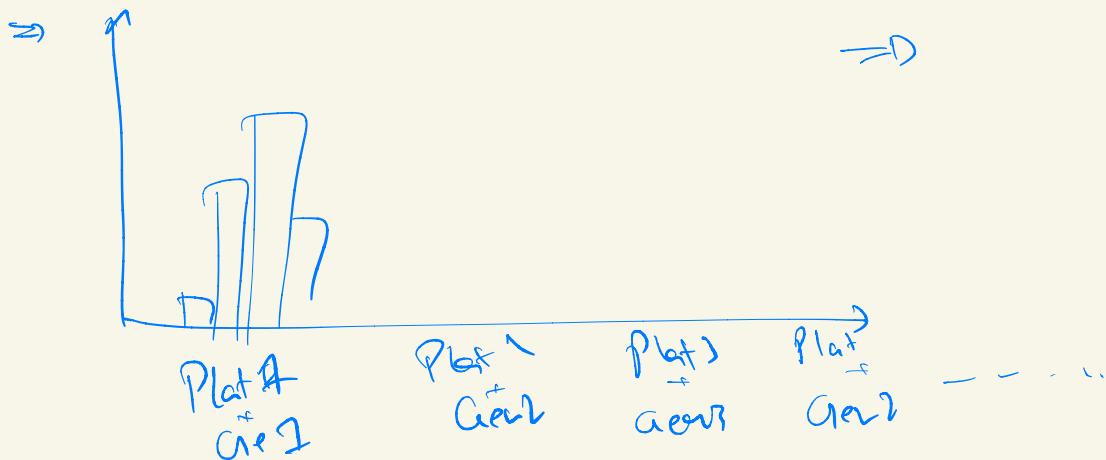
$\rightarrow C - C - N$

$\rightarrow C - N - N$

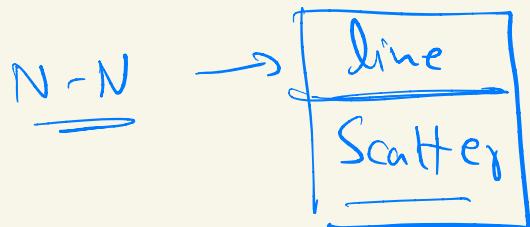
$\Rightarrow \underline{C-C} \rightarrow \text{dodged count},$

← more
Categorical
variables





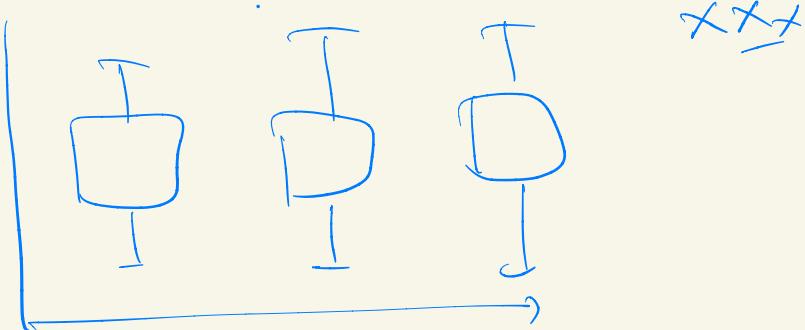
⇒ N-N-C



← add a
Categorical informa
→ Color

N-C → multi box
→ multi bar

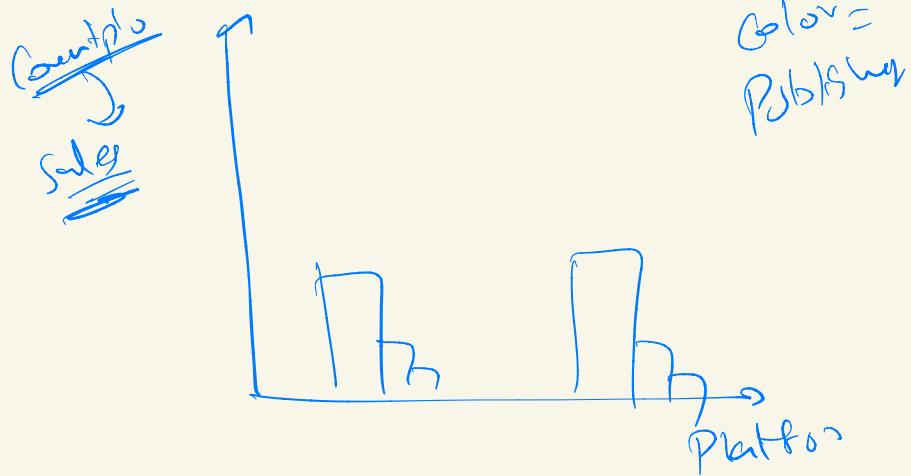
← a numerical informat



$\Rightarrow \underline{C-C-N} \rightarrow$

$\Rightarrow \underline{G-C} \rightarrow$ Stacked
Dodged

numerical value/info
to this

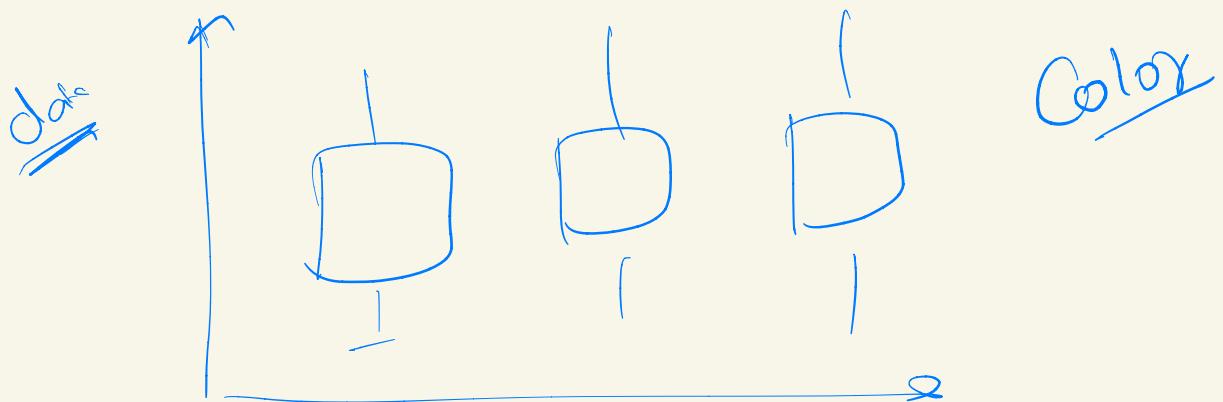


Color =
Publisher

\Rightarrow Sales vs Category, customer profile

\Rightarrow

$\Rightarrow \underline{G-N}$ $\xrightarrow{\text{Box}}$ add some
 $\xrightarrow{\text{multi box}}$ Cat. info

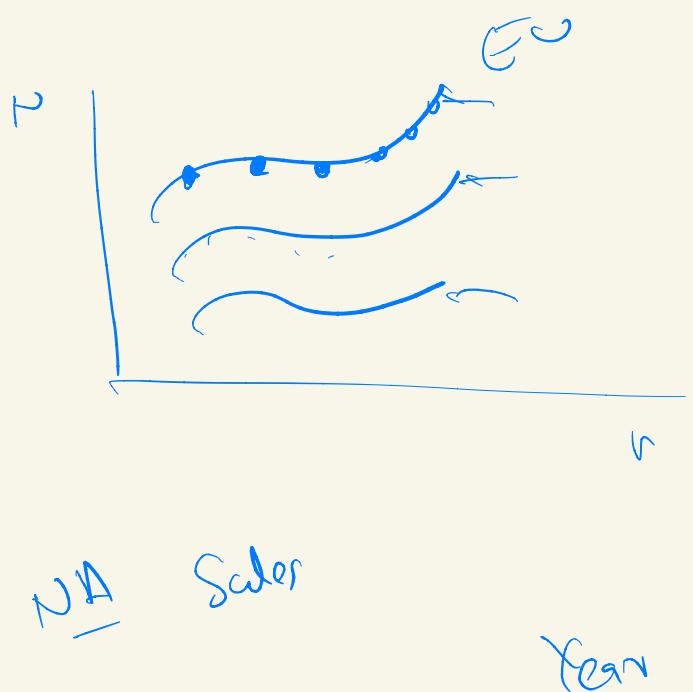
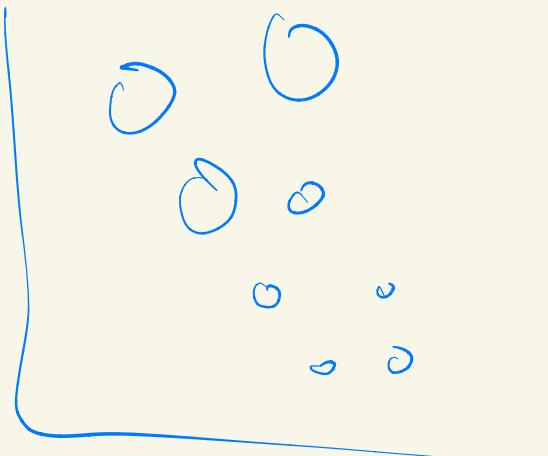


$\Rightarrow \underline{N-N-N} \rightarrow$

$N-N \rightarrow$ Scatter
line

— we want to
add numerical
info

$\rightarrow \underline{\text{Size}} \rightarrow \text{Scatter} \rightarrow S_i$



\downarrow

$\Rightarrow N_1 - N_2$

$N_2 - N_3$ $N_1 - N_3$

$N_1 - N_2 - N_3$

$\Rightarrow C - C - N \rightarrow \star \star C$

(Bivariate) \rightarrow

$= D \cdot \begin{cases} C - C - N \\ N N N \\ C - N - N \end{cases} \rightarrow \begin{cases} \text{colorbar, multi bar} \\ \text{Scatter with size} \\ \text{Scatter with Gbo} \\ \text{line with Gbo} \end{cases}$

\Rightarrow Joint Plot \rightarrow (N)

\rightarrow Scatter + histogram
 $(N-N)$

Scatter + Density $\rightarrow (N)$
 $(N-N)$

\Rightarrow 2 Num Variable



Pair plot $\rightarrow K$ N_i

K numerical variables ;

every single pair of them

N_1 n_2

n_2 n_3

n_1 n_3

n_1 n_4

→ Multivariate

→ subplots ↗

→ Joint plot

Pair plot

Heatmap