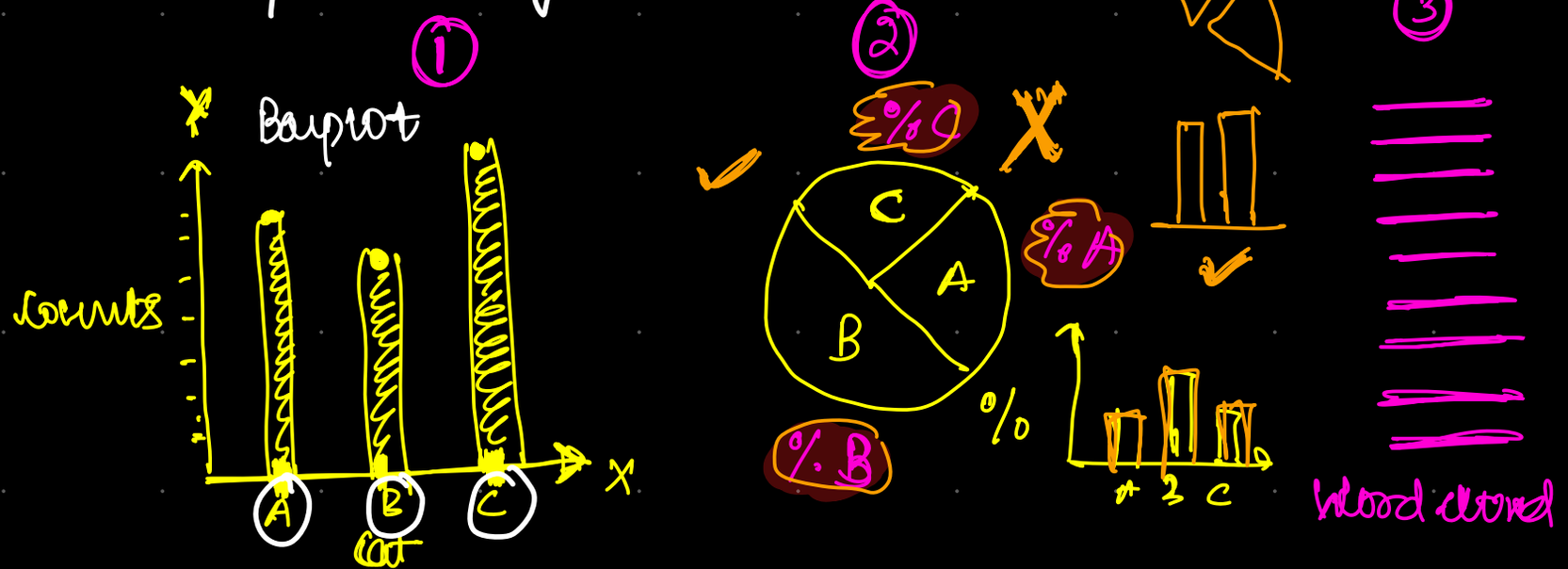
Data Visualization -2

How to choose the right plot ?

Univariate, Categorical.

- Distribution/frequency of diff categories – 1
- Percentage share/proportions of categories - 2
- Unique categories. - 3

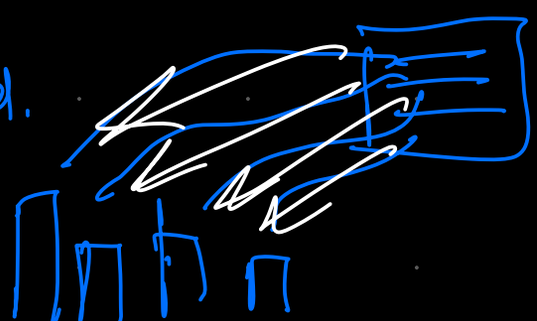


Overlapping x-ticks :-

1. ↑ the figure size. ✓
2. Change the angle of x-ticks - rotate ✓
3. Make the font smaller. ✗
4. Use the legend. ✗

→ long legend.

→ Colors.



cognitive load.

Univariate, Numerical (cont.)

- Distribution of different values.
- Skewed / outliers.
- Aggregates/Statistics - mean, med, max, min, %tiles.
- How much % data is below/above a certain values?

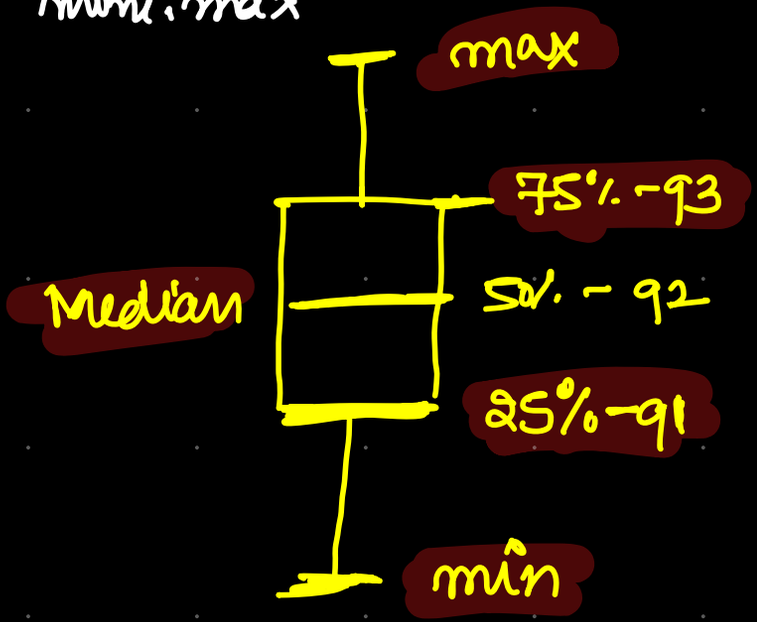
① Distribution/skewed

Histogram

A, B, C, D, E | A' B' C' D' E'

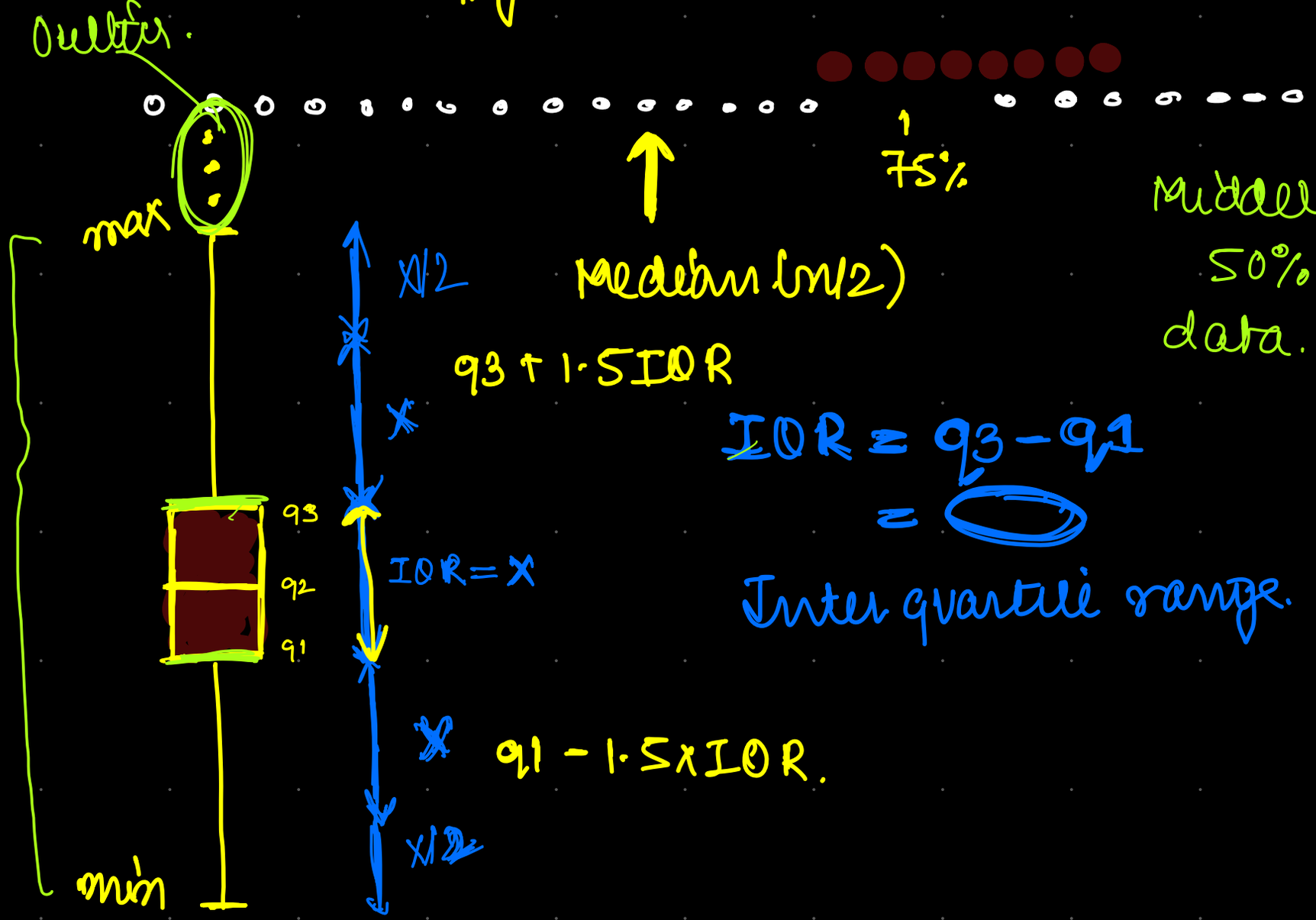
② Statistics.

median/mean, quartiles. min, max



smoothened version of histogram. Density Plot

How to identify outliers ?



Univariate

Categorical - Barchart, Piechart

Numerical - Histogram, KDE, Boxplot

Boxplot → Box-and whisker plot
→ Violin plot :- box, KDE plot
→ Bee-swarm plot.

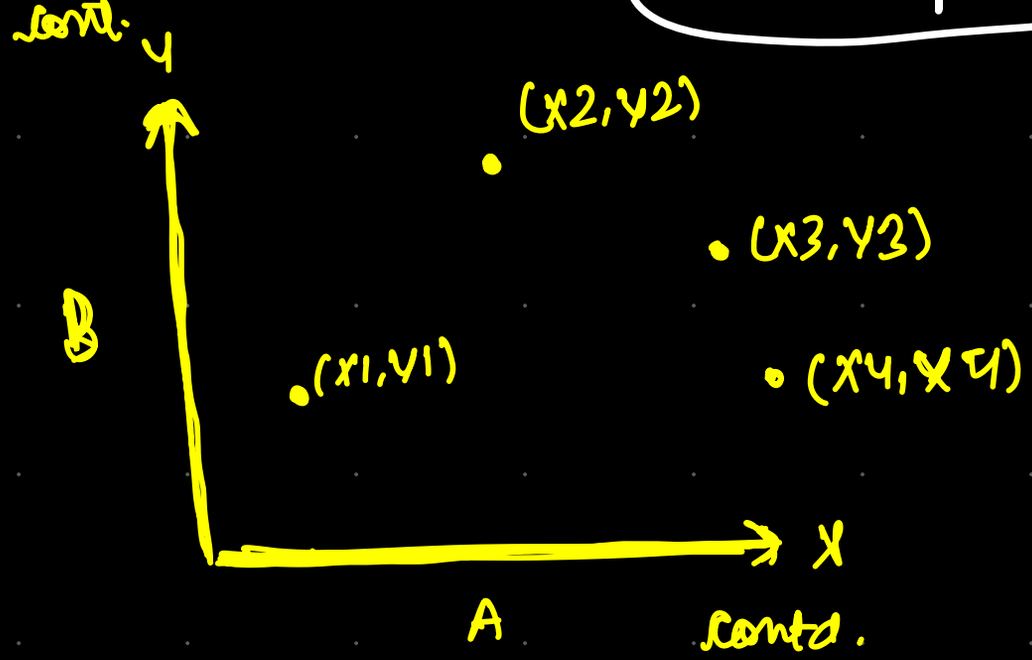
Bivariate Data Visualisation.

1. N - N ✓
2. C - C ✓
3. N - C

Bivariate, Num, Num

- Relationship / association.
- Correlations

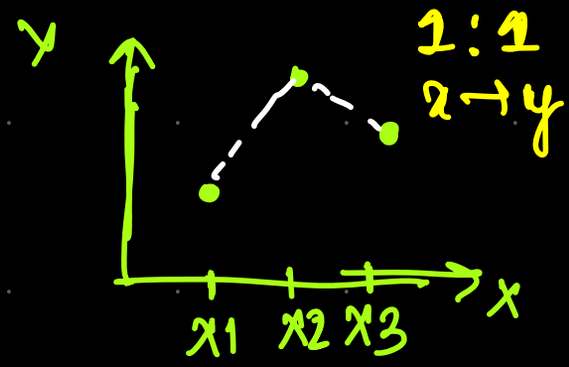
scatterplot

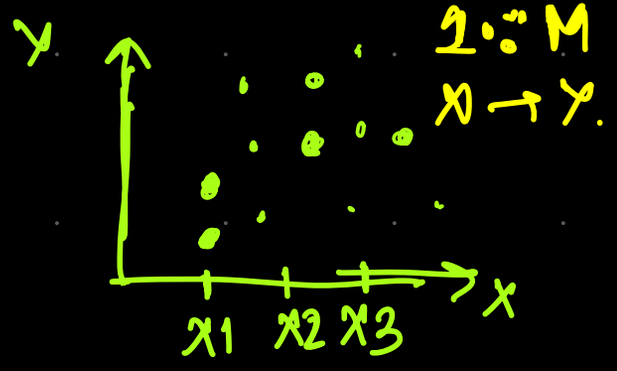


Num, Num

Scatter Lineplot

?? ??





Bivariate, Num. Num

1 : 1
$X \rightarrow Y$ → wheplot

1 : many
$X \leftrightarrow Y$ → scatterplot.