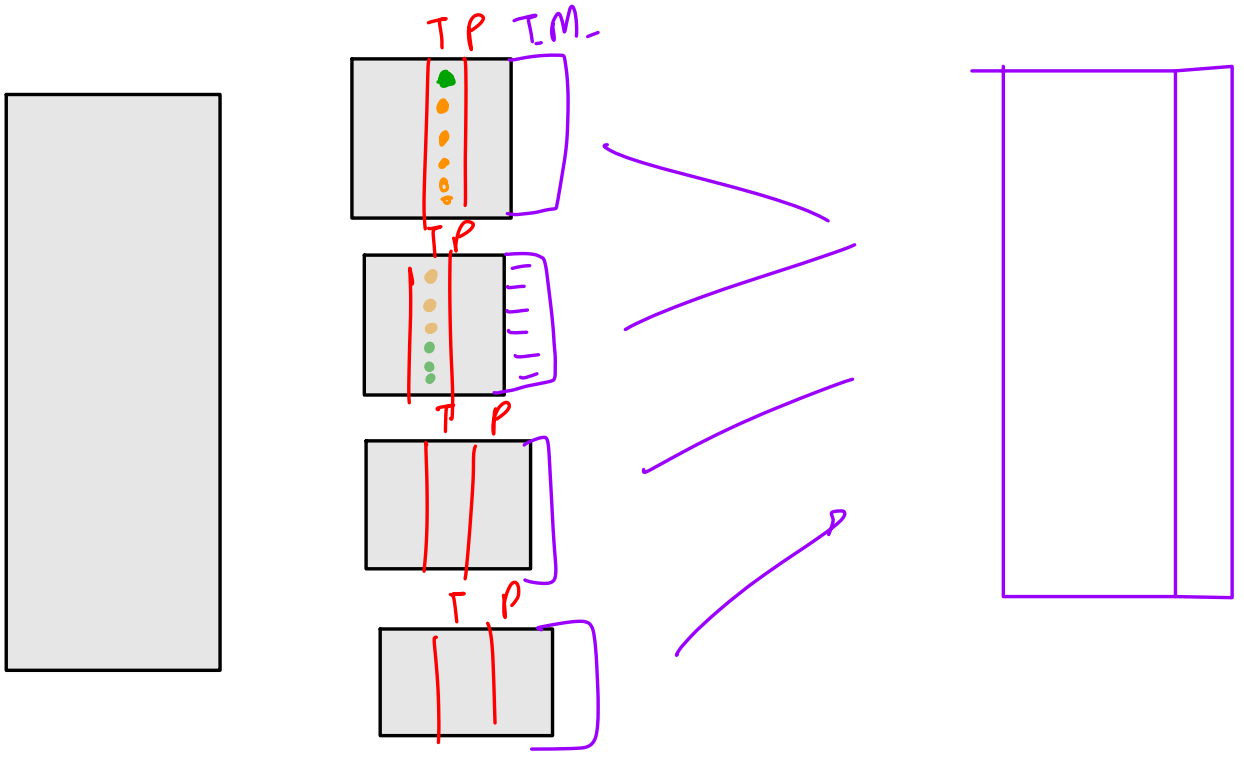
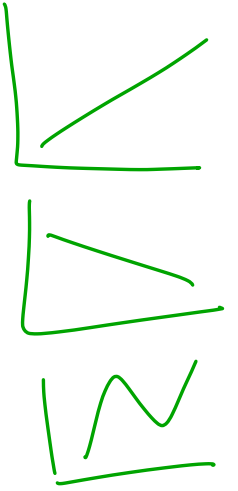

T P T.M.
T P
T P
T P

DATA VISUALISATION

- Exploratory → EDA → Exploratory data Analysis
- Explanatory → Showcase Insights.

DV
- Art → how?
- Science → why?



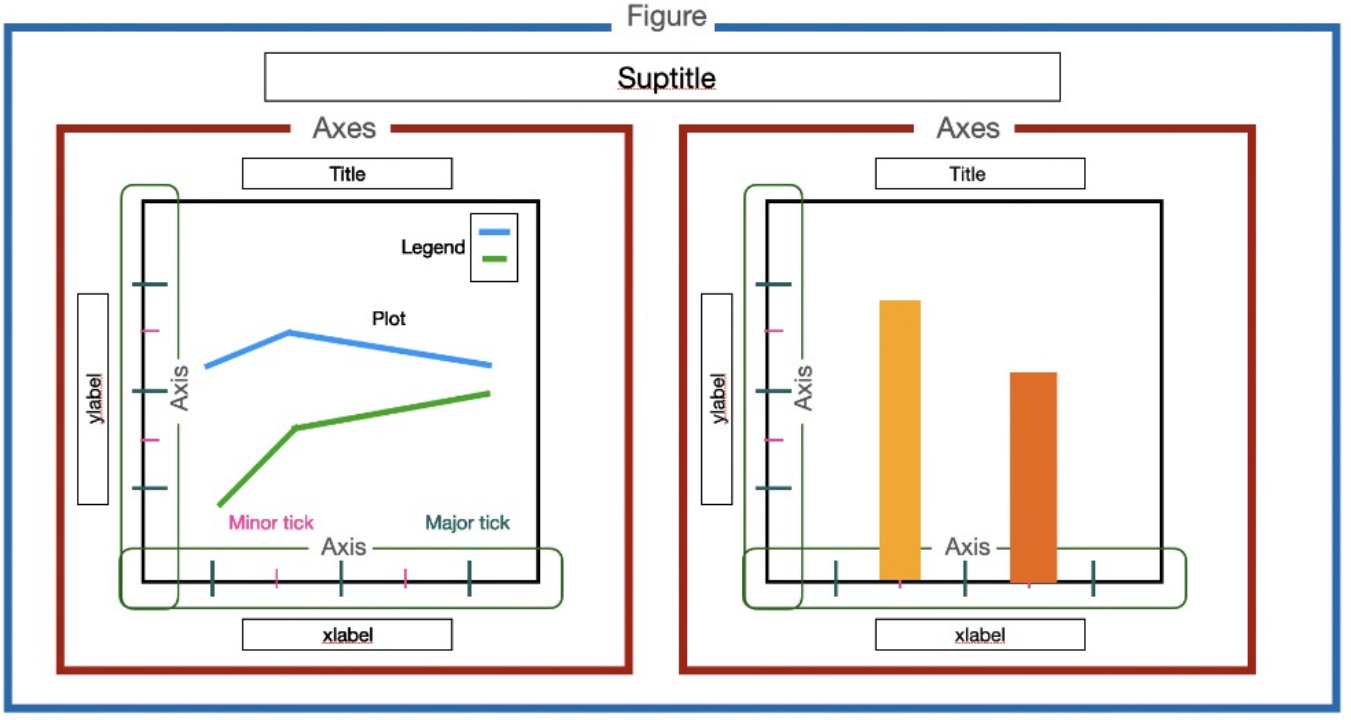


Figure
Suptitle
Axes
Title
Legend
Plot
ylabel
Axis
Minor tick
Major tick
Axis
xlabel
Axes
Title
ylabel
Axis
Axis
xlabel

* Univariate
  - C - Barplot with plt, Countplot (sns)
  - N - Histogram, KDE plot, BOXPLOT, VIOLINPLOT
* Bivariate
  - CC -
  - CN - SNS. BARPLOT
  - NN -
* Multivariate
  - NNN -
  - CNN - Scatterplot mw.
  - CCN -
  - CCC -

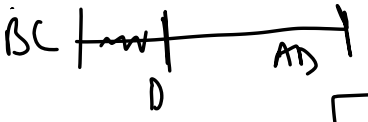
NOIR

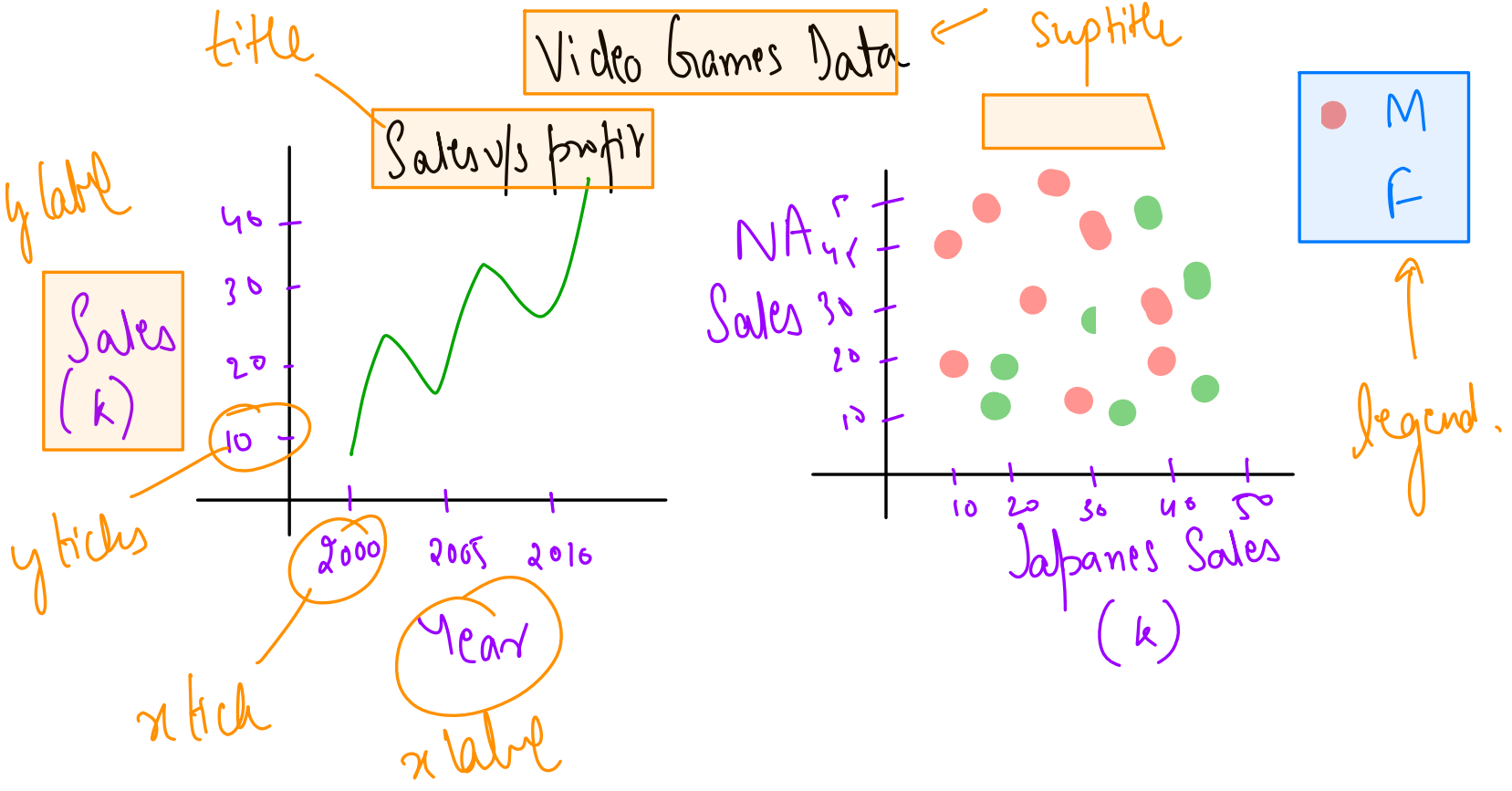
DATA

DISCRETE → 1, 2, 3, 4, 5

CONTINUOUS → 1.23, 1.94, 5.96, 5.45

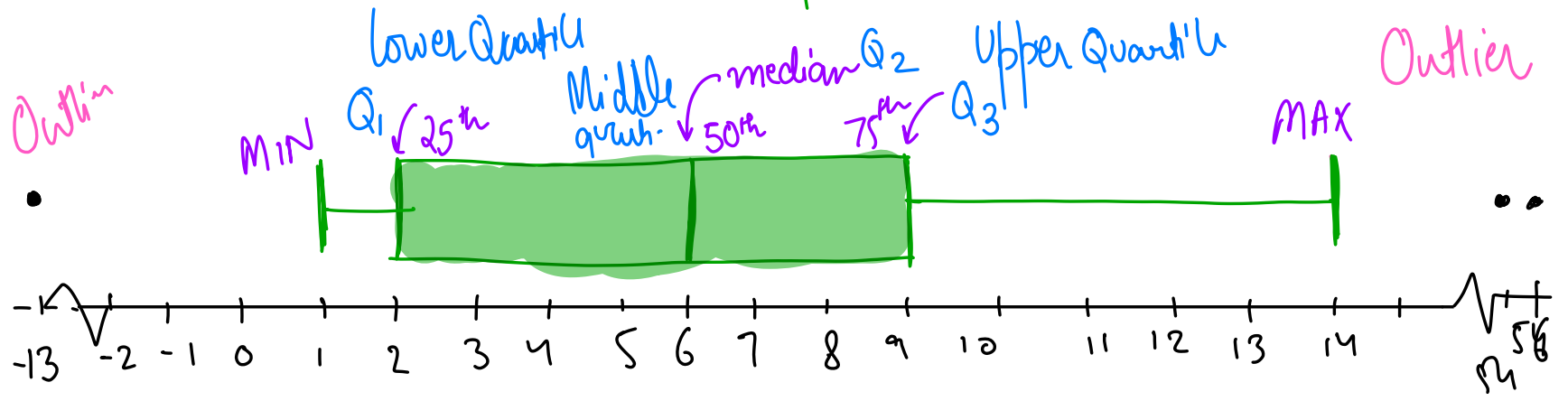
- CATEGORICAL
  - NOMINAL
    - { Electronic, fashion, Homedecor, Amazon pay } Books, Car Accessories }
  - ORDINAL
    - { HIGH, MEDIUM, LOW }
    - { UPPER MIDDLE LOWER }
- CONTINUOUS
  - Interval
    - Temp
    - Time
    - 0°C
  - Ratio
    - Weight
    - Height
    - 0cm
    - 0kg
    - 0 apples




title
Video Games Data
Sales v/s profit
subtitle
y label
Sales (k)
40
30
20
10
y ticks
2000
2005
2016
x tick
year
x label
NA Sales
40
30
20
10
10
20
30
40
50
Japanes Sales (k)
M
F
legend

DATA: 14, 51, 1, 1, 2, 6, 56, 7, 8, 9, 3, 4, -13, 6

Order/Sort: -13, 1, 1, 2, 3, 4, 6, 6, 7, 8, 9, 14, 51, 56



IQR → Inter Quartile Range ⇒ $Q_3 - Q_1$ ⇒ 9 - 2 ⇒ 7

MAX → Q3 + IQR(1.5) ⇒ 9 + (1.5)7 ⇒ 19.5

MIN → Q1 - IQR(1.5) ⇒ 2 - (1.5)7 ⇒ -8.5

Outlier → Uber

Websites 2 sec

Movies 80

2, 80, 3, 120, 18, 1, 3

R_1 R_2 R_3 R_4 R_5 R_6

= 2, 3, 4, 10, 12, 200

$$\frac{2+3+4+10+12+200}{6} = 38.5$$

podcasts

Clipping
Capping
% per
imputed
(DBSCAN)

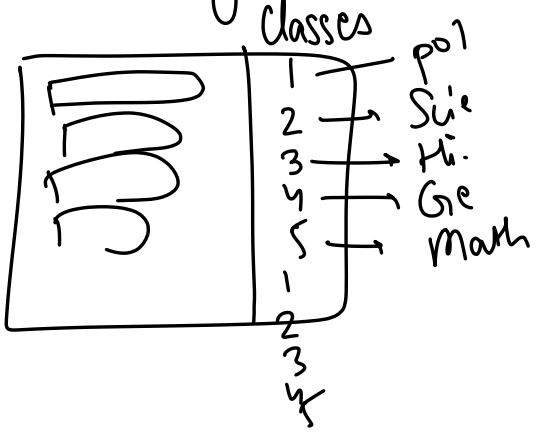
→ 2, 3, 3.5, 4, 5, 7, 60, 70, 120, 80, 65

Spotify → podcast

Year 1980 - - - - - - 2020

Numerical

→ histogram



Categorical

1980
1981
1982