

4) Shape Measure (Histogram, Violin, Box)

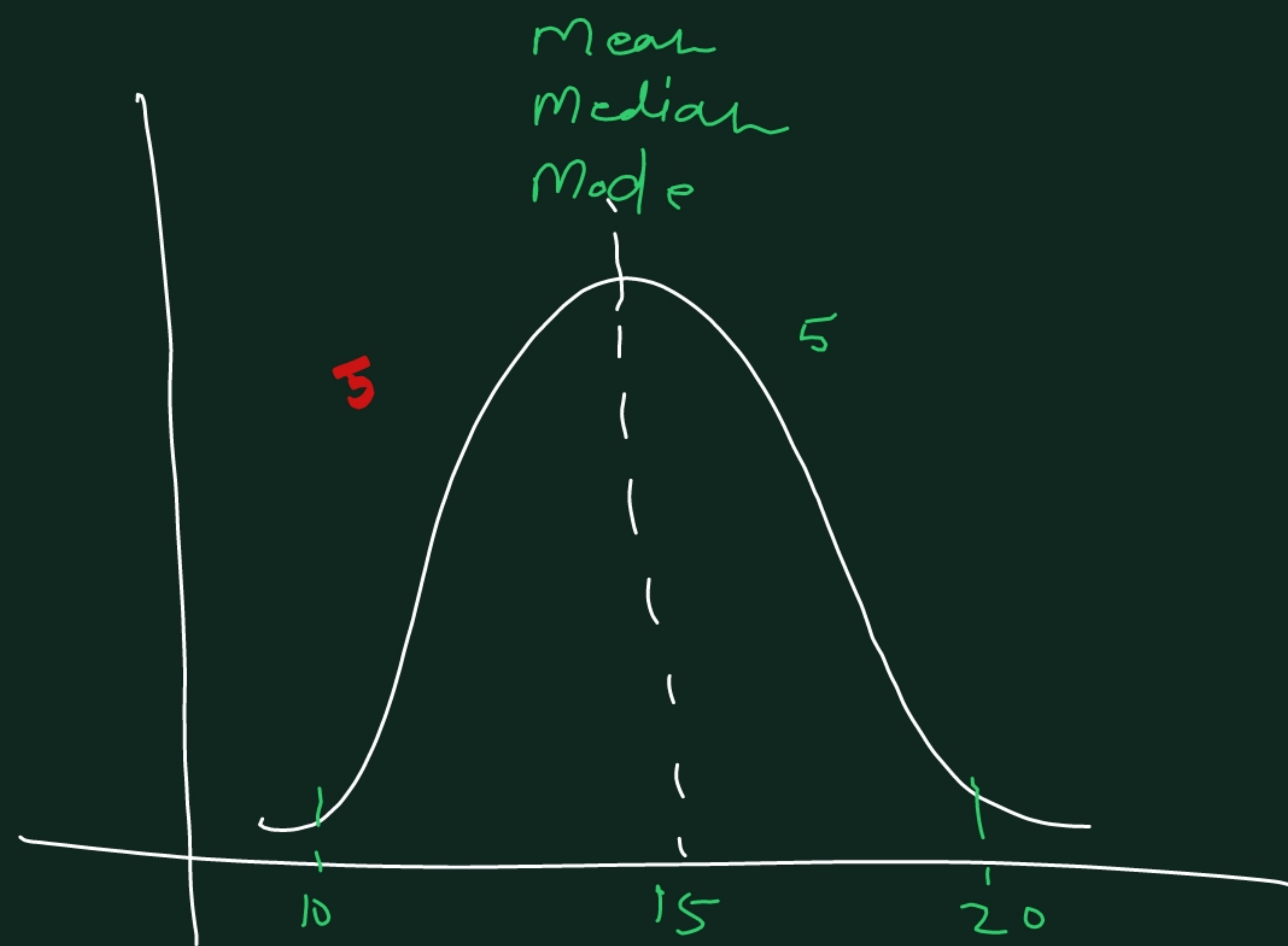
① Skewness → Data Direction

Positive

Negative

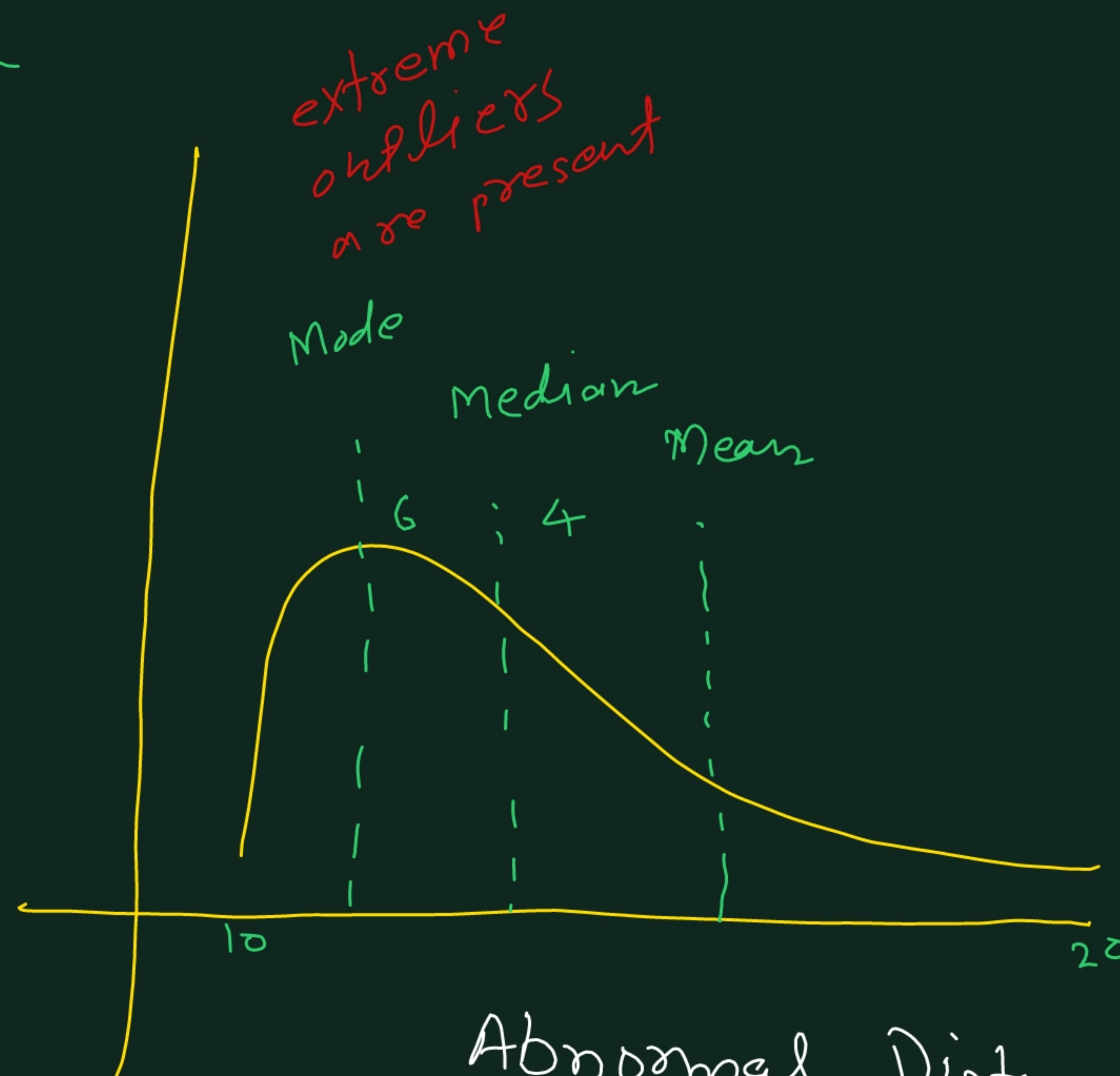
② Kurtosis → Strength of tail

①



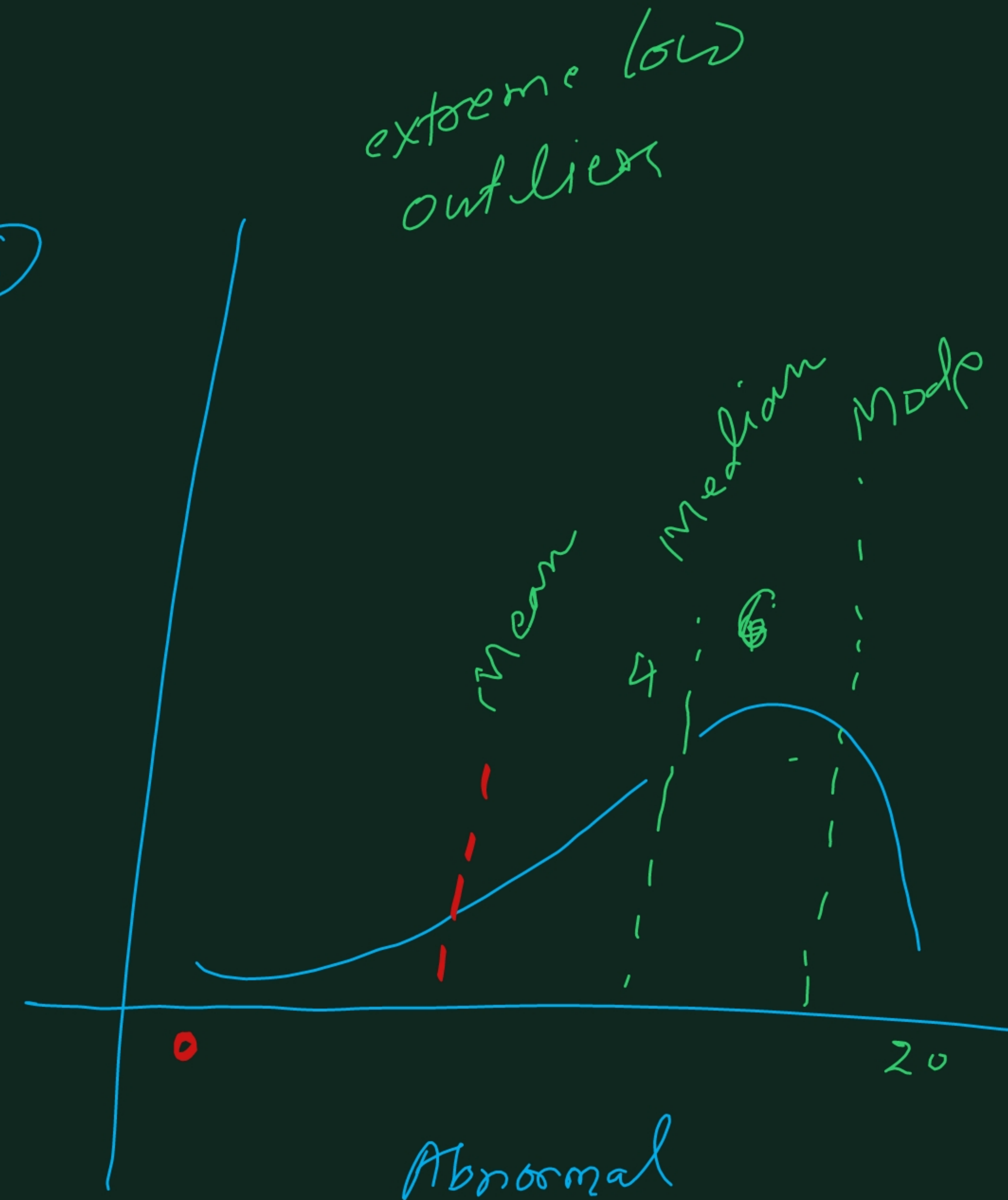
Normal Distribution
(Not Skewed)
(No Outliers)

②



Abnormal Dist
Right/+ve Skewed
Contains Outliers

③



Abnormal
Left/-ve skewed
Contains Outliers

5) Relationship Measure

1) Covariance \rightarrow Direction

2) Correlation \rightarrow Strength

Positive
Negative

Positive Correlation: Relationship: $B \uparrow G \uparrow$

Negative Correlation: Breakup: $B \downarrow G \uparrow$

Neutral Correlation: Freindzone: No effect

irrelevant
Column

Sales & Profit, Price

$$\text{corr} = 0.78$$

$$= -0.10$$

$$0.10$$

$$(0.001) \times$$

Range:

$$[-1, +1]$$

Problem → Solution

Process [Lifecycle]

① Problem Understanding

② Data Collection

③ Data Preprocessing

④ Data Analysis

⑤ Data Visualization

⑥ Optimization

⑦ Solution

70%

EDA

→ Understandin of data
4

Answering business queries.

Univariate, Bivariate, Multivariate