

② Handle Duplicates → ① duplicated ② drop_duplicates()

③ Missing Values → isnull(), fillna(), dropna()

④ Outliers → filter/condition, replace()

→ use central tendency measure to fill or replace.

Descriptive Stats

① Aggregate Measure

sum(), max(),
mean(), count()

② Central Tendency Measure

Mean(), Median(), Mode()

③ Dispersion Measure

range() $\left\{ \begin{array}{l} \text{max} \\ \text{min} \end{array} \right.$
IQR $\left\{ \begin{array}{l} Q_3 \\ Q_1 \end{array} \right.$
std()

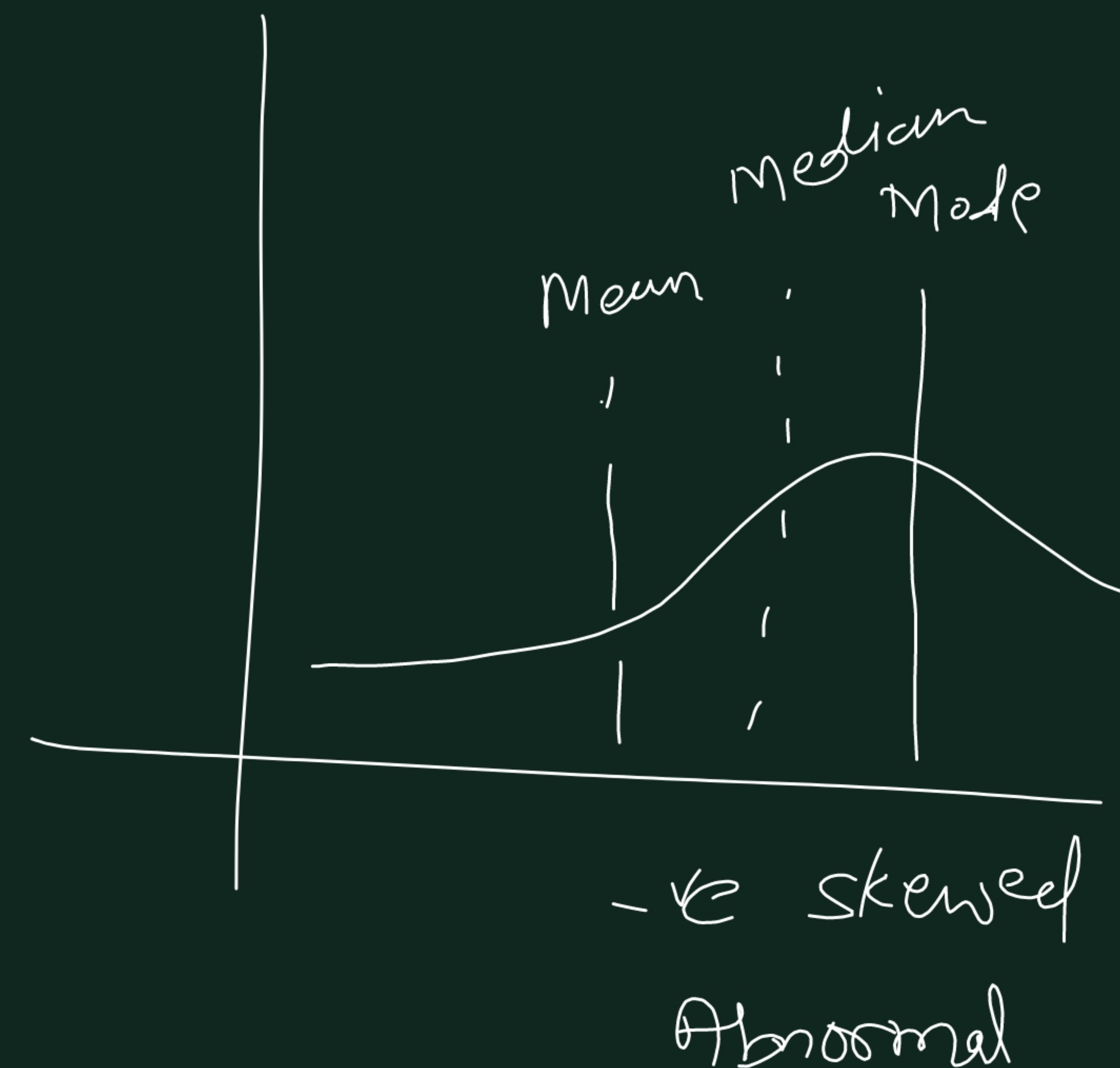
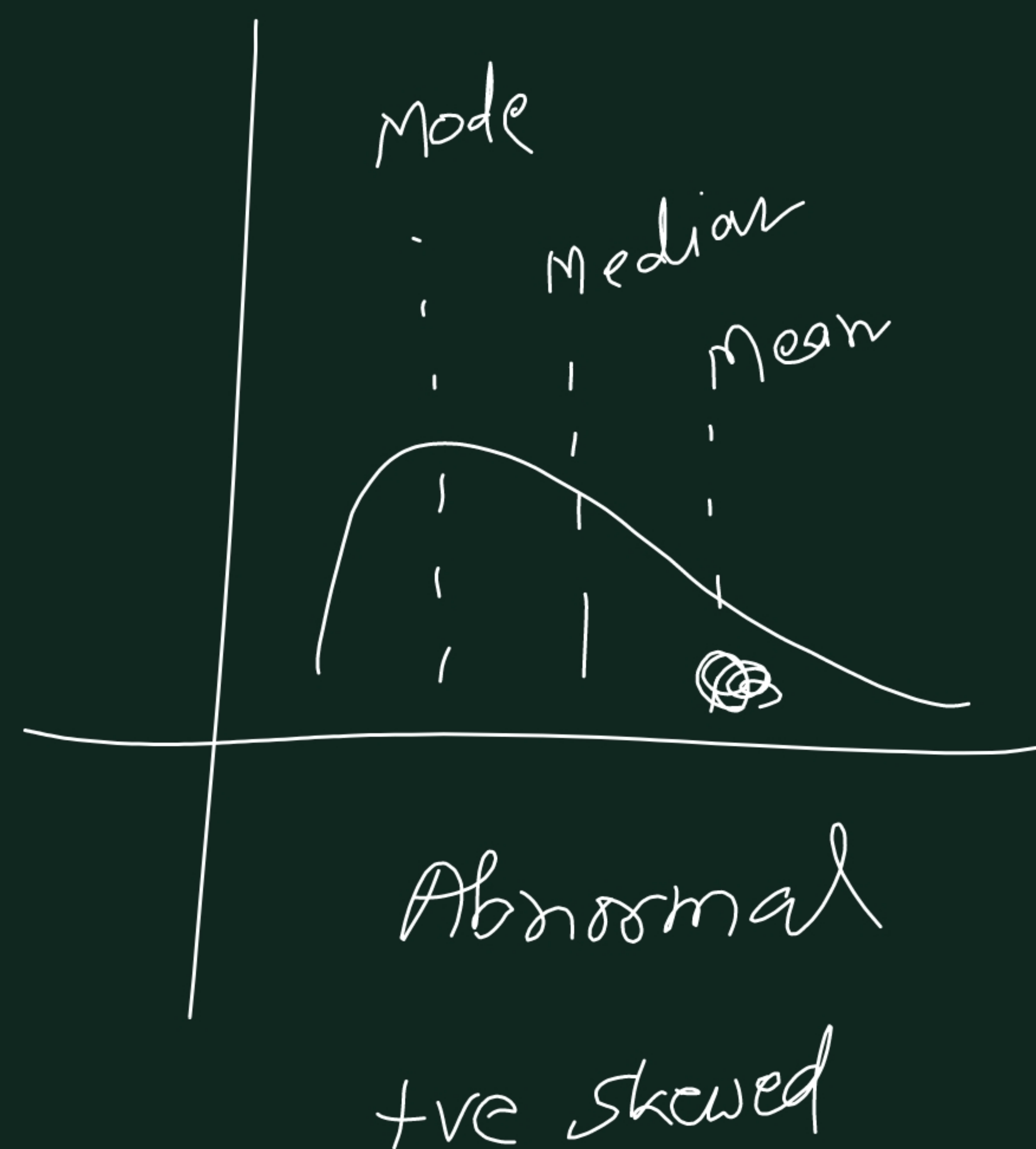
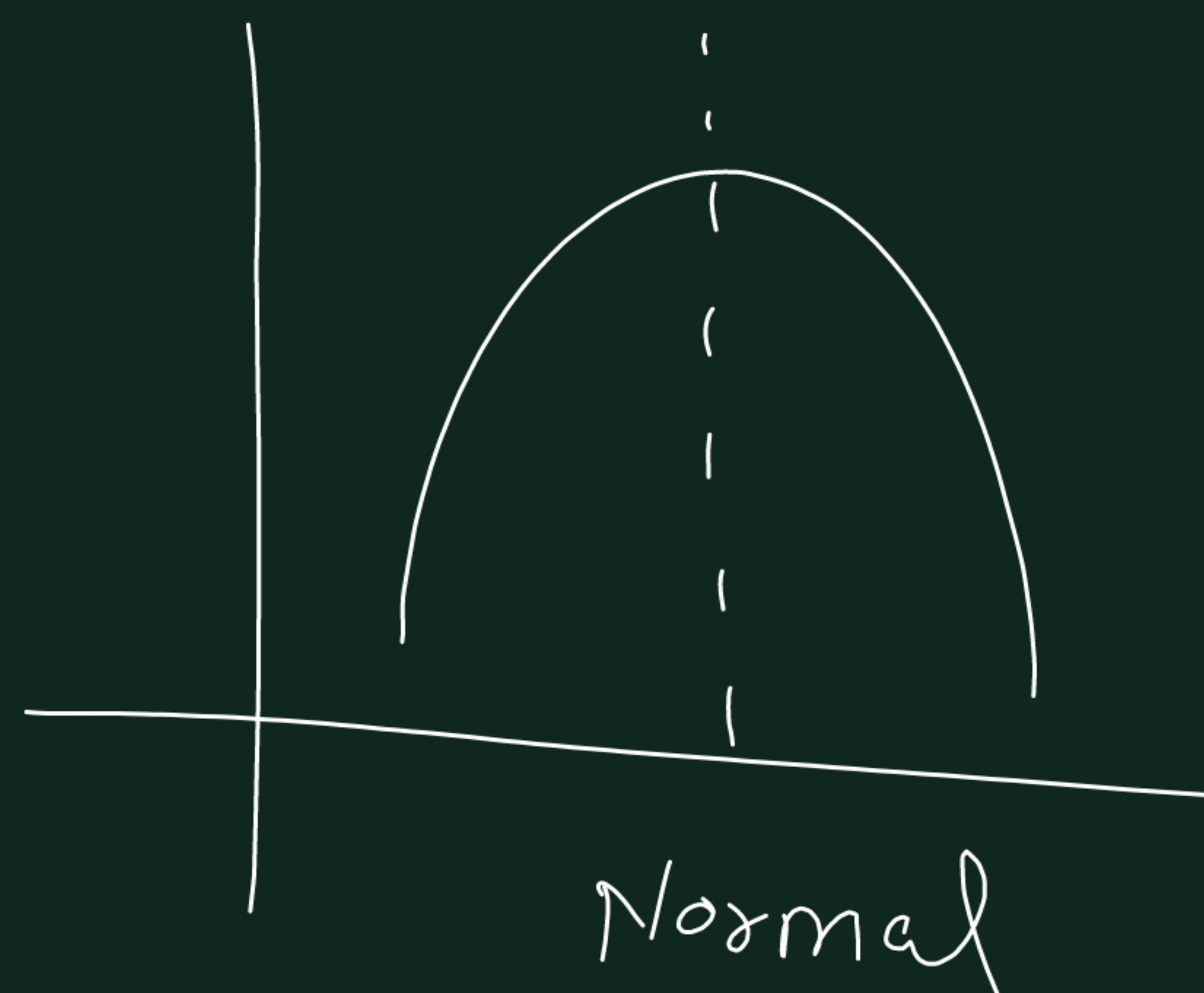
④ Shape Measure

Skewness () \rightarrow direction of abnormal

Kurtosis () \rightarrow strength of tail

Shape :

mean = Median = Mode



⑤ Relationship Measure

Covariance() \rightarrow Direction

Correlation() \rightarrow Strength

├ Positive

├ Negative

└ Neutral

$[-1, 1]$

EDA

① Univariate

② Bivariate

③ Multivariate

