Statistics for Data Science

Summarizing,
Analyzing &
visualization

Descriptive   **Inferential**

→ Population &
Sample

1) Measure of
central tendency → Random
variables

→ Mean
Median
Mode

→ Pdf, Cdf

2) Measure of Dispersion

→ Range
IQR
Standard deviation & variance
Box Plot
3) Outliers (IQR, Lower & upper Whisker)

4) Histogram, Bar graph, Scatterplot

5) Distribution — Normal/Gaussian Distribution
Central Limit Theorem
(Real Time Application)
Skewed/Non-Skewed Distribution

Poisson Distribution

Bernoulli/Binomial Distribution

Hypothesis Testing
1) Null vs
Alternate
Hypothesis
(α)

2) Type 1 &
Type 2 (β)
error

3) P-value

4) Confidence
Interval

↳ Correlation coefficient

.corr()

↳ covariance

↳ 2-score ⸺ Standard Normalization

↳ 2-test, T-test , Anova test

↳ Log Normal Distribution

↳ univariate, bivariate, multivariate

analysis