Complete Statistics for Data Science: Interview Notes

1. Introduction & Basics

Statistics: The science of collecting, organizing, and analyzing data.

Types of Statistics:

Descriptive: Summarizing data (Mean, Median, Mode, Graphs).

Inferential: Using sample data to draw conclusions about a population.

Population vs Sample:

Population (N): Entire group.

Sample (n): Subset of population.

2. Sampling Techniques

Simple Random Sampling

Stratified Sampling

Systematic Sampling

Convenience Sampling

3. Variables & Measurement Scales

Quantitative: Discrete, Continuous

Qualitative: Nominal, Ordinal

Measurement Scales: Nominal, Ordinal, Interval, Ratio

4. Measures of Central Tendency

Mean, Median, Mode

5. Measures of Dispersion

Variance, Standard Deviation

6. Percentiles, Quartiles & Outliers

Percentiles, IQR, Box Plot, Outlier Removal

7. Distributions

Normal Distribution, Empirical Rule, Z-Score

Other Distributions: Log-Normal, Bernoulli, Binomial, Pareto

8. Probability

Independent & Dependent Events

Permutation & Combination

9. Inferential Statistics

Hypothesis Testing, P-Value, Confidence Interval

Type I & Type II Errors

10. Statistical Tests

Z-Test, T-Test, Chi-Square, ANOVA

11. Correlation & Python

Covariance, Pearson, Spearman

Python: pandas, scipy, seaborn