Top 10 Statistics Topics for a Data Scientist

TIP-1: Dive Deep into Each Topic: For every subject you tackle, endeavor to delve into three primary components:
Concept**: Grasp the theory and the fundamentals that underpin the topic.**
Applications**: Understand how the concept is applied in real-world scenarios. Remember, showcasing how you've utilized these theories in practical situations can be even more impactful than merely reciting the theory.**
Interview Questions**: Familiarize yourself with common questions associated with the topic. This will not only prepare you for interviews but also solidify your understanding.**

TIP-2: Prioritize Understanding Over Memorization: While formulas can be critical in certain instances, it's the comprehension of underlying principles and their application that truly counts, especially during interviews. In the real world, you'll often rely on software to handle the calculations. What's crucial is your understanding of the concept at play.

TIP-3: Be Selective with Your Focus: Attempting to master everything about every topic can be a daunting and usually unnecessary task. Instead, concentrate on key concepts and areas most pertinent to the role you're targeting. It's far more advantageous to have an in-depth comprehension of a few relevant subjects than a superficial understanding of a wide range.

Remember, these tips are meant to guide your study and preparation process. The ultimate goal is to be able to convincingly demonstrate your knowledge and skills in the data science domain.

Good luck!

Descriptive Statistics

- Mean
- Median
- Mode
- Range
- Variance
- Standard Deviation
- Skewness
- Kurtosis
- Percentiles/Quartiles
- Box Plots

Inferential Statistics:

1. Hypothesis Testing
2. Null and Alternative Hypotheses
3. Type I and Type II Errors
4. P-Value
5. Confidence Intervals
6. Z-Score
7. T-Score
8. Chi-Square Test
9. F-Test
10. Analysis of Variance (ANOVA)

Probability Theory:

- Basic Probability Rules
- Conditional Probability
- Bayes' Theorem
- Random Variables
- Expectation and Variance of Random Variables
- Independence of Events
- Central Limit Theorem
- Law of Large Numbers
- Combinations and Permutations
- Probability Mass and Density Functions

Probability Distributions

- What is Probability Distribution
- Normal Distribution
- Binomial Distribution
- Poisson Distribution
- Exponential Distribution
- Uniform Distribution

Regression Analysis

- Simple & Multi Linear Regression
- Covariance vs Collinearity
- Assumptions of Regression Analysis
- Interpretation of Coefficients
- Residual Analysis, MSE, MAE
- Multicollinearity
- Heteroskedasticity
- Adjusted R-Squared, R-Squared
- F-Test for Overall Significance
- Log-Linear and Polynomial Regression

Sampling

- Sample vs. Population
- Random Sampling
- Stratified Sampling
- Cluster Sampling
- Systematic Sampling
- Sample Size Determination
- Sampling Distribution
- Sampling Error
- Non-Sampling Error
- Bias in Sampling

Experimental Design and A/B Testing

- Hypothesis Formulation
- Control Group and Treatment Group
- Random Assignment
- Statistical Significance
- Power Analysis
- P-Value Correction (Bonferroni, etc.)
- Confidence Intervals
- Effect Size
- Experiment Duration and Replication

Bayesian Statistics

1. Bayes' Theorem
2. Prior and Posterior Distribution
3. Bayesian Inference
4. Bayesian Update
5. Conjugate Prior
6. Maximum A Posteriori Estimation
7. Bayesian vs Frequentist Approach
8. Credible Intervals
9. Markov Chain Monte Carlo Methods
10. Bayesian Networks