Here’s a **detailed list of statistics topics** that are **essential for Data Analysis**, organized from beginner to advanced levels. These topics will help you **understand, analyze, and interpret data** effectively using tools like Python, Excel, R, or SQL.

**🟢 Beginner Level – Foundations of Statistics**

1. **Types of Data**
   * Qualitative vs Quantitative
   * Discrete vs Continuous
   * Nominal, Ordinal, Interval, Ratio
2. **Data Collection Methods**
   * Surveys, experiments, observational studies
   * Sampling techniques: Random, Stratified, Cluster, Systematic
3. **Descriptive Statistics**
   * Mean, Median, Mode
   * Range, Variance, Standard Deviation
   * Percentiles, Quartiles
   * Skewness and Kurtosis
4. **Data Visualization**
   * Histograms, Bar Charts, Pie Charts
   * Box Plots, Line Charts, Scatter Plots
   * Heatmaps and Correlation Matrix

**🟡 Intermediate Level – Inferential Statistics**

1. **Probability Theory**
   * Basic Probability Rules
   * Independent and Dependent Events
   * Conditional Probability
   * Bayes' Theorem
2. **Probability Distributions**
   * Uniform Distribution
   * Binomial Distribution
   * Normal (Gaussian) Distribution
   * Poisson and Exponential Distributions
3. **Sampling Distributions**
   * Central Limit Theorem (CLT)
   * Standard Error
   * Law of Large Numbers
4. **Confidence Intervals**
   * Z-score and T-distribution
   * Margin of Error
   * Confidence Level (90%, 95%, 99%)
5. **Hypothesis Testing**
   * Null and Alternative Hypothesis
   * p-value and Statistical Significance
   * One-tailed vs Two-tailed tests
   * Type I and Type II Errors
   * Z-test, T-test (one-sample, two-sample)
   * Chi-Square Test
   * ANOVA (Analysis of Variance)

**🔵 Advanced Level – Statistical Modeling & Multivariate Analysis**

1. **Correlation and Regression**

* Pearson and Spearman Correlation
* Simple Linear Regression
* Multiple Linear Regression
* Assumptions in Regression Analysis
* R-squared and Adjusted R-squared

1. **Time Series Analysis** *(If dealing with time-based data)*

* Trend, Seasonality, Noise
* Moving Averages
* Autocorrelation (ACF), Partial ACF
* ARIMA models

1. **Multivariate Statistics**

* Principal Component Analysis (PCA)
* Factor Analysis
* MANOVA (Multivariate ANOVA)
* Cluster Analysis (K-means, Hierarchical)

1. **Non-Parametric Tests**

* Mann-Whitney U test
* Wilcoxon Signed-Rank Test
* Kruskal-Wallis test

1. **Statistical Inference in Machine Learning**

* Cross-validation
* Bias-Variance Trade-off
* Overfitting and Underfitting
* Feature Selection and Multicollinearity

**🔍 Bonus Topics for Data Analysts**

* **A/B Testing** – used in product testing, marketing experiments.
* **Survival Analysis** – useful in medical or business lifetime studies.
* **Bayesian Statistics** – probabilistic approach to inference.
* **Monte Carlo Simulation** – simulate scenarios using random sampling.

Would you like:

* 📊 Recommended datasets to practice each concept?
* 🐍 Python notebooks for each topic?
* 📝 A PDF/Excel checklist to track your learning?

Let me know and I’ll prepare them for you.

1. Design, deploy, and manage Kubernetes clusters on Amazon EKS and Google Kubernetes Engine (GKE).  
2. Implement and maintain infrastructure as code using Terraform.