

DATA SCIENCE

INTERMEDIATE

5 DAYS

Statistics for Data Science and Data Analysis

Build a strong foundation in statistics for data analysis and data science by mastering descriptive statistics, probability, and hypothesis testing. Learn how to apply regression, categorical, and time series analysis while using visualization techniques and analytical tools to communicate data driven insights effectively.

Why Statistics for Data Science?

Informed Decision Making

Extract meaningful insights from your data and make informed choices that lead to better outcomes.

Predictive Analytics

Forecast future trends, demand patterns, and customer behavior.

Accuracy and Reliability

Identify patterns, relationships, and trends in the data, to gain deeper insights into operations, customer behavior, market trends, and more.

What you'll be able to do

- ✓ Develop a solid foundation in statistical concepts and terminology relevant to data science and data analysis.
- ✓ Understand the importance of data exploration and visualization for gaining insights and identifying patterns.
- ✓ Learn to summarize and describe data using descriptive statistics effectively.
- ✓ Gain proficiency in probability theory and its application to data analysis.
- ✓ Master hypothesis testing and confidence interval estimation for making data-driven decisions.
- ✓ Acquire the skills to perform regression analysis to model relationships and make predictions.
- ✓ Understand techniques for analyzing categorical data, such as chi-square tests.
- ✓ Explore time series analysis methods to identify patterns and forecast future values.
- ✓ Learn data visualization principles and techniques for effectively communicating data insights.
- ✓ Apply statistical techniques using popular software tools commonly used in data analysis.

Who this is for

- Data Scientists
- Data Analysts
- Business Analysts
- Data Professionals

Curriculum

01 Statistics and Data Science Foundations

- Connect statistical thinking to data science outcomes
- Frame analysis questions before choosing methods
- Distinguish descriptive vs inferential goals
- Spot common statistical misuse in business reporting

02 Data Types, Exploration, and Visualization

- Classify data types and measurement scales
- Explore datasets for shape, outliers, and structure
- Choose visuals that match the analysis question
- **Hands-on:** Build an exploration notebook/report

03 Descriptive Statistics

- Summarize center and spread with the right measures
- Interpret skewness and variability for decisions
- Compare groups with descriptive summaries
- Communicate summaries without overclaiming

04 Probability and Distributions

- Apply probability foundations to data problems
- Work with common discrete and continuous distributions
- Match distributions to real-world processes
- Use probability to quantify uncertainty

05 Sampling, Estimation, and Confidence Intervals

- Design sampling approaches for analysis goals
- Estimate population parameters from samples
- Construct and interpret confidence intervals
- Relate sample size to estimate precision

06 Hypothesis Testing

- State null/alternative hypotheses clearly
- Run hypothesis tests and interpret p-values carefully
- Control Type I/II errors in practice
- Translate test results into business language

07 Regression Analysis

- Model relationships with regression techniques
- Interpret coefficients and fit diagnostics
- Use regression for explanation and prediction
- Validate models before operational use

08 ANOVA and Chi-Square Tests

- Compare group means with ANOVA
- Analyze categorical associations with chi-square tests
- Check assumptions and choose alternatives when needed
- Report findings with effect context, not only significance

09 Non-Parametric Tests and Experimental Design

- Apply non-parametric tests when assumptions fail
- Design experiments that answer causal questions
- Control confounders and define treatment contrasts
- Plan analysis before collecting experiment data

10 Time Series Basics

- Identify trend and seasonality in time-ordered data
- Apply basic forecasting patterns for planning
- Validate forecasts with time-aware holdouts
- Communicate forecast uncertainty to stakeholders

11 Tools, Ethics, Case Studies, and Reporting

- Use R/Python/Excel appropriately for statistical work
- Apply ethics and privacy responsibilities in analysis
- Practice with case studies across domains
- Present results with clear visuals and caveats