

Statistical Analysis - Roadmap to AI & ML

Learn to apply statistical methods, from descriptive statistics and hypothesis testing to regression, time series and model evaluation for AI and ML.

40 Hours

DURATION

6

MODULES

32

TOPICS

18

HANDS-ON ACTIVITIES

COURSE OVERVIEW

This comprehensive course on Algorithms offers a structured progression from foundational concepts to real-world applications, tailoring skills for AI and ML. It begins with

WHO THIS COURSE IS FOR

AI/ML Enthusiasts & Data Scientists

Power BI & Data Analytics Professionals

Software Engineers Interested in AI/ML

Researchers & Business Analysts

TOOLS AND REQUIREMENTS

Tools and Lab Platform

- Python (Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Statsmodels)
- Jupyter Notebook/Colab

LEARNING OBJECTIVES

- Describe the role of statistics in AI and ML workflows
- Calculate measures of central tendency and dispersion
- Visualise data using histograms, boxplots and scatterplots
- Apply hypothesis testing methods such as t-test and ANOVA
- Interpret confidence intervals, p-values and significance levels
- Apply linear and logistic regression for predictive analysis
- Analyse time series data using ARIMA and SARIMA models
- Evaluate models using confusion matrix, ROC and R^2 metrics

6 Modules, 32 Topics

01

3 Practice exercises

Foundations of Statistical Analysis

TOPICS

- Role of Statistics in AI/ML
- Types of Data (Nominal, Ordinal, Interval, Ratio)
- Central Tendency (Mean, Median, Mode)
- Dispersion (Range, Variance, Std. Dev.)
- Data Visualization (Histogram, Boxplot, Scatterplot)
- Introduction to Probability

02

3 Practice exercises

Inferential Statistics & Hypothesis Testing

TOPICS

- Probability Distributions (Normal, Binomial, Poisson)
- Sampling Techniques
- Central Limit Theorem
- Confidence Intervals & Margin of Error
- **Hypothesis Testing:** t-test, ANOVA, Chi-square
- P-values, Significance Level, Type I/II Errors

03

4 Practice exercises

Advanced Statistical Analysis for AI/ML

TOPICS

Linear & Logistic Regression · Correlation vs Causation · Bayesian Statistics & Inference · Dimensionality Reduction (PCA, t-SNE) · Feature Engineering & Selection Techniques · Cross-Validation & Grid Search CV

04

4 Practice exercises

Time Series Analysis

TOPICS

- Components of Time Series (Trend, Seasonality, Noise)
- Stationarity & Differencing
- Autocorrelation (ACF, PACF)
- AR, MA, ARMA, ARIMA, SARIMA models
- Forecasting & Model Selection

Real-World Applications & Model Evaluation

TOPICS

- Markov Chains & Hidden Markov Models
- Experimental Design & AB Testing
- Evaluation Metrics (Confusion Matrix, ROC, Precision-Recall, R^2 , MSE)
- **Business case mapping:** Healthcare, Retail, Finance

Case Studies

TOPICS

- Customer Demographics Analysis in Retail - Segment customers based on purchasing patterns using summary stats & charts
- A/B Testing for Ad Campaign Effectiveness - Validate if the new campaign performs better using statistical tests
- Employee Attrition Prediction - Use logistic regression and feature engineering to identify why employees leave
- Forecasting Airline Passenger Demand - Use ARIMA to forecast ticket sales based on historical data
- Customer Churn Prediction for Telecom - Build and evaluate a model to reduce churn using statistical evaluation metrics

HANDS-ON ACTIVITIES

Practise What You Learn

#	MODULE	ACTIVITY
01	Foundations of Statistical Analysis	■ Descriptive statistics in Pandas ■ Visualizations using Seaborn ■ Exploratory Data Analysis (EDA) on real dataset
02	Inferential Statistics & Hypothesis Testing	■ Simulating distributions in Python ■ Hypothesis testing using SciPy ■ AB Testing Scenario (Email Campaign)
03	Advanced Statistical Analysis for AI/ML	■ Regression with scikit-learn ■ Feature selection via correlation, mutual info ■ PCA on multidimensional dataset ■ Cross-validation with GridSearch

#	MODULE	ACTIVITY
04	Time Series Analysis	■ Time Series Decomposition ■ ACF/PACF plots with statsmodels ■ Build & evaluate ARIMA model (Stock/Weather dataset) ■ Project: Forecast electricity demand
05	Real-World Applications & Model Evaluation	■ Simulating Markov Chains in Python ■ Confusion Matrix & ROC Curve ■ Mini Project: Customer churn prediction ■ Capstone Project with data storytelling

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training

Live Online Classes (ILO)

Fly-Me-A-Trainer (FMAT)

Flexi (Self-Paced)

1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com

✉ info@koenig-solutions.com