

€ TRAINING

Applied Data Analysis Masterclass:
Visualization, Statistics and Advanced
Programs



Applied Data Analysis Masterclass: Visualization, Statistics and Advanced Programs

Introduction:

Good data selection and/or collection are the foundation for effective data analysis. This demands a thorough understanding of all data kinds and the various sources that they come from. Additionally, by arranging it correctly, all of the data may be easily shown under various charts, and they can all be described using appropriate and effective descriptive statistics measures.

The foundational concepts of this course include developing a clever data collection procedure, choosing the optimum sampling strategy, confirming the accuracy of the data stored for analysis, and comprehending all the choices for visualization and the related descriptive statistical KPIs. Prior to beginning any work or even a career in the realm of data, this course addresses all approaches and instruments for thorough data analysis. The course also acts as a prerequisite for any course or program in machine learning.

Additionally, this course aims to give participants a clear and comprehensive understanding of data structuring for effective data analysis, scientific group profiling through intelligent and effective data analysis, and appropriate data manipulation using a variety of tools currently available on the market.

Course Objectives:

At the end of this conference, the participants will be able to:

- Understand and organize the phases of a successful data analysis project.
- Convert any industry into a thorough database.
- Analyze and assess the quality of the data.
- Basic data interpretation and description using full descriptive statistics
- Discover the full history of data analysis.

Targeted Audience:

- All practitioners of machine learning and artificial intelligence AI start with applied data analysis. It is fundamental information that applies to all fields and data-related jobs.

Course Outlines:

Unit 1: Data visualization and descriptive statistics

- The different types of Data
 - Data sources
 - Data
 - Variables
- Data visualization
 - Pies, Doughnuts, Bars
 - Histograms, Lines, Scatter plots
 - Heat maps and Tuckey boxes
 - Geographical maps
- Central tendency measurements

- Average
- Median
- Mode
- Scatter tendency measurements
 - Quartile
 - Variance
 - Standard deviation
- Estimations
 - Punctual
 - Confidence Interval

Unit 2: Comparing two groups

- Two men test
 - Equal variances t-test
 - Unequal variances t-test - Welch correction
- Two variance tests F-Test
- Two proportion test Chi-Square test
- Two distribution tests Chi-Square test
 - Attraction - Repulsion Matrix
 - Vertical and horizontal profiling

Unit 3: Comparing multiple groups

- Multiple mean tests
 - Equal variances F-Test and ANOVA Table
 - Unequal variances F-Test - Welch Correction
- Multiple Variance test
 - Levene test
 - Chi-Square test
- Multiple proportion test Chi-Square test
- Multiple distribution test Chi-Square test
 - Attraction - Repulsion Matrix
 - Vertical and horizontal profiling
- Mean pair comparisons methods:
 - General
 - Bonferroni
 - Tukey - Kramer

Unit 4: Simple regressions

- Simple linear regression
 - Line equation
 - Testing the regression line validity t-nullity test
 - R vs. R Square interpretation
 - ANOVA table analysis
- Simple logistic regression
 - Probabilistic model
 - Testing the model validity Chi-Square test
 - Predicting classification
 - Odds ratio interpretation

Unit 5: Data analysis project best practices

- Data analysis project best practices
 - Ask
 - Design
 - Preview
 - Analyze
 - Communicate
- Sampling methods
 - Random and systematic
 - Multilevel, stratified, and cluster
 - Convenient, quota, and judgmental
- PMP for research projects overview
 - Integration, cost, scope, time, cost, quality, communication
 - Risk, procurement, and stakeholders