LEARNBAY

DATA SCIENCE TRAINING

- **Course Features**
- Online And Classroom data science training in Bangalore by industry experts.
- Classes with 40% theory and 60% hands on
- Trainers having more than 10+ years of experience in multiple domains like finance,Healthcare ,Retails.
- Practical Approach With Mini Projects And Case studies.
- Job Assistance And Placement Support After end of Course.

- **Who Should Attend:**
- Those who want to become master in **data science and Data Analytics.**
- Business Analysts who want to learn **machine learning**
- Data Analysts who wants to improve their skills.
- Developers aspiring to become data scientist.
- Freshers/Experienced Professional,Managers,IT professional.

- **Course Overview:**
- Python for Data Science
- R Programming
- Introduction to SQL
- Introduction to Statistics
- Apache Spark
- Machine Learning And Introduction to Deep Learning.
- Project And Resume Prep Session
- Interview Guidance And Job Assistance

Duration Of Course - 12 Weeks Course (Weekend Only
(4 Hours Saturday And Sunday) - 100 Hours Course
3 Months Duration

Course Fee: RS. 40,000/-

Course Content

- **INTRODUCTION TO DATA SCIENCE:**
- What is data Science? - Introduction.
- Importance of Data Science.
- Demand for Data Science Professional.
- Brief Introduction to Big data and Data Analytics.
- Lifecycle of data science.
- Tools and Technologies used in data Science.
- Business Intelligence vs Data Science.
- Role of a data scientist.

PART A – R PROGRAMMING BASICS

1. **R Basics, background**

Comprehensive R Archive Network
Demo of Installing R On windows from CRAN Website
Installing R Studios on Windows OS
Setting Up R Workspace.
Getting Help for R-How to use help system
Installing Packages – Loading And Unloading Packages

2. Getting familiar with basics

Operators in R – Arithmetic,Relational,Logical and Assignment Operators
Variables,Types Of Variables,Using variables
Conditional statements,ifelse(),switch
Loops: For Loops,While Loops,Using Break statement,Switch

3. The R Programming Language- Data Types

creating data objects from the keyword.
How to make different type of data objects.
Types of data structures in R
Arrays And Lists- Create Access the elements
Vectors – Create Vectors,Vectorized Operations,Power of Vectorized Operations
Matrices- Building the first matrices,Matrix Operations,Subsetting,visualising subset
Data Frames- create and filter data frames,Building And Merging data frames.

4. Functions And Importing data into R

Function Overview – Naming Guidelines
Arguments Matching,Function with Multiple Arguments
Additional Arguments using Ellipsis,Lazy Evaluation
Multiple Return Values
Function as Objects,Anonymous Functions
Importing and exporting Data into R- importing from files like excel,csv and minitab.
Import from URL and excel Files
Import from database.

5. Data Descriptive Statistics,Tabulation,Distribution

Summary Statistics for Matrix Objects. apply() Command. Converting an Object into a Table
Histograms, Stem and Leaf Plot, Density Function.Normal Distribution

6. Graphics in R – Types of graphics

Bar Chart,Pie Chart,Histograms- Create and edit.
Box Plots- Basics of Boxplots- Create and Edit
Visualisation in R using ggplot2.
More About Graphs: Adding Legends to Graphs, Adding Text to Graphs, Orienting the Axis Label.

PART B – INTRODUCTION TO SQL

1. Introduction to SQL Server and RDBMS

Covers an overview of using relational databases. You'll learn basic terminology used in future modules,**SQL Server Management Studio is the primary tool used to create queries and manage objects in SQL Server databases**

2. SQL Operations

Single Table Queries - SELECT,WHERE,ORDER BY,Distinct,And ,OR
Multiple Table Queries: INNER, SELF, CROSS, and OUTER,oin, Left Join, Right Join, Full Join, Union and MANY MORE…..

3. SQL Advance -Operations

Data Aggregations and summarizing the data
Ranking Functions: Top-N Analysis
Advanced SQL Queries for Analytics

PART C- PYTHON FOR DATA SCIENCE

1. Python Programming Basics

Installing Jupyter Notebooks
Python Overview
Python 2.7 vs Python 3
Python Identifiers
Various Operators and Operators Precedence
Getting input from User,Comments,Multi line Comments.

2. Making Decisions And Loop Control

Simple if Statement,if-else Statement
if-elif Statement.
Introduction To while Loops.
Introduction To for Loops,Using continue and break,

3. Python Data Types: List,Tuples,Dictionaries

Python Lists,Tuples,Dictionaries
Accessing Values
Basic Operations
Indexing, Slicing, and Matrixes
Built-in Functions & Methods
Exercises on List,Tuples And Dictionary

4. Functions And Modules

Introduction To Functions – Why
Defining Functions
Calling Functions
Functions With Multiple Arguments.
Anonymous Functions - Lambda
Using Built-In Modules,User-Defined Modules,Module Namespaces,
Iterators And Generators

5. File I/O And Exceptional Handling

Opening and Closing Files
open Function,file Object Attributes
close() Method ,Read,write,seek.Exception Handling,the try-finally Clause
Raising an Exceptions,User-Defined Exceptions
Regular Expression- Search and Replace
Regular Expression Modifiers
Regular Expression Patterns,re module

6. Numpy

Introduction to Numpy. Array Creation,Printing Arrays
Basic Operations- Indexing, Slicing and Iterating
Shape Manipulation - Changing shape,stacking and spliting of array
Vector stacking

7. Pandas And Matplotlib

Introduction to Pandas
Importing data into Python
Pandas Data Frames,Indexing Data Frames ,Basic Operations With Data frame,Renaming Columns,Subletting and filtering a data frame.
Matplotlib - Introduction,plot(),Controlling Line Properties,Working with Multiple Figures,Histograms

PART D- INTRODUCTION TO STATISTICS

1. Fundamentals of Math and Probability

Basic understanding of linear algebra, Matrics, vectors
Addition and Multimplication of matrics
Fundamentals of Probability
Probability distributed function and cumulative distributed function.

Class Hand-on

Problem solving using R for vector manupulation
Problem solving for probability assignments

2 Descriptive Statistics

Describe or sumarise a set of data
Measure of central tendency and measure of dispersion.
The mean,median,mode, curtosis and skewness
Computing Standard deviation and Variance.
Types of distribution.

Class Handson:

5 Point summary BoxPlot
Histogram and Bar Chart
Exploratory analytics R Methods

3. Inferential Statistics

What is inferential statistics
Different types of Sampling techniques
Central Limit Theorem
Point estimate and Interval estimate
Creating confidence interval for population parameter
Characteristics of Z-distribution and T-Distribution
Basics of Hypothesis Testing
Type of test and rejection region
Type of errors in Hypothesis resting, Type-I error and Type-II errors
P-Value and Z-Score Method
T-Test, Analysis of variance(ANOVA) and Analysis of Co variance(ANCOVA)
Regression analysis in ANOVA

Class Hands-on:

Problem solving for C.L.T
Problem solving Hypothesis Testing
Problem solving for T-test, Z-score test
Case study and model run for ANOVA, ANCOVA

4. Hypothesis Testing

Hypothesis Testing
Basics of Hypothesis Testing
Type of test and Rejection Region
Type o errors-Type 1 Errors,Type 2 Errors
P value method,Z score Method

PART E – UNDERSTANDING AND IMPLEMENTING MACHINE LEARNING

1. Introduction To Machine Learning
What is Machine Learning?
What is the Challenge?
Introduction to Supervised Learning,Unsupervised Learning
What is Reinforcement Learning?

2. Linear Regression
Introduction to Linear Regression
Linear Regression with Multiple Variables
Disadvantage of Linear Models
Interpretation of Model Outputs
Understanding Covariance and Colinearity
Understanding Heteroscedasticity
Case Study – Application of Linear Regression for Housing Price Prediction

3. Logistic Regression
Introduction to Logistic Regression.– Why Logistic Regression .
Introduce the notion of classification
Cost function for logistic regression
Application of logistic regression to multi-class classification.
Confusion Matrix, Odd's Ratio And ROC Curve
Advantages And Disadvantages of Logistic Regression.
Case Study:To classify an email as spam or not spam using logistic Regression.

4. Decision Trees And Supervised Learning
Decision Tree – data set
How to build decision tree?
Understanding Kart Model
Classification Rules- Overfitting Problem
Stopping Criteria And Pruning
How to Find final size of Trees?
Model A decision Tree.
Naive Bayes
Random Forests and Support Vector Machines
Interpretation of Model Outputs
Case Study:
1 Business Case Study for Kart Model
2 Business Case Study for Random Forest
3 Business Case Study for SVM

5. Unsupervised Learning
Hierarchical Clustering
k-Means algorithm for clustering – groupings of unlabeled data points.
Principal Component Analysis(PCA)- Data
Independent components analysis(ICA)
Anomaly Detection
Recommender System-collaborative filtering algorithm
Case Study– Recommendation Engine for e-commerce/retail chain

6. Introduction to Deep Learning
INeural Network
Understaing Neural Network Model
Understanding Tuning of Neural Network

Case Study:
Case study using Neural Network

7. Natural language Processing

Introduction to natural Language Processing(NLP).
Word Frequency Algorithms for NLP
Sentiment Analysis

Case Study :

Twitter data analysis using NLP

8. Apache Spark Analytics

What is Spark
Introduction to Spark RDD
Introduction to Spark SQL and Dataframes
Using R-Spark for machine learning

Hands-on:

installation and configuration of Spark
Hands on Spark RDD programming
Hands on of Spark SQL and Dataframe programming
Using R-Spark for machine learning programming

9. Introduction to Tableau/Spotfire

Connecting to data source
Creating dashboard pages
How to create calculated columns
Different charts

Hands-on:

Hands on on connecting data source and data clensing
Hands on verious charts
Hands on deployment of Predictive model in visualisation

* Real Time Project
*Certificates On successful Completion of Project .
*Resume Preparation Tips
*Interview Guidance And Support
*Job Support And Placement Assistance

Course FAQ

1. **What are the profiles of Trainers?**
Our Trainers have relevant industry experience and are working in MNC as data scientist.

2.How many Case studies and Projects are covered in the course?
Course has multiple case studies and mini Project.Our course is designed by industry experts.Course features many real time problems.Please refer course content for more details.

3. Do i Need to carry my own laptop?What the the softwares required?
Yes,You need to carry your own laptop.To start with ,You need to install R And R studio installed in your system.
Both Of these are open source and in first class,trainer will help you to setup the environment in your system.

4. Can i Attend a Demo Session before enrolling for the course?
Of course,You can attend a Free live Demo Session before enrolling for the Course.

5. Are the session Online or Classroom?
We provide both live Online and classroom session.You can opt for online or classroom based on your
convenience.

6. Will i Get Class Recording if i Enroll for Classroom Session?
Our live Classroom Sessions are recorded and after the session ,Class recording will be shared to you.