

TRI-STARC's SOFTWARE TRAINING & PROJECT WORK ON DATA SCIENCE

With R Programming, R Studio & Python

•••

Click here for Registration https://forms.gle/Z1iN1PokMLE7Yapo7

OBJECTIVE

To train the students for real time project works and enabled them to handle Business Intelligence, Statistical Data Analysis, Big Data Analytics, etc. so that they can be prepared for good placements immediate after the completion of their courses.

TARGET GROUP



Students of Under Graduation (B.Sc. /B.Tech.) & Post-Graduation (M.Sc., M.B.A., M.Tech.) of Statistics/ Applied Statistics (Bio-statistics/ Agricultural Statistics/ Medical Statistics/ Actuarial Statistics/ Management Statistics), Data Science/ Data Analytics, etc.

MODE OF CLASSES:

- Student's Tutorial oriented class Training
- The conceptual understanding of the classical theory of data science
- Hands on experience with live data sets through means of project works
- Virtual Live Classes not the recorded videos.

TIMINGS:



7.00 PM to 8.30 PM IST from Monday to Friday

Registration Google Form https://forms.gle/Z1iN1PokMLE7Yapo7 TRI-STARC's SOFTWARE TRAINING & PROJECT WORK ON DATA SCIENCE With IT Industary experts & teaching experience

The Purpose of the training is to enable the students of B.Tech., B.Sc., M.Tech., M.Sc., M.B.A, and other similar programs in Statistics and Data Science for dealing the real time project works of Big Data Analytics.

Needles to say the significance of Machine Learning and Deep Learning Techniques in handling the tasks of Big Data Analytics, Statistical Data Modeling, Predictive Data Analysis, Business Intelligence, etc. with Data Science.

This training is intended to provide all the necessary insights of practical training and hands on experience orientation for getting the better job opportunities for the people who have the passion on working with Data Intelligence.

Interest Students are suggested to be registered by submitting their details through the following Google Form Link https://forms.gle/Z1iN1PokMLE7Yapo7 It will be facilitating them (exclusively registered) to have the clarifications after registration.



Ms. Rajya Lakshmi K Resource Person

Qualifications

M.Sc (Statistics) ; M.Phil (Statistics) & Certified Business Analytics Indian School of Business– Hyd

Experience:

HOD, Department of Statistics Pydah P.G courses – INDIA; Genpact & Ed- Ventures E-learning - INDIA; TCS, CTS, Google, Novartis – INDIA; ICON Docs, Watzman Consultancy - GERMANY & BCBS – USA.

Total Experience:

25 years both in Teaching and IT companies

Designations:

Senior Statistical Analyst; Senior Consultant ; Data Scientist; R/Python/SAS programmer, Asst Manager

Skill sets:

Python, R, SAS, SPSS, SQL, Excel-VBA, Tableau, Statistical modelling

Worked with various Data Science, AI projects and core experience in the following areas:

- Define the Problem for which we bring answer through data analysis.
- Data Acquisition means gathering relevant data from various sources, like databases, APIs, files, CSV, Excel, or web scraping etc.
- Data Cleaning and Preprocessing:
 - 1. Handle missing values: Impute or delete them based on the context.
 - 2. Handle outliers: Decide whether to remove, cap, or transform them.
 - Data normalization or standardization: Scale numerical features if needed.
 - 4. Data encoding: Convert categorical variables into a numerical format suitable for analysis.
- Exploratory Data Analysis (EDA):
 - 1. Explore the dataset using statistical summaries and visualizations.
 - 2. Identify patterns, trends, correlations, and relationships between variables.
 - 3. Generate insights that may inform the next steps in your analysis.
- Feature Engineering:
 - Create new features from existing ones that may enhance model performance.
 - 2. Select relevant features that contribute most to predicting the target variable.
- Model Selection and Training:
 - 1. Choose appropriate models based on the problem type (classification, regression,

clustering, etc.) and data characteristics.

- 2. Split the data into training and testing sets for model validation.
- 3. Train the selected models using appropriate techniques (e.g., crossvalidation, hyper parameter tuning).
- Model Evaluation:
 - Evaluate models using appropriate metrics (accuracy, precision, recall, F1-score for classification; RMSE, MAE regression).
 - 2. Compare different models and select the bestperforming one based on evaluation metrics.
- Model Deployment:
 - Deploy the chosen model into production, integrating existing systems or creating a new application interface.
- 2. Monitor the model's performance in production and update as necessary.
- Iterate and improve the model:
 - Gather feedback from stakeholders and monitor model performance over time.
 - 2. Iterate on the model by refining features, retraining with new data, or updating algorithms to improve performance.
- Documentation & Presentation:
 - Document the entire process, including data sources, cleaning steps, modeling techniques, and results.
 - 2. Present findings and insights to stakeholders using visualizations and clear explanations.

COURSES OFFERED

4

5

Data Analytics & Visualization :

Components of Excel, SQL Techniques, Data Visualization using Tableau & Project work on usage of Excel, SQL and data Visualization with real time data

Data Science with R-Programming :

2

3

Basic R Programming, Basic Statistical Analysis, Basic Python programming & Project work by using Basic Statistical Analysis, R and Python Programs with real time data

Big Data Analytics with R-Programming : Advanced R Programming, Advanced Statistical Analysis, Forecasting Analytics & Project work by using Advanced Statistical Analysis, R Programming with Forecasting Analytics using real time data

Machine Learning Through Python :

Machine Learning Techniques, Deep Learning Techniques, Advanced Python Programming & Project work by using Machine & Deep Learning Techniques with advanced Python Programming for real time data

Business Intelligence and Data Analytics with Python :

Machine Learning Techniques, Text Analytics – NLP, Advanced Python Programming & Project work by using Machine Learning Techniques, Text Analytics through NLP with advanced Python Programming for real time data

Course – I: Data Analytics & Visualization

S. No.	Training Module	Duration	Man Hours	Content Details	Fee
1	Components of Excel	2 weeks	1.30 hrs per day - (10 days)	Formulas, functions, shortcuts, cell references, formatting, style and Themes, filters, sorting, Graphs, Dashboards	
2	SQL Techniques	3 weeks	1.30 hrs per day - (15 days)	Intro, Installation, Retrieve, update, Inserting, Deleting, sorting, filtering, summarizing, grouping subqueries, Joining, stored procedures & views	₹ 5000/- only
3	Data Visualization using Tableau	2 weeks	1.30 hrs per day - (10 days)	Intro on DV, Intro on Tableau, connecting with Data, Visual graphs and insights, Analytics with Tableau, Dashboards, Storytelling, Mapping Calculations, KPI boards	
4	Project work on usage of Excel SQL and data Visualization with real time data				

Course – II Data Science with R-Programming

S. No.	Training Module	Duration	Man Hours	Content Details	Fee
1	Basic R Programming	2 weeks	1.30 hrs per day - (10 days)	 R/R-Studio as a statistical Software and Language Functions & Packages in R Data frames Fundamental of the R Language Basic commands in R Programming Data preparations with R 	
2	Advanced Statistical Analysis-I	2 weeks	1.30 hrs per day - (15 days)	 Probability distributions, BOXPLOT, statistical significance, ANOVA Correlation, Regression Analysis, Principles of Least Squares and Curve Fitting Sample, Populations, Statistic, Parameter, Sampling Distribution Large Sample Tests, Small Sample, Tests, Goodness of Fit, Test for Independence, Confidence interval 	₹ 6000/- only
3	Basic Python programming	4 weeks	1.30 hrs per day - (20 days)	 Significance and installation of Python Values, variables and statements Conditional executions Iterations like while, nested, for, infinite loops Functions Lists, objects, custom types, Imports and exports of files in Python 	

4

Project work by using Basic Statistical Analysis, R and Python Programs with real time data

Course – III: Big Data Analytics with R-Programming

S. No.	Training Module	Duration	Man Hours	Content Details	Fee
1	Advanced R Programming	4 weeks	1.30 hrs per day - (20 days)	1. R programming: Data preparations with R for ML, Installing packages, Reading and Writing the vartious formatting files in R, Numbers and different operators in R, practice on decisions through if-else clauses, Base R Graphics, statistics with R, Hypothesis Testing in R	
2	Advanced Statistical Analysis	3 weeks	1.30 hrs per day - (15 days)	Regression, Principal Component Analysis, Logistic Regression, Cluster Analysis, Factor analysisLinear Vs Non-Linear,	₹ 7500/- only
3	Forecasting Analytics	3 weeks	1.30 hrs per day - (15 days)	Intro for Forecasting Analytics, Why forecasting, Data collection, Time series components, Additive and Multiplicative model, Common Predictive Accuracy Measures, Autoregressive models, ARIMA, Smoothing, Naïve forecasts, Econometric models	
4	Project work by using Advanced Statistical Analysis, R Programming with				

Forecasting Analytics using real time data

Course – IV: Machine Learning Through Python

S. No.	Training Module	Duration	Man Hours	Content Details	Fee
1	Machine Learning Techniques	4 weeks	1.30 hrs per day - (20 days)	Intro to ML, Data Validation & Quality, Data Cleaning & Analyzing the data, Databases handling, Connect various databases, Patterns and models, EDA, Missing Value Mechanisms & Patterns, When can be Missing Values Ignored?, Missing Value Imputation Methods	
2	Deep Learning Techniquaes	4 weeks	1.30 hrs per day - (20 days)	Neural Networks, CNN, RNN, Deep learning Models, Non-convex optimization for deep networks	₹ 8000/- only
3	Advanced Python Programming	3 weeks	1.30 hrs per day - (15 days)	1. Python programming : Data preparations with Python for ML, Installing packages, Reading and Writing the vartious formatting files in Python, Numbers and different operators in R, practice on decisions through if-else clauses, Base R Graphics, statistics with R, Hypothesis Testing in R	
4	Project work by using Machine & Deep Learning Techniques with advanced				

Python Programming for real time data

Course – V: Business Intelligence and Data Analytics using ML&NLP Through Python

S. No.	Training Module	Duration	Man Hours	Content Details	Fee	
1	Machine Learning Techniques	4 weeks	1.30 hrs per day - (20 days)	Intro to ML, Data Validation & Quality, Data Cleaning & Analyzing the data, Databases handling, Connect various databases, Patterns and models, EDA, Missing Value Mechanisms & Patterns, When can be Missing Values Ignored?, Missing Value Imputation Methods		
2	Text Analytics - NLP	3 weeks	1.30 hrs per day - (15 days)	What is Natural Language Processing? Required packages in NLP in R, Building your Corpus, Tokenization, N-grams, Stemming and Lemmatization, Tokenizing POS Tagging and Stop words, Text "Features" and TF-IDF Classification, Sentiment Analysis, Dendograms, PCA scatterplots & k-means, Advanced Vector Analyses	₹ 7500/- only	
3	Advanced Python Programming	3 weeks	1.30 hrs per day - (15 days)	1. Python programming : Data preparations with Python for ML, Installing packages, Reading and Writing the vartious formatting files in Python, Numbers and different operators in R, practice on decisions through if-else clauses, Base R Graphics, statistics with R, Hypothesis Testing in R		
Project work by using Machine & Deep Learning Techniques with advanced Python						

Programming for real time data