

Certification Program on Business Analytics and Data Mining

**During November 2019 – January 2020
[16 days]**



Conducted by:
SQC & OR Unit, Mumbai
Indian Statistical Institute,
Room No 320, 3rd Floor Old C G O Building
101 Maharshi Karve Road, Mumbai 400 020:
Tel 22014588 / 22004574
email: sqcbombay@gmail.com
www.isimumbai.co.in

Introduction:

Business analytics refers to the skills, technologies, practices for continuous iterative exploration and investigation of past business performance to gain insight and drive business planning. Data mining is the set of methods and techniques for exploring and analyzing large data sets, in an automatic or semi-automatic way, in order to find among these data certain unknown or hidden rules, association or tendencies. Both the topic requires analysis of data using statistical and quantitative techniques using various software. This program is planned to impart basic skill of data analysis is using appropriate statistical techniques and machine learning techniques.

About the Institute:

The Indian Statistical Institute is a quasi central organization under the Ministry of Statistics & Programme Implementation. It is declared by an Act of Parliament as an Institute of National Importance. Over the years the Institute has grown as a multi-disciplinary organization. It functions as a University empowered to award degrees up to Ph.D.; as a Corporation in undertaking large scale projects; as a Firm of Consultants to industries to improve Quality, Reliability and Efficiency and as a meeting place.

Programme Objectives:

This program is designed to guide business analytics/ data mining professionals in extracting implicit, previously unknown and potentially useful knowledge from large data sets, developing performance models & usage of optimization techniques.

Program Benefits:

The participants will acquire the knowledge on, various statistical and analytical techniques requires for business analytics and data mining. All the topics will be covered using R-studio and MS-Excel.

Course Coverage:

The course is divided in 5 modules as follows:

Module – 1: Introduction, Basic Concepts and Visual Analytics using R

- Introduction to Analytics and concepts of statistical / machine learning / analytics problems.
- Introduction to random experiments and random variable, concepts of parameters, concepts of visual analytics using random variables and their parameters, different summary measures and presentations, examples and exercises
- Graphs and Charts for data visualisation
- Sampling Concept and Methods
- Introduction to R programming (writing R codes for data manipulation and simple graphics as well as basic statistical computations) with hands on exercises.

Module – 2: Probability distributions, Estimation and Hypothesis Testing

- Distribution as a model of a business process; usage of distribution for decision making.
- Probability and Distributions, understanding normal distribution, brief introduction to some useful distributions, probability computations
- Estimation and hypothesis Testing
 - Estimation of parameters like means, variances, proportions and model parameters in different circumstances and their usages in analytics;
 - Concepts of standard errors and confidence intervals
 - Theory of estimates; concept of likelihood, sufficiency and information criteria.
 - Formulating hypotheses in real life scenarios
 - Test for means, variances, proportions, odds ratios and relative risks

Module – 3: Prediction Modeling through Regression

- Simple and multiple linear regressions.
- Concepts of cross validation – usage of validation set, LOOCV, k fold cross validation and concepts of bootstrap.
- Logistic regression
- Introduction to censored data, Survival analysis models
- Nonparametric estimation of survival probability using Kaplan – Meier method

Module – 4: Tree based Method, Forecasting and Segmentation

- Classification and Regression Tree including concepts of bagging, random forests and boosting, fitting and validating tree based models
- Forecasting Models exponential smoothing (Holt and winters model) and ARIMA.
- Cluster analysis, carrying out non-hierarchical clustering using R, choosing the right solution for non-hierarchical clustering

Module – 5: Neural Network and Deep Learning:

- Association rule mining (Market Basket) analyses
- Naïve Bayes classifier
- Artificial Neural Network
- Introduction to deep learning including project and presentation
- Wrap up and examination

After each module assignments will be given. Participants need to carry out the assignments in team and present the finding in first day of next module.

Target Participants:

- Managers / Executives/ Professionals/ Students associated with data analysis or planning to make career in Analytics.

[Note: The participants should have a basic operating understanding of calculus, matrices and linear algebra. Attending the program without this background is possible, but likely to be more challenging.]

Schedule:

- Module – 1 November 8-10, 2019 (Friday, Saturday and Sunday)**
- Module – 2 November 22-24, 2019 (Friday, Saturday and Sunday)**
- Module – 3 December 6-8, 2019 (Friday, Saturday and Sunday)**
- Module – 4 December 20-22, 2019 (Friday, Saturday and Sunday)**
- Module – 5 January 10-12, 2020 (Friday, Saturday and Sunday)**

Overall Review and Exam: January 25, 2020

Faculty:

Experienced faculties of SQC & OR Division having in-depth experience in data analysis and its application in various industries.

Course Fee:

- INR 100000/- per participant + 18% GST (INR 118000/-) **for corporate/ sponsored persons**
- INR 80000 per participant + 18% GST (INR 94400/-) **for students and individual participants**
- US\$ 2000 for overseas participants (inclusive of course material, kit, lunch & snacks).

Registration:

- Please send your nomination along with appropriate course fee in the attached registration format. Before making online payment, please check availability of seats. Fees once paid will not be refunded under any circumstances
- All participants are required to bring laptop, for carrying out analysis using R-studio.
- Registrations are done on 'first come first served' basis.

Important Dates:

Last date of submission of nominations: **Nov-5, 2019**

Program dates: November 8-10, 22-24: December 6-8, 20-22: January 10-12 and 25

Timing: **9:30 hrs - 17:30 hrs**

Venue:

**Conference Room, SQC & OR Unit,
Indian Statistical Institute,
Room No 320, 3rd Floor Old C G O Building
101 Maharshi Karve Road, Mumbai 400 020**

Sponsored/Individual/Overseas

**Training Programme on
Certification program for Business Analytics and Data Mining
November 2019 to January 2020**

Organization:

Mailing Address:

Contact Person:

E-mail:

Phone: Fax :

Details of the participants attending the program:

1. Name :

Position :

2. Name :

Position :

Details of DD/Cheque attached (Rs 118000/- sponsored participants and Rs.94400/- for individual) per participant inclusive of GST @18% in favour of "Indian Statistical Institute" payable at Mumbai

Amount : _____

DD/Cheque No. : _____ Date : _____

Bank : _____ Branch : _____

Signature: _____ Name: _____ Date : _____

Bank Details for online payments [please inform details of online payment to office]

Bank Name: STATE BANK OF INDIA

Bank Account No: 10996682279

Account Name; INDIAN STATISTICAL INSTITUTE, Type: Current

Branch: MUMBAI MAIN BRANCH

Bank Address: MUMBAI SAMACHAR MARG, HORNIMAN CIRCLE, FORT
MUMBAI 400023

IFSC code: SBIN0000300

Signature :

Name :

Date

Xerox copy of this form may be used in case of more participants. Please e-mail the form to: sqcbombay@gmail.com or send to :

Program Director, SQC & OR Unit, Indian Statistical Institute, Room No,320, 3rd Floor, Old CGO Bldg,101, Maharshi Karve Road, Mumbai 400 020

Tel. 022 - 22014588 / 22004574

www.isimumbai.co.in