



Machine Learning with SAS

(33 hours class room + 30 hours of practice sessions)

About the Course

SAS is one of the most powerful statistical software used in the world of Data Science. It not only helps in Machine Learning algorithms, but also it has very robust 450 plus inbuilt functions that makes it very efficient in data management and data cleaning. SAS is a language which is easiest to learn among all the software being used for analytics. This course will be helpful for novice to professionals, who would be learning not only the power functionalities of SAS coding, but also the Machine Learning algorithms to add value to clients of different domain.

Overview of the course

Class 1: Import data into SAS and manipulate them

- Introduction to SAS
- 5 main windows of SAS
- Importing data into SAS
- Data step Vs Proc Step
- Conditional processing: If, else if, and else statement
- Boolean in if else statement
- Where statement
- All types of merging using data step

Class 2: Different Proc statements

- Proc Print
- Proc means and all the options
- Proc univariate
- Proc Freq
- Proc sort
- Removal of duplicates



Class 3: SAS Functions – Date, Numeric and Character

Difference between functions and Proc

Inbuilt Numeric functions of SAS

Inbuilt Character functions of SAS

Inbuilt Date functions of SAS

Class 4: SQL in SAS and other advanced functionalities

SQL queries

Merging with SQL

Macros in SAS

Output Delivery System in SAS

Class 5: Full Statistic refresher course

Everything you want to know about statistics....Well sort of!!

Mean, Median, Mode

Standard Deviation, Variance,

Normal Distribution

Hypothesis testing

T-test, Anova, Normality test

Class 6: Linear Regression with SAS

Predictive Analytics – Linear Regression

Concepts of Linear Regression

Simple and Multiple Linear Regression

Automatic Dummy Variables creation technique

Model Validation parameters

Model Assumption testing

Splitting of data for Validation and testing

Business Case Study with real data to model in SAS software



Class 7: Additional Case study on Linear Regression

Participants will be asked to develop a Linear Regression model on a real life data, in presence of the instructor. Time given is 2.5 hours. Participants will be treated like an industry employee, but in terms of help certainly the instructor will not be as ruthless as the boss. After completion of the model (with the help of the instructor wherever it is required), the instructor will show how to present a model to a real life client.

Class 8: Logistic Regression with SAS

Predictive Analytics – Logistic Regression
Concepts of Logistic Regression
Difference between Linear Regression and Logistic Regression
Automatic Dummy Variables creation technique
Model Validation parameters
Model Assumption testing
Splitting of data for Validation and testing
Business Case Study with real data to model in SAS software

Class 9: Additional Case study on Logistic Regression

Participants will be asked to develop a Logistic Regression model on a real life data, in presence of the instructor. Time given is 2.5 hours. Participants will be treated like an industry employee, but in terms of help certainly the instructor will not be as ruthless as the boss. After completion of the model (with the help of the instructor wherever it is required), the instructor will show how to present a model to a real life client.

Class 10: Time Series Forecasting with SAS

Time series forecasting: ARIMA
Difference between forecasting and prediction
Concepts of time series data
Concepts of ARIMA
Descriptive analytics for ARIMA
Development of model
Best model selection
Forecasting with the best model
Business Case Study with real data to model in SAS software



Analytics Educator

Bansdroni, Kolkata

Mb: 9163223228

www.analyticseducator.com

analyticseducator@gmail.com

Important points:

1. After each class, assignments will be given as homework which are needed to be completed before the next class. The first 15 minutes of every class will be reserved to answer the participant's queries.
2. After every session, the discussed codes, presentations, handouts will be emailed to all the participants. Participants are advised to carry it either in soft copy or as print outs in the class.
3. Participants are advised to bring their own computers so that they can practice the codes along with the instructor.
4. Normally the class duration would be 3 hours, with a break of maximum 5-10 minutes depending of the requirement of the participants. In case all the queries of the participants are not answered with in the stipulated time of 3 hours then the instructor will extend the class by 15 minutes to 30 minutes.
5. After the completion of the module, there will be an option for all the participants to work on other case studies on real life data for further practice. (This is optional and will not be considered for calculating your final grade)
6. If a participant feels that he/she requires further help on certain topic, then they can attend the same session of some other batch.