Basic Analytics

* **Overview of Analytics**
* **Models and Algorithms**
* **Analytics Methodology**
* **Problem Definition**
* **Data Exploration**
* **Data Preparation**
* **Linear Regression**
* **Logistic Regression**
* **Decision Trees**
* **Clustering**

Analytical tools

SAS

World Programming

Excell

SQL

Revolution analytics

Soft Skills

* **Resume Writing**
* **Interview Skills**
* **Presentation Skills**
* **Teleconference Etiquette**
* **E-mail writing**

Statistics

* **Introduction to statistics**
* **Summary statistics**
* **Probability**
* **Probability distribution**
* **- Binomial**
* **- Poisson**
* **- Normal**
* **Hypothesis testing**
* **T-test**
* **Chi-square test**
* **ANOVA**
* **Non parametric testing**

Overview of Analytics

• What is analytics?

• Types of problems in analytics

• Case studies of application of analytics in

business

• When analytics does not work

• Analytics vs. data warehousing, OLAP,

Statistics

• Widely used analytic software

• Companies using analytics

• Day in the life of a business analyst

• Career path in analytics

• Qualities of a business analyst

Models and Algorithms

• Modeling Terminology

• Linear Regression

• Logistics Regression

• Decision Trees

• MARS

• Rule Induction

• K-nearest

• Neural Network

• Genetic Algorithm

Analytic Methodology

• Problem definition

• Data selection

• Data exploration

• Data partition

• Data cleansing

• Data transformation

• Modeling

• Validation

• Deployment

• Assessment

• Re-start

Problem Definition

• Basics of problem definition

• Case study - Car Insurance

• Case study - Credit Cards

Data Exploration

• Overview

• Case study - Telecom data

• Preliminaries

• Types of data

• Data sanity checks

• Summary statistics

• Derived variables

• Cross tabs

• Graphical analysis

• Exceptions

• Correlations

Data Preparation

• Why data prep

• Outlier treatment

• Missing values treatment

• Telecom case study

• Categorical variables

• Dummy variables

• Derived variables

• Lag variables

• Interaction variables

• Variable transformation

• Quadratic variables

• Date, time variables

• Sampling and partitioning

• Case study - Auto manufacturer

Regression

• Basics of Regression

• Linear Regression

• Logistic Regression

• Interpretation of modeling results

• Violation of regression assumptions

• Insurance Case study

Decision Trees

• What are decision trees?

• Examples of trees

• Terminology in decision trees

• Data preparation for trees

• How to create a tree?

• Measure of effectiveness

– Gini

– Chi-square

– Information gain

– Reduction in variance

– Others

• Application of algorithms

• Case study - Fraud detection

• Case study - Car Insurance pricing

• Use of decision trees

• Pros and cons

• What makes a good tree?

• When to use Decision trees?

• Widely used software for Decision

Trees

Clustering

• What is clustering

• Types of clustering

• K-means clustering

• Measures of homogeneity

• Data prep

• Hierarchical clustering

• Cluster evaluation

• Cluster profiling

• When to use

• Important considerations

• Clustering in SAS - case study on

store clustering

Pitfalls to avoid while Modeling

• Misleading patterns

• Biased population

• Data at wrong level

• Already known insights

• Un-actionable insights

SAS

• Overview of SAS

• SAS Set Up

• DATA and PROC

• PDV and SAS Processing Rules

• SAS Syntax

• Reading Data into SAS

• Exporting Data from SAS

• Types of Variables

• Informat and Format options

• Data Manipulation in SAS

• Creating Variables in SAS

• SAS Functions

– Numeric

– Character

– Date

• SAS Procedures

– CONTENTS

– PRINT

– SORT

– FREQ

– MEANS

– UNIVARIATE

– CORR

– PLOT

– TRANSPOSE

– REG

– CLUSTER

• SAS Programming Basics

– If Then Else

– DO Loop

– Where

• SAS Best Practices

Excel

• Introduction to Excel

– Overview of Excel

– Basics

– Copy/Paste

– Formatting Cells

– Worksheets

– Formatting Worksheets

– Print

• Sort

• Filter

• Formulas

• Functions

– Math

– Text

– Statistical

– Date

• If statement

• Lookup

• Pivot Tables

• Charts

• Formatting Charts

• Conditional Formatting

• Name

• Ranges

• Conditional Sum

• Introduction to Excel Macros

Knowledge Studio

• Features of Knowledge Studio

• Case study using Knowledge Studio

Statistics

• Introduction to statistics

• Summary statistics

– Mean

– Median

– Mode

– Variance

• Probability

• Probability distribution

– Binomial

– Poisson

– Norma

• Hypothesis testing

• T-test

• Chi-square test

• ANOVA

• Non parametric testing

Placement Package

• Resume writing workshop

– How to write a resume for an

analytics role

• Interview prep workshop

– How to prepare for an interview

for an analytics role

• Presentation skills workshop

– How to build and deliver

presentations

– Presentation of analysis

• Other workshops

– Teleconference skills

– Listening skills

– E-mail writing skills

• Talks by industry experts

RETAIL ANALYTICS

• Scope of analytics in retail --sales,

merchandise, inventory, promotion and

supplier.

• Typical data structures in retail

• Analytical techniques with application

using case studies

– Customer segmentation

– Store segmentation

– Customer profitability analysis

– Vendor scorecard

– Category performance analysis

– Customer life time value

– Loyalty analytics

– Churn analysis

– Market basket analysis

– Trip mission analysis

– Promotion effectiveness analysis

– Cross-sell and up-sell

Emerging Trends

• Convergence in global retail

• Category management

• Customer centric retail –why

customization is the new mantra

• Localization

• Multi-channel retailing

• Role of analytics in winning in the

Future