R For Marketing Research and Analytics

Duration: 3 days

Course Overview

This course provides an introduction to marketing research and analytics using R. While the course assumes no particular domain, techniques and examples are provided for a variety of disparate verticals, including banking, retail, and telecommunications. The first day of the course covers an introduction to R and its applicability to marketing analysis, with particular emphasis on translating existing solutions from SAS to R. The second day provides a gentle introduction to marketing analytics topics in R such as brand perception, principal component analysis (PCA), exploratory factor analysis (EFA), multidimensional scaling (MDS). Finally, the last day of the course covers real-world applications such as clustering and classification, market basket analysis, and choice modeling.

Learning Objectives

- Map understanding of data analytics techniques from SAS to R
- Marketing analytics techniques in R
- Real-world applications of marketing analytics in R

Prerequisite

- Familiarity with data analytics in SAS or similar statistical software package
- Basic machine learning concepts

Course Outline

Day 1

What is R? (30 min)

- Comparison to SAS or other statistical packages
- Why R and When R
- Overview of R language

Basics of R (60 min)

- Basic Objects
- Data Frames
- Loading and Saving Data
- Visualizing Data

BREAK (15 min)

Writing R Functions and Sample Regression (60 min)

- Sample functions
- Regression Analysis of Exam Grades

Working with Vectors, Matrices, and Arrays (60 min)

- Scalars, Vectors, Arrays, and Matrices introduction
- Vector operations
- Filtering
- Vector functions
- Vector elements and equality
- Creating Matrices and Arrays
- Matrix Operations
- Higher-dimensional Arrays

Q&A (15 min)

LUNCH (60 min)

Hands-On Exercise (60 min)

- Participants will be asked to create Vectors, Matrices, and Arrays
- Participants will be asked to solve a simple regression problem
- Participants will be asked to translate from SAS to R for basic data objects

Lists and Data Frames (30 min)

- Creating Lists and Data Frames
- List and Data Frame Operations
- Accessing List and Data Frame elements
- Functions on Lists and Data Frames

BREAK (15 min)

Statistical Analysis and Linear Algebra in R (45 min)

- Functions for Statistical Distributions
- Linear Algebra Operations on Vectors and Matrices
- Set Operations

Hands-on Exercise (30 min)

- Participants will be asked to apply statistical analysis to a dataset
- Participants will be asked to apply linear algebra operations to a dataset
- Participants will visualize the results of their statistical analysis

Day 2

Fundamentals of Data Analysis (90 min)

- Simulating Data
- Functions to Summarize a Variable
- Summarizing Data Frames
- Single Variable Visualization
- Lattice vs ggplot2

BREAK (15 min)

Relationships Between Continuous Variables (120 min)

- Simulating Customer Data
- Simulating Satisfaction Survey Data
- Simulating Non-Response Data
- Scatterplots and Associations Between Variables
- Correlation testing

Q&A (15 min)

LUNCH (60 min)

Hands-On Exercise (60 min)

- Participants will be asked to simulate a customer dataset or use an existing one
- Participants will be asked to do scatterplots between the variables of the dataset
- Participants will be asked to run correlation tests on the dataset

Exploring Associations in Survey Responses (30 min)

- Jitter
- polychoric

BREAK (15 min)

Comparing Groups: Tables and Visualizations (45 min)

- Simulating Consumer Segment Data
- Finding Descriptives by Group

Hands-on Exercise (30 min)

- Participants will be asked to perform association analysis on survey response data
- Participants will be asked to compare groups using descriptive

Day 3

Comparing Groups: Statistical Tests (60 min)

- Chi-square testing
- Binomial testing and Confidence intervals
- P and T-testing
- ANOVA

BREAK (15 min)

Linear Models and Regression (30 min)

- Fitting linear models with lm
- Fitting linear models with multiple predictors
- Overfitting

Reducing Data Complexity (90 min)

- Brand perception and rescaling data
- Principal Component Analysis (PCA)
- Exploratory Factor Analysis (EFA)
- Multidimensional Scaling (MDS)
- Collinearity

Q&A (15 min)

LUNCH (60 min)

Customer Segmentation (60 min)

- Clustering using kmeans and other techniques
- Classification using naïve Bayesian and random Forest
- Identifying Potential Customers

BREAK (15 min)

Market Basket Analysis and Choice Modeling (75 min)

- Association Rules
- Non-Transactional Data
- Choice Modeling
- Customer Heterogeneity

Hands-on Exercise (30 min)

- Participants will be asked to create a customer segmentation model from simulated or existing data