

SAS Predictive Modeling Certification Training program

Enterprise Guide: ANOVA Regression and Logistic Regression

Course Description

This course focuses on the following key areas: statistical inference, analysis of variance, multiple regression, categorical data analysis, and logistic regression. You learn to construct graphs to explore and summarize data, construct confidence intervals for means, test hypotheses, apply multiple comparison techniques in ANOVA, assess and correct collinearity in multiple regression, use diagnostic statistics to identify potential outliers in multiple regression, use chi-square statistics to detect associations among categorical variables, and fit a multiple logistic regression model.

Course Contents

Introduction to Statistics

- examining data distributions
- obtaining and interpreting sample statistics
- describing categorical data

Inferences for the Mean

- constructing confidence intervals
- performing simple tests of hypothesis

Analysis of Variance

- performing one-way ANOVA
- performing multiple comparisons
- performing two-way ANOVA with and without interactions

Regression

- producing scatter plots
- producing correlations
- fitting a simple linear regression model
- understanding the concepts of multiple regression
- building and interpreting models

Regression Diagnostics

- examining residuals
- investigating influence and co linearity

Categorical Data Analysis

- producing frequency tables
- examining tests for general and linear association

- understanding the concepts of logistic regression

Applied Analytics Using SAS Enterprise Miner

Course Contents

Introduction

- touring SAS Enterprise Miner 5.3
- placing SAS Enterprise Miner in the analysis workflow
- application examples and case studies
- pattern discovery and predictive modeling

Accessing and Assaying Prepared Data

- defining a SAS Enterprise Miner 5.3 project
- defining a data source
- validating source data

Introduction to Predictive Modeling with Decision Trees

- defining a modeling data source
- partitioning data for model development
- growing a decision tree with the Desktop Tree Application
- running the Decision Tree node
- using Decision Tree node options
- understanding predictive modeling results data

Introduction to Predictive Modeling with Regressions

- comparing linear and logistic regression
- using the Regression node
- imputing missing values with the Impute node
- replacing data values with the Replacement node
- performing input selection
- understanding regression modeling output
- extending regression models with polynomial and interaction terms

Introduction to Predictive Modeling with Neural Networks and Other Modeling Tools

- introduction to neural network (multilayer perceptron) models
- using the Neural Network node
- performing model selection with the AutoNeural node (self-study)
- other SAS Enterprise Miner 5.3 modeling tools (self-study)

Model Assessment

- defining a prior vector
- effects of prior vectors on model development
- changing model selection criteria
- defining a profit matrix
- comparing models with the Model Comparison node
- introduction to model assessment statistics

Model Implementation

- defining a Score data set
- scoring a data set with the Score node
- using a SAS Code node to export scored data
- generating and using score code

Introduction to Pattern Discovery

- clustering and segmenting data
- using the Transform Variables node
- using the Clustering node for k-means cluster analysis
- applying association and sequence discovery (self-study)
- using the Associations node in a consumer banking example (self-study)
- quantifying the associations among items (self-study)
- exploring sequences among items (self-study)

Special Topics (Self-Study)

- reducing irrelevancy
- reducing redundancy
-
- combining models with the Ensemble node
- consolidating categorical inputs with the Decision Tree node
- explaining complex models with the Decision Tree node

Case Studies (Self-Study)

- segmenting bank customer transaction histories
 - association analysis on Web services data
 - creating a simple risk model from consumer loan data
 - predicting university enrollment
 - creating a response model for an insurance product
-