

Course Syllabus
STAT656 - Applied Analytics Using SAS Enterprise Miner
Spring 2012

Applied Analytics Using SAS Enterprise Miner. Credit 3. Introduction to data mining and will demonstrate the procedures; Optimal prediction decisions; comparing and deploying predictive models; neural networks; constructing and adjusting tree models; the construction and evaluation of multi-stage models. Prerequisite: Knowledge of multiple linear regression and logistic regression. ~~**Starting Spring 2012, students must have a Windows 64bit computer when enrolling in this course.~~

Instructor Information

Instructors: Mike Speed, Ph.D; Professor of Statistics
Edward Jones, Ph.D; Exec. Prof. & Dir. Statistical Consulting

Office Locations: 416B Blocker (Speed)
510D Blocker (Jones)

Telephone: (979) 845-3182 (Speed)
(979) 845-3151 (Jones)

E-mails: mspeed@stat.tamu.edu
ejones@stat.tamu.edu

Office Hours: Dr. Speed: TR 11:00am – 12:00pm or by appointment
Dr. Jones:

Course Identification

Course Number: STAT656

Course Name: ***Applied Analytics Using SAS Enterprise Miner***

Course Location: 457 Blocker

Class Times: 9:35 -10:50 TR

Prerequisites: 608 or 652

Course Description/Overview

This course is an introduction to the general concepts associated with Data Mining and Analytics Modeling. Data Mining is the modeling and analysis of data, usually very large datasets, for decision making. Although several software packages used for Data Mining will be reviewed and compared, the primary concepts will be illustrated using SAS Enterprise Miner.

Course Learning Objectives

By the end of this course will students will be able to:

- Identify data mining models
- Apply SAS Enterprise Miner to Develop Data Mining Models
- Identify appropriate data mining problems involving forecasting versus classification
- Identify appropriate use of multivariate regression, logistic regression, neural networks, cluster analysis and decision trees.
- Able to describe applications of text mining, social network analysis, web analytics and association rules.

Course Resources

Course Website(s)

- Course <<http://dostat.tamu.edu>>
- Personal Website <<http://www.stat.tamu.edu/~mspeed>>

Required Course Text

- Data Mining Techniques, 3rd Edition, by Linoff & Berry, ©2011 Wiley Publishing Inc., ISBN 978-0-470-65093-6

Course Software

SAS Ver9.3 & SAS Enterprise Miner – provided by CIS

Grading Scheme

Grading System

<i>Letter Grade</i>	<i>Percentage</i>	<i>Grade points/credit</i>	<i>Rating</i>
A	90% & above	4.00	Excellent
B	80% – 89%	3.00	Good
C	70% – 79%	2.00	Average
D	60% - 69%	1.00	Below Average
F	59% and below	0.00	Failure
I	Incomplete; given only when a student is unable to complete a segment of the course because of circumstances beyond the student's control. A grade of incomplete may be given only when approved in writing by the department chair or school dean.		

Grading Policy

Grades will be based on the following:

Homework	400
Exams (3 exams x 300 points each)	900
Class attendance/participation	300
Total Points	1600

Late Assignments

Late assignments are not accepted except under unusual circumstances such as illness or family emergency. Makeup examinations are not given. If an exam is missed, the points allocated for that exam (300) are added to the points allocated for the final exam. For example, currently the final exam is allocated 300 points. If one of the other exams is missed, the points allocated for the final exam would be increased to 600 points.

Course Policies

All students are expected to attend and participate in all classes. This requires students to prepare for class in advance. This includes completely assigned homework and reviewing the reading assignment for that class.

Collaboration/Plagiarism Rules

Collaboration on course assignments and homework is encouraged. However, plagiarism on assignments is not allowed and can result in a final grade of "F." Exams are taken individually and collaboration on exams is not allowed. Any collaboration on exams or plagiarism on exams or assignments is not allowed and can result in a failing grade.

Cell phones, Blackberries, iPods, or any other electronic communication devices are not to be used in the classroom. Laptops and iTables can be used in class taking notes.

University Policies

Academic regulations and procedures are governed by University policy. Academic dishonesty cases will be handled in accordance the University's policies.

Academic Integrity & Student Rules

<http://student-rules.tamu.edu/>

Aggie Code of Conduct:

<http://compliance.tamu.edu/CodeConduct.aspx>

Aggie Honor System:

<http://aggiehonor.tamu.edu/>

Aggie Core Values:

<http://www.tamu.edu/about/coreValues.html>

If you have a disability that could affect your performance in this class or that requires an accommodation under the Americans with Disabilities Act, please see me as soon as possible so that we can make appropriate arrangements.

Student Disability Services:

<http://disability.tamu.edu/>

Course Schedule

Week 1: Jan 17-20

Course introduction – Chapter 1 (Speed/Jones)

Introduction to SAS Enterprise Miner

Monday, Jan 18 MLK Holiday

Week 2: Jan 23-27

Course introduction – Chapter 2 (Speed)

Introduction to SAS Enterprise Miner

Monday, Jan 23 Last day to add/drop full semester courses

Week 3: Jan 30-Feb3

Course Introduction – Chapters 4-5 (Speed)

The Data Mining Process

SAS Enterprise Miner Modeling

Week 4: Feb 6-10

Forecasting vs. Classification Problems and Techniques (Jones)

Overview of SAS Techniques for Forecasting & Classification

Week 5: Feb 13-17

Chapter 6

Forecasting & Modeling Using Multivariate Regression (Speed)

Week 6: Feb 20-24

Chapter 6

Logistic Regression in SAS (Speed)

Week 7: Feb 28-Mar 2

Chapter 6

Applications Multivariate Regression in SAS (Speed)

Exam #1 (Chapters 1-6: Introduction & Regression)

Week 8: Mar 5-9

Chapter 8

Forecasting Neural Networks (Jones)

Spring Break – March 12-16

Week 9: Mar 19-23

Classification using Neural Networks in SAS (Jones)

Week 10: Mar 26-30

Validation of Neural Networks in SAS (Jones)

Exam #2 (Chapter 8: Neural Networks & Validation)

Week 11: Apr 2-5

Chapter 9

Cluster Analysis (Speed)

Week 12: Apr 9-13

Chapter 13

Cluster Analysis in SAS Enterprise Miner (Speed)

Monday Apr 2, Last Day for "Q" Drop

Friday, Apr 6, Reading Day-No Classes

Week 13: Apr 26-30

Chapter 7

Decision Trees (Jones)

Week 14: Apr 23-27

Application of Decision Trees in SAS (Jones)

Week 15: Apr 30-May 3

Overview of Other Data Mining Methodologies (Speed/Jones)

Association Rules & Text Mining, Social Network Analysis, & Web Mining

Finals Week May 4, 7-9

TDB