INSTRUCTOR INFORMATION
Instructors: Elizabeth Kolodziej, Ph.D; Senior Lecturer & Dir. of AP Summer Institute
Edward Jones, Ph.D; Exec. Prof. & Dir. Statistical Consulting
Office Locations: 446 Blocker (Kolodziej)
510D Blocker (Jones)
Telephone: (979) 862-1393 (Kolodziej – no voice mail)
(979) 845-3141 (Jones)
E-mails: eykolo@stat.tamu.edu
ejones@stat.tamu.edu
Office Hours: Dr. Kolodziej: 1:30-2:30 pm or any time my door is open
Dr. Jones: Drop in anytime my door is open or by appointment

COURSE IDENTIFICATION
Course Number: STAT 656
Course Name: Applied Analytics
Course Location: 457 Blocker for On Campus Section
Class Times: 3:00 – 3:50 pm    MWF: On Campus Section
Class Times: 6:00 – 7:00 pm    W: Optional Q&A, Especially Distance
Prerequisites: 608 or 652, plus some multiple linear regression and logistic regression

COURSE DESCRIPTION/OVERVIEW
This course is an introduction to the general concepts and methodologies 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. Models discussed include neural networks; multiple and logistic regression; decision trees; and clustering algorithms. Standards for making decisions about optimal models will also be discussed.

COURSE LEARNING OBJECTIVES
By the end of this course will students will be able to:
• Investigate data sets, identifying appropriate transformations, creating new variables, and interpolating missing values
• Identify and fit appropriate data mining models, such as multivariate regression, logistic regression, neural networks, cluster analysis and decision trees
• Interpret data mining models in context
• Distinguish between data mining problems involving forecasting and classification
• Assess data mining models for usefulness, predictive value, and financial gain

COURSE RESOURCES
Course Website(s)
• Homework: WebAssign: www.webassign.com
• Course Notes, Data Files, Discussion board: DoStat: dostat.tamu.edu/dostat
Course Texts:
- Optional: Getting Started with SAS Enterprise Miner 7.1 ©2011 SAS Institute Inc. ($7.95 from Amazon)

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

GRADING SCHEME
Grading Scale

<table>
<thead>
<tr>
<th>Letter Grade</th>
<th>Percentage</th>
<th>Grade points/credit</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>90% and above</td>
<td>4.00</td>
<td>Excellent</td>
</tr>
<tr>
<td>B</td>
<td>80% – 89%</td>
<td>3.00</td>
<td>Good</td>
</tr>
<tr>
<td>C</td>
<td>70% – 79%</td>
<td>2.00</td>
<td>Average</td>
</tr>
<tr>
<td>D</td>
<td>60% - 69%</td>
<td>1.00</td>
<td>Below Average</td>
</tr>
<tr>
<td>F</td>
<td>59% and below</td>
<td>0.00</td>
<td>Failure</td>
</tr>
<tr>
<td>I</td>
<td>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.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

GRADING SCHEME (continued)

Grading Policy: Grades will be based on the following:

<table>
<thead>
<tr>
<th></th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homework</td>
<td>15%</td>
</tr>
<tr>
<td>Participation</td>
<td>5%</td>
</tr>
<tr>
<td>Exams (2 exams x 25% each)</td>
<td>50%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>30%</td>
</tr>
<tr>
<td><strong>Total Points</strong></td>
<td><strong>100%</strong></td>
</tr>
</tbody>
</table>

Participation: Each student should either ask or answer a question on the discussion board every other week to receive full credit. Whether you’re asking questions or teaching someone else, you’re doing some learning.

Late Assignments: Late assignments are not accepted. To allow for University excused missing assignments, the lowest homework grade will be dropped. Makeup examinations are not given. If an exam is missed, the percentage grade allocated for that exam (25%) is added to the points allocated for the final exam. For example, currently the final exam is allocated 30%. If one of the other exams is missed, the grade allocated for the final exam would be increased to 55%.

Course Policies: All students are expected to attend and participate in all classes. This requires students to prepare for class in advance. This includes completing assigned homework and reviewing the reading assignment for that class. Class participation is participation on the discussion board in DoStat. Students are expected to either ask a question or answer another student’s question at least once every other week for full credit on class participation.

Any student caught cheating on an exam will receive a grade of 0 on that exam, and violations will be filed with the University.
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 tablets can be used in class taking notes. Be aware that other students behind you may be distracted by your screen.

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 Core Values: http://www.tamu.edu/about/coreValues.html
Aggie Honor System: “An Aggie does not lie, cheat, steal, or tolerate those who do.” Academic dishonesty includes the commission of any of the following acts: cheating, fabrication / falsification, multiple submissions, plagiarism, complicity, and other types of misconduct. For more details on the Aggie Honor system, please see: http://aggiehonor.tamu.edu/.

- Cheating is intentionally using or attempting to use unauthorized materials, information, notes, study aids or other devices or materials in any academic exercise. Unauthorized materials may include anything or anyone that gives a student assistance and has not been specifically approved in advance by the instructor. During an examination, for example, looking at another student’s examination or having a conversation with others is cheating unless specifically allowed in advance by the instructor.
- Fabrication is making up data or results, and recording or reporting them; or submitting fabricated documents. Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record. The intentional invention and unauthorized alteration of any information or citation in any academic exercise; failing to acknowledge the actual source from which cited information was obtained; changing information on tests, quizzes, examinations, reports, or any other material that has been graded and resubmitting it as original for the purpose of improving the grade on that material are examples of fabrication and falsification.
- Complicity is intentionally or knowingly helping, or attempting to help, another to commit an act of academic dishonesty. Examples include knowingly allowing another to copy from one’s paper during an examination or test; distributing test questions or substantive information about the test without the instructor’s permission; collaborating on academic work knowing that the collaboration will not be reported; taking an examination or test for another student; and conspiring or agreeing with one or more persons to commit, or attempt to commit, any act of scholastic dishonesty.

Disability Services: The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you believe you have a disability requiring an accommodation, please contact Disability Services, in Cain Hall, Room B118, or call 845-1637. For additional information visit http://disability.tamu.edu.

If you have a disability requiring an accommodation, please contact me as soon as possible so that we can make appropriate arrangements.

COURSE SCHEDULE (subject to change)
Week 1: Jan 14-18
- Course introduction – Chapters 1, 3 (Kolodziej / Jones)
- Introduction to SAS Enterprise Miner & The Data Mining Process
- Friday, Jan 18 Last day to add/drop full semester courses

Week 2: Jan 21-25
- NOTE: Monday, Jan 21 is MLK Day – No Classes
- S: Sample and E: Explore – Chapter 2 (Kolodziej)
- Introduction to SAS Enterprise Miner

Week 3: Jan 28-Feb 1
Week 4: Feb 4-8
M: Model - Chapters 6, 20 (Kolodziej)
Forecasting & Modeling Using Multivariate Regression
Variable Selection Procedures in SAS

Week 5: Feb 11-15
M: Model - Chapters 5 - 6
Applications Multivariate Regression in SAS (Kolodziej)

Week 6: Feb 18-22
Logistic Regression (Kolodziej)
A: Assess – Multivariate and Logistic Regression
Exam #1 (Chapters 1-6: Introduction & Regression) Friday, February 22

Week 7: Feb 25-Mar 1
M: Model - Chapter 8
Introduction to Neural Networks for Forecasting (Jones)
Enterprise Miner for Neural Network Forecasting (Jones)

Week 8: Mar 5-8
M: Model - Introduction to Neural Networks for Classification (Jones)
A: Assess: Validation of Neural Networks (Jones)
Spring Break – March 11-15

Week 9: Mar 18-22
M: Model – Model Selection Techniques (Jones)
Exam #2 (Chapter 8: Neural Networks & Validation) Friday, March 22

Week 10: Mar 25-29
M: Model - Chapter 13
Cluster Analysis (Kolodziej)
Friday, Mar 29, Reading Day-No Classes

Week 11: Apr 1-5
Cluster Analysis in Enterprise Miner (Kolodziej)
Tuesday Apr 2, Last Day for “Q” Drop

Week 12: Apr 8-12
M: Model - Chapter 7
Decision Trees (Jones)

Week 13: Apr 15-19
M: Model - Application of Decision Trees in SAS Enterprise Miner (Jones)
A: Assess – Decision Trees (Jones)

Week 14: Apr 22-26
A: Assess – Decision Costs – Profit & Loss (Jones)

Week 15: Apr 29-30 – Course Review
Tuesday Apr 30, Our last day of class

Finals Week May 3, 6-8
Final Exam: Tuesday May 7, 10:30 am - 12:30 pm

EXAM DATES:
- Exam 1: Friday, February 22 - 50 minutes
- Exam 2: Friday, March 22 - 50 minutes
- Final Exam: Tuesday, May 7: 10:30AM - 12:30pm

NOTE: THIS SYLLABUS CAN BE CHANGED WITHOUT ADVANCED NOTICE.