

PREDICTIVE ANAYTICS AND DATA MINING (MKTG 630)

Course Syllabus – 2015 Spring Semester

*Radford University
College of Business and Economics
Department of Marketing*

- INSTRUCTOR:** Dr. Wil Stanton
Professor of Marketing
- OFFICE:** **COBE Building -- BE374**
Tel: 540.831.5015 (Office)
540.633.2292 (Home)
Fax: 540.831.6438
E-mail: wstanton@radford.edu
 twitter.com/AnalyticsProfs
 <http://www.linkedin.com/in/drwilstanton>
- OFFICE HOURS:** Tuesday and Thursday, 9:00 AM – 10:30 am. I may also be around outside of my scheduled office hours. If you need to see me and you can't make it during my scheduled office hours, contact me and so a time can be set up for us to meet (or if you're in the neighborhood, just stop by -- if I'm in, my door is usually open).
- TIME/PLACE:** **Tuesday and Thursday**, 5:00 – 6:15 p.m. in either COBE 215 (Lab) or COBE 204
- Required TEXT:** Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management - third edition by Gordon S. Linoff and Michael J. A. Berry, Wiley (April 12, 2011) - **ISBN-10:** 0470650931 **ISBN-13:** 978-0470650936 About \$28 at Amazon
- Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die by Eric Siegel, Wiley (Feb 19, 2013) - **ISBN-10:** 1118356853 **ISBN-13:** 978-1118356852
- EXPECTATIONS:** Students enrolling in MKTG 630 must be able to:
- Use basic information technology tools (e.g., spreadsheets, word processors, web browsers, e-mail readers, presentation packages, etc.);
 - Demonstrate an ability and a willingness to learn the course content;
 - Behave in a professional manner; and
 - Work effectively with other class members.
- PREREQUISITES:** Graduate Standing, familiarity with basic business statistical tools, and of course, tenacity, and a desire to learn.

FOUNDATIONS FOR THIS COURSE

This course supports the mission, vision and learning outcomes sought for all MBA students in the College of Business and Economics at Radford University:

Vision and Mission of the College of Business and Economics:

Our vision is to be recognized for challenging minds, cultivating talents and connecting people in a technology-rich learning environment.

Our mission is to provide an active learning environment that develops analytical and innovative business professionals for a dynamic global economy.

Employer Expectations: This course also supports the needs and expectation of employers when hiring college graduates.¹ Specifically this course will focus on the needs of employers in the following areas:

Intellectual and practical skills, specifically

- The ability to communicate effectively, orally and in writing
- Critical thinking and analytical reasoning skills
- The ability to analyze and solve complex problems
- Teamwork skills and the ability to collaborate with others in diverse group settings
- The ability to innovate and be creative
- The ability to locate, organize, and evaluate information from multiple sources
- The ability to work with numbers and understand statistics

BACKGROUND FOR THE COURSE

Data is the unstoppable train barreling down on businesses. Companies are witnessing an exponential growth in data and find they are ill prepared to turn the data in to meaningful information for management decision making. In 2009, an estimated 800,000 petabytes of data existed worldwide, but today, there is 1.2 trillion gigabytes of digital data – and the amount of data is growing exponentially. With so much raw data, organizations urgently need tools and employees who know how to use them to effectively and efficiently extract actionable information to help optimize business decisions.

But, Big Data alone nets a company nothing. It's how the data are mined, models built, and predictions made that separate the truly successful world-class companies from the wannabes.

Analytics to the rescue!! Analytics is both an art and a science to discover and understand historical patterns in a company's data in order to predict and improve business performance under forecasted environmental, economic, and competitive conditions.

Companies want employees that understand the business domain in which decisions are to be made and depth and breadth of understanding of the analytical tools necessary for decision optimization. But the reality is: the demand for individuals grounded in Analytics, particularly in Data Mining and Predictive Analytics, far exceeds the supply of graduates.

The U.S. Bureau of Labor Statistics predicts that there will be a 24 percent increase in demand for professionals with management analysis skills over the next eight years; and McKinsey Global Institute Predicts there will be a shortage of talent necessary for organizations to take advantage of Big Data. By 2018, the United States alone could face a shortage of 140,000 to 190,000 people with deep analytical skills as well as 1.5 million managers and analysts with the know-how to use the analysis of big data to make effective decisions

¹ "Employers' Views On College Learning In The Wake Of The Economic Downturn: A Survey Among Employers," Conducted On Behalf Of The Association Of American Colleges And Universities, Hart Research Associates, January 2011.

NEED FOR THE COURSE

The growing global utilization of Big Data to assist in strategically positioning corporations for competitive advantage results in an expectation that a new employee holding an MBA be able to apply predictive analytics and data mining tools and techniques to solve strategic business problems. At no time has there been a greater need for quantitatively skilled and analytically minded managerial expertise. This need is being evidenced in a transformation of MBA programs across the country and around the globe. Universities are beginning to offer graduate courses or entire MBA programs focused on preparing graduates fully grounded in business principles, but who also possess the analytical skills to develop better decision models and create more accurate predictions from the stores of Big Data. Companies are collecting more data than ever before. In the past they were kept in different systems that were unable to talk to each other, such as finance, human resources or customer management. Now the systems are being linked, and companies are using Data Mining techniques to get a complete picture of their operations—"a single version of the truth", as the industry likes to call it. Data Mining and Predictive Analytics is the process of finding patterns in data that can be used to help optimize business decisions; and is becoming the bedrock of a new generation of companies and the individuals who lead them. Data Mining and Predictive Analytics has come to be essential to preparing new MBA students to assume leadership roles in an emerging data driven decision making environments.

Predictive Analytics and Data Mining. Predictive analytics and data mining can be used to seek out increasingly small patterns in the data (patterns that could not have been seen just a few years ago) to better understand a company's products, channels, partners, customers (down to the behaviors of an individual customer), and more.

GOALS AND OBJECTIVES OF THIS COURSE

In completing this course students will be able to:

- recommend the appropriate Predictive Analytics and Data Mining techniques for a variety of business decision problems;
- apply the processes of Predictive Analytics and Data Mining for formulating business objectives, data selection, preparation, and partition to successfully design, build, evaluate and implement analytic models for a variety of practical business applications;
- analyze large datasets typical in today's corporate setting using IBM SPSS, SPSS Modeler, SAS, SAS Enterprise Guide, and SAS Enterprise Miner advanced Data Mining software;
- apply predictive models such as classification and decision trees, neural networks, regressions, association analysis, and link analysis, to typical corporate Big Data;
- Interpret analyses produced by advanced analytical procedures and explain the results to better inform management decision-making.
- apply Knowledge Discovery processes across all stages of data mining including domain understanding, data collection and selection, data cleaning and transformation, dimensionality reduction, pattern discovery, evaluation, and knowledge extraction;
- apply data mining techniques to segmentation, classification, predictive modeling, association analysis, and sequential pattern discovery;
- gain hands-on experience with industry standard predictive analytic and data mining software including SPSS, SPSS Modeler, SAS, SAS Enterprise Guide, and SAS Enterprise Miner;
- determine which predictive analytics and data mining tool to use as well as understand the similarities and differences, which options affect the models most and how to verify and validate models;

- acquire skills in text analytics, text mining and sentiment analytics by understanding the "sentiment" conveyed in the vast and ever increasing amount of unstructured, user generated text from sources such as online reviews, tweets, forums, blogs, and internal customer data;
- develop and apply ensemble methods to leverage the power of multiple, complementary models to deliver additional business insight to ensure long term competitiveness;
- explore the applications of predictive analytics and data mining through a "hands-on" process where you will develop understanding mainly through conducting application projects and presenting results in such areas as: advertising/media; banking; customer acquisition and CRM; education; finance; fraud detection; healthcare; law enforcement; personnel/HR; political/politics; real estate; retail; sports; and web/Internet shopping.

Background readings from the two required texts, contemporary articles from industry leaders, academic journal articles, PowerPoint presentations, videos, YouTube tutorials, SAS and SPSS materials, and hand-on utilization of software will provide the foundational knowledge of the various Predictive Analytics and Data Mining techniques covered this semester.

ANALYTICAL TECHNIQUES COVERED

This semester we will explore various analytical techniques applied to predictive analytics and data mining. We will cover a few preliminaries before moving aggressively into Data Mining (Exploratory Data Analysis) - EDA and Predictive Modeling - PM. A Data scientist/Decision Scientist must be fully grounded in the tools, techniques, assumptions, identification of violations of assumptions, consequences of violating assumption, etc. BEFORE beginning any Data Mining and/or Predictive Modeling effort - particularly if the effort is focused on Big Data Applications. So we will begin with a review of some basic statistical tools and then move into more advanced and more complex algorithmic processes. Predictive Analytics and Data Mining are not as simple as introductory statistics courses may have led you to believe, i.e., EDA and PM cannot be accomplished by using only t-test, correlation, and regression analyses - and normality rarely exist. The techniques listed here are the ones planned for the semester, but the actual ones covered will depend on time, your abilities, and your desire to explore techniques specifically focused on your career goals.

However, Data Mining and Predictive Analytics are only part of the solution. Analytic insights gained from EDA and PM must be put to work and used properly to aid in optimizing business decisions. This course covers Predictive Analytics and Data Mining methods for both private and public sectors enterprises, but, as this is an MBA course, the analytic procedures covered will be used as a means to an end; with the end being better and more efficient decision making.

- | | |
|--|--|
| • Descriptive Statistics | • Factor and Discriminant Mapping |
| • Cross-Tabulation | • Classification and Regression Trees, e.g. CHAID |
| • Correspondence Analysis | • Cumulative Gains and Lift Charts and Other Visual Aids For Measuring Model Performance |
| • Correlation | • Survival Analysis, e.g., Cox Regression Models |
| • Canonical-Correlation Analysis | • Market Basket Analysis |
| • Factor Analysis and Principal Components Analysis | • Neural Networks |
| • Cluster Analysis | |
| • Regression Techniques Including Linear and Non-Linear Models | |
| • Logistic Regression | |
| • Discriminant Analysis | |

TEACHING METHODS

This is an MBA course designed to assist you in effectively contributing to your organization as a business professional. Of course, as a student in Radford University's College of Business and Economics' MBA program, you must be a responsible student. Therefore, I will expect you to come to class having completed all required assignments and fully prepared to actively engage in your education and the education of others. I will expect you to talk about what you are learning, write about it, and relate it to past experiences.

Everything we do in this class is designed to help you become a graduate qualified business professional; helping you think like a graduate qualified business professional. I will not ask you to simply memorize information and then recite it back in class or on an exam. Instead, I will expect you to internalize the principles, concepts, and theories of Analytics and then use those actively in every class and in every class assignment. Because not all students learn in the same ways, I will utilize a variety of teaching methods including lectures, case studies, readings, class discussion, assignments, and exams. I will emphasize critical thinking and analytical reasoning in all materials covered; and will expect a professional level of oral and written communication from you. As this is an MBA class, I will use Socratic questioning to foster learning; and more importantly to help you learn about the exponentially growing Analytics world and its impact on global industry. Thinking is driven not by answers but by questions, so I will focus on tools to help you reason through Predictive Analytics and Data Mining issues and problems demonstrating that the subject matter of this course is highly interrelated with that of other business subjects.

COURSE SOFTWARE

In this course you will be using IBM SPSS Statistics, SPSS Modeler, Base SAS, SAS Enterprise Guide, and SAS Enterprise Miner. These will all be available in the lab in BE 215. You will learn how to use these software packages and how to access your shared drive (the H drive) on Windows-based machines as this is what is provided in our classroom lab. Citrix can be used on your personal computer so that you can use most of these packages on your personal computer. Citrix can generally be used on Macs but, depending on your operating system, may require additional support from IT to run.

TEXTS USED OR CONSULTED IN DEVELOPING COURSE MATERIALS

In addition to the texts below, I may use materials from other texts as the semester progresses as the materials supporting predictive analytics and data mining are growing exponentially just as the subject area.

I wanted to have a solid and current foundation for this course, Predictive Analytics and Data Mining, so in addition to materials made available by IBM SPSS, IBM SPSS Modeler, SAS, SAS Enterprise Guide and SAS Enterprise Miner, I consulted a number of texts in preparing the materials to be covered. Among the texts used in developing this course included but was not limited to:

- [Analytics at Work: Smarter Decisions, Better Results](#) by Thomas H. Davenport, Jeanne G. Harris, Robert Morison, Harvard Business School Publishing Corporation (2010)
- [Applied Logistic Regression Analysis](#) by Scott Menard, Sage (1995)
- [Applied Predictive Modeling](#) by Max Kuhn and Kjell Johnson, Springer (2013)
- [Business Analytics for Managers: Taking Business Intelligence Beyond Reporting](#) by Gert H. N. Laursen and Jesper Thorlund, John Wiley & Sons, Inc. (2010)
- [Business Analytics for Sales and Marketing Managers: How to Compete in the Information Age](#) by Gert H. N. Laursen, John Wiley & Sons, Inc. (2011)
- [Business Intelligence: A Managerial Approach](#) (2nd edition) by Efraim Turban, Ramesh Sharda, Jay E. Aronson, and David King, Pearson – Prentice Hall (2011)

- Business Modeling and Data Mining by Dorian Pyle, The Morgan Kaufmann Series in Data Management Systems, Elsevier, Inc. (2003)
- Causal Analysis with Panel Data by Steven E. Finkel, Sage (1995)
- Competing on Analytics: The New Science of Winning by Thomas H. Davenport and Jeanne G. Harris, Harvard Business School Publishing Corporation (2007)
- Cult of Analytics: Driving Online Marketing Strategies Using Web Analytics by Steve Jackson, Elsevier, Inc. (2009)
- Customer Segmentation and Clustering Using SAS Enterprise Miner, second edition by Randall S. Collica, SAS (2011)
- Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis by Collen McCue, Butterworth-Heinemann (2006)
- Data Mining Explained: A Manager's Guide to Customer-Centric Business Intelligence by Rhonda Delmater and Monte Hancock, Digital Press, Butterworth-Heinemann (2001)
- Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management - third edition by Gordon S. Linoff and Michael J. A. Berry, Wiley (April 12, 2011)
- Data Mining: Concepts and Techniques, Third Edition by Jiawei Han, Micheline Jamber, and Jian Pei, The Morgan Kaufmann Series in Data Management Systems, Elsevier, Inc. (2012)
- Data Science for Business: What you need to know about data mining and data-analytic thinking by Foster Provost and Tom Fawcett O'Reilly Media (Aug 16, 2013)
- Data-Driven Marketing: The 15 Metrics Everyone in Marketing Should Know by Mark Jeffery, John Wiley & Son, Inc. (2010)
- Decision Management Systems: A Practical Guide to Using Business Rules and Predictive Analytics by James Taylor, IBM Press (2012)
- Excel 2010: Data Analysis and Business Modeling by Wayne L. Winston, Microsoft Press (2011)
- Handbook of Statistical Analysis and Data Mining Applications by Robert Nisbet, John elder IV, and Gary Miner, Elsevier (2009)
- IBM Cognos 10 Report Studio by Filip Draskovic and Roger Johnson, IBM Press (2011)
- Keeping Up with the Quants: Your Guide to Understanding and Using Analytics by Thomas H. Davenport and Jinho Kim, Harvard Business Review Press (2013)
- Loglinear Models with Latent Variables by Jacques A. Hagenaars, Sage. (1993)
- Marketing Metrics: The Definitive Guide to Measuring Marketing Performance, 2nd edition by Paul Farris, Neil T. Bendle, Phillip E. Pfeifer, and David J. Reibstein, Prentice Hall (2010)
- Maximum Likelihood Estimation: Logic and Practice by Scott R. Eliason, Sage (1993)
- Mining the Talk: Unlocking the Business Value in Unstructured Information by Scott Spangler, and Jeffrey Kreulen, IBM Press (2007)
- Modeling Techniques in Predictive Analytics: Business Problems and Solutions with R by Thomas W. Miller, Pearson (2013)
- Multivariate Data Analysis, 6th Edition by Joseph F. Hair, Jr., William C. Black, Barry J. Babin, Rolph E. Anderson, and Ronald L. Tatham, Pearson Prentice Hall (2006)
- Multivariate General Linear Models by Fichard F. Haase, Sage (2011)
- Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die by Eric Siegel, Wiley (Feb 19, 2013)

- Predictive Analytics: Unabridged Guide by Russell Nixon, Emero (2013)
- R Cookbook by Paul Teetor, O'Reilly (2011)
- R for Dummies by Andrie de Vries and Joris Meys, Wiley (2012)
- The Art of R Programming: A tour of Statistical Software Design by Norman Matloff, No Starch Press (2011)
- The Handbook of Data Analysis, Edited by Melissa Hardy and Alan Bryman, Sage (2009)
- The Little SAS Book for Enterprise Guide 4.2 by Susan J. Slaughter and Lora D. Delwiche, SAS (2010)
- The Little SAS Book: A Primer, Fifth Edition by Lora D. Delwiche and Susan J. Slaughter
- The New Direct Marketing: How to Implement t Profit-Driven Database Marketing Strategy, 3rd edition by David Shepard Associates, Inc. with contributions by Rajeev Batra, Richard Deere, Andrew Deutch, Robert Kendrick, George Orme, Bruce Rather, Russ Rempala, Dhiraj Sharma, and David Shepard, McGraw-Hill (1998)
- The New Know: Innovation Powered by Analytics by Thornton May, John Wiley & Son, Inc. Wiley & SAS Business Series (2009)
- The Performance Manager: Proven Strategies for Turning Information into Higher Business Performance by R. Mosimann, P. Mosimann, and M. Dussault, Cognos Press (2007)
- The Value of Business Analytics: Identifying the Path to Profitability by Evan Stubbs, John Wiley & Sons (2011)
- Web Analytics 2.0: The Art of Online Accountability & Science of Customer Centricity by Avinash Kaushik, Wiley Publishing (2010)

NOTE: Depending on your career goals, you may want to review some of these texts.

PREDICTIVE ANALYTICS AND DATA MINING INDIVIDUAL/SELF-SELECTED FOCUS

In addition to required and assigned readings, each of you will identify an area of predictive analytics and/or data mining which you believe will aid you in your future career plans, e.g., advertising/media; banking; customer acquisition and CRM; education; finance; fraud detection; healthcare; law enforcement; personnel/HR; political/politics; real estate; retail; sports; web/Internet shopping, etc.. Each week you will be expected to submit a review of a paper you selected. This will constitute a portion of your assignment grade for the course.

The format of each weekly submission of approximately 1 page is as follows:

- The title, author, and source of your paper/article. If the paper/article is accessible via the Internet, you should also provide the URL. If it is not, you should include a scanned copy of the paper/article and submit it with your write-up to D2L.
- The review of the paper/article selected should include:
 - A short abstract of an article/paper in your chosen area(s) of focus; and
 - A short description of what you found most interesting/informative about the paper.

At the end of the semester you will be expected to produce a single summative paper which synthesizes all the readings you submitted during the semester. Thus, I will expect you to use your weekly self-selected articles, white papers, reports, etc. and your weekly submissions to produce your **Final Individual Predictive Analytics and Data Mining Paper** (Final Paper) focused on the area(s) you chose at the beginning of the semester. I've posted the guidelines for this paper to D2L.

The format for both weekly and final paper submission should adhere to the following general style policies:

- I. Your weekly and Final Individual Paper shall be typed following a single spaced, 1" margins, and using Arial 11 pt. font – note: the default in some versions of Word is a spacing of 1.15. Change it to 1 for single space.
- II. Written work must be prepared in a professional manner. I will grade each work based on quality of the writing, thoroughness of the content, and soundness of the analytical reasoning.
- III. Each submission is to be clearly written, with proper grammatical construction and correct spelling. Work which is not presented in a professional manner will be returned as unacceptable. In addition, improper grammar, sentence construction and misspellings may result in a grade reduction.
- IV. Citations shall follow the APA style.
- V. Your Final Paper shall have a bibliography at the end.
- VI. Use graphs, figures and charts as appropriate. General guidelines For the Use of Visuals (Source: APA Guidelines) NOTE: complete style policies are posted on D2L.
- VII. Ensure your Final Paper is organized logically
- VIII. Provide a Table of Contents for your Final Paper
- IX. Use Headings to help the reading of your Final Paper

LINKEDIN: You will be expected to join LinkedIn and connect with me and others in the class. If you have not already done so, you must connect with me no later than the beginning of class on January 29, 2015. I may also create a group on LinkedIn just for the class. We will use LinkedIn to share information, articles, weekly self-selected papers, whitepapers, etc. I will regularly share one or more of the articles/papers found by you with my LinkedIn contacts (about 2,440). You will receive attribution for sharing the articles/papers. In that way you will begin to get broad exposure to individuals in key positions across many industries and geographic locations, including international locations.

I will also ask you to join several LinkedIn groups whose focus is on predictive analytics and data mining. Of course when the semester is complete you can drop out of any group you do not wish to continue in. You will be able to see what industry professionals are contributing and, to the extent you desire, contribute to the discussions.

NOTE: a Handout on best practices for LinkedIn will be provided

READINGS: The course will rely heavily on text, contemporary, and weekly readings to build a foundation for the understanding of Predictive Analytics and Data Mining. Reading will be more intense at the beginning of the semester. Other readings may be assigned throughout the semester based on the need and preparation of students. Any additional readings or changes in the schedule will be posted to D2L.

INDIVIDUAL ASSIGNMENTS: In addition to the readings, you will also be asked to individually complete a number of in-class (Lab) and out-of-class assignments. **NOTE:** All out-of-class assignments (group or individual) are due at the beginning of class on the date specified - **no late assignments or projects will be accepted.** While I encourage you to do any missed assignments for your own learning (and because some assignments build upon earlier assignments), you cannot make-up missed assignments for credit – this includes both out-of-class assignments and in-class assignments/activities. Because of the nature of the course, there will be no additional or extra credit assignments to increase your grade.

Please also note that while I have no problem with students discussing individual homework assignments and helping each other with problems (as this is part of the learning process), the work turned in for credit must be **your** work. Copying another student's work is unacceptable and in violation of the *RU Honor Code*. If there is an honor code violation, I will assign a grade of zero (0) for that assignment and report the honor code violation to the Dean of Students Office.

STUDENT ENGAGEMENT, PARTICIPATION, AND PROFESSIONALISM

You will be expected to be an active contributor to class discussions and in-class assignments, activities, etc. I will keep track of your attendance and participation throughout the semester. I also expect you to be professional in your class behavior (please see the *Professional Behavior* policy in the *Course Policies* for more detail). Your final class engagement grade for the semester will be based on a review of your daily participation and overall professionalism for the semester. Obviously you must be present in class in order to participate. Please keep in mind, however, that quantity in this area does not necessarily equate to quality. You will be evaluated on the quality of your contributions. Please do not assume that simply coming to class ensures you will receive a high grade in this area. Class attendance is a necessary but not sufficient condition to your ability to participate.

In order to reinforce key course concepts, you will be involved in participating in content-related assignments and exercises. Some of these may be assigned as homework; others may be in-class exercises. Some assignments will be individual while others may be done in teams. On assignments where you are allowed to work in teams, I expect all team members to **participate fully** in each and every exercise assigned.

COURSE EVALUATION, GRADING & IMPORTANT DATES: You will be evaluated on your knowledge of Predictive Analytics and Data Mining concepts as well as your ability to apply that knowledge effectively to typical business issues, situations, problems and opportunities. Your performance will be evaluated by several means including: examinations; in- and out-of- class assignments (both group and individual); Final Individual Paper; and your engagement, participation and professionalism in the class. Specifically, the weights assigned to each of these performance measures (as well as the associated dates) are:

- Submitted Assignments including weekly self-selected article (20%)
- Student Engagement, Participation, and Professionalism (5%)
- Mid-Term Exam (25%)
- Final Individual Paper (20%)
- Final Exam (30%)

Your final grade will be strictly determined as follows:

93 – 100	A
90 – 92.99	A-
87 – 89.99	B+
83 – 86.99	B
80 – 82.99	B-
70 – 79.00	C
Below 70	F

SEMESTER SCHEDULE

NOTE: topics, readings and dates are approximate and may change

Calendar for the Semester		
NOTE: All topics and dates may be altered as needed		
Week	Date	Requirement
1	20-Jan	Introduction, Syllabus, and PPT posted in D2L
	22-Jan	Continue with introductory PPT; Discussion of Readings for 22 Jan
2	27-Jan	Measurement; weekly paper due in D2L dropbox before beginning of class
	29-Jan	Choosing the Appropriate Analytical Technique to support Data Mining and Predictive Modeling; Discussion of Readings for 29 Jan
3	3-Feb	Issues with data, e.g., skewness, Kurtosis, and Multicollinearity; weekly paper due in D2L dropbox before beginning of class
	5-Feb	Using software appropriate to Data Mining and Predictive Analytics; homework due in dropbox before beginning of class
4	10-Feb	Data Exploration, Descriptive Statistics, and Frequencies; weekly paper due in D2L dropbox before beginning of class
	12-Feb	Crosstabulation; homework due in dropbox before beginning of class
5	17-Feb	T-Test and ANOVA; weekly paper due in D2L dropbox before beginning of class
	19-Feb	Correlation Analysis and Intro to Factor Analysis; homework due in dropbox before beginning of class
6	24-Feb	Factor Analysis Continued; weekly paper due in D2L dropbox before beginning of class
	26-Feb	Catch up day, Q&A, Lab Work; Distribution of take home mid-term exam which will be due in the dropbox before class on 17 March; homework due in dropbox before beginning of class
7	3-Mar	Cluster Analysis; weekly paper due in D2L dropbox before beginning of class
	5-Mar	Factor Mapping; homework due in dropbox before beginning of class
8	Spring Break 8 March – 16 March	
9	17-Mar	Regression Analysis; Mid-Term Exam due in dropbox before the beginning of class; weekly paper due in D2L dropbox before beginning of class
	19-Mar	Regression Analysis Continued; homework due in dropbox before beginning of class
10	24-Mar	Logistic Regression; weekly paper due in D2L dropbox before beginning of class
	26-Mar	Logistic Regression Analysis Continued; Lift and Gains Charts; homework due in dropbox before beginning of class
11	31-Mar	Discriminant Analysis; weekly paper due in D2L dropbox before beginning of class

	2-Apr	Discriminant Analysis Continued; Discriminant Mapping; homework due in dropbox before beginning of class
12	7-Apr	Classification and Regression Trees; Introduction to Modeler and Miner; weekly paper due in D2L dropbox before beginning of class
	9-Apr	Trees using Modeler and Miner continue; homework due in dropbox before beginning of class
13	14-Apr	Catch Up day; Q&A; Lab Work; weekly paper due in D2L dropbox before beginning of class
	16-Apr	Survival Analysis using Cox Regression; Neural Networks; homework due in dropbox before beginning of class
14	21-Apr	Combining all the techniques; weekly paper due in D2L dropbox before beginning of class; distribute the take home portion of the final exam
	23-Apr	More examples combining techniques
15	28-Apr	Introduction to R programming; weekly paper due in D2L dropbox before beginning of class
	30-Apr	More R programming; readings on R; Lab Work; Final Paper due
	5-May	Take Home portion of the Final exam due before 5:00 p.m.; In-Class portion of the Final Exam 5:00 – 7:00 p.m.

Read by	Readings
22-Jan	<ol style="list-style-type: none"> Schmarzo, "Business Analytics: Moving From Descriptive To Predictive Analytics"(posted on D2L) (6 pages); Guazzelli, "Predicting the future, Part 1: What is predictive analytics?" (posted on D2L) (9 pages); Guazzelli, Predicting the future, Part 2: What is predictive analytics?" (posted on D2L) (11 pages); Kantardzic, Data Mining: Concepts, Models, Methods, and Algorithms, Second Edition, Appendix B (posted on D2L) (14 pages); Linoff & Berry, Introduction, (3 pages); Linoff & Berry, Chapter 1 What Is Data Mining and Why Do It?, (26 pages); Siegel, Foreword Thomas H. Davenport xiii, (2 pages); Siegel, Preface xv, (2 pages); Siegel, Introduction: The Prediction Effect 1, (15 pages);
27-Jan	<ol style="list-style-type: none"> Berson, Smith and Thearling, "An Overview of Data Mining Techniques" (posted on D2L) (46 pages); Linoff & Berry, Chapter 2 Data Mining Applications in Marketing and Customer Relationship Management, (40 pages); Linoff & Berry, Chapter 3 The Data Mining Process., (34 pages); Siegel, Chapter 1: Liftoff! Prediction Takes Action (deployment), (20 pages); Siegel, Chapter 2: With Power Comes Responsibility: Hewlett-Packard, Target, and the Police Deduce Your Secrets (ethics), (30 pages);

Read by	Readings
3-Feb	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 4 Statistics 101: What You Should Know About Data, (50 pages); 2. Linoff & Berry, Chapter 5 Descriptions and Prediction: Profiling and Predictive Modeling, (44pages); 3. Siegel, Chapter 3: The Data Effect: A Glut at the End of the Rainbow (data), (36 pages); 4. Siegel, Chapter 4: The Machine That Learns: A Look Inside Chase's Prediction of Mortgage Risk (modeling), (30 pages);
10-Feb	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 15 Market Basket Analysis and Association Rules., (46 pages); 2. Linoff & Berry, Chapter 18 Building Customer Signatures, (38 pages); 3. Siegel, Chapter 5: The Ensemble Effect: Netflix, Crowdsourcing, and Supercharging Prediction (ensembles), (18 pages); 4. Siegel, Chapter 6: Watson and the Jeopardy! Challenge (question answering), (36 pages);
17-Feb	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 20 Too Much of a Good Thing? Techniques for Reducing the Number of Variables, (40 pages); 2. Linoff & Berry, Chapter 21 Listen Carefully to What Your Customers Say: Text Mining, (46 pages); 3. Siegel, Chapter 7: Persuasion by the Numbers: How Telenor, U.S. Bank, and the Obama Campaign Engineered Influence (uplift), (25 pages); 4. Siegel, Afterword: Ten Predictions for the First Hour of 2020, (3 pages); 5. Siegel, Appendices: A, B, and C;
24-Feb	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 6 Data Mining Using Classic Statistical Techniques, (42 pages); 2. Linoff & Berry, Chapter 19 Derived Variables: Making the Data Mean More, (42 pages);
3-Mar	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 9 Nearest Neighbor Approaches: Memory-Based Reasoning and Collaborative Filtering, (36 pages); 2. Linoff & Berry, Chapter 13 Finding Islands of Similarity: Automatic Cluster Detection, (40 pages); 3. Linoff & Berry, Chapter 14 Alternative Approaches to Cluster Detection, (36 pages);
26-Mar	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 7 Decision Trees., (44 pages); 2. Linoff & Berry, Chapter 8 Artificial Neural Networks., (40 pages);
2-Apr	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 10 Knowing When to Worry: Using Survival Analysis to Understand Customers, (40 pages); 2. Linoff & Berry, Chapter 11 Genetic Algorithms and Swarm Intelligence, (32 pages);
7-Apr	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 12 Tell Me Something New: Pattern Discovery and Data Mining, (30 pages); 2. Linoff & Berry, Chapter 16 Link Analysis., (32 pages);
14-Apr	<ol style="list-style-type: none"> 1. Linoff & Berry, Chapter 17 Data Warehousing, OLAP, Analytic Sandboxes, and Data Mining, (42 pages);

