

MIS 761: Business Analytics Statistical Methods and Tools

Instructor: Andrew Hardin, PhD

Full Professor of Information Systems

Lee Business School—Office: BEH 339

Email: andrew.hardin@unlv.edu Home page: [Faculty Site](#)

Course Description

Explores how data analytics relates to the scientific method, and then employs advanced statistical techniques for the contemporary analysis of organizational data. Taking an analytical approach, the applicability of particular statistical approaches/packages for gaining a business understanding in specific settings will be examined.

Learning Outcomes

By the end of this course, graduate students should be able to understand sophisticated processes associated with the analysis of organizational data sets. Specifically, students will first learn about the scientific method, and then about the CRISP-DM, a widely utilized framework for organizing the data mining process. In addition, graduate students will learn about the appropriate application of statistical models for answering specific research questions business problems. To accomplish this, students will gain knowledge of the statistical packages R and IBM SPSS Modeler.

Required Course Materials

The course requires the use of R and the R-Studio integrative development environment. R is open source, and students can download it free [here](#) along with R Studio [here](#). In the second half of the semester we switch to IBM SPSS Modeler which students will use for the final project. I will provide a license that can be used to install IBM SPSS Modeler software on your personal machine.

1. *Discovering Statistics Using R* (2012), Andy Field; SAGE: ISBN-13: 978-1446200469

2. *Data Mining with SPSS Modeler: Theory Exercises and Solutions*, (2016), Tilo Wendler and Soren Grottrup Springer, ISBN – 978-3-319-28707-2

Assignments:

	Points	Percentage
Participation (2 - introduction post, 5 - for discussion board posts)	7	5%
Technology Update	7	5%
R Packages and Script Example	7	5%
Andy Field In-Chapter Examples (2 points each * 8)	16	11%
IBM SPSS Modeler Application Guide Chapters (2 points each * 8)	16	11%
Wendler In-Chapter Examples (4 points each * 5)	20	14%
Project 1	40	27%
Final Project:	35	24%
Total	148	100%

Library Resources:

Patrick Griffis Email: patrick.griffis@unlv.edu Phone: 895-2231

Participation: Introduction, Course Questions Discussion Board Postings. Students receive two participation points for their introduction posting. Specifics on the content of the introduction post are included in the introduction discussion area. **Due by February 1** Students can receive up to five additional points for posting helpful hints and answering other students' questions in the course questions discussion boards. Posting thoughtful comments about other students' technology updates, or R-Script packages also count. Only frequent contributors will receive the full five points. If you contribute only occasionally, you will receive 2 points. If you do not contribute at all, you will not receive any points. So please contribute, we are all in this together! **Due by April 19** Students can also **earn seven extra credit points** by completing four surveys during the course.¹

In-Chapter Examples and Summaries: Students will choose and complete the in-chapter examples for any eight chapters in the Andy Field R text, **Due by March 1**,^{ii iii} any eight chapters of the IBM SPSS Modeler Application Guide, **Due by April 23**, and any five chapters from Chapter 2 through Chapter 8 in the Wendler text, **Also Due by April 23**. Students must also submit a one-page word document that summarizes when the statistical approach described in the chapter is appropriate, the theory underlying it, and a discussion of its assumptions. If the chapter is an overview chapter then students should submit a summary of the chapter concepts. Students should choose chapters based on their individual statistical backgrounds. The topics covered rapidly increase in complexity so the materials are suitable for any skill level. This is a graduate course, so please challenge yourselves accordingly! Note also that not all of the Andy Field script examples work because of updates to R since the 2012 publication of the textbook. Students can find documentation about these updates on the web (new packages, etc.). My videos demonstrating the Andy Field in-chapter examples also include some information about the updates.

R Packages and Script Examples Discussion Board Posting: Each student is required to post an example of a useful function in the R Packages and Script Examples Discussion area. I have provided an example of the minimum requirements for this assignment. **Due by April 16**

Technology Updates: Each student will post a one-page reports in Canvas on a contemporary topic related to business analytics (the topic can be about positive or negative aspects of business analytics). Students should base these updates on articles published in the popular press, recognized internet sources, etc. Please provide the article citation within your post. **Please submit your technology update early in the semester so that other students can comment on it. I will not accept any technology updates posted after April 16.**

Project 1: This projects substitutes for the exam in the course. I have provided simple and advanced project examples in Canvas. Students will utilize R for Project 1. Since this is basically an exam I will release the details of the assignment on March 1. **Due by March 15**

Final Project-SPSS Modeler: Students will design a study utilizing the Crisp-DM process, test it within IBM SPSS Modeler and write up the results. Students may use R to supplement the analysis. Students will receive a license for Modeler as part of the IBM MAP program. See the Final Project

Module for example projects. Please reach out to me as soon as possible with questions about the suitability of your project. Note that Patrick Griffis (listed above), can be a valuable resource! **Due by May 10**

Free Loading: To combat instances of freeloading, all students are required to complete a group assessment spreadsheet. I will consider these assessments as I compute final grades for the course. Please notify me if a group member is not participating so I can intervene as necessary, i.e., before it is too late. Failure to submit the team assessment spreadsheet will result in a seven-point deduction from your total score for the course.

Tentative Course Calendar

<i>Topic</i>	<i>Reading</i>	<i>Activity/Deliverable</i>
<u>Module 1: The Scientific Method and Statistics Review Using R</u>		
The Scientific Method	Slide Deck and Video	Review Slide Deck and Video
Statistics Review	Andy Field Text and Videos	Complete the in-chapter examples for eight of the 19 chapters. Due March 1
<u>Module 2: Project 1</u>		
Details released March 1: Due March 15		
Spring Break March 15 – 19		
<u>Module 3: The Crisp-DM and IBM SPSS Modeler</u>		
Crisp-DM	Slide Deck, Video, and Crisp-DM Guide	Read Crisp-DM Guide, Review Slide Deck and Video R-Script Examples: Due April 16 Tech Updates: Due April 16
IBM SPSS Modeler	IBM SPSS Modeler Application Guide, Wendler Textbook and Videos	Complete any eight chapters of the Modeler Application Guide and any five of Ch2 through Ch8 in the Wendler Text. Due 4/23
<u>Module 4: Course Project</u>		
Final Project	Andy Field Text, IBM SPSS Modeler Application Guide, Wendler text, Slide Decks and Videos	Course Project: Due 5/10

Class Policies:

Late Assignments: Late assignments are permissible only in cases of unavoidable personal or family emergencies and the student must notify me as soon as possible. In all other cases, there will be a significant reduction in points for late assignments.

Grade Appeals: If you believe there was a mistake made in the grading of your assignment please notify me promptly and I will determine whether a review of the assignment is warranted.

University Policies

See Information in Canvas Listed Under Syllabus

Endnotes:

ⁱ Students have the choice to complete the four surveys or complete an alternative assignment. Completing the four surveys should take about 60 minutes (students that choose this option must complete all four surveys to receive the seven extra credit points). The alternative assignment is to research an emerging topic on analytics such as machine learning applications, AI applications, etc. The deliverable is a two-page analysis (not summary) of the topic. Please cite your work in APA format.

ⁱⁱ Students that have taken the IS 372 course should choose different R chapters than those we covered in IS 372. Students should choose different Modeler Applications chapters as well. There are a limited number of Wendler chapters, but former IS 372 students should choose different chapters as much as possible.

ⁱⁱⁱ If you are an expert at R, or, have been away from statistics for a while, you can choose to complete the eight chapters and Project 1 using SPSS Statistics. There are a few caveats to taking this path. First, you will need to purchase the SPSS Statistics version of the Field textbook (either the 4th edition ISBN-13: 978-9351500827 or the 5th edition ISBN-13: 978-1526436566 will work). Second, IBM does not provide UNLV with student commuter licenses that allow students to install SPSS Statistics on their personal machine. Although you can remotely access the SPSS Statistics software by following these **instructions**, I am not sure how well it works as I have never needed to use the software in this fashion. On the positive side, I am an expert at using SPSS Statistics and IBM SPSS AMOS (a structural equation modeling package). Please let me know as soon as possible via email if you will be using IBM SPSS Statistics, and please also briefly explain why you will be using it rather than R. Note finally that this is for the R chapters and Project 1 only. All students must complete the Modeler Application chapters, the Wendler & Grottrup chapters and the final project using IBM SPSS Modeler.