

GSM 503 Data Analysis, Modeling and Decision Making
Fall, 2008
Professor Hand

Syllabus

Contact Information

Instructor: Professor Michael L. “Call Me Mike” Hand
Office: Atkinson 402
Hours: Tu, Th 11:30 a.m. – 12:30 p.m.,
and by appointment
Phone: 503-370-6056
E-mail: mhand@willamette.edu

Required Course Text

Bowerman, Bruce L., Richard T. O’Connell and Emily S. Murphree. *Business Statistics in Practice, Fifth Edition*. McGraw Hill/Irwin, Burr Ridge, Illinois. 2009. (ISBN: 007820097, includes Minitab Student Release 14.)

Course Philosophy and Learning Objectives

The techniques of quantitative analysis enjoy broad acceptance and wide-spread application in enterprises of all sizes and are deployed against problems in such diverse areas as – financial management, quality control, public budgeting, research and development, compensation analysis, and market research. Still, the many opportunities for fruitful application of powerful quantitative ideas and tools remain largely underexploited. We frequently hear ours described as an information economy, as if we were awash in the stuff (information) while, in actuality, high quality information remains a relatively scarce resource. Instead, we are merely up to our ears in data, facts and figures that by themselves provide little insight or basis for action. The challenge is to refine seemingly limitless quantities of raw data, to identify meaningful trends and patterns, to develop fundamental managerial insights, and to extract the essential information that we require to effectively manage and control our enterprises. Our firm grasp of quantitative concepts, particularly as they relate to the commonly recurring requirements of management, is the critical link between the data that we have and the information that we value.

Upon successful completion of this course, you should –

- Be an intelligent interpreter and user of data and statistical information.
- Understand different types of data and how they are obtained.
- Understand the range of statistical processes for transforming data into useful and actionable information.
- Develop a working familiarity with basic computational tools of data summarization, analysis, and modeling along with a level of self-sufficiency in their use.

- Be able to effectively communicate the results of quantitative analyses and models, especially through the use of statistical graphics and other visualization tools.
- Build confidence in your ability to creatively and intelligently exploit statistical opportunities to advance your professional objectives and those of your enterprise.

It is tempting to think of Data Analysis, Modeling and Decision Making as merely a general overview of selected statistical and analytical tools; in particular, as these tools relate to knowledge acquisition and to decision support in the context of management. On one level, we can and will resort to thinking of our course in this way. However, we are most interested in establishing and building the links between the quantitative tools that we know and the managerial decisions that we face so that, when we do encounter an opportunity, we are prepared to frame it in the analytical terms that allow us to work toward effective solutions. To paraphrase a favorite assertion of my good friend and esteemed (retired) colleague, Professor G. Marc Choate, this course is not just about statistics, but rather it is about “a way of thinking.”

To reinforce the point, consider the following small mathematical/statistical proposition – find the equation of the line connecting the coordinate pairs (100, 1300) and (300, 23000). You may scoff at the apparent simplicity of the question as posed mathematically. What if instead, we frame the situation noting that these coordinates represent data relating production cost to the level of production (in units.) To further nourish our insight into the nature of the production/cost relationship, we might (no, we most certainly would) also construct an appropriate visual representation of the data and perhaps be led to develop (at least as a first approximation) a simple linear model to represent the relationship. From there we should be prepared to address questions of how cost is related to production level. Can we interpret the estimated model coefficients to characterize the fixed and marginal components of cost? For a given level of production, what is the projected cost? For a given proposed production budget, what level of production can be supported?

Indeed, we are talking about a way of thinking; a way of thinking that applies not only when using simple linear models, but more generally to more complicated situations. As managers, we are interested in identifying, measuring and describing – the position of the enterprise, how that position is changing over time, what factors influence position, future projections of position, how we might control underlying drivers to improve or even optimize position. The course is laid out to address these critical questions while identifying and developing the underlying statistical concepts that support our fundamental information requirements. Among our immediate, short-term objectives is to equip all Willamette MBA students with the fundamental quantitative foundation essential to support your analytical work in subsequent courses, and ultimately in your career. Of even more lasting, long-term significance is that we learn to recognize and appreciate the value of quantitative and analytic thought in managing and controlling organizations. In this course, we strive to identify and exploit the many opportunities for the effective application of essential quantitative tools.

An additional theme of the course is that quantitative techniques are not the exclusive domain of specialists and technicians. While it would be naïve and borderline-reckless to dispute the value of quantitative tools in decision-making, it is still far too common that managers (in particular, and people in general) are willing to dismiss themselves as lacking the necessary academic

preparation and aptitude. The single most significant objective of this course is to overcome any latent anxieties that we may harbor relative to quantitative analysis and to develop confidence in our abilities to wield the weapons of knowledge, truth and beauty in the name of effective management.

Our quest to conquer the vagaries of an uncertain managerial environment is an odyssey that will begin by addressing the description and summarization of data (the tools of information extraction,) essential probability distributions and sampling distributions. Our initial work in data summarization and probability will form the foundation for our subsequent work in estimation, decision making, and statistical modeling.

Course Approach

The course is reasonably self-contained and assumes little at the outset, in the sense that we begin with the most fundamental of statistical concepts and proceed systematically to build a foundation of increasingly sophisticated ideas. However, we will move rather briskly across the statistical landscape in order to embrace a relatively broad range of ideas.

Our approach is a modern one, designed to exploit modern computing resources as essential tools of information extraction. Though much of our work is conceptual in nature and many of our modest computational objectives can be met using only paper and pencil (with perhaps a modest expenditure of cerebral energy,) we'll make frequent forays into the world of computing in search of the best ideas for extending our locus of control over "the numbers." Consistent with our view that, whenever possible, our analytical work should be conducted within the friendly confines of standard tools that are widely available, our principal software tool will be Microsoft Excel through its – basic computational and charting functionality, built-in statistical functions, and the Data Analysis Tool Pack add-in. In order to develop working familiarity with general-purpose statistical software, we will also rely on Minitab¹, a statistical exploration package that I have ordered in a bundle with our course text. We also have the option to make limited use of Mega-Stat (a McGraw-Hill Excel statistical add-in that is included on the resource CD that comes standard with our course text.) Finally, I may also supply occasional sample analyses conducted with other general-purpose statistical packages and data-mining tools such as SAS, SPSS, and Clementine.

Course Materials

The principal and organizing resource for the course is the text by Bowerman, O'Connell and Murphree (BOM, complete reference below.) BOM provides a fairly comprehensive traditional framework for the introduction of business statistics including strong cases and examples, problem sets and reference materials for key computational tools – Excel, Minitab, and Mega-Stat.

¹ Please see the syllabus section entitled Statistical Software – Minitab and Alternatives for Mac Users.

Required Course Text

Bowerman, Bruce L., Richard T. O'Connell and Emily S. Murphree. *Business Statistics in Practice, Fifth Edition*. McGraw Hill/Irwin, Burr Ridge, Illinois. 2009.
(ISBN: 007820097, includes Minitab Student Release 14.)

Additional course materials – readings, examples, and exercises – will be posted regularly to our Class Tools site over the Atkinson Intra-Web. You can navigate your way directly to the site through the link –

<http://agsm.willamette.edu/classtools/mba/siteinfo/> .

You will be prompted for your Willamette University login in order to gain access to course materials.

Course Activities

A detailed outline of the course – with a projected schedule of topics, resources, exercises, and examinations – follows. Minor adjustments to course materials, schedule, readings and problems, will be posted, as necessary, throughout the course. Background reading and problems are generally assigned for each class session. I suggest that you complete a first reading of all assigned material prior to class and make at least an earnest preliminary attempt at the assigned problems. This approach will help you to better identify learning objectives and areas of potential difficulty. When you prepare in this way, you will be primed to arrive for class – full of questions, curiosity, and infectious enthusiasm. Following class, reread the assigned material, this time for full comprehension, and complete (to your own satisfaction) any assigned problems. These assigned problems are designed to support your own self-assessment, to gauge comprehension, and to pinpoint any difficulties that may require additional attention.

There will be two required computing exercises and a final modeling project. The first exercise will give us practice in using the statistical tools available in Excel and Minitab for the graphical and numerical summarization of data. The second exercise provides a brief and relatively risk-free introduction to statistical modeling ideas and will help to develop the fundamental modeling and computing skills necessary for the final small-group modeling project. The final modeling project and presentation will be due at the end of our scheduled Final Examination time in lieu of an exam. Full project specifications will be distributed later in the course. Exercises, projects, and due dates are included in the course schedule that follows. As a vehicle for regular feedback and assessment, we may also have occasional in-class online surveys and mini-quizzes during class meetings, so be sure to bring your laptop computers to class.

Course Requirements and Grading

Course grades will be based upon two midterm examinations [25% each,] two computing exercises [5% and 10%, respectively,] the final modeling project [25%,] and class participation [10%,] where class participation scores will be determined by your participation in, in-class

surveys and quizzes. Consistent with the philosophy of emphasizing the application of fundamental statistical ideas, rather than formulas and computation, all examinations are open book and open notes. To assist in your review and preparation for exams, I will post sample examinations to our Class Tools site in advance of scheduled exam dates.

Laptop Computers in Class

You are advised to have your notebook computers with you in class and always at the ready, as we will be using them in a variety of ways. Our Class Tools site, available over the Atkinson Intra-Web, will be the sole means of distribution for essentially all course materials. While I will generally display current examples using the classroom projection system, you may wish to have copies of the examples open on your computer so that you can better follow along, to linger on selected points and zoom in as necessary. You will occasionally have the opportunity to look in and work along with me on in-class Excel examples, but we won't do this often, as it tends to consume a lot of our precious time together. As noted previously, to collect timely feedback during class, I plan to give occasional short-answer quizzes and surveys, where you will be asked to respond, using your laptop, through our Class Tools site. These in-class exercises will be one component of the assessment for the class participation portion of your course grade. Finally, if you are comfortable doing so, you may wish to use your laptop as a tool for taking notes. In between the times that we are formally using our laptops, please be considerate of your fellow class members and reserve Internet browsing, e-mail reading, and instant messaging for another time. I know that we all fancy ourselves brilliant multi-taskers, but these activities are a distraction, not only to you, but to your fellow class members. You may also risk injury to my tender feelings.

Waiver Policy

As previously noted, this course is not simply about the conveyance of mere technical content, but rather most importantly about establishing ways of thinking about analytics as a critical strategic resource of the enterprise. Our work will include significant experiential components to sharpen essential computational skills and, through project work, to develop a discipline for extracting actionable insights from available data. As such, we are not anxious to grant unconditional release from quantitative course requirements based solely on strong technical background. Instead, students with exceptional prior analytical training have the option of substituting a quantitative elective course – GSM 616 Business and Economic Forecasting (offered this semester Tu/Th 2:00 – 3:30 pm.) GSM 616 requires similar types of project work and covers at a higher level much of the modeling content of GSM 503, extending these ideas with additional techniques for analyzing time series data and forecasting. GSM 616 is recommended for students with two or more prior statistics courses (or perhaps one high-level course) including experience with statistical modeling software. Students wishing to substitute GSM 616 may do so essentially at their own discretion. However, I would like to confer briefly with each of you who choose to do so in order help you assess your own qualifications.

Statistical Software – Minitab and Alternatives for Mac Users

In much the same way that Apple has enjoyed a long and happy reign as the computer of choice among educators and creative professionals, Windows computers enjoy a similarly prohibitive market share edge in business and analytic applications. Though there are a few excellent statistical packages available for the Mac [I've always been a big fan of DataDesk], the choices are much more limited, especially among the major business analytics players. In the instructional setting – the best combination of price, performance, feature set, and ease of use is a statistical package called Minitab; the preceding optimality claim supported by the many (business) statistical texts (ours among them) that include coverage of Minitab, usually in addition to Excel. Because it is such a good choice and because it can be inexpensively bundled with our course text, I have ordered a text package that includes a student version of Minitab.

Unfortunately, Minitab is not available for the Mac. However, rather than being a problem, we should see this limitation as a feature, an opportunity; an opportunity to collaborate with other teammates, an opportunity for cross-platform computing experience, an opportunity to seek out and develop alternative computing solutions. While we all respect, even celebrate, the right of every individual to personal computing self-determination, when our course work calls for something more than Microsoft Excel (and occasionally it will, especially in the second half), Mac users may have a few choices to consider on how to proceed. However, I hasten to reiterate, this is not a problem, but an opportunity, a feature.

Perhaps the simplest short-term solution is to pool your computing resources and partner with a class member (you are encouraged to work in groups of two) who is Windows-capable to complete any required statistical computing. For possible cross-platform solutions, consider installing Minitab on an accessible Windows machine at work or home. If you have ready access to Willamette's Salem campus, Minitab is also available in all Willamette University computing labs; including in the large Smullin Hall lab and on the two public-access machines at Atkinson.

Thinking longer term, this will not be the only time that a software compatibility issue will rear its ugly head. In addition to many statistical and technical packages, other examples of PC-only software include MS Access and several software packages currently in use for various class simulations. It might make sense to look into alternatives for running Windows applications on your Mac. A variety of software products support dual boot capabilities or virtualization. Notable entries include a number of free solutions – Apple Boot Camp, Microsoft Virtual PC 2007, and VirtualBox. Other players in the market include Parallels and VMWare Fusion.

For alternative statistical software solutions, I am also including a link outlining statistical software available for the Mac. In particular, I draw your attention to MacANOVA that to my eye represents a reasonable combination of price (free) and features. You might try downloading MacANOVA and see what you can get it to do.

<http://www.macstats.org/statistics-software/regression.html>

Finally, if all else fails, do the best you can with whatever alternative computing resources you have available, noting in any submission any computing compromises that you make.

Atkinson School Expectation of Conduct for Students, Faculty and Staff

We are a community of learners. Our professional commitment is to create an environment that advances the science and practice of managing organizations. As individuals, we conduct ourselves with honor and integrity, treat everyone with respect, take responsibility for our actions, and fulfill our promises.

Examples of conduct consistent with this expectation include:

- Taking initiative
- Crafting solutions
- Completing assignments according to an agreed schedule
- Offering constructive criticism and accepting it appreciatively
- Taking responsibility for our own learning and that of others

We expect someone who observes or learns about behavior inconsistent with our expectations of conduct to initiate corrective action by clarifying the situation and encouraging the responsible party to act appropriately. In the case of a violation of School or University policies or the laws of relevant jurisdictions, notify the appropriate enforcement authorities.

Accommodations for Students with Disabilities

Students requesting accommodations for a disability must be registered and certified through the Willamette University Disability Services Office. Students seeking accommodations must contact the Disabilities Services Office in Bishop Health Center at 503-370-6471; or on the web at <http://www.willamette.edu/dept/disability/>. Atkinson students already approved for testing accommodations must notify Janet Jobes, Atkinson Recorder, at least two weeks prior to all scheduled tests for which they need accommodation.

Course Outline

The course schedule is a well considered plan, but it is only that, a plan. We will deviate as necessary, but I will keep the class informed of any changes or additions, mostly in the form of new materials that we develop throughout the course. Schedule updates, will be posted periodically as needed. Remember, this is only an exhibition. Please, no wagering.

Most readings and problem assignments are from the course text –

Bowerman, Bruce L., Richard T. O’Connell and Emily S. Murphree. *Business Statistics in Practice, Fifth Edition*. McGraw Hill/Irwin, Burr Ridge, Illinois. 2009. (ISBN: 007820097, includes Minitab Student Release 14.)

Additional examples and course readings will be available for download from the course Class Tools site – ClassTools: GSM-503:HandOuts:Examples:<subfolder>. Exercises will be distributed through – ClassTools: GSM-503:HandOuts:Exercises:<subfolder>.

I. Data Analysis, Probability, and Microsoft Excel

Tu 08/26 1. **Course Overview**

Graphical Summarization of Data

Resources: Why Math Will Rock Your World, *Business Week*
Competing on Analytics, *Harvard Business Review*
GraphicsGallery.ppt, BellEx.pdf
Reading: Chapters 1 and Sections 2.1 – 2.4
Problems: Chapter 2: 22, 37

Th 08/28 2. **Graphical Summarization of Data**

Resources: BellEx.pdf, Bell_Class.xls
Reading: Sections 2.1 – 2.4
Problems: Chapter 2: 75, 79, 80

Tu 09/02 3. **Numerical Summarization of Data**

Excel: Data Analysis ToolPak, Pivot Tables

Resources: BellEx, Bell_Class.xls, Describe.pdf, Describe.xls,
InvestClientEx.xls
Reading: Sections 3.1, 3.2
Problems: Chapter 3: 6, 14, 22, 24

Th 09/04 4. **Numerical Summarization of Data**

Minitab: Graphical Data Summary

Resources: BellEx, Bell_Class.xls, Describe.pdf, Describe.xls,
InvestClientEx.xls
Reading: Sections 3.3, 3.6
Problems: Chapter 3: 33, 52

**Exercise 1, Graphical and Numerical Data Summaries, Distributed
Due Tu 09/16**

- Tu 09/09 5. **Probability and Discrete Probability Distributions:**
Probability, Random Variables, Discrete/Continuous Distributions,
Expected Value, Binomial Distribution
Resources: DiscreteRV.pdf, DiscreteRV.xls, DealNoDeal.xls,
BinomialEx.pdf, BinomialWork.xls, BinTab.xls
Reading: Chapter 4 (light read), Sections 5.1 – 5.3, Appendices B and C
Problems: Chapter 5: 13, 15, 16
- Th 09/11 6. **Applications of the Binomial Distribution**
Resources: BinomialEx.pdf, BinomialWork.xls, BinTab.xls,
BinomialProblem.pdf, CaesarsGasp, CaesarsGasp.xls
Reading: Section 5.3, Appendices B and C
Problems: Chapter 5: 25, 27, 28
- Tu 09/16 7. **The Normal Distribution and Applications**
Resources: NormalProblem, NormalTable, NormalProbTool.xls
Reading: Sections 6.1, 6.3
Problems: Chapter 6: 29, 32, 34

Exercise 1 Due

- Th 09/18 8. **The Normal Distribution and Sampling Distributions**
Resources: NormalTable, NormalProbTool.xls, SampDistStock.pdf,
SampDistProb.pdf, CompFollow.xls
Reading: Sections 6.1, 6.3, 7.1, 7.2
Problems: Chapter 6: 29, 32, 34; Chapter 7: 11, 13

II. Statistical Estimation and Decision-Making

- Tu 09/23 9. **Interval Estimation, Large Sample Means**
Resources: CompFollow.xls, CircuitCityProb, BankWait.xls
Reading: Section 8.1
Problems: Chapter 8: 7, 11
- Th 09/25 10. **Interval Estimation, Small Sample Means**
Resources: BlueChip.xls, WholeFoods.xls, NormalTTables.xls,
ElementaryInference.pdf, EstimationTool.xls
Reading: Section 8.2
Problems: Chapter 8: 20, 23

Tu 09/30 11. **Interval Estimation, Sample Size and Proportions**
Resources: CompFollow.xls, WholeFoods.xls, PropEstimation.xls,
ExtWarr.xls, ColorSmart.xls, NormalTTables.xls,
ElementaryInference.pdf, EstimationTestingTool.xls
Reading: Sections 8.3, 8.4
Problems: Chapter 8: 29, 32, 41, 48

Th 10/02 12. **Midterm Examination I**

Tu 10/07 13. **Hypothesis Testing, Fundamentals and Large Sample Mean**
Resources: CompFollow.xls, BankWait.xls, BlueChip.xls,
NormalTTables.xls, ElementaryInference.pdf,
EstimationTestingTool.xls
Reading: Sections 9.1, 9.2
Problems: Chapter 9: 9, 38, 59, 60, 73, 75; Chapter 10: 7

Th 10/09 14. **Hypothesis Testing, Small Sample Mean and Proportions**
Resources: BlueChip.xls, WholeFoods.xls, PropEstimation.xls,
ColorSmart.xls, ExtWarr.xls, NormalTTables.xls,
ElementaryInference.pdf, EstimationTestingTool.xls
Reading: Sections 9.3, 9.4
Problems: Chapter 9: 9, 38, 59, 60, 73, 75; Chapter 10: 7

Tu 10/14 15. **IKEWS – No Class Meeting**

Th 10/16 16. **Estimation and Hypothesis Testing
Independent Two-Sample Situations**
Resources: ConcreteSet.xls, ConcreteBreak.xls,
ElementaryInference.pdf, EstimationTestingTool.xls
Reading: Sections 10.1, 10.2
Problems: Chapter 10: 22, 24

III. Statistical Modeling

Tu 10/21 17. **Estimation and Hypothesis Testing
Paired Sample Situations and Proportions**
Resources: ConcreteBreak.xls, AnnualFees.xls, Default.xls,
ElementaryInference.pdf, EstimationTestingTool.xls
Reading: Sections 10.3, 10.4
Problems: Chapter 10: 31, 34, 41

Th 10/23 18. **Statistical Modeling Overview**
Resources: FreshEx2, FreshSLREx, FreshSLR.xls, TargetBetaAnnual.pdf,
TargetBetaAnnual.xls
Reading: Sections 13.1, 13.2
Problems: Chapter 13: 9, 16

Statistical Modeling Project Requirements Distributed

Class Tools: GSM-503:HandOuts:Exercises:ModelingProject

**Exercise 2, Simple Linear Regression Exercise, Distributed
Due Tu 11/04**

- Tu 10/28 19. **Simple Linear Regression: Inferences and Prediction**
Simple Linear Regression: Estimation and Computing in Excel
Resources: TargetBeta.xls, WalmartBeta.xls, FreshSLREx, FreshSLR.xls,
F-Table.xls, FreshSLRClass.xls, FreshEx2
Reading: Sections 13.3, 13.4
Problems: Chapter 13: 24, 33
- Th 10/30 20. **Simple Linear Regression: Assessment**
Simple Linear Regression: Estimation and Computing in Minitab
Statistical Modeling with Multiple Predictors
Resources: FreshSLREx, FreshSLR.xls, F-Table.xls, FreshSLRClass.xls,
FreshEx2
Reading: Sections 13.5, 13.8, 14.1
Problems: Chapter 13: 40, 58; Chapter 14: 3
- Tu 11/04 21. **Multiple Regression: Model Assessment**
Resources: FreshEx2
Reading: Sections 14.2 – 14.4
Problems: Chapter 14: 9

Exercise 2 Due

- Th 11/06 22. **Multiple Regression: Prediction and Inference**
Estimation and Computing in Minitab
Resources: CCMClass.xls, FreshEx2
Reading: Sections 14.5, 14.6
Problems: Chapter 14: 14, 19
- Tu 11/11 23. **Multiple Regression: Residual Analysis and Assessment**
Resources: FreshEx2, AppraiseMLR(Minitab),
AppraiseMLR(Excel), AppraiseMLR08.xls
Reading: Section 14.9
Problems: Chapter 14: 37
- Th 11/13 24. **Midterm Examination II**

- Tu 11/18 25. **Model Building: Special Predictors**
Indicators Variables and Quadratic Models
Resources: FreshEx2, AppraiseMLR(Minitab),
AppraiseMLR(Excel), AppraiseMLR08.xls
Reading: Sections 14.7, 15.1, 15.2
Problems: Chapter 14: 26; Chapter 15: 3, 8, 9
- Th 11/20 26. **Multiple Regression: Multicollinearity and Variable Selection**
Multiple Regression Computations in Minitab and Excel
Resources: FreshEx2, AppraiseMLR(Minitab), AppraiseMLR(Excel),
AppraiseMLR08.xls
Reading: Section 15.4
Problems: Chapter 15: 16
- Tu 11/25 27. **Multiple Regression: Model Refinements**
Resources: AppraiseMLR(Minitab), AppraiseMLR(Excel),
AppraiseMLR08.xls
Reading: Sections 15.5, 15.6
Problems: Chapter 15: 25, 32
- Th 11/27 28. **Thanksgiving Break – No Class Meeting**
- Tu 12/02 29. **Multiple Regression: Recap of Model Building Process and Assessment**
Resources: AppraiseMLR(Minitab), AppraiseMLR(Excel),
AppraiseMLR08.xls
- Th 12/04 30. **Last Class Meeting**
Course Evaluations
Multiple Regression: Forecasting Applications
Resources: TargetAnalysis.xls
- Tu 12/09 **No Final Examination**
Final Project Presentations Due
Uploaded By End of Scheduled Final Exam Time, 1:00 pm