

QMB 6358 Data Analytics for Business

Instructor

Kiran Garimella, Ph.D.

M: 262-909-4192

E-mail: kgarimella@usf.edu

Class Information

Location: BSN 115

Time: Mondays, 6:30 – 10:15 pm ET

Office Hours: Wednesdays 6:00 - 8:00 pm ET, online (Teams)

[Links to an external site.](#)

Course Description

This course will provide an introduction to data analytics for managers. It is targeted for business students and provides a strong foundation for analyzing business data using statistics and machine learning.

Structure & Approach

This course builds off the foundational course (QMB 6305 Managerial Decision Analysis) by focusing on data analytics for business majors and MBA students. For those students who have not taken the foundation QMB 6305, the course recaps all of those important concepts. Those who have taken QMB 6305 will benefit by the extended treatment of those and additional topics with deeper hands-on experience through various exercises, projects, and in-class competitions. Several additional topics will include cluster analysis and classification.

We will work on assignments, class discussions, projects, and a final exam.

Textbook

We will use the following book as a reading text. **PDFs are available at the library.** Please use this links (if you are not already logged in, you will be asked to login with a NetID):

Competing on Analytics: The New Science of Winning by Thomas Davenport and Jeanne Harris, Harvard Business School Press, 2017:

<http://ezproxy.lib.usf.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1798662&site=eds-live>[Links to an external site.](#)

Data Analytic Software

We will use the software R and Microsoft Azure ML for analyzing and mining our data. They are both free! In fact, R is an open-source based software and as such grows much faster than any commercial software solution. R is one of the most powerful and popular software solutions for mining business data. In fact, large analytical companies such as Google or AT&T use R. R is also becoming more and more popular among government agencies (such as the DOD). R is powerful, yet it is free – this combination makes it a winning proposition for many companies, especially during times of tightening budgets and financial stress.

Microsoft Azure ML Studio is also available through a free tier with limits sufficient for classroom purpose. There is a very intuitive interface that allows you to build "workflows" for analyzing data.

Installing R

- Go to CRAN <http://cran.r-project.org/Links to an external site.>
- Locate your operating system (OS) in the first box of the welcome page. (Below, I will assume that you are using Windows; all other OS are very similar.)
- Click on "Windows", then click on "base"
- Then click on "Download R 4.0.0 for Windows" (Note: the release number may have changed by the time you are downloading this; that is not a problem; simply install the most recent version.)
- Follow all instructions during the installation process. On one of the screens, you will see an option to create an icon on your desktop. Check the box so that you can locate R quickly.
- After installation, locate the "R" icon on your desktop; double-click; DONE!!

Rather than using R (which is a command-line language) directly, we will use a more user-friendly graphical user interface (GUI). This GUI is called the "R Commander".

Installing the R Commander

- Right-click on the R desktop icon and 'Run as an administrator'.
- Go to "Packages" at the top of the R window.

- Click on “Install package(s)...”
 - A window will pop up asking you for your choice of an R mirror; select the location closest to you (e.g. USA TX1)
 - Another window will pop up. Scroll down until you find “Rcmdr”.
 - Select and click “OK”
 - If you are asked about installing from sources for compilation, choose 'No'.
 - You will now see several messages inside your R window; these messages confirm the installation of all the components necessary for the R Commander.
 - Upon successful installation, you will find the following message: “package 'Rcmdr' successfully unpacked and MD5 sums checked”
- Now, type the following inside your R window: `>library(Rcmdr)`
 - Then, hit “Enter”
 - If the system asks if you want to install any missing packages, choose 'Yes'.
 - A new window will pop up; this is the R Commander window.
 - For a quick tutorial on the R Commander, find the “Help” tab, and click on “Introduction to the R Commander”

Troubleshooting: Below are a few links that could help you trouble-shoot your installation; please visit them when you encounter any problems downloading or starting R Commander.

- <https://www.andrewheiss.com/blog/2012/04/17/install-r-rstudio-r-commander-windows-osx/>Links to an external site.
- <http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/installation-notes.html>Links to an external site.
- <http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/old-Mac-installation-notes.html>Links to an external site.

Course Website/Platform

I will use the Canvas to distribute various files as necessary (e.g., course documents, data, assignments and solutions). Please be sure to visit Canvas on a regular basis. It is best to install the mobile app also to see notifications in a timely manner. In communicating with the instructor, please only use the Canvas mail (not email).

Course Format

This class will meet face-to-face and will follow all guidelines and mandatory requirements as set forth by authorities.

There will be frequent breaks.

This is a very practice-oriented detailed course. We will discuss both hands-on implementation of data-driven ideas as well as more high-level, strategic concepts of data analytics. We will leverage useful tools, predominantly R, but also explore other tools (such as Azure ML Studio) to cover advanced concepts in data analytics.

We will not do hardcore programming in R or any other language, perform any complicated analysis, or use very large datasets. We will avoid as much as possible those technicalities that are the plight of professional statisticians and data analysts but do not add additional insights into data analytics for business itself. We will leverage simple tools that with a few commands - many through a visual interface - allow you to conduct sophisticated analysis without much coding in the R console.

Course Work & Grading

Deliverable	Points
Quizzes (8 * 10 pts)	80
Assignments (4 * 25 pts)	100
Individual Presentation	40
Group Project & Presentation	100
Final	80
Total	400

Quizzes: These will cover all material through presentations, videos, recordings, and assigned readings (if any) that are covered in class since the last quiz. Quizzes may include pre-work using given software (but not necessary during the quiz) and upload of screenshots of results appropriate questions in the quiz.

Assignments: These will cover integrated analysis of given datasets.

Individual Presentation: Each student will pick a company or tool and conduct search / research online to determine how the company is using analytics in its business or how the tool (vendor solution) can be used in data analytics.

Group Project: Students will be assigned randomly to a group and will conduct a complete data analysis by incorporating multiple tools and techniques.

Final: The final exam is cumulative from the beginning of the class.

Detailed Syllabus (All times in US Eastern Time)

Readings codes: COA (Competing on Analytics)

Week	Dates	Topic	Readings	Task / Deliverable
1	Feb 6 - Feb 12	Motivation, the process of data analytics, review of basic descriptive statistics and recap of statistical foundations: functions, slopes, coefficients, box and scatter plots, correlation R & Rcmdr		First Day Attendance
2	Feb 13 - Feb 19	Evaluation Metrics, Cost-Benefit Analysis	COA: Foreword, Introduction, Chapters 1 & 2	Quiz 1
3	Feb 20 - Feb 26	Model building, Least squares regression, Statistical Inference Linear regression insights	COA: Chapters 3 & 4	Quiz 2
4	Feb 27 - Mar 5	Dealing with Categorical Variables and Nonlinearities	COA: Chapter 5	Quiz 3
5	Mar 6 - Mar 12	Dealing with Dimensionality and Missing Data	COA: Chapter 6	Quiz 4
	Mar 13 - Mar 19	Spring Break		
6	Mar 20 - Mar 26	Multicollinearity, Stepwise regression	COA: Chapter 7	
7	Mar 27 - Apr 03	Cluster analysis and integrated analysis with regression	COA: Chapter 8	Quiz 5
8	Apr 4 - Apr 9	Predictions with Linear Regression, Evaluating Model Performance	COA: Chapter 9	Quiz 6
9	Apr 10 - Apr 16	Using Cluster Analysis Individual Presentations		Quiz 7

10	Apr 17 - Apr 30	Classification Problems & Confusion Matrix Individual Presentations		Quiz 8
11	May 1 - May 4	Project Reviews Finals Prep Individual Presentations		
12	May 5	Final Exam Week		Final Exam (Date TBD) Group Project

Quizzes and Exams

1. All quizzes will be a combination of True/False, Multiple Choice Questions, Multiple Answers, and possibly upload of data analysis screen shots.
2. Questions may include results of analysis or a visual picture of analysis and ask you to answer based on that information. So, it is important that you familiarize yourself with Rcmdr and visualizations with other tools, if any.
3. Most quizzes will be open book by default (unless explicitly stated otherwise by the instructor), but strictly no consultation with anyone in any form. You are to rely only on the textbook, your notes, or instructor's presentations. Quizzes will cover material from the reading texts or other articles that may be assigned during the course.
4. The Final exam, also open book, **is cumulative** and will cover the complete material from the beginning.

Letter Grades (% of total points)

Scores will be rounded up to the next whole point. To adhere to consistency, academic integrity, and fairness to everyone, requests for extra points in order to get the next higher letter grade for any reason will not be fulfilled. Reasons include, "I need the next letter grade otherwise my employer won't reimburse my tuition expenses", "I need to meet my scholarship requirements", or "It's a matter of cultural pride", etc.

Exceptions for makeup work are provided only for true, documented emergencies or based on guidance from the University due to any extraordinary circumstances.

Minimum Score	Maximum Score	Letter Grade
---------------	---------------	--------------

95%	100%	A+
90%	94%	A
85%	89%	B+
80%	84%	B
75%	79%	C+
70%	74%	C
60%	69%	D
0%	59%	F

Late Submission

One-day late submission on assignments incurs 10% penalty. No late submissions beyond 1 day.

Extra Credit and Make Up Work

There is no extra credit or make up work available. Quizzes or exams will not be re-opened in case of issues with your computer. It is your responsibility to keep your computer and wifi in full working order. Widespread network failures will be dealt with on a case-by-case basis and in accordance with University guidelines.

Policy Statements

<https://www.usf.edu/provost/faculty/core-syllabus-policy-statements.aspx> Links to an external site..