ISM 6358.902.F80694 Data Analytics for Business

Instructor

Kiran Garimella, Ph.D.

M: 262-909-4192

E-mail: kgarimella@usf.edu

Class Information

Location: BSN 2305

Time: Wednesdays, 6:30 pm – 10:15 pm ET

Office Hours: By Appt.

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 extended treatment of those and additional topics with deeper hands-on experience through various exercises, projects, and competitions. Several additional topics will include classification, neural networks, and decision trees.

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

Textbook

We will use the following books as foundational references and material for projects and discussions:

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

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/
- 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/
- http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/installation-notes.html
- http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/old-Mac-installation-notes.html

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 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 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 (5 * 10 pts)	50
Assignments (5 * 15 pts)	75
Individual Presentation (1 * 25)	25
Group Project	60
Final	40
Total	**250**

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. Quizzes are open book.

Assignments: These will cover integrated analysis of given datasets.

Individual Presentation: Students will research a company, tool, or vendor in data analytics and deliver a presentation.

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.

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.
4. You may start a quiz or exam at any time from the time it is available. However, you will have only the maximum allotted duration from the start to complete your quiz or exam.
5. The Final exam ***is cumulative*** and will cover the complete material.
6. **I strongly recommend** that you do **not** start your quizzes and exams right before the deadline. If you do not give yourself enough buffer time, anything that can go wrong

generally does go wrong, so don't cut it too fine. Remember, there is no makeup for *any* reason besides true, documented emergencies.

7. At the discretion of the instructor, quizzes and exams may be electronically proctored.

Late Submission

Late submission on assignments incurs 10% penalty per 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 Wi-Fi in full working order. Widespread network failures will be dealt with on a case-by-case basis and in accordance with the University guidelines.

Letter Grades (% of total points)

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

Scores will be rounded up or down using the usual rounding rules. To adhere to consistency, academic integrity, and fairness to everyone, requests for extra points in order to achieve the next higher letter grade for any reason will not be fulfilled. Exceptions for makeup work are provided only for true, documented emergencies or based on guidance from the University due to any extraordinary circumstances.

Syllabus

Reading code: COA (Competing on Analytics)

Week	Dates	Topic	Readings	Task / Deliverable
1	Aug 21 – Aug 27	Introduction and Foundations R & Rcmdr		First Day Attendance
2	Aug 28 – Sep 3	Exploratory Data Analysis	COA: Foreword, Introduction, Chapters 1 & 2	
3	Sep 4 – Sep 10	Understanding Data with Cluster Analysis	COA: Chapters 3 & 4	Quiz 1 Assignment 1
4	Sep 11 – Sep 17	Understanding Data with Linear Regression	COA: Chapter 5	Quiz 2

5	Sep 18 – Sep 24	Predicting with Linear Regression	COA: Chapters 6, 7	Quiz 3 Assignment 2
6	Sep 25 – Oct 1	Classification Problems & Confusion Matrix	COA: Chapter 8	Assignment 3
7	Oct 2 – Oct 8	Precision Classification and Prediction	COA: Chapter 9	Quiz 4
8	Oct 9 – Oct 15	Missing Data and Dimensionality Reduction		Quiz 5
9	Oct 16 – Oct 22	Cost Benefit Analysis (CBA) and Strategic Decision Making (SDM) (Online Class)		Assignment 4

10	Oct 23 – Oct 29	Case Studies in CBA and SDM		Assignment 5
11	Oct 30 – Nov 5	Case Studies in CBA and SDM		
12	Nov 6 – Nov 9	Exam Prep		Final Group Project Final Exam

Policy Statements

https://www.usf.edu/provost/faculty/core-syllabus-policy-statements.aspx.

Course Policy on Acceptable Use of Generative AI Tools

Purpose: The purpose of this policy is to foster a dynamic learning environment that encourages technological adaptation, innovative thinking, and the ethical use of AI resources in academic endeavors.

Policy:

1. *Definition of Generative AI Tools: Generative AI tools refer to any artificial intelligence-powered software, program or application that can generate content, including but not limited to text, visuals, music, and other creative outputs. Examples of these tools include AI text generators, AI content rewriters, AI graphic generators, etc.*
2. *Permitted Use: The use of generative AI tools is permitted for learning and presentations, but not for the individual lab work in projects (use in writing industry overviews is permitted), lab work in assignments, quizzes, and exams. For permitted use, students must responsibly use these tools, adhering to the guidelines outlined in this policy.*
3. *Student Responsibility: Students are responsible for appropriately using generative AI tools in their work. This includes:*
    - *3.1. Citing all AI-generated content used in their submissions.*
    - *3.2. Demonstrating a deep understanding of the subject matter, not solely relying on AI-generated content. Cross-reference claims and statements with original sources and providing appropriate citations are expected.*
    - *3.3. Using AI tools as a supplemental resource (i.e., as an editor), not as the primary means of completing assignments.*
    - *3.4. Understanding that generative AI tools, while powerful, are not infallible and can produce misinformation or inaccurate results. Students are responsible for the accuracy of their submissions and must cross-verify the information produced by these tools with reliable sources.*
4. *Violation Consequences: Misuse of AI tools, including use of AI that undermines the student learning objectives of the course or assignment, failing to cite AI-generated content, relying too heavily on AI for work completion, or submitting inaccurate information generated by AI tools, will be subject to academic penalties. Consequences may range from a reduction in an individual assignment grade to larger academic sanctions per USF policy, depending on the severity of the violation (USF Regulation 3.027).*
5. *Exceptions: If there are specific assignments where the use of AI tools is not appropriate, these will be clearly marked in the assignment guidelines. Students must adhere to these specific instructions.*
6. *Questions and Clarifications: If students are unsure whether a tool they wish to use qualifies as a generative AI tool, or if they have questions regarding the allowable use of such tools, they should consult with the course instructor before using it.*