**Course Syllabus** [4 May 2020]

**Data Analytics in Business**

**MGT 6203 Online**

**Summer 2020**

**Professors**:

Frederic Bien, PhD, MS.QCF

Email: [fvbien@gatech.edu](mailto:fvbien@gatech.edu)

Office: 496 (or 4161) in Scheller College of Business

Class Time and Location: Online in Canvas and EdX

Office Hours: Thursday 8:30-9:30 or by appointment in BlueJeans.com

**Teaching Assistants:**

To be announced. Teaching Assistants are very important to this course as you will see in Monday evening office hours (TBD). TAs will provide R code instruction and answers in Piazza online forums.

**Guest Lectures by:**

Prof. Sridhar Narasimhan, Prof. Jonathan Clarke, Prof. Bob Myers from GeorgiaTech Scheller College of Business

**Course Brief Description**

The primary objective of this course is to teach the scientific process of transforming data into insights for making better business decisions.

This course covers *basic methodologies, algorithms, and challenges* related to analyzing business data. We will also study *applications of data analysis* in:

1. Finance & Investments
2. Marketing & Advertising
3. Operations & Logistics.

**Prerequisite**

* Calculus and Linear Algebra
* Probability and Statistics
* Background in programming and willingness to learn R
* Introductory course in Analytics Modeling

**Course Goals**

After taking this course, students will be able to:

* approach business problems data-analytically. Students should be able to think carefully and systematically about whether and how data and business analytics can improve business performance.
* develop business analytics ideas, start projects to analyze data using business analytics software, and generate relevant business insights for decision-making.

**Textbooks**

* Required: (ISLR) *Introduction to Statistical Learning with Applications in R*, by Gareth James, Daniela Witten, Trevor Hastie & Robert Tibshirani. Publ. Springer, New York (2017). ISBN-10: 1461471370. ISBN-13: 978-1461471370

Downloadable for free at <http://faculty.marshall.usc.edu/gareth-james/ISL/>

Or available for purchase as a book at Amazon.com, [BN.com](http://BN.com), Ebay.com, etc.

* Recommended: (Galit) Data Mining for Business Analytics: Concepts, Techniques, and Applications in R, by Galit Shmueli, Peter C. Bruce, Inbal Yahav, Nitin R. Patel, Kenneth C. Lichtendahl Jr. Publ. Wiley, Hoboken, NJ (2018). ISBN-10: 1118879368. ISBN-13: 978-1118879368.

May be found on rogue sites on the internet (not recommended) or available for purchase from Amazon.com, [BN.com](http://BN.com), Ebay.com, etc.

* You will also need to purchase and download some case studies from Harvard Business School online library. Here is a link to a package that contains all 7: <https://hbsp.harvard.edu/import/708301>

**Course Description**

Today businesses, consumers, communities and societies create or leave behind massive amounts of data as a by-product of their activities. Leading-edge companies in every industry are using data analytics to add to, or replace, intuition and guesswork in their decision-making. As a result, business managers are collecting and analyzing enormous amounts of data to try to discover new patterns and insights, and/or to run controlled experiments to test various hypotheses.

This course prepares students to understand business analytics and become leaders in these areas in business organizations. This course teaches the scientific process of transforming data into insights for making better business decisions. It covers the methodologies, issues, and challenges related to analyzing business data.

This course will illustrate the processes of analytics by allowing students to apply business analytics algorithms and methodologies to various business problems.

The use of carefully selected examples places business analytics techniques in context and teaches students how to avoid the common pitfalls, emphasizing the importance of applying proper business analytics techniques. The course will also show that often there can be more than one “good answer” or one “good choice”. We need to be discerning in the type of data that we choose to analyze and how we analyze it.

**Hardware Requirements**

Please follow GeorgiaTech’s computer ownership guide at <http://sco.gatech.edu/>.

Make sure that you have admin rights on your laptop since occasionally you will need to install R, RStudio, many packages in R, and other software like Radiant, maybe Gephi. Note that tablets, Chromebooks, and old laptops may not work well for this class at this time. (As we move the course toward use of R notebooks, eventually they will work.)

**Software Requirements**

We will be learning business analytics with the help of open-source and free software applications that are provided for educational use. Please follow instructions provided in their respective websites and install the following software in your personal laptop:

1. R: <https://www.r-project.org/>
2. RStudio: <https://www.rstudio.com/>
3. Radiant: <https://radiant-rstats.github.io/docs/install.html>

There are many resources on how to learn R. We will discuss some in the course.

* *R for Datascience:* <http://r4ds.had.co.nz/>
* *DataCamp:* [www.DataCamp.com/courses/free-introduction-to-r](http://www.DataCamp.com/courses/free-introduction-to-r)
* *RStudio Education:* <https://education.rstudio.com/>
* *Swirl:*  www.SwirlStats.com

**Communication**

*Instructor/TA Communication*: All course announcements will be made via Canvas or EdX. You are expected to check Canvas/EdX a few times per week for important course-related information.  By following the instructions provided in Canvas/EdX, you can ensure that you do not miss important instructions, announcements, etc. If you want, you can adjust your Canvas/EdX account settings to receive important information directly to your email account or cellphone. To get started, log into the Canvas/EdX, click on this course, and see the section entitled *“Before You Begin: Instructions for Getting Started.”*

*Content Questions and Help*: Because questions can often be addressed for the good of the group, please do not email your questions directly to the instructor. Instead, course and content questions will be addressed on an online chat platform called [Piazza.com](http://www.piazza.com/).

Get an account in Piazza today. These online forums will be a VERY valuable source of information and hints about the course and problem sets. Note that you can set your post to “Private” to ask questions to the instructor and TA about issues unique to you.

*Office Hours*. Office hours will be conducted every week by the instructor and TAs. These sessions will be both an opportunity for you to ask questions and the TAs may discuss course logistics and content. Not all sessions may not be recorded. The ones that are recorded will be available via Canvas or via links posted in Piazza (for EdX students).

Please note that in recent semesters, a number of students commented very positively on the benefits of attending the online office hour videoconferences for this course. They are taught by the instructor and by TAs. TA office hours have been particularly helpful to learn programming skills. These videoconferences are part of the course. You’ve already paid for them with your tuition. We recommend you try to attend them as often as you can and PARTICIPATE. You can learn faster by being an active participant in these online office hours. Or you can attend them silently without sharing your audio & video.

**Student Effort**

Students are expected to devote about 13.5 to 14 hours per week to complete this course requirements. (That’s about 2 hours per day! This summer course goes 36% faster than Fall or Spring. It contains the SAME material as student learn during Fall and Spring.) This guideline encompasses all class activities, including reading the textbook and supplementary resources, watching lesson videos, participating in office hours and forum discussions, completing homework assignments, and studying for exams. Of course, students can spend as much time as necessary, but it is important to be careful not to fall behind.

**Grading**

Grades will be assigned on the following basis:

Homework Assignments (3; worth 10% each) 30%

Self-assessment Tests 10%

(Canvas for GT students/EdX for MicroMasters students)

Midterm Exam – Part 1 10%

Midterm Exam – Part 2 15%

Final Exam, Part 1 15%

Final Exam, Part 2 20%

Typically, the following grading scale will be used in the course:

* 90 – 100%: A
* 80 – 89%: B
* 70 – 79%: C
* 60 – 69%: D
* 0 – 59%: F

Scores will be rounded to the nearest integer. Please note that 80 – 89% yields a B, and a total score of 89.5% would round to 90% and get an A, while anything less than 89.5 yet more than 89 will still be a B. Similar rounding applies for the other grades.

Additional curving of the grades may be possible, depending how the course progresses and on the disparity of the students during this semester.

**Course Schedule**

Please see separate handout for the **Course Schedule** and see also copy on Canvas for GT Students / EdX for MicroMasters students.

**Readings**

The assigned pre-readings are crucial to your success in this course. Exams may include some material in readings that are not covered in the in-class lessons. Moreover, watching the video lessons alone will not sufficiently prepare you for the exams.

You must practice, try the programming exercises provided in lecture slides and in R Labs, and do the homework and online quizzes.

**Lessons**

Video lessons for this course will be housed in EdX.

**GT Students:** For more details on creating and linking your EdX account, log into the Canvas, go this course Canvas site, and complete all the steps in the section entitled *“Welcome to Data Analytics for Business.”*

**Assignments**

There will be three individual assignments to be submitted. Each assignment is equally weighted, each counting as 10% of the overall course grade. (The raw points for each assignment may vary. One assignment could have a total of 200 points and another a total of 50 points, but both carry equal weight as far as the overall course total score is concerned.)

You’ll have about two weeks to work on each assignment. Each assignment should be **submitted on Canvas for GT students/EdX for MicroMasters students by 11:59 pm EST on the days noted in the Course Schedule**. Each assignment must be submitted **no later than the deadline**.

Submission after this time (regardless of whether it is by minutes, hours or days) will received a score discount, unless you have exceptional circumstances, a TA was made aware of these, and you were granted in advance in writing a postponement. **Students are responsible** for making sure that their individual assignments are submitted in a timely manner according to the course guidelines.

Graded homework assignments will be released at least two weeks in advance, giving students opportunities to browse the assignment and organize their week’s plan accordingly. Graded assignments will be due on Wednesday evenings, two weeks later (except where noted on Course Schedule).

Students will have plenty of opportunity to ask questions during weekdays to the TAs. We’ll have LOTS of online discussions in the [Piazza.com](http://piazza.com/) platform.

Office hours will be scheduled every week on Monday with TAs, and Tuesday or Thursday nights with instructor, to address your questions about the assignments. During these “office hours” we will also answer any open questions.

**Peer-Grading Homework**

The three homework assignments will require peer-grading. This means that you will have to look at the work of 3 (random) students and grade their work. The median of the three grades will be assigned as the student’s grade.

You will have 5 days to perform this peer-grading task. If you do not do it, on time, your own grade for this assignment could be zero.

Peer-grading adds work to the students, and it also adds a lot of learning, for you’ll get to see other ways of thinking about problems and coding them. It is tedious but it helps you!

**Quizzes**

There are will be ten online quizzes or Self-Assessments that will be graded. They will be worth each 1% of your grade, for a total of 10% of your course grade. These quizzes will be administered online in Canvas for GT Students/EdX for MicroMasters students and you will have the opportunity to take them several times if you wish to improve your scores.

One of the early Self-Assessments will be peer-graded to give you a chance to practice your peer-grading skills.

**Exams**

The Midterm Exam will account for 25% of your course grade. The Final Exam will account for 35% of your overall course grade.

The Midterm Exam will cover Weeks 1 through 6 of the course. The Final Exam will be cumulative in scope and cover all of the course materials. The exams will cover concepts discussed in the readings, the lectures, and in the homework assignments.

Part 1 of each exam will be comprised of multiple-choice questions. Exams will be strictly-timed with proctoring software. No open books, notes, web browsers, or similar resources are allowed, unless otherwise stated by your professor.  The use of mobile phones and tablet devices is also prohibited. The questions will be mostly theoretical.

Part 2 of each exam will be comprised of application questions that require the use of R. In addition to answering multiple-choice questions in Part 2, you may have to upload your R code.

Review Course Schedule for release/due dates of Midterm Exam Part 1 and Part 2.

Review Course Schedule for release/due dates of Final Exam Part 1 and Part 2.

The midterm and final exams must be **submitted on Canvas for GT students/EdX for MicroMasters students by 11:59 pm EST on the days noted in the course schedule**. Any submission after this time (regardless of whether it is by minutes, hours, or days) will not be accepted. There is no grace period for taking the exam. If you have to travel on day an exam is due, please arrange to complete your work early. **It is the student’s responsibility to monitor their time and allow enough time to submit their exam before time is up.**

**Plagiarism**

Plagiarism is considered a serious offense. You are not allowed to copy and paste or submit materials created or published by others, as if you created the materials. All materials submitted and posted must be your own original work.

**Student Honor Code**

You are responsible for completing your own work. All students are expected and required to abide by the *letter* and the *spirit* of the Georgia Tech Honor Code. The teaching assistants and I will also abide by these honor codes. I am very serious about this expectation because ethical behavior is extremely important in all facets of life.

To review the Georgia Tech Honor Code, please visit <http://osi.gatech.edu/content/honor-code> . Any OMS Analytics degree student suspected of behavior in violation of the Georgia Tech Honor Code will be referred to Georgia Tech’s Office of Student Integrity. Please see also the GeorgiaTech Honor Advisory Council: <http://www.honor.gatech.edu>.

**Students with Learning Differences:**

This course offers accommodations to students with learning differences. If you need an online classroom accommodation, please contact GeorgiaTech’s ADAPTS office at

http://www.adapts.gatech.edu and let us know about your need and accommodation.

**General Comments**

* The Modules of this course follow a logical sequence
* You are responsible for completing your own work.
* Graded assignments should be completed by their due dates
* Self-Assessment tests must be completed within the time allotted

**Attendance Policy**

* Attendance in online office hours is not required, but it is recommended. You will learn better by being present, participating, and your questions in office hours could help other students learn better.
* Log in regularly into Canvas/EdX to check what’s new and complete your work, and so you do not have to spend a lot of time reviewing and refreshing yourself regarding the content.

**Communication**

* All students can and should ask questions. Online you can also all answer your fellow learners’ questions in the course discussion forums. Often, discussions with fellow learners are the sources of key pieces of learning.

**Netiquette**

* Netiquette refers to etiquette that is used when communicating on the Internet. Review the [Core Rules of Netiquette](http://www.albion.com/netiquette/corerules.html). When you are communicating via email, discussion forums or synchronously (real-time), please use correct spelling, punctuation and grammar consistent with the academic environment and scholarship.
* Conner, P. (2006-2014). Ground Rules for Online Discussions, Retrieved 4/21/2014 from <http://teaching.colostate.edu/tips/tip.cfm?tipid=128>
* Learners who do not adhere to this guideline may be removed from the course.

**Intellectual Property and Confidentiality**

We highly recommend that you avoid disclosing any confidential information in your assignments and discussion forum posts (including intellectual property and "third party" confidential information, such as information in relation to your employer that is not publicly available).

Although you are encouraged to draw on real-world experience, posting material or sharing links to material that is harassing, intimidating, or defamatory, or encourages or condones piracy or infringes on intellectual property rights is not appropriate. GeorgiaTech reserves the right to remove any postings that contravene the well-being of other students or goes against accepted integrity standards.

We would urge you to use only first names (or pseudonyms) wherever possible. You are entirely responsible for ensuring that you do not disclose any information that is protected by confidentiality undertakings – we will ensure that all information is treated in accordance with our privacy policy, but we will not sign any separate confidentiality agreements or non-disclosure agreements.

If, during the program, you disclose or create any intellectual property (for example, trading names, designs, written materials, know-how and other products of your independent thought, creativity and intellectual effort), then you accept all and any risks in relation to disclosure, including the risk that a fellow participant will use this intellectual property without your consent, or that disclosure weakens or erases any legal protections.

We won't use any intellectual property created by you and submitted in, or forming part of, your assignments without your written consent.

Our discussion forums operate on the basis of the **Chatham House Rule**: *“Participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.”* Please ensure that you take account of this Rule when posting on the discussion forum and using information learnt from discussion forum posts. <https://www.chathamhouse.org/chatham-house-rule>

Data Analysis in Business (MGT 6203)

Course Outline - Summer 2020

\*\*\*Course Outline subject to change during the semester to adapt to needs of instruction.\*\*\*

|  |  |  |
| --- | --- | --- |
| **Weeks** | **Course Topics** | **Release Dates (All dates on Friday except where noted)** |
|  | **Module 1: Basics (Weeks 1-4)** |  |
| Week 1 | Linear Regression: Simple and Multiple. Learning R Code. Real Estate Example. | 11 May 2020 (Monday) |
| Week 2 | Indicator Variables and Interaction Terms. Nonlinear Transformations and Log Models. Customer Analytics Example. | 15 May |
| Week 3 | Logistic Regression. Customer Default Example | 22 May |
| Week 4 | Treatment Effect, Randomized Controlled Experiments, and Natural Experiments. | 29 May |
|  | **Module 2: Finance (Weeks 5-6)** |  |
| Week 5 | Introduction, and Measuring Risk and Return. Measuring Risk Adjusted Performance | 5 June |
| Week 6 | Factor Investing | 12 June |
|  | **Midterm Exam: Parts 1 & 2 due on Sunday 28 June** | 19 (Mon)-28 (Sun) June |
|  | **Module 3: Marketing (Weeks 7-9)** |  |
| Week 7 | Marketing & Advertising : Traditional and Digital | 19 June |
| Week 8 | Implementing Integrated Digital Marketing | 26 June |
|  | **Module 4: Operation Management (Weeks 9-11)** |  |
| Week 9 | Predictive Marketing Across Channels.  Introduction and Managing Queues | 3 July |
| Week 10 | Statistical Process Control. Forecasting Demand | 10 July |
| Week 11 | Inventory Management  (Last 2 days of classes, then Reading Period, and Final Exams begin) | 17 July |
| Week 12 | **Final Exam: Parts 1 & 2 due on Thu 30 July** | 23 (Wed)-30 (Thu) July 2020 |

Data Analysis in Business (MGT 6203)

Course Outline - Summer 2020

\*\*\*Course Outline subject to change during the semester to adapt to needs of instruction. \*\*\*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Week | Activity | Assignment | | Peer Assessment | |
| Release Date | Due Date | Release Date | Due Date |
| Week 2 | Graded Homework #1 | May 21 @ 8:00am EST | Jun 3 @ 23:59 EST | Jun 4 @ 12:00am EST | Jun 8 @ 23:59 EST |
| Self-Assessment 1 | May 21 @ 8:00am EST | May 27 @ 23:59 EST |  |  |
| Week 4 | Graded Homework #2 | Jun 4 @ 8:00am EST | Jun 17 @ 23:59 EST | Jun 18 @ 12:00am EST | Jun 22 @ 23:59 EST |
| Self-Assessment 2 | Jun 4 @ 8:00am EST | Jun 17 @ 23:59 EST |  |  |
| Week 5 | Self-Assessment 3 | Jun 11 @ 8:00am EST | Jun 17 @ 23:59 EST |  |  |
| Week 6 | Self-Assessment 4 | Jun 18 @ 8:00am EST | Jun 24 @ 23:59 EST |  |  |
| Week 7 | Midterm | Jun 19 @ 5:00pm EST | Jun 28 @ 23:59 EST |  |  |
| Self-Assessment 5 | Jun 25 @ 8:00am EST | Jul 1 @ 23:59 EST |  |  |
| Week 8 | Graded Homework #3 | Jun 29 @ 8:00am EST | Jul 8 @ 23:59 EST | Jul 9 @ 12:00am EST | Jul 13 @ 23:59 EST |
| Self-Assessment 6 | Jul 2 @ 8:00am EST | Jul 8 @ 23:59 EST |  |  |
| Week 9 | Self-Assessment 7 | Jul 9 @ 8:00am EST | Jul 15 @ 23:59 EST |  |  |
| Self-Assessment 8 | Jul 9 @ 8:00am EST | Jul 15 @ 23:59 EST |  |  |
| Week 10 | Self-Assessment 9 | Jul 13 @ 8:00am EST | Jul 21 @ 23:59 EST |  |  |
| Self-Assessment 10 | Jul 13 @ 8:00am EST | Jul 21 @ 8:00am EST |  |  |
| Final | Final Exam | Jul 23 @ 8:00am EST | Jul 30 @ 23:59 EST |  |  |