

University of Pittsburgh  
College of Business Administration  
BUSQOM 1080: Data Analysis for Business

Prof. Michael Hamilton  
Fall Term, 2020 - 3 Credits

*E-mail:* [mhamilton@katz.pitt.edu](mailto:mhamilton@katz.pitt.edu)

*Office Hours:* Wednesday 1:00 pm - 4:00 pm

*Office:* 119A Mervis Hall (Closed)

*Zoom links and schedule will be provided on Canvas.*

---

*Web:* <http://canvas.pitt.edu>

*Class Hours:* Tu/Th 9:25-10:40 (Sec:1320)

12:40:1:55 (Sec:1300)

5:30-6:45 (Sec:1000)

*Class Room:* Online

## Course Description

Business Analytics is a strategic initiative for many companies and organizations. Many organizations use data analysis and decision science techniques to drive strategic decisions and to solve critical problems. This class will give students a background in the statistical techniques, tools, and processes necessary to solve analytic problems in business, including marketing, finance, and more. For a detailed list of the topics to be covered please see the course schedule on pg. 6 of this syllabus.

## COVID-19 Statement

**Due to the ongoing pandemic this course will be conducted fully online.**

In the midst of this pandemic, it is extremely important that you abide by public health regulations and University of Pittsburgh health standards and guidelines. While in class, at a minimum, this means you must wear a face covering and comply with physical distancing requirements; other requirements may be added by the University during the semester. These rules have been developed to protect the health and safety of all community members. Failure to comply with these requirements will result in you not being permitted to attend class in person and could result in a Student Conduct violation. For the most up-to-date information and guidance, please visit [coronavirus.pitt.edu](http://coronavirus.pitt.edu) and check your Pitt email for updates before each class.

## Recording notice for lectures

The class sessions will be recorded and made available for viewing by students at later dates. Participation in Zoom synchronous sessions grants permission to record the student's face and voice and to post such recording within the Canvas course website. While it is a better experience for all if we can see each other and you are encouraged to have your camera on, it is not mandatory. *If you are opposed to your likeness being recorded, you may turn your camera off.* If you do not want your voice to be recorded you do not have to speak and instead please use the chat function to interact with me during synchronous sessions.

## Learning Goals for the Course

The purpose of this course is to enhance students' statistical and analytical skills to the level necessary for them to deal with Data Mining and other essential material in the Business Analytics Certificate. The goals of this course are to:

- Cover material not currently included in STAT 1100, but which is preparatory to Data Mining.
- Provide software skills in a statistical package used in business (R/RStudio).
- Provide experiential exposure to business applications of the material.

Throughout the course of the semester, students will accomplish the following:

1. Develop an understanding of data analysis techniques beyond those covered in the introductory statistics courses (namely STAT 1100).
2. Develop a basic understanding of the statistical tools that can be used to implement these techniques, specifically the R programming language.
3. Develop a core understanding of analytic tools and methods that are used to solve strategic problems and how to apply them to a specific organizational objective through a series of focused data explorations.
4. Understand how to develop, organize, and deliver effective data driven reports.

## Course Prerequisites

Completion of STAT 1000 or STAT 1100 with a **B** or better.

## Course Materials

The primary material for this course will be covered in lecture (mainly via PowerPoint and supplementary documents) and presented over Zoom (with corresponding R Notebooks). All lectures and lecture materials will be made public on canvas. The material covered in this course is a curated selection of topics covered in the course texts. They are not mandatory, but may serve as a helpful reference.

## Course Texts

We will occasionally reference the following texts, they are meant to be supplemental to the main course material which are the lectures and notebooks. You are **not** required to purchase the first two books, the later two are available for free online.

1. *Business Analytics Using R - A Practical Approach* [BUR]  
Authors: Umesh R. Hodeghatta & Umesh Nayak  
ISBN: 978-1-4842-2513-4 (Print) 978-1-4842-2514-1 (Online)
2. *R for Marketing Research and Analytics* [RMRA]  
Authors: Chris Chapman & Elea McDonnell Feit  
ISBN: 978-3-319-14435-1 (Print) 978-3-319-14436-8 (Online)
3. *R for Data Science* [RDS]  
Authors: Garrett Grolemund & Hadley Wickham  
Available online: <https://r4ds.had.co.nz/>
4. *Exploratory Data Analysis with R* [EDA]  
Authors: Roger D. Peng  
Available online: <https://bookdown.org/rdpeng/exdata/>

## Software

This course will use the statistical programming language R and the interactive development environment (IDE) RStudio. Specifically we will use

1. R 4.0.2\* or the latest version
2. RStudio (Desktop)

Both R and RStudio can be downloaded for Mac, Windows, and Linux at [www.rstudio.com](http://www.rstudio.com). Detailed download and installation instructions and a walk through video will be provided prior to first use of R.

\*We will be using R to solve the problems in this course. I will use the above version for all demonstrations during class. However, if you already have a previous version installed on your device, it should have the same capabilities. If possible, I would encourage students to invest in a second monitor.

## Canvas

We will be using Canvas (<http://canvas.pitt.edu>), the University of Pittsburgh's new course management system (replacing courseweb/blackboard). I will post all of the materials necessary for this class as well as the assignments. You will find the following items there:

- Syllabus - Updated as needed throughout the term.
- Course schedule - Containing week-by-week readings, supplementary materials, assignments (subject to updates during the semester) and more.

- Course handouts - Lecture slides, R notebooks, assignment solutions, practice problems, class recordings, summaries of concepts, and more.
- Lecture recordings - Will be posted alongside the slides, they are meant to be viewed prior to the assigned online meeting time so that they can be discussed.
- Synchronously recorded zoom sessions - Will always be uploaded within 24 hours of the scheduled in class time.
- Course assignments - Will be submitted online using the links provided on Canvas.
- Your grades - Will be available through the *Grades* tab on Canvas.

## Course Method of Evaluation

Your course grade will be a combination of the following four categories.

Course Requirement	Percentage of Final Grade
Assignments	40%
Take Home Midterm	25%
Final Project	35%

There are ten assignments total, I will drop the lowest assignment grade automatically. Final grades will be curved at the discretion of the instructor.

### Assignments

Assignments will be submitted as R Notebooks (.Rmd files), and must be organized and complete to be acceptable. I expect you to clearly label your work and present all the steps you followed to obtain the final results (including all necessary data manipulation and code to arrive at your answer!).

All files submitted for each assignment should be saved with your name and the assignment number. I expect you to complete the assignments on time. Late assignments will not be accepted unless otherwise approved by the instructor - **for extensions related to your health or the health of a family member please contact me to request an extension.**

Please submit the online assignments using the Assignment tab on the Canvas page for this class. You are permitted and encouraged to discuss the problems with your classmates, **however you must complete the assignments on your own. If you received assistance from myself or a classmate, you must clearly and explicitly list your collaborators at the top of your submission,** failure to do so will result in a automatic 0 for the assignment. All work should be submitted independently unless otherwise instructed.

### Take Home Midterm

One take home exam will be given as identified in the course schedule. You will have 24 hours to complete the examination. All students are expected to take the midterm on the scheduled day. In general, there will be no make-up exams. However, in the event an exam is missed due

to either a pre-approved absence by the instructor or due to a documented illness, arrangements may be made to make-up the missed exam. Make-up examinations are given at the discretion of the instructor.

### **Term Project**

There will be one term project to be completed during the second half of the semester. Students will complete their projects *alone*. The project will consist of a written report, and will require students to demonstrate knowledge and mastery of various methodologies presented in the course. More details about both the project will be provided *after* the midterm.

### **Zoom Decorum**

Please have your webcam set up and ready at the designated time, and do not hesitate to (politely) mute or turn off your camera if it is necessary for whatever reason. Please attend your designated zoom session unless otherwise unable. Please be respectful to your fellow classmates and your instructor as we navigate this difficult time.

### **Code of Ethics and Academic Integrity**

Copying another student's work, using unauthorized aids during a quiz or exam or providing assistance to another student during a quiz or exam is cheating. Violations of academic integrity will be dealt with according to the CBA academic integrity policy. The minimum sanction will be receiving a zero for the assignment or exam. The CBA academic integrity policy can be found at <https://cba.pitt.edu/academics/standards-and-policies/> for your reference. Any violation of the Guidelines for Academic Integrity will lead to appropriate consequences. Please make sure you are familiar with these guidelines.

### **Students with Disabilities**

If you have a disability for which you are or may be requesting an accommodation, you are encouraged to contact both your instructor and Disability Resources and Services (DRS), 140 William Pitt Union, (412) 648-7890, [drsrecep@pitt.edu](mailto:drsrecep@pitt.edu), (412) 228-5347 for P3 ASL users, as early as possible in the term. DRS will verify your disability and determine reasonable accommodations for this course.

## Course Schedule (Subject to Change)

All dates and deadlines in the below course schedule are subject to change. The dates posted here are only to get a sense of how the course will progress. Almost surely, the course schedule will need to be adjusted to meet the needs of the class and the dates will become incorrect. I will not be updating this schedule, for due dates of assignments and projects see Canvas. Canvas deadlines will **always supersede** the deadlines listed here.

Lecture	Date (F'20)	Topic(s) to be Covered	Book Chapter(s)	Assignment Due (on the following Friday!)
1	8/20	Intro to Business Analytics	[BUR] Ch. 1	-
2	8/25	Intro to R	[BUR] Ch. 2, [RMRA] Ch. 1+2	Install R
3	8/27	Intro to R II	[BUR] Ch. 3, [RMRA] Ch. 2	Swirl Lesson
4	9/1	Intro to R III	[BUR] Ch. 3, [RMRA] Ch. 2	-
5	9/3	Descriptive Analytics and Graphical Representations	[BUR] Ch. 4, [RMRA] Ch. 3	Intro to R Assignment
6	9/8	Descriptive Analytics and Graphical Representations	[BUR] Ch. 4, [RMRA] Ch. 3	-
7	9/10	Comparing Groups	[RMRA] Ch. 6	Descriptive Analytics Assignment
8	9/15	Comparing Groups II	[RMRA] Ch. 6	-
9	9/17	Simple Linear Regression	[BUR] Ch. 8	Comparing Groups Assignment
10	9/22	Simple Linear Regression II	[BUR] Ch. 8	-
11	9/24	Multiple Regression	[BUR] Ch. 9, [RMRA] Ch. 7	Regression Assignment
12	9/29	Multiple Regression II	[BUR] Ch. 9, [RMRA] Ch. 7	-
13	10/1	Model Building, Variable Selection	[BUR] Ch. 9, [RMRA] Ch. 7	Multiple Regression Assignment
14	10/6	Model Building, Variable Selection II	[BUR] Ch. 9, [RMRA] Ch. 7	-
15	10/8	Review for Midterm	-	Midterm Review Problems
16	10/13	Take Home Midterm	-	-
17	10/15	Discuss Midterm, Final Project, Transformations	-	Take-home Portion of Midterm
18	10/20	Data Wrangling, Data-Driven Rhetoric, Markdown	[EDA] Ch. 4, 5	-
19	10/22	Data Processing and PCA	[RMRA] Ch. 8, [EDA] Ch. 13	-
20	10/27	Logistic Regression I	[BUR] Ch. 10, [RMRA] Ch. 9	Project Proposal
21	10/29	Logistic Regression II	[BUR] Ch. 10, [RMRA] Ch. 9	-
22	11/3	Introduction to Machine Learning	[RMRA] Ch. 8	Logistic Regression Assignment
23	11/5	Supervised Learning I	[RMRA] Ch. 8	-
24	11/10	Supervised Learning II	[RMRA] Ch. 8	-
25	11/12	Unsupervised Learning I	[EDA] Ch. 12	Supervised Learning Assignment
26	11/17	Unsupervised Learning II	[EDA] Ch. 11	-
27	11/19	Course Review, Example Term Projects	-	Unsupervised Learning Assignment
-	11/30	Term Project Final Report Due	-	-