Data Science and Business Intelligence
2 Credits

BU.330.780.T1

Tuesdays 8:15 AM – 11:15 AM
March 25, 2025 – May 13, 2025

Spring 2025
Washington DC, Bloomberg Center, room B232

Instructor
Changmi Jung, Ph.D.
Associate Professor, Carey Business School

Contact Information
E-mail Address: changmi@jhu.edu
The Johns Hopkins University Carey Business School
100 International Drive, Room 1331
Office Phone Number: (410) 234-9267

Office Hours
Tentative hours: Thursdays 11:30 AM – 1:30 PM (EST): please use the link on the Canvas (Home > Course Team) to reserve a slot.
I am available by email. Please give me 24 hours to respond.

Teaching Assistant
Keran Jiang, MS
E-mail: kjiang19@jh.edu

Binjun Jiang, MS
E-mail: bjiang13@alumni.jh.edu

All teaching assistants' office hours: please check Canvas (Modules > Course Overview > Course Team)

Required Texts & Learning Materials
- Foster Provost and Tom Fawcett, *Data Science for Business: What you need to know about data mining and data-analytic thinking,* O'Reilly Media, ISBN 978-1449361327

Technology Requirements
Software:
In this course, you will use the following software. Download links and guides are posted on the course Canvas site. This course assumes that students already have a good understanding of basic R programming and will not repeat R installation and basic syntax. Students are responsible for Installing the software before the first class and bringing their laptops for every class.
- **R**
  - Main software that will be used throughout the course
  - Downloadable and installation: please check course Canvas (Modules > Support Materials)
- **R Studio**

- Provide an easier interface for R programming
        - Downloadable and installation: please check course Canvas (Modules > Support Materials)
    • **Tableau**
        - Will be used mainly for exploratory analysis and data visualization.
        - Download and installation: please check course Canvas (Modules > Support Materials)
    • **Power BI**
        - Will be used mainly for exploratory analysis and data visualization.
        - Download and installation: please check course Canvas (Modules > Support Materials)

Course Description

This course introduces a set of fundamental principles and a framework that guides students through the process of extracting business insights from data using data science methodologies to generate competitive advantage. We will discuss how the ubiquity and massiveness of digital data and the application of data science have changed competitive landscapes. The data science methodologies covered in this course include data visualization, supervised segmentation, and text analysis, which are widely used in the real world. The topics and cases discussed in this course cover various fields, including marketing, finance, healthcare, and more.

This course is not a statistics or computer programming course. The emphasis will be on applications and interpretations of the results from data science projects to make business decisions. Students will apply these techniques in hands-on exercises as we analyze strategic concepts.

Prerequisite(s)

BU.510.650 Data Analytics or having basic R skills
Students are required to have knowledge of R and R Studio, either by completing a Data Analytics course as a pre-requisite or by demonstrating existing knowledge and background in R.

Learning Objectives

By the end of this course, students will be able to:

1. Understand the potential of data science in today's data-rich environment
2. Understand a data science project as a process
3. Identify appropriate data and select the most effective data science approach based on the problem and data at hand
4. Understand what the techniques do and how they work based on fundamental principles
5. Know how to represent data in various formats and visualize data for exploratory analysis
6. Assess the performance of different models and examine their effectiveness
7. Critically evaluate a big data project to assess business value/usefulness and its limitations

To view the complete list of the Carey Business School's general learning goals and objectives, visit the Carey website.

Attendance

Attendance and active participation in class are integral components of students' course grades. Students are expected to be present at every scheduled class. Should you need to miss a session, please inform the instructor as soon as possible. Students are accountable for keeping up with all class materials, both academic and administrative, which can be done by using the course Canvas and coordinating with classmates. Not attending classes hinders students' ability to meet the course objectives. Excessive absence (more than once) will be reported to student affairs, not to mention negatively impacting the grades. In cases of unavoidable commitments that conflict with class times, please notify the instructor at least 48 hours in advance and coordinate with your team members for any group work or assignment submissions.

Classroom Protocol
This is an in-person class, and therefore, it will be conducted exclusively on-site. Under any circumstances, there will be no option for remote participation or access to class recordings.

In the event of a class being shifted to an online format due to adverse weather conditions or instructor emergencies, it is mandatory for all students to keep their webcams active during the sessions. Failure to do so without prior approval will be treated as an absence. Active participation in class is strongly recommended. During these remote Zoom sessions, please refrain from private conversations with fellow students; instead, direct your queries and comments to the whole class for a more inclusive discussion.

Assignments
There will be several individual assignments. Each assignment should be submitted through Canvas by the specified deadlines. No late submission will be accepted.

Quizzes
The questions will cover data science topics that have been covered in class. If you develop a thorough understanding of the tasks in assignments and group projects, you will be ready to answer all questions. Thus, please double-check if you have a firm understanding of the tasks and can interpret the results.

Group Research Project Report & Presentation
Students will form groups (more details regarding the limit to the maximum number of students and others will be addressed in the first class). Each group will work on a data science project of their interest: You will select data, design a data science task, analyze the data, and describe the results. Teams will present their results in the last week and submit a final report. More details are provided in a separate instruction file (Final Project Instruction in Canvas).

Assignment	Group or Individual	Weight
Attendance and class participation	Individual	10%
Individual assignment	Individual	25%
Quizzes	Individual	30%
Group research project	Group	35%
Total		100%

Grading
The grade of A is reserved for those who demonstrate extraordinary performance as determined by the instructor. The grade of A- is awarded only for excellent performance. The grades of B+ and B are awarded for good performance. The grades of B-, C+, C, and C- are awarded for adequate but substandard performance. The grades of D+, D, and D- are not awarded at the graduate level. The grade of F indicates the student's failure to satisfactorily complete the course work. For Core/Foundation courses, the grade point average of the class should not exceed 3.35. For Elective courses, the grade point average should not exceed 3.45.

Tentative Course Calendar
Instructors reserve the right to alter course content and/or adjust the pace to accommodate class progress. Students are responsible for keeping up with all adjustments to the course calendar.

Week	Topic	Reading & In-class Exercise	Deliverables (due *before* each class)
1	Introduction & Data Preparation	• Chapter 1. Introduction: Data-Analytic Thinking • Chapter 2. Business Problems and Data Science Solutions	
2	Data Visualization	• No reading • R ggplot and Tableau exercise	• Assignment 1
3	Supervised Segmentation (I)	• Chapter 3. Introduction to Predictive Modeling: From correlation to supervised segmentation	• Assignment 2 • Group project proposal
4	Supervised Segmentation (II)	• Chapter 4. Fitting a Model to Data	Quiz 1
5	Model Validation	• Chapter 5. Overfitting and Its Avoidance	• Assignment 3
6	Model Performance: Evaluation and Probability Estimation	• Chapter 7. Decision Analytic Thinking	• Assignment 4 • Group project status report
7	Data-Driven Decision Making	• Might have to move the Text mining here • Latanya Sweeney, "Discrimination in Online Ad Delivery," acmqueue 11, no. 3 (April 2013), https://queue.acm.org/detail.cfm?id=2460278. • Lambrecht, A., and Tucker, C. (2019). Algorithmic Bias? An Empirical Study of Apparent Gender-Based Discrimination in the Display of STEM Career Ads, Management Science (65:7), pp. 2966–2981. (https://doi.org/10.1287/mnsc.2018.3093). • Berinato, S. (2016). With big data comes big	Quiz 2
8	**Team project presentation**		Presentation slides, Project reports, R Markdown file

Carey Business School Policies and General Information

Please note that failure to become acquainted with Carey policies will not excuse any student from adhering to these policies.

Canvas Site
A Canvas course site is set up for this course. Each student is expected to check the site throughout the semester as Canvas will be the primary venue for outside classroom communications between the instructor and students. Students can access the course site at https://canvas.jhu.edu/.

Technical Support
24/7 technical support for questions regarding Canvas, Zoom, and other technical issues is available. Please refer to Carey's Academic Resources webpage for contact information and other details.

Students with Disabilities - Accommodations and Accessibility
Johns Hopkins University values diversity and inclusion. We are committed to providing welcoming, equitable, and accessible educational experiences for all students. Students with disabilities (including those with psychological conditions, medical conditions, and temporary disabilities) can request accommodations for this course by providing an Accommodation Letter issued by Student Disability Services. Please request accommodations for this course as early as possible to provide time for effective communication and arrangements. For further information or to start the process of requesting accommodations, please contact Student Disability Services at the Carey Business School.

Academic Ethics Policy
Carey expects graduates to be exemplary global citizens in addition to innovative business leaders. The Carey community believes that honesty, integrity, and community responsibility are qualities inherent in an exemplary citizen. The objective of the Academic Ethics Policy (AEP) is to create an environment of trust and respect among all members of the Carey academic community and hold Carey students accountable to the highest standards of academic integrity and excellence.

It is the responsibility of every Carey student, faculty member, and staff member to familiarize themselves with the AEP and its procedures. Failure to become acquainted with this information will not excuse any student, faculty, or staff member from the responsibility to abide by the AEP. Please contact the Office of Student Affairs if you have any questions. For the full policy, please visit the Academic Ethics Policy webpage.

Student Conduct Code
The fundamental purpose of the Johns Hopkins University's regulation of student conduct is to promote and to protect the health, safety, welfare, property, and rights of all members of the University community as well as to promote the orderly operation of the University and to safeguard its property and facilities. Please contact the Office of Student Affairs if you have any questions regarding this policy. For the full policy, please visit the Student Conduct Code webpage.

Commitment to Respect
Respectful behavior creates an environment within the Carey Business School where all are valued and can be productive. Carey defines respectful behavior as conduct that, at a minimum, demonstrates consistent courtesy for others, including an effort to understand differences. As such, all in the community agree to the Carey Commitment to Respect, which states that we all strive to show that we value each other's human dignity and our differences, and to choose behavior and language that demonstrates mutual respect. Please visit the Commitment to Respect webpage to learn more about the expectations and resources available.

Classroom Policies for All On-Site and Remote-Live Classes
Carey is committed to maintaining the highest standards of excellence in all forms of instruction. To that end, we have developed policies and procedures for all classes offered in on-site and remote-live formats. These policies will govern all courses occurring in these formats, and all students are expected to familiarize themselves with and adhere to these policies.

Student Success Center
The Student Success Center offers assistance in core writing and quantitative courses. For more information, visit the Student Success Center webpage.

Other Important Policies and Services
Students are encouraged to consult the Student Handbook and Academic Catalog and Student Services and Resources for information regarding other policies and services. For your convenience, there is a singular website students can visit to learn about all JHU and Carey policies.

Copyright Statement
Unless explicitly allowed by the instructor, course materials, class discussions, and examinations are created for and expected to be used by class participants only. The recording and rebroadcasting of such material, by any means, is forbidden. Violations are subject to sanctions under the Academic Ethics Policy.