



INFO 4390: Advanced Predictive Modeling with R

<p><u>Term and Credits:</u> Winter 2022 4 Credits 4390-1 CRN 2860</p>	<p><u>Location:</u> T/Th 6:00pm – 7:50pm DCB 130 (First week on Zoom) You will be required to review some material outside of class which will be delivered through Canvas. Make sure you have a good internet connection during class for access to Canvas. Also make sure your Canvas notifications are set appropriately.</p>
<p><u>Instructor:</u> Name: Kellie Keeling Department: Business Information & Analytics Office Hours DCB 590/Zoom: T/Th 5-5:50p https://udenver.zoom.us/my/kkeeling Chat using MS Teams: (Join Link) and code: r8vuuuu) Make an appointment: https://bit.ly/KeelingBooking</p>	<p><u>Communication Conduct:</u> Feel free to refer to me as Dr. Keeling, Professor, or Kellie as you feel comfortable. Email or Teams is usually the best way to contact me. If I haven't responded in 36 hours, feel free to resend your message. I will send class level communications via Canvas announcements. I will typically initiate communication with individual students directly through your DU email or through Canvas email. Email: kkeeling@du.edu Office Phone: 303-871-2296 (forwards to my cell)</p>

COURSE DESCRIPTION:

This course serves as an introduction to advanced predictive modeling and statistical learning using the R statistical software. Specific topics include linear, non-linear, and logistic regression, classification, resampling methods, and non-linear regression, tree-based methods, support vector machines, and neural networks. The students will learn how to communicate their results (business reports, dashboards, etc.) of the various modeling exercises and projects using R Studio and the RMarkdown suite of tools.

PREREQUISITES/CO-REQUISITES:

Prerequisites: INFO 4300 or INFO 3200

LEARNING OUTCOMES:

Principal Content Elements:

1. R programming knowledge including cleaning, manipulating, and analyzing data.
2. Advanced business predictive modeling and statistical learning techniques and tools using R and Studio
3. Business presentation techniques including reports and dashboards using R and RStudio.

PROGRAM LEARNING OUTCOMES:

1. Ethical Decision-Making - Graduates will recognize and consider the legal and ethical implications of business decisions.



2. Technical Knowledge - Graduates will possess basic knowledge of the functional areas of business and how they integrate to achieve organizational goals.
3. Communication and Interpersonal Effectiveness - Graduates will communicate effectively and have the interpersonal skills necessary to collaborate effectively in a business environment.
4. Critical Thinking - Graduates will apply critical thinking skills to identify and resolve business issues.
5. Global/Intercultural Competence - Graduates will be prepared to meet the challenges and opportunities of intercultural differences that impact business organizations in a global society

STUDENT LEARNING OUTCOMES:

1. Apply predictive modeling and statistical learning methodology to a variety of business problems using R and RStudio.
2. Apply innovative, interactive data visualization tools/methods to both the raw data and model results as a means for communicating business insights.
3. Derive a set of conclusions and business recommendations from these models.
4. Learn how to effectively deliver the results of the modeling exercises and projects in technical and non-technical manners.
5. Continue to develop the critical thinking and problem solving skills necessary to produce a business decision or recommendation from a raw data set.

REQUIRED MATERIALS:

1. An Introduction to Statistical Learning with Applications in R, ISBN: 978-1461471387, Publisher: Springer. You can get a digital copy from the bookstore and/or it is also available online for free at <http://www-bcf.usc.edu/~gareth/ISL/ISLR%20Seventh%20Printing.pdf>
2. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data, ISBN 978-1491910399, Publisher: O'Reilly. (Note: This book is available online for free at <http://r4ds.had.co.nz/> .)
3. The R statistical software is often referred to as the lingua franca of data science. As such, we will rely extensively on this software, along with RStudio integrated development environment (IDE). Please download both R and RStudio prior to the first day of class or make sure you have gotten into the R Studio Server. Both R and RStudio are available for Windows, Macs, and Linux computers.

POLICIES, AND RESOURCES:

Class Policies.

Classroom Attendance: Please plan to join us on Zoom the first week and then in DCB 130 for the remainder of our class times.

Communication: If you are having difficulty with the course material, please contact me at your earliest convenience. If you are having an emergency situation, please let me know so we can plan accordingly. If you contact me with questions, I plan on responding within 36 hours. We will use **MS Teams** for course collaborations and this is often a quick way to get a response from me: [Join Link](#)



Individual Work: For individual work, you may discuss homework so long as each person is using their own technology to solve problems and turning in their individual work. Looking over someone’s shoulder or watching as they do work, no matter how much you feel like you are participating, is not working together. Turning in individual work done by someone else is academic dishonesty and can result in penalties ranging from zero points on the assignment to an “F” in the class and dismissal from the class.

Extra Credit: The syllabus reflects a fair and accurate assessment of your skills in the class. Under no circumstances will you be given the option to complete extra credit to make up for missing assignments and/or to raise your grade. But you will be given the opportunity for **1 Homework Oops Ticket and 2 Group Task Oops Tickets** which will allow you turn in late work without an excuse.

GRADING STRUCTURE, SCALE, AND POLICIES:

A: 93-100%; A-: 90-92.9%; B+: 87-89.9%, B: 83-86.9%; B-: 80-82.9%; etc.

Student performance will be evaluated and calculated on the items below:

ASSESSMENTS:

Activity	Percentage
Pre Class: DataCamp 7 Chapters	7%
Pre Class: Reading/Lecture Quizzes	20%
In Class: Group Task Mini Assignments	20%
Homework (4)	23%
Project	30%
Total	100%

DataCamp Chapters: At the beginning of the course, there will be 7 DataCamp chapters assigned to help everyone come up to speed with the R statistical programming language. Sign up for the class “Adv Predictive Analytics w R Wn 22” with the link below. I will automatically be able to see your progress on these assignments.

https://www.datacamp.com/groups/shared_links/6fe900b0e8718047e7d3b6ac31ef1d40a985d129dd32516888ee75f0cd29d782

Reading/Lecture Quizzes: Students are expected to read and study the assigned textbook pages in addition to watching lecture videos before class. Topics that are not covered in class, but are still in the text are eligible to appear on assignments. Reading is encouraged prior to class to improve learning. To help motivate you, there will be quizzes posted in Canvas that will be due before class.

Homework: There will be a homework posted about every 4 classes for a total of 4 during the quarter. The final deliverables will be made using R Markdown.

Project: The focus will be on data analysis and a dashboard presentation to your peers using the skills obtained in this class.

Group Tasks: Each class there will be at least one short assignment to complete to review the material from the reading and in class notes. Once you submit your answer, you will have access to my answers for the assignment. These will count as participation credit and you will automatically receive full credit. But note, I



will randomly check these and if you did not attempt all the work with a good faith effort, I will penalize you by adding a late penalty based on the following scale. I will encourage you to work with others during our class time on these assignments.

Strong Effort	Good Effort	Good Effort	Moderate Effort	Other
Completed All	Completed All	Completed Most	Completed Some	
100%	93%	85%	75%	0%

Course Rules/Requirements

1. Everything must be turned in by the indicated due dates on Canvas. If it is due by 11:59pm, I won't count it late as long as it is in by 8a the following morning. You might want to start early just in case you have a computer failure.
2. Late work will be accepted with a penalty of 12 percentage points a day (.5% per hour) unless you have requested an Oops ticket. To request an Oops ticket – put a comment with your submission on Canvas of the form "I request 1 of 1 HW Oops Ticket" or "I request 1 of 2 Group Task Oops Ticket".

Syllabus Policy: This syllabus is subject to change based on the needs of the class, and at the discretion of the instructor.

Tentative Schedule:

		TUE	THU
Week 0	Complete DataCamp Courses (Intro to Tidyverse/Intermediate R)	BEFORE CLASS STARTS	
Week 1	R/R Studio Intro/Tidyverse and Control Logic/ Statistical Learning	Jan 4	Jan 6
Week 2	Simple/Multiple Linear Regression	Jan 11	Jan 13
Week 3	Logistic Regression / Discriminant Analysis	Jan 18	Jan 20
Week 4	Cross Validation / Bootstrap Resampling	Jan 25	Jan 27
Week 5	Ridge Regression / Lasso	Feb 1	Feb 3
Week 6	Smoothers: Polynomials and Splines / Gen Additive Models	Feb 8	Feb 10
Week 7	CART Trees	Feb 15 (Zoom)	Feb 17 (No class)
Week 8	Bagging/Random Forest	Feb 22	Feb 24
Week 9	Support Vector Machines / Neural Networks	Mar 1	Mar 3
Week 10	Deep Learning	Mar 8	Mar 10
Week 11	Present Projects	Mar 15	



UNIVERSITY EXPECTATIONS, POLICIES, AND RESOURCES:

Students with Disabilities. A student who qualifies for academic accommodations because of a disability must submit a Faculty Letter to the instructor from the DU Disability Services Program (DSP) in a timely manner, so that the needs of the student can be addressed. Accommodations will not be provided retroactively, e.g., following an exam or after the due date of a project. DSP determines eligibility for accommodations based on documented disabilities. DSP is located in Ruffatto Hall, 1999 E. Evans Ave. (303-871-2278).

Inclusive Learning Environments.

- In this class, we will work together to develop a learning community that is inclusive and respectful. Our diversity may be reflected by differences in race, culture, age, religion, sexual orientation, socioeconomic background, and myriad other social identities and life experiences.
- The goal of inclusiveness, in a diverse community, encourages and appreciates expressions of different ideas, opinions, and beliefs, so that conversations and interactions that could potentially be divisive turn instead into opportunities for intellectual and personal enrichment.
- A dedication to inclusiveness requires respecting what others say, their right to say it, and the thoughtful consideration of others' communication.
- Both speaking up and listening are valuable tools for furthering thoughtful, enlightening dialogue. Respecting one another's individual differences is critical in transforming a collection of diverse individuals into an inclusive, collaborative and excellent learning community.
- Our core commitment shapes our core expectation for behavior inside and outside of the classroom. Office of Diversity, Equity, and Inclusion website (<https://www.du.edu/diversity-inclusion/index.html>).

University Expectations. Please review the University Expectations on the Daniels College of Business syllabus webpage (<http://daniels.du.edu/university-expectations/>)

- University of Denver Honor Code
- Policy Concerning Official Communication
- Students with Disabilities
- Policy Concerning Religious Accommodations
- Policy Concerning Emergency Procedures
- Policy Concerning Conflicts of Interest, Including Gifts from Students