

STA 235H - Introduction

Fall 2023

McCombs School of Business, UT Austin

Welcome to STA 235H
Data Science for Business Applications

Introductions

About the instruction team

Prof: Magdalena Bennett, Ph.D.

- Assistant Professor in the Stats Group (IROM department)
- Ph.D. in Economics of Education, Columbia University
- Research: Causal Inference (+ ML) applied to social policies (e.g. education).

T.A.: Pedro Lima (Ph.D. student)

T.A.: Emma Costa (3rd-year Honors)

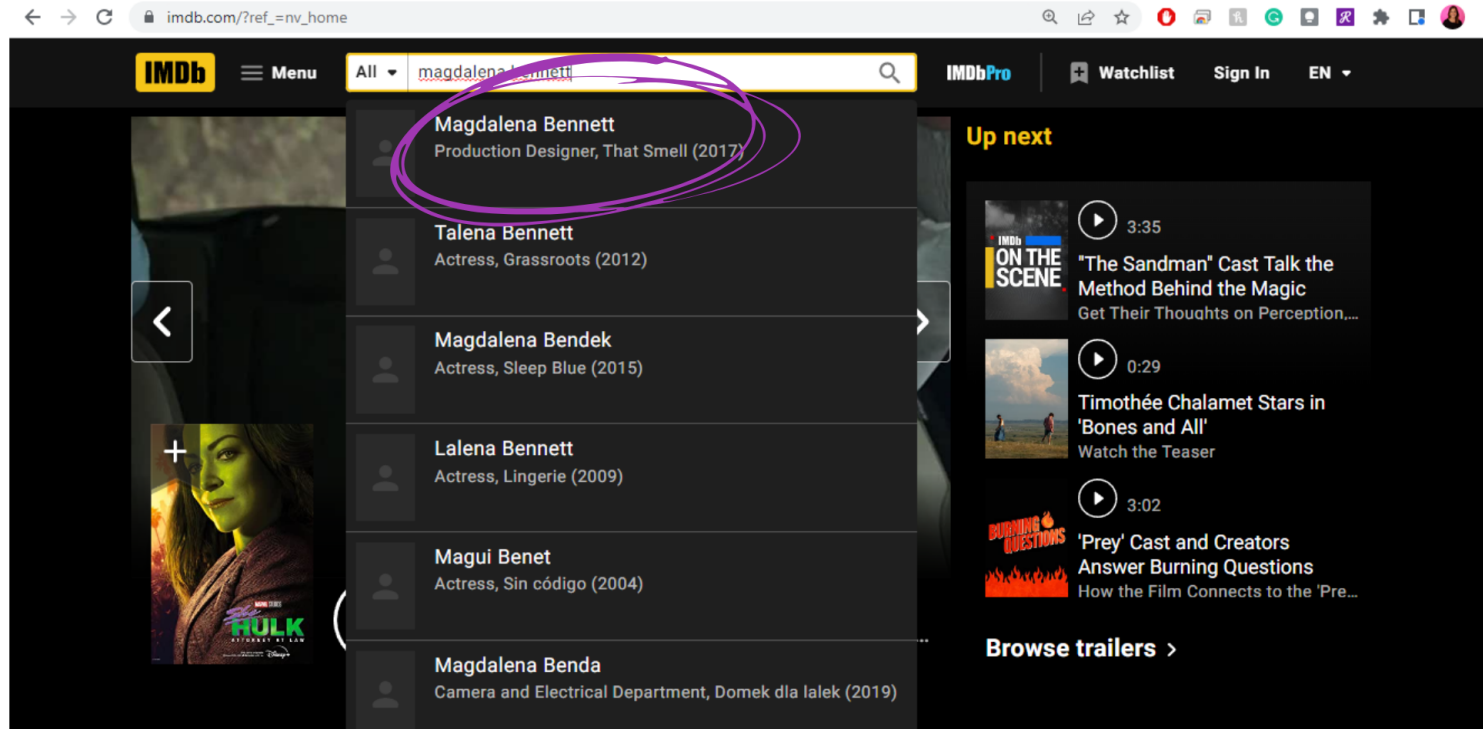
T.A.: Diego Robbins (3rd-year Honors)

Introduce yourself!

Interesting (or uninteresting) fact about yourself

Interesting fact about me?

... I have a credit on IMDB, the movie database.



Introduce yourself!

Interesting (or uninteresting) fact about yourself

Let's review the syllabus

Please, read the syllabus!

About this course

- **Objective:**

"[G]ain the tools you need to tackle real-world problems from a quantitative perspective."

You don't need to be a data scientist for this class to be useful!

How, when, and where?

- **In-person (Fall 2023):** 2 hrs/week in this classroom.
- **Drop-in Office Hours:**

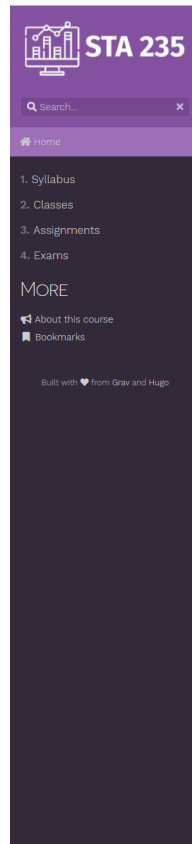
Prof. Bennett:
Wed 4:30 - 5:30 PM
Thu 4:00 - 5:30 PM

T.A.s:
Weekly + HW weeks (*TBD*)

- Other times available upon request

How, when, and where? (Cont.)

<https://sta235.com>



STA 235

search...

Home

1. Syllabus
2. Classes
3. Assignments
4. Exams

MORE

- About this course
- Bookmarks

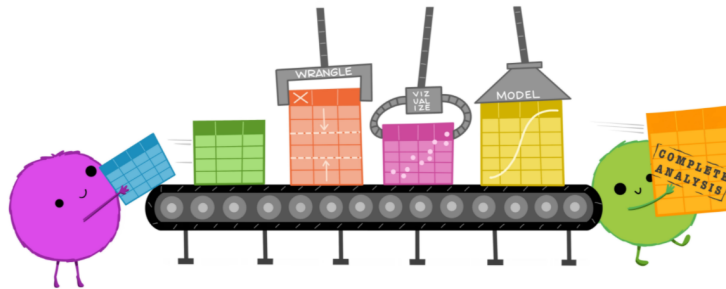
Built with ❤️ from Grav and Hugo

STA 235H

DATA SCIENCE FOR BUSINESS APPLICATIONS

Welcome to Data Science for Business Applications!

The objective of this course is for you to gain the tools you need to tackle real-world problems from a quantitative perspective. We will be covering topics on regression modelling, causal inference, and predictive modelling. You will have the opportunity to be exposed to an array of different real-world examples, get hands-on experience in working with data, and improve your R coding skills for data science.



Classroom Norms

- Please, **be on time**.
- **Participate and ask questions!** (cold-calling can be used to loosen the atmosphere)
- **Bring your laptop:** We will be doing in-class coding (let me know if you have any issues with this point).



What will you need?

- A **laptop** to bring to class.
- **R & R Studio**
- **Required Books:**
 - Angrist, J. & J. Pischke. (2015). "Mastering Metrics". Princeton University Press. (*Buy used or new*)
 - James, G et. al. (2021). "An Introduction to Statistical Learning with Applications in R". Springer. (*Available online*)



How to succeed in this course?

- **Attend class.**
- Slides are uploaded before class (*not self-contained*). Take notes but focus on **understanding**.
- **Ask questions** during class!
- Complete all **readings** and **assignments** by the assigned date.
- Get an **early** start on assignments and **follow the submission guidelines**.

Assignments, Exams, and Grading

- **Just in Time Teaching (JITT) assignments (10%):**
 - Short online quizzes about readings and/or material.
 - Submit by 11:59 PM on Saturday (for Mon class) or Monday (for Wed class) before that week's class.
 - Graded for completion (new material) and correctness (for material already seen). You can re-take it as many times as you want!
- **6 homework assignments (35%):**
 - All in Canvas.
 - Assignments include both written questions and code (submit R script).
 - You can drop one assignment (only 5 will count – can't drop the last HW).
 - You need to be responsible for your own work!

[Read submission guidelines](#)

Assignments, Exams, and Grading (Cont.)

- **Midterm (25%) and Final Exam (25%):**
 - In-class exams.
 - Open book (offline resources), but no online resources.
 - Final exam is cumulative.
- **Attendance/Participation (5%):**
 - Attendance will be taken on 5-7 random classes. You can be absent in one of them without penalty.
 - If you miss more than one (1) of those classes, you can make up with participation (see Syllabus for details).

Assignments, Exams, and Grading (Cont.)

- Assignments and exams are **usually curved**
 - Final grade **will not be curved**.
- Cutoffs for final letter grade:

Grade	A	A-	B+	B	B-	C+	C	C-	D	F
Cutoff	94%	90%	87%	84%	80%	77%	70%	65%	60%	<60%

- Cutoff scores are strict (no rounding)

Communicating with the instruction team

Email address: m.bennett@austin.utexas.edu

- Use the subject **STA 235H - Your subject**.
- Email me directly for questions related to course administration.
- Usually respond in 1 business day.
- General questions should be posted on Canvas (Discussion Board)
- **Please, do not send messages through Canvas.**

Collaborations and Academic Integrity

- **You are encouraged to form study groups!**
 - Studying or discussing class material with others does **not** mean you can copy other's work.
 - Students are responsible for their own work. All of it.
- **Do not share your files with other students**
 - If we find any evidence of copying or plagiarism, all students involved will be subject to disciplinary measures.
- **Remember to give credit where credit is due!**
 - Use citations and references when you use someone else's work.

What questions do you have?

A brief motivation

What is Data Science?



Data Science tasks

By **Hernán, Hsu, and Healy:**

Description

Prediction

Causal Inference

Data Science tasks

Can we classify our customers into different segments?

What is the probability of a shopper coming back to our website?

What is the effect of increasing our advertising budget on our total revenue?

We'll review all of these in this class!

After this course...

1) Bridge the gap between the "what" and the "how"

2) Be critical consumers of "Data Science"

Some notes before the break

- **Bootcamp** session on Tuesday 22nd (MEZ 1.306): Focus on R.
- "Disability & Access (D&A) is seeking the assistance of students to serve as **volunteer notetakers**."
 - Volunteers will be eligible to receive volunteer hours in appreciation for their time.
 - If you are a good notetaker and interested in helping other students, please contact me after class.