

Meetings:

Tuesdays and Thursdays

Room: LART 403

10:30 - 12:20 pm

Final Exam: December 14th (Thursday) 10am-12:45pm

Professor: Dr. Melissa Baker

Email: mnbaker@utep.edu

Office: Benedict 301

Office Hours: By appointment

Teaching Assistant: Lydia Torres

Email: Intorres4@miners.utep.edu

Office: Benedict 103

Office hours: Wednesdays and Thursdays 3:30-5pm

Catalog Description

This class is designed to provide an introduction to topics and skills necessary to consume, engage with, and critique social science research. We are in a heavily data-oriented world, so a large component of this course will be understanding choices made to obtain data (e.g., research designs) and how to process and analyze data. This course will be highly interactive and immersive so students can learn by doing and develop skills for the graduate school and/or the workforce.

Data analysis is quickly changing the way we understand and engage in politics, how we implement policy, and how organizations across the world make decisions. In this course, we will learn the fundamental principles of statistical inference and develop the necessary programming skills to answer a wide range of political and policy oriented questions with data analysis. Who is most likely to win the upcoming presidential election? Do countries become less democratic when leaders are assassinated? Is there racial discrimination in the labor market? These are just a few of the questions we will work on in the course.

Students are not expected to have any prior programming knowledge or experience. The course will be centered around bite-size assignments that will help build coding and statistical skills from scratch. Students will leave the course equipped for work in any setting that requires a social scientific approach to data analysis, from policy non-profits to government, from Silicon Valley to Wall Street and beyond.

Textbook: Llaudet, Elena and Kosuke Imai. *Data Analysis for Social Science: A Friendly and Practical Introduction*. United States, Princeton University Press, 2022.

Trochim, William and James P. Donnelly *Research Methods: The Essential Knowledge Base, 3rd Edition*. United States, Cengage Learning, 2006.

Statistical Softwares: R (www.r-project.org) and RStudio (www.rstudio.com/). Our class is located in a computer lab so you have the choice of using lab computers or your own personal devices.

Evaluation: The final grade will be based on the following:

1. *Participation (10%)*. Based on (i) participation in daily pollev quizzes, (ii) participation in class discussions, and (iii) completion of in-class work.
2. *Research Activities (10%)*. Based on (i) article analyses, (ii) colloquium series attendance (or arranged alternative), and (iii) SONA participation.
3. *Problem Sets (30%)*. Take-home problem sets will be given throughout the course. There are 9 required problem sets in total, plus 1 optional problem set for extra credit.
 - *Collaboration policy*: While working on problem sets, students are encouraged to consult with classmates and the instructor.
 - *Due date*: Problem sets are due *before* the class in which they are discussed. (See tentative schedule below.)
 - *Late policy*: Late submissions will not be accepted unless special permission is granted *in advance*. If something comes up, please send me an email petitioning an extension. No more than two extensions will be granted per student.
4. A *Midterm (20%)* and a *Final Exam (30%)*.
 - No collaboration is permitted during the exams.

Course Policies: This course adheres to policies and procedures that apply to all UTEP courses with regard to accommodations and academic misconduct. There is a zero-tolerance policy for plagiarism and cheating. For any accommodations related to disability or religious observation, with or without documentation, please speak with the instructor.

In addition:

- class attendance is expected; please discuss any potential absences with the instructor *ahead of time*
- students are expected to come to class on time and ready to participate (having done the readings and completed the problem sets listed for that day)
- make-up exams will only be offered in justified, special circumstances.

Student Resources: The university provides a range of academic, counseling, medical and administrative student resources and support services. To learn more, please review the resources document on Blackboard.

Course Goals and Learning Objectives:

Goals	Objectives	Assessments
1. Students will know how to recognize and interpret quantitative information	(a) Students will be able to read and understand quantitative data in various formats (b) Students will be able to communicate the meaning of quantitative data and the results of data analysis	Participation Problem sets Exams
2. Students will understand the theoretical basis of quantitative reasoning	(a) Students will be able to explain the basic concepts of quantitative reasoning, such as variables, constants, and estimates (b) Students will be able to understand how inferences are drawn from quantitative analysis (c) Students will be able to recognize the limitations of quantitative methods	Participation Problem sets Exams
3. Students will understand the practical application of quantitative data analysis	(a) Students will be able to determine and use appropriate quantitative methods to solve problems (b) Students will be able to accurately interpret the results of data analyses (c) Students will be able to assess results for reasonableness	Participation Problem sets Exams

Tentative Course Schedule: The tentative course schedule is detailed in the table below.

In the event of a class cancelation, students are expected to continue with readings as originally scheduled. Any assignments scheduled are due at the next class unless other instructions are posted on the course website or otherwise communicated.

Disclaimer: The schedule, policies, procedures, and assignments in this course are subject to change in the event of extenuating circumstances, by mutual agreement, and/or to ensure better student learning.

Tentative Course Schedule (Subject to Change as Semester Progresses):

Day	#	Topic	Readings	Key Concepts and R Code	PSet
T 08/29	1	Course Introduction			
R 08/31	2	Validity	T&D 2.1, 3.1, 7.1	external validity, internal validity, construct validity	
T 09/05	3	Research 101	T&D Ch 1	relationships, variables, hypotheses, research process ethics	
R 09/07	4	Surveys	T&D 4.1, 4.3	response type, question format, survey design ethics	
T 09/12	5	Guest Presentations			
R 09/14	6	Guest Presentations			
T 09/19	7	Introduction to R and RStudio	1-1.6	R: <code>+</code> , <code>-</code> , <code>*</code> , <code>/</code> , <code><-</code> , <code>"</code> , <code>()</code> , <code>sqrt()</code> , <code>#</code>	#0
R 09/21	8	Observations and Variables	1.7	dataframes, observations, variables, unit of observation, <i>i</i> , character vs. numeric variables, binary vs. non-binary vari- ables, <i>n</i> ; R: <code>setwd()</code> , <code>read.csv()</code> , <code>View()</code> , <code>head()</code> , <code>dim()</code>	
T 09/26	9	Computing and Interpreting Means	1.8-1.10	mean or average, \sum , unit of measurement; R: <code>\$</code> , <code>mean()</code>	#1
R 09/28	10	Estimating Causal Effects with Randomized Experiments	2-2.4 T&D 9.1-9.3	causal relationships, treatment (X) vs. outcome variables (Y), potential outcomes, factual vs. counterfactual outcomes, fun- damental problem of causal inference, individual vs. average causal effects, randomized experiments, treatment and control groups, pre-treatment characteristics, the difference-in-means estimator	
T 10/03	11	Does Social Pressure Increase the Probability of Turning Out to Vote?	2.5-2.7	R: <code>==</code> , <code>ifelse()</code> , <code>[]</code>	#2
R 10/05	12	Survey Research and Exploring One Variable at a Time	3-3.4 (skip: 3.2.2, 3.4.1, 3.4.2, 3.4.3, 3.4.5)	sample, representative sample, table of frequencies, table of proportions, histogram, descriptive statistics (mean, median, standard deviation, and variance); R: <code>table()</code> , <code>prop.table()</code> , <code>hist()</code> , <code>median()</code> , <code>sd()</code> , <code>var()</code> , <code>^</code>	

Day	#	Topic	Readings	Key Concepts and R Code	PSet
T 10/10	13	Flex Day			#3
R 10/12	14	Exploring the Relationship Between Two Variables	3.5-3.7	scatter plot, correlation; R: <code>plot()</code> , <code>cor()</code>	
T 10/17	15	Predicting Non-Binary Outcomes Using Linear Regression	4-4.4.1	prediction and correlation, predicted (\hat{Y}) vs. actual outcome (Y), prediction errors ($\hat{\epsilon}$), the least squares method, the linear regression model, $\hat{Y} = \hat{\alpha} + \hat{\beta}X$, interpretation of coefficients, intercept ($\hat{\alpha}$) and slope ($\hat{\beta}$), $\Delta \hat{Y} = \hat{\beta}\Delta X$; R: <code>lm(Y ~ X)</code> , <code>abline()</code>	#4
R 10/19	16	Flex Day			#5
T 10/24	17	Predicting Binary Outcomes Using Linear Regression	4.6-4.9 (skip 4.8)	R^2 , relationship between R^2 and correlation	
R 10/26	18	Estimating Causal Effects with Observational Data and the Problem of Confounders	5-5.3.1	observational studies vs. randomized experiments, confounders (Z), interpretation of $\hat{\alpha}$ and $\hat{\beta}$ when X is binary and identifies treatment assignment	
T 10/31	19	Controlling for Confounders Using Multiple Linear Regression	5.3.2-5.4.2	multiple vs. simple linear regression models, new interpretation of coefficients	#6
R 11/02	20	Internal vs. External Validity	5.5-5.7	internal validity, external validity	
T 11/07	21	Office Hours in Classroom			#7
R 11/09	22	MIDTERM EXAM			
T 11/14	23	Flex Day			
R 11/16	24	Probability	6-6.8 (skim) (skip 6.7 and ignore all code)	probability, random variables, probability distributions, Bernoulli vs. normal distribution, the standard normal distribution, population parameters vs. sample statistics, the law of large numbers, the central limit theorem	#8
T 11/21	25	<i>Holiday</i>			
R 11/23	26	<i>Holiday</i>			

Day	#	Topic	Readings	Key Concepts and R Code	PSet
T 11/28		Hypothesis Testing with Coefficients	7-7.1 (skim), 7.3-7.6 (skip 7.3.1)	hypothesis testing, test statistic, standard error of $\hat{\beta}$, R: <code>summary()\$coef</code>	
R 11/30	28	Do Small Classes Increase Probability of Graduating from High School?	7.7 (PDF on course website)		
T 12/05	29	Office Hours in Classroom			#9
R 12/07	30	Group study day			
R 12/14		FINAL EXAM at 10am			extra credit