

Datasci 203













Statistics for Data Science





[Return to All Courses](https://learn.berkeley.edu/courses/395/pages/welcome-to-mids-alumni-course-library)
(<https://learn.berkeley.edu/courses/395/pages/welcome-to-mids-alumni-course-library>)

This course aims to equip students with a foundational understanding of classical statistics within the broader framework of data science. It emphasizes the correct application of common statistical procedures, ensuring students can verify assumptions and address violations appropriately. The curriculum covers key aspects of effective analysis, including formulating research questions, operationalizing variables, exploring data, selecting appropriate hypothesis tests, and effectively communicating results. Beginning with an introduction to probability theory, students will engage in problem sets to build intuition for statistical models. The course progresses to the use of estimators for model parameters, explores the logic of hypothesis testing, and surveys tests for group comparisons. A significant portion of the course is dedicated to classical linear regression, highlighting its flexibility for both descriptive and causal inference. Practical analysis of real-world data using the open-source language R is integrated throughout the course.

[Click here for a sample syllabus.](https://learn.berkeley.edu/courses/395/files/16569?wrap=1) (<https://learn.berkeley.edu/courses/395/files/16569?wrap=1>). [↓](https://learn.berkeley.edu/courses/395/files/16569/download?download_frd=1)
(https://learn.berkeley.edu/courses/395/files/16569/download?download_frd=1)

Module	Videos
1 - Introduction to Probability	 (https://learn.berkeley.edu/courses/395/pages/datasci203-dot-1)

Module	Videos
2 - Defining Random Variables	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-2
3 - Summarizing Distributions	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-3
4 - Conditional Expectations and BLP	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-4
5 - Learning from Random Samples	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-5
6 - Hypothesis Tests	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-6
7 - Comparing Two Groups	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-7
8 - Ordinary Least Squares Regression	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-8
9 - Ordinary Least Squares Inference	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-9
10 - Descriptive Model Building	 https://learn.berkeley.edu/courses/395/pages/datasci203-dot-10

Module	Videos
11 - Explanatory Model Building	 (https://learn.berkeley.edu/courses/395/pages/datasci203-dot-11)
12 - The Classical Linear Model	 (https://learn.berkeley.edu/courses/395/pages/datasci203-dot-12)
13 - Reproducible Research	 (https://learn.berkeley.edu/courses/395/pages/datasci203-dot-13)
14 - Maximum Likelihood Estimation	 (https://learn.berkeley.edu/courses/395/pages/datasci203-dot-14)