Understanding and Visualizing Data with Python

by University of Michigan

**About this Course**

In this course, learners will be introduced to the field of statistics, including where data come from, study design, data management, and exploring and visualizing data. Learners will identify different types of data, and learn how to visualize, analyze, and interpret summaries for both univariate and multivariate data. Learners will also be introduced to the differences between probability and non-probability sampling from larger populations, the idea of how sample estimates vary, and how inferences can be made about larger populations based on probability sampling. At the end of each week, learners will apply the statistical concepts they’ve learned using Python within the course environment. During these lab-based sessions, learners will discover the different uses of Python as a tool, including the Numpy, Pandas, Statsmodels, Matplotlib, and Seaborn libraries. Tutorial videos are provided to walk learners through the creation of visualizations and data management, all within Python. This course utilizes the Jupyter Notebook environment within Coursera.

 Show less

* 

**Taught by:**[**Brenda Gunderson**](https://www.coursera.org/instructor/bkg), Lecturer IV and Research Fellow

Department of Statistics

* 

**Taught by:**[**Brady T. West**](https://www.coursera.org/instructor/~34237657), Research Associate Professor

Institute for Social Research

* 

**Taught by:**[**Kerby Shedden**](https://www.coursera.org/instructor/~34919608), Professor

Department of Statistics

|  |  |
| --- | --- |
| **Basic Info** | Course 1 of 3 in the [Statistics with Python Specialization](https://www.coursera.org/specializations/statistics-with-python) |
| **Level** | Beginner |
| **Commitment** | 4 weeks of study, 4-6 hours/week |
| **Language** | English, **Subtitles:**Arabic, French, Portuguese (European), Italian, Vietnamese, Korean, German, Russian, Spanish  [**Volunteer to translate subtitles for this course**](https://www.coursera.org/learn/understanding-visualization-data/home/info) |
| **Hardware Req** | None required - this course utilizes the Jupyter Notebook environment within Coursera. |
| **How To Pass** | Pass all graded assignments to complete the course. |
| **User Ratings** | Average User Rating 4.7 |

Syllabus

WEEK 1 - INTRODUCTION TO DATA

In the first week of the course, we will review a course outline and discover the various concepts and objectives to be mastered in the weeks to come. You will get an introduction to the field of statistics and explore a variety of perspectives the field has to offer. We will identify numerous types of data that exist and observe where they can be found in everyday life. You will delve into basic Python functionality, along with an introduction to Jupyter Notebook. All of the course information on grading, prerequisites, and expectations are on the course syllabus and you can find more information on our Course Resources page.

More

11 videos, 7 readings, 1 practice quiz

Expand week 1 material

**Graded:**Assessment: Different Data Types

WEEK 2 - UNIVARIATE DATA

In the second week of this course, we will be looking at graphical and numerical interpretations for one variable (univariate data). In particular, we will be creating and analyzing histograms, box plots, and numerical summaries of our data in order to give a basis of analysis for quantitative data and bar charts and pie charts for categorical data. A few key interpretations will be made about our numerical summaries such as mean, IQR, and standard deviation. An assessment is included at the end of the week concerning numerical summaries and interpretations of these summaries.

More

8 videos, 2 readings, 1 practice quiz

Expand week 2 material

**Graded:**Assessment: Numerical Summaries

**Graded:**Python Assessment: Univariate Analysis

WEEK 3 - MULTIVARIATE DATA

In the third week of this course on looking at data, we’ll introduce key ideas for examining research questions that require looking at more than one variable. In particular, we will consider both numerically and visually how different variables interact, how summaries can appear deceiving if you don’t properly account for interactions, and differences between quantitative and categorical variables. This week’s assignment will consist of a writing assignment along with reviewing those of your peers.

7 videos, 3 readings, 1 practice quiz

expandweek 3 material

**Graded:**Pizza Study Design Assignment

**Graded:**Python Assessment: Multivariate Analysis

WEEK 4 - POPULATIONS AND SAMPLES

In this week, you’ll spend more time thinking about where data come from. The highest-quality statistical analyses of data will always incorporate information about the process used to generate the data, or features of the data collection design. You’ll be exposed to important concepts related to sampling from larger populations, including probability and non-probability sampling, and how we can make inferences about larger populations based on well-designed samples. You’ll also learn about the concept of a sampling distribution, and how estimation of the variance of that distribution plays a critical role in making statements about populations. Finally, you’ll learn about the importance of reading the documentation for a given data set; a key step in looking at data is also looking at the available documentation for that data set, which describes how the data were generated.

More

15 videos, 8 readings

Expand week 4 material

**Graded:**Assessment: Distinguishing Between Probability & Non-Probability Samples

**Graded:**Generating Random Data and Samples

How It Works

**General**

**How do I pass the course?**

To earn your Course Certificate, you’ll need to earn a passing grade on each of the required assignments—these can be quizzes, peer-graded assignments, or programming assignments. Videos, readings, and practice exercises are there to help you prepare for the graded assignments.

**What do start dates and end dates mean?**

Once you enroll, you’ll have access to all videos, readings, quizzes, and programming assignments (if applicable). If you choose to explore the course without purchasing, you may not be able to access certain assignments. If you don’t finish all graded assignments before the end of the course, you can reset your deadlines. Your progress will be saved and you’ll be able to pick up where you left off.

**What are due dates? Is there a penalty for submitting my work after a due date?**

Within a course, there are suggested due dates to help you manage your schedule and keep coursework from piling up. Quizzes and programming assignments can be submitted late without consequence. However, it is possible that you won't receive a grade if you submit your peer-graded assignment too late because classmates usually review assignment within three days of the assignment deadline.

**Can I re-attempt an assignment?**

Yes. If you want to improve your grade, you can always try again. If you’re re-attempting a peer-graded assignment, re-submit your work as soon as you can to make sure there’s enough time for your classmates to review your work. In some cases you may need to wait before re-submitting a programming assignment or quiz. We encourage you to review course material during this delay.

 Show less

**Peer-graded assignments**

Peer-graded assignments require you and your classmates to grade each other’s work.

**How do peer graded assignments work?**

After you submit your assignment, you will review some of your peers’ assignments. The number of assignments you must review is set by the instructor of the course.

**I reviewed my peers’ assignments! What happens next?**

While you’re reviewing your peers’ assignments, they’ll review yours. If you submit your assignment on time, you’ll get your grade within a week, as long as at least one peer reviews your assignment. If you submit late, you’ll need all of the peer reviews the instructor requires. [Learn more about Peer Graded Assignments.](https://learner.coursera.help/hc/articles/209818803)

**How are grades calculated?**

You and your classmates will be asked to provide a score for each part of the assignment. Final grades are calculated by combining the median scores you received for each section.

**What kind of feedback should I give?**

Use the instructor’s criteria in the rubric to grade honestly and fairly. If your peers’ answers are excellent, score them highly and tell them what they did well. If their answers aren’t as good, give the score they deserve, and be sure to provide [respectful, useful feedback](https://coursera.community/study-tips-6/giving-feedback-116) so they can do better next time they attempt the assignment.

**Is there a penalty for submitting my work late?**

No, but it’s important to submit your work as close to the due date as you can. Classmates grade most of the assignments within three days of the due date. If you submit yours too late, there may not be anyone to review your work.

**If I fail an assignment, can I try again?**

Yes! You can always try again, but you’ll need to resubmit your work as soon as possible to make sure your classmates have enough time to grade your work.

**Can I edit my assignment?**

Yes, but you’ll need to re-submit your work and any grade you’ve already received will be deleted.