

# Introduction to Data Visualization, Fall 2017

This is a course in finding and telling visual stories from data. We will cover fundamental principles of data analysis and visual presentation, chart types and when to use them, and how to acquire, process and “interview” data. We will make interactive and static charts and maps using free software. We will be making graphics by writing code, as well as using point-and-click software. The emphasis is on gaining practical skills that students can apply in a newsroom setting.

## Weekly schedule

We will meet in **108/Lower NG** on Fridays from 9.30am - 12.30pm. Your instructor, [Peter Aldhous](#), will maintain office hours in **B1** from 1.30pm - 5.00pm, following each class. You are encouraged to arrange appointments to discuss your work.

Class time will also be scheduled for each of you to critique and lead class discussion of a recently published news graphic/interactive.

- **Aug. 25: What is data?**

Categorical and continuous variables; basic operations for interviewing a dataset; sampling and margins of error; plotting and summarizing distributions; choosing bins for your data; basic newsroom math; correlation and its pitfalls; exploring differences between groups; scatter plots.

- **Assignment:** Calculate summary statistics; subscribe to data viz blogs etc to follow latest developments in visualization.

- **Sept. 1: Data visualization: basic principles**

Encoding data using visual cues; choosing chart types to show comparisons, composition (parts of the whole) and connections; using color effectively; using chart furniture, minimizing chart junk and highlighting the story; avoiding pitfalls; good practice, including for interactive graphics.

- **Assignment:** Quiz on good practice in visualization and data analysis.

- **Sept. 8: Interviewing data: exploratory graphical analysis**

We will use [Tableau Public](#) to explore and visualize World Bank data on neonatal deaths across the globe, creating an interactive online dashboard.

- **Assignment:** Make another dashboard from the same data.

- **Sept. 15: Using GitHub, and starting your final projects**

We will explore how to use [GitHub](#) for version control of a project, before you pitch ideas for your final projects.

- **Assignment:** Submit a pitch for your final project; also make an appointment to discuss this with your instructor the following week.

## • **Sept. 22: Acquiring, cleaning and formatting data**

Data search and download tricks, including [Table2Clipboard](#) and [DownThemAll!](#) Firefox plugins; manipulating urls and using APIs to acquire data; scraping data from the web with [OutWit Hub](#); cleaning data with [Open Refine](#); converting data between different formats using [Mr. Data Converter](#).

- **Assignment:** Data cleaning task.

## • **Sep. 29: Let's apply what we've learned so far**

You will obtain and process World Bank data on life expectancy and GDP per capita for the world's nations from the web, and then create a Tableau visualization inspired by [this Gapminder video](#).

- **Assignment:** Download and process World Bank data on nations' carbon dioxide emissions per capita, and make a simple Tableau dashboard.

## • **Oct. 6: Manipulating data with R**

Introduction to [R](#), [R Studio](#) and the [dplyr](#) package for data manipulation.

## • **Oct. 13: Making static graphics with R**

Making charts with the [ggplot2](#) R package.

- **Assignment:** Process data and make charts using R.

## • **Oct. 20: Principles of mapping**

Basic mapping principles: projections, geocoding, geodata formats; approaches to putting data onto maps, including choropleth maps, scaled symbols, hexagonal binning and cartograms.

- **Assignment:** Continue work on final project, delivering a progress report.

## • **Oct. 27: Making static maps and processing geodata with GIS software**

We will use [QGIS](#) to make a multi-layered map. We will also learn how to use QGIS and its plugins to process geodata.

- **Assignment:** continue work on final project, delivering a progress report.

## • **Nov.3: Making interactive maps**

We will use [CARTO](#) to create interactive online maps.

- **Assignment:** Submit a full project update, via your GitHub account, for review in next week's meetings.

- **Nov. 10: No Class**

Instead, one-on-one meetings will be arranged with instructor to discuss progress with your final projects.

- **Nov. 17: From R to interactive charts and maps**

We will explore R packages called [htmlwidgets](#) that allow you to create JavaScript visualizations straight from your data with a minimum of coding.

- **Assignment:** Continue work on final project, delivering a progress report.

- **Nov 24: Thanksgiving — No Class**

- **Assignment:** Continue work on your final project.

- **Dec. 1: Iteration and animation: loops, GIFs, and videos**

We will explore how to make animation frames by writing loops in R, and use [ImageMagick](#), [FFmpeg](#) and the [gganimate](#) package to make animated GIFs and videos.

- **Assignment:** complete your final project, file via your GitHub account.

- **Dec. 8: Student presentations on final projects**

You will each present your final project work to the class, explain the process of creating the graphics, choices in analysis and design, and problems that arose and how these were addressed.

## Recommended reading

Alberto Cairo: *The Functional Art: An Introduction to Information Graphics and Visualization*

Nathan Yau: *Data Points: Visualization That Means Something*

Further reading/viewing will be recommended to support weekly class material.

## Attendance

Unexcused absence from two classes will drop you one letter grade; a third unexcused absence will result in an F. Excused absences will be permitted only in extraordinary circumstances. Regardless of the reason for an absence, students will be responsible for any assignments due and for learning material covered in class.

## Grading

Class participation, weekly assignments: **45%**

Final project: **45%**

Attendance: **10%**

## Good manners

Students must turn off the ringers on their cell phones before class begins. Students may not check e-mail, social media sites or other websites during lecture portions of class or while working on class exercises.

## Academic dishonesty and plagiarism

The high academic standard at the University of California, Berkeley, is reflected in each degree that is awarded. As a result, it is up to every student to maintain this high standard by ensuring that all academic work reflects his/her own ideas or properly attributes the ideas to the original sources.

These are some basic expectations of students with regards to academic integrity:

- Any work submitted should be your own individual thoughts, and should not have been submitted for credit in another course unless you have prior written permission to re-use it in this course from this instructor.
- All assignments must use “proper attribution,” meaning that you have identified the original source of words or ideas that you reproduce or use in your assignment. This includes drafts and homework assignments!
- If you are unclear about expectations, ask your instructor.

## Disability accommodations

If you need disability-related accommodations in this class, if you have emergency medical information you wish to share with the instructor, or if you need special arrangements in case the building must be evacuated, please inform the instructor as soon as possible by seeing him after class or making an appointment to visit during office hours. If you are not currently listed with DSP (Disabled Students' Program) but believe that you could benefit from their support, you may [apply online](#).