

Univeristy of Chicago  
Data Science  
DATA 227  
Autumn 2021

Course Title: Data Science 227  
Length of Course: 10 weeks

Instructor Name: Will Trimble  
Office Location: Crerar  
E-mail: [wtrimbl@uchicago.edu](mailto:wtrimbl@uchicago.edu)  
Class Time: Tu/Thursday 12:30-1:50  
Classroom: Stuart 104  
Course Website: not yet  
Enrollment: capped at 40

**Course Catalog Description:**

This course introduces best practices for presenting and communicating quantitative data. Principles of data visualization include the use of colors and negative spaces, drawing attention to important details, repetition of design motifs, appropriately using figures and tables, and combining different scales in a single figure. The course also discusses how to avoid common distortions resulting in misleading plots and figures and how to effectively communicate findings. Examples are chosen from a variety of fields, such as the biological sciences, the social sciences, and the media.

**Course Objectives:**

This class is intended to show the student some of the ways to make visualizations out of data, and some of the things that are known about how to do this well and badly. I will be giving example code using python, pandas, matplotlib and altair/vega-lite, but the use of specific tools is not required.

## TOPICS

Date	Date	Class topics	Workbooks
Week 1	Sep 28	Value of Visualization	pandas
	Sep 30	Tooling, data normalization	intro vega-lite, altair
Week 2	Oct 05	Data Types, Encoding Channels	data types
	Oct 7	Visualization Tools	
Week 3	Oct 12	Chart types, transformations	data transformation
	Oct 14	Case studies in Exploratory Data Analysis	
Week 4	Oct 19	Visual Encoding and Dark Patterns	scales, axes, and legends
	Oct 21	Ethics in Visualization	
Week 5	Oct 26	Color	
	Oct 28		project discussion
Week 6	Nov 02	Graphical Perception	
	Nov 04	Design Tips	
Week 7	Nov 09	Interaction	Interaction
	Nov 11	Animation	
Week 8	Nov 16	Cartography	Cartographic visualization
	Nov 18	Narrative	
Thanksgiving Break			
Week 9	Nov 30	Uncertainty	
	Dec 02	Networks	
Reading period			
Week 10	Dec 07		final projects

## **Texts, Materials, and Resources:**

The class does not have a fixed textbook, but a series of readings. Several readings are from Edward Tufte's *The Visual Display of Quantitative Information*, which should be on reserve at Crerar.

The notebooks at <https://github.com/uwdata/visualization-curriculum> provide a good tutorial on data viz using altair.

### **Week 1:**

Hadley Wickham. Tidy Data. *Journal of Statistical Software* 2014

The Power of Representation, Chapter 3 in *Things That Make Us Smart*. Don Norman.

Chapter 1: Graphical Excellence, In *The Visual Display of Quantitative Information*. Tufte.

### **Week 2:**

A Tour through the Visualization Zoo. Jeffrey Heer, Michael Bostock, and Vadim Ogievetsky. *ACM Queue*, 8(5). 2010.

Chapter 1: Information Visualization, in *Readings in Information Visualization*. Stuart Card, Jock Mackinlay, and Ben Shneiderman. 1999.

Enterprise Data Analysis and Visualization: An Interview Study. Sean Kandel, Andreas Paepcke, Joseph M. Hellerstein, and Jeffrey Heer. *IEEE InfoVis*. 2012.

Chapter 5: Chartjunk, In *The Visual Display of Quantitative Information*. Tufte.

### **Week 3:**

Exploring and Finding Information. Peter Pirolli. *HCI Models, Theories, and Frameworks: Toward a Multidisciplinary Science*. 2003.

D3: Data-Driven Documents. Michael Bostock, Vadim Ogievetsky & Jeffrey Heer. *InfoVis* 2011.

Postmortem of an Example. Jacques Bertin. *Graphics and Graphic Information-Processing*. 1981.

Chapter 3: Sources of Graphical Integrity, In *The Visual Display of Quantitative Information*. Tufte.

### **Week 4:**

Polaris: A System for Query, Analysis and Visualization of Multi-dimensional Relational Databases. Stolte, Tang, and Hanrahan. *IEEE TVCG* 2002.

A Nested Model for Visualization Design and Validation. Tamara Munzner. *IEEE InfoVis* 2009.

Chapter 2: Graphical Integrity, In *The Visual Display of Quantitative Information*. Tufte.

### **Week 5::**

Which color scale to use when visualizing data. Lisa Charlotte Muth

<https://blog.datawrapperr.de/which-color-scale-to-use-in-data-vis/>

Somewhere Over the Rainbow: An Empirical Assessment of Quantitative Colormaps. Yang Liu, Jeffrey Heer. ACM CHI 2018.

Week 6:

Ethical Dimensions of Visualization Research. Michael Correll. ACM CHI 2019.

How Deceptive are Deceptive Visualizations?: An Empirical Analysis of Common Distortion Techniques. Anshul Vikram Pandey, Katharina Rall, Margaret Satterthwaite, Oded Nov, Enrico Bertini. ACM CHI 2015.

Week 7:

The Work that Visualization Conventions Do. Helen Kennedy, Rosemary Lucy Hill, Giorgia Aiello & William Allen. Information, Communication & Society. 2016.

Dual-Scaled Axes in Graphs: Are They Ever the Best Solution? (Few, 2008)

The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, Shneiderman, Proc. IEEE Conference on Visual Languages, Boulder 1996.

Interactive Dynamics for Visual Analysis. Jeffrey Heer & Ben Shneiderman. 2012.

Effectiveness of Animation in Trend Visualization. George Robertson, Roland Fernandez, Danyel Fisher, Bongshin Lee, & John Stasko. InfoVis 2008.

Perception in Visualization. Christopher Healey. <https://www.csc2.ncsu.edu/faculty/healey/PP/39> Studies About Human Perception in 30 Minutes. Kennedy Elliott.

<https://medium.com/@kennelliott/39-studies-about-human-perception-in-30-minutes-4728f9e31a73>

Week 8:

Chapter 11: The Cartogram: Value-by-Area Mapping, in Cartography: Thematic Map Design. Dent.

Narrative Visualization: Telling Stories with Data. Edward Segel & Jeffrey Heer. InfoVis 2010.

Reinventing Explanation. Michael Nielsen, 2014.

[https://michaelnielsen.org/reinventing\\_explanation/](https://michaelnielsen.org/reinventing_explanation/)

Week 9:

The Visual Uncertainty Experience. Jessica Hullman. OpenVis Conf 2016.

<https://www.youtube.com/watch?v=pTVAn4oLvbc>

Error Bars Considered Harmful: Exploring Alternate Encodings for Mean and Error. Michael Correll, Michael Gleicher. IEEE InfoVis 2014.

Why Authors Don't Visualize Uncertainty. Jessica Hullman. IEEE InfoVis. 2019.

Squarified Treemaps. Mark Bruls, Kees Huizing & Jarke van Wijk. Eurographics Data Visualization 2000.

Hierarchical Edge Bundles: Visualization of Adjacency Relations in Hierarchical Data. Danny Holten. InfoVis 2006.

### Grade Distribution/Breakdown:

<u>Assessment</u>	<u>Number</u>	<u>Total points</u>	<u>% of the FINAL grade</u>
Assignments	5	50	50
Projects	1	20	20
Final project	1	30	30

### Rubrick:

Component	Excellent	Satisfactory	Poor
Data Question	An interesting question that the visualization clearly answers.	Good question, but visualization does not clearly answer it.	Unclear question posed of the data.
Encoding	Effective design choices. Visualization is easy to read and understand.	Good design choices, but minor flaws hinder comprehension.	Ineffective mark, encoding, or data transformation choices are distracting or potentially misleading.
Titles & Labels	Titles and labels helpfully describe and contextualize the visualization.	Most necessary titles and labels are present, but they could provide more context.	Many titles or labels are missing, or do not provide human-understandable information.
Design Rationale	Well crafted write-up provides reasoned justification for all design choices.	Most design decisions are described, but rationale could be explained at a greater level of detail.	Missing or incomplete. Several design choices are left unexplained.
Creativity & Originality	Exceeds the parameters of the assignment, with original insights or a particularly engaging design.	Meets all the parameters of the assignment.	Meets most of the parameters of the assignment.

### **Academic Honesty**

In this course, as in all your courses, you must adhere to the University-wide Academic Honesty policies of the Student Manual. <http://college.uchicago.edu/advising/academic-integrity-student-conduct> These are also described by the College under Academic Integrity & Student Conduct; expand the "Academic Integrity" section at the page bottom. To paraphrase:

1. Never copy work from any other source and submit it as your own.
2. Never allow your work to be copied.
3. Never submit work identical to another student's, excepting group projects
4. Document all collaboration
5. Cite your sources.

There are harsh consequences for breaking these rules. **Any student who is determined to have participated in academic dishonesty will not be allowed to withdraw and will receive a course grade no higher than a C.** You will also be reported to your adviser and may face further discipline as a result.

Some instructors forbid sharing your work publicly (such as posting it to the web) under the the second rule; for this class most of the work will be visualizations and the code and data that generate them, and are more creative works than technical exercises. You can share your work publicly if you wish. With respect to the third rule, you may discuss the general idea of how to solve a particular problem with another student (in which case, you must document it per the fourth rule), but you may not share your work directly. When it comes time to sit down and start writing or typing, you must do the work by yourself (or with your partner for that project). **Discussion of class work must be entirely voluntary and never transactional.** If you have any questions or concerns about this policy, or about the behavior of another student with respect to it, please ask your instructor as soon as possible. This statement of Academic Honesty is based on that of Adam Shaw.

**Advice** Writing code that does what it is supposed to do can be enlivening, joyful, even uplifting. By contrast, fighting for hours with broken code is the opposite: discouraging, frustrating, and generally miserable. We would like you to experience more of the former and less of the latter. Work methodically. Start your work well ahead of time. Beyond a certain point, it is not profitable to be stumped. If you have made no progress in some nontrivial chunk of time, say, one hour, it is time to stop and change your approach. Use one of our many support mechanisms to get some assistance. We will help you get going again when you are stuck.

### **Sexual Misconduct**

Our school is committed to fostering a safe, productive learning environment. Title IX and our school policy prohibits discrimination on the basis of sex. Sexual misconduct – including assault, harassment, stalking, and domestic and dating violence — is also prohibited. Harassment can take the form of, for example, any repeated unwelcome comments of a sexual nature, or any sexual advance associated with seeking help on class work. Review the Policy on Harassment, Discrimination, and Sexual Misconduct, in particular the Sexual Misconduct and Definitions, so you understand what this covers.

Our school encourages anyone experiencing sexual misconduct to talk to someone about what happened, so they can get the support they need and our school can respond appropriately.

If you wish to speak confidentially about an incident of sexual misconduct, want more information about filing a report, or have questions about policies, procedures, or support services, please contact our Title IX Coordinator; see the Reporting Options section of the policy. Our school is legally obligated to investigate reports of sexual misconduct after a formal complaint is filed or signed by the Title IX Coordinator, but a request for confidentiality will be respected to the extent possible. As a faculty member, I am required to report any harassment that I learn about to the Title IX Coordinator.

### **UChicago Health Pact**

All students on campus are required to adhere to the guidelines in the UChicago Health Pact in order to promote a safe environment in the classroom.

- Secure face coverings must be worn appropriately at all times at all times while in University buildings
- Maintain a distance of 6 feet from others

- Do not attend and in-person class if you feel unwell or are experiencing COVID-19 related symptoms

The complete text of the UChicago Health Pact along with additional information about COVID-19 protocols can be found at <https://goforward.uchicago.edu/health-requirements/>

Any concerns over inappropriate PPE usage, physical distancing, cleaning/disinfection, or other COVID-19 related public health concerns should be directed to UCAIR. If there is an emergency, call 773-702-8181 or dial 123 on any campus phone.

If you were potentially exposed to COVID-19 or your COVID-19 test results come back positive, reach out immediately to [C19HealthReport@uchicago.edu](mailto:C19HealthReport@uchicago.edu).