INFO 590 (Fall 2022)

Data Visualization

Sadamori Kojaku

skojaku@iu.edu

Office hours: Monday 2pm-3pm at https://iu.zoom.us/my/skojaku
(You can send a message anytime on Slack)

Assistant Instructors and office hours
See Canvas

COURSE DESCRIPTION

From news to cutting-edge scientific papers, from a home office to the largest companies in the world, data visualization is used as a critical step in understanding data. Because data visualization is indispensable in data analysis, data visualization has become an essential skill for every knowledge worker. This course is an introduction to basic statistical data analysis and visualization. We will learn the fundamentals of data visualization in the context of perception, integrity, design, statistics, types of data, and visualization techniques. The hands-on exercises using the Python stack aim to equip you with practical data processing and visualization skills.

Relationships with E583/Z637 Information Visualization (IVMOOC)   Compared with E583/Z637, this course is geared more towards producing fundamental statistical visualizations and exploratory data analysis by writing code using the Python data science and visualization stack. Therefore, this course may be more suitable for students who pursue their careers in research, development, engineering, and data analysis, and those who will directly process and analyze complex datasets.

COURSE OBJECTIVES

By the end of the course, you are expected to be capable of understanding, explaining, and manipulating basic types of data, analyzing them by applying exploratory visualization techniques, and creating explanatory visualizations. You are also expected to be able to evaluate and improve the effectiveness of data visualizations based on the fundamental visualization principles of human perception, design, data types, and visualization techniques. You will demonstrate your capacity by completing a course project that documents the detailed process of creating data visualizations.

COMMUNICATION

We will use Canvas and Slack for communication. **Canvas** is for official communications as well as for anything that contains personal and sensitive information. **Slack** is for day-to-day information sharing, Q&As, team discussions, and other casusal conversations.

Announcements, Q&As, and other communication will be sent via Canvas and Slack. Although the most critical announcements will be sent via both platforms, a lot of course-related information (as well as questions and answers) will be shared on Slack and thus you will miss most of useful—although not *essential* in completing the course—course-related information if you are not on Slack. When joining the course Slack, feel free to avoid using your full name (e.g., use "John D." instead of "John Doe") to protect your privacy. Also never post your personal information or sensitive data (e.g., grades) to the course Slack.

The address of the course slack is: `https://iu-dviz-course.slack.com`, and visit `https://join.slack.com/t/iu-dviz-course/signup` to signup. You can create an account by using one of the following IU email addresses: `indiana.edu`, `umail.iu.edu`, `iu.edu`, and `iupui.edu`. Please email the instructor if you want to use other email addresses.

Note that Email and Canvas will be much slower than Slack because the instructors are under a constant bombardment of emails about all kinds of things that you don't even want to know. Please expect *a few days* (up to a week) of waiting time for the response to your email. Please expect a few hours or a day for the response to your Slack message.

Whenever you have something to say about the course or have a suggestion for improving the course, please share your thoughts! We will be extremely grateful if you point out issues in the quizzes, discussions, grading, and so on. You can simply send a message on slack, or anonymously share your opinion:

`https://forms.gle/MzzNSV6Y8deJWGC77`

PREREQUISITES

Because producing visualizations using Python data & visualization stack is an integral part of the course, it is required to have a good understanding and working knowledge of programming (esp. Python), as well as working knowledge of using open-source libraries. It is also recommended to have a basic understanding of mathematics, statistics, and the Web (HTML, CSS, Javascript, and JSON).

For self-assessment, visit the following link: http://bit.ly/dvizselfassess. Contact the instructor if you are uncertain about your background.

EXPECTATIONS AND REQUIREMENTS

The primary assessment will be through the assignments, the final exam, and the final course project. The topic of the final project will be of your (team's) choice, but I encourage everyone to consult with the instructors. You are required to submit a final paper that contains not only the *results* but also detailed explanation of the visualization *process* to demonstrate your knowledge on visualization

principles and techniques, as well as your ability to apply them to create visualizations.

You are expected to complete all course modules (quizzes and discussions) and assignments, as well as visualization critiques through the "visualization of the week". You are also expected to engage in discussion on Canvas and Slack.

Residential students are expected to attend all course meetings to participate in quizzes and group discussions. If you cannot make it to class due to illness, you should contact the instructor and the AIs *before* class and let us know your situation. We can make accommodations for missed content such as quizzes. These are reviewed on a case-by-case basis.

GRADING

I sincerely hope that you focus on your learning and not on the grades! See `https://www.youtube.com/watch?v=u6XAPnuFjJc`

The grade may be curved at the end of the course. Moreover, the gradebook often has ungraded items. Therefore, *the grade that you can see on Canvas may not be a faithful reflection of your projected grade!*.

There will be extra credits based on your strong engagement in the course, in terms of sharing useful resources & interesting visualization-related articles, participating in discussions, and helping other students.

- · Attendance, Quiz, and Participation: 20%
- · Assignments: 20%
- · Exam: 30%
- · Final project: 30%

BOOKS AND KEY MATERIALS

There is no required textbook, but we will mainly use materials from the following books:

1. Fundamentals of Data Visualization by Claus O. Wilke (available online at `https://serialmentor.com/dataviz/`)
2. The Visual Display of Quantitative Information (2nd ed.) by E.R. Tufte: one of the foundational book on visualization. It contains a rich set of historical visualization, thoughtful discussion on visualization principles.

See also Visualization books and Data Visualization page on my wiki.

If you are still in the process of learning the basics of Python, the following books and websites may be helpful for you:

1. `https://docs.python.org/3/`: Python 3 Official Documentation
2. `http://www.diveintopython3.net/index.html`: Dive Into Python by Mark Pilgrim
3. `http://www.learnpython.org`: A web-based interactive tutorial

4. `http://ipython.rossant.net`: Learning IPython for Interactive Computing and Data Visualization by Cyrille Rossant: Introduction to IPython as well as lots of advanced analysis

If you are interested in web-based visualizations, you should check out the ObservableHQ and its tutorials for D3.js, Vega-Lite, and Observable Plot, all available at the following URL:

· `https://observablehq.com/tutorials`

FINAL PROJECT

See `https://github.com/yy/dviz-course/wiki/Projects` for the final project details, including the deliverables, types of projects, and some project ideas.

COURSE SCHEDULE

The schedule may change. See also IU Official Calendar.

Key dates

Mark your calendar and plan ahead!

· Project proposal due: **10/14**
· Project presentation files and final paper due: **12/5**
· Final Exam: During the final week of the semester.

Schedule

Week	Date	Topic
1	8/22–	Why visualization? | Visualization tools
2	8/29–	History and Integrity | Lilly Library visit
3	9/5–	Labor day | Perception
4	9/12–	Perception | Design principles
5	9/19–	Design principles | Data types and tidy data
6	9/26–	Data types and Tidy data | Project matchmaking
7	10/3–	Histogram and Boxplot
8	10/10–	Estimation
9	10/17–	Logscale and Beyond 1D
10	10/24–	Visualizing High-dimensional data
11	10/31–	Maps
12	11/7–	Maps | Text and networks
13	11/14–	Text and Networks | 11/16: guest lecture
14	11/21–	Thanksgiving break
15	11/28–	Project hack week
16	12/5–	Project Presentation
17	12/12–	Final exam week

POLICIES

1. *Be honest.* Don't be a cheater. Your assignments and papers should be your own work. If you find useful resources for your assignments, share them and cite them. If your friends helped you, acknowledge them. You should feel free to discuss both online and offline (except for the exam), but do not show your code directly. Any cases of academic misconduct (cheating, fabrication, plagiarism, etc) will be reported to the School and the Dean of Students, following the standard procedure. *Cheating is not cool.*

2. *You have the responsibility of backing up all your data and code.* Always back up your code and data. You should at least use Google Drive or Dropbox at the minimum. You can also use cloud services like Google Colaboratory. Ideally, learn version control systems and use `https://github.iu.edu` or `https://github.com`. Loss of data, code, or papers (e.g. due to malfunction of your laptop) is not an acceptable excuse for delayed or missing submission.

3. *Disabilities.* Every attempt will be made to accommodate qualified students with disabilities (e.g. mental health, learning, chronic health, physical, hearing, vision, neurological, etc.). You must have established your eligibility for support services through Disability Services for Students. Note that services are confidential, may take time to put into place, and are not retroactive. Captions and alternate media for print materials may take three or more weeks to get produced. Please contact Disability Services for Students at `http://disabilityservices.indiana.edu` or 812-855-7578 as soon as possible if accommodations are needed. The office is located on the third floor, west tower, of the Wells Library (Room W302). Walk-ins are welcome 8 AM to 5 PM, Monday through Friday. You can also locate a variety of campus resources for students and visitors who need assistance at `http://www.iu.edu/~ada/index.shtml`.

4. *Bias-based incidents.* Any act of discrimination or harassment based on race, ethnicity, religious affiliation, gender, gender identity, sexual orientation, or disability can be reported to `biasincident@indiana.edu` or to the Dean of Students Office at (812) 855-8188.

5. *Sexual misconduct and Title IX.* Title IX and IU's Sexual Misconduct Policy prohibit sexual misconduct in any form, including sexual harassment, sexual assault, stalking, and dating and domestic violence. If you have experienced sexual misconduct, or know someone who has, you can use university resources:
   a) The Sexual Assault Crisis Services (SACS) at (812) 855-8900 (counseling services)
   b) Confidential Victim Advocates (CVA) at (812) 856-2469 (advocacy and advice services)
   c) IU Health Center at (812) 855-4011 (health and medical services)
   It is also important that you know that Title IX and University policy require me to share any information brought to my attention about potential sexual misconduct, with the campus Deputy Title IX Coordinator or IU's Title IX Coordinator. In that event, those individuals will work to ensure that appropriate measures are taken and resources are made available. Protecting student privacy is of utmost concern, and information will only be shared with those that need to know to ensure the University can respond and assist. Visit *stopsexualviolence.iu.edu* to learn more.

6. If you have any mental health issues, don't hesitate to contact IU's Counseling and Psychological Services, which provides free counseling sessions. Also, please contact Disability Services for

Students at http://disabilityservices.indiana.edu or 812-855-7578 as soon as possible if accommodations are needed. See "Disabilities" section for more information.