

**INFO 422/I590**

# **Data Visualization**

**Yong-Yeol (YY) Ahn**

yyahn@iu.edu

Fall 2017. Info West 107 (M) / 109 (W)

MW 4:00pm–5:15pm.

Office hours: W 9am-10am

**Assistant Instructor**

TBD (TBD@indiana.edu)

Office Hours: TBD

## **COURSE HOMEPAGE**

This syllabus may contain outdated information. For the most accurate and up-to-date information, please check <https://yy.github.io/dviz-course/>.

## **COURSE DESCRIPTION**

From dashboards to cutting-edge scientific papers, from a home office to the largest companies in the world, we extensively use data visualization to reveal patterns in data. As our world becomes increasingly connected and digitized, and as decisions are increasingly driven by data, data visualization is becoming a critical skill for every knowledge worker. This course is an introduction to basic statistical data analysis and visualization. We will learn fundamentals of data visualization, in the context of visual perception, integrity, design, and statistics, as well as various types of data and visualization techniques. The hands-on exercises using Python stack and D3.js will be an integral part of the course.

## COURSE OBJECTIVES

By the end of the course, you are expected to understand, explain, and handle basic types of data, analyze them using basic exploratory visualization techniques, and create explanatory web-based visualizations. You will also be able to evaluate the effectiveness of data visualizations based on the principles of human perception, design, types of data, and visualization techniques.

## PREREQUISITES

This course is open to advanced undergraduate students as well as graduate students (I590). Because this course extensively uses programming (Python and Javascript), it is required to have good understanding and working knowledge of programming. The basic programming courses (I210 & I211, or equivalent) are required prerequisites. In addition, I308: “Information Representation” is a recommended prerequisite. It is also highly recommended to have working knowledge of statistics and web (HTML, CSS, Javascript).

Contact the instructor if you are uncertain about your background.

## REQUIREMENTS

You should attend and engage in every class. There will be reading materials that you are expected to read *prior* to the class. At the beginning of each class, there will be an *in-class quiz* based on assigned readings and materials from previous classes.

Assignments from classes and labs are integral part of the class. You cannot pass the class without completing these assignments.

The mid-term and final assessment will be based on team projects. The choice of project topic will be guided by the instructors and you have freedom to choose projects ranging from visualization-driven heavy data analysis projects to web-based explanatory visualization ones.

## BOOKS AND KEY MATERIALS

There is no required textbook, but the following books and websites are recommended.

### *Python and data analysis*

1. [Dive Into Python](#) by Mark Pilgrim (available online): a good Python book.
2. [Learnpython.org](#): A web-based interactive tutorial.
3. [An introduction to statistics \(with Python\)](#) by Thomas Haslwanter (available online): this book uses Python to explain basic statistics. It also contains a succinct tutorial for Python and data visualization using Python.
4. [Learning IPython for Interactive Computing and Data Visualization](#) by Cyrille Rossant: Introduction to IPython as well as lots of advanced analysis

### *Visualization and Design*

1. [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.
2. [Atlas of Knowledge: Anyone Can Map](#) by K. Börner: this book systematically analyzes vocabularies of visualization with a lot of great examples.
3. [Visualization Analysis and Design](#) by T. Munzner: a nice textbook that covers important topics of visualization.
4. [Visual Thinking for Design](#) by C. Ware: one of the best books on the role of perception in visualization.

### *D3.js*

1. [Interactive Data Visualization for the Web](#) by Scott Murray
2. [D3 documentation](#)

## POLICIES

1. *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>.

2. *No electronics—laptops, tablets, and smartphones—may be used in class*, unless the usage is specifically requested by the instructors. *It has been shown* that using laptops in class is not a good idea, *even if* you are using it to take notes. If you must have electronics due to a disability or other special reasons, please contact the instructor.
3. *Be honest*. Your assignments and papers should be your own work. First, if you find useful resources for your assignments, share them and cite them. If your friends helped you, acknowledge them. Second, feel free to discuss both online and offline, but you should not show your code (papers) nor see other's. Any cases of academic misconduct (cheating, fabrication, plagiarism, etc) will be immediately reported to the School and the Dean of Students, following the standard procedure. Cheating is not cool.
4. *You have the responsibility of backing up all your data and code*. Always use at least Box, Dropbox, or Google Drive. Ideally, learn version control systems and use <https://github.iu.edu> or <https://github.com>. Loss of data, code, or papers due to various reasons (e.g. malfunction of your laptop) is not an acceptable excuse for delayed or missing submission.
5. *Inform your excused absences prior to class*. Please contact the instructor until the previous day for an excused absence.
6. If you have any issues, don't hesitate to contact me or [IU's Counseling and Psychological Services](#).

## GRADING (TENTATIVE)

- Attendance, Quiz, and Participation (in-class and online): 20%
- Assignments (class and lab): 40%
- Mid-term and Final project: 40%

## COURSE SCHEDULE

The schedule is subject to change. Find the most up-to-date schedule at <https://github.com/yy/dviz-course/wiki/Schedule>

*Week 1: Why visualization? / Visualization tools*

*Week 2: History and integrity / Jupyter notebook / HTML*

*Week 3: Labor day / Python visualization tools / CSS*

*Week 4: Perception and Design / Javascript I*

*Week 5: Data Types and Fundamental Visualizations / Javascript II*

*Week 6: Fundamental Visualizations II / D3.js I*

*Week 7: Fundamental Visualizations III / D3.js II*

*Week 8: Fundamental Visualizations IV / Midterm Hack*

*Week 9: Midterm presentation / KDE, Regression*

*Week 10 High dimensional data*

*Week 11 Geospatial data and maps I*

*Week 12 Geospatial data and maps II*

*Week 13 Texts and Graphs*

*Week 14 Thanksgiving*

*Week 15 Graphs II*

*Week 16 Final project presentation*