CS-GY 6313: Information Visualization
Blended + Online — Fall 2020

Instructor Information

Name: Enrico Bertini
Office: Virtual
Email: enrico.bertini@nyu.edu
Office Hours: TBA

Teaching Assistants

Class Information

Dates: Sep 9, 2020 – Dec 7, 2020
Time: 5:00PM – 7:30PM
Classroom: Virtual on Zoom
Piazza Q&A: sign-up link, class home
Class material: google drive folder

Course Prerequisites

The coursework includes extensive programming with *JavaScript*, the *D3.js* library, and web technologies (*CSS*, *SVG*, etc.). While previous knowledge of these technologies is not required, being proficient and comfortable with extensive programming is a fundamental prerequisite for this course. If you are not comfortable with programming please contact the instructor before enrolling.

Course Description

Being able to analyze and present data visually has become one of the most important skills for students who want to work in data science and related fields. Information Visualization teaches you how to design effective interactive visualizations of complex data for data understanding, discovery, and presentation.

The course is a blend of theoretical knowledge and practical work aimed at developing a well-rounded set of skills to ideate, design, implement, and evaluate sophisticated data visualizations. The theoretical part aims at providing a mental model to think about the visualization design space in a principled manner. This includes the theory of visual encoding, human perception and visualization techniques.

The practical part aims at teaching the skills needed to develop effective interactive data visualizations for analysis and presentation. This includes teaching *D3.js* and *JavaScript* and practical labs on exploratory data analysis, sketching, graph design design, analysis and critique.

The course also includes a series of small practical projects which enable students to gain experience with the development of fully-working interactive visualizations to solve non-trivial assigned problems. The work is organized in a way to simulate conditions happening in real-world data analysis and communication projects and includes activities to gain feedback from the instructor and the teaching assistant.

Workload

The course is *intense*. We assign multiple assignments every week (see details below) and we do a lot of practical work during the class. It is very important for you to do all the work on time otherwise you may get lost and catching up is hard. The good news is that **you are going to learn a lot**! Most of our students are extremely happy about how much they have learned by the end of the course. You are going to acquire very solid intellectual and practical skills.

Course Objectives

After this course, you should be able to...

- Describe what problems visualization can solve and how.
- Describe the process needed to design effective visualizations.
- Explain how to evaluate a visualization.
- Use appropriate graphs for a data analysis/presentation problem.
- Describe any visualization in terms of its individual components ("deconstruction").
- Identify and solve effectiveness problems with a given visualization.
- Generate alternative solutions for a data visualization problem (design space exploration).
- Assess pros and cons of alternative solutions (comparative analysis).
- Describe visual perception mechanisms that impact visualization effectiveness.
- Make design decisions grounded on principles of visual perception.
- Design/develop effective visualizations for temporal, geographical, network data.
- Design/develop effective interactive functionalities for a data visualization.
- Design/develop interactive multiple views for data visualization (dashboards).
- Use JavaScript to manipulate and transform your data.
- Use JavaScript and D3 to develop and deploy interactive visualizations on the web.

Study Material and Textbooks

For this course there is **no specific textbook**. The following study material will be used instead:

- *Recorded video lectures:* You can find them in NYU Classes in the Lessons tool.
- *Lecture slides (corresponding to the video lectures):* You can find them in our Google Folder.
- *Research papers, articles and book excerpts:* You can find them in NYU Classes in the Lessons tool.
- *Observable notebooks:* You can find them in our Observable collection of notebooks we have for the class.

The material will be available through NYU Classes and a Google Drive folder we use for the course. Make sure to save a link to this folder somewhere. In any case, you can always find a link to it from the home page of the course in NYC Classes. The folder contains the following important sub-folders:

- *In-Class Activities:* The folder we use during the class to share work done in groups.
- *Lecture Slides:* The slides presented in the video lectures.
- *Exercises:* The instructions of the exercises assigned as homework.

The following texts are recommended (but not required):

- *Visualization Analysis and Design*, Tamara Munzner, CRC Press 2014.

- *Information Visualization: Perception for Design*, Colin Ware, Morgan Kaufmann 2019.
- *Interactive Data Visualization for the Web*, Scott Murray, O'Reilly Media 2017.
- *Fullstack Data Visualization with D3*, Amelia Wattenberger, Fullstack.io 2019.

Course Organization

The course is delivered using a **flipped classroom** format. This means that you will be watching (video-recorded) lectures and doing the readings at home, while you will be doing active work during our meetings. This is the best use of our time: (1) you first study at home and come in class with useful knowledge and questions; (2) when we meet you are ready to ask questions and do active work. We will act more like "coaches" for you than lecturers.

For this reason is very important for you to do the work required before our meetings take place and to be very active during our meetings. I can guarantee you that the more actively you participate during our meetings (asking questions, proposing ideas, sharing your thoughts, etc.), the more you are going to grow and learn.

Course Requirements

The course requires attendance, active participation and submission of all homework, including: exercises, quizzes and mini-projects.

Exercises. The first half of the course contains a set of exercises assigned for home as homework. The exercises are meant to teach you data visualization design/evaluation skills as well as the programming skills necessary to develop the solutions you design. The exercises count for 30% of your final grade. Each exercise can be resubmitted after grading. A re-submission can give you up to 75% of the lost points back, provided it addresses all the issues raised.

Quizzes. The main purpose of the quizzes is to test your knowledge *after* watching the lectures and doing the readings. Research has shown that answering questions (retrieval) is the best way to improve your memory and understanding. For this reason, we will keep releasing quizzes throughout the course so that you can consolidate your knowledge. For this same reason the quizzes will include both questions about the topics you studied for the current week as well as topics you studied in previous weeks. This is the best way to learn! On purpose the quizzes are "low-stake" assignments, that is, providing a wrong answer has a very minimal effect on your grade (if any).

Mini-Projects. In a mini-project you are asked to develop a data visualization for a specific data set and a specific set of questions. The assignment requires designing a solution and implementing it in D3/JavaScript.

The students will have to submit, for each assigned project, the following material: design sketches (developed before coding), short descriptions of the design and its rationale, working code showing the visualization implemented. More details will be provided at the time the projects are released.

All students are, on purpose, assigned the *same* projects. This will enable you to learn from the submissions your colleagues submit as well as the many discussions we will have during our meetings. The projects will be easier at the beginning and will gradually increase in complexity as the course progresses.

Each mini-project has one optional additional question and visualization to develop designed to give extra credits. Each question (one for each mini-project) will give you up to a total of 2 additional grade points to be added to your final grade. If your grade on this optional exercise is less than 100% the points will be scaled as $2 * p$, where p is the percentage points you received.

Class Attendance and Participation

Interacting with the instructor and with other students is crucial for your success in this course. Seriously. You are going to learn more from interaction than from watching lectures. For this reason, the course requires full attendance of classes, for the blended section, and of the weekly online meetings, for the on-line section. Attendance contributes to your final grade as 10% of the final grade (10 points out of 100). Missed meetings

result in reduction of attendance points according to the percentage of missed classes over the total number of classes. Note: class attendance is not recorded during the first 2 weeks of the course.

For special situations such as sickness and religious festivities, no attendance points will be removed *as long as you send your instructor and our TAs a note via email* **before** the class starts and follow the procedures defined by our school regarding absence. Please familiarize yourself with the NYU Tandon policies on excused absences. Attendance will be recorded at the beginning or end of each meeting every week.

Grading

The final grade in this course is calculated according to the following weights:

Grading breakdown:

- Attendance: 10%
- Exercises: 30%
- Quizzes: 10%
- Mini-Projects: 50%

Note: we do NOT have a final or a midterm exam in this course. The final grade will be computed from the grades you receive in individual assignments.

We have a number of *very important submission policies*. it is very important for you to familiarize yourself with these policies and to keep them in mind. I will assume you have read and understood these policies when grading your assignments.

Re-submission policy: In this course it is always possible to resubmit the exercises and the mini-projects after receiving a grade and feedback from your instructor. With a re-submission you can obtain up to 75% of the points lost in your first submission.

Late and no-submission policy: While we are pretty flexible on re-submissions and grading, we are strict with late and missing submissions. This is for two reasons: (1) We want to prevent you from falling behind. If you fall behind, it is going to be very hard to catch up in this course; (2) If you submit after the solutions have been released and discussed in class you would have an unfair advantage. For this reason we have find a way to penalize late submissions (no hard feelings). Keep in mind: it is always preferable to submit something that is not optimal that not submit at all.

These are the rules we will apply:

Late	**Penalty**
1-2 days	no penalty
2-7 days	25% reduction of assignment's grade
$>$ 7 days	50% reduction of assignment's grade
no-submission	2 points deducted **from final grade**

Important: it is your responsibility to check that you understand what it is to be submitted. When an assignment is posted make sure to read it carefully and ask for clarifications as soon as possible. If you ask at the very last minute we may not be able to give you the feedback that you need on time! Similarly, it is your responsibility to understand when a given assignment is due. It is never a justification to say: "*I made a mistake and thought the submission was due at a different time*".

Schedule

This is the tentative schedule for the course. I am going to add more details about the schedule during the second week of the course.

Week 1	Introduction and Evaluation
Week 2	Analytical Questions and Data Transformation
Week 3	Fundamental Graphs and Comparison
Week 4	Data Visualization Pitfalls
Week 5	Visual Encoding and Graphical Perception
Week 6	Color
Week 7	Multiple Views and Interactivity
Week 8–9	Geographic Data + Mini-projects (2 weeks)
Week 10–11	Temporal Data + Mini-projects (2 weeks)
Week 12–13	Network Data + Mini-projects (2 weeks)
Week 14–15	*Padding* (TBA)

Quoting Policy and Collaboration

The work students submit for individual assignments and class projects must be their own original work. When ideas are borrowed from existing work it is necessary to provide citations and a clear statement that describes which part has been adopted and which is original. For homework students are NOT allowed to collaborate with their peers. The submitted homework must be produced and submitted individually.

Academic Dishonesty

It's always annoying having to explain that copying work or cheating is not allowed. I like to totally trust each and everyone of you. But bad things happen and I have to warn you that academic dishonesty is a very serious thing and you might get in very serious trouble if caught cheating. Students caught in dishonest behavior will be reported to the school.

Moses Center Statement of Disability

If you are a student with a disability who is requesting accommodations, please contact New York University's Moses Center for Students with Disabilities at 212-998-4980 or mosescsd@nyu.edu. You must be registered with CSD to receive accommodations. Information about the Moses Center can be found at www.nyu.edu/csd. The Moses Center is located at 726 Broadway on the 2nd floor.