

DS594: Spark! Data Visualization X-Lab Practicum

Instructor Name: Anthony Chamberas
Office Location: 13th Floor
Contact Information: agc94@bu.edu
Office Hours: 206B TR 2:30-3:30

Course Dates: 09/05-12/12
Course Time: TR 3:30-4:45
Location: CDS 264
Course Credits: 4

Resources & Support:

Spark! Engineers and Experts in Residence: Find a list of all experts in residence and their officers here: <http://www.bu.edu/spark/resources/>

Course Description

The Data Visualization X-Lab Practicum offers students an opportunity to learn data visualization skills through course and project-based work. Projects will be completed on a schedule that aligns with topics being covered in class and assignments. This course provides an accurate experience of solving real-world problems with data visualization, and the various tradeoffs that need to be considered. Whether it's how to efficiently use color and space, effectively understand the profile of a dataset or cautiously avoid bias, this course will provide students with a solid understanding of applicable data visualization practices.

The course will encompass three distinct components related to data visualization:

1. Best practice instruction
2. Hands on development
3. Critique of public analyses

Ultimately, students will be assessed based on their progress in completing essential tasks within each part of the course as outlined above.

Students will learn the best practices necessary to tell a story with data, such as preparing data, selecting the correct chart type and colors, avoiding bias and using space effectively.

Project outcomes will be defined collaboratively between the student team, their project manager and the partner organization. Through this experience, students will learn the delicate art of managing client expectations and communicating technical processes to non-technical domain experts or users of data.

Critiques of public analyses will allow students to use what they have learned in class or in project work, to critically evaluate the work of professionals in the field.

Prerequisites

Students should come to the course with the following knowledge:

- Basic Excel/Google Sheets including formulas and pivot tables
- Basic algebra including equation of a line (slope, intercept)
- Basic Statistics, average, median, sum, count. Hypothesis testing a plus.
- Basic understanding of SQL a plus
- Basic Python or Javascript a plus

The prerequisites for this course are:

Data Science:

- CDS DS121
- CAS MA213 OR CAS MA214 OR CDS DS122
- CDS DS310

Computer Science equivalents:

- CAS CS132
- CAS CS237
- CAS CS460

Questrom equivalents:

- QS
- QS
- QS

Learning Outcomes:

This course will consist of methods and skills deriving mainly from computer science, engineering, mathematics, and statistics. The learning outcomes for this course are broadly defined as integrating the technical and soft skills necessary to produce impactful data visualizations that meet the expectations of real-world projects and partners. This includes:

- Evaluating data to answer or solve business and social problems
- Understanding how to communicate to technical and non-technical audiences
- Supporting hypotheses and analyses with data visualizations
- Leaving room for opposing points of view
- Being clear about outcomes and recommended actions

For data science majors, students should obtain approval from their advisors to determine if this course counts toward an In-the Field Elective or the Capstone Experience requirement. For computer science majors, students should obtain approval from their advisors to confirm this course counts toward one of two project-based courses under the Group D requirement.

Digital/Multimedia Expression

Learning Outcome 1:

Students will be able to craft and deliver responsible, considered, and well-structured arguments, statements, or expressions using appropriate digital media.

Digital media expression will be learned through various assignments and activities including: readings of data visualization best practices that discuss chart types, colors theory, layout design, bias and data storytelling; project work that addresses business or social problems through the analysis and synthesis of large volumes of data to identify patterns and anomalies; and critique of public and peer developed data visualizations.

This learning outcome will be evaluated through development of project-based data visualizations.

Learning Outcome 2:

Students will be able to reflect on the ethical use of digital media, considering relevant issues such as accessibility, intellectual property rights, citational practices, and other discipline-specific concerns.

The proper use of digital media will be addressed through readings on the importance of considering equity, inclusion and accessibility in data visualization. Students will also learn about how to use basic statistical inference to mitigate their own personal biases and prevent forgone conclusions. Lastly, all visualizations will be evaluated for appropriate citations, and responsible representation of referenced works.

This learning outcome will be evaluated through review of project-based data visualizations, and quizzes on readings about equity, inclusion and diversity.

Learning Outcome 3:

Students will be able to demonstrate an understanding of the capabilities of one or more digital communication technologies in their assignments.

Various data visualization technologies will be presented to students, and they will be asked to select one that will be their primary tool for the semester. The trade-offs

related to drag-and-drop vs. code-based interfaces, commercial vs. open source tools, and locally installed vs. cloud-based products will be discussed. Students will be encouraged to select or transition to tools within their comfort level that achieve the goals of the course.

This learning outcome will be evaluated through proficiency of the data visualization and presentation tools used for project work.

Learning Outcome 4:

Students will be able to demonstrate an understanding of the fundamentals of digital communication, such as principles governing design, time-based and interactive media, and the audio-visual representation of qualitative and quantitative data.

Students will be provided with a foundational understanding of graphic design as it applies to data visualization including basic color theory, information hierarchy, and the wide variety of methods to visualize data spanning charts, tables, graphs, maps, text, infographics, and dashboards. They will be taught how to select design frameworks that address the problem and audience.

This learning outcome will be evaluated through application of lecture topics to project-based data visualizations.

Oral and/or Signed Communication – OSC

Learning Outcome 1:

Students will be able to craft and deliver responsible, considered, and well-structured oral and/or signed arguments using media and modes of expression appropriate to the situation.

Project work will focus on building data visualizations for sponsor organizations. Students are expected to present their approach, design decisions and progress to the sponsor in bi-weekly meetings. These presentations will create an interactive cycle of presentation, discussion, feedback and improvement. These presentations will be practiced in class.

This learning outcome will be evaluated through in-class and client presentations.

Learning Outcome 2:

Students will demonstrate an understanding that oral/signed communication is generally interactive, and they should be able to attend and respond thoughtfully to others.

Students will be asked to explain their own work, as well as critique work of other students or data visualization professionals in the public domain. Emphasis will be placed on positioning data visualization as a subjective exercise that may have different points of view, which require open and interactive dialog.

This learning outcome will be evaluated through in-class and client presentations.

Learning Outcome 3:

Students will be able to speak/sign effectively in situations ranging from the formal to the extemporaneous and interact comfortably with diverse audiences.

Communication about data visualizations will range from full class presentations, to small group conversations, to project updates, to technical engineer feedback. Students will be taught to both defend their position and accept criticism for their work.

This learning outcome will be evaluated through in-class and client presentations.

Writing-Intensive Course – WIN

Learning Outcome 1:

Students will be able to craft responsible, considered, and well-structured writing that is appropriate to genre, intended.

Writing will be practiced across a number of genres, including: user guides, code documentation, data visualization critiques and written analysis of project outcomes.

This learning outcome will be evaluated through written analysis of public data visualizations and general project documentation.

Learning Outcome 2:

Students will be able to read and interpret texts, data, media, etc. with understanding, engagement, appreciation, and critical judgment.

As the result of studying data visualizations in the public domain, students will be asked to interpret the domain being analyzed, the author's intent, as well as provide subjective and objective critique of the work.

This learning outcome will be evaluated through study of public data visualizations.

Learning Outcome 3:

Students will be able to write clearly and coherently in a range of modes and styles, integrating graphic, multimedia, and other elements as appropriate to the genre.

While the intent of this course is to teach data visualization, students will be encouraged to use a variety of other media to complement their work, such as: images, videos, diagrams and the written word.

This learning outcome will be evaluated by the final project report which must include written analysis, visualizations, graphics and diagrams.

Books and Other Course Materials

Recommended:

- Tableau Software (licenses available)
- Flourish (freemium)

Required:

- [Storytelling with Data](#)

Courseware

- [Blackboard](#)

Assignments and Grading

Assignments serve the purpose of tracking individual and project progress. There will be a mix of assignments including:

- Quizzes on readings
- Writing and Digital Media Expression (DME)
 - Written or video critiques of public data visualizations with enhancements
 - Original data visualizations
- Project milestones reported to sponsor
- Project outcomes presented to sponsor
- Class participation including discussion and attendance

Assignments will contribute to the final grade with the following weights:

Category	%
Quizzes on readings	15%
Written and DME	30%
Project Milestones	15%
Project Outcomes	30%

Class Participation	10%
---------------------	-----

Extra Credit:

- Industry events
- > 10 writing or DME assignments
- exceptional final project or critiques

Assignments

Type	Topic	Assessment
Individual	Rank projects, data viz your life	Viz
Reading	SWD CH1: the importance of context	Quiz
Project	Write user requirements & storyboard	Paper (1p)
Reading	Data viz technologies	Quiz
Critique	What problem is being solved?	Paper (1p)
Reading	Basic Statistics	Quiz
Critique	Findings of analysis	Paper (1p)
Reading	SWD CH2: choosing an effective visual	Quiz
Critique	Chart selections OR Chart Champ	Paper (1p)
Reading	How to Choose Colors	Quiz
Critique	Color selections	Paper (1p)
Reading	SWD CH3: clutter is your enemy!	Quiz
Critique	Clutter or Too Simple	Paper (1p)
Reading	SWD CH 4: focus your audience's attention	Quiz
Critique	Engaging or distracting	Paper (1p)
Reading	IED, SWD CH5: Think like a designer (accessibility)	Quiz
Critique	Language, accessibility	Paper (1p)
Reading	Statistical Inference	Quiz
Critique	Misleading or accurate	Paper (1p)
Reading	SWD CH7: Lessons in storytelling	Quiz
Critique	What is the beginning, middle and end of the story	Paper (1p)
Reading	SWD CH9: Case studies	Quiz
Project	Final Presentation Prep	

Course Expectations

Project Expectations

- You should make yourself available to speak with your client on a bi-weekly basis (depends on client availability)
- You should meet with your Spark PM on a weekly basis
- For any team communication issues, please let your Spark PMs know ASAP - they are here to help. If the problem persists, please email faculty with a description of the situation.
- Projects may change over the course of the semester and students will be expected to meet client expectations as documented in the project document or refined scope as documented in team meeting notes and communications to the client/partner. While the components of a specific project's implementation scope may vary (data collection, data analysis, and data infrastructure), the total percentage of the grade will equal 45%.

Presentations

- Mock presentations will be given to the class and final presentations will be given to the customer.
- Presentations will be graded based on how effectively they utilize data visualization topics covered in lectures.

Documentation

- Thorough documentation of code, findings and guides. Labeling of ALL data files, folders, processes, and repositories are expected.

Class and University Policies

Attendance & Absences.

Due to the sequential nature of the project-based learning experience and the goal of completing the product by the end of the semester, attendance is required, including for all client meetings. Missing more than 3 sessions may affect your final grade. If you must miss class for any reason, please email the instructor ahead of time.

Assignment Completion & Late Work

Assignments, including scrum reports and new and refined project elements, are due the day before class at 11:59pm and may be submitted up to 24 hours late with a 5% late penalty for each day the assignment is late.

If you notice an issue with a grade you've received, you must email your TA within 48h of receiving this grade. Anything beyond 48h will not be accepted for a re-grade.

Religious Holidays & Mental Health

We understand that our student community is diverse and affirm the University's [Policy on Religious Observance](#). Please communicate in advance with faculty if you have a religious obligation that may impact attendance or meeting deadlines. Additionally, we

recognize that life is stressful and want to support your mental health. Students should feel supported by faculty to manage their personal mental health.

Borrowing Code and Academic Conduct

Software engineering is an inherently collaborative endeavor. In most cases, you will find open-source projects or code snippets on the internet that you might want to use in your own projects. While this is permitted, you *must* cite your sources appropriately. You are also responsible for ensuring that you have the original author's permission to use their work. The Open-Source Initiative maintains an excellent page on [the different types of software licenses](#) and what you can and cannot do with them. Using code you have borrowed from the internet without permission and/or attribution is an instance of plagiarism, which is a violation of the [Academic Code of Conduct](#). If you are in doubt about whether something might be construed as plagiarism, please check with course staff and in general—err on the side of caution. Collaboration on Assignments and Projects: Unless explicitly stated, collaboration on assignments and projects among teammates is both allowed and encouraged. Except by agreement with the partner, work product including documentation will be delivered on open source basis.

Disability Accommodations

Accommodations for Students with Documented Disabilities: If you are a student with a disability or believe you might have a disability that requires accommodations, please contact the Office for Disability Services (ODS) at (617) 353-3658 or access@bu.edu to coordinate any reasonable accommodation requests. ODS is located at 25 Buick Street on the 3rd floor. For more information, please see <http://www.bu.edu/disability>

Outline of Class Meetings

https://docs.google.com/spreadsheets/d/12llx2Dlf9r4YjnjqEKJ8Bc9rABCg2oDqVSve0-7Eax_M/edit#gid=1293722305