CUNY School of Professional Studies

Data 608 – Knowledge and Visual Analytics

Instructor: Lawrence Fulton, Ph.D. MSStat, CAP PStat
Class Meetup: Weekly Zoom (see announcements)
Office Hours: Saturdays and Sundays from 9-12 and by appointment (text 850-844-3178 to coordinate Zoom time)
Email: lawrence.fulton@sps.cuny.edu,

Course Description

This course is about the creation of high-quality data visualizations - diagrams, graphs/plots, maps/geographic, tables, etc. - that tell the story or stories hidden in the underlying data. You will learn the skills, techniques, and the art of "Story Telling with Data". You will gain the knowledge and experience necessary to source data and use the tools available to you as a Data Practitioner to build high quality Data Visualizations that tell a story, conveying the information that your audience needs to know and wants to hear.

Course Learning Outcomes:

By the end of the course, as a Data Practitioner you will:

- Be able to create high quality data graphics using various libraries (Python, R and/or Java Script), Web Applications, and/or Desktop Applications.
- Understand the quality factors that distinguish high quality data graphics.
- Have developed a "Critical Eye", the ability to discern poor quality and misleading data visualizations as differentiated from those of high quality.
- Be cognizant of the data story process and your role as a data practitioner in identifying the story audience and the required story content, determining the context of the story, and the iterative process of developing high quality data visualizations and accompanying text

Students will be required to:

- Demonstrate their ability to build high quality data visualizations.
- Demonstrate their ability to recognize poor quality and/or misleading visualizations published in online and /or print sources.
- Offer constructive critiques of data graphics produced by classmates in weekly class meetings to help each other improve their data visualization development skills

- Participate actively on Class Discussion Boards both initiating threads of discussion as well as participating in threads initiated by others.

Program Learning Outcomes addressed by the course:

- *Business Understanding*. Learn when analytical and/or probabilistic techniques apply to certain categories of business problems, and be able to build data stories that guide business decisions.
- *Foundational Data Visualization Skills*. Explore and analyze data, build probabilistic and statistical models, and create meaningful data stories using the craft of data visualization.
- *Presentation.* Complete and submit assignments using techniques from the course.

How is this course relevant for Data Practitioners?

Algorithms and models using probabilistic and/or analytical techniques, machine learning and Bayes analysis are the foundation of modern business analyses. This course will ensure that students have a strong understanding of these foundations as they relate to decisions making and the ability to communicate analysis results and conclusions using high quality data visualizations.

Course Prerequisites

The following courses are the listed prerequisites for this course:

- Data 602 - Advanced Programming Techniques, which in turn has as its prerequisite
- Data 607 - Data Acquisition and Management

Course Outline

Week	Starts	Ends	Discussion Topics (30%)	Major Assignments (70%)
1	1/25/2023	1/28/2023	Introduction & Setup Welcome, Syllabus Review, Course Overview, Assignments & Grading	**Story 1:** Infrastructure & Jobs Act Funding Allocation, Due by
2	1/29/2023	2/4/2023	Types of Data Visualizations	2/4/2023
3	2/5/2023	2/11/2023	Data Visualization Quality Features	**Story 2:** Can the FED Control Inflation and Maintain Full Employment? Due by
4	2/12/2023	2/18/2023	The Importance of Context & KYA (Know Your Audience)	2/18/2023
5	2/19/2023	2/25/2023	Pre-concious Vision	**Story 3:** Do Stricter Gun Laws Reduce Firearms Deaths? Due by
6	2/26/2023	3/3/2024	Sailence	3/3/2024
7	3/4/2024	3/10/2024	Emphasis -Effective Use of Color & Fonts	**Story 4:** How Much Do We Get Paid? Due by
8	3/11/2024	3/17/2024	Distributions & Histograms	3/17/2024
9	3/18/2024	3/24/2024	Project Score Cards & Burn-down Charts	**Story 5:** What Is The Effect of The Earth's Temperature On Cyclonic Storms? Due by
10	3/25/2024	3/31/2024	Multiple Figures	3/31/2024
11	4/1/2024	4/7/2024	Smoothing	**Story 6:** What Is The State of Food Security and Nutrition in the US? Due by
12	4/8/2024	4/14/2024	Visualizing Empirical Data	4/14/2024
13	4/15/2024	4/21/2024	Info-graphics	**Story 7:** Where Do Strategic Minerals Come From? Due by
14	5/1/2023	5/5/2023	Dynamic Visualizations	5/5/2023
15	5/6/2023	5/12/2023	Wrap up	

How This Course Works:

This course is entirely online. Participation in the discussion and optional Zoom meetings is essential and counts for 30% of your grade each week. Major assignments account for 70% of your grade.

Late Policies

Discussions must be completed in the week assigned or they become monographs. No late discussion work is accepted. Initial discussion posts are due on Wednesday of the academic week, and follow-up contributory responses to your fellow students are due by Sunday. Posts tantamount to 'good job' are not contributory. You should seek to expand the discussion and shared knowledge. Inappropriate or unprofessional commentary will

result in a zero grade for that week's discussion work.

Late major assignments are penalized at 33 1/3% per day without prior coordination. With coordination, that penalty may be reduced or waived.

Collaboration

Collaboration is encouraged, as it is essential in the workplace, but coding and visualization generation is the individual responsibility of each student. Do not collaborate on the assignments.

The use of ChatGPT and other AI tools is encouraged when fully documented and within reasonable bounds. If any work submitted contains analysis, text, or code that is tool generated it must be documented as such.

Grading

Each major assignment is worth 10 points (70 total). The grading rubric follows.

Criteria	Story Fails to Meet Expectations (0%)	Story Partially Meets Expectations (50%)	Story Meets or Exceeds Expectations (100%)
Fidelity (20 POINTS)	Visualization is not a true representation of the data.	Visualization is only partially true to the data.	Visualization is a high fidelity representation of the data.
Simplicity (10 POINTS)	Visualization is overly complicated and/or cluttered with chart junk.	Visualization can be simplified by eliminating unnecessary element.	Visualization is a simple as it can be.
Utility (20 POINTS)	The visualization is not fit for purpose. It doesn't tell the story of the data.	The visualization only partially tells the story of the data	The visualization tells the story of the data fully.
Saliency (10 POINTS)	The most important message of the story isn't told.	The most important message isn't emphasized.	The most important message is fully emphasized.
Efficacy (20 POINTS)	The visualization lacks clarity and is difficult to interpret.	The visualization tells parts of the story and is difficult to interpret.	The visualization tells the story fully and and can be interpreted straight forwardly.
Uniformity (10 POINTS)	The story lacks a theme.	The story visualizations violate the story theme unnecessarily.	The story has a theme and the visualizations consistently conform

			to it.
Amity (10 POINTS)	It is impossible for the viewer/reader to understand the story.	The viewer/reader has to work hard to understand the story told by the visualizations.	The viewer/reader will understand the story told by the visualization immediately without effort.

As noted previously, participation in the discussion accounts for 30% of your grade. The following Rubric will be used to value participation for each week's discussion.

Criteria	**Fails to Meet Expectations (0%)**	**Partially Meets Expectations (50%)**	**Meets or Exceeds Expectations (100%)**
Initiates Threads on Weekly Discussion Boards	Does not respond to the prompt	Responds partially to the prompt	Responds thoroughly to the discussion prompts
Comments Meaningfully on Threads Imitated by Others	Does not advance the discussion or respond to students	Advances the discussion but responds to only one student	Advances the discussion when responding to two or more students

Texts, Data Sources and Other References

Recommended Texts

Nussbaumer Naflic, Cole - *Story Telling With Data*; John Wiley & Sons ,(2015)
Available @ https://github.com/Saurav6789/Books-/blob/master/storytelling- with-data-cole-nussbaumer-knaflic.pdf

Wilke, Claus O. - *Fundamentals of Data Visualization*; O'Reilly (2019)
Available @ https://clauswilke.com/dataviz/boxplots-violins.html

Data Visualization Resources

The Data Visualization Catalogue @ https://datavizcatalogue.com/resources.html

Selected Data Sources

US Census @ https://www.census.gov/about/adrm/linkage/guidance.html Tyler

Technologies - Data Insights @ https://dev.socrata.com/

Centers for Disease Control https://data.cdc.gov/ United

Nations @ https://data.un.org/

NOAA @ https://www.noaa.gov/nodd/datasets

NOAA - Global Surface Temperature Dataset @

https://www.ncei.noaa.gov/access/metadata/landing-page

/bin/iso?id=gov.noaa.ncdc:C01585

US Bureau of Labor Statistics @ https://www.bls.gov/developers/home.htm Federal

Reserve Board @ https://www.federalreserve.gov/data.htm

Pew Research Center @ https://www.pewresearch.org/download-datasets/

Human Progress Datasets Archive @

https://www.humanprogress.org/datasets/#more-info

Contact

Dr. Larry Fulton
Lawrence.fulton@sps.cuny.edu

Accessibility and Accommodations

The CUNY School of Professional Studies is firmly committed to making higher education accessible to students with disabilities by removing architectural barriers and providing programs and support services necessary for them to benefit from the instruction and resources of the University. Early planning is essential for many of the resources and accommodations provided. Please see: http://sps.cuny.edu/student_services/disabilityservices.html

Online Etiquette and Anti-Harassment Policy

The University strictly prohibits the use of university online resources or facilities, including Blackboard, for the purpose of harassment of any individual or for the posting of any material that is scandalous, libelous, offensive or otherwise against the University's policies. Please see: http://media.sps.cuny.edu/filestore/8/4/9_d018dae29d76f89/849_3c7d075b32c268e.pdf

Academic Integrity

Academic dishonesty is unacceptable and will not be tolerated. Cheating, forgery, plagiarism and collusion in dishonest acts undermine the educational mission of the City University of New York and the students' personal and intellectual growth. Please

see: http://media.sps.cuny.edu/filestore/8/3/9_dea303d5822ab91/839_1753cee9c9d90e9.pdf

Student Support Services

If you need any additional help, please visit Student Support Services: http://sps.cuny.edu/student_resources/