CUNY School of Professional Studies

Data 608 – Knowledge and Visual Analytics

Instructor: Larry Fulton, Ph.D. MSStat
Class Meetup: Weekly Zoom (Posted in Announcements)
Location for Zoom: https://us02web.zoom.us/j/89708121650
Office Hours: Saturday and Sunday 8 am – 12 pm ET.
(Text 850- 844-3178 first)
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 Grading

Only two graded categories exist in this course, discussions and assignments. Their point values are shown below.

Category	Points
Discussions (20%)	15 discussions x 6.667 points x 20% weighting = 20 points
Assignments (80%)	7 stories x 14.289 points x 80% weighting = 80 points

How This Course Works

This course is entirely online. Discussions count for 20% of your grade; major assignments (stories) count for 80%. Weekly participation in the Zoom meetings can offset poor discussion performance at my discretion.

Collaboration

Collaboration is encouraged, as it is essential in the workplace, but coding and visualization generation is the individual responsibility of each student.

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.

Story Grading

Each major assignment (story) will be graded by the following rubric.

Criteria	**Story Fails to Meet Expectations (0%)**	**Story Partially Meets Expectations (50%)**	**Story Meets or Exceeds Expectations (100%)**
Fidelity 20%	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%	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%	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%	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%	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%	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%	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.

The Discussion Board accounts for 20% of your grade. The following Rubric will be used to value participation.

Criteria	Fails to Meet Expectations (0%)	Partially Meets Expectations (50%)	Meets or Exceeds Expectations (100%)
Addresses the Prompt by Wednesday	Fails to initiate threads	Initiate threads infrequently	Initiates threads almost every week
# of **Meaningful** Follow-up Posts	2	1	0

You must post a response by **Wednesday at midnight (ET)** and respond to at least one of your colleagues' contributions by **Saturday at midnight (ET)**, providing ***meaningful*** feedback on the analysis. By Sunday, the discussion board is a veritable ghost town, so no discussions occur.

Late Policy

Late assignments are penalized 20% per day. You have 5 weeks or more to complete each of the three assignments. All assignment due dates and times are shown in Blackboard. Late midterm examinations and final projects are not accepted. Late discussion posts are awarded no credit.

Texts, Data Sources and Other References

Recommended Texts

- Nussbaumer Naflic, Cole - *Story Telling With Data*; John Wiley & Sons ,(2115) 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

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_3c7d075b 32c268e.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_1753cee 9c9d90e9.pdf

Student Support Services

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