CSC 302

Introduction to Data Visualization (3 credits)- Summer 2022

Dr. Halil Bisgin

| Office Zoom: 2509859732 | Thu 12:30 - 2:30 PM, or by appointment |
| 810 762 3374 | bisgin@umich.edu | MSB 196 - (810) 762 3121 | csc302umf.slack.com |

Course Description

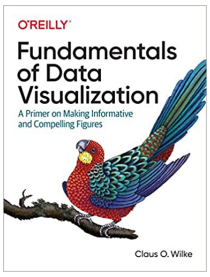
- Study of basic visualization techniques using different platforms ranging from scripting languages to off-the-shelf software packages along with necessary statistical measures. Includes an introduction to design and information literacy, fundamental data preprocessing techniques, dimension reduction, representation of time, spatial and network data, and ethical aspects in visualization. Examination of select topics in data science and machine learning.
- This class will not only introduce fundamentals of data visualizations, but also expose you several visualization tools, such as R, Python, Cytoscape, which will prepare you for any future project. Almost every class will be run like a workshop which will let you start practicing principles and methods covered on the same day. We will also cover some high-level statistical and linear algebra concepts through their applications on sample datasets which aim to broaden your vision regarding data preprocessing and manipulation steps before visualization.

Prerequisites

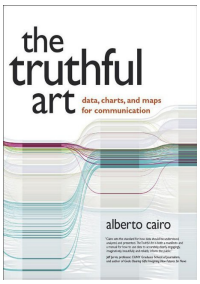
CSC 122; one from SCM 211, SOC 215, SWR 270, PSY 322, PHS 347, or prior or concurrent election of MTH 370 or MTH 372; or consent of instructor.

Textbook

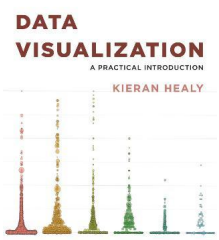
There is no required textbook for this course, but here are some recommended books which I will be using in course materials.



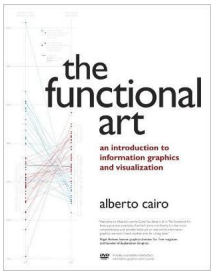
Fundamentals of Data Visualization:
A Primer on Making Informative and Compelling Figures
O'Reilly Media; 1st edition (2019)
by Claus O. Wilke.
ISBN: 978-1492031086



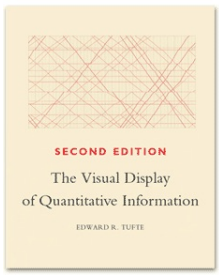
The Truthful Art,
Data, Charts, and
Maps for Communication
New Riders
by Albert Cairo
ISBN: 978-0321934079



Data Visualization: A practical introduction,
Princeton University Press; 1st edition (2019)
by Kieran Healy
ISBN: 978-0691181622



The Functional Art,
An introduction to
information graphics
and visualization
New Riders
by Albert Cairo
ISBN: 978-0321834737



The Visual Display of Quantitative Information,
Graphics Press; 2nd edition (2001)
by Edward R. Tufte
ISBN: 978-1930824133

Tools

- R (`https://www.r-project.org/`), RStudio(`https://www.rstudio.com/`)
- Python (`https://docs.continuum.io/anaconda/ide_integration`, `https://www.python.org/`, `https://colab.research.google.com`)
- PyCharm(`https://www.jetbrains.com/pycharm/`)
- Leaflet (`https://leafletjs.com/`)
- ArcGIS (`ArcGIS.com`)
- Cytoscape (`https://cytoscape.org/`)
- Gephi (`https://gephi.org/users/download/`)
- Tableau (`https://www.tableau.com/academic/students`, `https://public.tableau.com/s/`)

Course Objectives

- Understand the fundamental principles for effective data visualization.
- Familiarize with available visualization tools.
- Have an understanding of the distinction between good, bad, and ugly figures.
- Learn why and how figures can misinform or mislead.
- Know how to create a wide range of plots by using different platforms

Student Learning Outcomes

(Based on Bloom's taxonomy of cognitive domain as marked in parenthesis after each learning outcome) You will:

1. understand the foundations of the visualization processes, from basic building blocks to taxonomies and frameworks (knowledge, comprehension)
2. understand the visualization pipeline (comprehension, synthesis, evaluation)
3. understand the design considerations for the components of the good visualization (synthesis)
4. know the methods and algorithms used to map data to graphical depictions (synthesis, application)
5. understand the visualization design process (knowledge, evaluation)
6. know a wide range of interaction techniques and styles (knowledge, comprehension, analysis)
7. understand the characteristics and methods that are needed for the visualization of geospatial data (knowledge, comprehension, application)
8. know some commercial data visualization packages with functionality (knowledge, comprehension, application)

Tentative Schedule:

Week	Lecture	Topics	Assignments/Deadlines
Week 1	Lec 1	Introduction to the Course	Logistics, installations
	Lec 2	Introduction to Data Visualization	
	Lec 3	Introduction to Data Visualization (Cont.d)	
	Lec 4	Mapping Data onto Aesthetics, Coordinate Systems and Axes	
Week 2	Lec 5	Data Manipulation with Python	
	Lec 6		HW1 out
	Lec 7 Lec 8	Data Manipulation with R	Group formation HW2 out

Week 3	Lec 9	Color Scales, Directory of Visualizations	Slide 8
	Lec 10	Visualizing Amounts, Visualizing Distributions: Histograms And Density Plots	Slide 9 HW3 out
	Lec 11	Visualizing Many Distributions At Once	Slide 10
		Visualizing Proportions	
	Lec 12	Visualizing Nested Proportions & Associations	Slide 11
Week 4	Lec 13	Visualizing Time Series & Trends	Slide 12, HW4 out
	Lec 14	Visualizing Geospatial Data	Slide 13
Week 5	Lec 15	Visualizing Geospatial Data	Slide 14, HW5 out
	Lec 16	Multipanel Figures	Slide 15
	Lec 17	Titles, Captions, and Tables	Slides 16
	Lec 18	Network Visualization with Cytoscape - Gephi	Slides 17
Week 6	Lec 19		Slides 18, HW6 out
	Lec 20	Tableau - An overview	Slides 19-20
	Lec 21		HW7 out
	Lec 22	How to refine our Plots	Slides 21
Week 7	Lec 23		Slides 22
	Lec 24	Undestanding image formats & choosing right software Story telling with your data	Slides 23 Slides 24

Important Dates	
Group formation	Jul 8
Dataset Due	Jul 15
Midterm Presentations	Jul 22
Final Presentations	Aug 15
Drop individual class(es) 100% refund	Jul 11, 5pm.
Drop individual class(es) no refund w/W grade	Jul 12-22
Withdraw (drop all classes) 50% refund w/W grade	Jul 12-22

Assessment of Student Learning Outcomes

Group formation	5%
Dataset	5%
Homework	25%
In-class Assignments (ICAs*)	20%
Midterm Presentation	20%
Final Presentation	25%

*Almost every class, we will have hands-on practices or discussion threads and some of them will be asked to be submitted as your ICA. While regular homework assignments will be graded based on the quality of the work, ICAs will be granted 100 upon submission even if they're incomplete.

Grading Scale

A+	97-100	C+	78-79
A	92-96	C	72-77
A-	90-91	C-	70-71
B+	88-89	D+	68-69
B	82-87	D	60-67
B-	80-81	E	0-59

Grading may be curved if the class performance warrants it. Curving can only improve grades from the grading scale above

Homework Remarks

- These are individual home works, and not group assignments. Although you are allowed and even encouraged to discuss the general concepts behind the home works and even outline of solutions with your classmates, multiple

students MUST NOT work out one answer, and then submit this answer.

- I will accept late assignments. However, late assignments carry with them a 10% reduction in the grade, per week.
- All assignments must be turned in by three calendar weeks of their due date or you will receive a zero (0) grade for the assignment.
- Please note that if you are unable to turn an assignment in on the due date, then it will be considered late unless you have made arrangements with me at least twenty-four (24) hours before the assignment is due.
- Only submission via Canvas will be accepted.
- Some of the homework assignments may have to be demonstrated to the instructor.

Project Remarks

The project has two milestones, i.e., midterm and final, as indicated in the schedule. Both milestones require a presentation by the project groups where contribution of each member should be clear. While **midterm presentations** should have at least four different plots to describe the dataset and variables with your conclusions and inferences, final presentations should have more extensive visualizations with the new techniques and tools introduced. You can use different datasets for your your midterm and final presentations. Furthermore, you can use multiple datasets for the same topic to integrate and tell a story. In both stages, you are expected to provide the following content in your presentations:

1. You should organize each presentation to tell a story with the figures in it. This criterion will carry more weight for the final presentation.
2. Figures and their titles. Please make sure you use at least **four types of figures**, e.g., bar plot, scatter plot, etc. for the midterm and **12 figures (at least four types)** for the final by using every tool introduced so far. If your dataset doesn't have spatial and network data, you're not required to use Cytoscape and spatial data visualization tools.
3. Reasons to generate those figures. What was your purpose and why do you think these figures will capture the information you want to convey.
4. Your inferences for every figure. For instance, you can say "We observe a positive correlation between height and weight. As height increases, weight also increases according to the figure"
5. Your conclusions after presenting all visualizations.

The deadline to form a project group is **Jul 8**. Notice that forming a group is worth 5 points towards final score. You can use Canvas to manually enroll a project group.

You are encouraged to form groups with a maximum of three members. The scope of the project is such that a student by themselves can successfully complete the project. You must form your own groups. Talk to your peers (in class, using Slack etc) to form groups.

Here are some resources where you can find sample datasets and research problems:

1. UN Data:
   `http://data.un.org/`
2. Kaggle repository:
   `https://www.kaggle.com/`
3. UC Irvine Machine Learning Repository:
   `http://archive.ics.uci.edu/ml/index.php`
4. National Consortium for the Study of Terrorism and Responses to Terrorism:
   `http://start.umd.edu/data-and-tools/start-datasets`

Academic Integrity

(`http://catalog.umflint.edu/content.php?catoid=2&navoid=80#Student_rights/`)

Intellectual integrity is the most fundamental value of an academic community. Students and faculty alike are expected to uphold the highest standards of honesty and integrity in their scholarship. No departure from the highest standards of intellectual integrity, whether by cheating, plagiarism, fabrication, falsification, or aiding and abetting dishonesty by another person, can be tolerated in a community of scholars. Such transgressions may result in action ranging from reduced grade or failure of a course, to expulsion from the University or revocation of degree.

It is the responsibility of all students and faculty to know the policies on academic integrity in the instructional units at the University of Michigan-Flint. Information about these policies and the appeals process is available from the appropriate administrative office of the instructional units: in the College of Arts and Sciences, the Office of the Dean of the College of Arts and Sciences; in the School of Education and Human Services, the Office of the Dean of the School of Education and Human Services; in the School of Management, the Office of the Dean of the School of Management; in the School of Health Professions and Studies, the Office of the Dean of the School of Health Professions and Studies and for graduate students, the Office of the Dean of Graduate Programs. Departments and programs within these instructional units may have specific policies and procedures which further delineate academic integrity. In such cases students are bound by the University policy on academic integrity as well as these department or program policies.

Procedural Rights of the Accused Student

A student who is charged with academic dishonesty by an instructor, administrator, or another student may be assured that he/she has the right to a fair hearing of the charges and the evidence, the right to question witnesses, to invite witnesses on his/her behalf, and to introduce whatever other evidence may be relevant to the charge.

Code of Academic Conduct

The University, like all communities, functions best when its members treat one another with honesty, fairness, respect, and trust. Therefore, an individual should realize that deception for the purpose of individual gain is an offense against the members of the community. Such dishonesty includes:

- **Plagiarism:** taking credit for someone else's work or ideas, submitting a piece of work (for example, an essay, research paper, assignment, laboratory report) which in part or in whole is not entirely the student's own work without fully and accurately attributing those same portions to their correct source.
- **Cheating:** using unauthorized notes, or study aids, or information from another student or student's paper on an examination; altering a graded work after it has been returned, then submitting the work for regrading; allowing another person to do one's work, then submitting the work under one's own name.
- **Fabrication:** fabricating data; selectively reporting or omitting conflicting data for deceptive purposes; presenting data in a piece of work when the data were not gathered in accordance with guidelines defining the appropriate methods of collecting or generating data; failing to include a substantially accurate account of the method by which the data were gathered or collected.
- **Aiding and Abetting Dishonesty:** providing material or information to another person when it should reasonably be expected that such action could result in these materials or information being used in a manner that would violate this code of academic integrity.
- **Falsification of Records and Official Documents:** altering documents affecting academic records; forging a signature of authorization or falsifying or omitting necessary information on an official academic document, election form, grade report, letter of permission, petition, or any document designed to meet or exempt a student from an established College or University academic regulation; falsification or unauthorized altering of information in any official academic computer file.
- **Identity Theft:** Assuming another person's identity or role through deception or without proper authorization. Communicating or acting under the guise, name, identification, email address, signature, or indicia of another person without proper authorization, or communicating under the rubric of an organization, entity, or unit that you do not have the authority to represent.
- **Misrepresentation and Other Acts of Academic Dishonesty:** Fraudulently obtaining and/or using academic materials that would give oneself an unfair advantage over other students or would deceive the person evaluating one's academic performance.

An attempt to commit an act prohibited by this code may be punished to the same extent as a completed violation.

Accessibility Issues

It is my intention to support the full participation of all students in the learning process of this class. Students with disabilities that may restrict their full participation in course activities are encouraged to meet with the instructor or to contact the Office of Accessibility Services.

Accommodations: The University of Michigan?Flint strives to make learning experiences as accessible as possible and complies with Section 504 of the Rehabilitation Act of 1973 and the American with Disabilities Act. The university provides individuals with disabilities reasonable accommodations to participate in educational programs, activities, and services. Students with disabilities requiring accommodations to participate in class activities or meet course requirements must self-identify with Disability and Accessibility Support Services as early as possible at (810) 762-3456 or dassflint@umich.edu. The office is located in 264 University Center, inside the CAPS Office. Once your eligibility for an accommodation has been determined you will be issued an Accommodation Letter. Please present this letter to each faculty member in each class at the beginning of the term, or at least two weeks prior to the need for the accommodation (test, project, etc.).

Available Support Services

There is a plethora of support services available to students from tutoring to mental health services. Many times students are unaware of the services available to them. One such service is tutoring:

`http://www.umflint.edu/tutoring/student-success-center-tutoring`

For other services, please check Student Success Center at

`http://www.umflint.edu/studentsuccess/student-success-center`

Notes

I reserve the right to modify course policies, the course calendar, assignment point values, and due dates. Any extenuating circumstances that hinder your participation in the course should be discussed with me as soon as those circumstances are known. Make-ups for graded activities may be arranged if an absence is caused by documented illness or personal emergency. A written explanation, including supporting documentation, must be submitted to me; if the explanation is acceptable, then an alternative to the graded activity will be arranged. Whenever possible, make-up arrangements must be completed prior to the scheduled activity.