

EPPS6356 Data Visualization

Spring 2018 Wednesdays: 4:00 pm - 6:45 pm GR4.208

Instructor: Dr. Karl Ho; 972-883-2017; kyho@utdallas.edu
Office Hours: GR 3.203 Tuesdays and Thursdays 2 pm – 4 pm, by appointment.

Overview: This is a data method course focusing on data literacy through communicating data patterns, findings and insights via visual representation of data. This course is designed to equip future data scientists with data literacy including data theory, principles and concepts of visualizing data and best and newest data analytics tools. Topics on new developments and tools of data analytics in social science research will also be covered.

Learning Objectives:

At completion of this course, students will be able to:

1. Understand data theory
2. Be familiar with principles behind effective data visualization
3. Communicate message using visual form of data
4. Conduct independent research focusing on visual data analytics

Required Text:

Healy, Kieran. 2017. *Data Visualization for Social Science: A practical introduction with R and ggplot2* (online edition: <http://socviz.co>)
Murrell, Paul. 2016. *R graphics*. CRC Press. (‡)
Ware, Colin. 2012. *Information visualization: perception for design*. Elsevier. (‡)
Yau, Nathan. 2011. *Visualize This: The Flowingdata Guide To Design, Visualization, And Statistics*. John Wiley & Sons. (‡)
Shiny Tutorial online: <http://shiny.rstudio.com>

(‡) Available electronically at UTD Library

Recommended Reading:

Beeley, C., 2013. *Web application development with R using Shiny*. Packt Publishing Ltd.
King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Interpretation and Presentation." *American Journal of Political Science*. 44(2) 341-355.
Andrew Gelman, Cristian Pasarica, Rahul Dodhia. 2002. "Let's practice what we preach: Turning tables into graphs." *The American Statistician*. 56:2(May): 121-130.
Jake Bowers and Katherine W. Drake. 2005. "EDA for HLM: Visualization when Probabilistic Inference Fails." *Political Analysis*. 13(4): 301-326.
Tremmel, Lothar. 1995. "The Visual Separability of Plotting Symbols in Scatterplots." *Journal of Computational and Graphical Statistics*. 4(2): 101-112.
Ben Fry. 2007. *Visualizing Data*. O'Reilly. (‡)
Edward R. Tufte. 2001. *The Visual Display of Quantitative Information*. Graphics Press. 2nd ed.
Cleveland, W.S., 1985. *The elements of graphing data* (pp. 135-143). Monterey, CA: Wadsworth Advanced Books and Software.
Wong, D. (2014) *Wall Street Journal Guide to Information Graphics*, Norton.
Wills, Graham. 2011. *Visualizing time: Designing graphical representations for statistical data*. Springer Science & Business Media.
Lamigueiro, Oscar Perpinan. 2014. *Displaying Time Series, Spatial, and Space-Time Data with R*. Chapman & Hall/CRC.

Participation: Full attendance of all classes is required and imperative. Attendance however is not enough. All class members have to actively participate in class preparation and discussion. Participation entails full preparation for class including research of class materials, completion of assigned readings and full involvement in class discussion. Only medical

emergency (with documentation) is considered excused absence. If you have a special medical condition, please contact me in first week and special arrangements can be made. Participation is responsible for 10% of the final grade.

Class Conduct: Computer is only allowed for class content related activities. Prudent use of computer is strongly advised. More details on computer use policy will be announced in class.

Grading and Requirements:

Participation:	10%
Assignments	30%
Project presentation:	30%
Final project	30%
Hackathons	10%

Assignments:

There will be multiple assignments carrying different weights. Assignments must be individually and independently prepared and accomplished. Please read carefully the University's policy about cheating and plagiarism (see below for a brief version). More detail will be announced in class. All assignments will account for 30 percent of the final grade.

Project:

The final project must be no more than 15 pages in length, not including the codes (as appendix). Originality is the first quality I demand. Please read carefully the University's policy about cheating and plagiarism (see below for a brief version). Due date is listed on the schedule below. No late submission will be accepted. A proposal must be submitted in advance for approval of topic and methods by the instructor. Details and sample topics will be discussed in class. The final project must be preceded by the class presentation, which is of equal weight with the final project (30 percent each).

Hackathons

According to Wikipedia, a **hackathon** is a design sprint-like event in which [participants] collaborate intensively on software projects. Class members will be given a special assignment to accomplish within a timeframe (e.g. 48 hours). More details will be provided in class. Hackathon can be considered a bonus assignment but in a contest.

Document guideline:

All documents in this class must adhere to the following general guidelines:

- Must be typed or word-processed on letter size papers
- Document soft copy format is required to be in PDF.
- Required image format is PNG.
- If hard copy, document needs be stapled on upper left-hand corner and one inch on all margins
- No binders or plastic covers
- For final project, use a cover sheet with topic and name

Tentative Topics for entire course (not necessarily weekly)

Topic	Date	Reading
Introduction	1/10	
Data Theory	1/17	KH: 1
Messages in data	1/24	NY: 1, 2
Cognitive Science in DV	1/31	CW: 1, 4, 5
Tools survey+R workshop	2/7	NY: 3, PM: 1, 3, 6, 7
Tools survey+Stata workshop	2/14	TBA
Python workshop	2/21	TBA
R ggplot2	2/28	KH: 2, 3, 4
R Shiny	3/7	http://shiny.rstudio.com
Spring break	3/14	Spring Break

Brief D3	3/21	TBA
Data topic: Time Series Data	3/28	NY: 4; GW: 3, 4
Data topic: Spatial Data	4/4	NY: 8; OPL: 7, 8, 9
Data topic: Differences	4/11	NY: 7
Data topic: Relationships and Models	4/18	NY: 6
Project presentations	4/25	
Project presentations (Final Project Due)	5/2	

The following is an extract from the UTD Syllabus policies published by the Provost's office. Please read it thoroughly and if needed, refer to permanent URL at: <http://go.utdallas.edu/syllabus-policies>

Student Conduct and Discipline

The University of Texas System and The University of Texas at Dallas have rules and regulations for the orderly and efficient conduct of their business. It is the responsibility of each student and each student organization to be knowledgeable about the rules and regulations which govern student conduct and activities. General information on student conduct and discipline is contained in the UT Dallas printed publication, *A to Z Guide*, which is available to all registered students each academic year.

The University of Texas at Dallas administers student discipline within the procedures of recognized and established due process. Procedures are defined and described in the *Rules and Regulations, Series 50000*, Board of Regents, The University of Texas System, and in *Title V, Rules on Student Services and Activities* of the university's *Handbook of Operating Procedures*. Copies of these rules and regulations are available to students in the Office of the Dean of Students, where staff members are available to assist students in interpreting the rules and regulations (SU 1.602, 972/883-6391) and online at <http://www.utdallas.edu/judicialaffairs/UTDJudicialAffairs-HOPV.html>.

A student at the university neither loses the rights nor escapes the responsibilities of citizenship. He or she is expected to obey federal, state, and local laws as well as the Regents' Rules, university regulations, and administrative rules. Students are subject to discipline for violating the standards of conduct whether such conduct takes place on or off campus, or whether civil or criminal penalties are also imposed for such conduct.

[Added July 2010] Students are expected to be attentive during class and to participate actively in group activities. Students are expected to listen respectfully to faculty and to other students who are speaking. Racism, sexism, homophobia, classism, ageism, and other forms of bigotry are inappropriate to express in class. Classes may discuss issues that require sensitivity and maturity. Disruptive students will be asked to leave and may be subject to disciplinary action.

Academic Integrity

The faculty expects from its students a high level of responsibility and academic honesty. Because the value of an academic degree depends upon the absolute integrity of the work done by the student for that degree, it is imperative that a student demonstrate a high standard of individual honor in his or her scholastic work.

Scholastic Dishonesty: Any student who commits an act of scholastic dishonesty is subject to discipline. Scholastic dishonesty includes but is not limited to cheating, plagiarism, collusion, submitting for credit any work or materials that are attributable in whole or in part to another person, taking an examination for another person, or any act designed to give unfair advantage to a student or the attempt to commit such acts.

Plagiarism, especially from the web, from portions of papers for other classes, and from any other source, is unacceptable and will be dealt with under the university's policy on plagiarism (see general catalog for details). This course will use the resources of *turnitin.com*, which searches the web for possible plagiarism and is over 90% effective.

Copyright Notice

The copyright law of the United States (*Title 17, United States Code*) governs the making of photocopies or other reproductions of copyrighted materials, including music and software. Copying, displaying, reproducing, or distributing copyrighted works may infringe upon the copyright owner's rights and such infringement is subject to appropriate disciplinary action as well as criminal penalties provided by federal law. Usage of such material is only appropriate when that usage constitutes "fair use" under the Copyright Act. As a UT Dallas student, you are required to follow the institution's copyright policy ([Policy Memorandum 84.1.3-46](http://policy.memorandum.84.1.3-46)). For more information about the fair use exemption, see <http://copyright.lib.utexas.edu/copypol2.html>.

Email Use

The University of Texas at Dallas recognizes the value and efficiency of communication between faculty/staff and students through electronic mail. At the same time, email raises some issues concerning security and the identity of each individual in an email exchange. The university encourages all official student email correspondence be sent only to a student's UT Dallas email address and that faculty and staff consider email from students official only if it originates from a UT Dallas student account. This allows the university to maintain a high degree of confidence in the identity of all individuals corresponding and the security of the transmitted information. UT Dallas furnishes each student with a free email account that is to be used in all communication with university personnel. The Department of Information Resources at UT Dallas provides a method for students to have their UT Dallas mail forwarded to other accounts.

Class Attendance

Regular and punctual class attendance is expected. Students who fail to attend class regularly are inviting scholastic difficulty. Absences may lower a student's grade where class attendance and class participation are deemed essential by the instructor. In some courses, instructors may have special attendance requirements; these should be made known to students during the first week of classes.

Withdrawal from Class

The administration of this institution has set deadlines for withdrawal from any college-level courses. These dates and times are published in that semester's course inventory and in the academic calendar. Administration procedures must be followed. It is the student's responsibility to handle withdrawal requirements from any class. In other words, a professor or other instructor cannot drop or withdraw any student. You must do the proper paperwork to ensure that you will not receive a final grade of "F" in a course if you choose not to attend the class once you are enrolled.

Student Grievance Procedures

Procedures for student grievances are found in *Title V, Rules on Student Services and Activities*, of the university's *Handbook of Operating Procedures*.

In attempting to resolve any student grievance regarding grades, evaluations, or other fulfillments of academic responsibility, it is the obligation of the student first to make a serious effort to resolve the matter with the instructor, supervisor, administrator, or committee with whom the grievance originates (hereafter called "the respondent"). Individual faculty members retain primary responsibility for assigning grades and evaluations. If the matter cannot be resolved at that level, the grievance must be submitted in writing to the respondent with a copy of the respondent's School Dean. If the matter is not resolved by the written response provided by the respondent, the student may submit a written appeal to the School Dean. If the grievance is not resolved by the School Dean's decision, the student may make a written appeal to the Dean of Graduate or Undergraduate Education, and the dean will appoint and convene an Academic Appeals Panel. The decision of the Academic Appeals Panel is final. The results of the academic appeals process will be distributed to all involved parties.

Copies of these rules and regulations are available to students in the Office of the Dean of Students, where staff members are available to assist students in interpreting the rules and regulations.

Incomplete Grade Policy

As per university policy, incomplete grades will be granted only for work unavoidably missed at the semester's end and only if 70% of the course work has been completed. An incomplete grade must be resolved within eight (8) weeks from the first day of the subsequent long semester. If the required work to complete the course and to remove the incomplete grade is not submitted by the specified deadline, the incomplete grade is changed automatically to a grade of F.

Office of Student AccessAbility (OSA)

UT Dallas is committed to providing equal educational opportunities for students with documented disabilities to all University courses and programs. Any student with a documented physical, psychological, or learning disability, which affects his/her academic performance, is eligible for services. If you need classroom accommodations, please make an appointment with the Office of Student AccessAbility at: SSB 3.200 (3rd Floor in SSB), 972-883-2098. Office hours are Monday – Thursday, 8:30 a.m. - 6:00 p.m., Friday 8:00 a.m. – 5:00 p.m. Evening appointments are available by request. <http://www.utdallas.edu/studentaccess/>

OSA provides registered students with an accommodation letter to present to faculty members. The letter verifies that the student is qualified to receive certain accommodations. The accommodation letter should be presented to instructors of each course at the beginning of the semester. The approved accommodations should be discussed at that time. It is the student's responsibility to notify his/her professor of their needs.

The University of Texas at Dallas is proud to be an educational institution that welcomes and supports a diverse student body.

Religious Holy Days

The University of Texas at Dallas will excuse a student from class or other required activities for the travel to and observance of a religious holy day for a religion whose places of worship are exempt from property tax under Section 11.20, Tax Code, *Texas Code Annotated*.

The student is encouraged to notify the instructor or activity sponsor as soon as possible regarding the absence, preferably in advance of the assignment. The student, so excused, will be allowed to take the exam or complete the assignment within a reasonable time after the absence: a period equal to the length of the absence, up to a maximum of one week. A student who notifies the instructor and completes any missed exam or assignment may not be penalized for the absence. A student who fails to complete the exam or assignment within the prescribed period may receive a failing grade for that exam or assignment.

If a student or an instructor disagrees about the nature of the absence [i.e., for the purpose of observing a religious holy day] or if there is similar disagreement about whether the student has been given a reasonable time to complete any missed assignments or examinations, either the student or the instructor may request a ruling from the chief executive officer of the institution, or his or her designee. The chief executive officer or designee must take into account the legislative intent of *TEC 51.911(b)*, and the student and instructor will abide by the decision of the chief executive officer or designee.

Avoiding Plagiarism

[Adapted from Duke University's guidelines for writers; added July 2010]

Take time to make careful choices among--and learn to use--the research tools available to you. You will probably find that your favorite web search engine is not adequate by itself for college-level research. Consult with your professor or a librarian. You may need to use specialized research tools, some of which may require learning new searching techniques.

Expect to make trips to the library. While you can access many of the library's resources from your home computer, you may find that you need to make several trips to use materials or research tools that are not accessible remotely. Of course, you will be seeking the best information, not settling for sources simply because they happen to be available online.

Allow time for gathering materials that are not available at UT Dallas. The InterLibrary Loan Office can borrow articles and books from other libraries, but this process takes additional time.

Allow time for reading, rereading, absorbing information, taking notes, synthesizing, and revising your research strategy or conducting additional research as new questions arise.

Sloppy note-taking increases the risk that you will unintentionally plagiarize. Unless you have taken note carefully, it may be hard to tell whether you copied certain passages exactly, paraphrased them, or wrote them yourself. This is especially problematic when using electronic source materials, since they can so easily be copied and pasted into your own document.

Identify words that you copy directly from a source by placing quotation marks around them, typing them in a different color, or highlighting them. (Do this immediately as you are making your notes. Don't expect to remember days or weeks later what phrases you copied directly.) Make sure to indicate the exact beginning and end of the quoted passage. Copy the wording, punctuation and spelling exactly as it appears in the original.

Jot down the page number and author or title of the source each time you make a note, even if you are not quoting directly but are only paraphrasing.

Keep a working bibliography of your sources so that you can go back to them easily when it's time to double-check the accuracy of your notes. If you do this faithfully during the note-taking phase, you will have no trouble completing the "works cited" section of your paper later on.

Keep a research log. As you search databases and consult reference books, keep track of what search terms and databases you used and the call numbers and URLs of information sources. This will help if you need to refine your research strategy, locate a source a second time, or show your professor what works you consulted in the process of completing the project.

You must cite direct quotes.

You must cite paraphrases. Paraphrasing is rewriting a passage or block of text in your own words. If you paraphrase, you must still cite the original source of the idea.

You must cite ideas given to you in a conversation, in correspondence, or over email.

You must cite sayings or quotations that are not familiar, or facts that are not "common knowledge." However, it is not necessary to cite a source if you are repeating a well known quote or familiar proverb. Common knowledge is something that is widely known. For example, it is widely known that Bill Clinton served two terms as president; it would not be necessary to cite a source for this fact.

These types of sources should be cited as well. Printed sources: Books, parts of books, magazine or journal articles, newspaper articles, letters, diaries, public or private documents; *Electronic sources:* Web pages, articles from e-journals, newsgroup postings, graphics, email messages, software, databases; *Images:* Works of art, illustrations, cartoons, tables, charts, graphs; *Recorded or spoken material:* Course lectures, films, videos, TV or radio broadcasts, interviews, public speeches, conversations.