EPPS6302 Methods of Data Collection and Production

Fall 2020 Thursday: 4:00 pm - 6:45 pm Remote/Virtual Learning
Team link

Instructor: Dr. Karl Ho; 972-883-2017; kyho@utdallas.edu
Online Office Hours: On Microsoft Teams
Office Hours: Tuesday 2 pm – 4 pm, Wednesday, Thursday 10 am – 12 pm, by appointment.

Overview: This is a data method course introducing the theory, practices and issues of data collection and production. It aims at providing a comprehensive framework in understanding data, and how social scientists conduct research starting from the *data generation process*. We will cover data methods, data management, big data trends and how to prepare data for next phrases of research including modeling and reporting. This survey course is designed to equip social scientists with data generation concepts, tools and best practices. Topics on new developments and tools of big data 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 data methods by tasks and modes
3. Develop instruments and schemes for data collection

Course Modality and Expectations

Instructional Mode	**Remote/Virtual Learning**
Course Platform	**This course will meet in Microsoft Teams at the time period list in syllabus. Click on the Teams link to access the class at class time.**
Expectations	**Please see expectations below.**
Asynchronous Learning Guidelines	**Resources are provided for asynchronous students. All materials are available for access in Teams General channel under Files. Recorded videos will be provided via Microsoft Stream service. Both synchronous and asynchronous students are to fulfill all course requirements including exercises, proposal and final project presentations. For asynchronous students who cannot present in person, please submit recorded presentations according to guidelines provided by the instructor including presentation medium format and time limit.**

Required Text:
Foster, Ian, Rayid Ghani, Ron S. Jarmin, Frauke Kreuter, and Julia Lane (editors). 2016. Big Data And Social Science: A Practical Guide to Methods and Tools, Chapman and Hall/CRC Press. ‡
Groves, R. M., et al. 2009. *Survey Methodology* Second Edition. Wiley-Interscience. ‡
Guest, G., Namey, E.E. and Mitchell, M.L., 2012. *Collecting qualitative data: A field manual for applied research*. Sage. ‡
Krueger, R.A. and Casey, M.A., 2014. *Focus groups: A practical guide for applied research*. Sage publications. †

(‡) Available electronically at McDermott Library
(†) Older version available electronically at McDermott Library

Recommended Reading:
Ansolabehere, Stephen and Douglas Rivers. 2013. "Cooperative Survey Research." *Annual Review of Political Science* 16: 307-329.
Biemer, Paul P. Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly*, 74(5):817–848, 2010.
Biemer, Paul P. *Latent Class Analysis of Survey Error*. John Wiley & Sons, 2011.
Biemer, Paul P. and Lars E. Lyberg. *Introduction to Survey Quality*. John Wiley & Sons, 2003.

Breiman, Leo. 2001. "Statistical modeling: The two cultures (with comments and a rejoinder by the author)." *Statistical science* 16, no. 3 199-231.
Fowler, Jr., Floyd J. 2014. *Survey Research Methods*, 5th edition. Los Angeles, CA: Sage Publications. ISBN paperback 978-1-4522-5900-0.
Groves, Robert M. *Survey Errors and Survey Costs*. John Wiley & Sons, 2004.
Hansen, Morris H., William N. Hurwitz, and William G. Madow. *Sample Survey Methods and Theory*. John Wiley & Sons, 1993.
Kitchin, Rob. 2014. *The Data Revolution: Big Data, Open Data, Data Infrastructures & their Consequences*. Sage Publications.
Rudis, Bob. 2018. *21 Recipes for Mining Twitter Data with rtweet* (https://rud.is/books/21-recipes/)
Saltz, J.S. and Stanton, J.M., 2017. *An introduction to data science*. SAGE Publications.
Scheuren, Fritz and William E. Winkler. 1993. Regression analysis of data files that are computer matched. *Survey Methodology*, 19(1):39–58.
Valliant, Richard, Jill A Dever, and Frauke Kreuter. *Practical Tools for Designing and Weighting Survey Samples*. Springer, 2013.
Weisberg, Herbert, *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: University of Chicago Press, 2005

APSA Style Manual for Political Science Revised 2018 edition:
https://mk0apsaconnectbvy6p6.kinstacdn.com/wp-content/uploads/sites/43/2018/11/Style-Manual-for-Political-Science-2018.pdf

Expectations:

Camera-readiness: This online course requires students to attend with camera on. Camera-readiness is critical for future career in academia and industry alike. Data scientists must excel in communicating data projects and effective face-to-face presentation is very important part of data science training. Learn how to best learn and work remotely from this free Udemy course https://www.udemy.com/course/quick-guide-to-working-remotely-from-vp-of-learning/.

BYOD: Students are expected to have own functional, up-to-date computer equipped with features/hardware ready for the Remote/Virtual Learning mode including web camera, speaker and microphone. Personal computer (not mobile device) running MacOS, Linux or Windows operating systems is recommended.

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. Participation is responsible for 10% of the final grade.

Grading and Requirements:
Data Project:
Each student must design and implement a data project to product and/or collect original data using the ***systems and techniques covered in this course***. It will be fully implemented within the semester with a final report and an in-class presentation. The final report must be between 15-20 pages in length. Originality is the first quality the instructor demands. Please read carefully the University's policy about cheating and plagiarism (see below for a brief version). Due dates of the proposal and final reports are listed on the schedule below. No late submission will be accepted. Proposal of the project must be consulted with and subject to the instructor approval. It constitutes 30% of the final grade. The assignments account for 30% of final grade. The final report, with in-class presentation, is responsible for 30 percent of the final grade. In summary, the grade structure is as follows:

Participation:	10%
Data Project Proposal	30%
Assignments	30%
Project and presentation:	30%

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

- Must be typed or word-processed double-spaced on letter size papers
- Document soft copy format is required to be in PDF.
- Required image format on document is PNG.

- Required image format on web document is SVG or PNG.
- Use a cover page with topic and name

Tentative Topics (not necessarily weekly):

Week Topic

1 Introduction: Data Theory and Methods

2 Survey Modes and Design

Reading:

Groves 1, 2.

3 Questionnaire Design

Reading:

Groves 7, 8, 9

4 Sampling, Survey Errors and Weighting

Reading:

Groves, Chapters 3, 4, 5, 6, 10

5 Experiments

Reading:

Druckman, J.N., Green, D.P., Kuklinski, J.H. and Lupia, A., 2006. The growth and development of experimental research in political science. *American Political Science Review*, 100(4), pp.627-635.

6 Qualitative Data

Reading:

Guest, G., Namey, E.E. and Mitchell, M.L., 2012. Collecting qualitative data: A field manual for applied research. Sage.

7 Focus Group Research

Reading:

Krueger, R.A. and Casey, M.A., 2014. *Focus groups: A practical guide for applied research*. Sage publications.

Proposal due (October 1st, 2020)

8 Special lecture: US election data and forecast

TBA

9 Spatial data: Guest speaker

Reading:

Ballas, Dimitris et al., 2018. GIS and the social sciences : theory and applications , New York: Routledge. (‡)

Parker, R.N. and Asencio, E.K., 2009. *GIS and spatial analysis for the social sciences: Coding, mapping, and* modeling. Routledge.

10 Complex data

Reading:

Foster et al. Chapter 1, 3, 4

11 Text data

Reading:

Foster et al. Chapter 7, S&S Chapter 14,15

12 Web data

Reading:

Foster et al. Chapter 2

13 Social Media data

Reading: Bob Rudis. 2018. *21 Recipes for Mining Twitter Data with rtweet* (https://rud.is/books/21-recipes/)

14 Consultation week: individual meetings with instructor

15 Thanksgiving

16 Presentations (final project due)

COVID-19 Guidelines and Resources

The information contained in the following link lists the University's COVID-19 resources for students and instructors of record. Please see http://go.utdallas.edu/syllabus-policies.

Class Recordings

Students are expected to follow appropriate University policies and maintain the security of passwords used to access recorded lectures. Unless the Office of Student AccessAbility has approved the student to record the instruction, students are expressly prohibited from recording any part of this course. Recordings may not be published, reproduced, or shared with those not in the class, or uploaded to other online environments except to implement an approved Office of Student AccessAbility accommodation. Failure to comply with these University requirements is a violation of the Student Code of Conduct.

The instructor will record meetings of this course. Any recordings will be available to all students registered for this class as they are intended to supplement the classroom experience. Students are expected to follow appropriate University policies and maintain the security of passwords used to access recorded lectures. Unless the Office of Student AccessAbility has approved the student to record the instruction, students are expressly prohibited from recording any part of this course. Recordings may not be published, reproduced, or shared with those not in the class, or uploaded to other online environments except to implement an approved Office of Student AccessAbility accommodation. If the instructor or a UTD school/department/office plans any other uses for the recordings, consent of the students identifiable in the recordings is required prior to such use unless an exception is allowed by law. Failure to comply with these University requirements is a violation of the Student Code of Conduct.

Technical Requirements
In addition to a confident level of computer and Internet literacy, certain minimum technical requirements must be met to enable a successful learning experience. Please review the important technical requirements on the Getting Started with eLearning webpage.

Course Access and Navigation
This course can be accessed using your UT Dallas NetID account on the eLearning website.

Please see the course access and navigation section of the Getting Started with eLearning webpage for more information.

To become familiar with the eLearning tool, please see the Student eLearning Tutorials webpage.

UT Dallas provides eLearning technical support 24 hours a day, 7 days a week. The eLearning Support Center includes a toll-free telephone number for immediate assistance (1-866-588-3192), email request service, and an online chat service.

Communication
This course utilizes online tools for interaction and communication. Some external communication tools such as regular email and a web conferencing tool may also be used during the semester. For more details, please visit the Student eLearning Tutorials webpage for video demonstrations on eLearning tools.

Student emails and discussion board messages will be answered within 3 working days under normal circumstances.

Distance Learning Student Resources
Online students have access to resources including the McDermott Library, Academic Advising, The Office of Student AccessAbility, and many others. Please see the eLearning Current Students webpage for more information.

Server Unavailability or Other Technical Difficulties
The University is committed to providing a reliable learning management system to all users. However, in the event of any unexpected server outage or any unusual technical difficulty which prevents students from completing a time sensitive assessment activity, the instructor will provide an appropriate accommodation based on the situation. Students should

immediately report any problems to the instructor and also contact the online eLearning Help Desk. The instructor and the eLearning Help Desk will work with the student to resolve any issues at the earliest possible time.

Comet Creed
This creed was voted on by the UT Dallas student body in 2014. It is a standard that Comets choose to live by and encourage others to do the same:

"As a Comet, I pledge honesty, integrity, and service in all that I do."

Academic Support Resources
The information contained in the following link lists the University's academic support resources for all students.

Please go to Academic Support Resources webpage for these policies.

UT Dallas Syllabus Policies and Procedures
The information contained in the following link constitutes the University's policies and procedures segment of the course syllabus.

Please go to UT Dallas Syllabus Policies webpage for these policies.

The descriptions and timelines contained in this syllabus are subject to change at the discretion of the Professor.