Fundamentals of Survey and Data Science  
SURV 400  
3 credits/6 ECTS  
Spring 2019

Instructor  
Carolina Casas-Cordero, ccasascordero@uc.cl

Video lecture by Florian Keusch, Jennifer Sinibaldi

Short Course Description
The course introduces the student to a set of principles of survey and data science that are the basis of standard practices in these fields. The course exposes the student to key terminology and concepts of collecting and analyzing data from surveys and other data sources to gain insights and to test hypotheses about the nature of human and social behavior and interaction. It will also present a framework that will allow the student to evaluate the influence of different error sources on the quality of data.

Course and Learning Objectives
By the end of the course, students will...
- be able to apply the key terminology used by survey methodologists and data scientists.
- be able to assess the quality of data from different sources based on a data quality framework.
- be able to select an appropriate data source to answer different types of research questions.
- understand the influence of coverage, sampling, and nonresponse on data quality and know how to deal with deficiencies of the data.
- have a clear understanding of the steps involved in data preparation, data processing, data analysis, and data visualization.
- be able to comply with ethical standards in survey research and data science.

Prerequisites
Students are expected to be familiar with basic statistical concepts, such as mean, standard deviation, variance, and distributions (at the level of an undergraduate course), and have exposure to elements of social science perspectives on human behavior.
Class Structure and Course Concept
This is an online course using a flipped classroom design. It covers the same material and content as an on-site course but runs differently. In this course, you are responsible for watching video recorded lectures and reading the required literature for each unit and then “attending” mandatory weekly one-hour online meetings where students have the chance to discuss the materials from a unit with the instructor. Just like in an on-site course, assignment will be assigned and graded and there will be a final exam at the end of the course.

Although this is an online course where students have more freedom in when they engage with the course materials, students are expected to spend the same amount of time overall on all activities in the course – including preparatory activities (readings, studying), in-class-activities (watching videos, participating in online meetings), and follow-up activities (working on assignments and exams) – as in an on-site course. As a rule of thumb, for each credit offered by a course, students can expect to spend one hour per week on in-class activities and three hours per week on out-of-class activities over the span of a full 12-week term. This is a 3-credit course that runs for 12 weeks. Hence, the average workload is about 12 hours per week.

Mandatory Weekly Online Meetings
Monday, 1:00 PM – 1:50 PM EST/7:00 – 7:50 PM CET, starting March 4

Meetings will be held online through Zoom. Follow the link to the meeting sessions on the course website on https://www.elms.umd.edu/. If video participation via Internet is not possible, arrangements can be made for students to dial in and join the meetings via telephone.

In preparation for the weekly online meetings, students are expected to watch the lecture videos and read the assigned literature before the start of the meeting. In addition, students are encouraged to post questions about the material covered in the videos and readings of the week in the forum before the meetings (deadline for posting questions is Saturday, 1:00 PM EST/7:00 CET).

Daylight saving time begins in the US on March 10, 2019 and clocks are turned forward 1 hour. Daylight saving time begins in Europe on March 31, 2019. Therefore, look carefully at the times of meetings and deadlines! If in doubt, please consider the US Easter Time (e.g. in Washington DC) is the OFFICIAL time for all meetings and deadlines.

Students have the opportunity to use the Zoom meeting room set up for this course to connect with peers outside the scheduled weekly online meetings (e.g. for study groups). Students are encouraged to post the times that they will be using the room
to the course website forum to avoid scheduling conflicts. Students are not required to use Zoom and can of course use other online meeting platforms such as Google Hangout or Skype.

**Grading**

Grading will be based on:

- Participation in discussion during the weekly online meetings and contributions to the forum (deadline is Saturday, 1:00 PM EST/7:00 CET) demonstrating understanding of the required readings and video lectures is worth 10% of the final grade.
- Nine online quizzes, worth 100 points each, reviewing specific aspects of the material covered. The simple average of the points across all quizzes is worth 60% of the final grade.
- A final open-book online exam, worth 100 points, is worth 30% of the final grade.

Due dates for the online quizzes and assignments are indicated in the Class Schedule section. Quizzes and assignments will automatically close on the course website at the indicated time. Assignments can be completed at any time during the week leading up to the deadline. Therefore, extensions will be granted sparingly with penalty and are at the instructor’s discretion.

The final exam is open book and open notes, and students are on the honor system (see Academic Conduct below). Students will have 1 week to complete the exam online, within the course website. The exam will be open from Tuesday, May 28, 2019 at 6:00 PM EDT/00:00 CEST until Tuesday, June 4, 2019 at 5:59 PM EDT/23:59 CEST, at which point it will close. All students must submit their exams until June 4, 2019 at 5:59 PM EDT/23:59 CEST.

**Technical Equipment Needs**

The learning experience in this course will mainly rely on the online interaction between students and the instructor during the weekly online meetings. Therefore, we encourage all students in this course to use a web camera and a headset. Decent quality headsets and web cams are available for less than $20 each. We ask students to refrain from using built-in web cams and speakers on their desktops or laptops. We know from our experience in previous online courses that this will reduce the quality of video and audio transmission and therefore will decrease the overall learning experience for all students in the course. In addition, we suggest that students use a wire connection (LAN), if available, when connecting to the online meetings. Wireless connections (WLAN) are usually less stable and might be dropped.
**Long Course Description**

The fields of survey methodology and data science draws on theories and practices developed in several academic disciplines – mathematics, statistics, psychology, sociology, computer science, and economics. To become an accomplished professional in these fields requires a mastery of research literatures as well as experience designing, conducting, and analyzing surveys and data from other sources, such as administrative records, social media, or transactions.

This course introduces the student to a set of principles of survey design and data science that are the basis of standard practices in these fields. The course exposes the student to research literatures that use both observational and experimental methods to test key hypotheses about the nature of human behavior and other factors that affect the quality of data. It will also present important statistical concepts and techniques in sample design, execution, and estimation, as well as models of behavior describing errors in responding to survey questions. Thus, both social science and statistical concepts will be presented.

The course uses the concept of total error as a framework to discuss coverage properties of sampling frames and organic data, alternative sample designs and their impacts on standard errors of statistics, different modes of data collection and generation, the role of interviewers and respondents in surveys, impacts of nonresponse and missing data on statistics, measurement errors in data, data processing, and data/research ethics.

The course is intended as an introduction to the fields of survey methodology and data science, taught at a graduate level. Lectures and course readings assume that students understand basic statistical concepts (at the level of an undergraduate course) and have exposure to elements of social science perspectives on human behavior. For those lacking such a background, supplementary readings are recommended.

**Readings**

Primary readings will be from the following volumes:


Additional required and recommended readings will be made available on the course website: [https://www.elms.umd.edu/](https://www.elms.umd.edu/) None of the information in the recommended readings will be included on the online quizzes or the final exam.

**Academic Conduct**

Clear definitions of the forms of academic misconduct, including cheating and plagiarism, as well as information about disciplinary sanctions for academic misconduct may be found at

[https://www.president.umd.edu/sites/president.umd.edu/files/documents/policies/III-100A.pdf](https://www.president.umd.edu/sites/president.umd.edu/files/documents/policies/III-100A.pdf) (University of Maryland) and


Knowledge of these rules is the responsibility of the student and ignorance of them does not excuse misconduct. The student is expected to be familiar with these guidelines before submitting any written work or taking any exams in this course. Lack of familiarity with these rules in no way constitutes an excuse for acts of misconduct. Charges of plagiarism and other forms of academic misconduct will be dealt with very seriously and may result in oral or written reprimands, a lower or failing grade on the assignment, a lower or failing grade for the course, suspension, and/or, in some cases, expulsion from the university.

**Accommodations for Students with Disabilities**

In order to receive services, students at the University of Maryland must contact the Disability Support Services (DSS) office to register in person for services. Please call the office to set up an appointment to register with a DSS counselor. Contact the DSS office at 301.314.7682; [http://www.counseling.umd.edu/DSS/](http://www.counseling.umd.edu/DSS/).

Students at the University of Mannheim should contact the Commissioner and Counsellor for Disabled Students and Students with Chronic Illnesses at [http://www.uni-mannheim.de/studienbueros/english/counselling/disabled_persons_and_persons_with_chronic_illnesses/](http://www.uni-mannheim.de/studienbueros/english/counselling/disabled_persons_and_persons_with_chronic_illnesses/).

**Course Evaluation**

In an effort to improve the learning experience for students in our online courses, students will be invited to participate in an online course evaluation at the end of the course (in addition to the standard university evaluation survey). Participation is entirely voluntary and highly appreciated.
Class Schedule

Please note that assignments and dates are subject to change. Information (e.g., articles and assignments) posted to the course website supersedes the information noted here.

Unit 1: Introduction – How to do survey research and data science

Video lecture: available Monday, February 25, 2019

Online meeting: Monday, March 4, 2019, 1:00–1:50 PM EST/7:00–7:50 PM CET

Online quiz unit 1: due Monday, March 4, 2019, 6:00 PM EST/midnight CET

Readings:

Groves et al. (2009). Chapters 1.4 and 1.5

Peng & Matsui (2015). Chapters 1-3


Recommended:


Daylight saving time begins in the US on March 10, 2019 and clocks are turned forward 1 hour. Daylight saving time begins in Europe on March 31, 2019. Therefore, look carefully at the times of meetings and deadlines!

If in doubt, please consider the US Easter Time (e.g. in Washington DC) is the OFFICIAL time for all meetings and deadlines.

Unit 2: Quality of Data
Video lecture: available Monday, March 4, 2019

Online meeting: Monday, March 11, 2019, 1:00–1:50 PM EDT / 6:00–6:50 PM CET

Online quiz unit 2: due Monday, March 11, 2019, 6:00 PM EDT / 11:00 PM CET

Readings:
Groves et al. (2009). Chapter 2

Recommended:


Unit 3: Coverage

Video lecture: available Monday, March 11, 2019

Online meeting: Monday, March 18, 2019, 1:00–1:50 PM EDT / 6:00–6:50 PM CET

Online quiz unit 3: due Monday, March 18, 2019, 6:00 PM EDT / 11:00 PM CET

Readings:
Groves et al. (2009). Chapter 3


Unit 4: Modes of Survey Data Collection
Video lecture: available Monday, March 18, 2019

Online meeting: Monday, March 25, 2019, 1:00–1:50 PM EDT/6:00–6:50 PM CET

Online quiz unit 4: due Monday, March 25, 2019, 6:00 PM EDT/11:00 PM CET

Readings:
Groves et al. (2009). Chapters 1.3 and 5

Unit 5: Data Generation from Other Sources
Video lecture: available Monday, March 25, 2019

Online meeting: Monday, April 1, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Readings:


Recommended:


Unit 6: Sampling I
Video lecture: available Monday, April 1, 2019

Online meeting: Monday, April 8, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Online quiz unit 6: due Monday, April 8, 2019, 6:00 PM EDT/midnight CEST
Readings:
Groves et al. (2009). Chapters 4.1-4.6

Recommended:

**Unit 7: Sampling II**
Video lecture: available Monday, April 8, 2019

Online meeting: Monday, April 15, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Online quiz unit 7: due Monday, April 15, 2019, 6:00 PM EDT/midnight CEST

Readings:

Recommended:

**Unit 8: Questionnaires and Interviewing**
Video lecture: available Monday, April 15, 2019

Online meeting: Monday, April 22, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Online quiz unit 8: due April 22, 2019, 2018, 6:00 PM EDT/midnight CEST

Readings:

Recommended:
Unit 9: Nonresponse
Video lecture: available Monday, April 22, 2019

Online meeting: Monday, April 29, 2019, 1:00 – 1:50 PM EDT / 7:00 – 7:50 PM CEST

Online quiz unit 9: due Monday, April 29, 2019, 6:00 PM EDT/midnight CEST

Readings:
Groves et al. (2009). Chapters 6 and 10.6

Recommended:


Unit 10: Data Preparation, Data Processing, and Data Base Management
Video lecture: available Monday, April 29, 2019

Online meeting: Monday, May 6, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Online quiz unit 10: due Monday, May 6, 2019, 6:00 PM EDT/midnight CEST

Readings:
Groves et al. (2009). Chapters 10.1-10.5


One-week break, no online meeting on May 13, 2019!

Unit 11: Data Analysis and Data Visualization
Video lecture: available Monday, May 6, 2019

Online meeting: Monday, May 20, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Readings:
Peng & Matsui (2015). Chapters 4, 5, 7, 8, and 10


**Unit 12: Survey and Research Ethics**

Video lecture: available Monday, May 20, 2019

Online meeting: Monday, May 27, 2019, 1:00–1:50 PM EDT/7:00–7:50 PM CEST

Readings:

- Groves et al. (2009). Chapter 11

**Final Exam**

Due: Tuesday, June 4, 2019, 5:59 PM EDT/23:59 PM CEST

**Note:** Student access to the course website will be revoked two weeks after the final exam.