

COMM 6701: Working with Communication Data
Tuesday and Thursday 5:30 - 8:10 (7 weeks)
Journalism 224

Instructor: Robert Bond, Derby Hall 3072, bond.136@osu.edu

Office Hours: Monday 11-12 & Thursday 3-4 and by appointment.

Course Description

This class is intended to provide a bridge from your in-depth course work in Statistics and Methods to the process of data analysis for your own research. At times, this course will intentionally overlap with content covered in other courses. This course is focused on building a careful, thoughtful process of approaching research. Through this course, you will build on your existing data analytic and research skills to ensure that you have a solid foundation as you begin to work independently on data. At its best, this course will serve as a foundation of these skills for your graduate and subsequent professional career.

Objectives:

1. To learn to prepare data sets for analysis, including basics of secondary data analysis (data that others have collected).
2. To develop good professional practice with respect to documenting data, recoding, and statistical analyses, to support working with your advisor and other collaborators, and to ensure that you can retrieve and re-analyze data in the process of revision long after the data was originally analyzed.
3. To apply the principles learned in your research methods courses for operationalizing constructs.
4. To practice writing up the results of such efforts clearly, concisely, and professionally.
5. To increase facility in working with software for statistical analysis.
6. To have an initial experience of working with unstructured data (text as data, social media data, and so on).

To meet these objectives, we will have some readings (available via the library links provided on Carmen), and from a course text, but much of the course is “hands-on” and will involve working with data in class. By the end of the course you will be prepared to analyze your own data carefully and thoughtfully.

Science is a social activity. As such, the format of this course will be largely social. There will be a very large emphasis on working with, soliciting feedback from, and providing feedback to your peers.

Course Materials

- **Textbook:** Pallant, J. (2013). SPSS survival manual, 5th ed. NY: McGraw-Hill. (aka SM). Various required articles and book chapters as assigned.
- Additional, supplementary readings and PDFs distributed by CARMEN, as needed.
- SPSS for Windows or Mac.
- R for Windows or Mac.

Evaluation

Short Assignments (40% of the final grade)

In each of weeks 2 – 6 you will be assigned an exercise through which you will have the opportunity to practice the techniques we have gone over in class. These will be done as group projects and will receive a group grade (all group members will receive the same grade). I invite anonymous feedback from everyone about group members, so anyone not contributing fully to the group may be penalized. However, note that members of the class may come with various levels of skills and preparation. As such, if you are a relatively strong student in one particular area look at the group work as not only an opportunity to practice for yourself, but also as an opportunity to practice teaching others the skills you have. Conversely, if you feel that a concept is difficult or doesn't make sense, it is incumbent upon you to ask for assistance and extra explanation from the instructor, your peers, or both.

Class Participation/Attendance (20% of the final grade)

Each class will be focused on a discussion of the readings and working through various examples. The instructor will provide a basic overview of the material; students will make comments or ask questions about the material. The class participation grade will be an assessment of your contribution to the class as well as of your attendance.

A Final Project (40% of the final grade)

Students will complete an individual research project for the final assignment in the class (week 7). This research project serves the function of a final exam. I will provide a data set from which you can complete the project. The project will be a version of part of the methods portion of a research article (APA style) that describes the measurement and application of the data to a research question. You will need to identify and create at least two measures of some psychosocial variable, plus at least one variable to use as a criterion variable. The measurement of variables studied should be explained using examples from the items used to create the variable, any information regarding the properties of the variables (particularly if non-normal), any transformations conducted, descriptions of the dimensionality (mostly due to any factor analysis conducted), means and standard deviations of the variables, Cronbach's alpha calculations, and so on. You should provide a table showing the items used and their factor loadings on each dimension. You should also provide some further analysis describing the results of a regression analysis in which your measure is used as an independent variable for at least one outcome variable. This regression may be bivariate, or can be multivariate. Conclude the research project with at least one paragraph of discussion about the measure you created, the strength of weakness of the measurement, and a note about how measurement of this construct could be improved. Clarity of explanation and the appropriate use of the techniques from

the class are priorities. Unlike the weekly exercises, the final project may not involve any other students or faculty members in any way. This assignment is to be completed independently.

Derivation of Final Grade

My grading system is largely a percentage based system where:

93%+ = A, 90% - 92.9% = A-

87% - 89.9% = B+, 83% - 86.9%+ = B, 80% - 82.9% = B-

77% - 79.9% = C+, 73% - 76.9%+ = C, 70% - 72.9% = C-

67% - 69.9% = D+, 63% - 66.9%+ = D, 60% - 62.9% = D-

less than 60% = E

I reserve the right to modify this system *downward* depending on the distribution of grades. In other words, if only one student exceeds the 90% threshold, but five hit 89%, I may choose to move the cutoff for an A- to 89%.

Course Policies and Miscellaneous

Late or Absent Assignments and Missed Exams

Unless otherwise notified, assignments are due by the date due at the specified time.

An assignment will not be accepted more than 24 hours after the due date. The only exceptions to these rules are totally unforeseen circumstances that are convincingly documented no later than 24 hours after the due date.

Attendance

There is no formal attendance policy for this course. However, you should always attend class. If I believe attendance is slipping, I reserve the right to create an attendance policy. Not attending class will make learning the course material difficult.

COMM 6701 on CARMEN

I will upload data files, powerpoint slides, PDFs of extra readings, and other course-relevant material to CARMEN. Learn to use CARMEN, as you will be asked to use it during class to retrieve materials used in class that day.

Academic Misconduct

All students at Ohio State University are bound by the Code of Student Conduct (see http://studentaffairs.osu.edu/resource_csc.asp). Violations of the Code in this class, especially pertaining to 3335-23-04 Section A on Academic Misconduct, will be aggressively prosecuted through the procedures the university has set up to deal with violations of the Code. If I believe you have violated the student Code, your case will be referred to the Committee on Academic Misconduct (see <http://oaa.osu.edu/coam/html>). **Not following the rules of the course as outlined in this syllabus or provided orally is considered a violation of the Code of Student Conduct. Penalties for academic misconduct from a graduate student are especially stiff and are almost certain to include failure in this course and suspension from the university, even for a first offense. Graduate students found in violation of the Code can expect revocation of funding from the School of Communication and, potentially, expulsion from the graduate program.** Repeat offenses and especially egregious violations of the

Code can result in expulsion from the University, even on the first offense. Make sure that you are familiar with the Code of Student Conduct, and familiarize yourself with “Ten Suggestions for Preserving Academic Integrity” available online at <http://oaa.osu.edu/coamstensuggestions.html>.

I expect students who believe a classmate has violated this policy to come forth to me so the alleged violation can be investigated and appropriate action can be taken if needed. If possible, your identity will be protected. **You can be found in violation of the Code of Student Conduct for assisting others to violate the Code, including by turning a blind eye when you see it happening.** “Cheating” in graduate school simply cannot and will not be tolerated, and the consequences for doing so are severe.

Tentative Nature of the Syllabus

This syllabus represents a contract in the works. Events that transpire over the semester may require me to modify the administration of this course and therefore the syllabus. In the event I need to modify the syllabus, I will announce the modification in class and on CARMEN. Ultimately, it is your responsibility to keep up with any such modifications and be aware of current policies, deadlines, etc.

School of Communication Diversity Statement

The School of Communication at The Ohio State University embraces and maintains an environment that respects diverse traditions, heritages, experiences, and people. Our commitment to diversity moves beyond mere tolerance to recognizing, understanding, and welcoming the contributions of diverse groups and the value group members possess as individuals. In our School, the faculty, students, and staff are dedicated to building a tradition of diversity with principles of equal opportunity, personal respect, and the intellectual interests of those who comprise diverse cultures.

Students with Special Needs

Students with disabilities that have been certified by the Office for Disability Services will be appropriately accommodated and should inform the instructor as soon as possible of their needs. The Office for Disability Services is located in 150 Pomerene Hall, 1760 Neil Avenue; telephone 292-3307, TDD 292-0901; <http://www.ods.ohio-state.edu/>.

SPSS & R

SPSS and R are statistical analysis and data manipulation packages widely used throughout the scientific community. These are but two of many statistical software programs you should become familiar with during your days in graduate school. However, because SPSS is the one program most faculty and graduate students in this program are familiar with, we will use it. R, on the other hand, is quickly becoming a standard package for use in many disciplines both in and outside of the social sciences. Familiarity with R will equip you to both better understand the similarity of the various packages and also enable you to take courses in cognate departments that emphasize the package.

R is freely downloadable from <https://www.r-project.org/>. Students should download a copy as

soon as possible. R is available on lab computers. Many online resources are available to help students become more familiar with the computing and coding environment.

SPSS is available on all graduate student office computers in the School of Communication in Derby Hall as well as throughout public computing labs operated by OSU's Office of Information Technology. It also can be licensed for use on your personal machine by acquiring a copy and installation codes through the OSU Office of Information Technology. For most students in this class, there is no charge for the license. SPSS can be downloaded and installed from OSU, but it will not work without first acquiring the license codes. For details, see

<https://ocio.osu.edu/software/directory/slwin/#spss> for the Windows version

and

<https://ocio.osu.edu/software/directory/slmac/#spss1103> for the Mac version.

You can also contact the School of Communication's technical support personnel in Derby Hall, for assistance.

Schedule of Lectures and Readings

Note: scheduling, due dates, and details of activities and exercises subject to change/modification.

- Based on previous class surveys and informal email surveys of faculty, the focus of this class is on learning how to approach data analysis systematically and preparing data for more formal hypothesis testing. This includes loading and cleaning data, recoding variables, univariate analysis, factor analysis, and assessing the reliability and validity of measures. This process will be the primary focus of the class. As we do so, we will focus on proper documentation of this process.
- The course takes a “hands-on” approach. We will develop familiarity with various data sets in order to ensure that your skills translate from application to application.
- We also will look at issues regarding data structure, data merges, etc.
- Previous students have expressed interest in learning structural equation modeling. This is a course unto itself (perhaps one you should encourage the School to offer!), but as time permits I will demonstrate and discuss some of the basics of this.
- Students have expressed an interest in content analysis coding and reliability. We will include a section on this.
- Because of the differences in preparation and comfort with data analysis and software, I'm not sure how long each section will take. The schedule below is tentative and we will have to flex as necessary. It is your responsibility to be aware of changes announced in class and posted on Carmen and/or communicated via email.
- As time permits, I will demonstrate or discuss some more “advanced” topics related to data analysis. However, note that the main focus of this course is on becoming comfortable with basic concepts, so when needed we will spend more time on this rather than new material.

Week 1 – 1/7 & 1/9 – Introduction and basic data entry and manipulation

- Explanation of class project
- Class introduction
- Discuss types of data entry – Qualtrics or MediaLab vs paper and pencil, how to do data entry
- Data cleaning: getting to know your data set, labelling, missing values, distributions
- Recoding/reversing variables, documenting recodes

Activities: Divide into groups, perhaps by interest (health, political, new media, etc.). In class will check distributions, reverse variables using recode and compute (SPSS), and various functions in R, practice labeling syntax/code.

Readings:

1. SM: Chapters 1-5. [Note: each week, just skim the Survival Manual so you have a sense of what is in there and know where to look when you need to – even for experienced SPSS users, you will no doubt find a few features and tricks you didn't know.]
2. Van den Broeck J, Argeseanu Cunningham S, Eeckels R, Herbst K (2005) Data cleaning: Detecting, diagnosing, and editing data abnormalities. PLoS Med 2(10): e267.

Week 2 – 1/14 & 1/16 – Working with archival/industry data

- Working with archival data & large, national survey data
- Further practice at creating and assessing measures, using larger scale data

Exercise: Practice working with sample archival data set. I will provide class time to work on this. **Due Sunday midnight.**

Week 3 – 1/21 & 1/23 – Factor analyses

- Factor analyses
- Identifying possible dimensions
- Assessing reliability (internal consistency)

Exercise: Each group will get a set of items to factor analyze. Provide a cover memo summarizing whether orthogonal or oblique rotation was more appropriate, provide the tables with factor loadings, describe and name the dimensions, reverse items as needed and assess reliability (if we've gotten that far). I will try to save as much time as possible to work together in class on this on Thursday. **Due Sunday midnight.**

Readings:

1. SM: Chapter 6,7,8 (to page 90), chapter 15, chapter 9.
2. Fabrigar, L.R., et al. (1999). Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods, 272-299.

Week 4 – 1/28 & 1/30 – Creating and assessing indices

- Creating indices
- Univariate analyses of indices, checking for outliers, etc.

- Try data transformations
- Save data set with reversed items, new indices, merge data sets
- If time permits, I will demo confirmatory factor analysis this week

Exercise: Each group will provide a cover memo with the Cronbach alphas for the measure and for each dimension, and if alpha was improved by dropping any items. Talk about transformations. Provide properly documented and annotated syntax and output. Due Sunday midnight. I will reserve as much time as I can to work on this together in class on Thursday. **Due Sunday midnight.**

Readings:

1. SM: chapter 8 (p 96 to end)
2. Review Crano & Brewer ch. 15 on rating scale construction.
3. Review Hayes chapter 4.

Week 5 – 2/4 & 2/6 – Assessing validity

- Types of validity
- Establishing validity
- Tradeoffs between reliability and validity

Exercise: Use the criterion measures your group included in the data collection from the Methods class for at minimum a test of predictive validity for your measure including its various dimensions if any; preferably, you will have measures to help you assess convergent/discriminant validity as well. Conduct appropriate statistical analyses, report results, include properly annotated syntax/output. **Due Sunday midnight.**

Readings:

1. Review chapter 3 Crano & Brewer on reliability and validity.
2. Chapter on validity from Kerlinger & Lee (2000), Foundations of behavioral research.

Week 6 – 2/11 & 2/13 – Coding communication content

- What can be coded
- Developing definitions and rules for coding
- Conducting training
- Assessing intercoder reliability
- Assessing coded validity
- Automated coding procedures and reliability

Exercise: You and a partner will code an open-ended item from the class data collection. This will involve deciding what you will code for (I suggest you keep it simple for this class exercise!), developing definitions, trying to get reliable coding, running Krippendorff's alpha, working together to refine the definitions to improve reliability, running Krippendorff's alpha again. Hand in a report with the coding scheme, definitions, the alpha, the data matrix for testing the alpha, and the syntax, all appropriately documented. Obtaining adequate reliability in this short time frame is

not required (for a more ambitious code, it is not expected in this short time window – this is just for practice/learning purposes). **Due Sunday midnight.**

Readings:

1. Review Crano & Brewer chapter 13 on content analysis
2. Hayes, A.F. & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77-89.
3. Slater, M.D. (2014) Conducting a Content Analysis as a Foundation for Programmatic Research. *Communication Methods and Measures*.
4. Recommended for further reading: Krippendorff, K. (2013). *Content analysis: An introduction to its methodology*. Thousand Oaks, CA: Sage.

Week 7 – 2/18 & 2/20 – Mediation/moderation and final project.

- Lecture/discussion on mediation and moderation.
- Demo and in-class exercise on mediation testing.
- Discussion of final project.
- Writing the Methods and Results section.
- Catch-up if we have fallen behind.
- Review of any topics or problems, work sessions in lab on final project.

Alert: Project will be time-consuming and you only have this week through the weekend to do it! I'll try to hold as much class time as I can to support getting this done and so you can ask me questions. Project will be due before midnight on February 24.

Readings:

1. Hayes intro to mediation Chapter 1, on Carmen.
2. Review Bem and see examples in SM (e.g., for factor analysis, though SM overuses passive voice).

Final Project – See syllabus and class announcements for details on the project.