

---

**AHBR 548.01**

**Applied Data Management**

**Spring 2019**

**Syllabus**

Thursday, 4:30 to 7:00

**Instructor:**

Michael Elliott

Associate Professor

Salus Center Room 478

3545 Lafayette Ave.

Saint Louis, MO 63104-1314

314-977-4027 Phone

314-977-3234 Fax

elliottmb@slu.edu

**Office Hours:**

I am available to meet with students by appointment.

Please use [https://michael\\_elliott.youcanbook.me/](https://michael_elliott.youcanbook.me/)

---

## **Course Overview**

### *Course Description*

This course is an advanced course on data management for graduate students. Students will learn advanced concepts and techniques of research data management with particular emphasis on applications in health. Students will learn to use multiple data management and data analysis software packages including Excel, Access, MySQL SPSS, SAS and R.

## **Course Description**

***Course Format:*** (Lectures, class discussions, guest lectures)

### *Learning Objectives:*

1. To provide students with an in-depth understanding of important data management tools and techniques.
2. To give students hands-on experience with modern statistical and database programming environments.
3. To show students how data management practices are implemented in several real-world research settings.
4. To expose students to data management aspects of a research project from start to finish.

### ***At the end of the class, students will be able to:***

- Understand the importance of data documentation
- Design an effective data codebook
- Design an on-line survey
- Understand the strengths and weaknesses of major statistics packages for data management
- Import, export, and transfer data among various database and statistical platforms
- Fully document datasets
- Manage, manipulate, transform, and recode datasets
- Find and use national level data
- Design a simple relational database
- Perform simple SQL queries

***Course Format:*** Lectures, class discussions, guest lectures

***Course Elements and Requirements:***

1. Class participation, including readings as assigned. Participation is very important for a skills-based class such as Applied Data Management. Evidence of participation will include regular attendance, active involvement in class discussions, evidence of having read assignments, and active involvement with small groups. (20%)
2. Completion of weekly assignments. There will be 10 short homework assignments given out throughout the semester. These will focus on applied data management skills, such as data input, data cleaning, data import and export, and data transformations. **Homework will be due the week after it is assigned. Late homework will be docked 5 percentage points per week.** (50%)
3. Final presentations. There will be two final presentations: a group presentation and an individual presentation. The group presentation is the culmination of the group work that has been performed throughout the class. It is worth 10% of the final grade. The individual presentation will be a 15-20 minute Powerpoint presentation on a data management topic of the student's choosing. This topic can either be a novel data management technique (not discussed in class, such a cell phone data collection), or a novel data management environment/need (such as market research, "heat" maps). It is worth 20% of the final grade.

***Required Materials***

Watson, R. (2013). *Data Management Databases and Organizations* (6<sup>th</sup> Ed.). (Available only in Kindle version through Amazon for \$9.99).

Osborne J. (2013). *Best Practices in Data Cleaning: A Complete Guide to Everything You Need to Do Before and After Collecting Your Data*. Sage Publications. Can be ordered through Amazon for 28.61 paperback or 27.18 Kindle version.

You will need to be able to use a reliable computer.

***Grading Determination and Policy***

Final grades for this class will be assigned as following: A ( $\geq 90$  to 100 points), B ( $\geq 80$  to 89 points), C ( $\geq 70$  to 79 points) and F ( $< 70$  points).

**Course Calendar**

Session	Date	Topic	Assignment (due following week)
1	January 17	Introduction to data management Causal models Good questions	-create causal model (in class) <b>-find scales (#1)</b> <b>-create additional questions (#1)</b> <b>-read Watson Chapter 1</b>
2	January 24	Documentation/Codebooks	<b>-create codebook (in class) (#2)</b> <b>-read Watson Chapter 2</b>
3	January 31	Data Entry basics (Excel, ASCII)	-take survey (in class) <b>-enter data in Excel (#3)</b> <b>-read Osborne chapters 3 and 10</b>
4	February 7	On-line data collection/Census data	<b>-create on-line survey (#4) (in class)</b> <b>-read Watson chapters 7-9 and Reference 1</b>
5	February 14	DBMS theory	<b>-conduct group data collection (#5)</b> -create data model (in class) <b>-read Watson chapters 3-6</b>
6	February 21	Database development	-develop database (in class) <b>-read Watson chapter 10 and Ref. 2</b>
7	February 28	Database development	<b>-create query and report (in class) (#6)</b>
8	March 7	Data management in R	
<b><i>March 14 No Class Due to Spring Break!</i></b>			
9	March 21	Importing Data in SPSS	<b>- import data (#7)</b> <b>-read Osborne chapter 4</b>
10	March 28	Database presentations	
11	April 4	Using large datasets	<b>-design effect project (#8)</b> <b>-read Osborne chapters 1, 2, and 5</b> <b>-read Watson chapter 21</b>
12	April 11	Data quality/Cleaning	<b>-develop data cleaning plan (#9)</b>

**-read Osborne chapter 6**

13 April 18

Missing data

**-impute data (#10)**

**-read Osborne chapters 7 and 8**

14 April 25

Group presentations and Individual presentations