

CLRE 262 – Analyzing Medical Data Using R

MAS/CREST

Fall 2015

Course Director: Florin Vaida, PhD fvaida@ucsd.edu

Course Instructors: Florin Vaida
Anya Umlauf, MS aumlauf@ucsd.edu

Course Software: R open source software, at <http://www.r-project.org>

- Homework 0: Students should install R on their laptops prior to first lecture

Textbooks: Lecture notes

Useful Resources (see also BlackBoard > Resources):

1. Everitt BS and Hothorn T: A Handbook of Statistical Analyses Using R, Third Edition, Chapman & Hall/CRC, 2014 – see course website under *Resources*
2. A list of books on data analysis using R is at <http://www.r-project.org/doc/bib/R-books.html>
3. R website <http://www.r-project.org/> includes a wealth of information on R, including tutorials
4. Venables WN and Smith DM: An introduction to R – see course website (*Resources*)
5. Verzani J: simpleR – Using R for Introductory Statistics – see course website (*Resources*)

Course Objectives:

1. Introduce the students to the R statistical platform
2. Build proficiency using R for data analysis using the common statistical methods
3. Reinforce and expand the students' understanding and knowledge of commonly used statistical methods
4. Develop skills in analyzing statistical data for biomedical research, presenting and organizing analysis results in tabular and graphical forms, and preparing a manuscript for peer-review publication

Teaching Format:

Two 50-minute sessions. Case studies will be introduced and discussed for each key statistical concept. We will also conduct interactive data analysis sessions using R. Students will work in groups.

Groups:

1. The students will work in *groups of two* for homework and the final project.
2. If the course enrollment is of 13 or more students, the group size will be of 2-3 students (no more than 6 groups per class).
3. All students in the group will receive the *same grade for each assignment*.
4. The final course grade may differ within a group, however, e.g. due to absences during project presentations.
5. Students are responsible for forming their own groups. Students who did not find a group should inform the instructor(s) as soon as possible.
6. Homework 1 and/or 2 may be submitted individually.
7. All groups *must be set by homework 3* and cannot be changed afterwards.
8. Any issues regarding *lack of equal participation* of all members of the group should be brought to the attention of the instructor as soon as possible.

Homework:

1. Weekly assignments, to be submitted through Blackboard/TED.
2. Homework is due one week (7 days) from assignment. Late homework is not acceptable.
3. The only acceptable format is .pdf. Do not submit raw R output!
4. One homework submission for a group of two students.

Final Exam: *There will be no final exam*

Group Project: *There will be a group project, with documents and presentation due at the last session (Week 11), with a proposal due under Homework 3 and preliminary results and presentation due under Homework 7. This is discussed in detail in a separate document.*

Course Grade:

The course grade will have three components:

1. **50%** Homework (including project proposal and preliminary results)
2. **50%** Final group project paper and presentation

Attendance Policy:

The general CREST attendance policy applies, no more than 3 missed sessions are allowed. You have to be present in both halves of the class.

Time Commitment:

Expect to spend 4-6 hours a week outside of class.

Timeline and Syllabus

| Lecture | Date | Topic | Assignment |
|----------------------------|---------------|--|-----------------------------|
| Wk1 (Umlauf) | Sept 21 | Introduction to R | Hw0 due |
| Wk2 (Umlauf) | Sept 28 | Describing data using R | Hw1 due |
| Wk3 (Umlauf) | Oct 5 | Statistical methods for two groups | Hw2 due |
| Wk4 (Umlauf) | Oct 12 | Analysis of variance and covariance | Hw3 due (project) |
| Wk5 (Umlauf) | Oct 19 | Simple linear regression | Hw4 due |
| Wk6 (Vaida) | Oct 26 | Multiple linear regression | Hw5 due |
| Wk7 (Vaida) | Nov 2 | Logistic regression | Hw6 due |
| <i>Wk8 (Vaida)</i> | <i>Nov 9</i> | <i>Project Presentation: Preliminary Results</i> | <i>Hw7 due (project)</i> |
| Wk9 (Vaida) | Nov 16 | Analysis of survival data | Hw8 due |
| | <i>Nov 23</i> | <i>No class</i> | |
| Wk10 (Vaida) | Nov 30 | Analysis of longitudinal data | Hw9 due |
| <i>Wk11 (Vaida/Umlauf)</i> | <i>Dec 7</i> | <i>Final Project presentation</i> | <i>Project presentation</i> |