

Course Syllabus: Special Topic: Applied Statistics with R - STAT 290B

| | |
|--|--|
| Division | Computer, Electrical and Mathematical Sciences & Engineering |
| Course Number | STAT 290B |
| Course Title | Special Topic: Applied Statistics with R |
| Academic Semester | Spring |
| Academic Year | 2018/2019 |
| Semester Start Date | 01/27/2019 |
| Semester End Date | 05/23/2019 |
| Class Schedule (Days & Time) | 02:30 PM - 04:00 PM Sun Wed |

| Instructor(s) | | | | |
|------------------------|------------------------------------|-------|---------------------------------|--|
| Name | Email | Phone | Office Location | Office Hours |
| Joaquin Ortega Sanchez | JOAQUIN.ORTEGASANCHEZ@KAUST.EDU.SA | | 4297, 1, Al-Khwarizmi (bldg. 1) | Mondays 14:00 - 15:00 Wednesdays 10:00 - 11:00 Thursdays 15:00 - 16:00 |

| Teaching Assistant(s) | |
|-----------------------|-------|
| Name | Email |
| | |

| Course Information | |
|--|---|
| Comprehensive Course Description | This course is an introduction to practical applied statistics with R, a programming language and software environment for statistical computing, and with RStudio, an integrated development environment for R. Topics include introduction to R, data and programming, summarizing data, probability and statistics in R, simple and multiple linear regression, categorical predictors and interactions, model diagnostics, collinearity, variable selection and model building, selected data analyses. |
| Course Description from Program Guide | This course is an introduction to practical applied statistics with R, a programming language and software environment for statistical computing, and with RStudio, an integrated development environment for R. Topics include introduction to R, data and programming, summarizing data, probability and statistics in R, simple and multiple linear regression, categorical predictors and interactions, model diagnostics, collinearity, variable selection and model building, selected data analyses. |
| Goals and Objectives | This course will provide students with an introduction to applied statistics with R and RStudio. Specific objectives: first, to become experienced with the high-level programming language R, an environment for data analysis and graphics, and the use of the RStudio interface and its functionalities to perform reproducible research projects. Second, to become experienced with basic statistical techniques, rigorously implemented in R. Third, to begin the study of more advanced statistical methods on the basis of case studies from challenging real-world problems. |
| Required Knowledge | The material provided in introductory graduate courses on calculus, basic statistical theory and probability. |

| | |
|----------------------------------|--|
| Reference Texts | <p>Nicholas J. Horton, Ken Kleinman (2015). Using R and RStudio for Data Management, Statistical Analysis, and Graphics, Second Edition, CRC Press.</p> <p>Pierre-André Cornillon et al. (2012). R for Statistics, CRC Press.</p> <p>Michael J. Crawley (2013). The R Book, Second Edition, John Wiley & Sons.</p> <p>Peter Daalgard (2008). Introductory Statistics with R, Springer.</p> <p>Proposed reference texts for the final projects:</p> <p>John Maindonald, W. John Braun (2010). Data Analysis and Graphics Using R – an Example-Based Approach, Third Edition, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.</p> <p>Frank E. Harrell, Jr. (2015). Regression Modeling Strategies, Second Edition, Springer Series in Statistics, Springer.</p> <p>Christopher Gandrud (2015). Reproducible Research with R and RStudio, Second Edition, The R Series, Chapman & Hall/CRC Press.</p> <p>Thomas Rahlf (2017). Data Visualisation with R: 100 Examples, Springer.</p> <p>Yangchang Zhao (2013). R and Data Mining: Examples and Case Studies, Academic Press.</p> |
| Method of evaluation | <p>40.00% - Course Project(s)</p> <p>30.00% - Midterm exam</p> <p>30.00% - Homework /Assignments</p> |
| Nature of the assignments | <p>Homework exercises.</p> <p>Assigned reading of specific book chapters and papers.</p> <p>Individual project.</p> |
| Course Policies | <p>Attendance is required.</p> <p>Assignments will be handed out on the announced dates throughout the course.</p> <p>Late assignment submissions will not be accepted unless prior arrangements have been made (except in university established cases of illness or emergency). All homework assignments must be neatly typed (LaTeX or RMarkdown are recommended). All projects and homework assignments are required.</p> <p>Grades will be posted on the course website.</p> <p>If you dispute your grade on any homework or project, you may request a regrade (from the TA/instructor for the assignments or the instructor for the exams) only within 48 hours of receiving the graded exam.</p> <p>Incomplete (I) grade for the course will only be given under extraordinary circumstances such as sickness, and these exceptional circumstances must be verifiable. The assignment of an (I) requires first an approval of the Dean and then a written agreement between the instructor and student specifying the time and manner in which the student will complete the course requirements.</p> <p>Collaboration and checking answers on assignments are allowed and encouraged. Of course, copying assignments is not tolerated. In brief, you are allowed to collaborate on all homework problems according to the following rules: you must first attempt to solve each problem on your own. If you get stuck, you can then talk to any student currently enrolled in the class about the issue, as well as the instructor or the TA. However, solutions and R code should not be exchanged (i.e., you still must work through the details of the problem after you have gotten help, write the final answers alone, and understand them thoroughly).</p> |
| Additional Information | |

Tentative Course Schedule

(Time, topic/emphasis & resources)

| Week | Lectures | Topic |
|-------------|----------------------------------|---|
| 1 | Sun 01/27/2019 Wed 01/30/2019 | Getting started with R. Essentials of the R language. R objects. RStudio |
| 2 | Sun 02/03/2019 Wed 02/06/2019 | R programming structures. R functions. |
| 3 | Sun 02/10/2019 Wed 02/13/2019 | Data input and output. Data Management |
| 4 | Sun 02/17/2019 Wed 02/20/2019 | Graphics with R. Tools for reproducible research |
| 5 | Sun 02/24/2019 Wed 02/27/2019 | Probability distributions and random number generation. Matrix algebra. Summary statistics. |
| 6 | Sun 03/03/2019 Wed 03/06/2019 | Simulation. Introduction to the Montecarlo method. Bootstrap. |
| 7 | Sun 03/10/2019 Wed 03/13/2019 | Statistical models. Basic inference concepts. Classical statistical tests |
| 8 | Sun 03/17/2019 Wed 03/20/2019 | Regression analysis. |
| 9 | Sun 03/24/2019 Wed 03/27/2019 | Spring Break |
| 10 | Sun 03/31/2019 Wed 04/03/2019 | Regression analysis. |
| 11 | Sun 04/07/2019 Wed 04/10/2019 | Analysis of variance. |
| 12 | Sun 04/14/2019 Wed 04/17/2019 | Simple and multiple analysis of covariance. |
| 13 | Sun 04/21/2019 Wed 04/24/2019 | Simple and multiple analysis of covariance. |
| 14 | Sun 04/28/2019 Wed 05/01/2019 | Count data. Proportion data. |
| 15 | Sun 05/05/2019 Wed 05/08/2019 | Ordinal logistic regression. |
| 16 | Sun 05/12/2019 Wed 05/15/2019 | Parametric survival modeling. |
| 17 | Sun 05/19/2019 Wed 05/22/2019 | Final Exam Week |

Note

The instructor reserves the right to make changes to this syllabus as necessary.