

SyllabusLab

Honors Data Science Lab: Statistical Programming with R (IDS 3932L 1 credit)

Instructors: Kevin Lanning

Term: Spring 2018

Lanning office: WB 213; Office hours Mondays and Tuesdays 2-5 and by appointment

Class meeting days/time: IDS 3932L – W 2-5

Class location: SR 258

KL Phone: 561-799-8652

Emails: lanning@fau.edu Website: <https://kevinlanning.github.io/DataSciSpring2018/topics.html>

Course Description

IDS 3932L is a laboratory which may accompany IDS 3932, the introductory seminar on data science. The course receives honors credit by virtue of its small class size, by virtue of its dialectic structure, and by the fact that students receive extensive exposure to supplementary materials and primary sources.

Course Prerequisites/Co-requisites

STA 2023 or equivalent or permission of instructor is a prerequisite for IDS.

Enrollment in IDS 3932 is a corequisite for IDS 3932L. (Enrollment in IDS 3932L is not a corequisite for IDS 3932. See [prereqs.html](#) for one way to depict these relationships).

Texts and Materials

For the lab, there is one required text.

Wickham, H. & Golemund, G. (2016) *R for data science*. Sebastopol, CA: O'Reilly or online at <http://r4ds.had.co.nz>.

In addition, there are a number of sources which we are likely to access at least occasionally, including books, papers, and data science courses at other universities (see references). All of these are presently available online.

Course Objectives

Hochster (in Hicks & Irizarry, 2017) describes two broad types of data scientists: Type A (Analysis) data scientists, whose skills are like those of an applied statistician, and Type B (Building) data scientists, whose skills lie in problem solving or coding, using the skills of the computer scientist. Our course is closer to a Type A than a Type B treatment, one which is closer to Statistics than to Computer Science, but it is also essentially concerned with content in the concentrations in the arts, humanities, and natural and social sciences. As discussed in the first meeting, it is arguably best understood as a third (Type C) approach, one which has as its objective progress in the understanding of statistics and computing, but also skills such as collaboration and communication, providing an introduction to some of the methods and tools of reproducible science, and fostering a heightened sensitivity to the ethical challenges of the digital age. (*IDS* is an abbreviation for “InterDisciplinary Studies” as well as “Introduction to Data Science.” So our course number is particularly apt).

In addition to R, we'll use a range of other tools: We'll communicate on the Slack platform. We'll write using markdown editors such as Typora. We'll certainly use spreadsheets such as Excel or Google Sheets. Students working on data visualization may use additional tools for visualizing data such as Gephi and Tableau.

Course Evaluation: IDS 4932L

Grades will be based on a 100 point scale, with points earned by participation and a term project.

Participation (50 points). Attendance is a necessary but not sufficient part of class participation. Your participation grade will be based also on the extent to which you contribute to our class by asking constructive questions and helping us solve the numerous challenges which we collectively will face. In the lab, you are a teacher as well as a learner; a strong participation grade will be earned by leading the class towards interesting questions, resources, and solutions.

A term project (50 points). Group. Learning is social. The term project will be a collaborative, data based **project**, which you will write with one to four of your peers. The typical project will involve (a) finding an interesting dataset, (b) asking some interesting questions about it, (c) preparing your data for analysis, (d) applying appropriate techniques for data analysis, and (e) presenting your data in a way which is both reproducible (sharing your code) and edifying (communicated in a way that teaches your audience (where audience = me + your classmates). The project will ideally be a publishable document (on a repository such as rpubs, GitHub, or the OSF).

As in the lecture, in order for me to assess your individual contributions and to minimize social loafing, I ask that (1) all meetings and communications among group members be undertaken on the Slack platform, (2) that all writing be done in Google docs, and (3) in addition to the paper, that all group members sign a 1-page cover page describing the primary contribution and percent effort of each person. Your primary results may include a visualization, but should also include text and code.

Groups and topics will be developed in class.

Class Schedule

** for detail, please see <https://kevinlanning.github.io/DataSciSpring2018/schedule> **

Weeks 1-3: Introduction

Weeks 3-5: Towards data literacy

Weeks 5-10: Towards data proficiency

Weeks 10-14: Towards Data fluency

Weeks 14-15: Towards Data leadership

Etc.

Classroom etiquette policy

In order to enhance and maintain a productive atmosphere for education, personal communication devices, such as cell phones, are to be disabled in class sessions. Laptops and tablets should be used only for appropriate purposes.

Attendance Policy

Note that as course grades are based in part on classroom participation, regular attendance is effectively a course requirement. Absences may be excused due to illness or a medical emergency.

Policy on Accommodations

In compliance with the Americans with Disabilities Act (ADA), students who require reasonable accommodations to properly execute coursework must register with Student Accessibility Services (SAS) – in Boca Raton, SU 131 (561-297-3880); in Davie, LA 131 (954-236-1222); in Jupiter and all Northern Campuses, SR 111F (561-799-8585) – and follow all SAS procedures.

Code of Academic Integrity Policy Statement

Students at Florida Atlantic University are expected to maintain the highest ethical standards. Academic dishonesty is considered a serious breach of these ethical standards, because it interferes with the university mission to provide a high quality education in which no student enjoys an unfair advantage over any other. Academic dishonesty is also destructive of the university community, which is grounded in a system of mutual trust and places high value on personal integrity and individual responsibility. Harsh penalties are associated with academic dishonesty. For more information, see University Regulation 4.001.