Course Syllabus

CS 636: Data Analytics with R Programming
Spring 2019

Class Schedule: Saturday 9:00 am - 11:50 am, Fenster Hall 160
Instructor: David Li, email: dli@njit.edu, tel: 631-800-3381

TA: Shih-Chuan Weng, email: sw464@njit.edu, tel: 973-718-1822
Office Hours: Friday 11:00 am – 12:00 am, GITC 4th floor. Please schedule with Shih-Chuan so that he can reserve a seat.

Course Description and format:
This course will teach how to program in R and how to use R for effective data analysis. The students will learn how to install and configure R necessary for an analytics programming environment and gain basic analytic skills via this high-level analytical language. The course covers fundamental knowledge in R programming. Popular R packages for data science will be introduced as working examples. The format of the course will include lectures by the instructor, computing labs, class discussion, directed reading, and student presentation or project. The exact format will depend on the size of enrollment and student background.

Prerequisite: Some basic knowledge of programming, probability and statistics. If in doubt about the prerequisites, please consult with the instructor for permission to take the class.

Attendance: You are supposed to attend all the classes. Participation is highly encouraged to make the class more interactive. Class attendance and participation are taken into consideration by the instructor for the evaluation of the students. In general, students who attend class regularly perform much better than those who come only occasionally. If you miss one class be sure to consult one of your classmates about the content of the lecture and visit the course web page and moodle (moodle.njit.edu) to get notes, exercises, assignments, deadlines and announcements.

Textbooks:

- R Programming for Data Science, by Roger D. Peng, https://leanpub.com/rprogramming
- Using R for Introductory Statistics, by John Verzani, Chapman & Hall/CRC, 2004, ISBN 1584884509
- Advanced R, by Hadley Wickham, ISBN 9781466586963.

Grading:

The requirements of this course will consist of participating in lectures, homework, in class computing lab assignments, two exams and two projects. The grading breakdown is the following:

- Homework, computing lab exercise (5%)
- Quiz (20%)
- Projects (10%)
- Midterm Exam (25%)
- Final Exam (40%)

Collaboration and Honor Code Students may discuss problems together but must write up their own solutions. When writing up the solutions, students should write the names of people, if any, with whom they discussed the assignment. Note in particular that copying homework or programming assignments, in full or in part is forbidden. Students found cheating or plagiarizing will be immediately referred to the Dean of Students and the NJIT Committee on Professional Conduct and subject to Disciplinary Probation, a permanent marking on the record, possible dismissal, and an "F" grade in the course. All submitted assignments will be checked for similarities, and plagiarism and guilty students identified.

Tentative course topics (Subject to changes according to progress)

Date	Topic	Readings	Assignments	Comments
1/26/2019	Class overview + Introduction R Nults and Bolts (I)	R Programming: Chapters 3, 4; Using R: Chapter 1		
2/02/2019	R Nults and Bolts (II)	R Programming: Chapters 5, 10, 11; Using R: Chapter 2.1	Lab 1 HW 1	
2/09/2019	Getting Data In and Out of R	R Programming: Chapters 6-9	Lab 2 Quiz 1	
2/16/2019	Control Structures and Functions	R Programming: Chapters 14-17	Lab 3 HW 2	
2/23/2019	Loop Functions	R Programming: Chapter 18	Lab 4 Quiz 2	
3/02/2019	String Operations (stringr package)	R Programming: Chapter 19	Lab 5 Project 1	
3/09/2019	Data Manipulation (dplyr, reshape2 packages)	R Programming: Chapter 13	Lab 6 Quiz 3	
3/16/2019	***Midterm***			
3/23/2019	***Spring Recess***			
3/30/2019	Probability Basics & Data Exploration		Lab 7	
4/06/2019	Clustering	April 8 Monday Last Day to Withdraw	Project 2	
4/13/2019	Regression and Classification (linear models)		Lab 8 HW 3 Quiz 4	

4/20/2019	Model/Feature Selection		Lab 9 HW 4	
4/27/2019	SVM and other Representative Classifiers		Quiz 5	
5/04/2019	R Graphics and Debugging; Packaging & Profiling; Real Case Study	R Programming: Chapters 20,21	Final Review	Project Due
5/11/2019	***Final Exam***			TBD