MW 12:00 – 1:50 p.m.

Profesor : Cesar Acosta, Ph.D.
Office : GER 216
Office hours: M 3 – 4 p.m.
e-mail : acostame@usc.edu

Teaching Assistant: TBD
Office : TBD
Office hours : TBD
e-mail : TBD

Pre-requisites: ISE 225 Engineering Statistics

Course Objectives: This course provides an overview of data analytics methods and their application to engineering and decision problems. The focus is on learning from data through statistical model fitting. The course scope is on regression, classification, and, cluster analysis. Most of the methods reviewed start analyzing past history to discover hidden information. The goal is to make predictions about the response of interest. To apply these methods some computational tool is needed. In this course *R* will be used.

Week	Topic	
Jan 9	Introduction to R	
Jan 18	Simple & Multiple Linear Regression	Ch. 3
Jan 23	ANOVA	notes
Jan 30	Multiple linear Regression, categorical variables	Ch. 3
Feb 6	The bootstrap	Ch. 5
Feb 13	Cross-validation	Ch. 5
	Midterm Exam	**Feb 22**
Feb 27	Logistic Regression – introduction	Ch. 4
Mar 6	Logistic Regression, KNN, Loess	Ch. 4
Mar 20	Unsupervised Learning – Clustering & Principal Components	Ch. 10
Mar 27	Tree methods – Classification & Regression Trees (CART)	Ch. 8
Apr 3	Tree methods – Bagging, Random forest, boosting	Ch. 8
Apr 10,17	Support Vector Machines	Ch. 10
Apr 24	Hierarchical linear models	notes
	FINAL EXAM	**May 5 (11 a.m.)**

Course Details

- 40% course is made of computing sessions
- Datasets are expected to be multivariate and high dimensional
- A laptop is required during class sessions
- No textbook is required. References will be the choice of the students.

Text (free)

An Introduction to Statistical Learning with Applications in R.
James, Witten, Hastie, and, Tibshirani. Springer, 2015

Software

The *R language and environment for statistical computing and graphics* is the main computational tool. Many libraries (known as *R* packages) are expected to be used in this course. A WiFi connection is required to download and install them
In addition *R studio* is usually the most preferred interface but not required.
CSV files are sometimes the file type from which the data is loaded on to *R*.
Students will use their own laptop during exams and class sessions.

Grading Policy: attendance, assignments 50%, midterm 25%, final exam 25%.

Academic Integrity. The Viterbi School of Engineering adheres to the University's policies and procedures governing academic integrity as described in SCampus (www.usc.edu/dept/publications/SCAMPUS/). Students are expected to be aware of and to observe the academic integrity standards described in SCampus, and to expect those standards to be enforced in this course.

Students with Disabilities. Any Student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me (or to TA) as early in the semester as possible. DSP is located in STU 301and is open 8:30 a.m. - 5:00 p.m., Monday through Friday. The phone number for DSP is (213) 740-0776