CS/STAT 6313.0U1 – Statistical Methods for Data Science

Course Information

Course Number/Section	*CS/STAT 6313.0U1*
Course Title	*Statistical Methods for Data Science*
Term	*Summer 2016*
Days & Times	*MW 12:30 – 2:45pm*
Location	*SLC 2.302*

Professor Contact Information

Instructor	*Dr Bill Semper*
E-Mail	*WJS130130@utdallas.edu*
Office hours	*MW 3:00 - 4:00 pm, ECSS 4.706*
Phone	*972-883-4139*
Website	*www.utdallas.edu/~wjs130130/*

Course Pre-requisites, Co-requisites, and/or Other Restrictions:

Prerequisites: CS/SE/STAT 3341 or equivalent.

Course Description

Introduction to statistics, including simple descriptive statistics and graphical statistics. Statistical inference, including parameter estimation, confidence intervals, and hypothesis testing. Chi-square tests, nonparametric statistics, Bootstrap, and Bayesian Inference. Regression Analysis, including Least squares estimation, analysis of variance, ANOVA, and multivariate regression. Introduction to the R programming language, with hands on application.

Student Learning Objectives/Outcomes

Students will learn the basics of statistical analysis and be able to apply the fundamental methods of sampling, interval estimation, hypothesis testing, and regression analysis. Students will be able to use the R programming language to analyze data sets.

Textbooks and Materials

Text: Applied Statistics and Probability for Engineers, Montgomery and Runger, 6th Edition, Wiley. ISBN-13 9781118539712. We will cover chapters 1-5 (briefly), 6 – 12.

Text: Probability and Statistics for Computer Scientists, M. Baron, CRC Press (2007) or second edition (2013), ISBN 1584886412 or 1439875901.

Academic Calendar: ***These descriptions and timelines are subject to change at the discretion of the Professor.***

Date	Topic
May 23	Classes Begin
June 27	Mid-Term Exam
August 8	Classes End
August 9-10	Final Exam

Grading Policy

HW	10%
Mid-Term Exam	45%
Final Exam	45%

Please note: No extra credit assignments will be given.

Grading will be on a curve, with the median point for the class being used as the B+/A- cutoff point. Cutoffs for A and B+ grades will be roughly 2% from the median. If the median score is above 90%, the following scale is used:

$$\begin{array}{ll} 92 \leq score \leq 100 & A \\ 90 \leq score < 92 & A- \\ 88 \leq score < 90 & B+ \\ 82 \leq score < 88 & B \\ 80 \leq score < 82 & B- \\ score < 80 & C \end{array}$$

The instructor reserves to the right to assign B- grades to students with final scores less than 2 standard deviations below the median score. The instructor reserves the right to assign C grades to students who score less than 50% on the final.

For detailed information about University policies and procedures related to this syllabus, please refer to http://go.utdallas.edu/syllabus-policies.