40.004: Statistics

Course Description:
This course provides an introduction to applied statistics and data analysis. The course will cover the following topics: collecting and exploring data, estimation and hypothesis testing, linear regression, analysis of variance, and elementary methods in non-parametric statistics. Applications and examples based on engineering systems data will be given.

Learning Objectives:
At the end of the term, you will be able to:
1. Examine data and use tools to visualize data to uncover relationships.
2. Develop estimates and confidence intervals from a data sample; perform hypothesis testing and goodness-of-fit tests.
3. Build a regression model and estimate the parameters involved, and perform a diagnosis on the quality and validity of the model.
4. Design an experiment and conduct analysis of variance for a multi-treatment samples.

Textbook:
Ajit C. Tamhane and Dorothy D. Dunlop, *Statistics and Data Analysis: From Elementary to Intermediate*, Prentice-Hall, 2000.
Copies of the book are available from the library; the call number is QA276TAM.

Course assessment:

Exam 1 (3 March, Week 6, 2 hours)	30%
Exam 2 (28 April, Week 14, 2 hours)	30%
2D Project (with OM and/or SM)	20%
Problem Sets (4 sets in total)	15%
Class Participation	5%

Class format: 2 lectures + 1 recitation per week. Non-comprehensive slides will be provided, though students should consult the textbook for extra exercises and to gain a more complete understanding.

Prerequisite: 40.001 Probability or 03.001 Probability and Statistics.

Computing: *Excel* (with the Data Analysis Add-In) is the software that will be used in the course. For students interested in more computing, *R* is recommended. A set of tutorials on *R* will be offered during recitations.

Academic Integrity:
Students are held accountable for the integrity of the work they submit, which means copying and allowing others to copy are strictly prohibited. Working with others is fine and even encouraged, provided that each student writes the solutions by him/herself.

Teaching team:

Lecturer:

Assistant Professor James Wan   Email: james_wan@sutd.edu.sg

Teaching Assistant:

Qifang Bao   Email: qfbao@mit.edu

Office hours: by email appointment.

Course schedule:

A rough schedule of the material that will be covered is given below. We may deviate from the schedule from time to time, based on reactions to each week's material.

Week	Topics	Textbook reference
1	Introduction Collecting Data	Chapter 1, Chapter 3: Section 3.1 Chapter 3: Sections 3.2, 3.3, 3.4
2	Summarizing and Exploring Data	Chapter 4: Sections 4.1, 4.2, 4.3 Chapter 4: Sections 4.4, 4.5
3	Sampling Distribution of Statistics Basic Concepts of Inference	Chapter 5: Sections 5.1, 5.2 Chapter 6: Sections 6.1, 6.2
4	Basic Concepts of Inference Inference for Single Samples	Chapter 6: Sections 6.3 Chapter 7: Sections 7.1, 7.2
5	Inference for Single Samples Inference for Two Samples	Chapter 5: Section 5.1, 5.2, 5.3 Chapter 8: Sections 8.1, 8.2, 8.3
6	Revision Exam 1	
7	Recess	
8	Inferences for Proportions	Chapter 9: Section 9.1, 9.3, 9.4
9	Simple Linear Regression and Correlation	Chapter 10: Sections 10.1, 10.2, 10.4
10	Multiple Linear Regression	Chapter 11: Sections 11.1, 11.2, 11.3
11	Simple Linear Regression and Correlation Multiple Linear Regression	Chapter 10: Sections 10.3, 10.6 Chapter 11: Sections 11.4, 11.6, 11.7
12	Analysis of Single Factor Experiments Analysis of Two-Factor Experiments	Chapter 12: Sections 12.1, 12.2 Chapter 13: Sections 13.1
13	Nonparametric Statistical Methods Likelihood Methods	Chapter 14: Section 14.1 Chapter 15: Section 15.1
14	Exam 2	