MET CS 544 - Foundations of Analytics

Course Description

The goal of this course is to provide students with the mathematical and practical background required in the field of data analytics. Starting with an introduction to probability and statistics, the R tool is introduced for statistical computing and graphics. Different types of data are investigated along with data summarization techniques and plotting. Data populations using discrete, continuous, and multivariate distributions are explored. Errors during measurements and computations are analyzed in the course. Confidence intervals and hypothesis testing topics are also examined. The concepts covered in the course are demonstrated using R. Laboratory Course.

Course Prerequisites

MET CS 546 - Quantitative methods for Information Systems, or equivalent background

Course Grading Policy

The course grade will be based on active class participation and quizzes (10%), assignments (30%), mid term exam (30%), and a final exam (30%). Assignments are expected to be submitted by their respective due dates. Late submissions carry a penalty.

Course Topics

Module 1 - Introduction

- Probability and Statistics
- Basic concepts of R

Module 2 - Probability

- Counting methods
- Conditional probability and Independence
- Random variables
- Functions of random variables

Module 3 - Data Description

- Types of Data
- Univariate, Bivariate, and Multivariate Data
- Data summarization and presentation

Module 4 - Distributions

- Bernoulli, Binomial, Normal Distribution
- Uniform Distribution
- Central Limit Theorem

Module 5 - Errors

- Propagation of Errors
- Measurement of Error
- Linear combination of measurements
- Type-I and Type-II errors
- Biases, Noise, Data Dredging

Module 6 - Estimation of Performance

- Confidence Intervals, Hypothesis Testing
- P-Values, Correlation
- Statistical Tests (T-Test, etc.)
- Resampling (bootstrapping, cross-validation, randomization)

Reference Textbooks

J. Verzani, Using R for Introductory Statistics, Chapman & Hall, 2005
http://cbb.sjtu.edu.cn/~mywu/bi217/usingR.pdf

G. Jay Kerns, Introduction to Probability and Statistics Using R, 2010.
http://cran.r-project.org/web/packages/IPSUR/vignettes/IPSUR.pdf

M. J. Crawley, Statistics: An Introduction Using R, Wiley, 2005.
http://www3.imperial.ac.uk/naturalsciences/research/statisticsusingr

Course Web Site

http://lms.bu.edu

Student Conduct Code

http://www.bu.edu/met/for-students/met-policies-procedures-resources/academic-conduct-code/