3.1.4 MSTA 1101 STATISTICS FOR DATA SCIENCE (42 HOURS)

Instructional Hours: 28 Lecture hours and 42 practical hours

Prerequisite: None

Purpose

This course exposes the learner to many different types of quantitative and statistical techniques and software tools for analyzing data.

Expected learning outcomes:

At the end of this course, the student should be able to:

1. Analyse qualitative data using R/Python;

2. Apply quantitative techniques and methods in data analysis using R/Python;

3. Apply qualitative techniques and methods in data analysis using R/Python;

Course description

Introduction to statistical techniques for analyzing data. Variables and Measurements. Descriptive, inferential statistics and causal inference.

Sampling versus experimental design.

Parametric and non-parametric tests of difference in Means and Relationships.

Regression: Ordinary least squares regression and logistic regression. Use of R and Python Software.

Instructional Methods

Lectures, Lab sessions, discovery learning, problem-based learning, experiential learning, group-based learning, independent studies, illustrations, and online tutorials /exercises.

Instructional Materials

Whiteboards, markers, handouts, presentation software, LCD projectors and computers, Flipcharts, televisions,videos.

Course Assessment

CATs, Assignments, Presentations 40%

Final Examination 60%

Total 100%

Course Text Books

1. James D. Miller (2017) Statistics for Data Science: Leverage the power of statistics for Data Analysis, Classification, Regression, Machine Learning, and Neural Networks Kindle Edition Packt Publishing ASIN: B06Y2XX2LH

2. Peter Bruce & Andrew Bruce (2017) Practical Statistics for Data Scientists: 50 Essential Concepts 1st Edition, Kindle Edition ASIN: B071NVDFD6

3. John Slavio (2017) Statistics: Practical Concept of Statistics for Data Scientists Kindle Edition ASIN: B07768JKTB