

MDS503 Statistics for Data Science

PRE-REQUISITE(S)

Basic statistics

CREDIT HOURS

4

SYPNOSIS

This module provides an introduction to basic statistical concepts and methods which include: simple and

multiple linear regression, classification, decisions trees, support vector machines, and unsupervised learning.

In addition, this module will serve as an introduction to implementing these methods through the use of

statistical software.

LEARNING OUTCOMES

- Design solutions using statistical methods to solve real-world data science problems. (C6, PLO7, MQF3e)
- Explain results obtained from statistical data using data story telling techniques. (A4, PLO5, MQF3c)
- Construct programs using contemporary statistical software to import and analyze datasets (C6, PLO6, MQF3d)

TOPICS

Topic 1 Introduction to Statistical Learning

Topic 2 Linear Regression

Topic 3 Classification

Topic 4 Resampling Methods

Topic 5 Linear Model Selection and Regularization

Topic 6 Polynomial Regression

Topic 7 Decision Trees

Topic 8 Support Vector Machines

Topic 9 Unsupervised Learning

Topic 10 Lab Sessions

Topic 11 Final Presentation

ASSESSMENT DETAILS

Test 1 25%

Test 2 25%

Assignment 30%

Final Presentation 20%

PRESCRIBED TEXT

Gareth, J. (2017). An Introduction to Statistical Learning with Applications in R (7th printing).
Springer Publishing

Company.