SDSC2102: STATISTICAL METHODS AND DATA ANALYSIS

Effective Term
Semester A 2024/25

Part I Course Overview

Course Title
Statistical Methods and Data Analysis

Subject Code
SDSC - Data Science
Course Number
2102

Academic Unit
Data Science (DS)

College/School
College of Computing (CC)

Course Duration
One Semester

Credit Units
3

Level
B1, B2, B3, B4 - Bachelor's Degree

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
MA2506 Probability and Statistics or MA2510 Probability and Statistics

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
Nil

Part II Course Details

Abstract

This course aims to introduce essential statistical methods and analytical tools used to analyze data, gain insights, and make informed decisions. In this course, core topics in probability and statistics are reviewed and expanded through on the study and practice of data analytics. Students will learn how to: describe and visualize data; test our understanding against data; and create statistical models based on domain knowledge. This course will let students gain hands-on experience using different statistical techniques and tools. Upon completion of this course, students should be able to think critically about data and apply standard statistical inference procedures and statistical modelling methods to draw conclusions about a real-world problemfrom such analyses.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Decide probability distribution of a random variable and calculate probabilities based on given distribution	20	x	x	
2	Discuss statistical estimation and hypothesis testing in different cases	20	x	x	
3	Compare the two cultures of statistical modeling	20	x	x	
4	Build linear regression, logistic regression and tree models	20	x	x	
5	Apply statistical inference and modelling methods to gain insights from real-world data	20	x	x	x

A1: Attitude
Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability
Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to real-life problems.

A3: Accomplishments
Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

Learning and Teaching Activities (LTAs)

	LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	In the lectures, students will learn theories and concepts on statistical methods and data analysis, and work on small-data examples.	1, 2, 3, 4, 5	3 hours/week
2	Laboratory work	In the labs, students will learn programming tools to apply statistical inference and modelling methods taught in lectures to real-world datasets	2, 4, 5	1 hour/week

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1	Homework assignments	1, 2, 4, 5	20	
2	Project	5	20	

Continuous Assessment (%)

40

Examination (%)

60

Examination Duration (Hours)

2

Additional Information for ATs

Note: To pass the course, apart from obtaining a minimum of 40% in the overall mark, a student must also obtain a minimum mark of 30% in both continuous assessment and examination components.

Assessment Rubrics (AR)

Assessment Task

Homework assignments

Criterion

Based on submitted written work to evaluate understanding of subject matter, evidence of knowledge base, capacity to analyse and synthesize, and evidence of original and critical thinking.

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)

Moderate

Marginal (D)

Basic

Failure (F)

Not even reaching marginal levels

Assessment Task

Project

Criterion

Based on project presentation and submitted project report to evaluate understanding of subject matter, evidence of knowledge base, capacity to apply data analysis methods and interpret results, and evidence of original and critical thinking.

Excellent (A+, A, A-)

High

Good (B+, B, B-)

Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Assessment Task
Examination

Criterion
Based on submitted written work to evaluate understanding of subject matter, evidence of knowledge base, capacity to analyse and synthesize, and evidence of original and critical thinking.

Excellent (A+, A, A-)
High

Good (B+, B, B-)
Significant

Fair (C+, C, C-)
Moderate

Marginal (D)
Basic

Failure (F)
Not even reaching marginal levels

Part III Other Information

Keyword Syllabus

- Probability theory and Expectations
- Common families ofprobability distributions
- Descriptive statistics
- Estimation and hypothesis testing
- Linear regression and logistic regression
- Logistic regression and generalized linear regression
- Decision trees

Reading List

Compulsory Readings

	Title
1	Applied Statistics and Probability for Engineers
2	Data Analytics: A Small Data Approach

Additional Readings

	Title
1	Nil