SDSC1001: INTRODUCTION TO DATA SCIENCE

Effective Term
Semester A 2023/24

Part I Course Overview

Course Title
Introduction to Data Science

Subject Code
SDSC - School of Data Science
Course Number
1001

Academic Unit
School of Data Science (DS)

College/School
School of Data Science (DS)

Course Duration
One Semester

Credit Units
2

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

Medium of Instruction
English

Medium of Assessment
English

Prerequisites
Nil

Precursors
Nil

Equivalent Courses
Nil

Exclusive Courses
GE1356 Introduction to Data Science

Part II Course Details

Abstract
Data is everywhere. This is the first introductory course for the first-year students without backgrounds in college mathematics or statistics or computer programming. The course aims to provide the training of the important mind-

set and unique perspective of data-driven modelling: how to identify, formulate, process and interpret the role of data and data techniques when modelling the real problems. In this course, we will describe the roles of data analytics in solving various scientific, engineering, business, societal problems, by presenting a large set of empirical case studies and successful applications across various industries. This course will be co-taught by multiple lecturers on selected topics related to the data science from the traditional applications to the most state-of-the-art industrial applications. Rather than developing abstract and specific theoretical knowledge, the focus of the course will be the overview of the data science and the participation of versatile projects and applications from real world. Through interaction with guest lecturers, pilot group projects and individual written reports and oral presentations, students will gradually build up the unique viewpoints of data resource behind many successful businesses and the understanding of the vital functions of big data in modern societies.

Course Intended Learning Outcomes (CILOs)

	CILOs	Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Identify the presence of big data in examples from daily life and business practice; elaborate the logics of why and how the use of data could change the way of solving these problems.	30	x		
2	Build and refine analytical concepts to understand the synthesis of domain knowledge and data science; Understand the pro and con of data-driven modelling approach.	30	x	x	
3	Develop the concept on the connection of data science with other technical fields such as mathematics, statistical learning, computer science, artificial intelligence.	10	x	x	
4	Develop critical thinking skills regarding to the application of data sciences and societal benefit.	30	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 self-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.

Teaching and Learning Activities (TLAs)

	TLAs	Brief Description	CILO No.	Hours/week (if applicable)
1	Lecture	Learn the development of data science during lecturing time	1, 2, 3, 4	2 hours per week
2	Tutorial	Learn basic programming skills and review the methods taught in lecture	1, 4	1 hour per week

3	Hand-in Assignment	Summary and reflection on the learning materials.	1, 2, 4	after class

Assessment Tasks / Activities (ATs)

	ATs	CILO No.	Weighting (%)	Remarks
1	Test	1, 2	10	
2	In-class discussion	2, 3, 4	20	
3	Assignments/Reports	1, 2, 4	50	

Continuous Assessment (%)

80

Examination (%)

20

Examination Duration (Hours)

1

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

Test

Criterion

Ability to explain and analyse in detail on the issues of data sciences.

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

In-class discussion

Criterion

Attitude of learning and participating the seminar topics

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
Assignments/Reports

Criterion
Ability to explain and analyse in depth on the various aspects of data sciences.

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
The examination will consist of one 1-hour paper. The questions have multiple and flexible formats to assess the general understanding of the course materials.

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

data science in a nutshell: what and why data science; fundamental principles guiding data science; impact of data science across fields;
introduction of dataset, data collection, causality ad experiments, data types, sequences of data, structure of data, sampling and randomness.
core concepts of inference and computing: computational tools and introductory statistical techniques; brief introduction of testing hypothesis, estimation, prediction, regression, classification.
tutorials on applying and illustrations of techniques and algorithms on various public datasets.
hands-on experience of handling real data and real-world issues, such as data visualization
introduction and review of recent developments and application of data sciences;

Reading List

Compulsory Readings

	Title
1	Lecture notes will be provided

Additional Readings

	Title
1	Introducing Data ScienceBy Davy Cielen, Arno D. B. Meysman, and Mohamed Ali2016, ISBN 9781633430037
2	Data Science For DummiesBy Lillian Pierson , Foreword by Jake PorwayPublisherJohn Wiley & Sons, 2017
3	Data Science for Business : What You Need to Know About Data Mining and Data-Analytic ThinkingBy Foster Provost, By Tom FawcettPublisher O'Reilly Media, Inc, USA, 2015
4	Computational and Inferential Thinking: The Foundations of Data ScienceBy Ani Adhikari and John DeNeroThis is a free online textbook that includes interactive Jupyter notebooks and public data sets for all examples. The textbook source is maintained as an open source project.
5	Data Science from Scratch: The #1 Data Science Guide for Everything A Data Scientist Needs to Know: Python, Linear Algebra, Statistics, Coding, Applications, Neural Networks, and Decision Trees.By Steven Cooper. 2018
6	Foundations of Data Scienceby Avrim Blum John Hopcroft Ravindran Kannan https://www.cs.cornell.edu/jeh/book.pdf