

**City University of Hong Kong
Course Syllabus**

**offered by School of Data Science
with effect from Semester A 2019/20**

Part I Course Overview

Course Title: Social Data Processing and Modelling

Course Code: SDSC3011

Course Duration: One Semester

Credit Units: 3

Level: B3

- Arts and Humanities
 Study of Societies, Social and Business Organisations
 Science and Technology

Proposed Area:
(for GE courses only)

Medium of Instruction: English

Medium of Assessment: English

SDSC1001 Introduction to Data Science* and SDSC2001 Python for Data Science

Prerequisites: * Pre-requisite SDSC1001 will be exempted for students who are enrolled in Minor in Data Science
(Course Code and Title)

Precursors: Nil
(Course Code and Title)

Equivalent Courses: Nil
(Course Code and Title)

Exclusive Courses: Nil
(Course Code and Title)

Part II Course Details

1. Abstract

(A 150-word description about the course)

This course provides students with an extensive exposure to the elements of data processing and modelling for social media. Topics include human error detection, missing data handling, record aggregation, data integration, categorical variable modelling, multivariate data modelling, multilevel data modelling, latent data modelling, temporal data modelling, and spatial data modelling.

2. Course Intended Learning Outcomes (CILOs)

(CILOs state what the student is expected to be able to do at the end of the course according to a given standard of performance.)

No.	CILOs [#]	Weighting* (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	A2	A3
1.	Explain clearly fundamental principles and methods of social media data processing and modelling	20%	√		
2.	Classify various properties of social media data and the corresponding modelling methods	20%	√	√	
3.	Evaluate existing practices in processing and modelling of social media data and seek ways to improve the existing practices	30%	√	√	√
4.	Apply appropriate processing/modelling methods to solve given practical problems in social media data	30%	√	√	√
		100%			

* If weighting is assigned to CILOs, they should add up to 100%.

Please specify the alignment of CILOs to the Gateway Education Programme Intended Learning outcomes (PILOs) in Section A of Annex.

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.

3. Teaching and Learning Activities (TLAs)

(TLAs designed to facilitate students' achievement of the CILOs.)

TLA	Brief Description	CILO No.				Hours/week (if applicable)
		1	2	3	4	
Lecture	Learning through teaching is primarily based on lectures.	√	√	√	√	39 hours in total
Case studies	Describe and critique classic cases of social media data processing and modelling.		√	√	√	in or after classes
Take-home assignments	Learning through in-class or take-home assignments is primarily based on interactive problem solving and hands-on exercises allowing instant feedback.		√	√	√	in or after class

4. Assessment Tasks/Activities (ATs)

(ATs are designed to assess how well the students achieve the CILOs.)

Assessment Tasks/Activities	CILO No.				Weighting*	Remarks
	1	2	3	4		
Continuous Assessment: <u>40%</u>						
Test	√	√	√	√	20-40%	Questions are designed for basic characteristics of social data to see how well the students have learned fundamental concepts and methods, and applications of social data processing.
Hands-in assignments			√	√	0-20%	These are skills based assessment to enable students to demonstrate the basic concepts and methods of social data modelling, and applications of the models in some applications.
Examination: <u>60%</u> (duration: 2 hours)	√	√	√	√	60%	Examination questions are designed to see how far students have achieved their intended learning outcomes.
					100%	

*The weightings should add up to 100%.

For a student to pass the course, at least 30% of the maximum mark for the examination must be obtained.

5. Assessment Rubrics

(Grading of student achievements is based on student performance in assessment tasks/activities with the following rubrics.)

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Test	Ability to understand and apply fundamental concepts and methods of social media data processing.	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Hands-in assignments	Ability to learn the basic concepts, apply methods and algorithms of social data modelling, and develop applications of modelling algorithms.	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Examination	Ability to solve learning tasks using social media data modelling methods.	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

(An indication of the key topics of the course.)

Human error detection, missing data imputation, data transformation, record aggregation, data integration, multivariate data modelling, multilevel data modelling, latent data modelling, temporal data modelling, and spatial data modelling

2. Reading List

2.1. Compulsory Readings

(Compulsory readings can include books, book chapters, or journal/magazine articles. There are also collections of e-books, e-journals available from the CityU Library.)

1.	Müller, H., & Freytag, J. C. (2005). Problems, methods, and challenges in comprehensive data cleansing. Professoren des Institute für Informatik.
2.	Osborne, J. W. (2013). Best practices in data cleaning: A complete guide to everything you need to do before and after collecting your data. Sage.
3.	Buttrey, S. E., & Whitaker, L. R. (2017). A Data Scientist's Guide to Acquiring, Cleaning, and Managing Data. John Wiley & Sons.

2.2. Additional Readings

(Additional references for students to learn to expand their knowledge about the subject.)