City University of Hong Kong
Course Syllabus

offered by School of Data Science
with effect from Semester B 2020/21

Part I   Course Overview

Course Title: **Introduction to Data Science**

Course Code: **GE1356**

Course Duration: **One Semester**

Credit Units: **3**

Level: **B1**

☐	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**

Prerequisites:
(Course Code and Title)   **Nil**

Precursors:
(Course Code and Title)   **Nil**

Equivalent Courses:
(Course Code and Title)   **Nil**

Exclusive Courses:
(Course Code and Title)   **SDSC1001 Introduction to Data Science**

Part II   Course Details

1. Abstract
(A 150-word description about the course)

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.

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.	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.	20%	√		
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.	20%	√	√	
3.	Build an elementary knowledge of mathematical, probabilistic and statistical foundations for data science and develop elementary programming techniques of handling simple data processing and mining.	20%	√		
4.	Develop the concept on the connection of data science with other technical fields such as mathematics, statistical learning, computer science, artificial intelligence.	20%	√	√	
5.	Develop critical thinking skills regarding to the application of data sciences and societal benefit.	20%	√	√	
	If weighting is assigned to CILOs, they should add up to 100%.	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	5	
Lecture	Learn the development of data science during lecturing time	√	√	√		√	2 hours per week
Tutorial	Learn basic programming skills and review the methods taught in lecture	√		√	√	√	1 hour per week
Hand-in Assignment	Summary and reflection on the learning materials.	√	√	√	√	√	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	5		
Continuous Assessment: <u>80%</u>							
Test/Quiz	√	√	√			10%	
In-class discussion		√		√	√	20%	
Assignments/Reports	√	√	√	√	√	50%	
Examination: <u>20%</u> (duration: 1 hours)							
Examination	√	√		√	√	20%	
*The weightings should add up to 100%.						100%	

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/Quiz	Ability to explain and analyse in detail on the issues of data sciences.	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. In-class discussion	Attitude of learning and participating the seminar topics	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Assignments/Reports	Ability to explain and analyse in depth on the various aspects of data sciences.	High	Significant	Moderate	Basic	Not even reaching marginal levels
4. Examination	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.	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.)

- 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;

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.	Lecture notes will be provided

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

1.	Introducing Data Science By Davy Cielen, Arno D. B. Meysman, and Mohamed Ali 2016, ISBN 9781633430037
2.	Data Science For Dummies By Lillian Pierson , Foreword by Jake Porway Publisher John Wiley & Sons, 2017
3.	Data Science for Business : What You Need to Know About Data Mining and Data-Analytic Thinking By Foster Provost, By Tom Fawcett Publisher O'Reilly Media, Inc, USA, 2015
4.	Computational and Inferential Thinking: The Foundations of Data Science By Ani Adhikari and John DeNero This 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 Science by Avrim Blum John Hopcroft Ravindran Kannan https://www.cs.cornell.edu/jeh/book.pdf

A. Please specify the Gateway Education Programme Intended Learning Outcomes (PILOs) that the course is aligned to and relate them to the CILOs stated in Part II, Section 2 of this form:

GE PILO	Please indicate which CILO(s) is/are related to this PILO, if any *(can be more than one CILOs in each PILO)*
PILO 1: Demonstrate the capacity for self-directed learning	1,4,5
PILO 2: Explain the basic methodologies and techniques of inquiry of the arts and humanities, social sciences, business, and science and technology	1,3,5
PILO 3: Demonstrate critical thinking skills	2,4
PILO 4: Interpret information and numerical data	1,2,3
PILO 5: Produce structured, well-organised and fluent text	1,2
PILO 6: Demonstrate effective oral communication skills	5
PILO 7: Demonstrate an ability to work effectively in a team	4,5
PILO 8: Recognise important characteristics of their own culture(s) and at least one other culture, and their impact on global issues	2
PILO 9: Value ethical and socially responsible actions	1,2,5
PILO 10: Demonstrate the attitude and/or ability to accomplish discovery and/or innovation	1,4,5

GE course leaders should cover the mandatory PILOs for the GE area (Area 1: Arts and Humanities; Area 2: Study of Societies, Social and Business Organisations; Area 3: Science and Technology) for which they have classified their course; for quality assurance purposes, they are advised to carefully consider if it is beneficial to claim any coverage of additional PILOs. General advice would be to restrict PILOs to only the essential ones. (Please refer to the curricular mapping of GE programme: http://www.cityu.edu.hk/edge/ge/faculty/curricular_mapping.htm.)

B. Please select an assessment task for collecting evidence of student achievement for quality assurance purposes. Please retain at least one sample of student achievement across a period of three years.

Selected Assessment Task
To ensure the quality assurance, the following evidence of student achievements will be connected: • The written reports as the outcome of the students' group or individual projects • The presentation slides or the posters of the students with achievements • The collection of online quiz and exam papers for all students