City University of Hong Kong
Course Syllabus

offered by Department of Management Sciences
with effect from Semester A 2018/19

Part I   Course Overview

Course Title: **Big Data Analytics**

Course Code: **MS4252**

Course Duration: **One Semester**

Credit Units: **3**

Level: **B4**

Proposed Area:
(for GE courses only)

☐ Arts and Humanities

☐ Study of Societies, Social and Business Organisations

☐ Science and Technology

**Medium of
Instruction:**   **English**

**Medium of
Assessment:**   **English**

Prerequisites:
(Course Code and Title)   **MS3251 Analytics using SAS**

Precursors:
(Course Code and Title)   **MS3252 Regression Analysis**

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

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

Part II   Course Details

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

    This course aims to:
    - Provide students with knowledge of the key concepts of big data analysis to enhance structured and unstructured data availability for enterprise strategic decision making;
    - Enable students to apply relevant knowledge for defining big data frameworks and formulating logical and physical designs for statistical analysis in business organizations;
    - Develop students' hands-on experience of construction of big data analysis using professional software packages;
    - Prepare students to demonstrate generic skills in interpersonal interaction, communication, working both individually and in teams.

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 the importance of big analytical data compared to traditional method in the efficiency and effectiveness of information extraction.	20%		✓	
2.	Assess the efficiency and effectiveness of big analytical data in business organizations	20%		✓	✓
3.	Formulate and design a statistical-oriented data for business solutions	30%		✓	✓
4.	Perform big data analysis using professional software (e.g. SAS/DIS, SAS/EM, Python, R)	30%	✓	✓	✓
	* If weighting is assigned to CILOs, they should add up to 100%.	100%			

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	Concepts and general knowledge of big data analytics using SAS or Python	✓	✓	✓				

	or R; Introduce the subject-oriented data model; Data preparation for big data analysis; Introduce big data analytics Generate a single view of data;						
Tutorial	Hands-on practice to enhance their skills in big data analytics using SAS or Python or R so that learning difficulties can be identified and tackled. Identify the business case issues regarding how to enhance the data preparation for big data analysis, design a star schema to deliver a feasible solution in single view of data to the problem identified, performing different analytics technique to analyze the big data and generate different levels of statistical reporting.	✓	✓	✓	✓		
…							

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: __60__%								
Mid-term Test	✓	✓	✓				20	
Group project	✓	✓	✓	✓			20	
Individual presentation and Q&A	✓	✓	✓	✓			20	
Examination: __40__% (duration: 3 hours , if applicable)								
Examination	✓	✓	✓				40	
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. Assignment	1.1 ABILITY to UNDERSTAND the knowledge of data mining	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Mid-term Test	2.1 ABILITY to EXPLAIN the key concepts and fundamental knowledge of data mining	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Group project and presentation	3.1 CAPACITY for COLLABORATING with students to carry out problem-based activities based on real world problems.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	3.2 ABILITY to EXPLAIN in DETAIL and with ACCURACY methods in analysing the relationship between business and sustainability solutions.	High	Significant	Moderate	Basic	Not even reaching marginal levels
	3.3 CAPACITY for SELF-DIRECTED LEARNING to find solutions to the problems and make recommendations for implementing the solutions	High	Significant	Moderate	Basic	Not even reaching marginal levels
4. Examination	4.1 ABILITY to EXPLAIN the key concepts and fundamental knowledge of data mining	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.)

    1. **Issue of big data analytics for business**

        Success factors for big data analytics, The Analysis Process, Business Point of view in big data, Analytic Complexity;

    2. **Structured and unstructured big data management**

        Unstructured and structured big data management; Probabilistic matching for unstructured data; Map Reduce and Hadoop; Analytics Process Model

    3. **Big data Analytics technique**

        Structured and unstructured (web logs, e-mails, twitter, and so on) big data for business analysis; Deriving customer segmentation measures from transactional data; Algorithm for unstructured data; Statistical modeling for solving big data problems; Information Retrieval and Web search; Vector Space Model; web analytics, text analytics, Opinion Mining and sentiment analytics, link analysis, social network analysis, similarity algorithm, expectation-maximization.

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.	EMC Education Services, 2015. Data Science & Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data. WILEY
2.	Liu, B., 2013. Web Data Mining: Exploring Hyperlinks, Contents and Usage Data, Springer.
3.	Goutam Chakraborty at el, 2013. *Text Mining and Analysis: Practice Methods, Examples, and case studies using SAS.* Cary, NC: SAS Institute Inc.
4.	Robert Blanchard at el, 2016. *Data Mining Techniques: Predictive Analytics on Big Data Course Notes.* Cary, NC: SAS Institute Inc.
5.	David Ghan, 2016. *Introduction to SAS and Hadoop: Essentials Course Notes.* Cary, NC: SAS Institute Inc.
6.	Madhavan Samir 2015. *Mastering Python for Data Science.* Packt Publishing.

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

1.	Lin, Jimmy. 2010, *Data-Intensive Text Processing with MapReduce,* Morgan & Claypool Publishers.
2.	Svolba, Gerhard. 2006. *Data preparation for analytics using SAS.* Cary, NC: SAS Institute Inc.
3.	Michael Berry, & Gordon Linoff, *Data mining techniques: For marketing, sales, and customer support*, John Wiley & Sons, 2004