COURSE INFORMATION

1 .	**Name of Course**	Introduction to Data Science
2 .	**Course Code**	TDS2101
3 .	**Type of Course** (e.g. : Core, major, elective etc.)	Specialization Core for BCS (DS) and elective for all other specializations
4 .	**Synopsis**	1. Overview of Data Science This topic provides an overview to the definition and importance of data science. It also helps to prepare students with understanding of role and responsibilities of data scientists in academia and industry. 2. Data Preprocessing, storage and analysis This topic covers methods to preprocess data and managing stored data. Students can also understand characteristics of different methods used to analyze data. 3. Basic concepts of programming language used to support data science. This topic helps students to build the fundamental understanding in using programming language for data science.
5 .	**Version** (State the date of theSenate's approval - previous and the current approval date)	Current: January 2018 Previous: June 2016
6 .	**Name(s) of Academic Staff**	Bhawani A/P S.Selvaretnam Ho Chiung Ching
7 .	**Semester and Year Offered**	Trimester 2 (Gamma)
8 .	**Credit Value**	4
9 .	**Pre-Requisite**	None

10 . **Objective of the course in the programme:**
• To equip students with fundamental concepts of data science.
• To equip students with fundamental ways to pre-process, store and analyze data.
• To equip students with basic programming language used to support data science.

11 . **Justification for including the course in the programme:**
This course prepares students with fundamental concepts of data science and programming to manage and analyze data. It is an introductory course and a pre-requisite course for selected courses offered to students who intend to specialize in Data Science.

12 . **Course Learning Outcomes (CLO)**

	Domain	Level
CLO1: Explain about concepts of data science	Cognitive	2
CLO2: Demonstrate understanding about dealing with data, data storage and data analysis.	Cognitive	3
CLO3: Demonstrate understanding about basic concepts of programming language used to support data science.	Cognitive	3

13 . **Mapping of the Course Learning Outcomes to the Programme Learning Outcomes, Teaching Methods and Assessment:**

Course Learning Outcomes (CLO) (Must tally with CLOs in item 12)	Programme Learning Outcomes (PLO)												Teaching Methods	Assessment Method
	PLO1	PLO2	PLO3	PLO4	PLO5	PLO6	PLO7	PLO8	PLO9	PLO10	PLO11	PLO12		
CLO1							✓						Lecture / Practical	Quizzes / Test / Final Exam
CLO2							✓						Lecture / Practical	Assignment / Quizzes / Test/ Final Exam
CLO3									✓				Lecture / Practical	Assignment / Quizzes / Final Exam
Total							**2**		**1**				Indicate the relevancy between the CLO and PLO by ticking "✓" the appropriate relevant box (This description must be read together with standards 2.1.2, 2.2.1, and 2.2.2 in Area 2 – pages 16 & 18 of COPPA 2.0)	

14 . **Transferable Skills:**
Critical Thinking, Time Management , Communication and Delegation

15 . **Distribution of Student Learning Time (SLT)**

	Course Content Outline	**CLO	Teaching and Learning Activities Guided Learning (F2F)*				Guided Learning (NF2F)*	Independent Learning (NF2F)*	Total SLT
			*L	*T	*P	*O			
1	**Overview** Overview of Data Science; The Needs for Data Science; Responsibilities of data scientists; Roles of Data Scientists in Academia and Industry; Data science Process.	1	2	2				4	8
2	**Dealing with Data** Overview of Data Model; Data Characterization; Data Cleaning (The Need for Data Cleaning, Data Quality); Data Integration (Problems and Solutions) ; Data Storage (Models, Scalability Basics and Challenges)	2	4	6			4	10	24
3	**Data Analysis and Getting Started with Programming** Categories of Data Analysis; Data Analysis Steps; Basic Concepts to Programming (Data Types, Operations and Structures)	3	8		6		8	14	36
4	**Data Visualization** The Need to Visualize Data; Design Considerations; Graphs and Charts	3	8		6			14	28

No.	Topic	*L	*T	*P	*O		SLT		Total SLT
5	**Ethics, security and privacy issues related to data science** Definition of ethics, security and privacy; Importance of ethics and privacy; Big Data best practices, Code of Conduct for data scientists	1	2	2			8	4	16
								Total SLT	**112**

SUMMATIVE ASSESSMENT

1. Continuous Assessment	Percentage %	Total SLT
Quizzes	10%	4
Test	20%	6
Assignments	30%	16
Total SLT for Continuous Assessment		**26**

2. Final Assessment	Percentage %	Total SLT F2F	Total SLT ILT
Final Exam	40%	2	20
Total SLT for Final Assessment (F2F + NF2F)		**22**	

Grand Total	**100%**	**160**

**Indicate the CLO based on the CLO's numbering in Item 12.
*L= Lecture, *T= Tutorial, *P= Practical, *O= Others, F2F*= Face to Face, NF2F*= Non Face to Face

16 .	**Identify Special Requirement to Deliver the Course (e.g., software, nursery, computer lab, simulation room):** R Studio / Python Software
17 .	**Main References:** EMC Education Services (Editor). (2015). Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data.
18 .	**Additional References:** O'Neil, C. & Schutt, R. (2013). Doing Data Science Straight Talk from the Frontline. O'Reilly Media. Lutz, M. (2013). Learning Python. 5th Edition. O'Reilly Media. Adler, J. (2009). R in a Nutshell: A Desktop Quick Reference.

Note:

Cells shaded light grey contain formulas / fixed values. Edit these formulas only if needed.