Subject Code	COMP5703
Subject Title	Turning Data Insights to Business Actions
Credit Value	3
Level	5
Pre-requisite / Co-requisite/ Exclusion	Nil
Objectives	The objective of this subject is to introduce students the process, techniques and technologies for managing big data so as to have a re-evaluation of business strategies and formulate new business plans.
Intended Learning Outcomes	Upon completion of the subject, students will be able to: a) understand the concept and challenge of big data and why existing technology is inadequate to analyze the big data; b) collect, manage, store, query, and analyze various form of big data; and c) understand how big data analytics can guide business decisions and strategies; and d) gain hands-on experience on large-scale analytics tools to solve some open big data problems; and
Subject Synopsis/ Indicative Syllabus	• Introduction to Big Data: The different V's, their challenges and application domains. • Collection of Big Data: eventual consistency and NoSQL systems MongoDB, Google BigTable • Large-scale data analytics systems: auto-parallel data programming; MapReduce, Hive, and parallel databases • Machine learning systems for Big Data • Basic statistical analysis, graph analytics and sentiment analysis • Data visualization: data types and dimensions; visual encoding and perception • Business Intelligence (BI)
Teaching/Learning Methodology	A mix of lectures and lab sessions is used to deliver the various topics in this subject. Lectures are conducted to initiate students with the concepts and techniques of big data. Case studies from big data domains such cybersecurity, finance, marketing, health care and education will also be discussed. Relevant online materials will be developed for students to enable blended learning and flipped classroom arrangement is considered when appropriate. Students are given the opportunity to gain hands-on experience on both open-source and commercial big data analytics software during the

<table>
<tr><td></td><td>laboratory sessions.

Relevant online materials will be developed for students to enable blended learning and flipped classroom arrangement under a SPOC (small private online class) set up is considered when appropriate.</td></tr>
<tr><td>Assessment Methods in Alignment with Intended Learning Outcomes</td><td>
<table>
<tr><td rowspan="2">Specific assessment methods/tasks</td><td rowspan="2">% weighting</td><td colspan="6">Intended subject learning outcomes to be assessed (Please tick as appropriate)</td></tr>
<tr><td>a</td><td>b</td><td>c</td><td>d</td><td></td><td></td></tr>
<tr><td>1.Class exercises, assignments and term projects</td><td>60</td><td>√</td><td>√</td><td>√</td><td>√</td><td></td><td></td></tr>
<tr><td>2. Examination</td><td>40</td><td>√</td><td>√</td><td>√</td><td></td><td></td><td></td></tr>
<tr><td>Total</td><td>100 %</td><td colspan="6"></td></tr>
</table>
</td></tr>
<tr><td>Student Study Effort Expected</td><td>
<table>
<tr><td>Class contact:</td><td></td></tr>
<tr><td>▪ Lecture</td><td>20 Hrs.</td></tr>
<tr><td>▪ Tutorials/Labs</td><td>10 Hrs.</td></tr>
<tr><td>Other student study effort:</td><td></td></tr>
<tr><td>▪ Self-study</td><td>60 Hrs.</td></tr>
<tr><td>Total student study effort</td><td>90 Hrs.</td></tr>
</table>
</td></tr>
<tr><td>Reading List and References</td><td>Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data, EMC Education Services, Jan 2015.

Jay Liebowitz (2013), Big Data and Business Analytics, Auerbach Publications.

Frank J. Ohlhorst (2012) Big Data Analytics: Turning Big Data into Big Money, Wiley.

Han, J., and Kamber, M., (2011), Data Mining: Concepts and Techniques, 3rd Ed., Morgan Kaufmann, San Francisco, CA.

Techniques, 3rd Ed., Morgan Kaufmann, San Francisco, CA.Cohen et al.“MAD Skills: New Analysis Practices for Big Data”, 2009

Dean and Ghemawat, “MapReduce: A Flexible Data Processing Tool”, Communications of the ACM, January 2010.

Rick Cattell, “Scalable SQL and NoSQL Data Stores”, SIGMOD Record, December 2010 (39:4)

Koudas, et. al. "Record Linkage: Similarity Measures and Algorithms"

Jeffrey Heer, Michael Bostock, Vadim Ogievetsky, A Tour through the</td></tr>
</table>

	Visualization Zoo, Communications of the ACM, Volume 53 Issue 6, June 2010 Howard Wen, *"Big Ethics for Big Data",* O'Reilly Media