

# IS3100: TECHNIQUES FOR BIG DATA

---

## Effective Term

Semester B 2024/25

## Part I Course Overview

### Course Title

Techniques for Big Data

### Subject Code

IS - Information Systems

### Course Number

3100

### Academic Unit

Information Systems (IS)

### College/School

College of Business (CB)

### Course Duration

One Semester

### Credit Units

3

### Level

B1, B2, B3, B4 - Bachelor's Degree

### Medium of Instruction

English

### Medium of Assessment

English

### Prerequisites

IS3230 Java Programming for Business, or  
IS2240 Python Programming for Business, or  
CB2240 Introduction to Business Programming in Python, or Other programming courses

### Precursors

Nil

### Equivalent Courses

Nil

### Exclusive Courses

Nil

## Part II Course Details

### Abstract

Nowadays, the rapid development in social media, smart devices, and the Internet of Things (IoT) resulted in an explosion of data available for analysis. The unprecedented abundance of data gives both technical challenges in processing and value creation opportunities in analytics, which together form the concept of Big Data. In this course, students will learn to appreciate the superior value of data through analytics in different applications. To handle the large amount data for analysis, new technology solutions need to be adopted. This course introduces the practical skills to process and analyse data, which can facilitate more innovative data analysis and decision making applications. The course will use Python and third party packages in data collection, preparation, preliminary modelling, and visualization. The course will also study issues related to Big Data storage and implementation using Hadoop, Spark and related technologies. Upon successful completion of this course, students should be able to better facilitate data-related applications and generate business value.

### Course Intended Learning Outcomes (CILOs)

CILOs		Weighting (if app.)	DEC-A1	DEC-A2	DEC-A3
1	Describe the challenge brought by big data and the value of combining different data to develop innovative business solutions.	35	x		
2	Apply Python packages and the tools in the big data application stack, such as Hadoop, Spark, etc.	30	x	x	
3	Formulate solutions with data mining algorithms by examples.	10	x	x	x
4	Create and run large-scale data processing demo programs.	25	x	x	x

#### 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 real-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.

### Learning and Teaching Activities (LTAs)

LTAs	Brief Description	CILO No.	Hours/week (if applicable)
1	LTA1: Lecture Students will learn the conceptualization of Big Data. Introduce different applications in the Big Data area. Teach programming techniques and the MapReduce framework. Small Exercises on programming implementation and big data techniques.	1, 2, 3	1 Hour/Week

2	LTA2: Laboratory/ Practice	Students will conduct the following activities to reinforce the knowledge of big data techniques learnt in lectures: <ul style="list-style-type: none"> <li>· Exercises: Hands-on activities in using the tools, and/or through programming.</li> <li>· Discussions: Students will be assigned self-learning tasks and they will demonstrate their work to other students.</li> </ul>	2, 3, 4	2 Hours/Week
---	-------------------------------	--	---------	--------------

**Assessment Tasks / Activities (ATs)**

ATs	CILO No.	Weighting (%)	Remarks (e.g. Parameter for GenAI use)
1 AT1: Continuous Assessment Students are encouraged to discuss and reflect on the materials covered in the lecture and attempt extra laboratory exercises.	1, 2	10	
2 AT2: Project & Tutorials Tutorials will be given to assess the students' capability to creatively apply the skills to solve problems in big data analytics. Group project will be conducted to give students more chance to deepen their understanding of the technique concepts and the programming skills introduced during the lectures and the tutorials.	1, 2, 3, 4	40	

**Continuous Assessment (%)**

50

**Examination (%)**

50

**Examination Duration (Hours)**

2

**Additional Information for ATs**

(20% Mid Term Exam; 30% Final Exam)

**Assessment Rubrics (AR)**

**Assessment Task**

AT1: Continuous Assessment

**Criterion**

CILO 1-2 Demonstrate evidence of active learning through participating in the class discussion, asking critical questions and completing extra-credit activities.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT2: Project and Tutorials

**Criterion**

CILO 1 Ability to discover and describe the big data challenges and solutions.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT2: Project and Tutorials

**Criterion**

CILO 4 Capability to run large-scale data processing applications.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT2:Project and Tutorials

**Criterion**

CILO 2-3Capability to accomplish computation and machine learning tasks with big data tools.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT2:Project and Tutorials

**Criterion**

CILO 1-2Ability to communicate model and algorithm design with oral and written forms.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT3:Examination (Mid-term and Final Examines)

**Criterion**

CILO 1 Understanding of big data opportunities and big data challenges.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT3:Examination (Mid-term and Final Examines)

**Criterion**

CILO 2-4 Capability to accomplish computation tasks and machine learning with big data tools.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

---

**Assessment Task**

AT3:Examination (Mid-term and Final Examines)

**Criterion**

CILO 1-2 Ability to communicate system and algorithm design.

**Excellent (A+, A, A-)**

High

**Good (B+, B, B-)**

Significant

**Fair (C+, C, C-)**

Moderate

**Marginal (D)**

Basic

**Failure (F)**

Not even reaching marginal levels

## Part III Other Information

### Keyword Syllabus

Big Data Concept; Big Data Applications; Python; Data Analysis Process; Data Collection; Data Modelling; Big Data Infrastructure; Distributed File System; Parallel Data Processing.

### Reading List

#### Compulsory Readings

	Title
1	Wes McKinney, Python for Data Analysis, Wes McKinney, 3e, O' Reilly, 2022.
2	Alberto Boschetti, Luca Massaron, Python Data Science Essentials, 3e, Packet Publishing, 2018.
3	Benjamin Bengfort & Jenny Kim, Data Analytics with Hadoop, O' Reilly, 2016.
4	Dirk Deroos et al., Hadoop for Dummies, Dreamtech Press, 2014.
5	Bart Baesens, Analytics in a Big Data World: The Essential Guide to Data Science and its Applications, Wiley, 2014.

#### Additional Readings

	Title
1	Erik Brynjolfsson et al., The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, W. W. Norton & Company, 2014.
2	John W. Foreman, Data Smart: Using Data Science to Transform Information into Insight, Wiley, 2013.