

# SC1015 : Course Information

## Introduction to Data Science and Artificial Intelligence

<b>Course Instructors</b>	Associate Prof Bo AN	boan@ntu.edu.sg	Office : N4-02b-55
LAMS and Reviews	Dr Sourav SEN GUPTA (Co-Ord)	sg.sourav@ntu.edu.sg	Office : N4-02b-45

<b>Lab Instructors</b>	Graduate Teaching Assistants	Note email of your Lab's TA
Lab Exercises	(will work closely with Co-Ord)	Details posted in TimeTable

### General Information

#### Description

In today's era of Information, 'Data' is the new driving force, provided we know how to extract relevant 'Intelligence'.

This course will start with the core principles of Data Science, and will equip you with the basic tools and techniques of data handling, exploratory data analysis, data visualization, data-based inference, and data-focused communication. The course will also introduce you to the fundamentals of Artificial Intelligence - state space, search problems, computer vision and natural language processing.

The course will motivate you to work closely with data and make data-driven decisions in your field of study. The course will also touch upon ethical issues in Data Science and Artificial Intelligence, and motivate you to explore the cutting-edge applications related to Big Data, Neural Networks and Deep Learning. Python will be the language of choice to introduce hands-on computational techniques.

#### Outcome

By the end of this course, you should be able to identify and define data-oriented problems and data-driven decisions in real life, discuss and illustrate the problems in terms of data exploration and visualization, apply basic machine learning tools to extract inferential information from the data, compose an engaging "data-story" to communicate the problem and the inference, outline the roles and requirements of artificial intelligence in practical applications, apply basic artificial intelligence techniques in search problems and game playing, and discuss and explain concepts of computer vision and natural language processing.



### Course Material

There is no single textbook for the course. The following books and resources will be used as references time and again.

1. **Python Data Science Handbook** : Jake VanderPlas : O'Reilly (1st edition)
2. **An Introduction to Statistical Learning** : James, Witten, Hastie, Tibshirani
3. **Artificial Intelligence: A Modern Approach** : Russell and Norvig (3rd edition)

Additional resources, if required, will be shared with you in the LAMS videos, face-to-face lectures, and example classes. Almost all of these resources will either be online (free and open source) books or online (freely available) videos/codes.

## Course Structure

### Instruction

13 Online LAMS Lectures	1 to 2 hours of Videos on NTU Learn + Quizzes	Weeks 01 to 13
+6 Review Lectures on DS	1 hour of face-to-face Review Lecture (online)	Weeks 01 to 07
+3 Review Lectures on AI	1 hour of face-to-face Review Lecture (online)	Weeks 09 to 12

### Interaction

10 Physical Lab Sessions	2 hours of Hands-on Exercises and Discussions	Weeks 01 to 12
+2 Project Consultations	Group-wise consultation time on appointment	Weeks 10 to 13

*Check the Course Calendar for a detailed schedule of the semester. Check Weekly TimeTable for your Lab timings.*

## Course Evaluation

Quizzes within LAMS	5%	Based on objective questions embedded within LAMS sequences.
Theory Quizzes	40%	2 theory quizzes will be held after Weeks 7 and 13, respectively.
Lab Exercises	25%	5 lab sessions will have submittable Exercise problems to solve.
Mini-Project	30%	8-week Group Project to be completes during the Weeks 6 to 13.

*This is a 100% Continuous Assessment (CA) course, with NO Final Examination. Graded components of the CA will be spaced out throughout the semester. You will complete all assessed components by Week 14, after the classes end.*

## Course Content

The Theory (LAMS Videos + Lectures)			In Practice (Lab Sessions)
Module 01	1 Weeks	Data-Analytic Thinking and the Data Pipeline	Basic Data Handling in Python
Module 02	2 Weeks	Basic Statistics and Exploratory Data Analysis	Statistics and EDA in Python
Module 03	1.5 Weeks	Data-driven Prediction - Fitting a Linear Model	Linear Regression in Python
Module 04	1.5 Weeks	Data-driven Classification - Using a Decision Tree	Classification Trees in Python
Module 05	1 Week	Digital Storytelling - Visualization and Dashboards	Data Dashboards in Python
Module 06	1 Week	Artificial Intelligence - Current State-of-the-Art	No Lab Session for this Module
Module 07	2 Weeks	Intelligent Agents and Search Space Solutions	Uninformed and Informed Search
Module 08	1 Week	Constraint Satisfaction and Game Playing	Game with Constrained Search
Module 09	e-Learning	Miscellaneous topics in Artificial Intelligence	No Lab Session for this Module

*Miscellaneous topics in AI may include basic introduction to cutting-edge practical domains like Computer Vision, Natural Language Processing, Reinforcement Learning, as well as Ethics in the context of Data Science and AI.*