

CORE FOUNDATIONS FOR DATA SCIENCE (DSC1010): INTRODUCTION TO ANALYSIS AND VISUALISATIONS WITH PYTHON

Course Details

1. Introduction to Data Science

There are 2.5 quintillion bytes of data created each day at our current pace, but that pace is only accelerating with the growth of the Internet of Things (IoT). And what do we do with the vast amount of data we possess? This is where Data Science come into play, where we can harness the power of data to generate valuable and insightful information for a plethora of purposes. It is a fact that the number of jobs requiring Data Science skills is growing exponentially, and hence, it is our duty to guide and empower you with the basic skills of Data Science, one step at a time.

2. Course Description

This course will expose you to the Data Science practices executed in the business world. We will explore key areas as the analytical process, how data is created, stored, accessed, computed, and how we can work with Data Science. What you learn in this course will give you a strong foundation in areas that requires analytics and will give you an edge over your peers.

You'll develop skills and a perspective that will make you more productive faster and allow you to become a valuable asset. This programme also acts as a bridging course for going deeper into advanced investigative and predictive methods, which you have an opportunity to explore in our future courses of the Data Science for Business specialization and Machine Learning.

3. Learning Objectives

Upon completion of this course(with efforts!), the student should be able to:

- Explain the big picture and core concepts of Data Science
- Understanding the role of Data Science in different businesses
- Tackle problems using computational thinking
- Basic programming
- Derive insights using visualisation

4. Materials

On our end, Singapore Data Science Academy will provide all course materials required, with the exceptions of the materials stated below.

Required:

- Personal Laptop
- Anaconda installed in laptop(refer to guide)

Supplementary:

- Data Science for Business (Very good book on Data Mining/Machine Learning for Business)
Authors: Foster Provost, Tom Fawcett, Publisher: O'Reilly Media
- Data Structure and Algorithmic Thinking with Python: Data Structure and Algorithmic Puzzle
Authors: Narasimha Karumanchi

5. Assignments & Capstone Project

Homework and Lab Assignments:

- Four homework/lab assignments to be completed individually. Each assignment will reinforce concepts related to the relevant class.
- Lab assistance is provided specifically to help you. It is not wrong to seek clarifications with your peers and TAs.
- Each lab assignment will require you to spend time on the computer. Please plan on it. Each submission assignment will contribute a certain percentage towards your final score.

Capstone Project:

- In week five, students are required to complete a capstone project.
- This project will be the most time consuming of all the homeworks, so please allocate sufficient time on that particular week!
- Completion of the project is compulsory for the eligibility of Singapore Data Science Academy's certification.

6. Tentative Course Schedule

Week	Topic/Objective	Breakdown	Due
1	❖ Have a better understanding on how Data Science plays a part in today's society	<ul style="list-style-type: none"> • Self Introduction • Introduction to Data Science • Why Python? 	
	❖ Basics of computational thinking and python	<ul style="list-style-type: none"> • Primitives, variables and types • String slicing • Operators, combination, abstraction • Booleans, Truth Tables • Week 1 Bring-home lab 	
2	❖ How to solve errors with available resources ❖ Basic sequences and loops	<ul style="list-style-type: none"> • Bring-home lab discussion • Conditionals • Sequence <ul style="list-style-type: none"> ◦ List and set • Iterations <ul style="list-style-type: none"> ◦ For, while ◦ Continue, break 	Week 1 bring home lab
	❖ Hashing values	<ul style="list-style-type: none"> • Dictionary • Nested Dictionary of lists • Week 2 Bring-home lab 	
3	❖ Conditions	<ul style="list-style-type: none"> • Bring-home lab discussion • Functions 	Week 2 bring home lab
	❖ Higher Order Functions	<ul style="list-style-type: none"> • Abstraction • Brief on first project (LTA cars dataset) • Week 3 Bring-home lab 	
4	❖ Manipulation of real datasets using simple operations ❖ Making use of mathematical libraries for analysis	<ul style="list-style-type: none"> • Bring-home lab discussion • How all of these are integral parts of data science • Introduction of libraries <ul style="list-style-type: none"> ◦ Math ◦ Statistics ◦ Numpy ◦ Pandas • Reading Data; Selecting and Filtering Data; Data manipulation, sorting, grouping and rearranging • Descriptive Statistics • Visualisations 	Week 3 bring home lab

		<ul style="list-style-type: none"> • Week 4 bring-home lab 	
5	<ul style="list-style-type: none"> ❖ Machine Learning ❖ Linear Regression 	<ul style="list-style-type: none"> • Lab Discussion • What is Machine Learning? • Machine Learning Process • Types of Machine Learning 	Week 4 bring home lab
		<ul style="list-style-type: none"> • Linear Regression • Capstone Project 	
6	<ul style="list-style-type: none"> ❖ Discussions on Capstone Project ❖ Bonus topics on advanced libraries <ul style="list-style-type: none"> ➤ Sentiment Analysis 	<ul style="list-style-type: none"> • Common mistakes • Using twitter feeds to predict sentiments of Apple's products • Example of different prediction use cases • Ending note: Takeaways and bridge to DSC1020 Machine Learning 	Capstone Project