DASC2020 Applied Linear Algebra for Data Analytics in Science (2025-26 Spring)

Course Title: Applied Linear Algebra for Data Analytics in Science

Course Code: DASC2020

1. **Instructor (s) – Name and Contact Details**

Name: Dr. PANT, Nidhi
E-mail: pantnidhi@ust.hk

Office Hour: Tuesday (3 pm - 4 pm), Room no. 4451

2. **Teaching Assistant (s) - Name and Contact Details**

Name: ZHANG Yangyang

Email: yzhangse@connect.ust.hk

Help desk: Thursday (4 pm - 5 pm), Room 4475

3. **Meeting Time and Venue – Lectures, Tutorials/ Laboratory**

Section	Date & Time	Venue
L1	Mon 15:00–16:20 Friday, 10:30–11:50	Rm 2302
T1	Friday, 16:30–17:20	Rm 1409

4. **Course Description - Credit Points, Pre-requisite, Exclusion, Brief Information/synopsis**

This course is an introduction to linear algebra with applications in least square optimization. Topics include vectors, matrices, and eigenproblems. These ideas will be applied to different applications including data fitting, machine learning, pattern recognition, and finance, etc. Numerical computation is used throughout the course as a learning tool. No previous knowledge of linear algebra is assumed.

Credit Points: 3

Pre-requisite: MATH 1014 OR MATH 1020 OR MATH 1024

5. **Intended Learning Outcomes**

(State what the student is expected to be able to do at the end of the course according to a given standard of performance)

On successful completion of this course, students are expected to be able to:

1. Explain fundamental linear algebra concepts and techniques, including vector and matrix operations, determinants, matrix inverse, decomposition, eigen decomposition, and singular value decomposition.
2. Apply linear algebra techniques to solve least squares optimization problems in machine learning, including regression, classification, inversion, constrained problems, regularization, and model validation.
3. Organize and communicate the results of linear algebra computations in the context of machine learning applications, including visualizing data and identifying patterns and trends.
4. Design and analyze machine learning algorithms including dimensionality reduction techniques using singular value decomposition.

6. **Assessment Scheme**

Homework	10%
Class exercises	15%
Quiz	15%
Midterm	25%
Final Exam	35%

7. **Criterion-Referenced Grading**

Grading in this course will be based on **criterion-referencing**, meaning students will be graded against a defined standard or rubric rather than being compared to other students.

a. The difficulties of the assessment will be considered when determining the range of overall course grade in each final grade, such that the final grade reflects the criteria that students achieved in the course. Thus, the range of overall course grades of each final grade may vary semester to semester depending on the difficulties of the assessment, including homework, quizzes and exams.

8. **Grading Rubrics**

a. **Grading rubrics for class exercises:**
   i. 0 for no, or irrelevant, or failed attempt
   ii. 1 for correct attempt but with mistakes and incorrect answers

iii. 2 for correct answer

Grading rubrics for homework, midterm and final exam questions.

Steps	Maximum Points	Criteria for Grading	Performance Levels
1. Identify Key Principles	3	Correct equations and concepts identified. -Logical reasoning and no irrelevant equations.	**Excellent (3):** All correct principles identified with no errors. **Very Good (2.5):** Minor mistakes (e.g., 1 irrelevant equation). **Good (2):** Some correct principles, with multiple minor mistakes.**Fair (1):** Many mistakes, with little relevance to the problem. **Poor (0):** No correct principles identified.
2. Apply Problem-Solving Steps	3	Correct steps followed in the algorithm. Logical flow of steps with no major errors.	**Excellent (3):** Logical steps with no errors. **Very Good (2.5):** Minor procedural errors. **Good (2):** Some correct steps, but major gaps in logic. **Fair (1):** Few correct steps or mostly irrelevant steps. **Poor (0):** No meaningful steps applied.
3. Solve Mathematical Problems	2	Correct mathematical procedures used. Relevant techniques applied, with no mistakes.	**Excellent (2):** No math errors; all intermediate steps shown. **Very Good (1.75):** Minor arithmetic or formatting errors. **Good (1.5):** Some mistakes, but major steps are correct. **Fair (1):** Frequent errors or major steps missing. **Poor (0):** No meaningful math applied.
4. Numerical Calculations	1	All required numbers are calculated correctly.	**Excellent (1):** All calculations correct and units properly used. **Very Good (0.8)**: Minor calculation or unit errors. **Good (0.6)**: Some

			calculation errors, but partial correctness. **Fair (0.4):** Frequent calculation errors or missing units. **Poor (0):** No correct calculations.
5. Interpret and Justify Results	1	Final answer is correct, clear, and justified.	**Excellent (1):** Clear, correct answer with complete justification. **Very Good (0.8)**: Mostly correct but missing small details. **Good (0.6):** Correct answer but unclear justification. **Fair (0.4):** Some justification, but reasoning is incomplete. **Poor (0):** No justification or correct results.
Total: 10 points (100 %)			

9. **Final Grade Descriptors:**

[As appropriate to the course and aligned with university standards]

Grades	**Short Description**	**Elaboration on subject grading description**
A	Excellent Performance	Students with excellent performance in the course demonstrate a strong grasp of lecture materials, critical thinking ability, and excel in various course assessments.
B	Good Performance	Students with good performance in the course demonstrate a solid grasp of lecture materials, critical thinking ability, and good performance in various course assessments.
C	Satisfactory Performance	Students with satisfactory performance demonstrate an adequate understanding of lecture materials, critical thinking ability, and satisfactory performance in various course assessments.

D	Marginal Pass	Students with a marginal pass show limited understanding of lecture materials, critical thinking ability, and minimal performance in various course assessments.
F	Fail	Students who fail the course display a lack of understanding of lecture materials, critical thinking ability and poor performance in the various course assessments

10. **DASC2020 policy on using Generative AI (GAI) tools including ChatGPT:**

a. You are encouraged to use any GAI tools in assisting your learning, including working on homeworks and tutorial exercises, just like we encourage you to work in groups with your classmates. But you MUST write up your own solution and acknowledge help you received from any source. Copying from GAI tools or any sources, or using them without proper acknowledgement is considered plagiarism.
b. In our assessments I usually expect short and concise reasoning from you based on what you learned in this course. Reasons given by GAI tools are typically long and tedious, which I will not accept.
c. No GAI tool is allowed in close-book assessments including midterm and final exams.

11. **Student Learning Resources - Lecture Notes, Readings**

Textbook:

Introduction to Applied Linear Algebra: Vectors, Matrices, and Least Squares by
Stephen Boyd, Lieven Vandenberghe
Cambridge University Press; 1st edition (August 23, 2018)

12. **Teaching and Learning Activities -**

a. Lectures:
- Lecture slides provided courtesy of the authors of the textbook, together with a modified version of Python notebooks of the Python Companion to the textbook which will be published in the course JupyterHub.
- Students are encouraged to bring their own device to class to practice with Python notebooks (No need NOT submit these notebooks).

b. Tutorials:

A tutorial session consists of two parts, both are counted towards student's final grade:

1. A **quiz** on Canvas. This will be counted in your final grade.
2. A worksheet on JupyterHub that you have to submit on or before the next tutorial. This will be counted in your final grade.

13. **Communication and Feedback**

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on assignments will include areas for improvement via canvas. Students who have further questions about the feedback including marks should consult the TA within one week after the assessment feedback is received.

14. **Course Schedule:**

Lecture	Topic	Textbook chapter
1	Vectors	1
2	Linear functions	2
3	Norm and distance	2
4	Linear independence	5
5	Matrix	6
6	Matrix Examples	7
7	Linear equations	8
8	Linear dynamic equations	9
9	Matrix multiplication	10
10	Matrix Inversion I	11
11	Matrix Inversion II	11
12	Least Square	12
13	Least Square data fitting I	13
14	Least Square data fitting II	13
15	Least Square Classification	14
16	Multi-objective Least Squares and Inversion	15.1, 15.3
17	Regularization	15.4
18	Constrained Least Squares	16
19	Portfolio Optimization	17.1
20	Clustering	4
21	Determinant and Eigenvalue Problem	
22	Diagonalization and dynamical system	
23	Matrix Decomposition	
24	Singular value decomposition (SVD) and dimensionality reduction	

The Hong Kong University of Science and Technology

Course Syllabus (2025-2026 Spring)

Object-Oriented Programming for Data Analytics in Science

DASC2110

3 credits

Pre-requisites: COMP1021

Exclusions: COMP2012 & COMP2012H

Name: Wilson Woon

Email: wilsonwoon@ust.hk

Office: Room 3433

Office Hours: By email appointment

Course Description

This course is designed to provide students with a solid understanding of Object-Oriented Programming (OOP) in Python and the skills to apply these concepts to practical data analytics scenarios. Through a structured, hands-on approach, students will progress from foundational Python programming to advanced OOP concepts and graphical user interface (GUI) development.

The course is divided into three key sections:

1. Core Python Refresher

Students will begin with a brief review of essential Python programming concepts they have previously learned. This section serves as a refresher to reinforce their knowledge and boost confidence before delving into OOP.

2. Object-Oriented Programming (OOP)

As the core focus of the course, this section introduces students to the principles and techniques of the OOP paradigm. Topics include classes, objects, inheritance, polymorphism, encapsulation, and other foundational OOP concepts. Students will also explore real-world applications of OOP in Python.

3. Graphical User Interface (GUI) Development

Students will apply their OOP knowledge to build interactive GUI applications using Python libraries. This section provides a practical understanding of how OOP principles are used in software interfaces.

In addition to these core sections, the course will cover other essential Python programming concepts, including exception handling, user input validation, menu design and implementation, and file I/O. By the end of the course, students will have the confidence and skills to design and develop data analytics applications using OOP techniques.

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Demonstrate proficiency in foundational and intermediate Python programming concepts.
2. Design and develop efficient solutions using Python Object-Oriented Programming (OOP), including the principles of classes, objects, inheritance, polymorphism, and encapsulation.
3. Apply OOP principles to solve real-world data analytics problems.
4. Build interactive Python applications with graphical user interfaces (GUIs) by leveraging OOP techniques and GUI libraries.
5. Implement robust Python programs with additional programming concepts such as exception handling, user input validation, menu-driven programming, and file I/O operations in the context of data analytics.

Assessment and Grading

This course will be assessed using criterion-referenced grading, and grades will not be assigned on a curve. Detailed rubrics for each assignment are provided in the assignment specification. Students must achieve a minimum score of 50% to obtain a marginal pass (D grade).

Assessments:

Assessment Task	Contribution to Overall Course grade (%)	Due date
Problem Set 1	10%	12 March*
Problem Set 2	10%	2 April*
Project	20%	21 April*
Class Exercises & Tutorials	20%	Various dates*
Mid-Term Examination	15%	19 March*
Final examination	25%	

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Problem Set 1	ILO1	This assignment will include several programming questions that require students to apply their basic and intermediate Python knowledge to solve data analytics problems.
Problem Set 2	ILO2, ILO3	In this assignment, students will apply OOP programming techniques to implement a simple data-collection application.
Project	ILO1, ILO2, ILO3, ILO4, ILO5	Students will develop a GUI-based data-collection application using Python OOP and intermediate-level Python programming techniques.

Class Exercises & Tutorials	ILO1, ILO2, ILO3, ILO4	Complete easy to intermediate programming exercises to reinforce understanding of course topics.
Mid-Term Examination	ILO1, ILO2, ILO3	A closed-book examination mainly covering the theoretical and practical aspects of core Python programming. Some questions on OOP will also be included.
Final Examination	ILO2, ILO3, ILO5	A closed-book examination covering the theoretical and practical aspects of intermediate Python programming and OOP.

Grading Rubrics

Please refer to the specifications of each assessment task.

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	• Demonstrate a solid understanding of key theoretical concepts. • Demonstrate a solid competency in solving problems effectively using Python Object Oriented Programming (POOP) concepts. • Demonstrate a solid competency in intermediate Python programming. • Demonstrate a solid competency in applying key concepts taught in class in data analytics workflow and applications. • Demonstrate a solid competency in designing and implementing a graphical user interface (GUI) application for scientific purposes.
B	Good Performance	The elaboration is the same as that of an A grade, but students who fall into this category demonstrate a **good** understanding and competency. In short, their outputs are not perfect, but they are good. Students may also demonstrate solid understanding or competency in only some parts of the assessment tasks.
C	Satisfactory Performance	Students are demonstrating **adequate** understanding and competence in all aspects of the assessment tasks.
D	Marginal Pass	Students are demonstrating **minimal** understanding and competence in all aspects of the assessment tasks.
F	Fail	Students are demonstrating little to no understanding and competence in all aspects of the assessment tasks.

Course AI Policy

In general, the use of Generative AI is **encouraged**. However, students must understand its limitations and drawbacks. Students are encouraged to use it to solve problems and improve their knowledge and competence. It should never be used to generate answers to assessment tasks. Generative AI is not permitted in closed-book examinations.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on these assignments may include strengths and areas for improvement. Students who have further questions about the feedback, including marks, should consult the instructor within 5 days of receiving the feedback. No appeals regarding marks will be accepted after the deadline. For a mid-term examination, a special tutorial session will be held specifically for paper checking and to conduct relevant discussions. No paper checking and marks review will be entertained after this tutorial.

The instructor may not have enough time to enter the final examination marks in Canvas. Students should take the initiative to contact the instructor via email three days after the final examination if they would like to review their final examination paper and discuss their performance. Request for final examination paper checking will not be entertained 7 days after the exam.

After the final grade has been released by the Academic Registry Office (ARO), students have 7 days to appeal their grades if they wish to do so. Please email the instructor and specify why you think there is an error in the grade awarded to you. It should be noted that a grade change will be made only if there is a clerical error or if the grading rubrics are incorrectly applied. Reasons such as "I don't think you should have taken off so many points for that error" and "I think I deserve more credits for the amount of work that I put in" are not encouraged. For further information about grade appeal, please visit Grade Review, Student Appeals and Grade Changes | HKUST - Academic Registry. No appeals regarding marks or grades will be accepted after the deadline.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Academic Integrity | HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

Reading and Learning Resources

To excel in this course and gain a strong command of Python programming and Object-Oriented Programming (OOP) concepts, students are encouraged to make full use of the provided readings and resources. These materials have been carefully curated to complement the course content and include a mix of textbooks, online articles, and instructional videos. They serve as essential tools for reinforcing classroom learning, deepening your understanding of key concepts, and developing practical programming skills.

1) Object-Oriented Programming in Python. Download the e-book from https://python-textbok.readthedocs.io/en/latest/index.html

2) A Survey of Popular Python IDEs in 2023. Available from https://www.simplilearn.com/tutorials/python-tutorial/python-ide

3) Object-Oriented Programming in Python – Explained in Plain English by Tiago Monteiro. Available from https://www.freecodecamp.org/news/object-oriented-programming-python/#where-to-go-from-here

DASC2220 Statistics and Probability for Data Analytics in Science

Course outline - 2025-2026 Spring

Credits: 3 units

Instructor: Dr. Xiaoran Wu

Email: wuxiaoran@ust.hk

Office Hours: Tue 3 to 5 pm, Rm3489

Teaching Assistant: Mr. Changlin He

Email: cheaq@connect.ust.hk

Course Objectives

This course aims to provide students with a comprehensive understanding of statistics within data analytics and equip them with the skills to apply these concepts in practical data analytics tasks.

Pre-requisite (Official)

1. COMP1021 Introduction to Computer Science **and**
2. MATH1014 Calculus II **or** MATH1020 Accelerated Calculus **or** MATH1024 Honors Calculus II

Exclusions

1. MATH 2411 Applied Statistics
2. IEDA 2520 Probability for Engineers
3. IEDA 2540 Statistics for Engineers
4. ISOM 2500 Business Statistics
5. LIFS 3150 Biostatistics

Course Intended Learning Outcomes (CILOs)

After completing this course, students should be able to

1. Demonstrate a good understanding of statistics and probability concepts in data analytics.
2. Apply statistics and probability concepts using modern software.
3. Solve real-life data analytics problems using statistical concepts and tools.

Bring Your Device (BYD)

Students are required to bring their laptops or tablets to class. It should be equipped with R or Python, a web browser and access to the Internet. Students are responsible for ensuring their devices are fully charged and functional for class activities and assignments. The instructor will provide or recommend any additional software or tools required for the course.

Teaching and Learning Format

We will be using the face-to-face teaching mode. No attendance will be taken during the lessons. Students are responsible for their academic performance.

1) Lectures will be held at the venue indicated above. Besides lectures, students may be required to perform hands-on activities.
2) Post-lesson readings & activities: Students should expect some take-home readings, exercises, and other relevant activities.
3) Assignments and Projects. They will enhance your understanding of the course topics and provide valuable experience on how statistics is applied in real-life data analytics projects and applications.

Performance Evaluation

Assessment	Percentage of Grade	CILOs assessed
Assignment	15%	1, 2, 3
Review Quizzes	10%	1, 2
Project	20%	1, 2, 3
Mid-Term Exam*	20%	1, 2
Final Exam*	35%	1, 2, 3
Total	100%	

* Closed-book examination

Readings and References

Probability and statistics for data science. Fernandez-Granda, C. (2025). Cambridge University Press.

HKUST Code of Honour

Students must be aware of University policy and regulations on honesty in academic work and of the disciplinary guidelines and procedures applicable to breaches of such policies and

regulations, as contained in https://registry.hkust.edu.hk/resource-library/academic-integrity.

Code of Conduct on the Use of Generative AI

In general, the use of Generative AI is encouraged. However, students must understand its limitations and drawbacks. Generative AI is not permitted in the mid-term and final exams. These examinations will be closed book.

Accommodation for Students with Special Needs

Students with special needs should inform the instructor of their needs at the beginning of the semester so that appropriate accommodations can be provided.

Tentative Teaching Plan

The following teaching plan is **subject to change**. Please check back from time to time.

Date	Topic
Tue, 3 Feb 2026	Course Outline & Introduction
Thu, 5 Feb 2026	Introduction to Probability
Tue, 10 Feb 2026	Introduction to Probability
Thu, 12 Feb 2026	Introduction to Probability
Tue, 17 Feb 2026	Chinese New Year Holiday (No Class)
Thu, 19 Feb 2026	Chinese New Year Holiday (No Class)
Tue, 24 Feb 2026	Modelling Data with Probability Distributions
Thu, 26 Feb 2026	Modelling Data with Probability Distributions
Tue, 3 Mar 2026	Modelling Data with Probability Distributions
Thu, 5 Mar 2026	Modelling Data with Probability Distributions
Tue, 10 Mar 2026	Review & Review Quiz 1
Thu, 12 Mar 2026	Descriptive Statistics
Tue, 17 Mar 2026	Descriptive Statistics
Thu, 19 Mar 2026	Descriptive Statistics
Tue, 24 Mar 2026	Descriptive Statistics
Thu, 26 Mar 2026	Estimation of Population Parameters
Tue, 31 Mar 2026	No Regular Class / Mid-Term Exam
Thu, 2 Apr 2026	Estimation of Population Parameters
Tue, 7 Apr 2026	Mid-Term Break (No Class)
Thu, 9 Apr 2026	Estimation of Population Parameters
Tue, 14 Apr 2026	Estimation of Population Parameters
Thu, 16 Apr 2026	Estimation of Population Parameters
Tue, 21 Apr 2026	Estimation of Population Parameters Review Quiz 2

Thu, 23 Apr 2026	Drawing Conclusions with Hypothesis Tests
Tue, 28 Apr 2026	Drawing Conclusions with Hypothesis Tests
Thu, 30 Apr 2026	Drawing Conclusions with Hypothesis Tests
Tue, 5 May 2026	Project Presentation
Thu, 7 May 2026	Project Presentation

Performance Evaluation: Assessment & Grading Rubrics (Tentative and subject to change)

Overall Grading Rubrics

Outcome	Grade A: Excellent	Grade B: Good	Grade C: Satisfactory	Grade D: Marginal Pass	Fail
CILO 1	Demonstrate solid understanding. Went above and beyond in grasping the course concepts	Demonstrate good understanding.	Demonstrate adequate understanding.	Demonstrate minimal understanding.	Poor understanding.
CILO 2	Demonstrate expert-level competency	Demonstrate a good level of competency	Demonstrate some competencies	Minimal competencies	Poor competencies
CILO 3	Demonstrate creativity and a significantly high ability to analyse and evaluate cases critically.	Demonstrate good creativity and an ability to analyse and evaluate cases critically.	Demonstrate some creativity and an ability to analyse and evaluate cases critically.	Minimal competencies.	Poor competencies

DASC 3240 Data Visualization in Science
Spring 2025-26

1. Instructor(s)

Name: Masayuki USHIO

Contact Details: Room CYT-2013 (Lift 35/36), ushio@ust.hk,

2. Meeting Time and Venue

Lectures:	**Date/Time:**	Monday and Wednesday (10:30–11:50)
	Venue:	Room 2302 (Lift 17-18)
Tutorials:	**Date/Time:**	Wednesday (17:30–18:20)
	Venue:	Room 2406 (Lift 17-18)

3. Course Description

Credit Points: 3 Pre-requisite: DASC 2220 Exclusion: COMP 4462

Data visualization is the graphical representation of applied data science. It can also provide us with a powerful way to communicate data-driven findings, motivate analyses, and detect flaws in an infographic or dashboard. This course illustrates how to use the techniques of data visualization and discovery tools to explore, visualize and analyze data. By the end of the course, students will be able to utilize tools and packages in R or Python to enhance their skills on science communication.

4. Intended Learning Outcomes

Upon successful completion of this course, students should be able to:

No.	ILOs
1	Elaborate on the fundamental concepts of exploratory and explanatory data visualization.
2	Perform data processing for visualization using various data formats in R or Python.
3	Implement effective data visualizations using R or Python by considering design principles such as color usage, layout, format, and font.
4	Interpret and critically analyze various graphs.
5	Apply visualization techniques, including static figures, interactive figures, animations, and web applications, to effectively visualize data across various scientific disciplines.

5. Course Assessments

- Quiz (20%)
- Assignments (20%)
- Final Project (60%)

Assessment Task	Contribution to Overall Course (%)	Due Date
Quiz	20%	By the end of each lecture
Assignments	20%	To be announced
Final Project	60%	17:00, 28 April 2026

6. Major References

- Hadley Wickham, Danielle Navarro, and Thomas Lin Pedersen, "*ggplot2: Elegant Graphics for Data Analysis (3rd edition)*", https://ggplot2-book.org/
- Carson Sievert, "*Interactive web-based data visualization with R, plotly, and shiny*" https://plotly-r.com/

7. Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Assignments	ILO1, ILO2, ILO3, ILO4	This task assesses students' knowledge of and ability to implement basic data visualization (ILO1, ILO2, ILO3, and ILO4).
Final Project	ILO1, ILO2, ILO3, ILO4, ILO5	This task assesses students' knowledge of and ability to implement basic data visualization (ILO1, ILO2, ILO3, and ILO4), and application of the conceptual and practical skills to interpret data (ILO5).

8. Final Grade Descriptors

Grades	Short Description	Elaboration on Subject Grading Description
A	Excellent Performance	Students demonstrate a deep and thorough understanding of the fundamental principles of data visualization, including layout, color usage, licensing, and tools for visualizing data. They are able to produce graphs that accurately and effectively represent the key features of raw data. Their code is correct, well-formatted, and easily understandable by others. They can accurately interpret graphs in any format.
B	Good Performance	Students demonstrate a strong understanding of the fundamental principles of data visualization, including layout, color usage, licensing, and tools for visualizing data. They are able to produce graphs that accurately represent the key features of raw data. Their code is generally correct and well-formatted, and it can be understood by others with careful review. They can accurately interpret graphs in most formats.
C	Satisfactory Performance	Students demonstrate a satisfactory understanding of the basic principles of data visualization, including layout, color usage, licensing, and tools for visualizing data. They are able to produce graphs that represent the key features of raw data, although they may overlook some important details. Their code is generally correct, but the graphs may not always be reproducible by others. They can accurately interpret graphs in standard formats.
D	Marginal Pass	Students demonstrate a minimal understanding of the basic principles of data visualization, including layout, color usage, licensing, and tools for visualizing data. They are able to produce graphs, but their graphs may lack some key features of the raw data. Their code may work, but it is difficult for others to read and often fails to reproduce the graphs. They can interpret graphs in common formats, but they often misinterpret graphs in less common or more complex formats.
F	Fail	Students have not met the minimum requirements for the course. They demonstrate a lack of understanding of the core concepts and tools in data visualization. They are unable to produce graphs in most formats. Their code does not work, and graphs cannot be generated. They are unable to interpret graphs in most formats.

9. Communication and Feedback

Marks for the assessments and Final Project will be communicated via Canvas within two weeks of the assessment date.

10. Course AI Policy

In this course, students are allowed to use generative AI to assist them in various ways. However, appropriate credit must be given for any use of generative AI. Additionally, students must review, analyze, and revise the output from AI to ensure it is better suited for the assessments or projects. Students must not copy, paste, and submit the output as if it is entirely their own work.

11. Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Academic Integrity – HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

DASC 3240 Spring 2025-26 Tentative Course Schedule

Lectures: Monday and Wednesday (10:30–11:50)

Venue: Room 2302 (Lift 17-18)

Wk	L#	Date	Topic	Format	Quiz/Assignments
1	1	2 Feb (Mon)	Course introduction R and RStudio – Overview and installation	Lecture + hands-on	Quiz
	Part I. Static plots				
	2	4 Feb (Wed)	Basics in figure and data presentation I	Lecture	Quiz
2	3	9 Feb (Mon)	Basics in figure and data presentation II	Lecture	Quiz
	4	11 Feb (Wed)	tidyverse and ggplot2	Lecture + hands-on	Quiz
3	5	16 Feb (Mon)	Plotting data – Basics and more	Lecture + hands-on	Quiz
		18 Feb (Wed)	Lunar New Year	No class	
4	6	23 Feb (Mon)	Scatterplot, regression, and visualization of high dimensional data	Lecture + hands-on	Quiz
	7	25 Feb (Wed)	Tips for reproducibility and transparency: File names, Git, and license issues	Lecture	Quiz
5	8	2 Mar (Mon)	Static plots: Assignments		Assignments
	Part II. Interactive plots				
	9	4 Mar (Wed)	Maps and images I: ggmap	Lecture + hands-on	Quiz
6	10	9 Mar (Mon)	Maps and images II: leaflet	Lecture + hands-on	Quiz
	11	11 Mar (Wed)	Interactive plots: plotly I	Lecture + hands-on	Quiz
7	12	16 Mar (Mon)	Interactive plots: plotly II	Lecture + hands-on	Quiz
	13	18 Mar (Wed)	Animations: gganimation	Lecture + hands-on	Quiz
8	14	23 Mar (Mon)	Interactive plots: plotly III	Lecture + hands-on	Quiz
	15	25 Mar (Wed)	Sharing plots with markdown and quarto	Lecture + hands-on	Quiz
9	16	30 Mar (Mon)	Integrating Git, GitHub and RStudio	Lecture + hands-on	
	17	1 Apr (Wed)	Interactive plots: Assignments (GitHub)		Assignments
10		6 Apr (Mon)	Mid-Term Break	No class	
		8 Apr (Wed)	Mid-Term Break	No class	
	Part III. Publishing figures using shiny application				
11	18	13 Apr (Mon)	Shiny I: Static figures	Lecture + hands-on	Quiz
	19	15 Apr (Wed)	Shiny II: Interactive figures	Lecture + hands-on	Quiz
12	20	20 Apr (Mon)	Shiny III: Shiny and markdown	Lecture + hands-on	Quiz
	21	22 Apr (Wed)	Shiny IV: Sharing applications	Lecture + hands-on	Quiz
13	22	27 Apr (Mon)	Preparation for Final Presentation		
	23	29 Apr (Wed)	Preparation for Final Presentation		
14	24	4 May (Mon)	Final presentation I		
	25	6 May (Wed)	Final presentation II		

DASC 3240 Spring 2025-26 Tentative Schedule for Tutorials

Tutorials: Wednesday (17:30–18:20)

Venue: Room 2406 (Lift 17-18)

Wk	T#	Date	Topic	
1	1	4 Feb (Wed)	R and RStudio – R basics	Hands-on
2	2	11 Feb (Wed)	tidyverse and ggplot2	Hands-on
3		18 Feb (Wed)	Lunar New Year	No Class
4	3	25 Feb (Wed)	Scatter plot and more	Hands-on
5	4	4 Mar (Wed)	Maps and images	Hands-on
6	5	11 Mar (Wed)	Plotly hands-on I	Hands-on
7	6	18 Mar (Wed)	Plotly hands-on II	Hands-on
8	7	25 Mar (Wed)	gganimation hands-on	Hands-on
9	8	1 Apr (Wed)	Markdown, Quarto, and GitHub	Hands-on
10		8 Apr (Wed)	Mid-Term Break	No Class
11	9	15 Apr (Wed)	Shiny hands-on I	Hands-on
12	10	22 Apr (Wed)	Shiny hands-on II	Hands-on
13	11	29 Apr (Wed)	Q&A	
14	12	6 May (Wed)	Q&A	

DASC 3250 Numerical Methods for Data Analytics in Science

Couse Outline (2025-26 Spring)

Instructor: Prof. Zhichao Peng (pengzhic@ust.hk)

Teaching Assistant: CHAO, Zitong (zchaoaa@connect.ust.hk)

Meeting Time and Venue:

Lectures: Monday 4:30PM - 5:50PM & Friday 12:00PM - 1:20PM, Room 2304

Tutorials: Thursday 6:00PM - 6:50PM, Room 2304

Course Description:

Credits: 3

Prerequisite: COMP1021 and one of DASC2020 / MATH2111 / MATH2121 / MATH2131 / MATH2350

Exclusions: MATH3312, MECH4740, PHYS3142

This course introduces numerical algorithms for data and scientific computing. Topics include linear systems, eigenvalues and eigenvectors, SVD/PCA, QR decomposition and least squares/regression, interpolation, nonlinear equations, numerical integration and differentiation, ODEs, basic optimization algorithms, and a brief introduction to neural networks and deep learning. Students will also learn how to apply and implement selected methods using Python.

Course Objectives and Intended Learning Outcomes:

By the end of the course, students are expected to:

1. Know basic algorithms in numerical analysis, numerical linear algebra, and data analysis.

2. Understand the basic mathematical concepts and intuition behind these algorithms.

3. Know when and how to apply and implement these algorithms in Python.

Relevant competencies include mathematical reasoning, numerical problem solving, algorithmic implementation, and application of numerical methods to data/scientific computing problems.

Teaching and learning activities:

Lectures: mathematical concepts, algorithmic intuition, worked examples, and Python demonstrations. These support ILOs 1–3.

Tutorials/laboratory sessions: problem solving, implementation details, discussion of homework/exam problems, and Python practice. These support ILOs 2–3.

Assessment and grading:

The course is assessed using criterion-referencing; grades are not assigned using a curve.

Assessment tasks:

- HW1: 4%, due Feb 28 at 11:59pm, assessing ILOs 1–3

- HW2: 4%, due Mar 23 at 11:59pm, assessing ILOs 1–3

- HW3: 4%, due May 1 at 11:59pm, assessing ILOs 1–3

- Final project: 8%, due May 17 at 11:59pm, assessing ILOs 2–3

- Mid-term exam: 20%, assessing ILOs 1–3

- Final exam: 60%, assessing ILOs 1–3

Tentative final grade descriptors:

A, score >= 90: excellent performance; comprehensive grasp of mathematical theories for numerical algorithms and strong ability to implement basic numerical methods.

B, score >= 75: good performance; good understanding of mathematical theories for numerical algorithms and ability to implement basic numerical methods.

C, score >= 60: satisfactory performance; reasonable understanding of mathematical theories for numerical algorithms and ability to implement basic numerical methods.

D, score >= 45: marginal pass; some knowledge of mathematical theories for numerical algorithms and rough ability to implement basic numerical methods.

F, otherwise: fail; insufficient understanding of mathematical theories for numerical algorithms and insufficient ability to implement basic numerical methods.

Weekly topics and readings:

Required readings are the lecture notes/Jupyter notebooks and their corresponding PDF versions. Recommended readings are:

- Gilbert Strang, Introduction to Linear Algebra

- D. E. Stewart, Numerical Analysis: A Graduate Course

Week 1: Course overview; Gaussian elimination and LU decomposition, Cholesky decomposition;

Week 2: Basic iterative methods for linear systems

Week 3: Determinants; eigenvalues and eigenvectors.

Week 4: Eigenvalues and eigenvectors II (page rank); singular value decomposition

Week 5: Principal component analysis; QR decomposition

Week 6: Least squares and regression; interpolation

Week 7: Nonlinear equations, one-dimensional case; nonlinear equations, high-dimensional case

Week 8: Numerical integration; numerical ODEs

Week 9: Optimization algorithms I; optimization algorithms II

Week 10: Optimization algorithms III; Optimization IV.

Week 11: Neural networks; Convolutional neural networks;

Week 12: Recurrent neural networks. Attention and transformers;

Week 13: Rough introduction to large language models; Course review and final examination preparation

The Hong Kong University of Science and Technology

Course Syllabus (2025-2026 Spring)

SQL for Data Analytics

DASC4020

3 credits

Pre-requisites: COMP1021

Exclusions: ISOM3260 and COMP3311

Name: Wilson Woon

Email: wilsonwoon@ust.hk

Office Hours: By email appointment

Course Description

This course is designed to provide students with a solid foundation in database design principles and the practical skills to apply these concepts in data analytics. The course is divided into three key parts:

1. Database Design Principles: Students will learn the fundamentals of database design, including Entity-Relationship (ER) modelling, the Relational Model, Normalization, and the Dimensional Model for analytical databases.
2. Structured Query Language (SQL): Basic and Advanced SQL will be taught.
3. Applications in Data Analytics: The course will culminate in applying database concepts and SQL to real-world data analytics tasks, focusing on data preparation, transformation, and exploratory analysis.

By the end of this course, students will be able to design databases, write simple and complex SQL queries, and integrate database techniques into data analytics workflows.

Intended Learning Outcomes (ILOs)

After completing this course, students should be able to

1. Describe the database design principles and produce effective database designs.
2. Produce effective SQL queries for data cleaning, retrieval, manipulation, and aggregation.
3. Apply data models in practical data analytics workflows.

Assessment and Grading

This course will be assessed using criterion-referencing, and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided in the assignment specification.

Assessments

The course will be graded using the conventional letter grading system found HERE. Students must achieve a minimum score of 50% to obtain a marginal pass (D grade).

Assessment	Percentage of Grade	ILOs assessed
Assignment 1	10%	1
Assignment 2	10%	2
Assignment 3	10%	3
Class Exercises & Tutorials	20%	1, 2, 3
Final Exam	50%	1, 2, 3
Total	100%	

Mapping of Course ILOs to Assessment Tasks

Assessed Task	**Mapped ILOs**	**Explanation**
Assignment 1	ILO1	Utilise effective and efficient data modelling concepts and normalization techniques to design databases.
Assignment 2	ILO2	Write effective SQL statements to solve data analytics problems.
Assignment 3	ILO3	Integrating key concepts learned in the course to solve a practical data analytics application.
Class Exercises & Tutorials	ILO1, ILO2, ILO3	Contains in-class and tutorial exercises meant to reinforce students' understanding of the materials taught.
Final Exam	ILO1, ILO2, ILO3	Closed-book examination covering all the course materials.

Grading Rubrics

Detailed rubrics for each assignment will be provided in the assignment specification. These rubrics clearly outline the criteria used for evaluation. Students can refer to these rubrics to understand how their work will be assessed.

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A+, A, & A-	Excellent Performance	Demonstrate a solid understanding of fundamental concepts and principles taught in the course. Able to apply suitable database concepts, techniques, and SQL statements to solve data analytics problems efficiently and effectively. Demonstrate a strong desire and commitment to do well and submit high-quality work.
B+, B, & B-	Good Performance	Demonstrate a good understanding of fundamental concepts and principles taught in the course. Able to apply database concepts, techniques, and SQL statements to solve data analytics problems efficiently and effectively. Demonstrate a good desire and commitment to do well and submit high-quality work.
C+, C, & C-	Satisfactory Performance	Demonstrate a satisfactory understanding of fundamental concepts and principles taught in the course. Able to solve data analytics problems, but the methods used can be improved significantly. Demonstrate a desire and commitment to do well and submit high-quality work.
D	Marginal Pass	Demonstrate a lack of understanding of fundamental concepts and principles taught in the course. Not able to solve a large majority of data analytics problems.
F	Fail	No understanding of data analytics and algorithms concepts, principles, and implementations. Not able to solve any of the data analytics problems. Academic dishonesty actions such as plagiarism and cheating in exams.

Course AI Policy

The use of Generative AI is generally encouraged. However, students must understand its limitations and drawbacks. Students are encouraged to use it to solve problems and improve their knowledge and competence. It should never be used to generate answers to assessment tasks. Generative AI is not permitted in examinations.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on these assignments *may* include strengths and areas for improvement. Students who have further questions about the feedback, including marks, should consult the instructor within 5 days of receiving the feedback. No appeals regarding marks will be accepted after the deadline. For a mid-term examination, a special tutorial session will be held specifically for paper checking and to conduct relevant discussions. No paper checking and marks review will be entertained after this tutorial.

The instructor may not have enough time to enter the final examination marks in Canvas. Students

should take the initiative to contact the instructor via email three days after the final examination if they would like to review their final examination paper and discuss their performance. Request for final examination paper checking will not be entertained 7 days after the exam.

After the final grade has been released by the Academic Registry Office (ARO), students have 7 days to appeal their grades if they wish to do so. Please email the instructor and specify why you think there is an error in the grade awarded to you. It should be noted that a grade change will be made only if there is a clerical error or if the grading rubrics are incorrectly applied. Reasons such as "I don't think you should have taken off so many points for that error" and "I think I deserve more credits for the amount of work that I put in" are not encouraged. For further information about grade appeal, please visit Grade Review, Student Appeals and Grade Changes | HKUST - Academic Registry. No appeals regarding marks or grades will be accepted after the deadline.

Recommended Texts and Materials

1. Modern database management by Jeffrey A. Hoffer, V. Ramesh, Heikki Topi. Pearson Education Limited. 2020 Edition.
2. Star schema: the complete reference by Christopher Adamson. McGraw-Hill. 2010 Edition.
3. The data warehouse toolkit: the definitive guide to dimensional modelling by Ralph Kimball. John Wiley & Sons. 2013 Edition

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Academic Integrity | HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

The Hong Kong University of Science and Technology

Course Syllabus (2025-2026 Spring)

Capstone Project for Data Analytics in Science

DASC4300

3 credits

Pre-requisites: For DASC Final Year students only

Name: Wilson Woon

Email: wilsonwoon@ust.hk

Office Hours: By email appointment

Course Description

This course provides final-year DASC students in the School of Science with a platform to apply and integrate their acquired knowledge, technical skills, and problem-solving abilities. By completing this capstone project in a chosen area of interest and under the supervision of a faculty member, students will gain practical experience in managing data analytics challenges.

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Identify and describe a data-intensive field.
2. Describe a suitable data collection method.
3. Produce a comprehensive dashboard for performing descriptive analytics.
4. Apply suitable data-driven models to analyse data and identify problems and opportunities.
5. Communicate data-driven insights, findings, and recommendations effectively through clear, structured, and professional oral presentations.

Assessment and Grading

This course will be assessed using criterion-referencing, and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided in the assignment specification.

Assessments

Assessment Task	Contribution to Overall Course grade (%)	Due date
Deliverable 1: Project Outline, Data Cleaning, and Description of Domain Area and Dataset	10%	8 March *
Deliverable 2: Data Collection Methodology	20%	22 March *
Deliverable 3: Project Dashboard	25%	19 April *
Deliverable 4: Data Analyses	25%	10 May *
Project Presentations	20%	19 April & 10 May *

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Deliverable 1	ILO1	First, identify an area of interest where you have sufficient domain knowledge. Next, select and clean an appropriate dataset, then describe its structure. This dataset will be used in all phases of the project and cannot be changed once selected.
Deliverable 2	ILO2	Provide a comprehensive description of how the data was collected and compiled. Discuss relevant ethical issues and limitations of the data collection.
Deliverable 3	ILO3, ILO5	Create a dashboard to track key metrics of your dataset, and then deliver an in-video presentation.
Deliverable 4	ILO4, ILO5	Analyse your dataset using suitable data analytics tools to identify problems and/or opportunities, and recommend the optimal course of action or strategy for the future. Your analysis must be data-driven and supported by strong evidence. Finally, deliver an in-video presentation.

Grading Rubrics

Detailed rubrics for each assignment will be provided in the deliverable specifications. These rubrics clearly outline the criteria used for evaluation. Students can refer to these rubrics to understand how their work will be assessed.

Final Grade Descriptors

Grades	Short Description	Elaboration on subject grading description
A+, A, & A-	Excellent Performance	Demonstrate a passionate desire to submit top-quality deliverables. Made impeccable contributions to group submissions (if any). Demonstrate excellent abilities in all the CILOs. In addition, the Project Dashboard in Deliverable 3 is visually appealing, highly interactive, adheres to standard data visualization principles, and presents important insights into the dataset and domain. Finally, Deliverable 4 is data-driven and has a significant impact. The analyses are conducted using the correct tools and methodologies. They are clear, meaningful, thorough, and supported by strong evidence. The proposals are logical, constructive, and, if possible, supported by evidence or existing practices. Innovative proposals are desired, but not compulsory.
B+, B, & B-	Good Performance	Demonstrate a strong desire to learn and a commitment to submitting good-quality or better deliverables. Made impeccable contributions to group submissions (if any). Other criteria are similar to those in the Excellent Performance category, although some are rated Excellent and others Good.
C+, C, & C-	Satisfactory Performance	Demonstrate a satisfactory desire to learn and submit good-quality deliverables. Made satisfactory contributions to group submissions (if any). Demonstrate satisfactory abilities in all the CILOs. In addition, the Dashboard in Deliverable 3 is somewhat meaningful and provides satisfactory insights into the dataset and domain. Finally, Deliverable 4 includes some data-driven analyses, but the overall output is less impactful. Students in this category mainly achieved satisfactory performances across all criteria.
D	Marginal Pass	Demonstrate an adequate desire to learn and submit quality deliverables. Made adequate contributions to group submissions (if any). Demonstrate adequate abilities in all the CILOs. Deliverables 3 and 4 are largely substandard and lack impact.
F	Fail	Demonstrates a lack of desire to submit adequate quality work in most assessment tasks. The instructor is not convinced that the student wants to do well academically. Plagiarism and other academic offences. Non-submission of both Deliverables 3 and 4.

Course AI Policy

The use of Generative AI is generally encouraged. However, students must understand its limitations and drawbacks. Proper acknowledgements about the use of this tool must be provided in each submitted assessment task.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on these assignments *may* include strengths and areas for improvement. Students who have further questions about the feedback, including marks, should consult the instructor within 5 days of receiving the feedback.

After the final grade has been released by the Academic Registry Office (ARO), students have 7 days to appeal their grades if they wish to do so. Please email the instructor and specify why you think there is an error in the grade awarded to you. It should be noted that a grade change will be made only if there is a clerical error or if the grading rubrics are incorrectly applied. Reasons such as "I don't think you should have taken off so many points for that error" and "I think I deserve more credits for the amount of work that I put in" are not encouraged. For further information about grade appeal, please visit Grade Review, Student Appeals and Grade Changes | HKUST - Academic Registry.

No appeals regarding marks or grades will be accepted after the above deadlines.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance for academic misconduct. Please refer to Academic Integrity | HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.