

Lingnan University
Course Syllabus

Course Title	:	Mathematical Fundamentals for Data Science
Course Code	:	CDS1004
Recommended Study Year	:	1
No. of Credits/Term	:	3
Mode of Tuition	:	Sectional Approach
Class Contact Hours	:	2-hour lecture and 1-hour tutorial per week
Category	:	Required course
Discipline	:	-
Prerequisite(s)	:	Nil
Co-requisite(s)	:	Nil
Exclusion(s)	:	Nil
Exemption Requirement(s)	:	Nil

Brief Course Description:

The Mathematical Fundamentals for Data Science course offers a comprehensive introduction to the essential mathematical foundations required for data science. The course covers fundamental concepts such as sets, functions, linear algebra and calculus, while delving into probability theory, descriptive and inferential statistics, and common probability distributions,. By the end of the course, students will have the skills necessary to apply these mathematical principles to real-world data analysis scenarios, making them well-prepared for careers in data science, analytics, or any field requiring mathematical reasoning.

Aims:

This course aims:

1. To provide a solid foundation in essential mathematical concepts: The unit aims to ensure that students have a strong understanding of fundamental mathematical principles, including sets, functions, basic calculus and linear algebra, which are crucial for understanding more advanced topics in data science.
2. To develop proficiency in probability and statistics: The unit aims to help students gain a thorough understanding of basic probability theory, descriptive and inferential statistics, and common probability distributions. Students will learn to calculate probabilities, draw conclusions from sample data, and interpret statistical results in various contexts.
3. To foster critical thinking and problem-solving skills: The unit aims to develop students' ability to apply mathematical principles to real-world data analysis scenarios, encouraging critical thinking and problem-solving skills. Through hands-on exercises and case studies, students will learn to analyse data, draw conclusions, and communicate their findings effectively.

4. To prepare students for careers in data science and analytics: The unit aims to equip students with the mathematical tools and skills necessary for successful careers in data science, analytics, or any field requiring mathematical reasoning. By the end of the unit, students will have a solid understanding of the mathematical foundations of data science, enabling them to make informed decisions and analyse data effectively.

Learning Outcomes:

On completion of the course, student will be able to:

1. Demonstrate a comprehensive understanding of fundamental mathematical concepts, including sets, functions, basic calculus and linear algebra, and apply these principles to solve problems in data science.
2. Calculate and interpret probabilities, understand common probability distributions, and apply these concepts to real-world scenarios.
3. Perform descriptive and inferential statistics, including measures of central tendency and variability, and draw conclusions from sample data.
4. Analyse and interpret data using appropriate mathematical methods and communicate their findings clearly and effectively.
5. Apply the mathematical principles learned in the course to real-world data analysis scenarios, demonstrating critical thinking and problem-solving skills.

Indicative Contents:

- **Sets and Functions:** Understanding the mathematical concepts of sets and functions is fundamental to grasping algorithms and data manipulation in machine learning.
- **Basic Calculus:** Fundamental concept of derivatives and integrals, and their use in machine learning.
- **Basic Linear Algebra:** Linear spaces, vectors, matrices, matrix operations, bases, matrix transformations, projections, and their use in data science applications.
- **Probability and Statistical Inference:** These concepts are essential for making predictions and drawing conclusions from data.
- **Software Tools:** Tools like Python and Google Colab are invaluable for implementing advanced data science algorithms. Students will get a headstart on using these tools to prepare them for higher-level courses.

Teaching Method:

The teaching method for this course emphasizes a blend of theoretical instruction and practical application, incorporating traditional lecture-based instruction to present fundamental concepts, interactive learning activities to engage students and reinforce concepts, hands-on projects that require students to apply mathematical and statistical principles to real-world data analysis scenarios, the use of technology for data analysis and visualization, and regular assessments to monitor progress and provide feedback. This approach aims to provide students with a comprehensive understanding of mathematics and statistics in data science while fostering active participation, critical thinking, and problem-solving skills.

Students will meet every week for:

1. A two-hour lecture session
2. A one-hour tutorial session

Measurement of Learning Outcomes:

	Class Attendance and Participation	Assignments	Midterm Examination	Final Examination
1. Demonstrate a comprehensive understanding of fundamental mathematical concepts, including sets, functions, basic calculus and linear algebra, and apply these principles to solve problems in data science.		X	X	X
2. Calculate and interpret probabilities, understand common probability distributions, and apply these concepts to real-world scenarios.		X		X
3. Perform descriptive and inferential statistics, including measures of central tendency and variability, and draw conclusions from sample data.		X		X

4. Analyze and interpret data using appropriate mathematical methods, and communicate their findings clearly and effectively.		X	X	X
5. Apply the mathematical principles learned in the course to real-world data analysis scenarios, demonstrating critical thinking and problem-solving skills.	X	X	X	X

- Class Attendance and Participation: Measure student engagement, communication skills, and ability to contribute to discussions, indirectly supporting all learning outcomes, especially LO5.
- **Homework assignments**: Students will be regularly assigned problem sets to check their understanding of concepts they learned in class, apply them to real-world scenarios and reflect on their learning.
- Midterm and Final Examinations: Comprehensive evaluation of student learning, assessing understanding and application of course material across all learning outcomes.

Assessment:

Class Attendance and Participation	10%
Assignments	40%
Midterm Examination	20%
Final Examination	30%
Total	100%

Required/Essential Readings:

1. Strang, G. (2022). *Introduction to linear algebra*. Wellesley-Cambridge Press.
2. Stewart, J., Clegg, D., & Watson, S. (2021). *Calculus: early transcendentals*. Cengage.
3. Hogg, R. V., Tanis, E. A., & Zimmerman, D. L. (2020). *Probability and statistical inference*. Hoboken.

Recommended/Supplementary Readings:

1. Nield, T. (2022). *Essential Math for Data Science: Take Control of your data with fundamental linear algebra, probability, and statistics*. O'Reilly Media, Inc.
2. Strang, G. (2019). *Linear algebra and learning from data* (Vol. 4). Cambridge: Wellesley-Cambridge Press.
3. Cohen, M. X. (2022). *Practical linear algebra for data science*. O'Reilly Media, Inc.
4. Bruce, P., Bruce, A., & Gedeck, P. (2020). *Practical statistics for data scientists: 50+ essential concepts using R and Python*. O'Reilly Media, Inc.

Important Notes:

- (1) Students are expected to spend a total of 9 hours (i.e. 3 hours of class contact and 6 hours of personal study) per week to achieve the course learning outcomes.
- (2) Students shall be aware of the University regulations about dishonest practice in course work, tests and examinations, and the possible consequences as stipulated in the Regulations Governing University Examinations and Course Work. In particular, plagiarism, being a kind of dishonest practice, is “the presentation of another person’s work without proper acknowledgement of the source, including exact phrases, or summarised ideas, or even footnotes/citations, whether protected by copyright or not, as the student’s own work”. Students are required to strictly follow university regulations governing academic integrity and honesty.
- (3) Students are required to submit writing assignment(s) using Turnitin.
- (4) To enhance students’ understanding of plagiarism, a mini-course “Online Tutorial on Plagiarism Awareness” is available on <https://pla.ln.edu.hk/>.
- (5) Students must adhere to the University’s guidelines and practices when using Generative Artificial Intelligence (AI) tools. The official documents “Guidelines for Using GAI Tools at Lingnan University” and “Best Practices for Ethical and Responsible Use of GAI Tools in Course Assessments” can be found here: <https://www.ln.edu.hk/cital/generative-artificial-intelligence/gai-guidelines-best-practices>.

Rubric for Midterm and Final Examinations (50%) of CDS1004 – Mathematical Fundamentals for Data Science

Criteria	Excellent (A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Poor (D+, D)	Fail (F)
Foundational Mathematics: Sets and Functions, Basic Calculus and Linear Algebra	Demonstrates an exceptional understanding of mathematical concepts and theories, including complex operations and relationships. Can apply these concepts in diverse and challenging problem-solving scenarios with ease and accuracy.	Shows a strong grasp of mathematical concepts and theories, with the ability to apply them correctly in most situations. Occasionally encounters difficulties with more complex or abstract concepts.	Has a basic understanding of mathematical concepts and theories and can apply them in straightforward situations. However, there may be some confusion or errors when dealing with more complex or nuanced concepts.	Struggles with mathematical concepts and theories, often making errors or misunderstandings in their application. May require significant guidance or support to understand and apply these concepts.	Lacks a fundamental understanding of mathematical concepts and theories, struggling to apply them even in straightforward situations. May require extensive support and remediation to develop basic proficiency.
Probability and Statistical Inference	Demonstrates an exceptional understanding of probability and statistical inference concepts, including advanced techniques and problem-solving strategies. Can apply these concepts in diverse and challenging scenarios with ease and accuracy.	Shows a strong grasp of probability and statistical inference concepts, with the ability to apply them correctly in most situations. Occasionally encounters difficulties with more advanced or abstract concepts, but generally performs well.	Has a basic understanding of probability and statistical inference concepts and can apply them in straightforward situations. However, there may be some confusion or errors when dealing with more complex or nuanced concepts.	Struggles with basic probability and statistical inference concepts, often making errors or misunderstandings in their application. May require significant guidance or support to understand and apply these concepts.	Lacks a fundamental understanding of probability and statistical inference concepts, struggling to apply them even in straightforward situations. May require extensive support and remediation to develop basic proficiency.
Software Tools	Demonstrates exceptional proficiency in using various statistical and data analysis software tools, including advanced features and functionalities. Can apply complex techniques and tools with ease and accuracy, drawing meaningful conclusions from diverse datasets.	Shows strong proficiency in using most statistical and data analysis software tools, correctly applying relevant techniques and tools to analyze data. Occasionally encounters difficulties with more advanced or nuanced features, but generally performs well.	Has a basic understanding of statistical and data analysis software tools and can use them to perform straightforward analyses. However, there may be some confusion or errors when dealing with more complex or nuanced features.	Struggles with basic software tools, often making errors or misunderstandings in their application. May require significant guidance or support to use software tools effectively.	Lacks a fundamental understanding of statistical and data analysis software tools, struggling to use them even in straightforward situations. May require extensive support and remediation to develop basic proficiency.

Rubric for Assignment (40%) of CDS1004 – Mathematical Fundamentals for Data Science

Criteria	Excellent (A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Poor (D+, D)	Fail (F)
Data Analysis Techniques	Demonstrates thorough understanding and application of advanced data analysis techniques, effectively drawing meaningful insights from complex datasets.	Applies appropriate data analysis techniques to draw reasonable conclusions from datasets, with minor errors or limitations.	Demonstrates basic understanding of data analysis techniques, but struggles to apply them correctly or draw meaningful insights.	Shows limited understanding of data analysis techniques, often applying them incorrectly or drawing incorrect conclusions.	Demonstrates no understanding or incorrect application of data analysis techniques.
Algorithm Implementation	Implements algorithms efficiently and accurately, with a clear understanding of underlying concepts and optimization techniques.	Implements algorithms correctly, with some understanding of underlying concepts and optimization, but with minor errors or inefficiencies.	Attempts to implement algorithms but struggles with underlying concepts or makes significant errors.	Shows limited understanding of algorithm implementation, often struggling to implement correctly or missing key concepts.	Demonstrates no understanding or incorrect implementation of algorithms.
Statistical Inferences	Draws accurate and meaningful statistical inferences from data, with a deep understanding of underlying concepts and principles.	Draws reasonable statistical inferences from data, with some understanding of underlying concepts, but with minor errors or limitations.	Makes basic statistical inferences but struggles with underlying concepts or draws incorrect conclusions.	Shows limited understanding of statistical inferences, often drawing incorrect conclusions or missing key concepts.	Demonstrates no understanding or incorrect drawing of statistical inferences.
Evaluating Assumptions and Biases	Critically evaluates assumptions and biases in data and analysis, demonstrating a deep understanding of potential sources of error and methods to mitigate them.	Evaluates assumptions and biases with some understanding of potential sources of error, but with minor limitations or oversights.	Attempts to evaluate assumptions and biases but struggles with underlying concepts or makes significant errors.	Shows limited understanding of evaluating assumptions and biases, often missing key sources of error or mitigation methods.	Demonstrates no understanding or incorrect evaluation of assumptions and biases.

Rubric for Class Attendance and Participation (10%) of CDS1004 – Mathematical Fundamentals for Data Science

Criteria	Excellent (A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Poor (D+, D)	Fail (F)
Class Attendance	Attends all classes and is punctual.	Attends all classes but not punctual.	Misses ≤ 1 classes per semester.	Misses ≤ 2 classes per semester.	Misses > 3 classes per semester.
Preparedness	Always comes to class prepared with necessary materials.	Comes to class prepared with necessary materials most of the time.	Comes to class prepared with necessary materials some of the time.	Often comes to class unprepared.	Never comes to class prepared.
Participation	Actively engages in class discussions and contributes valuable insights.	Participates in class discussions and offers relevant input.	Occasionally participates in class discussions.	Rarely participates in class discussions.	Never participates in class discussions.
Engagement	Shows enthusiasm and a willingness to learn.	Shows a generally positive attitude towards learning.	Shows a somewhat positive attitude towards learning.	Shows a lack of enthusiasm for learning.	Shows a negative attitude towards learning.
Respectfulness (Listening Attentively, Polite Interruptions, Respecting Differences)	Treats classmates and instructor with respect and courtesy.	Generally treats classmates and instructor with respect.	Occasionally treats classmates or instructor with disrespect.	Frequently treats classmates or instructor with disrespect.	Always treats classmates and instructor with disrespect.