SEMESTER STUDY PLAN
DATA SCIENCE
(ELECTIVE COURSES)
(Case-Based Method)

DEPARTMENT OF MATHEMATICS AND DATA SCIENCE
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
UNIVERSITAS ANDALAS
2024

	SEMESTER STUDY PLAN (SSP) STUDY PROGRAM: BACHELOR OF MATHEMATICS FACULTY OF MATHEMATICS AND NATURAL SCIENCES UNIVERSITAS ANDALAS

SEMESTER STUDY PLAN

Course Name	Course Code	*I-Learn* URL	Credits	Semester	Compilation Date
Data Science	MAT62254	https://sci.ilearn.unand.ac.id	3	6	12 May 2024

Person in Charge	Study Plan Creator	Head of Research Group	Head of the study program
	Prof. Dr. Ferra Yanuar, M.Sc Yudiantri Asdi, M.Sc	Yudiantri Asdi, M.Sc	Dr. Noverina Alfiany

Intended Learning Outcomes (ILO) and Performance Indicator (PI)	Intended Learning Outcome	
	ILO-2	Possesses profound knowledge of the basic concept mathematics PI-1: An ability to explain basic mathematical concepts PI-2: An ability to provide examples that are relevant to basic mathematical concepts PI-3: An ability to determine solutions to simple problems using basic mathematical concepts
	ILO-3	An ability to identify, explain and generalize simple mathematical PI-1: An ability to identify simple mathematical problems PI-2: An ability to explain simple mathematical problems PI-3: An ability to generalize simple mathematical problems
	ILO-4	An ability to use concept and fundamental technique of mathematics in solving simple mathematical problems PI-1: An ability to choose appropriate basic mathematical concepts and techniques in solving simple mathematical problems PI-2: An ability to illustrate simple mathematical problems based on appropriate basic mathematical concepts and techniques PI-3: An ability to solve simple mathematical problems using appropriate basic mathematical concepts and techniques

	ILO-5	An ability formally and correctly proves a simple mathematical statements using facts and methods that have been studied. PI-1: An ability to identify formal structures and analogous forms in mathematics PI-2: An ability to use facts and apply methods to prove simple mathematical statements
	ILO-6	Have ability data literacy and technology and can apply them in solving simple mathematical problems or other relevant fields PI-1: An ability to identify the right data and technology to solve simple mathematical problems or other fields PI-2: An ability to use data and technology and apply them to solve simple mathematical statements or other areas PI-3: An ability to process data using available technology in simple mathematical problems or other fields
	ILO-7	An ability to communicate effectively especially in the area of mathematics in with diverse communities PI-1: An ability to convey ideas or study results orally, especially in the field of mathematics PI-2: An ability to present ideas or study results in writing, especially in the field of mathematics PI-3: An ability to respond to feedback given
	Course Learning Outcome	
	1	Mastering the understanding and terminology in data science (**ILO-2**, **ILO-3** : PI-3, **ILO-4**:PI-1, **ILO-5** , **ILO-6**)
	2	Master how to visualize data , calculate statistics from data , perform *basic* SQL queries , and the basics of Machine Learning (**ILO-2**, **ILO-3**: PI-3, **ILO-4**: PI-1, **ILO-5**, **ILO-6**)
	3	Mastering simple & multiple linear regression analysis with Python (**ILO-2**, **ILO-3** : PI-3, **ILO-4** : PI-1, PI-2, **ILO-5**:PI-1, **ILO-6**)
	4	Mastering the Classification method with Logistic Regression (ILO-2, ILO-3 : PI-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1, **ILO-6**)
	5	Mastering the Clustering method (K-Means & Hierarchical Clustering) (**ILO-2**, **ILO-3** : PI-3, **ILO-4** : PI-2, **ILO-5**:PI-1, **ILO-6**)
	6	An ability to create and present research reports (**ILO-2**, **ILO-3** : PI-3, **ILO-4** : PI-1, PI-2, **ILO-5**:PI-1, **ILO- 6**, **ILO-7**)
Brief Description		This course , which applies the CBM (*Case Based Method*) learning method , discusses the basic principles, tools and methods used in data science & analytics in extracting information or knowledge from data. The principles and methods discussed are based on various fields of science, including: statistics, probability, databases, machine

	learning. Important concepts discussed include data exploration analysis, information extraction, data visualization, regression, classification, clustering . This course emphasizes the integration and synthesis of the principles and methods provided to be applied in solving problems.
Course Materials	1. Understanding the terminology in data science . 2. Visualization of data and calculating statistics from data 3. Perform *basic* SQL queries to pull data from *the database* 4. Get to know the basics of Machine Learning . 5. A simple linear regression analysis with phyton . 6. Logistic Regression Classification Method using phyton 7. K-Means Clustering Method: Algorithm and Examples in phyton
References	**Main :** 1. Shah, Chirag. 2020. A Hands-On Introduction to Data Science. Cambridge University Press,. 2. Wickham, Hadley and Garret Grolemund. 2017. R for Data Science. O'Reilly Media, Inc.. **Additional:** 1. Forsyth, David. 2018. Probability and Statistics for Computer Science. Springer International Publishing AG.
Learning Media	**Software :** Zoom, Software related projects such as R and Python — **Hardware :** Computer/Laptop and LCD Projector
Team Teaching	Prof. Dr. Ferra Yanuar , M.Sc Yudiantri Asdi, M.Sc
Assessment	Assignment, midterm exam, mini research, final exam
Required courses	Calculus Elementary Linear Algebra Data Analysis Mathematical Statistics I

Weekly Study Plan

WEEK TO - (1)	CPMK / SUB-CP-MK (2)	ASSESSMENT INDICATORS (3)	FORM OF ASSESSMENT (4)	LEARNING ACTIVITIES [ESTIMATED TIME] (5)				MEDIA (9)	LEARNING MATERIALS / STUDY MATERIALS (10)	WEIGHT (11)
				Synchronous		Asynchronous				
				Face to Face Offline (5)	Face to Face Online (6)	Independent (7)	Collaborative (8)			
1-2	CLO 1: Mastering the understanding and terminology in data science (ILO-2, ILO-3 : PI-3, ILO-4:PI-1, ILO-5)	● Ability know principal discussion, method learning, achievements learning, reference And evaluation ● Ability understand about understanding and terminology in data science	Assignment (5%) Mini research progress report (2.5%)		● Studying : - Introduction to Study Plan - discussion and question and answer course material [2 x 3 x 50 minutes]	Students look for references related to understanding and terminology in data science [2 x 3 x 60 minutes]	Students discuss in teams regarding understanding and terminology in data science [2 x 3 x 60 minutes]	• I-learn UNAND • Zoom	● Study Plan ● Understanding and terminology in data science	7.5%
3-7	CLO-2: Mastering how to visualize data , calculate statistics from	● Accuracy in data visualization , calculating	Assignment (10%)	● Studying : - concept explanation		Students look for references	Students discuss in teams related to data	LMS (ilearn UNAND)	● Data visualization , ● calculate the k statistic from	12.5%

	data , perform *basic* SQL queries , and the basics of Machine Learning . (ILO-2, ILO-3: PI-3, ILO-4: PI-1, ILO-5)	statistics from data , performing *basic* SQL queries , and the basics of Machine Learning .	Mini research progress report (2.5%)	- discussion and question and answer course material [5 x 3 x 50 minutes]		[5 x 3 x 60 minutes]	visualization , calculating statistics from data , carrying out *basic* SQL queries , and the basics of Machine Learning . [5 x 3 x 60 minutes]		the data , ● *basic* SQL queries , ● Machine Learning basics .	
8 (Midterm)	CLO 1: Mastering the understanding and terminology in data science (ILO-2, ILO-3 : PI-3, ILO-4:PI-1, ILO-5)	● Accuracy in explaining meaning and terminology in data science.	Midterm scores (5%)	-	-	-	-	LMS (ilearn UNAND)	● Midterm Answers	5%
	CLO-2: Mastering how to visualize data , calculate statistics from data , perform *basic* SQL queries , and the basics of Machine Learning . (ILO-2, ILO-3: PI-3, ILO-4: PI-1, ILO-5)	• Accuracy in visualizing data , calculating statistics from data , carrying out *basic* SQL queries , and the basics of Machine Learning .	Midterm scores (10 %)	-	-	-	-	ILearn (UNAND)	● Midterm Answers	10%

9-11	CLO-3: Mastering simple and multiple linear regression analysis methods with Python (ILO-2, ILO-3 : PI-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1)	● Ability to perform simple and multiple linear regression analysis with Python	Mini research progress report (15 %)	- Review and discussion of project design alternatives and evaluation results - Discussion of the project data collection process [3 x 3 x 50 minutes]		Student propose alternative project designs along with selected case data and project implementation schedule. [3 x 3 x 60 minutes]	Students discuss in teams regarding simple linear regression analysis with Python [3 x 3 x 60 minutes]	LMS (ilearn UNAND)	● Simple and multiple linear regression analysis	15%
1 2 -13	CLO- 4 : Mastering the implementation of the Classification method using Logistic Regression (ILO-2, ILO-3 : PI-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1)	● Ability implementing the Classification method with Logistic Regression	Mini research progress report (15 %)	Discussion of the best design alternatives between lecturers and students [2 x 3 x 50 minutes]		Students carry out final design activities based on the results of team discussions [2 x 3 x 60 minutes]	Students discuss in teams to determine the best design alternative [2 x 3 x 60 minutes]	● ILearn (UNA ND)	● Classification Method with Logistic Regression	15%
1 4-15	CLO- 5 : Mastering the Clustering method (K-Means & Hierarchical Clustering)	● Student accuracy in applying the Clustering method (K-	Mini research progress report (15 %)	Discussion of the best design alternatives between lecturers		Students revise the written mini research progress report [2 x 3 x 60 minutes]	Students discuss in teams to improve the mini research progress	● ILearn (UNA ND)	● Clustering method using the K-Means method ● Clustering method using	15%

	(ILO-2, ILO-3 : PI-3, ILO-4 : PI-2, ILO-5:PI-1)	Means & Hierarchical Clustering)		and students [2 x 3 x 50 minutes]			report based on input during the discussion [2 x 3 x 60 minutes		the Hierarchical Clustering method)	
16 (Final Exam)	CLO-6: Able to write research reports (ILO-2, ILO-3 : ILO-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1, ILO-7:P1-2)	● Student accuracy in writing Mini research report	Mini research report (15%)	Compilation of Final Mini Research Report						15%
				Total Weight						**100%**

1.1. Indicators, Criteria and Assessment Weights

1. **Assessment Weight** :

   - Assignment   : 20 %
   - Midterm exam   : 15 %
   - Mini research progress report   : 50%
   - Final Exam (mini research report)   : 15%

2. **Assessment Weight for each Course Learning Outcome**

   - CLO 1: 12.5%
   - CLO 2 : 27.5 %
   - CLO 3: 15%

- CLO 4: 15 %

- CLO 5: 15 %

- CLO 6 : 15%

1.2. Plan Assessment

Assessment Plan Table

Form of Assessment Course Learning Outcomes	Assignment	Midterm	Mini research progress report	Final Exam	Total Weight
1. Mastering the understanding and terminology in data science .	5%	5%	2.5%		**12 .5%**
2. Master how to visualize data , calculate statistics from data , perform *basic* SQL queries , and the basics of Machine Learning	10 %	10 %	2.5%		**22.5 %**
3. simple and multiple linear regression analysis methods with Python			1 5%		**15%**

Form of Assessment Course Learning Outcomes	Assignment	Midterm	Mini research progress report	Final Exam	Total Weight
4. Mastering the implementation of the Classification method with Logistic Regression			15 %		**15 %**
5. Mastering the Clustering method (K-Means & Hierarchical Clustering)			15 %		**15 %**
6. Able to write research reports				20 %	**20 %**
Total Weight	**15 %**	**15 %**	**50%**	**20 %**	**100%**

1.4. Mini Research Report Format

STUDY PROGRAM	Mathematics (Bachelor Degree)
FACULTY	MIPA
COLLEGE	Andalas University
COURSE / CODE	Data Science / MAT622 54
LECTURER	Prof. Dr. Ferra Yanuar, M.Sc
TEAM TEACHING	Prof. Dr. Ferra Yanuar, M.Sc Yudiantri Asdi, M.Sc

CASE STUDY MEMBER	Group of 4 people
ASSIGNMENT THEME	Application of Methods (Multiple Linear Regression Analysis / Logistic Regression / K-Means* Clustering) in Case Studies

COURSE LEARNING OUTCOMES to be achieved from this Mini Research
CLO
1. Mastering the understanding and terminology in data science (ILO-2, ILO-3 : PI-3, ILO-4:PI-1, ILO-5 , ILO-6)
2. Master how to visualize data , calculate statistics from data , perform *basic* SQL queries , and the basics of Machine Learning (ILO-2, ILO-3: PI-3, ILO-4: PI-1, ILO-5, ILO-6)
3. Mastering simple & multiple linear regression analysis with Python (ILO-2, ILO-3 : PI-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1, ILO-6)
4. Mastering the Classification method with Logistic Regression (ILO-2, ILO-3 : PI-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1, ILO-6)
5. Mastering the Clustering method (K-Means & Hierarchical Clustering) (ILO-2, ILO-3 : PI-3, ILO-4 : PI-2, ILO-5:PI-1, ILO-6)
6. Able to create and present research reports (ILO-2, ILO-3 : PI-3, ILO-4 : PI-1, PI-2, ILO-5:PI-1, ILO- 6, ILO-7)
MINI RESEARCH DESCRIPTION
In this mini research , each group consisting of 4 students is asked to apply one of the statistical data analysis methods: multiple linear regression analysis, classification analysis, or cluster analysis. The statistical data analysis method is applied to empirical data chosen by the students themselves. Please carry out a complete data analysis by determining the best model that is adapted to the selected analysis method so that a logical and theoretically acceptable conjecture /clusterization/classification model is produced.

Case studies are written in the form of Mini Research Reports . |
| MINI RESEARCH STEPS |
| 1. Identification and formulation of problems. 2. Mini Research work schedule 3. Implementation of Mini Research: 4. Monitoring and evaluation during the implementation of Mini Research by lecturers |

including the obstacles and challenges faced .
5.   Mini Research final report .
6.   L report presentation end of Mini Research.
7.   Final analysis and evaluation of Mini Research.

OUTPUT FORM AND FORMAT

REPORTS AND OTHER RELEVANT FORMS

INDICATORS, CRITERIA AND ASSESSMENT WEIGHTS

• Assignment : 20 %
• Midterm Exam : 15 %
• Mini research progress report : 50%
• Final exam (mini research report) : 15%

Assessment Weight for Each Learning Outcome
• CLO 1: 12.5%
• CLO 2 : 27.5 %
• CLO 3: 15%
• CLO 4: 15 %
• CLO 5: 15 %
• CLO 6 : 15%

IMPLEMENTATION SCHEDULE (synchronized with the week in the Study Plan)

Duration: 16 weeks

List of Referrals/References

[1], a, b, c, d, and e.

Matrix CLO and ILO

CLO	ILO																	
	1		2			3			4				5			6		
	PI		PI			PI			PI				PI			PI		
	1	2	1	2	3	1	2	3	1	2	3	4	1	2	3	1	2	3
1			✓	✓	✓			✓	✓				✓	✓	✓			
2			✓	✓	✓			✓	✓				✓	✓	✓			
3			✓	✓	✓			✓	✓				✓	✓	✓			
4			✓	✓	✓			✓	✓	✓			✓	✓	✓	✓	✓	✓
5			✓	✓	✓			✓	✓	✓						✓	✓	✓
6			✓	✓	✓			✓	✓	✓			✓	✓	✓	✓	✓	✓

1.5. Assessment rubric

A. Assignment Grading Rubric

Task: Perform simple data analysis on data drawn using basic SQL *queries* .

Group	Criterion 1			Criterion 2				FINAL SCORE
	3	2	1	4	3	2	1	
1								
2								
3								
4								
⋮								

$$\text{Grades (criteria 1 and 2)} = \frac{\square\square\square\square}{\square\square\square\square\ \square\square\square\square\square\square\square\square\square\square} \times 100, \text{Final Grade} = \frac{\square\square\square\square\square\ 1 + \square\square\square\square\square\ 2}{2}$$

Criterion 1	Score	Indicator
Perform *basic* SQL queries	3	Write all *basic stages of* SQL queries accurately and correctly.
	2	Write several *basic* stages of SQL queries accurately and correctly.
	1	Writing *basic* SQL query steps but not correctly, or vice versa.

Criterion 2	Score	Indicator
Simple data analysis	4	Carry out data analysis properly and completely, and the estimation results are correct
	3	Performing incomplete data analysis , but estimation results are correct
	2	incomplete data analysis and incorrect estimation results

	1	Only perform data analysis, there is no estimation process

B.   Midterm Assessment Rubric

Student name	Score (0 – 100)			Mark
	Explaining the meaning and terminology in data science (Weight 2 0%)	Visualize and calculate statistics from empirical data (Weight 4 0%)	Do *basic* SQL queries , and the basics of Machine Learning (Weight 40%)	
1				
2				
3				
4				
⋮				

C. Mini Research Progress Report Assessment Rubric

Group	Score			Mark
	Structure of Writing a Mini Research Progress Report (Grade 1)	Use of Scientific Language (Grade2)	Accuracy in Writing Theory (Grade 3)	
1				
2				
3				
4				
⋮				

$$\text{Value} = \frac{\square\square\square\square1 + \square\square\square\square2 + \square\square\square\square3}{3} \text{(after conversion, namely Value(1,2,3)} = \frac{\square\square\square\square}{\square\square\square\square \ \square\square\square\square\square\square\square\square} \times 100)$$

Scoring Guide Table

Score	Indicator		
	Writing Structure of a mini research progress report	Use of Scientific Language	Accuracy of Writing Theory

4	Complete and ordered: consisting of - Cover (Title, writing team), - Summary, - Introduction, - Literature review - Research methods, - Results and Discussion, - Conclusion - Bibliography	Using active sentences, clear and correct use of SPOK, correct use of terms (both in foreign languages), no typos or spelling mistakes	There is no confusion in writing a definition or theory, there are no double meanings, there are no repeated sentences with the same meaning
3	Incomplete: only 75% of what is listed in the highest score	Using active sentences less than 25% of all sentences) passive sentences, there are inaccurate SPOKs (less than 25% of all sentences), there are incorrect foreign terms, there are 25% typos/spelling errors	There are 25% errors in writing definitions or theories, double meanings and repetitive sentences
2	Incomplete: only 50% of what is listed in the highest score	There are 50% passive sentences, 50% incorrect SPOKE, and 50% incorrect foreign terms, there are 50% typos/spelling errors	There are 50% errors in writing definitions or theories, double meanings and repetitive sentences
1	Incomplete: only 25% of what is listed in the highest score	More than 75% use passive sentences, more than 75% of SPOK are incorrect, more than 75% of foreign terms are incorrect, there are typos/spelling errors	There are 75% errors in writing definitions or theories, double meanings and repetitive sentences

D. Mini Research Final Report Assessment Rubric

Group	Score			Mark
	Structure of Writing a Mini Research Progress Report (Grade 1)	Use of Scientific Language (Grade2)	Accuracy in Writing Theory (Grade 3)	
1				
2				
3				
4				
⋮				

$$\text{Value} = \frac{\square\square\square\square\square 1 + \square\square\square\square\square 2 + \square\square\square\square\square 3}{3} \text{(after conversion, namely Value(1,2,3)} = \frac{\square\square\square\square}{\square\square\square\square \ \square\square\square\square\square\square\square\square} \times 100)$$

Scoring Guide Table

Score	Indicator		
	Writing Structure of a mini research progress report	Use of Scientific Language	Accuracy of Writing Theory

4	Complete and ordered: consisting of - Cover (Title, writing team), - Summary, - Introduction, - Literature review - Research methods, - Results and Discussion, - Conclusion - bibliography	Using active sentences, clear and correct use of SPOK, correct use of terms (both in foreign languages), no typos or spelling mistakes	There is no confusion in writing a definition or theory, there are no double meanings, there are no repeated sentences with the same meaning
3	Incomplete: only 75% of what is listed in the highest score	Using active sentences less than 25% of all sentences) passive sentences, there are inaccurate SPOKs (less than 25% of all sentences), there are incorrect foreign terms, there are 25% typos/spelling errors	There are 25% errors in writing definitions or theories, double meanings and repetitive sentences
2	Incomplete: only 50% of what is listed in the highest score	There are 50% passive sentences, 50% incorrect SPOKE, and 50% incorrect foreign terms, there are 50% typos/spelling errors	There are 50% errors in writing definitions or theories, double meanings and repetitive sentences
1	Incomplete: only 25% of what is listed in the highest score	More than 75% use passive sentences, more than 75% of SPOK are incorrect, more than 75% of foreign terms are incorrect, there are typos/spelling errors	There are 75% errors in writing definitions or theories, double meanings and repetitive sentences