

SYLLABUS

Data Collection and Modeling

University year 2025-2026

1. Information regarding the programme

1.1. Higher education institution	Babeş Bolyai University
1.2. Faculty	Faculty of Mathematics and Computer Science
1.3. Department	Department of Computer Science
1.4. Field of study	Computer Science
1.5. Study cycle	Master
1.6. Study programme/Qualification	Data Science for Industry and Society
1.7. Form of education	With Frequency

2. Information regarding the discipline

2.1. Name of the discipline		Data Collection an Modeling					Discipline code		MME8182		
2.2. Course coordinator					Lect. Dr. Ing. Grebla Horea Adrian						
2.3. Seminar coordinator					Lect. Dr. Ing. Grebla Horea Adrian						
2.4. Year of study		1	2.5. Semester		1	2.6. Type of evaluation		E	2.7. Discipline regime		Compulsory

3. Total estimated time (hours/semester of didactic activities)

3.1. Hours per week	4	of which: 3.2 course	2	3.3 seminar/laboratory/project	11+1p
3.4. Total hours in the curriculum	56	of which: 3.5 course	28	3.6 seminar/laboratory/project	28
Time allotment for individual study (ID) and self-study activities (SA)					hours
Learning using manual, course support, bibliography, course notes (SA)					36
Additional documentation (in libraries, on electronic platforms, field documentation)					28
Preparation for seminars/labs, homework, papers, portfolios and essays					36
Tutorship					9
Evaluations					10
Other activities:					
3.7. Total individual study hours	119				
3.8. Total hours per semester	175				
3.9. Number of ECTS credits	7				

4. Prerequisites (if necessary)

4.1. curriculum	
4.2. competencies	<ul style="list-style-type: none"> Average computer usage skills

5. Conditions (if necessary)

5.1. for the course	<ul style="list-style-type: none"> Projector
5.2. for the seminar /lab activities	<ul style="list-style-type: none"> Computers, Python programming language environment, Excel or Libre Calc

6.1. Specific competencies acquired ¹

¹ One can choose either competences or learning outcomes, or both. If only one option is chosen, the row related to the other option will be deleted, and the kept one will be numbered 6.

Professional/essential competencies	<ul style="list-style-type: none"> • Understanding and working with basic concepts in data analysis; • Modeling and solving real-life problems in interdisciplinary teams with complementaries in knowledge;
Transversal competencies	<ul style="list-style-type: none"> • Ethic and fair behavior, commitment to professional deontology; • Professional communication skills; concise and precise description, both oral and written, of professional results , negotiation abilities;

6.2. Learning outcomes

Knowledge	<p>The student knows:</p> <ul style="list-style-type: none"> • Collection, representation, analysis and visualization of data from various fields (eg economics, finance, biology, natural sciences) • Classification of structured and unstructured data and information
Skills	<p>The student is able to:</p> <ul style="list-style-type: none"> • Evaluation of data quality • Use of specific techniques and methods for obtaining and extracting information from unstructured, semi-structured or structured documents and digital sources
Responsibility and autonomy:	<p>The student has the ability of</p> <ul style="list-style-type: none"> • Administration of data collection and storage systems • Creating data models and knowledge models (conceptual, logical and physical models)

7. Objectives of the discipline (outcome of the acquired competencies)

7.1 General objective of the discipline	<ul style="list-style-type: none"> • Learn the modes of collecting local and network data using various tools and programming languages and use the basic data structures to store, process and present data
7.2 Specific objective of the discipline	<ul style="list-style-type: none"> • Study data collection techniques both from local and network (even web) storage and learn the concepts and methods used for data collection. • Study possible simple data models for storing and processing data.

8. Content

8.1 Course	Teaching methods	Remarks
1. Introduction to data (modeling and modeling)	<ul style="list-style-type: none"> • Interactive exposure 	

2. File types and the way we can store data in them; data sharing using files	<ul style="list-style-type: none">● Presentation● Explanation● Practical examples● Case-study discussions	
3. Data Collection Local Data Acquisition Network Data Retrieval		
4. Data Transport/Interchange formats: XML, JSON		
5. Data cleaning		
6. Web crawling		
7. APIs: REST vs SOAP		
8. SQL Basics: modeling		
9. SQL Basics: queries		
10. Simple Data types: list, dictionary, data frames; Python implementation		
11-14. Applications of presented topics using Python		
Bibliography		
<ol style="list-style-type: none">1. Graeme Simsion Graham Witt, Data Modeling Essentials 3rd Edition, https://www.elsevier.com/books/data-modeling-essentials/simsion/978-0-12-644551-02. Steve Hoberman, Data Modeling Made Simple, https://technicspub.com/data-modeling-made-simple/3. Andy Oppel, Data Modeling, A beginner's guide , https://www.oreilly.com/library/view/data-modeling-a/9780071623988/4. Marcelo Arenas, Pablo Barceló, Filip Murlak, Foundations of Data Exchange, https://www.cambridge.org/core/books/foundations-of-data-exchange/FABEE52F4A503AA9DEFC54A9382AF54C		
8.2 Seminar / laboratory	Teaching methods	Remarks
1. Setup working environment and libraries used (PostgreSQL, Python beautifulsoup4,)	<ul style="list-style-type: none">● Interactive exposure● Explanation● Conversation● Didactical demonstration	
2. Working with various file types: csv, XML, JSON, BAM, FASTA		
3. Data storage and retrieval using SQL		
4. Web data crawling		
5. Preparation of individual projects: collect data, clean data, store data.		
6. Project progress and feedback		
7. Project demonstration and presentations		
Bibliography		
<ol style="list-style-type: none">1. Wes McKinney, Python for Data Analysis, https://wesmckinney.com/pages/book.html2. SOAP UI, https://www.soapui.org/learn/api/soap-vs-rest-api/3. Ryan Mitchell, Web Scraping with Python, 2nd edition, https://www.oreilly.com/library/view/web-scraping-with/9781491985564/		

9. Corroborating the content of the discipline with the expectations of the epistemic community, professional associations and representative employers within the field of the program

<ul style="list-style-type: none"> • The course exists in the curriculum of many universities in the world. • The topics of this course are considered to be a major trend in industry and research.
--

10. Evaluation

Activity type	10.1 Evaluation criteria	10.2 Evaluation methods	10.3 Percentage of final grade
10.4 Course	Demonstrate basic knowlwdge of concepts for data collection and modeling	Written exam	50%
10.5 Seminar/laboratory	Apply various techniques for data collection and modeling for a real life problem	Project implementation and presentation	50%
10.6 Minimum standard of performance			
<ul style="list-style-type: none">Each student should obtain minimum 5 for the written exam and for the final grade			

11. Labels ODD (Sustainable Development Goals)²

Not applicable.

Date:

...

Signature of course coordinator

Lect. dr. ing. Grebla Horea Adrian

Signature of seminar coordinator

Lect. dr. ing. Grebla Horea Adrian

Date of approval:

...

Signature of the head of department

Assoc.prof.phd. Adrian STERCA

² Keep only the labels that, according to the [Procedure for applying ODD labels in the academic process](#), suit the discipline and delete the others, including the general one for *Sustainable Development* – if not applicable. If no label describes the discipline, delete them all and write „*Not applicable.*”.