MD-2068, CHISINAU, 9/7 STUDENTILOR STR, PHONE: 022 50-99-63, www.utm.md

DATA ANALYSIS AND VISUALIZATION

1. Information about the discipline/module

Facultaty	Computers, Informatics, and Microelectronics				
Chair/department	Informatic and Systems Engineering				
Cycle of studies	Bachelor's degree, cycle II				
Study program	**Data Science**				
Year of study	**Semester**	**Type of evaluation**	**Formative category**	**Optionality category**	**ECTS credits**
II (full-time education)	3	E	F – Fundamental course unit	O – Obligatory course unit	5

2. Total estimated time

Total hours in the curriculum plan	Including				
	Auditory hours		Individual work		
	Course	Laboratory/seminar	Term paper	Study of theoretical material	Application preparation
150	20	20	-	80	30

3. Preconditions for access to the discipline/module

According to the curriculum plan	Mathematical Analysis, Probability and Applied Statistics, Data Structures and Algorithms, Discrete Mathematics, Computer Programming, Interactive Programming.
According to competences	The use of domain-specific theories and tools (algorithms, methods, techniques, etc.) for analyzing fundamental algorithms in statistics, mathematical analysis, and graphical data visualization.

4. Conditions for conducting the educational process

Course	For presenting theoretical material in the classroom, a board, projector, and computer are required.
Laboratory/seminar	Students will complete reports in accordance with the conditions specified in the methodological guidelines. The deadline for submitting the laboratory work is two weeks after its completion. Late submissions will be penalized by 1 point per week of delay. For conducting seminars in the study auditorium, a whiteboard, projector, and computer are required.

5. Specific competencies acquired

Professional competencies	**CPM1** Elaboration and design of the system architecture **CPM2**. Monitoring technology trends. Innovation. Sustainable development. **CPM3** Application development. Component integration. Systems engineering. **CPM5.** Process improvement.

6. Course/Module objectives

General objective	Describe the basic concepts. Determine the role of data visualization. Determine the tools and techniques used
Specific objectives	To understand measures of central tendency: mean, median, mode, as well as measures of dispersion: variance, standard deviation, interquartile range. To define the concept of regression and evaluate regression models.

	To apply design principles for data visualization, use visualization tools such as Matplotlib, Seaborn, and Plotly, and create interactive charts.

7. Course/Module content

Syllabus of teaching activities	Number of hours Full-time education
Course topics	
Topic 1. Introduction to Statistical Analysis and Data Visualization.	2
Topic 2. Data Collection and Preprocessing.	2
Topic 3. Descriptive Statisticians.	2
Topic 4. Probability and Probability Distributions.	2
Topic 5. Regression Analysis.	2
Topic 6. Factor Analysis.	2
Topic 7. Clustering Analysis.	2
Topic 8. Data Visualization Techniques.	2
Topic 9. Visualization of geospatial data.	2
Topic 10. Data Visualization Projects.	2
Total lectures:	**20**

Syllabus of teaching activities	Number of hours Full-time education
Topics of laboratory work/seminars	
Practical work no. 1. Preprocessing and Data Mining	4
Practical work no. 2. Basic Statistical Views	4
Practical work no. 3. Visualization and Analysis of Probability Distributions	4
Practical work no. 4. Regression Analysis Project	4
Practical work no. 5. Interactive Visualization Project	4
Total laboratory/practical work:	**20**

8. Using generative AI

Permission to use	The use of generative AI in assignments and projects is permitted, provided that students adhere to the following rules: • Generative AI may be used to generate ideas, text structures, or code, but all generated materials must be reviewed and adjusted by the student to ensure that they meet academic requirements. • Any use of generative AI must be declared in the appendix section of each paper, using the phrase: "During the preparation of this paper, the author used [NAME OF TOOL / SERVICE] for the purpose of [REASON]. After using this tool / service, the author reviewed and edited the content as necessary and assumes full responsibility for the content of the paper."

Restrictions to use	Students ***MUSTN'T consider generative AI as a reliable source of information***, as it does not provide clear references or documented sources. • ***Direct citation of AI-generated content*** in academic papers as if it were a primary source ***isn't permitted.*** • Activities in which the use of **generative AI is prohibited** are specified by the teacher and are usually ***intermediate and final assessments*** or that don't involve professional competence development activities.

9. Bibliographic references

Main	1. Lucia Căbulea, Nicoleta Breaz. INTERPRETAREA STATISTICĂ A INFORMAŢIILOR. ELEMENTE DE DATA MINING ŞI PROGNOZĂ, Universitatea 1 Decembrie 1918" Alba Iulia. 2. http://www.uab.ro/cursuri_perfectionare/program_perfectionare_profesionala/pagini/cursuri/Modul-7-Cabulea-Breaz/curs_modul7.doc 3. T. Andrei, Statistică şi econometrie, Ed. Economică, 2003. 4. C. Anghelache, E. Bugudui, S. Gresoi, E. Niculescu, Statistică aplicată-indicatori, sinteze şi studii de caz, Ed. Economică, 2006 5. E.N. Bâzdoacă, S. Matei, N.G. Bâzdoacă, Iniţiere în Excel, Ed.Arves, Colecţia Iniţiere în calculator 2, 2002 6. Liviu Ciortuz , Alina Munteanu , Elena Badarau Exercitii de invatare automata ed UNIVERSITATII "ALEXANDRU IOAN CUZA" (2015) ISBN: 606-714-197-9 7. MARCUS, S., NICOLAU, E., STATI, S., 1966. Introducere în lingvistica matematică. Bucureşti, Editura ştiinţifică, p.336. 8. An estimate of an upper bound for the entropy of English. Brown, Della Pietra, Mercer, Della Pietra, Lai. Computational Linguistics, 18(1), pp31-40, 1992. 9. VLAD, A., MITREA, A., MITREA, M., 2005. Limba română scrisă ca sursă de informaţie. Paideia, România, 286 p. 10. Roman Feldman, James Sauger. The Text Mining Handbook. Advanced Approaches in Analyzing Unstructured Data. Cambridge 2008.
Additional	1. MIHALCEA, R., 2002. Diacritics Restoration: Learning from Letters versus Learning from Words. in Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics, Mexico, p.339-348. 2. Christopher D.Manning, Hinrich Schutze Foundations of Statistical Natural Language Processing Mit Press, 1999. 3. Ronald Rosenfeld. Two decades of Statistical Language Modeling: Where Do We Go From Here? Proceedings of the IEEE, 88(8), 2000. 4. Lillian Lee. "I'm sorry Dave, I'm afraid I can't do that": Linguistics, Statistics, and Natural Language Processing. Computer Science: Reflections on the Field, Reflections from the Field, pp. 111--118, 2004.

10. Evaluation

Form of education	Periodic		Current	Individual work	Final examination
Full-time	Mid-term 1	Mid-term 2			
	15%	15%	15%	15%	40%
Minimum performance standard					
Attendance and participation in lectures, seminars, and laboratory work. Obtaining a minimum grade of "5" in each assessment and laboratory work.					

Demonstrating knowledge of fundamental algorithms in formal language theory, finite automata, and pushdown automata in the final examination.

11. Evaluation criteria

Activity	Evaluation components	Evaluation method, evaluation criteria	Weight in final grade for the Activity	Weight in course evaluation
Full-time education				
Mid term I	Theoretical content, topics 1-3	Test	100%	**15%**
Mid term II	Theoretical content, topics 4-5	Activities during practical work/seminar	100%	**15%**
Current evaluation	Practical activity	Attendance and participation in classes	50%	**15%**
Individual study	Classification of research by activity type	Presentation/Discussion on the topic	100%	**15%**
Final examination	Theoretical and practical content	Oral exam. Grading according to grading scale	100%	**40%**