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

DATA ANALYSIS TOOLS

1. Course/Module information

Faculty	Computers, Informatics, and Microelectronics				
Department	Informatics and Systems Engineering				
Study cycle	Cycle II, Master's Studies				
Study program	**Data Science**				
Year of study	**Semester**	**Evaluation type**	**Formative category**	**Optionality category**	**ECTS credits**
1st year *(full-time education)*	II	PA	S – specialty course unit	O - obligatory course unit	5

2. Estimated total time

Total hours in the curriculum	**Including**				
	Auditory hours		**Individual work**		
	Lecture	**Laboratory**	**Term paper**	**Study of theoretical material**	**Application development**
Full-time education	-	40	-	110	-

3. Prerequisites for access to the course/module

According to the curriculum plan	Project Management, Artificial Intelligence / Neural Networks and Deep Learning, Machine Learning and Data Mining, Data Analysis and Visualization, Exploratory Data Analysis and Data Modeling.
According to competencies	Students must have knowledge of the specifics of various data analysis tools and be able to apply them to solve various practical problems.

4. Conditions for conducting the educational process

Lecture	To present theoretical material in the classroom, a board, chalk, projector, and computer are required. Delays from students, laptop use, and phone conversations during the lecture will not be tolerated.
Practice	For conducting laboratory work in the classroom, a board, chalk, and computers are required. Student tardiness and phone conversations during the class will not be tolerated.

5. Specific competencies acquired

Professional competencies	**CPM 1.** Development and Design of Architecture K1 Architecture models, methodologies, and system design tools K2 System architecture requirements: performance, maintainability, extensibility, scalability, availability, security, and accessibility K3 The costs, benefits, and risks of a system architecture K4 Enterprise architecture and internal company standards K5 New emerging technologies (e.g., distributed systems, virtualization models, datasets, mobile systems) S1 Provides expertise to help solve complex technical problems and ensure the implementation of the best architecture solutions S2 Uses technological knowledge from various fields to develop and implement the enterprise architecture S3 Understands the company's objectives that impact the components of the architecture (data, applications, security, development, etc.) S4 Assists in communicating the enterprise architecture and standards, principles, and objectives to the various teams involved S5 Develops design models and architectural models to assist system analysts in designing cohesive applications. **CPM 2.** Monitoring Technological Trends. Innovation. Sustainable Development. K1 Existing and emerging technologies and their relevant applications in the market K2 Business, societal, and research objectives, trends, and needs

	K3 Relevant sources of information (e.g., journals, conferences and events, opinion leaders, online forums, etc.) K4 Concrete approaches to applied research programs K5 Innovation process techniques K6 Criteria and indicators for sustainable development K7 Corporate social responsibility (CSR) of stakeholders within the information system infrastructure S1 Monitors information sources and continuously tracks the most promising ones S2 Identifies vendors and suppliers of the most promising solutions; evaluates, justifies, and proposes the most suitable ones S3 Identifies the advantages and improvements brought by adopting emerging technologies S4 Thinks without preconceived ideas S5 Applies recommendations within projects that support the latest sustainable development strategies **CPM 3.** Application Development. Component Integration. Systems Engineering. K1 Appropriate programs/modules, DBMS, and programming languages. Cutting-edge technologies. K3 The impact of system integration on the organization or existing system K4 Interface techniques between modules, systems, and components K5 Integration testing techniques K7 Hardware components, tools, and hardware architectures K8 Functional and technical design K9 Information security fundamentals K10 Prototyping S1 Applies appropriate software and/or hardware architectures S2 Measures system performance before, during, and after system integration S3 Identifies and records activities, issues, and corrective actions related to maintenance S4 Adapts customer needs to existing products S5 Secures and backs up data to ensure its integrity during data or system integration S6 Explains and communicates to the client regarding design/development S7 Launches and evaluates test results according to product specifications S9 Applies data models and processes to develop efficiently and productively **CPM 4.** Staff Development K1 Competence development methods K2 Methodologies for analyzing competence needs K3 Methods for supporting learning and development (e.g., coaching, teaching) K4 Relevant information technologies and processes K5 Accountability and empowerment techniques S1 Identifies competence and qualification gaps S2 Identifies and recommends development opportunities based on work activities S3 Incorporates competence development opportunities into routine work processes
Transversal competencies	CT 1. Autonomy and responsibility. CT 2. Social interaction. CT 3. Personal and professional development.

6. Course/Module objectives

General objective	Development of practical skills in using modern data analysis tools for interpretation, visualization, and data-driven decision-making.
Specific objectives	• Familiarization with the fundamental concepts of data analysis. • Development of skills in using specialized software for data processing and analysis. • Application of data visualization and interpretation methods for practical purposes. • Use of modeling and prediction techniques in data analysis. • Mastering the process of reporting and presenting analytical results.

7. Course/Module content

Syllabus of teaching activities	**Number of hours**
	Full-time education
Course topics	
Topic 1. Stages in Data Analysis	**4**

Syllabus of teaching activities	Number of hours Full-time education
Defining data analysis and its importance. The data analysis process: collection, processing, analysis, interpretation. Data analysis using commonly used software tools (Excel, Python, Google Colab).	
Topic 2. Presentation of the KNIME Analytical Platform Introduction to KNIME. Elements of a workflow. Basic nodes. Basic statistical analysis in KNIME.	**4**
Topic 3. Data Processing: Import and Cleaning Importing data from various sources (CSV, Excel, databases). Detecting and handling missing values. Data cleaning techniques (removing duplicates, standardizing formats). Data processing using Excel, Python, KNIME, Google Colab.	**4**
Topic 4. Exploratory Data Analysis (EDA) Using descriptive statistics to understand datasets. Identifying patterns and relationships between variables. Creating charts (histograms, scatter plots, boxplots). Exploratory analysis using Excel, Python, KNIME, Google Colab.	**4**
Topic 5. Data Visualization Basic principles of data visualization. Creating customized dashboards. Data visualization using Excel, Python, KNIME, Google Colab.	**4**
Topic 6. Data Analysis using Python Using Pandas and NumPy libraries for data processing. Creating visualizations with Matplotlib and Seaborn. Practical examples using real datasets.	**4**
Topic 7. Analysis of AI/ML/DS Projects using KNIME Studying the KNIME repository. Analyzing different types of models.	**8**
Topic 8. Regression and Prediction Models Introduction to regression models (linear and logistic). Building and evaluating models. Using a dataset for predictions with Excel, Python, KNIME, Google Colab.	**4**
Topic 9. Presenting and Reporting Results Writing analytical reports. Presenting a final lab project.	**4**
Total practical work:	**40**

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	**10. McKinney, Wes.** *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython.* O'Reilly Media, 2018. ISBN 978-1-491-95761-5. **11. Nisbet, Robert, Gary Miner, and Ken Yale.** *Handbook of Statistical Analysis and Data Mining Applications.* Academic Press, 2017. ISBN 978-0-12-416632-5.

	12. **Cielen, Davy, Arno D. B. Meysman, and Mohamed Ali.** *Introducing Data Science: Big Data, Machine Learning, and More, Using Python Tools.* Manning Publications, 2016. ISBN 978-1-61729-156-2. 13. **Aggarwal, Charu C.** *Data Mining: The Textbook.* Springer, 2015. ISBN 978-3-319-14142-8. https://doi.org/10.1007/978-3-319-14142-8
Supplementary	1. **Knime Community.** *KNIME Beginner's Luck: A Guide to KNIME Analytics Platform for Beginners.* KNIME Press, 2020. Available at: https://www.knime.com 2. **Shmueli, Galit, et al.** *Data Mining for Business Analytics: Concepts, Techniques, and Applications in Python.* Wiley, 2020. ISBN 978-1-119-54984-0.

10. Evaluation

Periodic		Current	Individual study	Project/thesis	Exam
Mid term 1	**Mid term 2**				
Full-time education					
20%	20%	-	-	60%	-

11. Evaluation criteria

Activity	Evaluation components	Evaluation method, Evaluation criteria	Weight in the final grade of the activity	Weight in the course evaluation
Full-time education				
Mid term I	Laboratory work, topics 1-4	Presentation of completed lab assignments	100%	**20%**
Mid term II	Laboratory work, topics 5-9	Activities in practical work/seminar	100%	**20%**
Creating a project	Theoretical and practical content	Oral presentation. Grading according to grading scale	100%	**60%**