|  |
| --- |
| **WSB University** |
| **Field of study: Management** |
| **Module / course: Introduction to Data Mining Methods** |
| **Educational profile: General** |
| **Education cycle: II cycle studies** |
| **Number of hours per semester** | 1 | 2 |
| I | **II** | III | IV |
| **Full time studies****(L/C/lab/pr/e)** |  | **16L/16lab** |  |  |
| **Part-time studies****(L/C/lab/pr/e)** |  |  |  |  |
| **LECTURER** | Karol Jędrasiak PhD Eng., Damian Skipioł, M.Sc. |
| **FORM OF CLASSES** | Lectures, laboratories |
| **COURSE OBJECTIVES** | To familiarize students with the practical aspects of analysis, visualization and methods of business data mining. Familiarization with the R Project, Anaconda and Qlik Sense Environments. |
| **Course objectives** | **References to course objectives:** | **Description of educational results**  | **Verification of educational results** |
| **Major-related learning outcomes**  | **Area-related learning outcomes**  | **Knowledge** |
| InDatExM\_K01 | Z2\_W12Z2\_W20 | S1A\_W06 | Student understanding the mechanisms of analysis, integration, visualization and methods of business data mining. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_K02 | Z2\_W12Z2\_W20 | S1A\_W06 | Student correctly defines the concepts of analysis, integration, visualization and methods of business data mining. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_K03 | Z2\_W12Z2\_W20 | S1A\_W06 | Student correctly describes the most important techniques used in the processes of analysis, integration, visualization and methods of business data mining. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| InDatExM\_K04 | Z2\_W12Z2\_W20 | S1A\_W06 | Student indicates the dangers and potential problems resulting from low quality of data in the processes of analysis, integration, visualization and methods of business data mining. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_K05 | Z2\_W12Z2\_W20 | S1A\_W06 | Student correctly describes available methods and techniques of analysis, integration, visualization and methods of business data mining. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| **Abilities** |
| InDatExM\_A06 | Z2A\_U04Z2A\_U10 | S2A\_U01S2A\_U02S2A\_U03S2A\_U04S2A\_U06 | Student should demonstrate the ability to understand and implement solutions in the field of analysis, integration, visualization and methods of business data mining in R Project, Anaconda and Qlik Sense Desktop / Cloud environments. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_A07 | Z2A\_U04 | S2A\_U01S2A\_U02S2A\_U03S2A\_U04 | Student has the ability to properly design the CRISP-DM process | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_A08 | Z2A\_U04 | S2A\_U01S2A\_U02S2A\_U03S2A\_U04 | Student implements the process of analysis, integration, visualization and methods of data mining in R Project technologies, Anaconda and Qlik Sense Desktop / Cloud | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| **Social competences:** |
| InDatExM\_S14 | Z2A\_K03 | S2A\_K03 | Awareness of the consequences of proper design and application of analysis, integration and data reporting systems. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_S15 | Z2A\_K02 | S2A\_K02 | Collaborates with other team members, he can convey the results of data quality analysis and the design of the reporting process to people who do not know the technology but have knowledge about the data. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| InDatExM\_S16 | Z2A\_K03 | S2A\_K03 | The student can correctly interpret the requirements and needs reported by the users of the solution | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| InDatExM\_S17 | Z2A\_K03 | S2A\_K03 | The student shows creativity in the evaluation of data and the possibility of the process, critically evaluates the results. | Written testEvaluation of solutions for simple design tasksControl questions about the knowledge gained during the classes. |
| **Students’ workload (in didactic hours 1h did.=45 minutes)\*\***  |
| **Full- time**Participation in lectures = 16hParticipation in laboratory = 16hPreparation to laboratory = 14hPreparation to lectures = 14hPreparation to an examination = 15hProject tasks =e-learning =Credit/examination = 4hothers (indicate which) = **TOTAL: 79h****ECTS points: 3****Including practical classes:3** | **Part - time**Participation in lectures = Participation in laboratory = Preparation to laboratory = Preparation to lectures = Preparation to an examination = Project tasks =e-learning =Credit/examination = others (indicate which) = **TOTAL:** **ECTS points:****Including practical classes:** |
| **PREREQUISITES** | Knowledge of the basics: computer science, statistics, mathematical analysis, operating systems and databasesKnowledge of methods of modeling multidimensional data in associative models.Knowledge of data warehouse topics and analytical systems architecture.Knowledge of SQL language at the basic level. |
| **COURSE CONTENT****(Division to contact classes and e-learning)** | Content implemented in the direct form:Laboratory:* Introduction to the subject matter.
* Theoretical foundations of analysis, integration, visualization and methods of data mining.
* Theoretical basics of data design.
* Practical aspects of the implementation of analytical solutions in the R Project, Anaconda and Qlik Sense Desktop / Cloud tools
* Practical aspects of application and design of neural networks, associative analysis, APRIORI algorithm, basic statistical analysis, ARIMA / SARIMA model, prediction models, Holt Winters model, decision trees
* Other issues from the area.
* Responding to errors and problems.

Content implemented in the form of e-learning: not applicable |
| **LITERATURE****(compulsory reading)** | * Philipp K. Janert, Data Analysis with Open Source Tools. O’Reilly 2011.
* Niall M. Adams, Perspectives on Data Mining. International Journal of Market Research 2010, vol. 52, 1. (Sage Journals).
* SPSS Data Entry Builder 3.0 User's Guide / SPSS Inc. - Chicago : SPSS Inc, 2001.
 |
| **OPTIONAL LITERATURE** | * GLENN J. MYATT WAYNE P. JOHNSON, MAKING SENSE OF DATA I A Practical Guide to Exploratory Data Analysis and Data Mining Second Edition, John Wiley & Sons, Inc., Hoboken, New Jersey 2016.
* Pace Larry, Beginning.R., Apress, Oct.2012., ISBN.1430245549.
* Mico Yuk Stephanie Diamond, Data Visualization For Dummies, John Wiley & Sons, New Jersey 2014.
* R. Kimball, M. Ross, W. Thornthwaite, J. Mundy, B. Becker “The Data Warehouse Lifecycle Toolkit. 2nd Edition”. John Wiley & Sons 2008.
* New to Qlik Sense - Topics You Need to Know, <https://community.qlik.com/t5/Qlik-Sense-Enterprise-Documents/New-to-Qlik-Sense-Topics-You-Need-to-Know/ta-p/1530582>
* Qlik Sense First Step Tutorial from <https://help.qlik.com/en-US/sense/April2019/Subsystems/Hub/Content/Sense_Hub/Apps/create-first-app.htm>
* Getting started using Qlik Sense, <https://help.qlik.com/en-US/sense/April2019/Subsystems/Hub/Content/Sense_Hub/Introduction/get-started.htm>
* Step-by-Step GuideGet Started with Qlik®, https://help.qlik.com/en-US/onboarding
 |
| **TEACHING METHODS****(Division to contact classes and e-learning)** | In the direct form:Instructive method.In the form of e-learning: not applicable |
| **TEACHING AIDS** | * Multimedia presentation.
* Source files containing data for exercises.
* R Project, Anaconda and Qlik Sense Desktop / Cloud software
* Sample text files containing data.
 |
| **PROJECT****(if implemented in the framework of a classes module)** | The analytical application project implemented in the R Project environment, Anaconda and Qlik Sense Desktop / Cloud available for inspection and evaluation. |
| **METHOD Of ASSESSMENT****(Division to contact classes and e-learning**  | Credit with grade |
| **FORM AND CONDITIONS OF ASSESSMENT** | * Implementation of practical tasks within the laboratory.
* The condition for obtaining a pass is obtaining a positive grade for all forms of credit provided for in the curriculum, including the quantitative criteria for assessment set out in the Student Assessment Framework at the WSB University.
 |

*\* L-lecture, C- classes lab- laboratory, pro- project, e- e-learning*