

## Subject card

|                                             |                                                                                                                                      |                                                          |                                                                                        |                                     |                                                                                              |            |     |
|---------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------|----------------------------------------------------------------------------------------|-------------------------------------|----------------------------------------------------------------------------------------------|------------|-----|
| Subject name and code                       | Data Mining – Data Sets Exploration, PG_00156751                                                                                     |                                                          |                                                                                        |                                     |                                                                                              |            |     |
| Field of study                              | Informatics and Econometrics                                                                                                         |                                                          |                                                                                        |                                     |                                                                                              |            |     |
| Date of commencement of studies             | October 2024                                                                                                                         |                                                          | Academic year of realisation of subject                                                |                                     | 2025/2026                                                                                    |            |     |
| Education level                             | Bachelor's studies                                                                                                                   |                                                          | Subject group                                                                          |                                     | Optional subject group<br>Subject group related to scientific research in the field of study |            |     |
| Mode of study                               | full-time studies                                                                                                                    |                                                          | Mode of delivery                                                                       |                                     | at the university                                                                            |            |     |
| Year of study                               | 2                                                                                                                                    |                                                          | Language of instruction                                                                |                                     | Polish                                                                                       |            |     |
| Semester of study                           | 4                                                                                                                                    |                                                          | ECTS credits                                                                           |                                     | 7.0                                                                                          |            |     |
| Learning profile                            | academic                                                                                                                             |                                                          | Assessment form                                                                        |                                     | exam                                                                                         |            |     |
| Conducting unit                             | Department of Statistics -> Faculty of Management -> Rector                                                                          |                                                          |                                                                                        |                                     |                                                                                              |            |     |
| Name and surname of lecturer (lecturers)    | Subject supervisor                                                                                                                   |                                                          | dr hab. Krzysztof Najman                                                               |                                     |                                                                                              |            |     |
|                                             | Teachers                                                                                                                             |                                                          | dr hab. Krzysztof Najman<br><br>mgr Katarzyna Raca<br><br>dr hab. Kamila Migdał-Najman |                                     |                                                                                              |            |     |
| Lesson types                                | Lesson type                                                                                                                          | Lecture                                                  | Tutorial                                                                               | Laboratory                          | Project                                                                                      | Seminar    | SUM |
|                                             | Number of study hours                                                                                                                | 30.0                                                     | 30.0                                                                                   | 15.0                                | 0.0                                                                                          | 0.0        | 75  |
|                                             | E-learning hours included: 0.0                                                                                                       |                                                          |                                                                                        |                                     |                                                                                              |            |     |
| Learning activity and number of study hours | Learning activity                                                                                                                    | Participation in didactic classes included in study plan |                                                                                        | Participation in consultation hours |                                                                                              | Self-study | SUM |
|                                             | Number of study hours                                                                                                                | 75                                                       |                                                                                        | 30.0                                |                                                                                              | 70.0       | 175 |
| Subject objectives                          | The aim of the course is to familiarize students with the basic methods of Exploratory Data Analysis and its practical applications. |                                                          |                                                                                        |                                     |                                                                                              |            |     |

|                                 |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                               |                                                                                                                              |
|---------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------|
| Learning outcomes               | Course outcome                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         | Subject outcome                                                                                                                                               | Method of verification                                                                                                       |
|                                 | [liEL3_K02] The student can communicate with the public on computer science and econometrics topics in and outside the workplace, communicate his knowledge and share his skills through various media.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Is able to discuss, communicate and share knowledge regarding the theory and practice of applying Data Mining models.                                         | [SK1] oral statement/conversation/discussion<br>[SK2] presentation/project/paper/report<br>[SK4] test/exam - oral or written |
|                                 | [liEL3_U02] The student is able to acquire basic information about economic processes and phenomena by direct observation, planned experiment or database queries, and collect and process it using modern information technology tools.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | The student is able to collect statistical data from primary and secondary sources for the purposes of implementing social research using Data Mining models. | [SU2] presentation/project/paper/report                                                                                      |
|                                 | [liEL3_W05] The student has advanced knowledge of the sources of socio-economic data, their databases and ways of creating them.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Student knows the sources and methods of searching for secondary data used in social research analyzed using Data Mining methods.                             | [SW4] test/exam - oral or written<br>[SW1] oral statement/conversation/discussion<br>[SW2] presentation/project/paper/report |
|                                 | [liEL3_W04] The student has advanced knowledge of mathematical, statistical, econometric and computer methods of obtaining, processing and analysing data depicting the functioning and development of the national economy and its elements, as well as the phenomena and processes occurring in their environment.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | Has in-depth knowledge of advanced data analysis methods using selected Data Mining models.                                                                   | [SW4] test/exam - oral or written<br>[SW2] presentation/project/paper/report                                                 |
|                                 | [liEL3_K01] The student understands the need to constantly supplement and deepen acquired knowledge; is open to new ideas and teaching methods.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | Understands the need to constantly update knowledge in the theory and applications of Data Mining.                                                            | [SK1] oral statement/conversation/discussion<br>[SK2] presentation/project/paper/report                                      |
| Subject contents                | <p>[liEL3_U05] The student can make a statistical description of advanced economic categories, formulate and verify simple hypotheses on their formation.</p> <p>The student is able to independently conduct an analysis of selected social phenomena using selected Data Mining models.</p> <p>[SU1] oral statement/conversation/discussion<br/>[SU2] presentation/project/paper/report</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                               |                                                                                                                              |
|                                 | <p>1. Basic concepts, purpose and functions of data mining. Data mining (EDA), data mining, knowledge discovery in databases (KDD), data mining as a process, stages of the data mining process.2. Multidimensional data sets, the concept of Big data. Analytical software.3. Data mining tasks: description, estimation, prediction, classification, grouping, rule discovery. Examples of data mining applications in socio-economic research. Data sources.4. Data pre-processing: data preparation for analysis, the problem of variable and case selection, feature space dimension reduction, the problem of missing data, graphical methods for identifying distant units, analytical methods for identifying distant units. Variable and feature normalization procedures.5. Basic tabular and graphical EDA methods.6. Partial and multiple correlation. Interdependence matrices.7. Association rules, association concept, binary and quantitative rules, one-dimensional and multi-dimensional association rules, one-level and multi-level association rules, association rule support and confidence, elements of basket analysis, basket problem, sequence analysis, association analysis, other applications of association rules in EDA.8. Cluster analysis and data classification. The process of clustering and classifying data. Similarity of objects and its measurement. Measurement of object diversity according to quantitative features - distance measures. Measurement of object similarity according to quantitative features. Similarity of features and its measurement. Measurement of similarity of quantitative features based on correlation.9. Hierarchical classification methods. The essence of agglomerative hierarchical classification. Nearest neighbor method, farthest neighbor method, group mean method, centroid method, median method, Ward method. Lance and Williams scheme. Visualization of group structure. Evaluation of clustering based on dendrogram.10. Selected EDA methods: classification trees, regression trees, k-nearest neighbors algorithm.</p> |                                                                                                                                                               |                                                                                                                              |
| Prerequisites and co-requisites | Advanced knowledge of algebra and statistics.                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |                                                                                                                                                               |                                                                                                                              |
| Assessment methods and criteria | Subject passing criteria                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Passing threshold                                                                                                                                             | Percentage of the final grade                                                                                                |
|                                 | semester project                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 51.0%                                                                                                                                                         | 50.0%                                                                                                                        |
|                                 | written exam                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 51.0%                                                                                                                                                         | 50.0%                                                                                                                        |

|                                                                |                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|----------------------------------------------------------------|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Recommended reading                                            | Basic literature         | <p>1. M.Lasek, Data mining, Biblioteka Menedżera i Bankowca, Warszawa, 2002</p> <p>2. D.T.Larose, Okrywanie wiedzy z danych. Wprowadzenie do eksploracji danych, Wydawnictwo Naukowe PWN, Warszawa, 2006</p> <p>3. E.Gatnar, Nieparametryczna metoda dyskryminacji i regresji, Wydawnictwo Naukowe PWN, Warszawa 2001</p> <p>4. A. Balicki, Statystyczna analiza wielowymiarowa i jej zastosowania społeczno-ekonomiczne, Wydawnictwo UG, Gdańsk, 2009</p> |
|                                                                | Supplementary literature | Nong Ye, The Handbook of Data Mining, Lawrence Erlbaum Associates, New Jersey, 2003                                                                                                                                                                                                                                                                                                                                                                        |
|                                                                | eResources addresses     |                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Example issues/<br>example questions/<br>tasks being completed |                          |                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| Work placement                                                 | Not applicable           |                                                                                                                                                                                                                                                                                                                                                                                                                                                            |

Document generated electronically. Does not require a seal or signature.