#### card of course

|  |  |
| --- | --- |
| Subject name | Data Set Analysis – Data Set Programming |

1. The placement of the subject in the study system

|  |  |
| --- | --- |
| 1.1. Field of study | Computer science |
| 1.2. Form and path of study | Full-time/Part-time |
| 1.3. Level of education | First-cycle studies |
| 1.4. Study profile | Optional/practical |

|  |  |
| --- | --- |
| 1. 5. Specialty | Artificial intelligence |
| 1.6. Subject Coordinator | Dr Michał Kalisz |

2. General characteristics of the subject

|  |  |
| --- | --- |
| 2.1. Belonging to a subject group | Optional/practical |
| 2.2. Number of ECTS | 3 |
| 2.3. Language of lectures | English |
| 2.4. Semesters in which the subject is taught | III |
| 2.5.Criteria for selecting course participants | For specialization: Artificial Intelligence |

1. Learning outcomes and course delivery
	1. Subject Objectives

|  |  |
| --- | --- |
| No. | Subject Objectives |
|
| C1 | Acquiring knowledge of tools and techniques for data exploration, analysis and visualization. |
| C2 | Learning programming and data processing using libraries and tools such as Python, Pandas, NumPy, and SQL. |
| C3 | Implementation of projects involving data preparation, analysis and interpretation of results in a professional environment. |

* 1. Subject-specific learning outcomes, divided into knowledge , skills and competences , with reference to the directional learning outcomes

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Description of subject learning outcomes | Reference to directional effectslearning (symbols) | Method of implementation (mark "X") |
| ST | NST |
| Classes at the University | Activities on the platform | Classes at the University | Activities on the platform |
| After passing the course, the student knows and understands **the knowledge** |
| W1 | Knows concepts related to data analysis, such as preprocessing, data mining, and predictive models. | INF\_W03INF\_W04 | X |  | X |  |
| W2 | Understands data manipulation and analysis methods in Python, including the use of libraries such as Pandas, NumPy, and Matplotlib. | X |  | X |  |
| W3 | Possesses knowledge of the structure and processing of large data sets. | X |  | X |  |
| W4 | Knows basic data visualization methods and tools for creating charts and dashboards. | X |  | X |  |
| W5 | Understands the basics of SQL and its applications in working with databases. | X |  | X |  |
| After passing the course, the student is **able** to: |
| U1 | Is able to process and analyze data sets using programming tools. | INF\_U02INF\_U03 | X |  | X |  |
| U2 | Is able to apply data mining techniques, including grouping, aggregating and filtering data. | X |  | X |  |
| U3 | Can design and implement data analyses using Pandas, NumPy libraries and SQL tools. | X |  | X |  |
| U4 | Can create clear data visualizations using Matplotlib, Seaborn, or Tableau. | X |  | X |  |
| U5 | Is able to interpret the results of data analysis and present them in the form of reports. | X |  | X |  |
| After completing the course, the student is ready to take part in **social competences.** |
| K1 | Understands the importance of reliable data analysis and its impact on decision-making. | INF\_K02 | X |  | X |  |
| K2 | Is aware of the responsibility for processing and protecting data in accordance with applicable regulations. | X |  | X |  |

3.3. Forms of teaching and their number of hours - Full-time studies (ST), Part-time studies (NST)

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Path | Lecture | Exercises | Design | Workshop | Laboratory | Seminar | Lecturer | Classes conducted using distance learning methods and techniques in the form of a lecture | Other | **ECTS points** |
| **ST** |  |  |  |  | 30 |  |  |  |  | 3 |
| **NST** |  |  |  |  | 15 |  |  |  |  | 3 |

3.4 . Content of education (separately for each form of classes: (W, ĆW, PROJ, WAR, LAB, LEK, OTHER). It should be marked (X) how the given content will be implemented (classes at the university or classes on the e-learning platform conducted using distance learning methods and techniques)

TYPE OF CLASS: LABORATORY

|  |  |  |  |
| --- | --- | --- | --- |
| No. | Content of the course | Reference to subject-specific learning outcomes | Method of implementation (mark "X") |
| ST | NST |
| **Classes at the University** | **Activities on the platform** | **Classes at the University** | **Activities on the platform** |
| 1. | Introduction to data analysis , basic concepts, steps of data analysis. | W1, K1 | X |  | X |  |
| 2. | Preparing data for analysis , data cleaning, dealing with missing values, data transformations. | W1, U1 | X |  | X |  |
| 3. | Data mining with Pandas and NumPy, statistical analysis, aggregation, filtering, grouping. | W2, W3, U2 | X |  | X |  |
| 4. | Data visualization with Matplotlib and Seaborn , creating charts, selecting appropriate forms of visualization. | W4, U4 | X |  | X |  |
| 5. | Basics of SQL and its use , creating queries, database operations. | W5, U3 | X |  | X |  |
| 6. | Advanced analysis , predictive models, regression, classification. | W3, U2, U5 | X |  | X |  |
| 7. | Working with large data sets , optimizing analysis, introduction to big data. | W3, K2 | X |  | X |  |
| 8. | Presentation of final projects , summary and evaluation of students' work. |  | X |  | X |  |

3.5. Methods of verifying learning outcomes (indication and description of methods of conducting classes and verification of achievement of learning outcomes and method of documentation)

|  |  |  |  |
| --- | --- | --- | --- |
| Subject Effects | Teaching methods | Methods of verifying learning outcomes | Documentation methods |
| KNOWLEDGE |
| W1-W5 | Knowledge provided during the laboratory | Final Project: Comprehensive analysis of a selected data setDescription:Students analyze a selected data set (e.g., public data, financial, healthcare, or technology data). The project should include:* Data preparation (cleaning, missing data, transformation).
* Data mining (descriptive statistics, visualizations).
* Conducting problem analysis, e.g. correlations, regression models, classification.
* Presentation of results in the form of visualization and report.
 | The project was placed on the platform |
| SKILLS |
| U1-U5 | Practical classes performed at computer stations | Final Project: Comprehensive analysis of a selected data setDescription:Students analyze a selected data set (e.g., public data, financial, healthcare, or technology data). The project should include:* Data preparation (cleaning, missing data, transformation).
* Data mining (descriptive statistics, visualizations).
* Conducting problem analysis, e.g. correlations, regression models, classification.

Presentation of results in the form of visualization and report. | The project was placed on the platform |
| SOCIAL COMPETENCES |
| K1-K2 | Practical classes performed at computer stations | Final Project: Comprehensive analysis of a selected data setDescription:Students analyze a selected data set (e.g., public data, financial, healthcare, or technology data). The project should include:* Data preparation (cleaning, missing data, transformation).
* Data mining (descriptive statistics, visualizations).
* Conducting problem analysis, e.g. correlations, regression models, classification.

Presentation of results in the form of visualization and report. | The project was placed on the platform |

3.6. Assessment criteria for the achieved learning outcomes

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Learning effect | For a grade of 3 or "pass."the student knows and understands/is able to/is ready to | For a grade of 3.5, the student knows and understands/is able to/is ready to | For a grade of 4, the student knows and understands/is able to/is ready to | For a grade of 4.5, the student knows and understands/is able to/is ready to | For a grade of 5, the student knows and understands/is able to/is ready to |
| W | 51-60% of knowledge indicated in learning outcomes | 61-70% of knowledge indicated in learning outcomes | 71-80% of knowledge indicated in learning outcomes | 81-90% of knowledge indicated in learning outcomes | 91-100% of knowledge indicated in learning outcomes |
| U | 51-60% of skills indicated in learning outcomes | 61-70% of skills indicated in learning outcomes | 71-80% of skills indicated in learning outcomes | 81-90% of skills indicated in learning outcomes | 91-100% of skills indicated in learning outcomes |
| K | 51-60% of skills indicated in learning outcomes | 61-70% of skills indicated in learning outcomes | 71-80% of skills indicated in learning outcomes | 81-90% of skills indicated in learning outcomes | 91-100% of skills indicated in learning outcomes |

3.7. Literature

**Basic**

1. Baka Benjamin; Python Data Structures and Algorithms. Improve application performance with graphs, stacks, and queues; Packt Publishing; Birmingham 2017
2. Finch Victor Data Analytics For Beginners: Your Ultimate Guide To Learn And Master Data Analysis. Get Your Business Intelligence Right – Accelerate Growth And Close More Sales Createspace Independent Publishing Platform; Scotts Valley; 2017
3. McKinney W., "Python w analizie danych. Przetwarzanie danych za pomocą pakietów pandas i NumPy oraz środowiska Jupyter. Wydanie III", Helion, 2023.

**Supplementary:**

1. Gagolewski M., "Przetwarzanie i analiza danych w języku Python", Wydawnictwo Naukowe PWN, 2021.
2. Muller Andreas C. ; Guido Sarah; Machine Learning : Python i data science; Helion; Gliwice; 2023

4. Student workload - ECTS points balance

|  |  |
| --- | --- |
| **Types of student activity** | **Student Load** |
| **ST** | **NST** |
| **Classes requiring direct contact between the student and the academic teacher at the university premises** | **30** | **15** |
| Classes included in the study plan | 30 | 15 |
| **Student's own work** | **45** | **60** |
| Ongoing preparation for classes, preparation of project work/presentations/etc. | 25 | 30 |
| Preparation for passing classes | 20 | 30 |
| **TOTAL STUDENT HOURLY LOAD** | **75** | **75** |
| **Number of ECTS points** | **3** | **3** |

|  |  |
| --- | --- |
| Last change date | 30/09/2024 |
| The changes were introduced | INF Education Quality Team |
| The changes were approved | Arkadiusz Gwarda, M.A. |