|  |  |
| --- | --- |
| **Prowadzący** | dr Beata Żukowska |
| **Oferta PJO\*** | ~~TAK /~~ NIE\*\* |
| **Oferta PJOE\*** | TAK / ~~NIE\*\*~~ |
| **Kierunek, rok, stopień dla PJO (\*obowiązkowe)** |  |
| **Semestr roku 2024/2025** | zimowy / ~~letni\*\*~~ |

\* PJO – przedmiot w języku obcym dla studentów polskich / PJOE – przedmiot w języku obcym dla studentów Erasmus+
\*\* zostawić właściwe

BASIC INFORMATION ABOUT THE SUBJECT (INDEPENDENT OF THE CYCLE)

|  |  |
| --- | --- |
| **Module name** | Introduction to Data Science |
| **Erasmus code** |  |
| **ISCED code** |  |
| **Language of instruction** | English |
| **Website** | [https://www.umcs.pl/en/courses-in-english-2021-2022,21582.htm](https://www.umcs.pl/en/courses-in-english-2021-2022%2C21582.htm) (dla PJOE) |
| **Prerequisites** | Basic knowledge of statistics |
| **ECTS points hour equivalents** | Contact hours (work with an academic teacher): 30Total number of hours with an academic teacher: 30Number of ECTS points with an academic teacher: 4Non-contact hours (students' own work): 40Total number of non-contact hours: 40Number of ECTS points for non-contact hours: 2Total number of ECTS points for the module: 6 |
| **Educational outcomes verification methods** | In-class activityPractical exercicesData analysis project |
| **Description** | Lecture with workshop introducing main concepts of data science. Students will be provided with practical tools on how to prepare and analyze data for machine learning models. During the course basics of R-programming will be covered. No previous knowledge of R is necessary.  |
| **Reading list** | 1. R.A. Irziarry (2019), Introduction to Data Science, CRC Press.
2. A. Shipunov (2019), Visual Statistics. Use R!, available: <https://cran.r-project.org/>
3. H. Wickham, G. Grolemund (2017), R for Data Science, available: <https://r4ds.had.co.nz/>,
4. <http://www.cookbook-r.com/>
5. R Documentation – available: <https://cran.r-project.org/>
6. Other articles and data provided by lecturer.
 |
| **Educational outcomes** | KNOWLEDGEA student will know:* what is the aim of data science projects
* the difference between data analysis and data science
* the concept of machine learning

SKILLSA student will be able to:* import and manipulate data in R
* clean and prepare data for modelling
* visualize and discuss data

ATTITUDESA student will be:* ready to deal with big datasets and conclude about them
* prepared to work as a member of data science project team
 |
| **Practice** | n/a |

INFORMATION ABOUT CLASSES IN THE CYCLE

|  |  |
| --- | --- |
| **Website** | [https://www.umcs.pl/en/courses-in-english,21103.htm](https://www.umcs.pl/en/courses-in-english%2C21103.htm) (dla PJOE) |
| **Educational outcomes verification methods** | In-class activityPractical exercisesData analysis project |
| **Comments** |  |
| **Reading list** | 1. R.A. Irziarry (2019), Introduction to Data Science, CRC Press.
2. A. Shipunov (2019), Visual Statistics. Use R!, available: <https://cran.r-project.org/>
3. H. Wickham, G. Grolemund (2017), R for Data Science, available: <https://r4ds.had.co.nz/>,
4. <http://www.cookbook-r.com/>
5. R Documentation – available: <https://cran.r-project.org/>
6. Other articles and data provided provided by lecturer.
 |
| **Educational outcomes** | KNOWLEDGEA student will know:* what is the aim of data science projects
* the difference between data analysis and data science
* the concept of machine learning

SKILLSA student will be able to:* import and manipulate data in R
* clean and prepare data for modelling
* visualize and discuss data

ATTITUDESA student will be:* ready to deal with big datasets and conclude about them
* prepared to work as a member of data science project team
 |
| **A list of topics** | 1. Understanding data science – roles and tools
2. Getting started with R and RStudio
3. Basics of R – vectors, matrices, factors, lists, data frames
4. Programming basics – functions and loops
5. Importing data to R (with elements of webscrapping)
6. Data wrangling with tidyverse packages
7. Data cleaning, working with dates and time
8. Data visualization – basics and best practices
9. Exploratory data analysis in R
10. Methods of sampling
11. Machine learning – general introduction
12. Communicating the data and models
 |
| **Teaching methods** | lecture, case studies, exercises, gamification |
| **Assessment methods** | Participation and in-class activity – 40%Practical exercises – 10%Data analysis project – 50% |