

**Course title:** **Data Mining with R**  
**Institute/Division:** F-3, Institute of Computer Science

**Course code:**

**Erasmus subject code:**

**Number of contact hours:** 45 hours

**Course duration:** 1 semester

**ECTS credits:** 6

**Course description:** The students will learn the most current techniques to process large datasets using the very popular and very simple R language, which will be used to extract valuable knowledge from data, to classify and to cluster it within datasets and to perform analytical predictions, based on the available datasets.

**Literature:**

1. Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer, Sixth edition, Springer, 2015, available on the Internet: [http://www-bcf.usc.edu/~gareth/ISL/ISLR\\_Sixth\\_Printing.pdf](http://www-bcf.usc.edu/~gareth/ISL/ISLR_Sixth_Printing.pdf)
2. Yanchang Zhao, R and Data Mining: Examples and Case Studies, Elsevier, 2012-2014, available on the Internet together with source codes of all examples: <http://www.RDataMining.com>
3. Many additional materials and links will be given to all students during the lectures.

**Course type:** Lectures and exercises

**Assessment method:** There will be a few homeworks. Each one should be returned to the teacher up to 2 weeks after distribution.

**Prerequisites:** None

**Primary target group:** computer science students of the 3rd or 4th year

**Lecturer:** Barbara Borowik, PhD

**Contact person:** Barbara Borowik, PhD; [bborowik@pk.edu.pl](mailto:bborowik@pk.edu.pl)

**Deadline for application:** 15th of September