GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	DATA ANALYSIS - A TEAM PROJECT, PG_00060793
Field of study	Economic Analytics

Date of commencement of studies	October 2023	**Academic year of realisation of subject**	2023/2024
Education level	second-cycle studies	**Subject group**	Obligatory subject group in the field of study
			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	1	**Language of instruction**	English
Semester of study	1	**ECTS credits**	5.0
Learning profile	general academic profile	**Assessment form**	assessment

Conducting unit	Katedra Statystyki i Ekonometrii -> Faculty of Management and Economics

Name and surname of lecturer (lecturers)	Subject supervisor	dr inż. Karol Flisikowski
	Teachers	dr inż. Karol Flisikowski

Lesson types and methods of instruction	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	45.0	0.0	0.0	60
	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	60	5.0	60.0	125

Subject objectives	Uses advanced tools for processing raw economic and social data, which are then used in in-depth statistical analysis, carrying out tasks in the form of a team project

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K7_U01] creates innovative solutions to complex and unstructured problems, taking into account the variability of the environment by synthesising information from many sources	creates innovative solutions to complex problems, taking into account the influence of many factors on the studied phenomenon, synthesizing data from many sources	[SU3] Assessment of ability to use knowledge gained from the subject
	[K7_U05] cooperates with other people in the implementation of teamwork, both as a leader and a team member, effectively achieving the assumed goals	performs analytical work demonstrating the ability to work in a team	[SU4] Assessment of ability to use methods and tools
	[K7_W03] demonstrates in-depth preparation in the application of analytical methods and techniques for formulating and solving problems	correctly apply modern methods of pre-processing and raw data processing before applying advanced analytical methods, ensuring the reliability of the results	[SW1] Assessment of factual knowledge

Subject contents	Introduction to R, R-studio. Basic operations. Data import from various formats. Measuring scales vs data types in R (vector, dataframe, matrix, list, etc.) Functions, variables, operators, constants. Loops. Conditional expressions and their use in data analysis Basic commands - descriptive statistics Basic commands - mathematical statistics Reporting in R-Markdown Basic data processing (new variables, filters, combining frames: reshape, split, combine) Imputation methods for missing cross-sectional and temporal data Dirty data - missing observations; duplicates; outliers; format errors Data cleaning using Dplyr and Tidyr Data cleaning outliers Transformations and discretization of variables Data sources: downloading data from databases (sqlite); web scraping; downloading data to R (Yahoo Finance; Quandl; Google Trends, Eurostat etc.) Dimensional reduction using principal component analysis (PCA). Example applications Graphics in R basic and advanced graphical presentation of data (packages: ggplot2; Lattice; Grid) Publishing reports directly from R introduction to R-Markdown (notebook; presentations R and Powerpoint; HTML slides; PDF beamer etc.) Final project. Presentations

Prerequisites and co-requisites	

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Project	60.0%	60.0%
Test	60.0%	40.0%

Recommended reading

Basic literature	Podstawy statystyki z przykładami w R, Tomasz Górecki, Wydawnictwo BTC, 2011 Przewodnik po pakiecie R, Przemysław Biecek, GIS, 2014
Supplementary literature	https://cran.r-project.org/web/packages/IPSUR/vignettes/IPSUR.pdf - G. Jay Kerns, Introduction to Probability and Statistics using R, Third Edition, 2018
eResources addresses	Adresy na platformie eNauczanie: Data Analysis 2023/24 - Moodle ID: 31969 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=31969

Example issues/ example questions/ tasks being completed	Final project: preparation of a report and presentation in R-Markdown after data processing and analysis in R
Work placement	Not applicable