



Subject card

Subject name and code	DATA ANALYSIS, PG_00049639						
Field of study	Economic Analytics						
Date of commencement of studies	October 2021	Academic year of realisation of subject			2021/2022		
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			Polish		
Semester of study	1	ECTS credits			4.0		
Learning profile	general academic profile	Assessment form			assessment		
Conducting unit	Department of Economic Sciences -> 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	0.0	0.0	30.0	0.0	0.0	30
	E-learning hours included: 0.0						
Analiza danych S 2021_22 - Moodle ID: 17612 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=17612							
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	30		8.0		62.0	100
Subject objectives	Students will be familiar to the R ecosystem and learn how to use R for the most common data analysis tasks, including loading, cleaning, transforming, summarizing and visualizing data.						
Learning outcomes	Course outcome	Subject outcome			Method of verification		
	[K7_U06] has a broad knowledge of methods and tools for acquiring and collecting data, as well as analysing, explaining and reasoning on socio-economic phenomena and processes.	Student has the ability to pre-process and process raw data (conversion, change of dimensions, cleaning, etc.).			[SU2] Assessment of ability to analyse information [SU4] Assessment of ability to use methods and tools		
	[K7_W07] has an in-depth knowledge on methods of social and economic phenomena description, including market information acquisition techniques and methods of analysis and modelling of economic processes	Student has knowledge of the methods of statistical analysis of data and methods of processing, converting, cleaning raw data.			[SW1] Assessment of factual knowledge		
	[K7_K01] understands the need for continuous learning and, in particular, for advanced and modern tools for data analysis	Student is aware that for the correct data analysis it is necessary to properly format the data, variables, clean observations from outliers, dirty data, etc.			[SK5] Assessment of ability to solve problems that arise in practice [SK1] Assessment of group work skills		
	[K7_U09] has the ability to use advanced mathematical tools to analyse and assess economic phenomena and to make decisions by economic operators	Student has the ability to use advanced raw data processing tools in order to then subject them to appropriate statistical analysis.			[SU2] Assessment of ability to analyse information [SU4] Assessment of ability to use methods and tools		

Subject contents	<ol style="list-style-type: none"> 1. Introduction to R, R-studio. Basic operations. Import of data from various formats. Measuring scales and data types in R (vector, data frame, matrix, list, etc.). 2. Functions, variables, operators, constants. Loops. Conditional expressions and their application in data analysis. 3. Basic commands - descriptive statistics. 4. Basic commands - mathematical statistics. 5. Reporting in R-Markdown. 6. Basic data processing (new variables, filters, joining frames: reshape, split, combine). 7. Imputation methods for missing cross-sectional and time series data. 8. Dirty data - missing observations; duplicate; outliers; formatting errors. Data cleaning using Dplyr and Tidy. 9. Data cleaning - outliers. 10. Transformations and discretization of variables. 11. Data sources: downloading data from databases (sqlite); web scraping; downloading data to R (Yahoo Finance; Quandl; Google Trends, Eurostat etc.). 12. Dimension reduction using principal component analysis (PCA). Application examples. 13. Graphics in R - basic and advanced graphic data presentation (packages: ggplot2; Lattice; Grid). 14. Publishing reports directly from R - introduction to R-Markdown (notebook; presentations - R and Powerpoint; HTML slidy; PDF beamer etc.). 15. Final project. Presentations. 											
Prerequisites and co-requisites	Descriptive and mathematical statistics											
Assessment methods and criteria	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr> <th style="width: 33%;">Subject passing criteria</th> <th style="width: 33%;">Passing threshold</th> <th style="width: 33%;">Percentage of the final grade</th> </tr> </thead> <tbody> <tr> <td>moodle quizzes</td> <td>60.0%</td> <td>50.0%</td> </tr> <tr> <td>final project</td> <td>60.0%</td> <td>50.0%</td> </tr> </tbody> </table>			Subject passing criteria	Passing threshold	Percentage of the final grade	moodle quizzes	60.0%	50.0%	final project	60.0%	50.0%
Subject passing criteria	Passing threshold	Percentage of the final grade										
moodle quizzes	60.0%	50.0%										
final project	60.0%	50.0%										
Recommended reading	Basic literature	<ol style="list-style-type: none"> 1. Discovering statistics using R, Andy Field, Jeremy Miles, Zoe Field, Sage, 2012. 2. Statistics (The easier way) with R, Nicole M. Radziwill, 2016. 										
	Supplementary literature	<ul style="list-style-type: none"> • 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											
Example issues/ example questions/ tasks being completed	Final project: micro-data analysis and presentation prepared in R-Markdown after fully preprocessed data in R (with codes).											
Work placement	Not applicable											