



Subject card

Subject name and code	Descriptive statistics, PG_00045293						
Field of study	Data Engineering						
Date of commencement of studies	October 2022	Academic year of realisation of subject			2022/2023		
Education level	first-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	2	ECTS credits			3.0		
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	15.0	0.0	0.0	30
	E-learning hours included: 0.0						
Descriptive Statistics - summer 2023 - Moodle ID: 27804 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=27804							
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		4.0		41.0	75
Subject objectives	The main aim of the course is to acknowledge students with the statistical analysis based on the sample data using R & R-studio.						
Learning outcomes	Course outcome	Subject outcome			Method of verification		
	[K6_W09] has basic knowledge of the nature of economic sciences and ways of its description with IT tools	Student has a basic knowledge of descriptive statistics, methods of presenting data of various types and statistical modeling (for cross-sectional and time series data) and syntax of the R programming language.			[SW1] Assessment of factual knowledge		
	[K6_K05] understands the need for self-improvement through systematic acquisition of knowledge and skills.	The student is aware of the diversity of statistical variables and types of data in R & R-studio environment . The student consistently improves his programming skills.			[SK2] Assessment of progress of work [SK3] Assessment of ability to organize work		
[K6_U11] is able to use mathematical and IT tools in economics.	Student selects the appropriate statistical methods to perform the analysis for the selected statistical variable using the R programming language.			[SU4] Assessment of ability to use methods and tools [SU1] Assessment of task fulfilment			

Subject contents	<p>Part 1. Introduction to R & R-studio.</p> <p>1. Types of data, basic programming skills (R)</p> <p>2. Types of statistical variables, basic operations, graphs</p> <p>3. Methods of raw data aggregating</p> <p>Part 2. Describing univariate data</p> <p>4. Measures of central tendency, dispersion, assymetry (skewness) and kurtosis, inequality & concentration + aggregating data (in respect to the type of variable) + graphs</p> <p>Part 3. Describing bivariate data</p> <p>5. Correlation analysis for pair of quantitative variables + graphs</p> <p>6. Rank correlation</p> <p>7. Two-way tables, correlation analysis of qualitative variables + graphs</p> <p>Part 4. Regression analysis</p> <p>8. OLS. Simple regression, multiple linear regression.</p> <p>9. Nonlinear regression.</p> <p>10. Regression analysis for qualitative dependend variable (optional)*.</p> <p>Part 5. Time series</p> <p>11. Describing time series graphs, tables, dynamics & indices.</p> <p>12. Decompositions. Time series models (MA, AR, ARMA).</p>											
Prerequisites and co-requisites	Mathematics, English (intermediate level), basic programming skills.											
Assessment methods and criteria	<table border="1"> <thead> <tr> <th data-bbox="453 1590 794 1619">Subject passing criteria</th> <th data-bbox="799 1590 1141 1619">Passing threshold</th> <th data-bbox="1145 1590 1473 1619">Percentage of the final grade</th> </tr> </thead> <tbody> <tr> <td data-bbox="453 1626 794 1655">Laboratory - final test and projects</td> <td data-bbox="799 1626 1141 1655">60.0%</td> <td data-bbox="1145 1626 1473 1655">50.0%</td> </tr> <tr> <td data-bbox="453 1662 794 1691">Lecture (final exam)</td> <td data-bbox="799 1662 1141 1691">60.0%</td> <td data-bbox="1145 1662 1473 1691">50.0%</td> </tr> </tbody> </table>			Subject passing criteria	Passing threshold	Percentage of the final grade	Laboratory - final test and projects	60.0%	50.0%	Lecture (final exam)	60.0%	50.0%
Subject passing criteria	Passing threshold	Percentage of the final grade										
Laboratory - final test and projects	60.0%	50.0%										
Lecture (final exam)	60.0%	50.0%										
Recommended reading	<p>Basic literature</p> <p>Supplementary literature</p> <p>eResources addresses</p>	<p>1. Statistics (The easier way) with R, Nicole M. Radziwill, 2016.</p> <p>2. Discovering statistics using R, Andy Field, Jeremy Miles, Zoe Field, Sage, 2012.</p> <p>1. Statistics for Business and Economics, McClave Benson Sincich, Pearson, 2008.</p> <p>2. Using R for Introductory Statistics, John Verzani, Chapman and Hall, 2000.</p>										
Example issues/ example questions/ tasks being completed	<p>1. How can we recognize which model of time series decomposition we analyze - additive or multiplicative? Give one example for each of them.</p> <p>2. Sketch the bimodal distribution of the discrete variable.</p> <p>3. How can we deal with a large number of missing cases? List three methods.</p>											
Work placement	Not applicable											