GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	DATA ANALYSIS IN PYTHON, PG_00070513		
Field of study	Economic Analytics		
Date of commencement of studies	October 2026	Academic year of realisation of subject	2026/2027
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	2	ECTS credits	4.0
Learning profile	general academic profile	Assessment form	assessment
Conducting unit	Department of Statistics and Econometrics -> Faculty of Management and Economics -> Faculties of Gdańsk University of Technology		
Name and surname of lecturer (lecturers)	Subject supervisor	dr inż. Karol Flisikowski	
	Teachers		

Lesson types	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	0.0	0.0	45.0	0.0	0.0	45
	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	45	3.0	52.0	100

Subject objectives	preparing students to develop innovative analytical solutions using Python for exploratory data analysis, while fostering independence, creativity, and a responsible approach to data analysis in the context of complex, unstructured socio-economic problems.

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K7_W03] demonstrates in-depth knowledge of the applications of analytical methods and techniques for formulating and solving socio-economic problems.	has knowledge and understanding of contemporary methods of preprocessing and processing raw data prior to the application of advanced analytical techniques, in order to ensure the reliability of results.	[SW1] Assessment of factual knowledge
	[K7_K01] is ready to critically evaluate his/her knowledge in economic analytics and seek expert opinions when facing difficulties in solving a problem independently.	is ready to critically reflect on data analysis results and their own skills in Python, and to seek support from experts and available sources when facing difficulties.	[SK2] Assessment of progress of work [SK5] Assessment of ability to solve problems that arise in practice
	[K7_U03] formulates research hypotheses and select appropriate methods for their verification using advanced it tools.	is able to formulate research hypotheses and select appropriate methods for their verification using analytical tools in Python.	[SU4] Assessment of ability to use methods and tools [SU5] Assessment of ability to present the results of task

Subject contents	Course content – laboratory 1. Introduction to Python basic operations 2. Data sources and importing data from various formats in Python 3. Variables and data types in Python (vectors, data frames, matrices, lists, arrays, sets, dictionaries) 4. Basic functions descriptive statistics and mathematical operations in Python 5. Basic data processing (creating new variables, filtering, merging data frames: transformation, splitting, joining) in Python 6. Dirty data issues missing values, duplicates, outliers, formatting errors; Python libraries for data handling (e.g., naniar equivalent tools) 7. Data processing and transformation in Python (tidy data principles and data reshaping) 8. Data cleaning handling outliers and missing values, imputation, transformations, discretization, optimal binning in Python 9. Data visualization in Python basic and advanced graphical presentation (e.g., Matplotlib, Plotly) 10. Reporting analyses in Python introduction to Quarto 11. Interactive applications in Python (e.g., Shiny for Python) 12. Machine learning linear models, generalized linear models, iterative model fitting, and reporting 13. k-Nearest Neighbors (KNN) method 14. Classification and clustering linear discriminant analysis, decision trees, logistic regression 15. Bagging and random forests 16. Boosting methods, including XGBoost
Prerequisites and co-requisites	

Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	Team project	60.0%	50.0%
	Test	60.0%	50.0%

Recommended reading	Basic literature	1. Chantal D. Larose, Daniel T. Larose, Data Science Using Python and R, Wiley, 2019 2. Rick J. Scavetta, Boyan Angelov, Python and R for the Modern Data Scientist, O'Reilly Media, 2021
	Supplementary literature	1. Ajay Ohri, Python for R Users: A Data Science Approach, Wiley, 2017 2. Hadley Wickham, Garrett Grolemund. R for Data Science, https://r4ds.had.co.nz 3. J. Hathaway, Katie Larson, Python for Data Science, https://byuidatascience.github.io/python4ds/
	eResources addresses	

Example issues/ example questions/ tasks being completed	• Data preprocessing in Python languages • Exploratory Descriptive Data Analysis (DEA) report in Python • Statistical inference from a sample in Python
Practical activites within the subject	Not applicable

Document generated electronically. Does not require a seal or signature.