



Subject card

Subject name and code	DATA ANALYSIS IN PYTHON, PG_00070337						
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	Part-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	27.0	0.0	0.0	27
	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	27	3.0	70.0	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_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		
	[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		

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	<table border="1" data-bbox="448 624 1487 730"> <thead> <tr> <th data-bbox="448 624 798 658">Subject passing criteria</th> <th data-bbox="802 624 1139 658">Passing threshold</th> <th data-bbox="1144 624 1487 658">Percentage of the final grade</th> </tr> </thead> <tbody> <tr> <td data-bbox="448 665 798 694">Team project</td> <td data-bbox="802 665 1139 694">60.0%</td> <td data-bbox="1144 665 1487 694">50.0%</td> </tr> <tr> <td data-bbox="448 701 798 730">Test</td> <td data-bbox="802 701 1139 730">60.0%</td> <td data-bbox="1144 701 1487 730">50.0%</td> </tr> </tbody> </table>			Subject passing criteria	Passing threshold	Percentage of the final grade	Team project	60.0%	50.0%	Test	60.0%	50.0%
Subject passing criteria	Passing threshold	Percentage of the final grade										
Team project	60.0%	50.0%										
Test	60.0%	50.0%										
Recommended reading	<table border="1" data-bbox="448 736 1487 1021"> <tbody> <tr> <td data-bbox="448 736 798 837">Basic literature</td> <td colspan="2" data-bbox="802 736 1487 837">           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         </td> </tr> <tr> <td data-bbox="448 844 798 992">Supplementary literature</td> <td colspan="2" data-bbox="802 844 1487 992">           1. Ajay Ohri, Python for R Users: A Data Science Approach, Wiley, 2017            2. Hadley Wickham, Garrett Grolemund. R for Data Science, <a href="https://r4ds.had.co.nz">https://r4ds.had.co.nz</a>            3. J. Hathaway, Katie Larson, Python for Data Science, <a href="https://byuidatascience.github.io/python4ds/">https://byuidatascience.github.io/python4ds/</a> </td> </tr> <tr> <td data-bbox="448 999 798 1021">eResources addresses</td> <td colspan="2" data-bbox="802 999 1487 1021"></td> </tr> </tbody> </table>			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, <a href="https://r4ds.had.co.nz">https://r4ds.had.co.nz</a> 3. J. Hathaway, Katie Larson, Python for Data Science, <a href="https://byuidatascience.github.io/python4ds/">https://byuidatascience.github.io/python4ds/</a>		eResources addresses		
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, <a href="https://r4ds.had.co.nz">https://r4ds.had.co.nz</a> 3. J. Hathaway, Katie Larson, Python for Data Science, <a href="https://byuidatascience.github.io/python4ds/">https://byuidatascience.github.io/python4ds/</a>											
eResources addresses												
Example issues/ example questions/ tasks being completed	<ul data-bbox="448 1028 1487 1126" style="list-style-type: none"> <li>• Data preprocessing in Python languages</li> <li>• Exploratory Descriptive Data Analysis (DEA) report in Python</li> <li>• Statistical inference from a sample in Python</li> </ul>											
Practical activities within the subject	Not applicable											

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