



Subject card

Subject name and code	Business Data Analytics, PG_00053096						
Field of study	Business Data Analytics						
Date of commencement of studies	October 2023		Academic year of realisation of subject		2025/2026		
Education level	first-cycle studies		Subject group		Optional subject group 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	3		Language of instruction		English		
Semester of study	5		ECTS credits		4.0		
Learning profile	general academic profile		Assessment form		assessment		
Conducting unit	Department of Informatics In Management -> Faculty of Management and Economics -> Wydział Politechniki Gdańskiej						
Name and surname of lecturer (lecturers)	Subject supervisor		dr Nina Rizun				
	Teachers		dr Nina Rizun				
Lesson types	Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
	Number of study hours	15.0	0.0	30.0	0.0	0.0	45
	E-learning hours included: 0.0						
	eNauczanie source addresses: Moodle ID: 46763 Business Data Analytics-2025/2026 https://enauczanie.pg.edu.pl/moodle/course/view.php?id=46763						
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		5.0		50.0	100
Subject objectives	The aim of the course is to prepare students to apply methods and tools of data analytics to generate actionable business insights from structured and unstructured data, based on knowledge of text analytics, clustering, topic modeling, sentiment analysis, and social network analysis, and to foster attitudes of critical thinking, ethical responsibility, sensitivity to data biases, and readiness for interdisciplinary collaboration, in the context of data-driven decision-making for solving business and societal challenges.						
Learning outcomes	Course outcome		Subject outcome		Method of verification		
	[K6_W08] Knows the models and structure of the data mining process and their multidimensional analysis and can assess the results of such analyses		The student knows the models and structure of data analytics processes, multidimensional analysis techniques, and is able to explain their application for evaluating the quality and reliability of business and text analytics results		[SW2] Ocena wiedzy zawartej w prezentacji [SW3] Ocena wiedzy zawartej w opracowaniu tekstowym i projektowym		
	[K6_K05] understands the need for self-improvement through systematic acquisition of knowledge and skills.		The student demonstrates readiness to continuously develop knowledge and skills in data analytics, critically evaluates their own work, and improves solutions during laboratory work and projects		[SK1] Ocena umiejętności pracy w grupie [SK2] Ocena postępów pracy [SK5] Ocena umiejętności rozwiązywania problemów występujących w praktyce		
	[K6_U01] programs in procedural, object, functional and logic programming languages, codes programs at the processor instruction level, runs and tests programs.		The student is able to program and implement selected data mining algorithms in R, run and test the code, and apply it to the analysis of textual and numerical datasets		[SU2] Ocena umiejętności analizy informacji [SU4] Ocena umiejętności korzystania z metod i narzędzi		

Subject contents	LECTURE:		
	1. Introduction to Data Analytics: What is Data Analytics? Data can help us solve specific problems. KDD Process. What techniques do people apply to data? Data Biases. 2. Introduction to Text Analytics. What is Text Analytics? Why is Analytics relevant? Why do we study it? Application domains. The complexity of unstructured text (the origin of Analytics challenges). Bag-of-words representation of text. Vector Space Model. Methods/techniques for text pre-processing. Zipfs Law. 3. F-IDF Transformation. Clustering. Term frequency and weighting. TF-IDF weighting. Clustering: Theoretical background. Euclidian Distance. Hierarchical Clustering. K-means Clustering. 4. Measures of Semantic Similarity Between the Documents. Vector Space Model. Cosine Similarity. Similarity Measure for Text Mining. Tf-IDF and Cosine similarity Example. Adjacency Matrix. Cosine Similarity as a Weight of Graph Edges. Community Detection Algorithms Within the Cosine Similarity Graph. 5. Multidimensional Scaling. Latent Semantic Analysis. Latent Dirichlet Allocation. Topic Modeling. Latent Semantic Analysis: Motivation. Singular Value Decomposition (SVD). LSA-based Similarity Search. Latent Dirichlet allocation. Topic modeling. 6. Introduction to Sentiment Analysis. 7. Structural Topic Modeling. 8. Social Network Analysis: Theory and Applications		
	Laboratoria:		
	1. Basic Text Analytics in R 2. TF-IDF transformation. Cluster analysis using R 3. Cosine similarity. Network graphs using R 4. Topic Modeling. Multidimensional Scaling and Latent Semantic Analysis 5. Opinion Mining 6. Structural Topic Modelling 7. Social Network Analysis. Basic Network Measures		
Prerequisites and co-requisites			
Assessment methods and criteria	Subject passing criteria	Passing threshold	Percentage of the final grade
	Project	60.0%	40.0%
	Final test	60.0%	40.0%
	Excercises classes	60.0%	20.0%
Recommended reading	Basic literature	1. Data mining and knowledge discovery handbook. Editors Oded Maimon, Lior Rokach. Springer, New York, 2005. 2. Making Hard Decisions with Decision Tools 3rd Edition, Robert T. Clemen and Terence Reilly, South-Western Cengage Learning, 2014.	
	Supplementary literature	1. Bramer, M. A. Principles of data mining. Springer, London, 2007. Additional: 2. Han, J., Kamber, M. Data mining: concepts and techniques. 2nd ed. Morgan Kaufmann, San Francisco, 2006. 3. Berry, M., Linnof, G. Data mining techniques: for marketing, sales, and customer relationship management. 2nd ed. Wiley, Indianapolis, 2004. 4. Cox, E. Fuzzy modelling and genetic algorithms for data mining and exploration. Morgan Kaufman, Amsterdam, 2005. 5. Advances in knowledge discovery and data mining. Editors Usama M. Fayyad et al. AAAI, Menlo Park, 1996.	
	eResources addresses		
Example issues/ example questions/ tasks being completed	Examples of Test Questions: • Choose the main properties of Euclidian distance • Choose the right value of the Cosine similarity measure between Document 1 and Document 2 • K-means algorithm allows ... • What's the difference between Euclidian distance and Cosine similarity measure? • Person with high Eigenvector centrality is ... Examples of Open Questions: • Please calculate the IDF for the terms from the following Corpus • Build the Zipfslaw chart for the following Corpus • Give the general characteristics of the actors of the following Social network		

Practical activities within the subject	Not applicable
---	----------------

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