University of Gdańsk

Subject card

Subject name and code	Data processing, PG_00135485		
Field of study	Physical geography and geoinformation		
Date of commencement of studies	October 2024	Academic year of realisation of subject	2024/2025
Education level	postgraduate studies	Subject group	Obligatory subject group 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	2.0
Learning profile	academic	Assessment form	
Conducting unit	Pracownia Rekonstrukcji Geomorfologicznych -> Katedra Geomorfologii i Geologii Czwartorzędu -> Faculty of Oceanography and Geography		
Name and surname of lecturer (lecturers)	Subject supervisor	dr Maurycy Żarczyński	
	Teachers	dr Maciej Markowski dr Maurycy Żarczyński	

Lesson types	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						

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	15.0	15.0	60

Subject objectives	1. Familiarization with basics of the R language and methods of data processing; 2. Familiarization with ETL (Extract, Load, Transformation) tools on example of FME. Rules for model creation, data processing and data manipulation processes automatization.

Learning outcomes

Course outcome	Subject outcome	Method of verification
[GFGMU2_U02] precisely and appropriately use terminology in the field of physical geography and geoinformation in oral statements and written works	Can prepare a work using specialist language used in physical geography and geoinformatics.	[SU2] presentation/project/paper/ report
[GFGMU2_W04] theoretical foundations of research methods used in physical geography and closely related sciences, descriptive and mathematical statistics, as well as advanced methods of analyzing spatial phenomena	Knows statistical methods used in analyses of the environmental data.	[SW4] test/exam - oral or written
[GFGMU2_U04] describe and analyze the causes and course of physical and geographical processes and phenomena, selecting and applying advanced techniques and research tools in the field of statistical and geoinformation methods, interpreting the results, using theoretical knowledge to formulate own opinions and conclusions	Can carry statistical and spatial analysis of spatial and numeric data.	[SU2] presentation/project/paper/ report
[GFGMU2_U03] effectively use selected scientific literature in the field of physical geography and geoinformation, both in Polish and English	Can choose appropriate statistical method based on the literature.	[SU2] presentation/project/paper/ report
[GFGMU2_K01] critical assessment of knowledge in the field of Earth and environmental sciences and geoinformation, its completion and verification through critical analysis of scientific literature	Exhibits critical perception of their achievements.	[SK2] presentation/project/paper/ report
[GFGMU2_K03] accepting responsibility for group work assuming various roles in it, participating in preparation of scientific projects, taking responsibility for the equipment and safety rules, active developing of professional competences and knowledge in Earth and environmental sciences and geoinformation, including interdisciplinarity, as well as developing the principles of professional ethics, respecting copyright rules	Exhibits ability to work in a group.	[SK2] presentation/project/paper/ report [SK8] observation of student's independent or team work

Subject contents

1. Introduction to R language and RStudio environment;
2. Data input and import (text data, tabular data and spatial data);
3. Data cleaning (tidyr, dplyr and other);
4. Basics programming in R language;
5. Indexing, loops, functions;
6. Data preparation and processing, basic statistical modeling approaches;
7. Model assessment procedures (model appraisal, cross-validation, bootstrapping);
8. Practical use of selected statistical methods;
9. Introduction and familiarization with basic functions of the ETL tool FME (Feature Manipulation Engine) by SAFE Software;
10. Integration of the various spatial data types as an example of batch data transformation process into single output;
11. Automatization of the data processing using variable data sources (raster, vector, tabular) with geometry and attributes manipulation with the FME Workbench.

Prerequisites and co-requisites

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Term works (FME)	51.0%	23.34%
Term work (R)	51.0%	46.66%
Test	51.0%	30.0%

Recommended reading	Basic literature	Komsta, 2004, Wprowadzenie do środowiska R - R Project - https://cran.r-project.org/doc/contrib/Komsta-Wprowadzenie.pdf Biecek P., 2014, Przewodnik po pakiecie R, GIS, Wrocław - http://biecek.pl/r/przewodnikpopakiecierwydanieiiiinternet.pdf Wickham H., Çetinkaya-Rundel M., Grolemund G., 2024, R for Data Science. 2nd Edition, O'Reilly Media
	Supplementary literature	FME Transfomer Reference Guide: https://cdn.safe.com/resources/fme/FME-Transformer-Reference-Guide.pdf Syed Muhammad Fahad Akhtar, 2017, Big Data Architect's Handbook, Packt Publishing Ltd.
	eResources addresses	Podstawowe https://knowledge.safe.com/page/documentation/index.html - Documentation of FME from the SAFE Software Adresy na platformie eNauczanie:
Example issues/ example questions/ tasks being completed	1. Data pivoting from the long to wide format and vice versa; 2. Transforming data structure and content. 3. Changing file formats and geoprocessing using FME (Feature Manipulation Engine).	
Work placement	Not applicable	

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