GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	R language in chemistry: data analysis and visualization, PG_00069260						
Field of study	Chemical Technology, Chemistry, Biotechnology, Engineering and Technologies of Energy Carriers, Corrosion , Green Technologies, InfoBioChem						
Date of commencement of studies	February 2026		Academic year of realisation of subject	2026/2027			
Education level	second-cycle studies		Subject group				
Mode of study	Full-time studies		Mode of delivery	at the university			
Year of study	1		Language of instruction	Polish			
Semester of study	2		ECTS credits	3.0			
Learning profile	general academic profile		Assessment form	assessment			
Conducting unit	Department of Physical Chemistry -> Faculty of Chemistry -> Faculties of Gdańsk University of Technology						
Name and surname of lecturer (lecturers)	Subject supervisor		dr hab. Agnieszka Gajewicz-Skrętna				
	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						
	eNauczanie source addresses: Moodle ID: 5995 Język R w chemii: analiza i wizualizacja danych https://enauczanie.pg.edu.pl/2025/course/view.php?id=5995						
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	25.0	75	
Subject objectives	The course aims to introduce students to the R programming language for processing, analysing and visualising data, including chemical data. The course focuses on equipping students with the practical skills necessary for working with data obtained from laboratory experiments, instrumental analyses and environmental studies. Students will learn to use R independently for data processing and statistical, chemometric and/or cheminformatics analysis, as well as for creating professional plots, including interactive ones.						

Learning outcomes

Course outcome	Subject outcome	Method of verification
[K7_W02] identifies analytical techniques appropriate for solving specific analytical tasks – also in the production plant	knows and identifies analytical techniques appropriate for solving specific research and practical tasks and understands how data obtained from laboratory experiments, instrumental analyses, and environmental studies can be processed and analysed statistically, chemometrically and/or cheminformatically using the R language.	[SW1] Assessment of factual knowledge [SW3] Assessment of knowledge contained in written work and projects
[K7_U04] develops and transmits technical information in the form of text documents, spreadsheets, graphs, technological diagrams and multimedia presentations, and prepares a speech including a multimedia presentation	is able to develop and communicate the results of chemical data analysis in the form of clear and professional plots, including interactive plots, and to prepare automated data analysis reports in PDF, Word, or HTML format using the R language and appropriate packages.	[SU1] Assessment of task fulfilment [SU2] Assessment of ability to analyse information [SU3] Assessment of ability to use knowledge gained from the subject [SU4] Assessment of ability to use methods and tools [SU5] Assessment of ability to present the results of task
[K7_W01] recognizes problems of modern chemistry, including properties and obtaining chemical compounds, necessary for making calculations, including the dependence of the compound's structure and its reactivity	knows and recognises that the R language can be used to process, analyse and visualise chemical data, including data related to the properties of chemical compounds, structure–reactivity relationships and the results of chemical calculations.	[SW1] Assessment of factual knowledge [SW3] Assessment of knowledge contained in written work and projects
[K7_U02] prepares detailed documentation of the results of independently conducted experiments and analyzes the obtained results, uses professional vocabulary with understanding and prepares and communicates information	is able to independently import, process and prepare chemical data for further analysis, perform basic and advanced statistical, chemometric and/or cheminformatics analysis in R, and prepare documentation of research results using appropriate specialist terminology.	[SU1] Assessment of task fulfilment [SU2] Assessment of ability to analyse information [SU3] Assessment of ability to use knowledge gained from the subject [SU4] Assessment of ability to use methods and tools [SU5] Assessment of ability to present the results of task
[K7_U101] is able to formulate complex research problems and adopts appropriate methods, obtaining innovative solutions, cooperating with other people, both as a leader and a team member	is able to formulate complex research problems related to the analysis of chemical data and select appropriate statistical, chemometric and/or cheminformatics methods available in the R environment in order to obtain useful and innovative solutions, while working both independently and collaboratively.	[SU2] Assessment of ability to analyse information [SU3] Assessment of ability to use knowledge gained from the subject [SU4] Assessment of ability to use methods and tools [SU5] Assessment of ability to present the results of task [SU1] Assessment of task fulfilment
[K7_K02] is able to cooperate and work in a group, taking on different roles	is ready to work effectively as an integral part of a team on tasks related to the processing, analysis, and visualization of chemical data using R, taking on various roles and assuming responsibility for assigned tasks.	[SK1] Assessment of group work skills [SK2] Assessment of progress of work [SK3] Assessment of ability to organize work [SK4] Assessment of communication skills, including language correctness [SK5] Assessment of ability to solve problems that arise in practice

<table>
<tr><td>Subject contents</td><td>Course content – laboratory

The following areas are covered during the laboratory classes:

- Introduction to the R language and the RStudio environment, including its history, capabilities and applications, as well as R syntax, data types, data structures and simple data operations;
- Introduction to working with data in R, including reading and saving data, data import and export (CSV, Excel and TXT), data cleaning, handling missing data, and filtering, sorting and merging datasets;
- Basic statistical analysis of research results, including mean, median and standard deviation, as well as correlations;
- Introduction to chemometrics and cheminformatics in R, including correlation analysis, similarity analysis and dimensionality reduction methods;
- The tidyverse package as the foundation of modern data analysis, including data transformation and manipulation using dplyr and tidyr;
- Data visualisation - theory and good practices, the principles of creating scientific plots (plot and ggplot2), types of plots (scatter plot, line plot, histogram, box plot, bar chart and pie chart), colour schemes and layered plot construction.Customisation, faceting and interactive data presentation using Shiny and Plotly.
- First scripts in R: writing simple functions and user scripts.
- Overview of selected applications of R in the analysis of real chemical and environmental data, such as the analysis of spectroscopic, toxicological and environmental monitoring data.
- Project: comprehensive analysis and visualisation of real chemical data, and independent completion of a full workflow involving data import, statistical analysis, visualisation and report writing.</td></tr>
<tr><td>Prerequisites and co-requisites</td><td></td></tr>
<tr><td>Assessment methods and criteria</td><td><table><tr><th>Subject passing criteria</th><th>Passing threshold</th><th>Percentage of the final grade</th></tr><tr><td>Laboratory classes: project-based assignment</td><td>50.0%</td><td>100.0%</td></tr></table></td></tr>
<tr><td>Recommended reading</td><td><table><tr><td>Basic literature</td><td>- Przemysław Biecek. Przewodnik po pakiecie R. Oficyna wydawnicza GIS. 2014. ISBN: 9788389020987.
- Gillespie Colin Lovelace Robin. Wydajne programowanie w R. Praktyczny przewodnik po lepszym programowaniu. Helion 2018, ISBN: 9788375413526.</td></tr><tr><td>Supplementary literature</td><td>- Marek Gągolewski. Programowanie w języku R: Analiza danych, obliczenia, symulacje. Wydawnictwo Naukowe PWN 2016. ISBN: 9788301189396
- Jared P. Lander. Język R dla każdego. Zaawansowane analizy i grafika statystyczna. Pearson Addison-Wesley 2018. ISBN: 978-83-7541-336-6</td></tr><tr><td>eResources addresses</td><td></td></tr></table></td></tr>
<tr><td>Example issues/ example questions/ tasks being completed</td><td>Examples of theoretical topics:

- Explain the difference between a vector and a list in R.
- What are the basic data structures in R? Provide examples.
- How can variables be created and accessed in R?
- Explain the difference between the read.csv() and read.table() functions.
- How does the mutate() function work in the dplyr package?
- Which dplyr functions are used for filtering and grouping data?
- How can dplyr functions be combined using the pipe operator (%>%)?
- What does the 'layered structure of a plot' mean in ggplot2?
- What is faceting, and what is it used for?
- In what ways does plot() differ from ggplot()?
- How can a simple function be created in R?
- What is a for loop, and how can it be replaced with apply() functions?
- How can custom scripts be stored and called up?
- How can a data analysis report be prepared in R using R Markdown?</td></tr>
<tr><td>Practical activites within the subject</td><td>Not applicable</td></tr>
</table>

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