GDAŃSK UNIVERSITY OF TECHNOLOGY

Subject card

Subject name and code	Statistics and data analysis, PG_00060847		
Field of study	Chemical Technology		
Date of commencement of studies	October 2025	Academic year of realisation of subject	2025/2026
Education level	first-cycle 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	2	ECTS credits	2.0
Learning profile	general academic profile	Assessment form	assessment
Conducting unit	Department of Physical Chemistry -> Faculty of Chemistry -> Wydziały Politechniki Gdańskiej		
Name and surname of lecturer (lecturers)	Subject supervisor	prof. dr hab. inż. Adam Kloskowski	
	Teachers		

Lesson types and methods of instruction

Lesson type	Lecture	Tutorial	Laboratory	Project	Seminar	SUM
Number of study hours	15.0	0.0	15.0	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	5.0	15.0	50

Subject objectives

The student will be able to explain the basic concepts of statistics. After completing the laboratories, the student is able to use Excel and Python (Orange Data Mining) software. for analysis of experimental data After completing the course, the student should:

1) be proficient in using advanced functions of Excel

2) know the basics of using Python

3) be able to create mathematical (regression) models to solve basic technological problems

Learning outcomes	Course outcome	Subject outcome	Method of verification
	[K6_W01] Possesses knowledge of mathematics and physics necessary to analyze and describe technological processes, including differential and integral calculus, numerical methods, statistics and elements of vector analysis.	The student has knowledge in the field mathematics necessary for proper statistical description data sets. Student has knowledge of physics necessary in data analysis process regarding issues technical and technological.	[SW1] Assessment of factual knowledge [SW2] Assessment of knowledge contained in presentation [SW3] Assessment of knowledge contained in written work and projects
	[K6_U01] Is able to independently plan the learning process and acquire, analyse and interpret information from various sources, also in English.	The student is able to prepare interesting way of presentation statistical data. Fluently uses tools for creating charts and presentations dependencies between variables. The student has the ability to analyze information in the context of impact decisions made on environment. He has consciousness responsibility for the actions taken decisions. Able to work in a group as well as individually and is aware of necessity keeping the assumptions deadlines	[SU1] Assessment of task fulfilment [SU3] Assessment of ability to use knowledge gained from the subject [SU5] Assessment of ability to present the results of task

Subject contents	During the classes, issues will be discussed that introduce the student to the basic concepts in the field statistics aimed at equipping the student with the tools necessary at later stages of education the scope of developing measurement results, planning and conducting experiments and metrology. IN During the lectures, students become familiar with the basics of statistical description in terms of concepts and tools computational (with an introduction to metrology). In the laboratory, they solve practical tasks related to the use of statistical tools and appropriately selected software (Excel, Python) in a general and technical context. Content discussed during the lecture and laboratory cover the following areas: Statistical description of the data set - statistics of a one-dimensional random variable - absolute error, relative error, precision, accuracy - determining the uncertainty of measuring instruments, uncertainty of the calibration stage of experimental methods, uncertainty propagation method, uncertainty estimation methods (types A and B), Ishikawa diagram - correct recording of experimental measurement results with uncertainty and unit, introduction to the concept of measurement consistency, - position measures (with particular emphasis on such as arithmetic mean, geometric mean, mode, median, quantiles); measures of dispersion (with particular emphasis on measures such as deviation standard, coefficient of variation, range); asymmetry measures, - data representation methods: histogram, box diagram - normal, t-Student, chi-square, Poisson (small numbers), Boltzmann, uniform, triangular, skewness of distribution, distribution function, central limit theorem - determination of the confidence interva - concepts commonly used in laboratory practice: repeatability, reproducibility, accuracy, correctness, linearity, measurement range, sensitivity, calibration, noise and detection limit Statistical inference Verification of hypotheses - concepts: null hypothesis, alternative hypothesis, significance level, critical test area, tests parametric and non-parametric - statistical inference procedures - types of errors: systematic errors, random errors, gross errors, - estimating the probability of an event occurring - statistical tests: Q-Dixon, Grubbs, F-Snedecor, t-Student, Aspin-Welch, Cochran-Cox, - 3 sigma rule - p-value Data analysis - presenting two-dimensional data: scatter chart, line chart - presenting the confidence interval and uncertainty values on the chart

	- correlation, Pearson's and Spearman's correlation coefficient - linear regression, non-linear relationships - introduction to multidimensional data analysis, multiple regression - introduction to big data analysis: processing and cleaning of data sets, determining and predicting patterns and relationships in data sets Applications of statistical methods and tools - Applications in the analysis of experimental data (procedure calibration, QSAR methods, analysis clusters) - Validation of the measurement method - Elements of experimental optimization - Planning experiments, taking into account factorial and minimal designs - Statistical criteria for assessing the validity of results and comparing experimental methods
Prerequisites and co-requisites	Basic knowledge of mathematics

Assessment methods and criteria

Subject passing criteria	Passing threshold	Percentage of the final grade
Laboratory – problem task	50.0%	50.0%
Lecture - test	50.0%	50.0%

Recommended reading

Basic literature	Wojciech Hyk, Zbigniew Stojek, Analiza statystyczna w laboratorium badawczym, PWN, Warszawa 2019 - Andrzej Balicki, Wiesława Makać, Metody wnioskowania statystycznego, Wydawnistwo UG, Gdańsko 2006 - Felix Zumstein, Python i Excel. Nowoczesne środowisko do automatyzacji i analizy danych, Helion, Warszawa 2021] - James Miller, Jane Miller, Statystyka i chemometria w chemii analitycznej, PWN, Warszawa 2016 - YouTube: Orange Data Mining tutorials [dostępne online]
Supplementary literature	- P. Konieczka Ocena i kontrola jakości wyników pomiarów analitycznych, WNT, Warszawa 2007 - J. Mazerski Podstawy chemometrii, Wydawnictwo PG, Gdańśk 2004 - A. Navlani, A. Fandango, I. Idris, Python i praca z danymi. Przetwarzanie, analiza, modelowanie i wizualizacja, Helion, Warszawa 2022 - Joel Grus, Data science
eResources addresses	

Example issues/ example questions/ tasks being completed	Examples of theoretical issues: - How many digits should be presented in the measured result? What methods are used to estimate measurement uncertainty? - What is the R2 coefficient? Can it take negative values and if so, when? what's going on? - What is correlation? Is high correlation of variables in a regression model beneficial? - Uncertainty and error - discuss the meaning of these phrases. - Explain the concepts: precision, accuracy, sensitivity, specificity. - Discuss the tools available in Excel to determine standard deviation. Discuss the differences between them (e.g. giving formulas according to which each function calculates them) - Discuss selected methods for identifying distant observations. - What is linearization? Examples of calculation problems: - Using the multiple regression method (linear model), find the relationship between toxicity and the variable values given in the table for the given sets of ingredient examples cosmetics.Using linear regression, determine the pH value of a river water sample using calibration results of the electrochemical pH meter presented in the table - Assess whether the tested fermentation conditions have an impact on the composition (obtaining efficiency, etc.) of the wine using the selected statistical test. - Based on the data set, assess the accuracy and precision of the measurement technique
Work placement	Not applicable

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