



## Subject card

|   |  |  |                                    |                                     |  |            |     |
|---|--|--|------------------------------------|-------------------------------------|--|------------|-----|
| Subject name and code                       | Statistics and data analysis, PG_00060847  |  |                                    |                                     |  |            |     |
| Field of study                              | Chemical Technology  |  |                                    |                                     |  |            |     |
| Date of commencement of studies             | October 2026   | Academic year of realisation of subject                  |                                    |                                     | 2026/2027                                      |            |     |
| 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 -> Faculties of Gdańsk University of Technology   |  |                                    |                                     |  |            |     |
| Name and surname of lecturer (lecturers)    | Subject supervisor   |  | prof. dr hab. inż. Adam Kloskowski |                                     |  |            |     |
|   | Teachers   |  |                                    |                                     |  |            |     |
| Lesson types                                | 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                          | <p>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<br/>After completing the course, the student should:</p> <ol style="list-style-type: none"><li>1) be proficient in using advanced functions of Excel</li><li>2) know the basics of using Python</li><li>3) be able to create mathematical (regression) models to solve basic technological problems</li></ol> |  |                                    |                                     |  |            |     |

|                                 |   |  |   |
|---------------------------------|---|--|---|
| 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 technical and technological.  | [SW3] Assessment of knowledge contained in written work and projects<br>[SW2] Assessment of knowledge contained in presentation<br>[SW1] Assessment of factual knowledge  |
|                                 | [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   | [SU5] Assessment of ability to present the results of task<br>[SU3] Assessment of ability to use knowledge gained from the subject<br>[SU1] Assessment of task fulfilment |
| Subject contents                | <p>Course content – lecture</p> <p>During the classes, issues will be discussed that introduce the student to the basic concepts of statistics in order to equip the student with the tools necessary at later stages of education in the field of developing measurement results, planning and conducting experiments, and metrology. During the lectures, students learn the basics of statistical description in terms of concepts and computational tools (with an introduction to metrology). Within the laboratory, they solve practical tasks related to the use of statistical tools and appropriately selected software (Excel, Orange, Python) in a general and technical context. The content discussed during the lecture covers the following areas:</p> <p>Opis statystyczny zbioru danych</p> <ul style="list-style-type: none"> <li>- statistics of a one-dimensional random variable</li> <li>- absolute error, relative error, precision, accuracy</li> <li>- uncertainty propagation method, uncertainty estimation methods</li> <li>- data representation methods: histogram, normal distribution, Student's t, Poisson (small numbers),</li> <li>- determination of the confidence interval,</li> <li>- statistical inference, hypothesis verification - concepts: null hypothesis, alternative hypothesis, significance level, critical test area, parametric and non-parametric tests - statistical inference procedures</li> <li>- introduction to big data analysis: processing and cleaning data sets, determining and predicting patterns and relationships in data sets</li> </ul> <p>Course content – laboratory</p> <p>Applications of statistical methods and tools</p> <ul style="list-style-type: none"> <li>- data analysis: descriptive statistics, statistical tests (statistical inference)</li> <li>- examples of calculations regarding error transfer/propagation</li> <li>- Using the Orange program in data analysis:</li> <li>- introduction to the Orange program</li> <li>- regression/linearized regression</li> <li>- PCA analysis/regression models and neural networks/SHAP methodology</li> </ul> |  |   |
| Prerequisites and co-requisites | Basic knowledge of mathematics  |  |   |
| Assessment methods and criteria | Subject passing criteria  | Passing threshold  | Percentage of the final grade   |
|                                 | Lecture - test  | 50.0%  | 50.0%   |
|                                 | Laboratory – problem task   | 50.0%  | 50.0%   |
| Recommended reading             | Basic literature  | <p>Wojciech Hyk, Zbigniew Stojek, Analiza statystyczna w laboratorium badawczym, PWN, Warszawa 2019</p> <ul style="list-style-type: none"> <li>- Andrzej Balicki, Wiesława Makać, Metody wnioskowania statystycznego, Wydawnictwo UG, Gdańsko 2006</li> <li>- Felix Zumstein, Python i Excel. Nowoczesne środowisko do automatyzacji i analizy danych, Helion, Warszawa 2021]</li> <li>- James Miller, Jane Miller, Statystyka i chemometria w chemii analitycznej, PWN, Warszawa 2016</li> <li>- YouTube: Orange Data Mining tutorials [dostępne online]</li> </ul> |   |
|                                 | Supplementary literature  | <ul style="list-style-type: none"> <li>- P. Konieczka Ocena i kontrola jakości wyników pomiarów analitycznych, WNT, Warszawa 2007</li> <li>- J. Mazerski Podstawy chemometrii, Wydawnictwo PG, Gdańsk 2004</li> <li>- A. Navlani, A. Fandango, I. Idris, Python i praca z danymi. Przetwarzanie, analiza, modelowanie i wizualizacja, Helion, Warszawa 2022</li> <li>- Joel Grus, Data science</li> </ul>  |   |
|                                 | eResources addresses  |  |   |

|   |  |
|---|--|
| <p>Example issues/<br/>example questions/<br/>tasks being completed</p> | <p>Examples of theoretical issues:</p> <ul style="list-style-type: none"> <li>- How many digits should be presented in the measured result? What methods are used to estimate measurement uncertainty?</li> <li>- What is the R2 coefficient? Can it take negative values and if so, when? what's going on?</li> <li>- What is correlation? Is high correlation of variables in a regression model beneficial?</li> <li>- Uncertainty and error - discuss the meaning of these phrases.</li> <li>- Explain the concepts: precision, accuracy, sensitivity, specificity.</li> <li>- 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)</li> <li>- Discuss selected methods for identifying distant observations.</li> <li>- What is linearization?</li> </ul> <p>Examples of calculation problems:</p> <ul style="list-style-type: none"> <li>- 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</li> <li>- Assess whether the tested fermentation conditions have an impact on the composition (obtaining efficiency, etc.) of the wine using the selected statistical test.</li> <li>- Based on the data set, assess the accuracy and precision of the measurement technique</li> </ul> |
| <p>Practical activities within the subject</p>                          | <p>Not applicable</p>  |

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