Semmelweis University, Faculty of Medicine

Pharmaceutical Innovation and Business Administration Master of Science

Name of the host institution (and any contributing institution):

Department of Pharmacology and Pharmacotherapy in collaboration with the Institute of Biostatistics and Network Science of Semmelweis University

Name of subject:

Data management and statistics

in English: Data management and statistics

in German: Not applicable

Credit value: 5

Semester: 2025/2026 1st Semester

in which the subject is taught according to the curriculum

Hours per semester	Lecture	Course work	Consultation
150	15	130	5

Hours per week	Lecture	Course work	Consultation
Course blocks tailored to the students' employment obligations on Fridays and Saturdays Course dates: 4th October 9.00-12.00 17th October 9.00-12.00 24th October 14.00-18.00 7th November 14.00-18.00 13rd December 9.00-12.00			

Type of course:

compulsory

Academic year:

2025/2026

Language of instruction (for optional and elective subjects):

English

Course code:

(in the case of a new course, to be completed by the Dean's Office, following approval

Course coordinator name: Dr. Roland Molontay

Course coordinator location of work, telephone availability: Semmelweis University, 1082 Budapest, Baross u. 22 tel.: +36205913228

Course coordinator position: Director

Course coordinator Date and number of habilitation:

Objective of instruction and its place in the curriculum:

The aim of the course is to equip students with the skills to recognize business and operational challenges within the pharmaceutical and healthcare sectors where data analysis and data science can provide strategic value. Through practical prototyping, students will learn to demonstrate and communicate the competitive advantages of data-driven solutions. The course covers both theoretical and practical foundations of data analysis methods relevant to economic and managerial decision-making in the life sciences industry. Students will acquire quantitative tools for analyzing and predicting industry-relevant phenomena. Beyond delivering essential theoretical knowledge, the course emphasizes practical problem-solving, real-world case studies, and the cultivation of a data-oriented mindset tailored to the pharmaceutical business environment.

Method of instruction (lecture, group work, practical lesson, etc.): real-time online lectures

---

Competencies acquired through completion of course:

1. Understands the key tasks of business data analysis, the main areas of expertise, and the tools applicable in each.
2. Understands the technical details of the main steps in data analysis: data collection, data preparation, modeling, evaluation, and application.
3. Has knowledge of the most important theoretical models and algorithms in data science, including the basic paradigms of supervised and unsupervised machine learning.
4. Knows the fundamental tools and methods of data visualization.
5. Is familiar with the basic operation of data-driven decision support tools.
6. Understands the most important micro- and macroeconomic applications of data science, data analysis, and data visualization, particularly in the field of business intelligence.
7. Is aware of the learning, knowledge acquisition, and data collection methods used in data analysis, as well as their ethical limitations and problem-solving techniques.
8. Can identify business problems to which data science or machine learning solutions can be applied.
9. Can prototype possible solutions, visualize results, and identify business value to inform decision-making and guide further analysis.
10. Can apply learned theories and methods to explore, systematize, and analyze facts and relationships; formulate independent conclusions and critical observations; propose and evaluate decisions in both routine and partially unknown domestic and international contexts.
11. Are able to determine the complex consequences of economic processes and organizational events.

12. Can apply data analysis problem solving techniques, problem solving methods, their application conditions and limitations.
13.
14. Collaborate effectively with instructors and peers to expand collective knowledge.
15. Continuously develops expertise through independent and ongoing learning.
16. Demonstrates openness to and proficiency in using information technology tools.
17. Shows problem sensitivity, proactive behavior, and constructive cooperation in projects and group tasks to ensure high-quality outcomes.
18. Strives for accuracy and error-free problem solving.
19. Works independently with responsibility, including the selection of appropriate methodologies and techniques, as well as the organization, planning, and management of tasks.
20. Collects, systematizes, analyzes, and evaluates data effectively while fostering both general and professional growth.
21. Applies a systems-oriented approach to thinking and problem solving.
22. Takes full responsibility for analyses, conclusions, and decisions.

Course outcome (names and codes of related subjects):

none

Prerequisites for course registration and completion: (CODE): -

none

In the case of multi-semester courses, position on the possibility of and conditions for concurrent registration:

none

The number of students required to start the course (minimum, maximum), student selection method:

all students admitted

Detailed course syllabus (if the course can be divided into modules, please indicate):

(Theoretical and practical instruction must be broken down into hours (weeks), numbered separately; names of instructors and lecturers must be listed, indicating guest lecturers/instructors. It cannot be attached separately! For guest lecturers, attachment of CV is required in all cases!)

Each lecture - 45 minutes

Lecture 1:

- Introduction to Data Science – History, key concepts, objectives
- Overview of job roles, tools, and fields of application (with examples from pharmaceutical and business contexts)

Lecture 2:

- Data Discovery – Identifying and sourcing relevant datasets

- Examples from pharmaceutical innovation, healthcare, and market analysis

Lecture 3:

- Data Preparation – Cleaning, transforming, and organizing data
- Practical exercise with a sample dataset

Lecture 4:

- Data Visualization I – Principles of effective visual communication
- Introduction to visualization tools (e.g., Tableau, Power BI, Python libraries)

Lecture 5:

- Data Visualization I (continued) – Basic chart types, best practices, and common pitfalls
- Hands-on: Creating clear and informative visuals

Lecture 6:

- Supervised Machine Learning I – Concept of supervised learning
- k-Nearest Neighbors (kNN): Theory and example applications

Lecture 7:

- Supervised Machine Learning I (continued) – Decision Trees: concept, strengths, and limitations
- Short practical demo with sample data

Lecture 8:

- Supervised Machine Learning II – Introduction to ensemble methods (bagging, boosting, random forests)

Lecture 9:

- Supervised Machine Learning II (continued) – Neural Networks: basic architecture and applications in pharma/business

Lecture 10:

- Model Evaluation and Validation – Performance metrics (accuracy, precision, recall, F1-score)

Lecture 11:

- Model Evaluation and Validation (continued) – Cross-validation, train/test splits, avoiding overfitting
- Short lab exercise to compare models

Lecture 12:

- Unsupervised Machine Learning – Concept and applications
- k-Means clustering: theory and process

Lecture 13:

- Unsupervised Machine Learning (continued) – Practical example with clustering in a business or pharma dataset

Lecture 14:

- Integrated Case Study – From data discovery to visualization and modelling
- Applying both supervised and unsupervised methods in a real-world scenario

Lecture 15:

- Review and Discussion – Key takeaways, open Q&A, industry applications
- Guidance for further study and project work

Other courses with overlapping topics (obligatory, optional, or elective courses) in interdisciplinary areas. To minimalize overlaps, topics should be coordinated. Code(s) of courses (to be provided): -

none

Requirements for attendance, options for making up missed sessions, and method of absence justification:

Full attendance is required. Completing additional e-learning materials are required to make up missed courses.

Assessment methods during semester (number, topics, and dates of midterms and reports, method of inclusion in the course grade, opportunities for make-up and improvement of marks):

test and project work submitted at the end of the semester

Number and type of individual assignments to be completed, submission deadlines: -

---

Requirements for the successful completion of the course: successful completion of the test at the end of the semester (>50%)

Type of assessment: test

Examination requirements (list of examination topics, subject areas of tests, lists of mandatory parameters, figures, concepts and calculations, practical skills, optional topics for the project assignment recognized as an exam and the criteria for its completion and evaluation)

same as course syllabus

Method and type of grading (Share of theoretical and practical examinations in the overall evaluation. Inclusion of the results in the end-of-term assessment. Possibilities of and conditions for offered grades.):

Written end-term test: during the semester, the course material will be tested with a written end-term test. The test consists of theoretical questions and calculations. At least 50% of the points of the mid-term test must be obtained in order to obtain the signature and pass the course.

test: 70% theoretical, 30% practical

Grading Scale

- Excellent: 90–100
- Good: 80–89
- Satisfactory: 65–79
- Pass: 50–64
- Fail: <50

Signature of habilitated instructor (course coordinator) announcing the course:

Prof. Dr. Péter Ferdinandy	Dr. Roland Molontay
Head of Department	Director

Signature of the director of the host institution:

Prof. Dr. Péter Ferdinandy
Head of Department

Date of submission:

11th August 2025