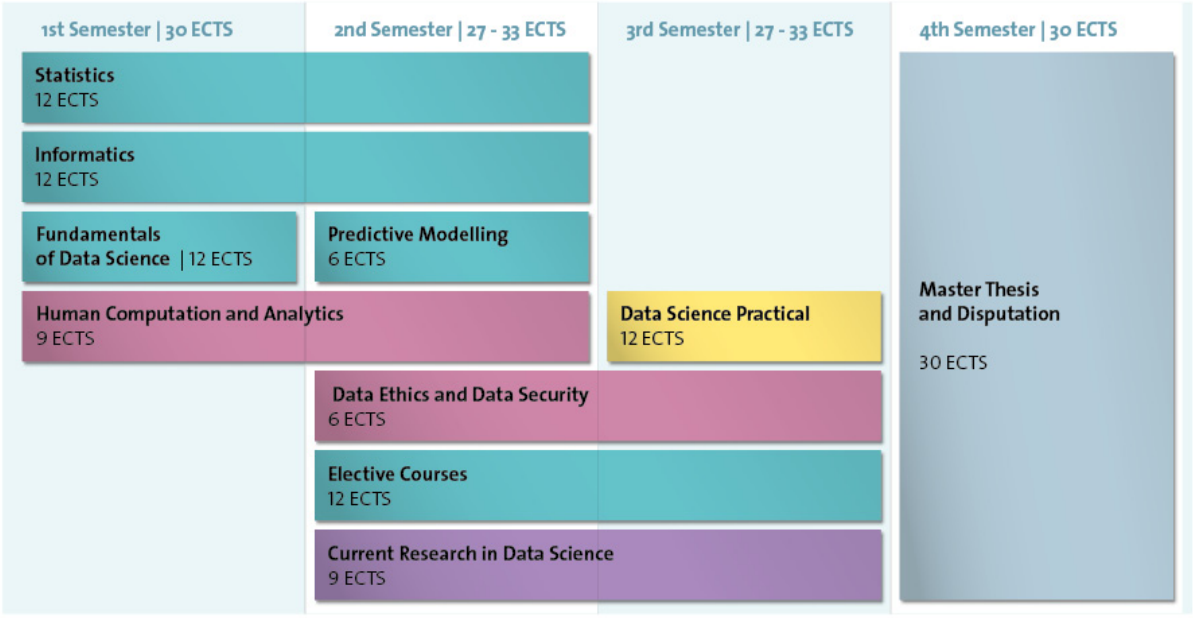
MSc Data Science@LMU: Overview of Modules



Statistics: Inference and Sampling (core module)

Semester 1–2, 12 ECTS

Prof. Kauermann

Type	2 lectures, each with a tutorial: 1. Statistical Reasoning and Inference (1) 2. Statistical Reasoning and Inference (2): Sampling and Experimental Design
Duration	2 courses, each 4 SWS: 3 SWS lecture and 1 SWS tutorial.
Content	The core module *Statistics* covers fundamental statistical concepts and methods and consists of two courses. The first course, *Statistical Reasoning and Inference (1)*, comprises (i) traditional and modern methods of statistical inference (maximum likelihood, composite likelihood, multiple testing, false discovery rate, etc.) and (ii) Bayesian approaches including computer intensive Markov-Chain-Monte-Carlo (MCMC) methods. The lecture is accompanied by a tutorial, in which the content of the lecture will be consolidated and numerical tools such as R will be applied. In the course *Statistical Reasoning and Inference (2): Sampling and Experimental Design*, fundamental ideas of sampling, bootstrapping, missing data, multivariate models and experimental design are introduced A second focus of the lecture is on the analysis of "observational data" and related problems of potential biases. Finally, fundamental concepts and ideas of experimental design will be introduced.
Objective	Students learn fundamental concepts of statistical inference. They are able to use these concepts with new data and draw samples from complex data, and they know how to design experiments in order to draw conclusions from the data.
Assessment	Written or oral exam at the end of the second semester.

Informatics: Knowledge Discovery and Big Data Management (core module)

Semester 1–2, 12 ECTS

Prof. Schubert

Type	2 courses, each consisting of a lecture and a tutorial: 1. Knowledge Discovery and Data Mining 2. Big Data Management
Duration	2 courses, each 5 SWS: 3 SWS lecture and 2 SWS tutorial.
Content	The core module *Informatics* covers all tasks within each step of the knowledge discovery process and consists of two courses. The first course, *Knowledge Discovery and Data Mining*, introduces feature representations and similarity measures as core concepts of data analysis. Based on these concepts, the course covers various methods from the area of data mining and pattern extraction (e.g. lazy learning, density-based clustering, k-medoid clustering, local outlier factor, a-priori algorithm, FP-growth, frequent subsequence mining). The second course, *Big Data Management*, focuses on the implementation of analysis methods and information systems for large, complex, and volatile data sets. First, modern data processing frameworks are presented that are used for managing, processing, and distributing data in data science applications. These systems include batch processing (e.g. Hadoop, Spark), streaming systems (e.g. Storm, Flink), and NoSQL database systems (e.g. MongoDB, Cassandra). In addition, the implementation of established data mining methods in parallel, distributed, and streaming systems is introduced.
Objective	Students get an overview of the data analysis process as a whole as well as the individual tasks within this process and the different methods available to handle these tasks. They learn central techniques of data mining and pattern extraction and will be able to manage analysis processes by using suitable data management systems. This includes not only being able to use specific methods in a given system, but also being able to choose suitable methods and systems.
Assessment	Written or oral exam at the end of the second semester.

Fundamentals of Data Science (Individual Modules)

Semester 1, 12 ECTS

Prof. Kauermann, Prof. Seidl

Type	Individually chosen courses, typically as inverted classroom (two courses).
Duration	4 SWS
Content	Each student is assigned to two courses from a variety of courses in advanced methods of statistics and computer science. These comprise lectures on statistical modelling, multivariate data analysis, advanced programming, and multivariate data analysis, among others.
Objective	At the end of the module students will be on a homogeneous level of expertise in advanced methods in both statistics and computer science.
Assessment	Written or oral exam or written assignment.

Human Computation and Analytics

Semester 1–2, 9 ECTS

Prof. Butz

Type	Lecture "Human Computation and Visual Analytics" with a tutorial, and a practical project in Human Computation and Visual Analytics.
Duration	2 SWS lecture, 2 SWS tutorial und 2 SWS practical.
Content	The module *Human Computation and Analytics* covers those aspects of data science, in which humans either produce data, and process and analyse it with the help of algorithms, or in which data are presented to humans by a computer system. In the area of Human-Computer Interaction (HCI), the basics of human perception and cognition are introduced as well as some approaches for the design of usable systems. The lecture part on Visual Analytics (VA) covers the visual analysis of data by the human user as well as some visualization techniques. The lecture part on Human Computation (HC) gives an introduction to distributed data collection by humans (crowdsourcing), and the processing of data by humans, for example in the form of online games (HC). The course includes lab meetings, in which students develop their own concepts based on what they have learned in the lectures. In the practical, students will implement their own concepts for HC/VA systems in the form of a working prototype.
Objective	Students should learn to understand that the human user is an integral part of the chain of collecting data, processing data and evaluating data. At the same time students should develop a consciousness for the effects of data science on the individual person as well as on society as a whole.
Assessment	Oral exam or written assignment at the end of the second semester.

Predictive Modelling

Semester 2, 6 ECTS

Prof. Bischl

Type	Lecture with a tutorial.
Duration	4 SWS: 3 SWS lecture, 1 SWS tutorial.
Content	Predictive Modelling, in particular by means of non-linear, non-parametric methods, has become a central part of modern data analysis both in computer science and statistics in order to uncover complex patterns and relationships in data. The module covers models such as decision trees, support vector machines, and ensembles (random forest, bagging, boosting) and concludes with advanced techniques regarding model selection, feature selection, and hyperparameter optimization.
Objective	Students acquire theoretical as well as practical competences regarding the most important models of learning from data. The students should be able to conduct a data analysis project themselves, including understanding and interpreting the data, in order to critically judge advantages and disadvantages of the different methods. The accompanying tutorials are a mix of theoretical and practical assignments. The latter will be conducted in R and will cover all methods introduced during the lecture.
Assessment	Written exam.

Data Ethics and Data Security

Semester 2–3, 6 ECTS

Prof. Kranzlmüller

Type	Lecture "Data Security and Data Anonymisation" with a tutorial, and a lecture series "Data Ethics".
Duration	1 SWS lecture, 1 SWS tutorial; 2 SWS lecture.
Content	The module *Data Ethics and Data Security* covers basic legal and ethical questions and challenges of data security. The module comprises two courses. The first course looks into methodological questions of data anonymisation and technical aspects of data security. The second course is a lecture series with (invited) talks by different speakers on ethical and legal aspects of data security. Students are introduced to the technical, legal, and ethical issues of data security, especially when dealing with personal data or when planning experiments in data science.
Objective	Students will reflect on standard procedures and problems of data protection and learn technical methods to handle data responsibly.
Assessment	Oral or written exam at the end of the third semester.

Elective Modules

Semester 2-3, 12 ECTS

Prof. Kauermann, Prof. Butz

Type	Lectures, tutorials, and practical courses from the regularly offered master courses in statistics, informatics, and computer linguistics as well as lectures offered by the partner universities TUM and Universität Augsburg.
Duration	7–10 SWS, depending on the combination of lectures (and tutorials).
Content	In the elective modules, students may choose courses in specialized fields from the regularly offered master courses in statistics, informatics, and computer linguistics. In addition, students may also attend master level courses at the partner universities. These include courses on image processing and mathematical statistics at TUM and computational finance at Universität Augsburg.
Objective	Students acquire theoretical and practical knowledge and skills in specialized areas related to data science.
Assessment	Exams of the chosen courses as stated in the relevant programme regulations.

Current Research in Data Science

Semester 2–3, 9 ECTS

Prof. Schmid

Type	Seminar, workshops and summer schools
Duration	3 SWS: 2 SWS seminar, 1 SWS workshops etc.
Content	In the seminar, publications of current research in data science will be discussed. This module also comprises the summer schools, the focused tutorials, and the DataFest, among others.
Objective	New methodologies will be presented and discussed. Students will deepen their skills of working with scientific publications and will learn to present newly acquired scientific knowledge. The summer schools and other events will provide an insight into the different fields of data science.
Assessment	Presentation and written assignment in the third semester.

Data Science Practical

Semester 3, 12 ECTS

Prof. Kauermann, Prof. Schubert

Type	Practical project
Duration	ca. 3 months
Content	The module *Data Science Practical* plays a central role in the curriculum of the master program. Practical experience with data-analytic methods, which are taught in the core and elective modules, is essential in order to generate knowledge from data. Students will work on practical problems in the field of data science. The problems are typically concrete projects provided by non-university partners. The focus of the course is not only on tackling methodological challenges in the analysis of massive data, but on communicating the results and findings to the client.
Objective	Students learn to work in teams with large datasets.
Assessment	Presentation of results and final report.

Master Thesis and Disputation

Semester 4, 30 ECTS

Prof. Kauermann, Prof. Seidl

Type	Master thesis
Duration	1 semester
Content	The master thesis concludes the study program. The thesis may be either research-orientated or stimulated by a practical problem, e.g. as an extension of the Data Science Practical. The thesis will be defended in a disputation.
Objective	Consolidation of acquired knowledge; independent scientific study.
Assessment	Master thesis and disputation.