Data Science with Python

Winter

(in German: Data Science with Python)

Module-ID: FIN-INF-120513

Abbreviation	Credit Points	Semester	Term	Duration	Language	Level
DSWP	6	Winter	2.	1 Semester	english	Master

Link:	LSF
Limited accessability:	false
Responsibility:	Dr. Christian Beyer
Lecturer:	Dr. Christian Beyer
Classes:	Data Science with Python
Applicability in curriculum:	- M.Sc. INF: Informatik - M.Sc. INF: Schlüssel- und Methodenkompetenzen - M.Sc. INGINF: Informatik - M.Sc. INGINF: Schlüssel- und Methodenkompetenzen - M.Sc. WIF: Informatik - M.Sc. WIF: Schlüssel- und Methodenkompetenzen - M.Sc. DKE: Applied Data Science - M.Sc. VC: Computer Science - M.Sc. VC: Schlüssel- und Methodenkompetenzen - M.Sc. DE: Methoden der Informatik - M.Sc. DE: Interdisziplinäres Team-Projekt

Intended learning outcomes:
The course is about learning from data to perform predictions and obtain useful insights. In the seminar, we will use the programming language Python.
Necessary skills to manage and analyze data will be taught and practiced on real-world applications. Programming knowledge from other courses is helpful but not mandatory. However, students are expected to have a profound knowledge of fundamental data-analysis techniques, such as classification, regression and clustering. After successful completion of this course, the

student will be able to perform the following tasks in Python:

- Import and preprocess raw data (files, databases, web APIs)
- Transform data for modelling
- Perform exploratory data analysis with summary statistics and visualization
- Understand, build and evaluate predictive classification and regression models, including tree-based models, ensembles and boosted models
- Communicate and disseminate results and findings through reproducible documents, presentations, websites and interactive web applications

Content:
Part Fundamentals & Visualization: Basics, scripts, workflows, vectors & functions in Python Explorative data visualization Data transformation Part Data Management & Exploratory Data Analysis: Data cleaning & scraping Generating hypotheses and an intuition about the data with exploratory data analysis Data import Data management Relational data Strings, categorical data, dates & time Iteration: imperative & functional programming Part Modeling: Linear regression Classification Evaluation Model selection & regularization (LASSO, Ridge) Feature selection & model interpretation Decision trees Ensembles: random forests Boosting: gradient boosted trees Unsupervised learning, e.g. k-means, hierarchical clustering, self-organizing maps, principal component analysis Topic modeling with simple graphical models Statistical testing Part Communication: Communication and dissemination of results through visualization and interpretable summaries with documents, notebooks, presentations & websites

Workload:
Attendance time = 28 h: - 2 SWS weekly seminar; Independent work outside the actual Seminar time = 152 h: - 76 h preparation and follow-up of the seminar topics - 76 h solving the tasks, incl. work in the laboratory 180h = 28h attendance time + 152h independent work

Pre-examination requirements:	**Type of examination:**	**Teaching method / lecture hours per week (SWS):**
	Referat	Seminar (2 SWS)

Prerequisites according to examination regulations:	**Recommended prerequisites:**
keine	Area 1: Data Mining, Machine Learning, Artificial Intelligence Area 2: Databases Area 3: Programming Languages and Software Engineering Area 4: Stochastics, Applied Statistics

Media:	**Literature:**
	Will provided during the seminar

Comments:

Revision #1
Created 2026-10-20 16:59:07 UTC by Fakultät für Informatik
Updated 2026-10-20 16:59:07 UTC by Fakultät für Informatik