Title of course	**Data Science**
Responsible instructor	Prof Diego d'Andria, PhD
Learning objectives	▪ Learn about how data are stored and represented for the purposes of data analysis ▪ Learn how to manipulate data in tabular form ▪ Learn how to make data storage and access more efficient via database optimizations and indexing ▪ Be introduced to the SQL coding language and learn how to perform queries to retrieve selected data from a relational database ▪ Learn how to install and use the open-source statistical software *R* ▪ Learn about computational methods to solve for multiple regression models: maximum-likelihood, "robust" regression models ▪ Understand the concept of "bootstrapping" and learn how to apply it for a variety of purposes ▪ Learn how to represent and analyse network data ▪ Learn about specific challenges and pitfalls met when analysing microdata representing households, firms or Web users ▪ Learn commonly used methods for matching data and for cluster analysis ▪ Learn about different strategies to visualise data by means of graphs ▪ Learn about the basic principles and methods behind computational techniques to find the minimum/maximum of a function or data set ▪ Be introduced to computational models used to perform economic analysis. In particular: general-equilibrium models (CGE, DSGE), microsimulation models, agent-based models (ABM).

Course contents	1. An introduction to data science
	2. Data types and data structures
	  a. Data types and storage
	  b. Data in tabular form
	  c. Database optimization
	3. Principles of statistics
	  a. Distributions and their moments
	  b. Sampling
	4. An introduction to *R*
	5. Structured Query Language (SQL)
	6. Data visualisation
	  a. Two-dimensional (2D) graphs
	  b. Plotting multidimensional data
	  c. Using graphs to diagnose regression models
	7. Multiple regression analysis
	  a. Computational methods to solve regression models
	  b. "Robust" regression methods
	8. Bootstrapping
	9. Network analysis
	10. Using microdata
	11. Cluster analysis
	12. Optimization algorithms
	13. Matching algorithms
	14. Simulation models for economic analysis

Teaching methods	▪ Lectures ▪ Exercises

	▪ In-class coding ▪ Discussion ▪ Self-study
Prerequisites	There are no formal requirements.
Suggested reading	Handouts and references will be given during the classes.
Applicability	This course is applicable to all economics- and business-oriented Bachelor programmes offered by Schmalkalden University of Applied Sciences.
Workload	Total workload: 240 hours, of them: ▪ Lecture: 60 ▪ Self-study: 180, of them: ▪ Course preparation (in particular reading): 45 ▪ Follow-up: 45 ▪ Readings and exam preparation (including mid-term): 90
ECTS credit points and weighting factor	5 ECTS credit points; weighting factor: 5/180 for International Business and Economics 5/210 for Economics or Business Administration
Basis of student evaluation	▪ Comprehensive written examination, 60 minutes (80%) ▪ Mid-term exam, 60 minutes (20%)
Time	First academic year
Frequency	Each academic year
Duration	One semester
Course type	Elective course
Remarks	Teaching language is English.