

Big Data Analytics Overview

Prof. Dr. Philipp Baumann

University of Bern

Fall 2023

(tentative, as of July 7, 2023)

Outline

- 1 Lecturers
- 2 Course
- 3 Exam
- 4 Content
- 5 Literature

Outline

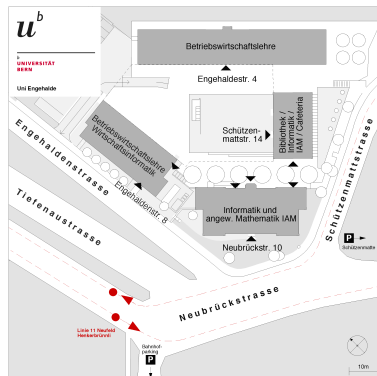
- 1** Lecturers
 - Curriculum Vitae
 - Contact
- 2** Course
- 3** Exam
- 4** Content
- 5** Literature

CV Philipp Baumann

- Born in Bern
- 2009: MSc in Business Administration, University of Bern
- 2013: PhD in Business Administration, University of Bern
- 2013–14: Research Scholar at IEOR Department, UC Berkeley
- 2014–15: Postdoc at IEOR Department, UC Berkeley
- 2015–18: Assistant Professor in Quantitative Methods/Operations Research, University of Bern
- Since 2018: Associate Professor in Quantitative Methods/Operations Research, University of Bern
- Research:
 - Machine learning/Data mining
 - Mathematical programming in finance and operations

Contact

- Department of Business Administration
- Chair in Quantitative Methods
 - Engehaldenstr. 4,
3012 Bern, office 209 (PB)
- E-Mail: philipp.baumann@unibe.ch
- Website: <http://www.pqm.unibe.ch>



Outline

- 1 Lecturers
- 2 Course
 - General information
 - Components of course
- 3 Exam
- 4 Content
- 5 Literature

Course in context of master studies

- 2015 course scheme: course can be selected as elective course
- 2007 course scheme: course is
 - part of the choice area for MSc BA and MSc B&Ec students
 - not part of any focus area
- No registration required for participation
- Required knowledge: completion of Bachelor degree in Business Administration or Economics
 - Introduction to Mathematics
 - Introduction to Statistics
 - Quantitative Methods in Business Administration I

Lecture

- Time: Monday, 8:15am to 10:00am
- Location: main building, room 120
- Start: Sep 18, 2023
- Lecture material: ILIAS
- Lecture includes
 - Explanations
 - Examples
 - Discussion of case studies
- Podcast will be provided

Exercises

- Time: Tuesday, 8:15am to 10:00am
- Location: main building, room 120
- Download of exercises: ILIAS
- Types of exercises
 - Review
 - Modelling
 - Manual application of discussed methods
 - Application of discussed methods using Software Python
- Podcast will be provided

Project

- Available on ILIAS on Tuesday Nov 7
- Project is conducted in groups of two students
- Access to Software Python required
- Project tutorial: Tuesday Nov 14, 2023 (instead of exercises)
- Deadline for submission of solution: Tuesday Nov 21, 2023
- Grading: 8 extra points can be achieved
- Discussion of solutions on Tuesday Nov 28

Outline

- 1 Lecturers
- 2 Course
- 3 Exam**
 - Overview
 - Dates
 - Exam additional information
- 4 Content
- 5 Literature

Overview exam

- 6 ECTS
- Written exam
- Exam will cover lecture, exercises, and project
- Permitted aids:
 - Non-programmable hand-held calculator
 - Formulary (will be distributed before the exam)
- Grading:
 - Maximal attainable score in exam: 90 points
 - Maximal attainable score in project: 8 points
 - At most 90 points required for highest grade
 - Project points can be credited to exams in Dec 2023 and Feb 2024

Dates for final exam

- Dates (duration of final exam is 90 minutes)
 - 1 Monday Dec 18, 2023 from 8:15am to 9:45am
 - 2 Monday Feb 12, 2024 from 8:15am to 9:45am
- Location: will be announced on ILIAS after closing date for exam deregistration
- Q&A on Tuesday Dec 12, 2023

Final exam: additional information

Exam on Dec 18, 2023

Closing date registration	Dec 11, 2023
Closing date deregistration	Dec 16, 2023
First date to access graded exam	Jan 11, 2024 (9am–10am)
Second date to access graded exam	Jan 23, 2024 (4pm–5pm)

Exam on Feb 12, 2024

Closing date registration	Feb 5, 2024
Closing date deregistration	Feb 10, 2024
First date to access graded exam	Feb 22, 2024 (9am–10am)
Second date to access graded exam	Feb 26, 2024 (3pm–4pm)

Registration and deregistration only via KSL

Outline

- 1 Lecturers
- 2 Course
- 3 Exam
- 4 Content**
- 5 Literature

Learning outcomes

The students are able to

- understand the benefits and challenges involved in big data-driven decision making
- preprocess and visualize data using the software Python
- apply methods for clustering and dimensionality reduction
- select and fine tune methods for regression and classification problems in Finance, Marketing, and Operations
- build highly-scalable learning systems using stochastic gradient descent and map reduce
- use big data to build product recommender systems, discover communities in social networks, and analyze text




Structure

- 1 Big data overview and introduction to Python
- 2 Descriptive analytics
 - Clustering
 - Dimensionality reduction
- 3 Predictive analytics
 - Regression
 - Classification
- 4 Analytics at scale
 - Map reduce and online learning
 - Stochastic gradient descent
- 5 Applications
 - Recommender systems
 - Social network analytics
 - Text analytics

Overview

- 1 Lecturers
- 2 Course
- 3 Exam
- 4 Content
- 5 Literature**

Literature (Textbooks)

-  McKinney (2022): Python for data analysis, O'Reilly (available online)
-  Geron (2019): Hands-on machine learning with Scikit-Learn and TensorFlow, O'Reilly
-  James, Witten, Hastie, Tibshirani (2021): An introduction to statistical learning, Springer (available online)