

Seminar in Business Analytics: Data Retrieval with R and Python Summer Term 2020

Course Description:

Prior to the start of the Information Age in the late 20th century, companies were forced to collect data from non-automated sources manually. Companies back then lacked the computing capabilities necessary for data to be analyzed, and as a result, decisions primarily originated not from knowledge but from intuition. With the emergence of ubiquitous computing technology, company decisions nowadays rely strongly on computer-aided “**Data Mining**”.

In this seminar, students will acquire two kinds of skills. First, students will learn to build a web-scraper to collect their own dataset from the web. Second, students will review different strategies for data analysis, and data visualization. Students are asked to describe and visualize the content of their dataset. Optionally, committed students can pick a statistical method / data mining algorithm of their choice and perform a descriptive or predictive data mining task on their dataset.

Basic programming skills are recommended (preferably, you will work in R and/or Python). For students who are less familiar to programming, it is also possible to work with a precollected dataset and focus on the data mining part. Please indicate in your application if you would prefer this setup. **Please also indicate your level of programming expertise.**

Among others, we will cover material from the following books:

- Wickham, Hadley, and Garrett Golemund. R for data science: import, tidy, transform, visualize, and model data. O'Reilly Media, Inc., 2016.
- Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. The elements of statistical learning. Vol. 1. New York: Springer series in statistics, 2001.

Target Group:

This Seminar specifically addresses students all IMP disciplines, as well as in the M.Sc. Economics and M.Sc. VWL programs.

Organization:

Registration: Until May 3, 2020

Application via email to gunther.gust@is.uni-freiburg.de. **Make sure the following information is contained in the email body:**

- First name, last name
- Matriculation number (Matrikelnummer)
- Current overall grade average
- Attended lectures and seminars at our chair and grade obtained
- Email, phone number
- Study program, semester
- Short description of experience level in Python or R

In addition, you need to send your **transcript of records in the attachment**.

Response whether application was successful will be sent out shortly after the registration deadline

First meeting: **May 12, 2020 at 17ct (Online Meeting)**

Paper due: **Jul 3, 2020**

Final presentation: **Jul 10, 2020, 16 ct.**

Revised paper due: **Aug 31, 2020**

Communication:

All announcements, handouts, etc. will be sent via email.

Topics:

Exact topics along with hints on literature will be announced at a later point.

Policies and Procedures

Grading:

Paper (about 15 pages, 50%) and revised final paper (50%). In addition, you have to hand in your programming code and datasets (if applicable). In addition, there will be

a presentation session (online), where the participants get feedback on their preliminary work. The seminar paper can be written in English only.

Credit points: 6

Credit points are applicable to:

- M.Sc. BWL PNP: Allgemeine BWL, Wirtschaftsinformatik
- M.Sc. VWL (2011): BWL, Wirtschaftsinformatik
- M.Sc. VWL (2014): Business Analytics
- M.Sc. Economics: Elective in Information Systems and Network Economics profile
- M.Sc. Computer Science: Wahlmodule BWL und VWL

Chair: Prof. Dr. Dirk Neumann
Albert-Ludwigs-Universität
Chair of Information Systems
Platz der Alten Synagoge
79085 Freiburg
Germany