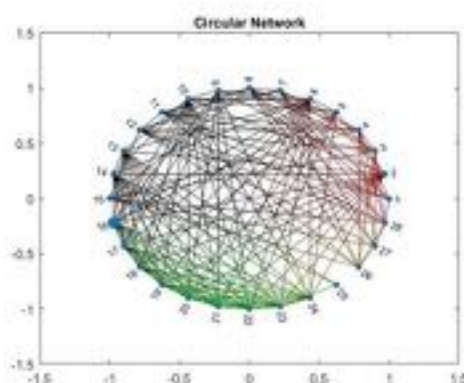
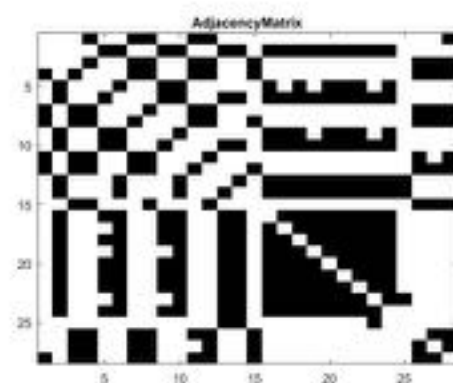
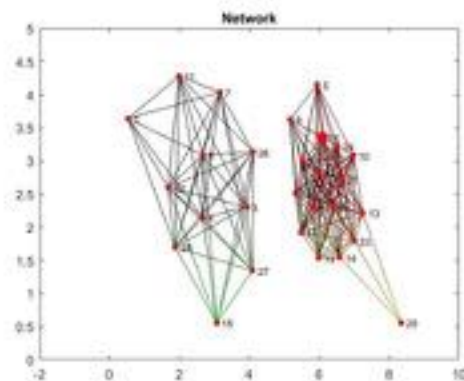
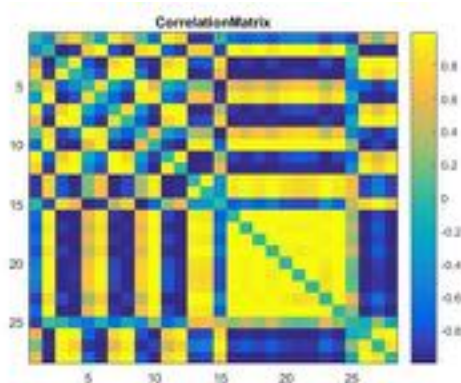


Smart Data Analytics

Wolfgang Karl Härdle

Cathy Yi-Hsuan Chen



Wolfgang Karl Härdle

Wolfgang Karl Härdle completed his Dr. rer. nat. in Mathematics at Heidelberg University and received his habilitation in Economics at Friedrich Wilhelm Universität Bonn.

He was the founder and Director of Collaborative Research Center CRC 373 “Quantification and Simulation of Economic Processes” (1994 - 2003) and also the Director of C.A.S.E. (Center for Applied Statistics and Economics) (2001 - 2014). He founded the CRC 649 “Economic Risk” (2005 - 2016) and is now directing the Sino-German International Research Training Group IRTG1792 “High dimensional non stationary time series analysis” (2013-2018). He has been teaching Master courses at Ladislaus von Bortkiewicz Chair of Statistics at Humboldt-Universität zu Berlin for more than twenty years.

His research focuses on dimension reduction techniques, computational statistics and quantitative finance. He has published many books and more than 300 papers in top statistical, econometrics and finance journals and has a high citation indices across different ranking platforms.

He is among the top 1% of economists registered at REPEC and has similar top notch rankings in other scales, such as the Handelsblatt ranking, Scopus and Google scholar.

His professional experience includes financial engineering, analysis of unstructured data and dynamic decision analytics. He currently focuses his research on sentiment distillation, crypto currencies and DEDA - Digital Economy & Decision Analytics. He has supervised more than 50 PhD students and is holding up long-term research relations to partners in the USA, Singapore, Prague, Warsaw, Paris, Cambridge, Beijing, Xiamen and Taipei among others.



Cathy Yi-Hsuan Chen

Cathy Yi-Hsuan Chen received her PhD diploma in Finance at National Cheng-Chi University. She is since 2015.8 Associate Professor at the School of Business & Economics in Humboldt-Universität zu Berlin, and a principal investigator of the International Research Training Group 1792 – High Dimensional Non Stationary Time Series. She is the member of Princeton-Humboldt Cooperation and Collective Cognition Network (CoCCoN), and visiting fellow of Sim Kee Boon Institute for Financial Economics, Singapore Management University.

She has been teaching Master courses at Ladislaus von Bortkiewicz Chair of Statistics at Humboldt-Universität zu Berlin since 2015. The courses she has taught include *Statistics of Financials markets*, *Quantitative Finance*, *Economic Risk Seminar* and *Q-Kolleg* (a joint course collaborated with the National University of Singapore).

She had a five-years of consulting experience in the insurance industry. Through internal rating-based models, developed and customised programming training (Matlab), a company she consulted can benefit from superior risk management, employee training and fulfil all the obligations and regulations required by the authorities. She is currently heading a “transfer project” between Humboldt-Universität and Deutsche Bank, and focussing on credit risk modelling and stress testing.



Her research focuses on quantitative finance, risk modeling and management, behavior finance and text mining in finance. She has published the papers in reputable journals in the past 10 years. She currently focuses her research on text extraction, text analytics, textual sentiment distillation, textual sentiment modeling, lexicon construction, investor decision and sentiment.

SDA Learning Objectives

The SDA course presents tools and concepts for unstructured data with a strong focus on applications and implementations. It presents the decision analytics in a way that is understandable for non-mathematicians and practitioners who are confronted with day to day number crunching statistical data analysis. All practical examples may be recalculated and modified: software and Quantlets are in www.quantlet.de. The SDA course endows the practitioner with ready to use practical tools for smart data analytics.

SDA Structure

Data are everywhere and the ubiquitous availability of huge amounts of data makes it necessary to develop smart data analytics. Out of the plethora of tools that are available for many scientific disciplines this course offers for the common data analyst an easy access to all levels of analysis without deep computer programming knowledge. SDA provides a wide variety of exercises. In addition a full set of slides is provided making it easier for the participants to reanalyze the presented material. The R and Python programming language are becoming the lingua franca of computational data analysis. They are the common smart data analysis software platforms used inside corporations and in academia. Both are OS independent free open-source programs which are popularized and improved by hundreds of volunteers all over the world.

SDA Literature

Franke J, Härdle WK and Hafner C (2015) Statistics of Financial Markets: an Introduction. 4th ed., Springer Verlag, Heidelberg. ISBN: 978-3-642-54538-2

Chen C YH, Härdle WK and Overbeck L (2017) Applied Quantitative Finance. 3rd extended ed., Springer Verlag, Heidelberg.

Härdle WK, Simar L (2015) Applied Multivariate Statistical Analysis. 4th ed., Springer Verlag, Heidelberg. ISBN 978-3-662-45170-0

Härdle WK, Okhrin O and Okhrin Y (2017) Basics of Computational Statistics, Springer Verlag, Heidelberg.

All examples are presented in R or Python. The Quantlets are available here:

www.quantlet.de



Schedule

| Unit 1 | |
|-----------------------|---|
| What do we see? | <ul style="list-style-type: none">• Basic concepts• Data Management• Structuring Data elements |
| Unit 2 | |
| Data Analysis | <ul style="list-style-type: none">• Sentiment extraction• Stemming, lemmatizing• DTM Dynamic Topic Modeling |
| Unit 3 | |
| Modern Data Analytics | <ul style="list-style-type: none">• Cluster Analysis and Classification• TEDAS Tail Event Driven Asset Allocation• CRIX a CRYPTO currency Index |
| Unit 4 | |
| Smart Data Analytics | <ul style="list-style-type: none">• R tools for text mining• text mining in Quantitative Finance• Applications & Empirics |
| Unit 5 | |
| Smart Data Analytics | <ul style="list-style-type: none">• Network Geometry• Tail Event driven Network AutoRegression• DYTEC DYnamic Tail Event Curves |
| Unit 6 | |
| Smart Data Analytics | <ul style="list-style-type: none">• Financial Risk Meter• DDI Networks Topology• Q3 D3 LSA |

Contact

Wolfgang Karl Härdle
Ladislaus von Bortkiewicz Chair of Statistics
School of Business and Economics
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin, Germany

Telefone **+49 30 2093 5631**

FAX **+49 30 2093 5649**

E-Mail **stat@wiwi.hu-berlin.de**

Links



hu.berlin/93629



hu.berlin/irtg1792



<http://crix.hu-berlin.de>



hu.berlin/rdc

financialriskmeter **hu.berlin/frm**



<http://quantlet.de>

Appendix

Homework examples:

Unit 1 (3h)

1. Calculate the increase of memory of PCs over the last 30 years and check whether the FMRI analysis could have been done 20 years ago
2. prepare 2-5 slides explaining logistic regression
3. install R and run simple programs from Quantlet.de, make sure you have a Github account.

Unit 2 (3h)