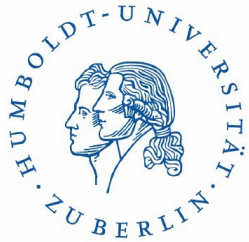
# Scientometrics on Quantnet 2.0 @ GitHub

Aleksandr Bartuli
Alona Zharova

Ladislaus von Bortkiewicz Chair of Statistics and Economics
Master's Program in Economics and Management Science
Collaborative Research Center 649: Economic Risk
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://sfb649.wiwi.hu-berlin.de
http://wiwi.hu-berlin.de/de/international/mems

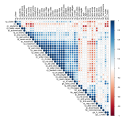





# Scientometrics

- Studies the development of science and the mechanism of scientific research in a quantitative way
- Covers measurement of articles impact, reference sets of articles, the impact of journals and institutes, scientific citations
- Uses Bibliometrics
  - Statistical analysis of written publications
  - Citation and content analyses as commonly used methods
  - Basis for Academic Rankings





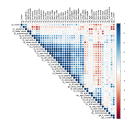
# Objectives

(i) Provide the transparency of used data and research results

(ii) Improve the availability of used codes

(iii) Enhance collaboration in scientometrics



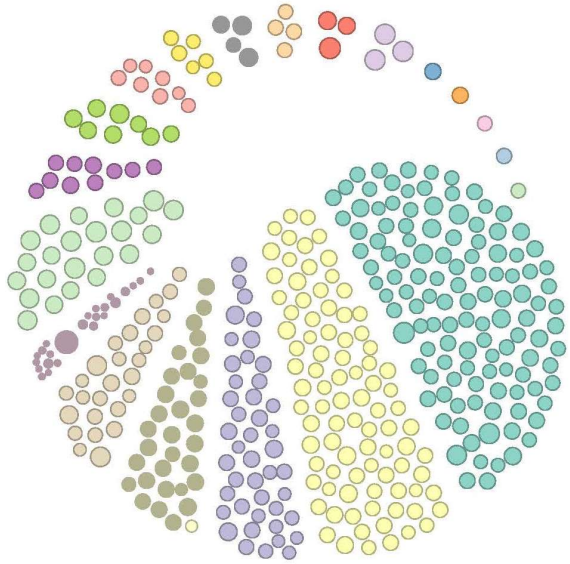


# Outline

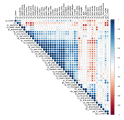
1. Motivation ✓
2. Methodology
3. Data
4. Empirical Results
5. Conclusions

# QuantNet 2.0

- QuantNet – open access code-sharing platform
- Quantlet – statistics-related document and program code
- Knowledge discovery via information visualization

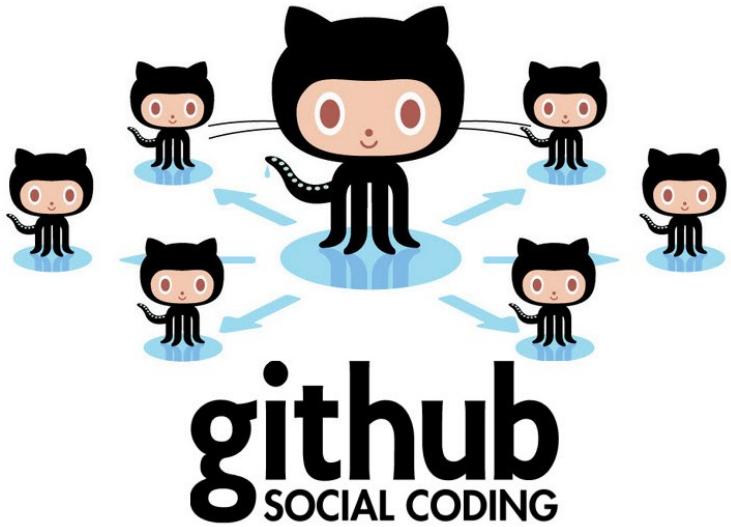






# GitHub

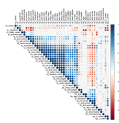
- One of the largest Web-service for IT projects with more than 11 million users
- Free service for open code projects
- Git – distributed version control system
- Hub – collaboration platform



# QuantNet 2.0 @ GitHub

- Fully integrated with GitHub
- Transparency of the used methodology
- Easy of discovery, usage and improvements



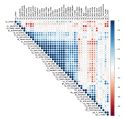


# Data

- Academic Rankings
  - Handelsblatt (HB)
  - Research Papers in Economics / RePEc (RP)
  - Google Scholar (GS)

- Data matrix
  - 3011 researchers
  - HB: yearly ranks from 2009 to 2014, age, subject, etc. (42 parameters)
  - RP: rankings from January 2015 (77 par.)
  - GS: yearly scores from 2007 to 2014 (16 par.)





# Hexagon plot

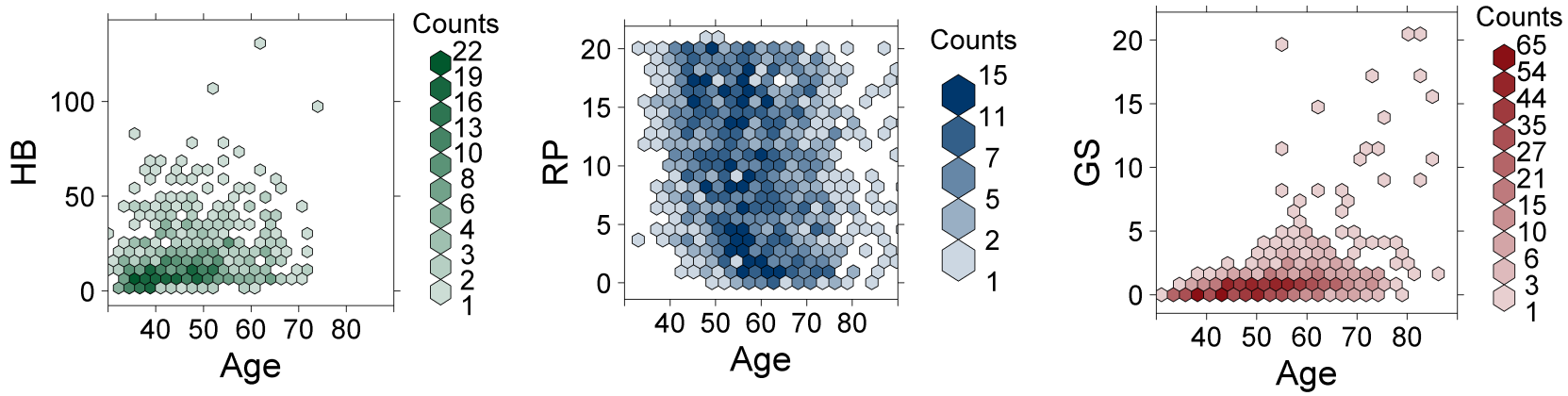
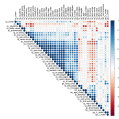



Figure 1. Hexagon plots between age and main ranking scores of HB, RP and GS rankings. ARRhexage





# Scatter and Hexagon plots

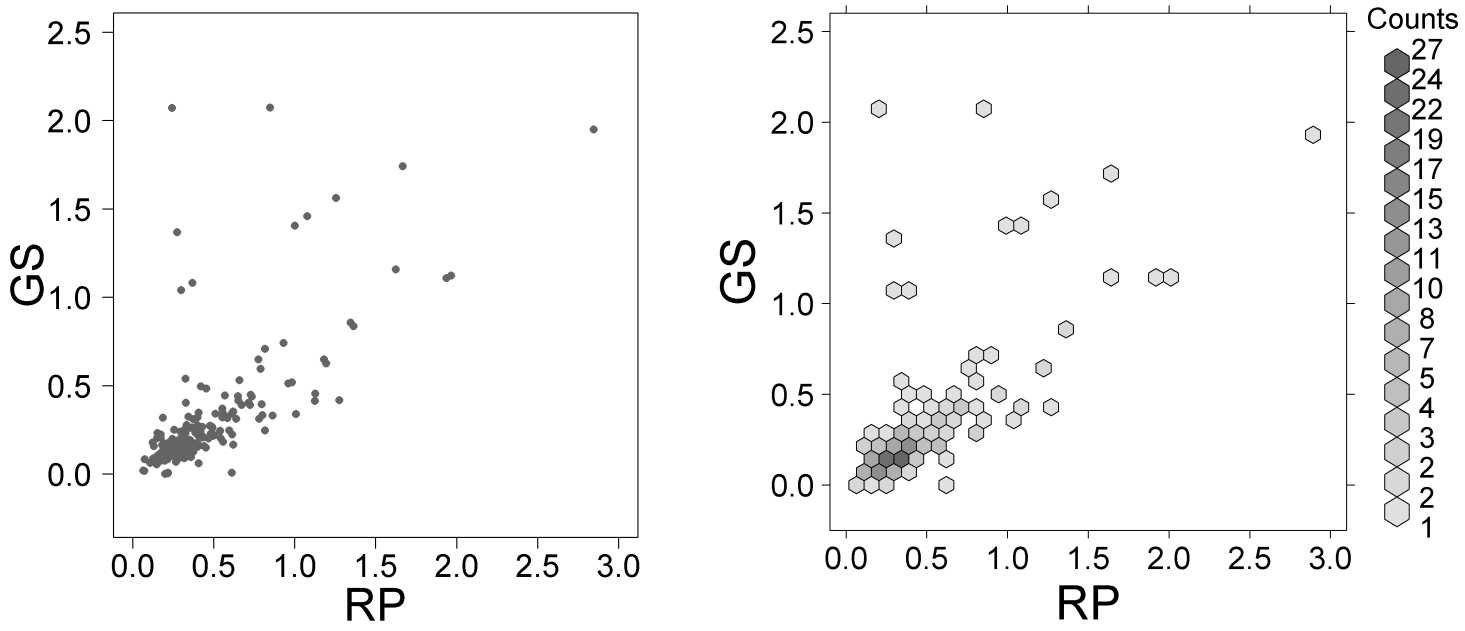
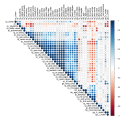



Figure 2. Scatter and hexagon plots between the number of citations listed in RP and GS rankings. ARRhexcit





# Scatter and Hexagon plots

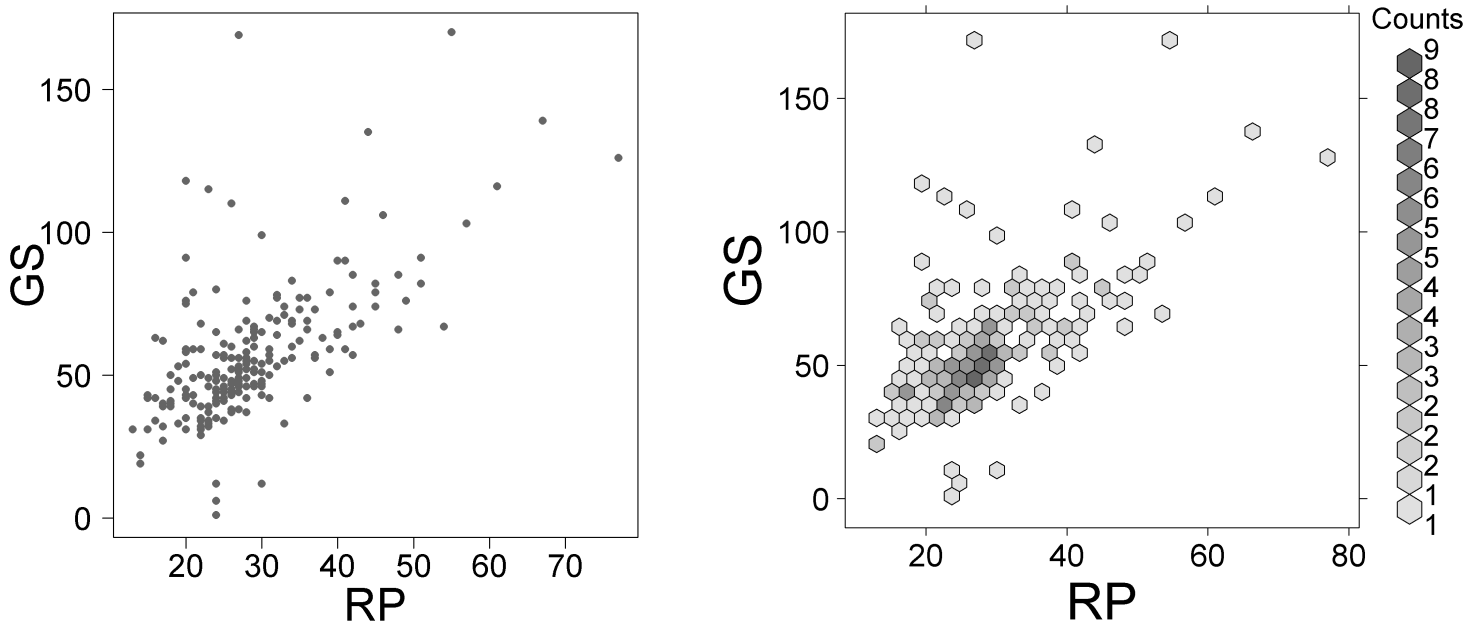
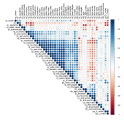



Figure 3. Scatter and hexagon plots between the h-index score listed in RP and GS rankings. ARRhexhin





# Histogram

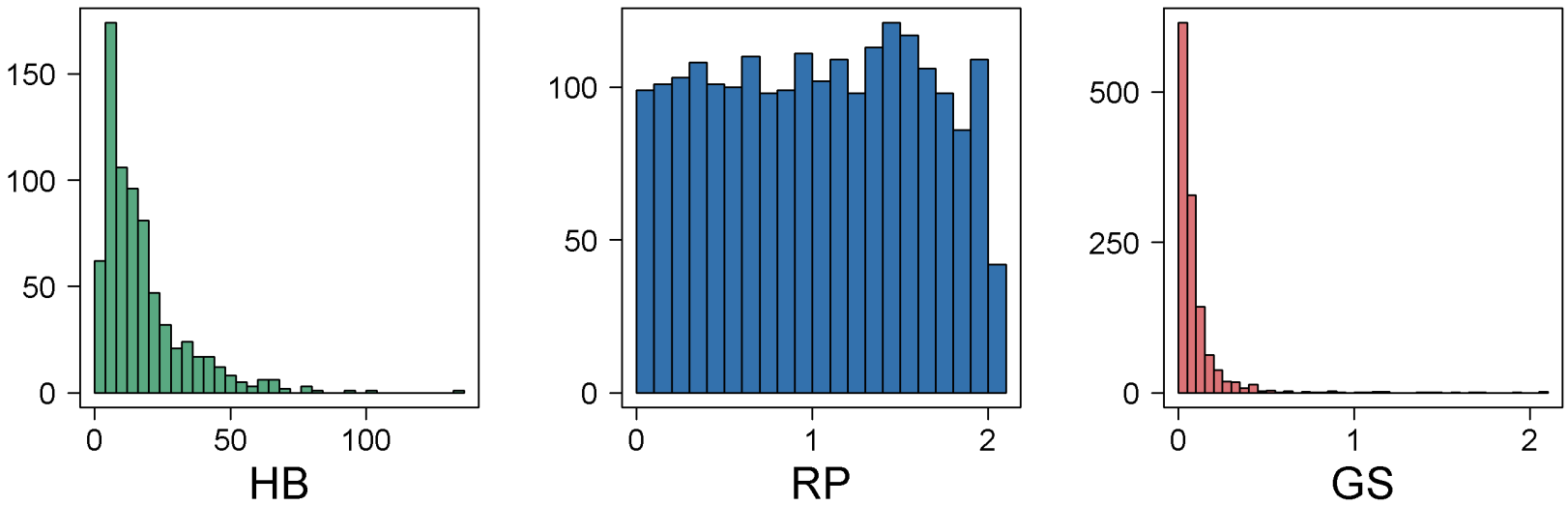
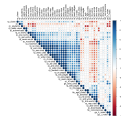



Figure 4. Histograms of HB common score, RP average rank score and GS citations.

ARRhismer





# Mosaic plot

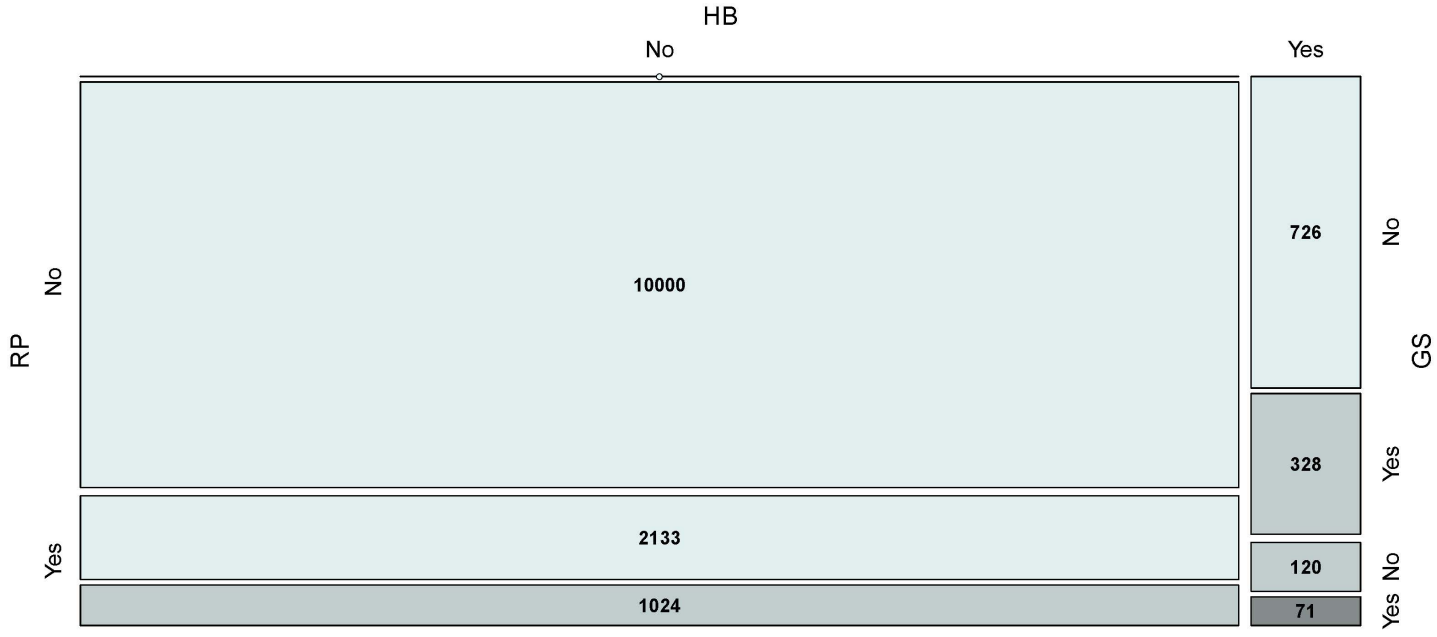
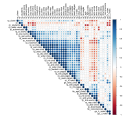



Figure 5. Mosaic plot of number of researchers, when merging of HB, RP and GS rankings takes place. ARRmosage





# Mosaic plot

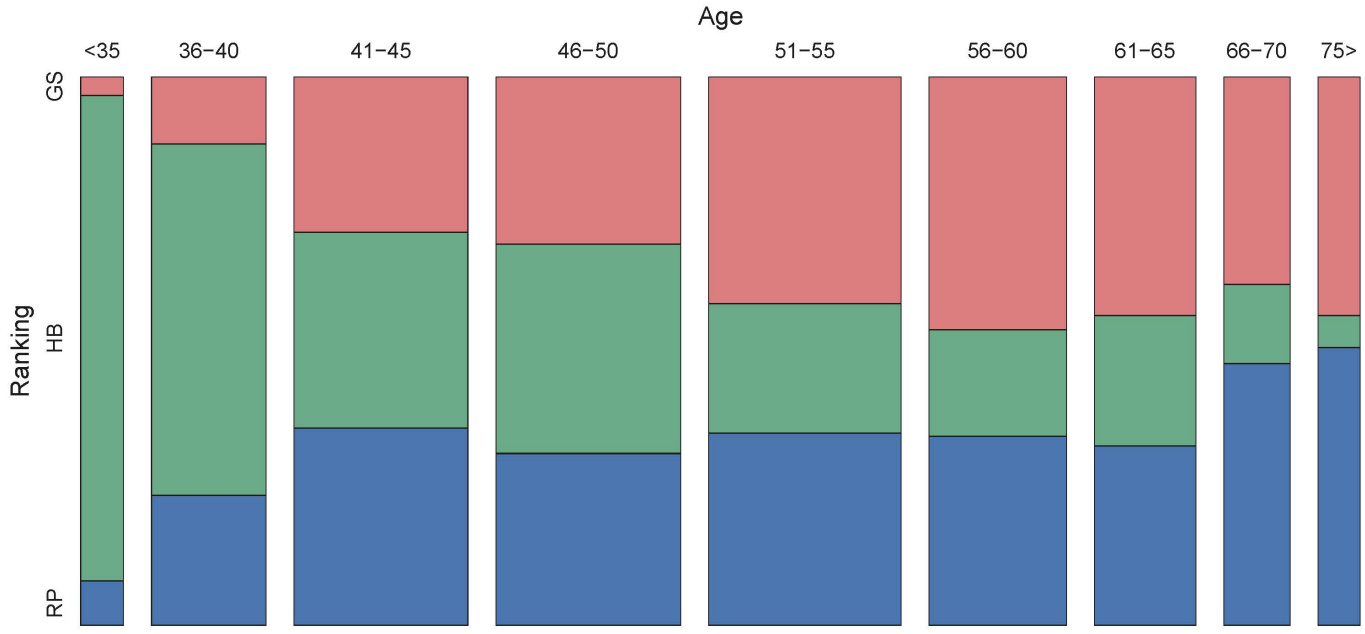



Figure 6. Mosaic plot of Top-700 scientists within different age groups, when merging of HB, RP and GS rankings takes place.

ARRmosagegr

# Correlation matrix

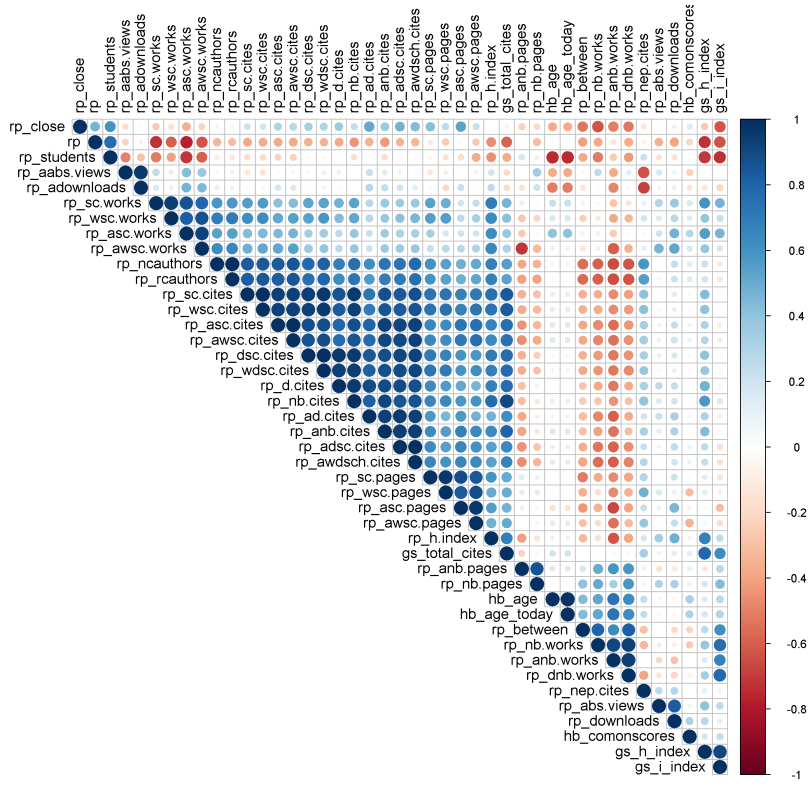
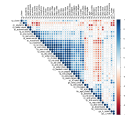



Figure 7. Correlation matrix for 43 variables of HB, RP and GS rankings in an upper triangular matrix. The values are clustered.

ARRcormer

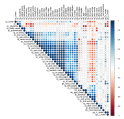


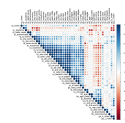


# Conclusions

- Knowledge discovery and collaboration in scientometrics
- Boosting transparent and reproducible science
- Contributing to good scientific practice

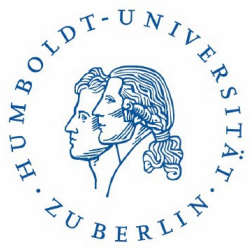






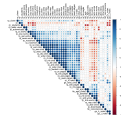
Aleksandr Bartuli
Alona Zharova

Ladislaus von Bortkiewicz Chair of Statistics and Economics
Master's Program in Economics and Management Science
Collaborative Research Center 649: Economic Risk
Humboldt–Universität zu Berlin
http://lvb.wiwi.hu-berlin.de
http://sfb649.wiwi.hu-berlin.de
http://wiwi.hu-berlin.de/de/international/mems





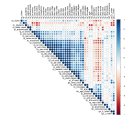
# References


Franke, J., Härdle, W., and Hafner, C.
*Statistics of Financial Markets: an Introduction. Third edition.*
Springer, 2011

Härdle, W. and Simar, L.
*Applied multivariate statistical analysis. Second edition.*
Springer Science & Business Media, 2007

Borke, L. and Neuhoff, D.
*Implementing Collaborative Reproducible Research. QuantNet 2.0 @ GitHub.*
Presentation at CRC 649 Jour Fixe Fall 2015, 2015



# References


Borke, L. and Härdle, W.
*Q3-D3-LSA (Quantlets, Quantnet, Quantmining – Data-Driven Documents – Latent Semantic Analysis).*
Presentation at CRC 649 Conference 2015, 2015

Zharova, A.
*R codes to ARR (Academic Rankings Research).*
Quantlet Repository on Github, 2015