



Data Mining Organisation

Prof. Dr. Stephan Trahasch
Offenburg University of Applied Sciences

Vorlesung und Übungen

- Folien, Aufgaben und Materialien in GitHub
<https://github.com/dminfm/datamining-ws1819>
<https://github.com/dminfm/r-tutorial>
- Erstellen Sie einen kostenlosen Account und senden Sie mir Ihren Account-Name zu.

Prüfungsleistungen

Vorlesung → Klausur

Praktikum

- Aufgaben und Termine in GitHub
- Es besteht Anwesenheitspflicht
- Alle zwei Wochen Aufgaben. Siehe Zeitplan in GitHub.
- kontinuierliche und sinnvolle Bearbeitung der Aufgaben

Themen

- Einführung bzw. Wiederholung Data Mining
- Klassifikation
Random Forests, Support Vector Machines, XGBoost
- Evaluation von Modellen
- Feature Selection, PCA, SVD
- Outlier Detection
- Recommender Engines
- Assoziationsregeln
- ...

Übungen mit R

■ R

<https://cran.r-project.org>

■ RStudio

<https://www.rstudio.com/>

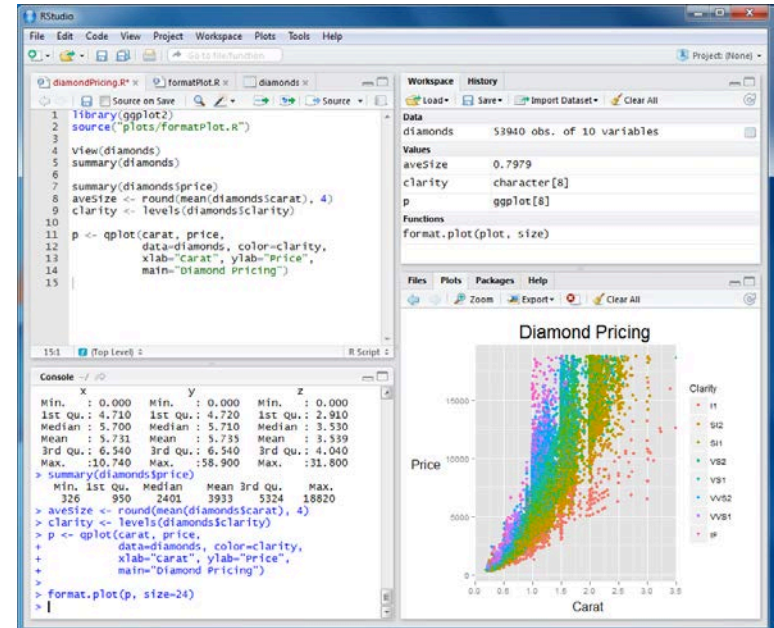
■ R Introduction

- Use tidyverse!
- R for DataScience by Hadley Wickham

<http://r4ds.had.co.nz>

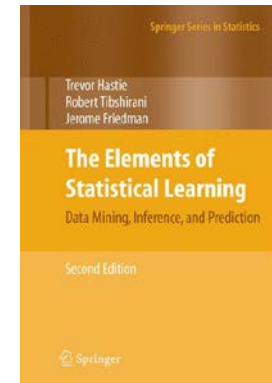
- Take a free R course

<https://www.datacamp.com/courses/free-introduction-to-r>



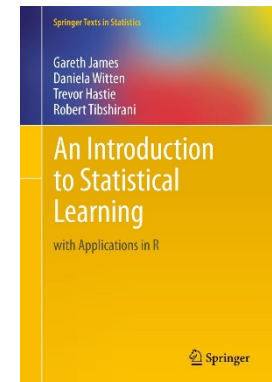
Hastie, T., Tibshirani, R., and Friedman, J. H., *The elements of statistical learning: Data mining, inference, and prediction*, 2nd ed. Springer series in statistics. New York: Springer, 2009.

Online: <https://web.stanford.edu/~hastie/ElemStatLearn/>



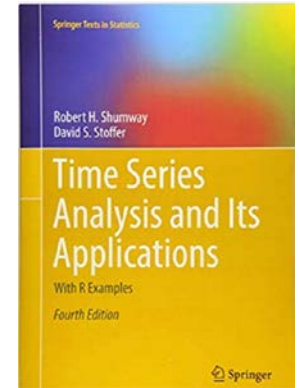
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2014). *An introduction to statistical learning: With applications in R*. Springer texts in statistics. New York: Springer.

Online: <https://www-bcf.usc.edu/~gareth/ISL/>



Shumway, R. H., & Stoffer, D. S. (2017). *Time series analysis and its applications: With R examples* (Fourth Edition). *Springer texts in statistics*. Cham, Switzerland: Springer.

Online: <https://www.stat.pitt.edu/stoffer/tsa4/>

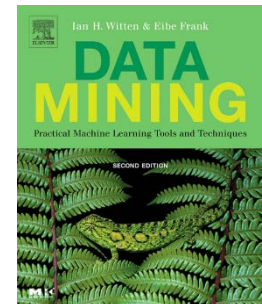
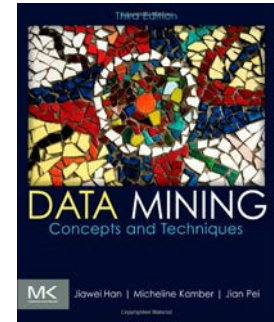


Links zu weiterer Literatur unter

<http://ks.informatik.hs-offenburg.de/lehre/data-mining/free-data-mining-ebooks>

Weitere Literatur

- Han, J., Kamber, M., and Pei, J., *Data Mining: Concepts and Techniques*, 3rd ed. Burlington: Elsevier Science, 2011.
- Witten, I. H. and Hall, M. A., *Data mining: Practical machine learning tools and techniques*, 3rd ed. Burlington, MA: Morgan Kaufmann, 2011.
- Alpaydın, E., *Maschinelles Lernen*. München: Oldenbourg, 2008.



Im Web

Das Portal für DM

<http://www.kdnuggets.com/>

Reale Wettbewerbe von Unternehmen

<http://www.kaggle.com/>

...

2016 Magic Quadrant for Advanced Analytics Platforms



2017 Magic Quadrant for Data Science Platforms

