References

1. Ricardo Baeza-Yates and Berthier Ribeiro-Neto. *Modern Information Retrieval*. ACM Press, 2nd edition, 2011.
2. Carlo Batini and Monica Scannapieca. *Data Quality: Concepts, methodologies and techniques*. Springer, 2006.
3. Richard Belew. *Finding Out About*. Cambridge University Press, 2008.
4. Michael Berthold, Christian Borgelt, Frank Höppner, and Frank Klawonn. *Guide To Intelligent Data Analysis*. Texts in Computer Science. Springer, 2010.
5. Tamraparni Dasu and Theodore Johnson. *Exploratory Data Mining and Data Cleaning*. John Wiley and Sons, 2003.
6. AnHai Doan, Alon Halevy, and Zachary Ives. *Principles of Data Integration*. Morgan Kaufmann, 2012.
7. Xin Luna Dong and Divesh Srivastava. Big data integration. *Synthesis Lectures on Data Management*, 7(1):1–198, 2015.
8. Joseph M. Hellerstein. Quantitative data cleaning for large databases. Technical Report, UC Berkeley, 2008. https://dsf.berkeley.edu/jmh/papers/cleaning-unece.pdf
9. Jeroen Janssens. *Data Science at the Command Line*. O'Reilly, 2015.
10. Daniel Lemire and Anna Maclachlan. Slope one predictors for online rating-based collaborative filtering. *CoRR*, abs/cs/0702144, 2007.
11. Gordon Linoff. *Data Analysis Using SQL and Excel*. Wiley, 2008.
12. Christoper D. Manning, P. Raghavan, and H. Schutze. *Introduction to Information Retrieval*. Cambridge University Press, 2008.
13. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. The MIT Press, 2012.
14. Kevin Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012.
15. David Skillcorn. *Knowledge Discovery for Counterterrorism and Law Enforcement*. Chapman and Hall/CRC Data, 2008.
16. Pang-Ning Tan, Michael Steinbach, Anuj Karpatne, and Vipin Kumar. *Introduction to Data Mining*. Pearson, 2nd edition, 2019.
17. Robert Trueblood and John Lovett. *Data Mining and Statistical Analysis Using SQL*. Apress, 2001.
18. Hadley Wickham. The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1), 2011.
19. Hadley Wickham. Tidy data. *Journal of Statistical Software*, 51(10), 2014.
20. Hadley Wickham and Garret Grolemund. *R for Data Science*. O'Reilly, 2017.

A. Badia, *SQL for Data Science*, Data-Centric Systems and Applications,
https://doi.org/10.1007/978-3-030-57592-2

Index

A. Badia, *SQL for Data Science*, Data-Centric Systems and Applications,
https://doi.org/10.1007/978-3-030-57592-2