References

DATA SCIENCE ESSENTIALS

References

C. C. Aggarwal, *Data Classification: Algorithms and Applications*. CRC Press, 2015.

C. C. Aggarwal, *Data Mining: The Textbook*. Cham: Springer, 2015.

C. C. Aggarwal, C. K. Reddy, *Data Clustering: Algorithms and Applications*. CRC Press, 2014.

P. Boily, J. Schellinck, *Data Understanding, Data Analysis, and Data Science*. Data Action Lab, 2022.

D. Brin, *The Transparent Society: Will Technology Force Us to Choose Between Privacy and Freedom?* Perseus, 1998.

Coursera, "Introduction to Data Engineering."

References

T. H. Davenport and D. J. Patil, "Data Scientist: The Sexiest Job of the 21st Century," *Harvard Business Review*, Oct. 2012.

T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed.* Springer, 2008.

H. Konsek, "Automating Data Pipelines: Types, Use Cases, Best Practices." Soft Kraft.

J. Kunigk, I. Buss, P. Wilkinson, and L. George, *Architecting Modern Data Platforms: A Guide to Enterprise Hadoop at Scale*. O'Reilly Media, 2018.

T. Malaska and J. Seidman, *Foundations for Architecting Data Solutions: Managing Successful Data Projects*. O'Reilly Media, 2018.

References

C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown, 2016.

T. Orchard and M. Woodbury, *A missing information principle: theory and applications*. University of California Press, 1972.

R. W. Paul and L. Elder, *Understanding the Foundations of Ethical Reasoning*, 2nd ed. Foundation for Critical Thinking, 2006.

What is Data Engineering? Everything You Need to Know in 2022. phData, 2022.

F. Provost and T. Fawcett, *Data Science for Business*. O'Reilly, 2015.

References

T. Raghunathan, J. Lepkowski, J. Van Hoewyk, and P. Solenberger, "A multivariate technique for multiply imputing missing values using a sequence of regression models," *Survey Methodology*, vol. 27, no. 1, pp. 85–95, 2001.

D. B. Rubin, *Multiple imputation for nonresponse in surveys*. Wiley, 1987.

R. Schutt and C. O'Neill, *Doing Data Science: Straight Talk from the Front Line*. O'Reilly, 2013.

simplystatistics.org, "An interactive visualization to teach about the curse of dimensionality."

S. van Buuren, *Flexible imputation of missing data*. CRC Press, 2012.

A. Watt, *Database Design*. BCCampus, 2014.

Data Governance in the GoC

Central point of reference for GoC (Digital Government website):

- Strategic plans, policies, standards and guidelines related to government digital services

Report to the Clerk of the Privy Council:

- A Data Strategy Roadmap for the Federal Public Service

Treasury Board Secretariat (selected):

- Policy on Service and Digital
- Government of Canada Strategic Plan for Information Management and Information Technology 2017 to 2021
- Digital Operations Strategic Plan: 2018-2022
- Government of Canada Cloud Adoption Strategy: 2018 update

Innovation, Science and Economic Development Canada:

- Canada's Digital Charter in Action: A Plan by Canadians, for Canadians