

# THE MATHEMATICS OF DATA SCIENCE

MSRI Undergraduate Program

June 16 to July 29, 2018

- [1] G. GROLEMUND and H. WICKHAM, *R for data science: Visualize, model, transform, tidy, and import data*, O'Reilly, Sebastopol, CA, 2017, ISBN 978-1-491-91039-9. <http://r4ds.had.co.nz>.
- [2] T. HASTIE, R. TIBSHIRANI, and J. FRIEDMAN, *The elements of statistical learning: Data mining, inference, and prediction*, 2nd ed., Springer, New York, NY, 2009, ISBN 978-0-387-84857-0; 978-0-387-84858-7. MR 2722294. Zbl 1273.62005. doi: 10.1007/978-0-387-84858-7.
- [3] G. JAMES, D. WITTEN, T. HASTIE, and R. TIBSHIRANI, *An introduction to statistical learning: With applications in R*, Springer Texts in Statistics **103**, Springer, New York, NY, 2013, ISBN 978-1-4614-7137-0; 978-1-4614-7138-7. MR 3100153. Zbl 1281.62147. doi: 10.1007/978-1-4614-7138-7.
- [4] W. MCKINNEY, *Python for data analysis: Data wrangling with pandas, NumPy, and IPython*, 2nd ed., O'Reilly, Sebastopol, CA, 2017, ISBN 978-1-491-95766-0. <http://wesmckinney.com/pages/book.html>.
- [5] M. E. J. NEWMAN, *Networks: An introduction*, Oxford University Press, Oxford, 2010, ISBN 978-0-19-920665-0. MR 2676073. Zbl 1195.94003. doi: 10.1093/acprof:oso/9780199206650.001.0001.
- [6] T. PETROU, *Pandas cookbook: Recipes for scientific computing, time series analysis and data visualization using Python*, Packt, Birmingham, UK, 2017, ISBN 978-1-78439-387-8. <https://www.packtpub.com/big-data-and-business-intelligence/pandas-cookbook>.
- [7] G. STRANG, *Introduction to linear algebra*, 5th ed., Wellesley-Cambridge, Wellesley, MA, 2016, ISBN 978-09802327-7-6. MR 575349. Zbl 1351.15002. <http://math.mit.edu/~gs/linearalgebra>.