
data journalism reading

Release 0.1

Aidan O'Donnell

Sep 12, 2023

CONTENTS

1	Basics	3
2	Census 2021	5
3	Covid-19	7
4	Digital security	9
5	Ethics	11
6	Excel & Sheets	13
7	FOI & EIR	15
8	Journalism	17
9	Legal	19
10	Machine Learning & CS	21
11	Mapping	23
12	Network analysis	25
13	Numbers & statistics	27
14	OSINT	29
15	Scraping & APIs	31
16	Storytelling with data	33
17	Visualisation	35

This is the reading list for data journalism and computational journalism classes at [Jomec](#) (Cardiff University) for undergraduate and postgraduate journalism students. Some references have a link to the Cardiff University library that is only available to registered students but most of the works below are online and the others are generally available from book dealers.

BASICS

Bounegru, Liliana & Gray, Jonathan. *The data journalism handbook: towards a critical data practice*. Amsterdam University Press, 2021

Bounegru, Liliana, Chambers, Lucy & Gray, Jonathan. *The data journalism handbook (One)*. 2011 Datajournalism.com

Dasgupta, Subrata. The ‘stuff’ of computing, *Computer Science. A very short introduction*. OUP, 2016, p. 1-12

Felle, Tom et al. *Data journalism: inside the global future*. Abramis, 2015. [Library resource](#)

Global Investigative Journalism Conferences. 2019 | 2021

GIJN Introductory videos.

Harford, Tim. Ponder your personal experience. *How to make the world add up. Ten rules for thinking differently about numbers*. Bridge Street Press, 2020. [Library resource](#)

Houston, Brant. *Data Journalism. A brief history*. GIJN, 2015 and Datajournalism.com, 2021

Houston, Brant. *Data for journalists: a practical guide for computer-assisted reporting*. Routledge, 2019. [Library resource](#)

International Journalism Festival (Perugia) 2019. [Video playlist](#)

Kovach, Bill & Rosenstiel, Tom. *The Elements of Journalism. What newspeople should know and the public should expect*. Three Rivers Press, 2012. [Library resource](#)

Mair, John et al. *Data journalism: past, present and future*. Abramis, 2017. [Library resource](#)

Mayhew, Henry. *London labour and the London poor [1851-52]*. Penguin, 1985. [Library resource](#)

Meyer, Philip. *The journalism we need. Precision Journalism*. Rowman & Littlefield, 1973, 2002, 4th ed., p. 1-17. [Library resource](#)

Mutsaers, Bruce et al. *Data Journalism in the Global South*. Springer, 2020. [Library resource](#)

News impact summit: democratising data (June 2021). [Video](#), 5h & programme

Pilhofer, Aron. *Getting started with datajournalism*. 2017. 40' [Video](#)

Reilley, Mike & Sunne, Samantha. *Data + Journalism: A story-driven approach to learning data reporting*. Routledge, 2023. [Library resource](#)

Richards, Alex. *Foundations of data and digital journalism*. Routledge, 2023. (US focused)

Rogers, Simon. *Facts are sacred*. Faber & Faber, 2013.

Rukmini, S. *Whole numbers and half truths: what data can and cannot tell us about modern India*. Context, 2021. [Library resource](#)

SPJ. [Data tools & tipsheets](#) [Overpowering list but worth browsing]

Tong, Jingrong. Data Journalism. *Journalism in the data age*. Sage, 2022, p. 59-88.

CENSUS 2021

BBC. Census to provide snapshot of life during pandemic. March 2021

Benton, Pete. Census 2021: ask me anything. Reddit. October 2020

Dunn, Ed. “Census 2021: the count is done, the data is in, so what happens next?” ONS.

Hutchinson, Roger. *The Butcher, the baker, the candlestick maker*. Abacus, 2017

Jack, Andrew & Giles, Chris. Academics warn census in pandemic risks distorting government grants. Financial Times. March 2021

COVID-19

DATA SOURCES

- [ECDC](#)
- [ONS](#)
- [PHW](#) (headline summary / download data)
- [UK.GOV](#)
- [Wales vaccination numbers](#)

[Barr, Caelinn et al., Counting the human cost of Covid-19: Numbers tell a story words can't. Guardian, May 2020](#)

[Bradshaw, Paul. How to brainstorm Covid-19 data story ideas. Datajournalism.com, April 2020](#)

[Bronner, Laura et al., Why it's so freaking hard to make a good Covid-19 model, Fivethirtyeight. March 2020](#)

[Burn-Murdoch, John. Coronavirus trajectory tracker explained. 2020. 6' Video](#)

[Burn-Murdoch, John. EJC Trust conference: The FT's trajectory tracker. 2020. 20' Video](#)

[Byrd, Josh et al. These charts track how coronavirus is spreading around the world. ABC \(au\), 2020](#)

[Chang, Kenneth. A different way to chart the spread of coronavirus. NY Times. 2020](#)

[Field, Kenneth. Mapping coronavirus data responsibly. Feb 2020](#)

[First Draft News. Covering coronavirus: an online course for journalists. 2020](#)

[Frumkin, Rebekah. How to draw the coronavirus. The Paris Review. 2020](#)

[FT. Coronavirus tracker.](#)

[Georgia public health. What is a moving average and why is it useful? 2020](#)

[GIJN. Evaluating the evidence. 2020. 1h 20' Video](#)

[GIJN. Covid-19 hub](#)

[Harford, Tim. Statistics, lies and the virus: Tim Harford's five lessons from a pandemic. Financial Times. September 2020](#)

[Johnson, Oliver. How to read coronavirus graphs. The Spectator. April 2020](#)

[Krakov, Andy. 13 things to visualise about Covid-19 besides case loads. Medium. March 2020](#)

[Makulic, Amanda. Ten considerations before you create another chart about Covid-19, Medium. March 2020](#)

[Meers, Jelta, How data journalists exposed the real Covid-19 death toll in Brazil. GIJN, July 2020](#)

[ONS. Project Cairo \(overall portfolio but includes Covid-19 work\)](#)

Philip, Rowan. How to tackle the global undercount in Covid-19 deaths: reporters offer tips and techniques, GIJN, May 2020

Reuters. Breaking the wave: How to graph the curve

Reuters. A deluge of death in northern Italy, March 2020

Riva, Catherine & Tinari, Serena. Guide for journalists: Investigating health & medicine. GIJN

Spiegelhalter, David. Communicating statistics, risk and uncertainty in the age of Covid. 40' Video

Spiegelhalter, David & Masters, Anthony. Covid by Numbers: making sense of the pandemic with data. Penguin, 2021.

Stevens, Harry. Why outbreaks like coronavirus spread exponentially, and how to 'flatten the curve'. Washington Post. March 2020

Tableau. Covid-19 hub

DIGITAL SECURITY

— See first section of Bazzell in **OSINT** section on setting up machines for online investigation

Committee to Protect Journalists. Digital safety kit for journalists

Global Cyber Alliance. Cybersecurity toolkit for journalists

McGregor, Susan. *Information security essentials: a guide for reporters, editors, and newsroom leaders*. Columbia University Press, 2021 [Library resource](#)

Rory Peck Trust. Digital security resources for freelancers

Reporters Without Borders. Digital security guide

Totem project

— See **Journalism** section for Codes and Guidelines

Alsudais, Abdulkareem. Incorrect Data in the Widely Used Inside Airbnb Dataset. 2020

Belmont Report (US govt)

Cabreros, Irineo. Side-stepping safeguards, data journalists are doing science now. Undark. 2021

Chalabi, Mona. Mapping kidnappings in Nigeria. 538. 2014

Christian, Sue Ellen. Cognitive Biases and Errors as Cause — and Journalistic Best Practices as Effect. *Journal of Mass Media Ethics*. 28:3, 2013. p. 160-74. Library resource

Coddington, Mark & Lewis, Seth. Do journalists hide behind sources when they use numbers in the news? Nieman Lab, 2021 & Lawson, B.T. Hiding Behind Databases, Institutions and Actors: How Journalists Use Statistics in Reporting Humanitarian Crises. Journalism Practice, 2021

Craig et al., To Post or Not to Post: Online Discussion of Gun Permit Mapping and the Development of Ethical Standards in Data Journalism. *Journalism & Mass Communication Quarterly*, January 2017

Craig et al. Ethical questions in data journalism and the power of online discussion. Datajournalism.com, 2017

Criado-Perez, Caroline. *Invisible women: data bias in a world designed for men*. Vintage, 2019. Library resource

Diakopoulos, Nick. BuzzFeed's pro tennis investigation displays ethical dilemmas of data journalism. CJR, 2016.

Diakopoulos, Nick. The ethics of predictive journalism. CJR, 2020

Gold, Zachardy & Latonero, Mark. Robots Welcome? Ethical and Legal Considerations for Web Crawling and Scraping. *Washington Journal of Law, Technology & Arts*. 13/3, 2018, p. 275. Library resource

Golumbia, David. Fair Game: commonly used by researchers and journalists, data scraping is an underacknowledged privacy concern. 2022

Harris, Jacob. Connecting with the dots

Heydt, Michael. Scraping. Code of conduct. *Python web scraping cookbook*, Packt, 2018. Library resource

Knight Center. Ethics and equity in data journalism. Online course. 2020

Marconi, Francesco. *Newsmakers: Artificial Intelligence and the Future of Journalism*. Columbia University Press, 2020. Library resource

Mitchell, Ryan. Legalities & ethics of web scraping (p. 265-79); Note on ethics (p. 217-18). *Web scraping with Python*. O'Reilly, 2nd ed., 2018. Library resource

NPR. Did George Floyd die or was he murdered? 2020

ONS. Web scraping policy

Open Data Institute. Assessing risk when sharing data: a guide. Feb 2022

Reuters. Microsoft's LinkedIn loses appeal over access to user profiles, Sept 2019

Rogers, Richard. The API and the ethics turn. *Doing Digital Methods*. Sage, 2019, p. 206-09

Scrapinghub. Web Scraping Best Practices Guide

Sellers, Andrew. Twenty years of web scraping and the Computer Fraud and Abuse Act. *Boston University Journal of Science & Technology Law*. 24, 2018

Shiab, Nael. On the ethics of web scraping. GIJN, 2015

Singel, Ryan. Netflix spilled your Brokeback Mountain secret, lawsuit claims. *Wired*, 2009

Tameez, Hanaa. What type of local news stories should be automated?, Nielman Lab, 2022

UK Statistics Authority. Ethical considerations in the use of geospatial data. May 2021 [includes checklist]

Velotio. Scraping guidelines & best practices

Wilson, Greg et al. Good enough practices in scientific computing. *PLOS*. 2017

EXCEL & SHEETS

Berkeley Media Lab. A quick and easy guide to spreadsheets

Bradshaw, Paul. *Finding stories with spreadsheets*. Leanpub, 2016

Free Code Camp. 3-hr Google Sheets youtube video

GCF Global. Google spreadsheets

Google. Sheets training & help

Groeger, Lena. Spreadsheets lab

Harford, Tim. The tyranny of spreadsheets. FT, June 2021

Houston, Brant. Let the spreadsheet do the math. GIJN. 2019

Katz, Abbott. Excel dynamic array functions: what datajournalists need to know. Datajournalism.com

Kelion, Leo. Excel: Why using Microsoft's tool caused Covid-19 results to be lost. BBC News, Oct 2020.

Lambrechts, Maarten. Cleaning data in Excel: video course. Datajournalism.com. 1h 30' video

MacDonald, Matthew. Excel 2013: the missing manual. O'Reilly

Muth, Lisa Charlotte. How to match data with VLOOKUP.

NY Times. Data course spreadsheets cheatsheets

Quartz. Bad data

Webster, Mary-Jo. Excel basics

Webster, Mary-Jo. Excel magic. 2019

FOI & EIR

11KBW. Pantopicon (blog)

Amin, Lucas. Access denied: the UK government attack on Freedom of Information. Open Democracy, October 2021

Amin, Lucas & Montague, Brendan. *FOIA without the lawyer: freedom, information and the press*. TCIJ, 2012

Amin, Lucas & Montague, Brendan. *EIRs without the lawyer: the journalist's guide to the Environmental Information Regulations*. TCIJ, 2013

Bartlett, Rachel. How to submit a freedom of information request. Journalism.co.uk, 2012

Basnett, Guy & McNamara, Paul. FOI at ten. How all journalists can use the freedom of information act to find great exclusives. Press Gazette. 2015

BBC & Miller, Charles. Five ways to stop your FOI request being rejected. 2015

Brooke, Heather. *Your right to know. A citizen's guide to the Freedom of Information Act*. 2nd ed., Pluto, 2007. Library resource

Burgess, Matt

— FOI directory

— *Freedom of Information: A Practical Guide for UK Journalists*. Routledge, 2015. Library resource

— FOIA. 103 stories that prove Chris Grayling wrong. Guardian, 2015

— Some of the best stories based on FOI requests in 2017

Campaign for Freedom of Information

Committee Evidence on Cabinet Office FOI Clearing House, Oct 2021

Harrison, Sian & Hanna, Mark. The Freedom of Information Act 2000. *McNae's Essential Law for Journalists*, 26th ed., 2022, Chapter 37 (online). Library resource

Gibbons, Paul. The Freedom of Information Officer. Open Data Institute. 2019. 40' Video

GIJN. FOIA this! (article series)

Geoghegan, Peter. FOI. London Review of Books. 43/4, 2021

Goldberg, Kevin. FOI resources. GIJN. 2019

Harris, Kelly, Metcalfe, Eric & Wadham, John. *Blackstone's guide to the Freedom of Information Act 2000*. 5th ed., OUP, 2013. Library resource

Higgerson, David. FOI ideas. 2016

Information Commissioner's Office

— Guide to authorities that respond to FOI. 2017

— How to get information from a public body (for individuals)

— A guide to freedom of information (for organisations)

Legal-leaks. Legal leaks toolkit. A guide for journalists on how to access government information

McIntosh, Toby. Walking the FOI beat. Tipsheet. GIJN. 2019

Miller, Claire, A great big list of FOI ideas. Jan 2013

Morrison, James. Freedom of Information. *Essential public affairs for journalists (7th ed.)*. OUP, 2021, p. 436-52. Library resource (6th ed)

Open Democracy. Freedom of Information (article series)

Reid, Alastair. Ten years of FOI. Resources and tips for journalists. Journalism.co.uk. 2015

Rosenbaum, Martin. 10 things we found out because of FOI. BBC Online. 2015

TBIJ. Bureau launches action over hidden council finances. October 2020

UCL. Chronology of FOI in the UK

UK government guide to FOI

What do they know? FOIs already placed

JOURNALISM

ABC. Style guide

Associated Press. News values & principles

Associated Press. Data Journalism. *The Associated Press Stylebook*. 55th ed., 2020-22. p. 358-63. [Library resource](#)

BBC Academy: journalism students

BBC. Editorial guidelines

BBC. Style guide

Bloomberg. *The Bloomberg Way*. 2014 edition. [Library resource](#)

Carpenter, Siri. *The craft of science writing: selections from the Open Notebook*. Madison, 2020. [Library resource](#)

Corbett, Philip B. Copy edit this! (complete list of quiz links)

Editors' Code of Practice Committee. The Editors' Codebook or Editors' Code

Guardian. Style guide

Impress. Standards code

IPSO. Rulings

NCTJ. Skills videos

NUJ. Code of conduct

Ofcom. The Ofcom broadcasting code

ONS. Style guide

Orwell, George. *Politics and the English Language*. 1946

Publica. "Data style guide". Guide to working with data and publishing apps

Reuters. Handbook of journalism

Baranetsky, Victoria. Data Journalism and the Law [US]. Columbia Journalism Review. 2018

Beriain et al. Data protection in journalism: a practical handbook

Bhandari, Esha & Goodman, Rachel. Data Journalism and the Computer Fraud and Abuse Act: Tips for Moving Forward in an Uncertain Landscape. Paper from Computation+Journalism Symposium (Northwestern), 2017

Borger, Julian. NSA files: why the Guardian in London destroyed hard drives of leaked files. Guardian. 2013

Coventry University. The Computer Misuse Act (R v. Gold & Schifreen (1988))

Defence and Security Media Advisory (DSMA) Committee. Standing Notices

Harrison, Sian & Hanna, Mark. *McNae's Essential Law for Journalists*. 26th ed., 2022. Library resource [Esp. Chapter 28, 'Data protection', p. 433-41]

Georgiev, Peter. A robot commits libel. Who is responsible?, RJI, 2019

Github. The legal side of open source

Github. Licensing a repository

Glanville, Jo. The journalistic exemption (Data Protection Act). LRB. 2018

ICO. Data protection and journalism. A quick guide. 2018

IPSO. Decision 03350-16 InFacts v The Sun. 2016

Lorentzen, Christian. I wasn't just a brain in a jar. Review of Edward Snowden's Permanent Record. London Review of Books. 41/18, Sept 2019

Murray, Andrew. *Information technology law: the law and society*. 4th ed., OUP, 2019. Library resource

MACHINE LEARNING & CS

- Aldhous, Peter. We trained a computer to search for hidden spy planes: this is what it found. Buzzfeed. 2017 with code
- Beckett, Charlie. New powers, new responsibilities: A global survey of journalism and artificial intelligence. 2019
- Bradshaw, Paul. Ten ways AI can be used for investigations (slides). 2019
- Cheung, Paul. Journalism's superfood: AI? Knight Foundation. 2019
- DataFlair. Detecting fake news with python
- Deepnews. Article scorer for journalistic quality
- Fussell, Sidney. AI Shows ExxonMobil Downplayed Its Role in Climate Change. 2021
- Gajtkowski, Adam. Predicting FT trending topics. 2021
- Google News Initiative. Introduction to machine learning
- ICIJ. Datashare (tool for sorting large amounts of documents)
- Merrill, Jeremy B. Chamber of Secrets: teaching a machine what Congress cares about. Propublica. Oct 2017
- Merrill, Jeremy B. Searching Bill de Blasio's emails with the universal sentence encoder. Quartz
- Roberts, Brandon. Dissecting a machine learning powered investigation: Uncovering local property tax evasion using machine learning and statistical modeling. Artificial Informer. April 2019
- Ronderos, Maria Teresa. How innovative newsrooms are using artificial intelligence. GIJN, January 2019
- Soma, Jonathan. Practical data science for journalists
- Stray, Jonathan. Beyond the hype: Using AI effectively in investigative journalism. GIJN, September 2019
- Thurman, Neil et al. (eds.). *Algorithms, Automation & News*. Routledge, 2021. Library resource See also *Digital Journalism*, vol 7/8, 2019. Library resource
- Quartz. Helicopters circling (code)
- Wu, Floris. How machine learning can (and can't) help journalists. March 2019
- Wu, Floris. The future of machine learning in journalism. Storybench. 2019

MAPPING

- Au, Yung. Thinking critically about maps. Exposing the invisible
- Barney, Timothy. How maps show and hide key information about the Ukraine war
- Cutts, Andrew. *QGIS quick start guide*. Packt, 2019 Library resource
- Datawrapper. Walkthrough for live choropleth map
- De Groot, Len. QGIS basics for journalists. Berkeley. 2014
- D'Efilippo, Valentina. Sketching the world. An icebreaker to working with data. Medium, 2020
- Dodge, Martin, Kitchin, Rob & Perkins, Chris. Thinking about maps. *Rethinking Maps: New Frontiers in Cartographic Theory*. Routledge, 2009, p. 1-25
- Ericson, Matthew. When maps shouldn't be maps. 2011
- ESRI. US election 2016. Battle of the maps. 2016
- Field, Kenneth & Dorling, Danny. UK election cartography. *International journal of cartography*. 2/2, 2016
- Field, Kenneth. Thematic map design
- GIS Geography. Choropleth maps: a guide to data classification
- Giulani, Diego & Napolitano, Maurizio. Online course: Geospatial analysis and representation for data science. University of Trento, 2020
- Graser, Anita. *Learning QGIS*. Packt, 2016 Library resource
- Graser, Anita & Peterson, Gretchen. *QGIS map design*. Locate Press, 2018
- Green, Chris. Mapping the Brexit vote. Blogpost. 2017
- House of Commons. MSOA names.
- Knight Center. Online course: Intro to mapping for journalists. 2018
- Lambrechts, Maarten. The essential lies in news maps. 2019.
- Manson, Stephen. Mapping, society & technology. University of Minnesota, 2017. <<https://open.lib.umn.edu/mapping/>> _
- Mapschooll. A free introduction to GEO
- Marcoux, Jacques. Geographic information systems, a use case for journalists. Datajournalism.com. 2019
- Mason, Betty. Election maps are everywhere. Don't let them fool you. NY Times. 2020
- McConchie, Alan. How we watched the [U.S.] election. 2020
- Monmonier, Mark. *How to lie with maps*. 3rd ed., University of Chicago Press, 2018 Library resource

Morphocode. The use of color in maps

NHS. Summary of geographical areas in Wales

ONS. A beginner's guide to UK geography. 2019

ONS. All geographic codes for the UK. 2022

Robinson, Arthur H. *The look of maps. An examination of cartographic design*. Esri press classics, 1952; 2010

Slocum, Terry et al. *Thematic cartography and geovisualisation*. 3rd ed., Prentice Hall, 2010

Tableau. Overview of maps in Tableau

Vermeulen, Maite, de Korte & van Houtum. How maps in the media make us more negative about migrants. De Correspondent. September 2020

NETWORK ANALYSIS

Tools:

- Flourish
- NodeXL (excel). Windows only
- Gephi
- Hyphe
- Network X (python)
- Neo4j

Aldhous, Peter. NodeXL for network analysis. 2014

Bloomberg. Who marries whom? 2016

Bounegru et al. Visual network exploration for data journalists. 2017

Bounegru et al. Networks as storytelling devices in journalism. 2016

Bradshaw, Paul. A journalist's introduction to network analysis. 2020

Bradshaw, Paul. Empathy as an investigative tool. 2020

Brown, Meta S. Why Panama Papers journalists use graph databases. 2016

Eastwood, Jones & Wolfe. Trump, his children and 500+ potential conflicts of interest. WSJ, 2017

ICIJ. Neo4j sandbox and ICIJ datasets.

Hill, Sophie. My little crony. 2021

Hunger, Michael & Lyon, William. Analyzing the Panama Papers with Neo4j: Data Models, Queries & More. 2016

Intal, Carla & Yasseri, Taha. Dissent and Rebellion in the House of Commons: a social network analysis of Brexit-related divisions. 2020

Lindenberg, Frederich. Here's why investigative reporters need to know knowledge graphs. GIJN. 2020

Marshall, Alex. Radiohead's corporate empire. 2016

Mohr, Jeff. Celebrity ice bucket challenge. 2021

Neo4j case studies

Romano, Tricia. In Seattle art world, women run the show. The Seattle Times. 2016

Wild, Johanna. Digital investigations for journalists: How to follow the digital trail of people and entities. Module 4: Network analysis. Knight Center. 2020

NUMBERS & STATISTICS

BBC. Editorial guidelines: statistics.

Bell, Andrew et al. *Making sense of data in the media*. SAGE, 2020. [Library resource](#)

Bergstrom, Carl T. & West, Jevin D. *Calling Bullsh-t. The art of scepticism in a data-driven world*. Allen Lane, 2020 [Library resource](#)

Best, Joel. *Damned lies and statistics*. University of California Press, 2012 (updated ed.) [Library resource](#); *More damned lies and statistics*, University of California Press, 2004 [Library resource](#)

Blastland, Michael & Dilnot, Andrew. *The Tiger that isn't. Seeing through a world of numbers*. Profile, 2008 [Library resource](#)

Blauw, Sanne. *The number bias*. Sceptre, 2020

Chivers, Tom & David Chivers. *How to read numbers. A guide to statistics in the news*. W&N, 2021

Cohen, Sarah. *Numbers in the Newsroom: Using Math and Statistics in News*, 2nd ed., IRE, 2014

Cohn, Victor & Cope, Lewis. *News and numbers*. Wiley-Blackwell, 3rd ed., 2012 [Library resource](#)

Cushion, Stephen, Lewis, Justin & Callaghan, Robert. Data Journalism, Impartiality And Statistical Claims. Towards more independent scrutiny in news reporting. *Journalism Practice*, 11/2017, p. 1198-1215

Davies, William. How statistics lost their power. *Guardian*, 2017

Goldacre, Ben. *Big Pharma. How drug companies mislead doctors and harm patients*. 4th Estate, 2012

Goldacre, Ben. *I think you'll find it's a bit more complicated than that*. 4th Estate, 2014

Gould, Stephen Jay. The median isn't the message. *Ceylon Medical Journal*. Vol. 49(4), 2004, p. 139-40

Hand, David J. *Statistics. A very short introduction*. OUP, 2008. [Library resource](#)

Hand, David J. *Dark Data. Why what you don't know matters*. Princeton, 2020. [Library resource](#)

Harford, Tim. *How to make the world add up*. Bridge Street Press, 2020

Harkness, Timandra. *Big data. Does size matter?* Bloomsbury, 2016. [Library resource](#)

Huff, Darrell. *How to lie with statistics*. Norton, 1954

Jerven, Morten. Preface. *Poor Numbers*. Cornell University Press, 2013, p. ix-xiv. [Library resource](#)

Jerven, Morten & Wilkinson, Kate. Hard data and soft statistics: A guide to critical reporting. *Datajournalism.com*, 2019

Kahneman, Daniel. *Thinking, Fast and Slow*. Penguin, 2011. [Library resource](#)

Kaplan, Daniel. *Statistical Modelling*. 2017

Lievesley, Denise. A guide to statistics for journalists. Reuters Institute, 2020. 1h Video

- Marshall, Sarah. How to correctly report numbers in the news. 2012
- Matthews, Robert. Storks deliver babies. Teaching statistics, 22/2, 2000
- Meyer, Philip. *Precision Journalism. A Reporter's Introduction to Social Science Methods*. Rowman & Littlefield, 4th ed., 2002. [Library resource](#)
- Nguyen, An (ed.). *News, numbers and public opinion in a data-driven world*. Bloomsbury, 2018. [Library resource](#)
- Paulos, John Allen. *A Mathematician Reads the Newspapers*. Penguin, 1995
- Paulos, John Allen. *Innumeracy. Mathematical illiteracy and its consequences*. Penguin, 1988
- Porter, Theodore. *Trust in numbers: the pursuit of objectivity in science and public life*. Princeton University Press, 1996. [Library resource](#)
- Reinhart, Alex. *Statistics Done Wrong. The woefully complete guide*. No Starch Press, 2015
- Reuben, Anthony. *Statistical. Ten easy ways to avoid being misled by numbers*. Constable, 2019
- Rosling, Hans. The joy of stats. 1h Video
- Rowntree, Derek. *Statistics without tears. An introduction for non-mathematicians*. Penguin, 2018
- Sense About Science. Making sense of statistics. What statistics tell you and how to ask the right questions. Sense About Science. 2010
- Silver, Nate. *The Signal and the Noise. The Art and Science of Prediction*. Penguin, 2012
- Spiegelhalter, David. *The Art of Statistics. Learning from Data*. Pelican, 2019
- Stray, Jonathan. Drawing conclusions from data. 2013.
- Stray, Jonathan. The curious journalist's guide to data. Tow Center, 2018
- Taylor, Paul. Insanely complicated. Hopelessly inadequate. London Review of Books. 43/2, January 2021
- Wheelan, Charles. *Naked statistics. Stripping the dread from the data*. Norton, 2013

Ahmad, Mohammed Idriss. Bellingcat and how open source reinvented journalism. *The New York Review of Books*. June 2019

Bazzell, Michael. *Open Source Intelligence Techniques*. Inteltechniques. 8th ed., 2021

Berkeley Law Human Rights Center. Berkeley Protocol on Digital Open Source Investigations. Berkeley & UN, 2020

Carrieri, Matt. Open-source evidence and the International Criminal Court. *Harvard Human Rights Journal*. April 2019.

Chauhan, Sudhanshu & Kumar Panda, Nutan. *Hacking Web Intelligence: open source intelligence and web reconnaissance concepts and techniques*. Syngress, April 2015. [Library resource](#)

Creta, Sara. Open source investigations handbook. Al Jazeera

Digital Investigation. Academic journal. [Library resource](#)

Dubberly, Sam, Koenig, Alexa & Murray, Daragh. *Digital Witness. Using Open Source Information for Human Rights Investigation, Documentation, and Accountability*. OUP, 2020

Economist. Open-source intelligence challenges state monopolies on information. August 2021.

Edwards, Maxim. Open-source journalism in a wired world. Nieman report, 2022

Hahn, Oliver & Stalph, Florian Hahn, Oliver, editor.; Stalph, Florian, editor. *Digital investigative journalism: data, visual analytics and innovative methodologies in international reporting*. Palgrave Macmillan, 2018. [Library resource](#)

Hassan, Nihad A & Hijazi, Rami. *Open Source Intelligence Methods and Tools: A Practical Guide to Online Intelligence* Apress, 2018. [Library resource](#)

Higgins, Eliot. *We are Bellingcat. An intelligence agency for the people*. Bloomsbury, 2021. [Library resource](#)

McDermott, Yvonne. Open Source Information's Blind Spot: Human and Machine Bias in International Criminal Investigations. *Journal of International Criminal Justice*, 2021

Pool, Hans. *Bellingcat. Truth in a post-truth world*. 2018. 60' documentary [Library resource](#) (Box of Broadcasts)

Silverman, Craig. *Verification Handbook 2*

Tett, Gillian, The keyboard warriors on Ukraine's digital front line, FT. Feb 2023

SCRAPING & APIS

— See also **Ethics** and **Legal**

Bradshaw, Paul. *Scraping for Journalists*. Leanpub, 2017

Bradshaw, Paul. What Data Journalists Need to Know About APIs. GIJN, 2022

Carter, Laura. There is always an element of judgement. Datakind UK

Google. Robots.txt files overview

Harlow, Max. Fetch and enrich data with APIs.

Heydt, Michael. *Python web scraping cookbook*. Packt, 2018 [Library resource](#)

Jarmul, Katherine & Lawson, Richard. *Python web scraping: fetching data from the web*. Packt, 2nd ed., 2017 [Library resource](#)

Kouzis-Loukas, Dimitrios. *Learning Scrapy: learn the art of efficient web scraping and crawling with Python*. Packt, 2016 [Library resource](#)

McCarthy, Kieran. Web scraping for me, but not for thee. 2023.

Mitchell, Ryan. *Web scraping with Python*. O'Reilly, 2nd ed., 2018 [Library resource](#)

Ni, Daniel. Five Tips for web scraping without getting booted. 2019

Schacht, Kira. A web scraping toolkit for journalists, 2019

Shiab, Nael, Web scraping. A journalist's guide, GIJN, 2015

Smith, Madolyn. APIs for journalism. Datajournalism.com

Sweigart, Al. "Web scraping", *Automate the boring stuff*, No starch press, 2nd ed, 2020, p. 267-300.

The Markup. Why web scraping is vital to democracy. 2020

STORYTELLING WITH DATA

- Bastien, Karen. Story angles from data. 2020. 1hr Video (in French)
- Blinderman, Ilia. How to make dope shit: Part 3 Storytelling. The Pudding
- Bradshaw, Paul. Empathy as an investigative tool. Blogpost, 2020
- Bradshaw, Paul. It's not all about numbers: six ways data can give you a story lead. 2020
- Bradshaw, Paul. Seven story structures. Blogpost, 2019
- Constantaras & Valeeva. A data journalist's guide to building a hypothesis. Datajournalism.com, 2021
- Flowers, Andrew. The six types of data journalism stories. 48' Video (from 10'). 2017
- Gottschall, Jonathan. *The storytelling animal*. First Mariner, 2013
- NY Times, Finding stories in data: tipsheets
- Paulos, John Allen. *Once upon a number: The hidden mathematical logic of stories*. Penguin, 1998
- Storr, Will. *The science of storytelling*. William Collins, 2019
- Thomas, Amber. How to go from idea, to data, to story. The Pudding. 2020
- Webster, Mary-Jo. Bringing the power of data to deadline stories. Datajournalism.com. 2020

VISUALISATION

Guides to what kind of graph to use:

- [Andy Kriebel's visual vocabulary](#)
- chartmaker.visualisingdata.com/
- data-to-viz.com/caveats.html
- datavizcatalogue.com/search.html
- datavizproject.com
- [Datawrapper: dos and don'ts series](#)
- [EEA: dos and don'ts of data viz](#)
- flowingdata.com/chart-types
- ft-interactive.github.io/visual-vocabulary/
- policyviz.com/2021/02/08/five-charts-youve-never-used-but-should/
- policyviz.com/2016/11/30/style-guides/
- style.ons.gov.uk/category/data-visualisation/

Anscombe, F.J. Graphs in Statistical Analysis. *The American Statistician*, 27/1, 1973, pp. 17-21

Blinderman, Ilia. How to make dope shit. Part 2: Design. *The Pudding*

Cairo, Alberto. *The functional art*. New Rider, 2013 Library resource

Cairo, Alberto. *The truthful art*. New Rider, 2016 Library resource

Cairo, Alberto. *Datasaurus*. 2016

Chalabi, Mona. *Instagram*

Dick, Murray. *The infographic: a history of data graphics in news and communications*. MIT Press, 2020. Library resource

Engelbrechtsen, Martin & Kennedy, Helen. *Data Visualisation in Society*. 2020

Financial Times. *Chart advice blogs*.

Frost, Adam et al. *Communicating with data visualisation: a practical guide*. Sage, 2022. Library resource

House of Commons. *Library guide to designing effective charts*

Kirk, Andy. *Data visualisation: a handbook for data driven design*. SAGE, 2019 (2nd ed.) Library resource

Knight Center. *Module 5. Visualising*. Data journalism and visualisation with free tools. Online course. 2019

Levontin, Polina & Walton, Jo Lindsay, *Visualising Uncertainty*. Sad Press, 2020

Matejka, Justin & Fitzmaurice, George. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing, CHI, 2017 [‘the Datasaurus dozen’]

Muth, Lisa Charlotte. How to choose more beautiful colors for your data visualisation. Blogpost, Datawrapper, 2020

Muth, Lisa Charlotte. In defense of simple charts. Blogpost, Datawrapper, 2021

Posavec, Stefanie & Lupi, Giorgia. *Dear Data*. Particular Books, 2016 & Library resource

Rue, Jeremy. Visualising data: a guide to chart types. Berkeley. 2019

Schwabisch, Jonathan. *Better data visualisations*. Columbia, 2021. Library resource

Smith, Alan. *How charts work*. Pearson, 2022. Library resource

Sunlight Foundation. Dataviz style guidelines

Tufte, Edward. *Beautiful Evidence*. Graphics Press, 2006

Tufte, Edward. *Envisioning information*. Graphics Press, 1990. Library resource

Tufte, Edward. *The visual display of quantitative information*. 2nd ed., Graphics Press, 2001

Tableau resources:

- [Reference gathering](#)
- [Tableau gallery](#)
- [Makeover Monday: data & gallery & forum & youtube](#) (live walkthroughs)
- [BBC Style guide for Tableau](#)

Wilkinson, Leland. *The grammar of graphics*. Springer, 1999 Library resource

Wong, Dona M. *The Wall Street Journal Guide to Information Graphics*. Norton, 2013. Library resource (2010 ed.)