

Essential Websites for Learning Digital Skills and Biotechnology

William Okech

2024-04-04

Table of contents

Welcome	11
Introduction	12
Why did I write this book?	12
About the author	12
Topics Covered	12
Feedback	13
Summary	13
I Programming	14
A curated collection of resources	15
1 R and RStudio	16
1.1 Personal Websites and Blogs	16
1.2 Basics	17
1.2.1 Curated list of resources	17
1.2.2 Books	17
1.2.3 R and GPT	18
1.2.4 Conference Workshops	18
1.2.5 Package Development	18
1.2.6 Newsletters	18
1.2.7 Training Courses and Tutorials	18
1.2.8 Tools	20
1.2.9 Troubleshooting / Tips and Tricks	20
1.2.10 Hackathons and Competitions	20
1.3 Tidyverse	20
1.3.1 Basics	20
1.3.2 Books	21
1.3.3 Workshops	21
1.3.4 Tips and Tricks	21
1.3.5 Tutorials	21
1.4 Tidymodels in R	21
1.4.1 Basics	21
1.4.2 Books	22

1.4.3	Conferences	22
1.4.4	Tutorials	22
1.5	R Shiny	22
1.5.1	Basics	22
1.5.2	Books	22
1.5.3	Tutorials	23
1.5.4	Tools	23
2	Python	24
2.1	Python	24
2.1.1	Curated collection of resources	24
2.1.2	Personal Websites and Blogs	24
2.1.3	Books	24
2.1.4	Basics	25
2.1.5	Tutorials	25
2.1.6	Applications	25
2.1.7	Projects	26
3	Quarto	27
3.1	Introduction and Tutorials	27
3.1.1	Curated collection of Quarto Resources	27
3.1.2	Beginner's Guide	27
3.1.3	R Markdown and Distill	27
4	Git, GitHub, Command Line, and SQL	28
4.1	Git	28
4.1.1	Basics and Summaries	28
4.1.2	Books	28
4.1.3	Tutorials	28
4.1.4	Git Troubleshooting	29
4.2	GitHub	29
4.2.1	Basics	29
4.2.2	Training and Tutorials	29
4.2.3	Applications	29
4.3	Command Line (Bash, UNIX, and Linux)	29
4.3.1	Books	29
4.3.2	Lectures and Tutorials	30
4.4	SQL	30
4.4.1	Collection of SQL Resources	30
4.4.2	Tutorials	30

II	Data Science, Machine Learning, and Artificial Intelligence	31
	A curated collection of resources	32
5	Data Science, Machine Learning, and Artificial Intelligence	33
5.1	Data Science	33
5.1.1	Courses	33
5.1.2	Books	33
5.1.3	Tutorials, Resources, and Projects	34
5.2	Machine Learning	34
5.2.1	Books and Readings	34
5.2.2	Tutorials, Resources, and Projects	35
5.3	Artificial Intelligence	36
5.3.1	Books and Readings	36
5.3.2	Tutorials and Resources	36
5.3.3	Tools	36
5.4	Statistics and Computer Science	37
5.4.1	Books and Readings	37
5.4.2	Tutorials and Resources	37
5.5	Blogs	37
5.6	Forecasting and Prediction	37
5.7	Reproducibility	38
5.8	Health	38
5.9	Journals	38
5.10	Newspaper Articles and Newsletters	38
6	Statistics	39
6.0.1	Books	39
6.0.2	Personal Websites and Blogs	39
6.0.3	Lectures and Courses	39
6.0.4	Tutorials	40
6.0.5	Troubleshooting	40
6.0.6	Biostatistics	40
7	Public Datasets	41
7.1	Working with Data	41
7.2	Free and Open Data	41
7.3	Machine Learning Training Data	42
7.4	Industry-Specific Data	42
8	Sports Data Science	43
8.1	Sports Data Science Blogs	43
8.1.1	Multiple Sports	43
8.1.2	Soccer	43

8.1.3	Basketball Analytics	44
8.2	Sports Data Science Stories	44
8.3	Sports Data Science Journals and Readings	44
8.4	Tools and Databases	44
8.4.1	Multiple Sports	44
8.4.2	Soccer	45
8.4.3	Tutorials	46
9	Social Media and Music Analysis	47
9.1	Social Media Analysis	47
9.1.1	Books	47
9.1.2	Tutorials	47
9.1.3	Courses	48
9.1.4	Projects	48
9.2	Music and Lyric Analysis	48
9.2.1	Basics	48
9.2.2	Tutorials	48
10	Data Science Interview	50
10.1	Job Descriptions	50
10.2	Curated collection of Data Science Resources	50
10.3	General Interview Questions	50
10.4	Language- and Field-Specific Interview Questions	51
III	Data Visualization and Storytelling	52
	A curated collection of resources	53
	Visualization	53
	Journalism and Storytelling	53
11	Data Visualization (General)	54
11.1	Data Visualization (Blogs)	54
11.1.1	People	54
11.1.2	Companies	54
11.2	Data Visualization (General)	54
11.2.1	Selecting the appropriate chart types	54
11.2.2	Selecting the appropriate colors	55
11.2.3	Data Viz Challenges	55
11.2.4	Books	55
12	Data Visualization (R)	56
12.1	Curated Links by other data viz specialists	56
12.2	Books	56

12.3	Templates	56
12.4	Courses	57
12.5	Tutorials and Notes	57
12.5.1	Basics	57
12.5.2	Troubleshooting	57
12.5.3	Enhancing Plots (with Labels, Axes, Themes, Scales, and Other Options)	58
12.5.4	Various graph types	58
12.5.5	Data Viz Examples	59
12.6	Data aRtistry	59
13	Data Visualization (Python)	60
13.1	Books	60
13.2	Tutorials	60
14	Data Visualization (Tableau)	61
14.1	Data Visualization (Tableau)	61
15	Data Journalism Tools and Stories	62
15.1	Data Journalism	62
15.1.1	Blogs and Websites	62
15.1.2	Guides	62
15.1.3	Writing	62
15.1.4	Tools	63
15.1.5	Tutorials	63
15.1.6	Organizations	63
15.2	Data Stories (Longform and In-depth articles)	63
15.2.1	Story aggregators	63
15.2.2	Top Data Story Examples	64
IV	Generative AI and Large Language Models	65
16	Generative AI and Large Language Models (Part 1)	67
V	Teaching Resources	68
	A curated collection of resources	69
17	General Teaching Resources	70
17.1	Teaching STEM Subjects	70
17.1.1	A curated collection of resources	70
17.2	Teaching Aids	70
17.2.1	Typing Test	70
17.2.2	Simulations	70

17.2.3	Online Textbooks	70
17.2.4	Design and Illustrations	71
17.2.5	Miscellaneous Tools	71
18	RStudio and Carpentries Certification	72
18.1	RStudio Certification	72
18.1.1	Tutorials and Resources	72
18.1.2	Study notes from RStudio Certified Trainers	72
18.2	Carpentries Certification	73
19	Medical, Science, and Technical Writing	74
19.1	Grammar Basics	74
19.2	Medical Writing	74
19.3	Science Writing	74
19.4	Technical Writing	75
VI	Bio-X: All things Bio	76
20	Bioinformatics and Computational Biology	78
20.1	Personal and Lab Websites	78
20.2	Bioinformatics News Stories	78
20.3	Bioinformatics Basics	79
20.3.1	Courses and Lectures	79
20.3.2	Books and Resources	79
20.3.3	Tools	80
20.4	Bioinformatics Tutorials and Projects	81
20.4.1	Tutorials	81
20.4.2	Projects	82
21	Biomechanics and Sports Engineering	83
21.1	Biomechanics	83
21.1.1	Curated collection of resources	83
21.1.2	Personal Blogs and Websites	83
21.1.3	Labs	83
21.1.4	Books	83
21.1.5	Tutorials and Lectures	83
21.1.6	Sports Engineering and Movement Science	84
22	Biotechnology Startups	85
22.1	Resources for Biotechnology Startups	85

23 Open-Source Bio	86
23.1 Science, Medicine, and Technology	86
23.1.1 Personal Blogs and Websites	86
23.1.2 Tools	86
23.1.3 Labs and Groups	86
23.1.4 Projects	86
23.1.5 Publications	87
23.1.6 mHealth	87
24 Biomedical Image Analysis and Microscopy	88
24.1 E-Books and Tutorials	88
24.2 BioImaging Software	89
24.3 Microscopy	89
VII Recommended Readings	90
STEM students	91
For Everyone	91
25 Recommended Readings (STEM)	92
25.1 Best Practices	92
25.2 Challenges	92
25.3 Transformative Research Papers in Biomedical Science	93
25.4 Statistics	93
25.5 Writing	93
25.6 Science Integrity	94
26 Recommended Readings (General)	95
26.1 Interesting Articles and Speeches	95
26.2 My Favorite Short Stories	95
26.3 Classics	95
26.4 Mental Health / Counseling / Peace of Mind	95
26.5 Religious	96

List of Figures

List of Tables

Welcome

This is the website for the book “Essential Websites for Learning Digital Skills and Biotechnology” written by [William Okech](#). This book is primarily intended to provide links to essential websites for learning digital skills and biotechnology.

[Essential Websites for Learning Digital Skills and Biotechnology](#) © 2024 by [William Okech](#) is licensed under [CC BY-NC-ND 4.0](#). The online version of this book is free to use.

Cover image was designed using [Canva](#).

This book was built with [Quarto](#).

Introduction

Why did I write this book?

This book is primarily intended to provide links to essential websites for learning digital skills and biotechnology. It is hoped that the websites provided will reduce the time spent by readers searching for relevant information in the fields of programming, data science, data visualization, and biotechnology. Additionally, I have provided some bonus readings for STEM graduate students and the general public.

About the author

The author of this book is a Certified Carpentries instructor and a trainer with the Digital Research Academy. Additionally, the author holds a PhD in Biomedical Engineering and has completed postdoctoral fellowships in vascular biology and infectious diseases. Lastly, the author is passionate about using R and RStudio to generate data-driven visualizations to allow for a more in-depth understanding of public policy issues.

Topics Covered

- i. Programming
- ii. Data Science, Machine Learning, and Artificial Intelligence
- iii. Data Visualization
- iv. Generative AI and Large Language Models
- v. General Teaching Resources
- vi. Biotechnology
- vii. Recommended Readings

Feedback

Feedback can be provided using a GitHub pull request: [Link](#)

Summary

Overall, I believe that this book will increase both the knowledge and confidence levels of novice programmers and allow them to perform basic statistical analysis and simplify everyday computational tasks at home or in their workplaces. In the next chapter, I will provide a basic overview of the R and RStudio ecosystem.

Part I

Programming

This page provides a comprehensive collection of resources on programming tools with a specific focus on R/RStudio (data analysis and visualization), Python, Git/GitHub (version control), Command Line, SQL, and Quarto (scientific and technical publishing).

A curated collection of resources

0.0.0.0.1 * 1. [R and RStudio](#)

0.0.0.0.2 * 2. [Python](#)

0.0.0.0.3 * 3. [Git, GitHub, Command Line, and SQL](#)

0.0.0.0.4 * 4. [Quarto](#)

1 R and RStudio

This page provides a comprehensive collection of R/RStudio resources. On this page, you will find a list of personal websites and blogs that do some great work in R/RStudio. Additionally, there is a list of resources (books, tutorials, and blogs) focused on the basics of R/RStudio and Shiny.

1.1 Personal Websites and Blogs

- 1) [Silvia Canelon](#)
- 2) [Meghan Hall](#)
- 3) [Paul Villanueva](#)
- 4) [Shannon Pileggi](#)
- 5) [Katherine \(Kat\) Hoffman](#)
- 6) [Deisy Gysi](#)
- 7) [Simon J](#)
- 8) [Ryo Nakagawara](#)
- 9) [Kieran Healy - Some Favorite Things](#)
- 10) [François Husson](#)
- 11) [Rebecca Barter](#)
- 12) [Alboukadel Kassambara](#)
- 13) [Michael Toth](#)
- 14) [Sam Abbott](#)
- 15) [Shel Kariuki](#)
- 16) [Austin Wehrwein](#)
- 17) [Nicola Rennie](#)
- 18) [R for Ecology](#)
- 19) [Stats and R](#)
- 20) [Roland Krasser](#)
- 21) [Martin Chan - Musings on R](#)
- 22) [Rsquared Academy](#)
- 23) [Olivier Gimenez](#)

More to come...

1.2 Basics

1.2.1 Curated list of resources

- 1) [Big Book of R](#)
- 2) [FREE R READING MATERIAL](#)
- 3) [An R Community Public Library](#)
- 4) [Awesome R Learning Resources](#)
- 5) [R Cheatsheets](#)
- 6) [Posit Cheatsheets](#)
- 7) [R4theRestofUs](#)
- 8) [EDUCATORS R LEARNERS](#)
- 9) [R Resources](#)
- 10) [R and Data Mining](#)

1.2.2 Books

- 1) [R4DataScience](#)
- 2) [R4DataScience - Solutions](#)
- 3) [R Workflow A](#)
- 4) [R Workflow B](#)
- 5) [Urban Informatics](#)
- 6) [Modern Statistics with R](#)
- 7) [R for Researchers: An Introduction](#)
- 8) [Supervised Machine Learning for Text Analysis in R](#)
- 9) [Reproducible Analytical Pipelines - Master's of Data Science](#)
- 10) [R for Data Analysis](#)
- 11) [Data Analysis and Prediction Algorithms with R](#)
- 12) [Hands-On Programming with R](#)
- 13) [R for Health Data Science](#)
- 14) [R for Health Data Science Resources](#)
- 15) [Data Science in Education Using R](#)
- 16) [A Succinct Intro to R](#)
- 17) [Data Science: A First Introduction](#)
- 18) [The Epidemiologist R Handbook](#)
- 19) [Introduction to R, version 2](#)
- 20) [Technical Foundations of Informatics](#)
- 21) [Data Science for Economists and Other Animals](#)
- 22) [Efficient R programming](#)
- 23) [An Introduction to R](#)
- 24) [YaRrr! The Pirate's Guide to R](#)
- 25) [Modern Data Science with R](#)

- 26) [Twitter for R programmers](#)
- 27) [Modern Dive](#)
- 28) [Advanced R](#)
- 29) [Yet Again: R + Data Science](#)
- 30) [Public Policy Analytics](#)
- 31) [Hands-On Machine Learning with R](#)

1.2.3 R and GPT

- 1) [R and gpttools](#)

1.2.4 Conference Workshops

- 1) [Talk recordings and workshop materials from rstudio::conf\(2022\)](#)
- 2) [Links to slides from rstudio::conf 2020](#)
- 3) [NHS-R](#)
- 4) [useR! 2019 Highlights](#)
- 5) [ToulouseR 2019](#)

1.2.5 Package Development

- 1) [R Packages](#)
- 2) [Introduction to Package Development](#)
- 3) [Your first R package in 1 hour](#)
- 4) [Developing R packages](#)
- 5) [Reflections on building my first few R packages](#)

1.2.6 Newsletters

- 1) [RWeekly](#)

1.2.7 Training Courses and Tutorials

- 1) [RStudio Education](#)
- 2) [RStudio Education Beginners](#)
- 3) [RStudio Education Intermediates](#)
- 4) [RStudio Education Experts](#)
- 5) [Quick-R Tutorial](#)
- 6) [Learn R in R](#)

- 7) [Learning R With Education Datasets](#)
- 8) [STAT 545 Data wrangling, exploration, and analysis with R](#)
- 9) [STHDA](#)
- 10) [datascienceplus](#)
- 11) [Statistical Computing in R](#)
- 12) [Statistical Programming Paradigms and Workflows](#)
- 13) [Teaching Statistics and Data Science Online](#)
- 14) [An Introduction to Docker for R Users](#)
- 15) [R Bootcamp](#)
- 16) [Kane's Data Science Course](#)
- 17) [Andrew Heiss R-based courses](#)
- 18) [STREAMLINING WITH R](#)
- 19) [R Bootcamp](#)
- 20) [Advanced R Topics](#)
- 21) [R for Water Resources Data Science](#)
- 22) [Reproducible science workshop](#)
- 23) [Think R](#)
- 24) [R 4 Beginners](#)
- 25) [Applied Economics with R](#)
- 26) [Ready for R](#)
- 27) [Interpretable Machine Learning](#)
- 28) [R basics: a practical introduction to R](#)
- 29) [Introduction to R](#)
- 30) [fasterR: Fast Lane to Learning R!](#)
- 31) [Analysis of community ecology data in R](#)
- 32) [R Screencasts](#)
- 33) [R for better science in less time](#)
- 34) [R and Stats](#)
- 35) [Palmer Archipelago Penguins Data](#)
- 36) [We are R-Ladies](#)
- 37) [Web Scraping in R: A Quick Guide](#)
- 38) [Word Clouds in R](#)
- 39) [Data science tutorials](#)
- 40) [MARKET ANALYSIS COURSE](#)
- 41) [RTutor: Interactive R Problem Sets](#)
- 42) [Data Wrangling and dplyr](#)
- 43) [Intro to R for Programming](#)
- 44) [purrr tutorial](#)
- 45) [Intro to R: Shiny App](#)
- 46) [Spatial Data Science](#)
- 47) [Generalized Additive Models in R](#)
- 48) [Visualizing the Bayesian workflow in R](#)
- 49) [Create machine learning models: An R version](#)

- 50) [Wrangling Unruly Data](#)
- 51) [Cleaning UK Office for National Statistics data in R](#)
- 52) [How to Automate EDA with DataExplorer in R](#)
- 53) [Rachael's R Tutorials](#)
- 54) [Manipulating Data with dplyr](#)
- 55) [Get me to the church on time with R spatial](#)
- 56) [Multivariate data analysis with R and vegan](#)

1.2.8 Tools

- 1) [R Universe Authors and Packages](#)
- 2) [R-hub](#)
- 3) [R in Visual Studio Code](#)
- 4) [R Libraries Every Data Scientist Should Know](#)
- 5) [Packages for Exploratory Data Analysis in R](#)
- 6) [15 Essential packages in R for Data Science](#)

1.2.9 Troubleshooting / Tips and Tricks

- 1) [How to Use “not in” operator in Filter](#)
- 2) [Common R Mistakes in Data Viz](#)
- 3) [R Examples](#)
- 4) [R Tips and Tricks](#)

1.2.10 Hackathons and Competitions

- 1) [Kaggle Housing Price Tutorial](#)

1.3 Tidyverse

1.3.1 Basics

- 1) [A very short introduction to Tidyverse](#)
- 2) [The Tidyverse in Action](#)
- 3) [A Gentle Guide to Tidy Statistics in R](#)
- 4) [An Introduction to R through the tidyverse](#)
- 5) [C'est quoi, le tidyverse?](#)
- 6) [Basic usage of tidycensus](#)

1.3.2 Books

- 1) [Tidyverse Skills for Data Science](#)
- 2) [Modern R with the tidyverse](#)
- 3) [Text Mining with R: A Tidy Approach](#)

1.3.3 Workshops

- 1) [Tidyverse Workshop Series](#)

1.3.4 Tips and Tricks

- 1) [Eight R Tidyverse tips for everyday data engineering](#)
- 2) [Tidyverse Tips](#)

1.3.5 Tutorials

- 1) [Data Science LADAL Tutorials](#)
- 2) [Pivoting data from columns to rows \(and back!\) in the tidyverse](#)
- 3) [Reshaping data frames using pivot functions from {tidyr} and tally from {dplyr}](#)
- 4) [Finding the modal school district](#)
- 5) [Teaching the tidyverse in 2021](#)
- 6) [Tidy analysis of cocktails - Part I – EDA](#)
- 7) [Tidy analysis of cocktails - Part II - Exploratory k-means clustering](#)
- 8) [Transitioning into the tidyverse \(part 1\)](#)
- 9) [Transitioning into the tidyverse \(part 2\)](#)
- 10) [Introduction to using Databases in R, with Tidyverse tools](#)
- 11) [Wrangling penguins: some basic data wrangling in R with dplyr](#)
- 12) [Data manipulation with the tidyverse](#)
- 13) [The TidyVerse in Action](#)

1.4 Tidymodels in R

1.4.1 Basics

- 1) [A Gentle Introduction to tidymodels](#)
- 2) [Tidymodels](#)
- 3) [The Case for tidymodels](#)
- 4) [Caret vs. tidymodels - comparing the old and new](#)

- 5) [Tidymodels: tidy machine learning in R](#)

1.4.2 Books

- 1) [Tidy Modeling with R](#)
- 2) [ISLR tidymodels labs](#)
- 3) [Modeling Data in the Tidyverse](#)
- 4) [Feature Engineering](#)

1.4.3 Conferences

- 1) [Introduction to Machine Learning with the Tidyverse](#)
- 2) [RStudio Conf Workshop Materials](#)
- 3) [Tidymodels package for survival models](#)

1.4.4 Tutorials

- 1) [Sample Story: When Will NYC's Subway Ridership Recover?](#)
- 2) [#TidyTuesday and tidymodels](#)
- 3) [tidypredict](#)
- 4) [Feature Engineering and Selection: A Practical Approach for Predictive Models](#)
- 5) [Experimenting with machine learning in R with tidymodels and the Kaggle titanic dataset](#)

1.5 R Shiny

1.5.1 Basics

- 1) [a gradual introduction to Shiny](#)
- 2) [Shiny Learning Resources](#)
- 3) [How to Start a Career as an R Shiny Developer](#)
- 4) [Shiny Apps development](#)
- 5) [The World's Most Advanced R Shiny Dashboards](#)

1.5.2 Books

- 1) [Mastering Shiny](#)
- 2) [Engineering Production-Grade Shiny Apps](#)
- 3) [Outstanding User Interfaces with Shiny](#)

1.5.3 Tutorials

- 1) [Shiny Tutorials](#)
- 2) [Building Web Applications WITH SHINY](#)
- 3) [Building Shiny apps - an interactive tutorial](#)
- 4) [Building Interactive World Maps in Shiny](#)

1.5.4 Tools

- 1) [HTML widgets flexdashboard](#)
- 2) [htmlwidgets for R](#)
- 3) [RintErface: HTML templates for Shiny](#)
- 4) [Loading screens for Shiny: waiter](#)

2 Python

This page provides a comprehensive collection of resources for those interested in Python. On this page, you will find a list of personal websites and blogs that do some great work in Python. Additionally, there is a list of resources (books, tutorials, and blogs) focused on the basics of Python.

2.1 Python

2.1.1 Curated collection of resources

- 1) [Python for Non-Programmers](#)
- 2) [49 Essential Resources To Learn Python](#)
- 3) [A Gallery of Interesting Jupyter Notebooks](#)
- 4) [A dozen ways to learn Python](#)

2.1.2 Personal Websites and Blogs

2.1.3 Books

- 1) [Automate the Boring Stuff with Python](#)
- 2) [Elements of Data Science](#)
- 3) [Scientific Computing with Python](#)
- 4) [Computational and Inferential Thinking: The Foundations of Data Science](#)
- 5) [Python for Data Analysis, 3E](#)
- 6) [Python Programming And Numerical Methods: A Guide For Engineers And Scientists](#)
- 7) [100 Page Python Intro](#)
- 8) [Think Python 2e](#)
- 9) [A Whirlwind Tour of Python](#)
- 10) [Google's Python Class](#)
- 11) [Python Data Science Handbook](#)

2.1.4 Basics

- 1) [Numpy](#)
- 2) [Python Cheatsheet](#)
- 3) [Comprehensive Python Cheatsheet](#)
- 4) [Python Curriculum Map](#)
- 5) [Data School: My top 25 pandas tricks](#)
- 6) [10 Python Pandas tricks that make your work more efficient](#)
- 7) [Python built-in functions to know](#)
- 8) [Data Science Best Practices with pandas](#)
- 9) [Tour of Python Itertools](#)

2.1.5 Tutorials

- 1) [Real Python](#)
- 2) [Resolving Python](#)
- 3) [Just into Data](#)
- 4) [Christina Levensgood Blog](#)
- 5) [Python Machine Learning Notebooks \(Tutorial style\)](#)
- 6) [Dataquest Tutorials](#)
- 7) [VIB Tutorials](#)
- 8) [Data Flair Training](#)
- 9) [Top 10 IPython Notebook Tutorials for Data Science and Machine Learning](#)
- 10) [Python for Beginners Course](#)
- 11) [Python Crash Course](#)
- 12) [Python for Everybody \(PY4E\)](#)
- 13) [Python Machine Learning Jupyter Notebooks](#)
- 14) [Practical Business Python](#)
- 15) [Learning Python For Data Science](#)
- 16) [Learn Python in 3 days: Step-by-Step Guide](#)
- 17) [A Complete 26 Week Course to Learn Python for Data Science in 2022](#)
- 18) [Python for Data Science: 8 Concepts You May Have Forgotten](#)
- 19) [What is Exploratory Spatial Data Analysis \(ESDA\)?](#)
- 20) [Getting started with data science using Python](#)

2.1.6 Applications

- 1) [Stock Portfolio Analysis using Python](#)
- 2) [Neural networks fundamentals with Python](#)
- 3) [Python News Webscraper](#)
- 4) [Web Scraping with Scrapy](#)

- 5) [Introduction to Market Basket Analysis in Python](#)

2.1.7 Projects

- 1) [Python Projects for 2023 – Work on Real-time Projects to Head Start Your Career](#)
- 2) [Python Projects with Source Code](#)
- 3) [45+ Data Analysis Projects with Python](#)

3 Quarto

The Quarto page contains a list of resources that are focused on the relatively new Quarto software. It includes various beginner guides to help one get started on their Quarto journey. Additionally, I provide links to resources for other tools such as RMarkdown and Distill.

3.1 Introduction and Tutorials

3.1.1 Curated collection of Quarto Resources

- 1) [Awesome Quarto](#)

3.1.2 Beginner's Guide

- 1) [Quarto](#)
- 2) [Tutorial: Hello, Quarto](#)
- 3) [What is Quarto? RStudio rolls out next-generation R Markdown](#)
- 4) [We don't talk about Quarto](#)
- 5) [Build a simple website with R and Quarto in under 10 minutes](#)
- 6) [Getting Started with Quarto](#)
- 7) [The ultimate guide to starting a Quarto blog](#)
- 8) [Creating a blog with Quarto in 10 steps](#)
- 9) [Making Slides in Quarto with reveal.js](#)

3.1.3 R Markdown and Distill

- 1) [The Distillery](#)
- 2) [R Markdown for Web Pages](#)
- 3) [Porting a distill blog to quarto](#)
- 4) [RMarkdown](#)

4 Git, GitHub, Command Line, and SQL

This page provides a comprehensive collection of resources focused on Git, GitHub, the Command Line, and SQL. On this page, you will find tutorials, books, and training materials that can get you started on the tools listed above.

4.1 Git

4.1.1 Basics and Summaries

- 1) [Git Basics](#)
- 2) [Git Cheatsheet](#)
- 3) [Git/GitHub Cheatsheet](#)
- 4) [A Visual Git Reference](#)
- 5) [Visualizing Git Concepts with D3](#)
- 6) [Git Immersion](#)
- 7) [Learn Git Branching](#)
- 8) [Introduction to Git & Github](#)
- 9) [How to Use Git/GitHub with R](#)
- 10) [git ready](#)

4.1.2 Books

- 1) [Pro Git Book](#)
- 2) [Happy Git and GitHub for the user](#)

4.1.3 Tutorials

- 1) [Understanding Git Conceptually](#)
- 2) [Become a git guru](#)
- 3) [Git and RStudio](#)
- 4) [Good Enough Practices in Scientific Computing](#)
- 5) [Version Control \(Git\)](#)

4.1.4 Git Troubleshooting

- 1) [On undoing, fixing, or removing commits in git](#)
- 2) [git pretty](#)
- 3) [Fixing Common Git Mistakes](#)

4.2 GitHub

4.2.1 Basics

- 1) [Getting started with Git and GitHub: the complete beginner's guide](#)
- 2) [An Intro to Git and GitHub for Beginners \(Tutorial\)](#)
- 3) [Get started with Git and GitHub](#)
- 4) [Mastering GitHub in 4 Steps](#)
- 5) [GitHub for supporting, reusing, contributing, and failing safely](#)

4.2.2 Training and Tutorials

- 1) [Introducing GitHub Global Campus](#)
- 2) [Collaboration, version control and learning to commit](#)
- 3) [Making Your First GitHub R Project](#)

4.2.3 Applications

- 1) [Introduction to GitHub Actions to R users](#)
- 2) [How to effortlessly create a website for free with GitHub Pages](#)
- 3) [Introducing Jekyll Now for GitHub-hosted blogs](#)
- 4) [Software development practices](#)

4.3 Command Line (Bash, UNIX, and Linux)

4.3.1 Books

- 1) [Command Line Basics for R Users](#)

4.3.2 Lectures and Tutorials

- 1) [Intro to UNIX](#)
- 2) [The Unix Shell 1](#)
- 3) [The Unix Shell 2](#)

4.4 SQL

4.4.1 Collection of SQL Resources

- 1) [21 of the best free resources to learn SQL](#)

4.4.2 Tutorials

- 1) [PostgreSQL Tutorial](#)
- 2) [The SQL Tutorial for Data Analysis](#)
- 3) [SQL Hackerrank Channels](#)
- 4) [SQL Tutorial 1](#)
- 5) [SQL Tutorial 2](#)
- 6) [SQLBolt](#)
- 7) [SQL queries don't start with SELECT](#)
- 8) [A Visual Explanation of SQL Joins](#)

Part II

Data Science, Machine Learning, and Artificial Intelligence

The data science, machine learning, and artificial intelligence page provides a varied list of general resources. Additionally, there is also a focus on datasets, statistics, interview prep, and specific topics such as sports and social media.

A curated collection of resources

4.4.0.0.1 * 1. [Data Science, Machine Learning, and Artificial Intelligence](#)

4.4.0.0.2 * 2. [Datasets for Statistical Analysis and Machine Learning](#)

4.4.0.0.3 * 3. [Statistics](#)

4.4.0.0.4 * 4. [Data Science Interview Prep](#)

4.4.0.0.5 * 5. [Sports Data Science](#)

4.4.0.0.6 * 6. [Social Media and Music Analysis](#)

5 Data Science, Machine Learning, and Artificial Intelligence

This page provides a comprehensive collection of resources on data science, machine learning, and artificial intelligence. It includes links to courses, books, and tutorials.

5.1 Data Science

5.1.1 Courses

- 1) [Data 8: The Foundations of Data Science](#)
- 2) [Introduction to Data Science](#)
- 3) [My Data Science Master's](#)
- 4) [Data Science Master's](#)
- 5) [School of Data](#)
- 6) [Data Science Courses and Resources](#)
- 7) [How To Learn Data Science If You're Broke](#)
- 8) [Computing for Information Science](#)
- 9) [Data Science Learning Path For Complete Beginners](#)
- 10) [Data Science Berens Lab](#)

5.1.2 Books

- 1) [54 Best Data Science Books in 2023](#)
- 2) [Critical Dataset Studies Reading List](#)
- 3) [Elements of Data Science](#)
- 4) [The Art of Data Science](#)
- 5) [Readings in Applied Data Science](#)
- 6) [AI/ML/DS Book Club](#)
- 7) [Data Science Resources](#)
- 8) [Data Science Learning Resources](#)
- 9) [Data Science Live Book](#)

5.1.3 Tutorials, Resources, and Projects

- 1) [List of Data Science Blogs](#)
- 2) [Data Science Primer](#)
- 3) [Data Science Cheatsheet](#)
- 4) [30 Essential Data Science, Machine Learning & Deep Learning Cheat Sheets](#)
- 5) [Coding habits for data scientists](#)
- 6) [Data Science by Design \(DSxD\)](#)
- 7) [The Data Scientist](#)
- 8) [The 10 Data Mining Techniques Data Scientists Need for Their Toolbox](#)
- 9) [Data Science Tutorials](#)
- 10) [Our Coding Club](#)
- 11) [Data Science - Resources for Python & R](#)
- 12) [Preparing for the Transition to Data Science](#)
- 13) [Free Data Science Resources for Beginners](#)
- 14) [Data Science Glossary on Kaggle](#)
- 15) [HOW TO START YOUR FIRST DATA SCIENCE PROJECT](#)
- 16) [DS/ML/AI Blog](#)
- 17) [mljar: Outstanding Data Science Tools](#)

5.2 Machine Learning

5.2.1 Books and Readings

- 1) [Machine Learning University](#)
- 2) [Yellowbrick: Machine Learning Visualization](#)
- 3) [Visual Intro to ML](#)
- 4) [Machine Learning Basics 1](#)
- 5) [Machine Learning Basics 2](#)
- 6) [End-to-End Machine Learning Library](#)
- 7) [Hands-On Machine Learning with R](#)
- 8) [FOUNDATIONS OF MACHINE LEARNING](#)
- 9) [ML for Everyone](#)
- 10) [Machine Learning Notebooks](#)
- 11) [Interpretable Machine Learning](#)
- 12) [Machine Learning for Humans](#)
- 13) [Machine Learning Explained: Algorithms Are Your Friend](#)
- 14) [A Guide to Learning with Limited Labeled Data](#)
- 15) [Kaggle Winner's Blog](#)
- 16) [Machine Learning and Data Science Applications in Industry](#)
- 17) [The Illustrated Machine Learning website](#)

- 18) [Machine Learning Mastery](#)
- 19) [ML for Science](#)
- 20) [In-depth introduction to machine learning in 15 hours of expert videos](#)
- 21) [Python ML Tutorial](#)
- 22) [Visualize Machine Learning data](#)
- 23) [Made With ML](#)
- 24) [Amazon Machine Learning](#)
- 25) [Azure Machine Learning](#)
- 26) [Google Prediction API](#)
- 27) [BigML](#)

5.2.2 Tutorials, Resources, and Projects

- 1) [How to choose a good evaluation metric for your Machine learning model](#)
- 2) [ML Data Repository](#)
- 3) [The Best Machine Learning Resources](#)
- 4) [ML Google](#)
- 5) [Santander ML Explainability](#)
- 6) [Best GitHub Repositories for ML](#)
- 7) [101 machine learning algorithms for data science with cheatsheets](#)
- 8) [Commonly used Machine Learning Algorithms](#)
- 9) [modelStudio and The Grammar of Interactive Explanatory Model Analysis](#)
- 10) [Supervised ML in R](#)
- 11) [MACHINE LEARNING QUESTIONS](#)
- 12) [ML101 Tutorial](#)
- 13) [ML Tutorial](#)
- 14) [Scikit learn in Machine Learning](#)
- 15) [Machine Learning Tutorials](#)
- 16) [The Ultimate FREE Machine Learning Study Plan](#)
- 17) [100 Days of Machine Learning Coding](#)
- 18) [150 of the Best Machine Learning, NLP, and Python Tutorials](#)
- 19) [Code for Workshop: Introduction to Machine Learning with R](#)
- 20) [Machine Learning Crash Course with TensorFlow APIs](#)
- 21) [Intro to ML with R Slides](#)
- 22) [A Machine Learning Course with Python](#)
- 23) [Mathematics for Machine Learning](#)
- 24) [Matrix Methods In Data Analysis, Signal Processing, And Machine Learning](#)
- 25) [Machine Learning Zoomcamp](#)
- 26) [Feature Selection Techniques in Machine Learning with Python](#)
- 27) [Machine Learning Projects](#)
- 28) [Randy Olson's data analysis and machine learning projects](#)
- 29) [Class Imbalance](#)

- 30) [Anomaly Detection for Dummies](#)

5.3 Artificial Intelligence

5.3.1 Books and Readings

- 1) [Applied AI](#)
- 2) [Microsoft Artificial Intelligence for Beginners](#)
- 3) [AI Reading List](#)
- 4) [Elements of AI](#)
- 5) [Introduction to Deep Learning](#)
- 6) [12 Steps to Applied AI](#)
- 7) [AI Experiments](#)
- 8) [Deep Learning](#)
- 9) [Neural Networks and Deep Learning](#)
- 10) [Artificial Intelligence in Future and Present](#)
- 11) [Stanford AI Lab](#)
- 12) [AI Research Experiences](#)
- 13) [Deep Learning Fundamentals](#)
- 14) [TensorFlow, Keras and deep learning, without a PhD](#)
- 15) [Convolutional Neural Networks for Visual Recognition](#)
- 16) [Introduction to Natural Language Processing](#)
- 17) [Solving 90% of NLP problems](#)

5.3.2 Tutorials and Resources

- 1) [How to not overfit?](#)
- 2) [HOW TO USE DEEP LEARNING TO IDENTIFY WILDLIFE](#)
- 3) [AI Summer](#)

5.3.3 Tools

- 1) [Face2Face](#)
- 2) [Voice Cloning](#)

5.4 Statistics and Computer Science

5.4.1 Books and Readings

- 1) [From Zero to Research Scientist full resources guide](#)
- 2) [Intro to Stat Learning 1](#)
- 3) [Intro to Stat Learning 2](#)
- 4) [Introduction to Causal Inference](#)
- 5) [Visual Guides for Causal Inference](#)
- 6) [A Gentle Introduction to Graph Neural Networks](#)
- 7) [Turi Create - development of custom machine learning models](#)

5.4.2 Tutorials and Resources

- 1) [Principal component analysis](#)
- 2) [Understanding UMAP](#)
- 3) [t-Distributed Stochastic Neighbor Embedding \(t-SNE\)](#)

5.5 Blogs

- 1) [Sebastian Raschka](#)
- 2) [Aurimas Racas](#)
- 3) [The Analytics Lab](#)
- 4) [Shirin Elsinghorst](#)
- 5) [Christina Levengood](#)
- 6) [Ethan Rosenthal](#)
- 7) [Rachel Thomas](#)
- 8) [LIIP](#)
- 9) [Data Science Dojo](#)

5.6 Forecasting and Prediction

- 1) [Forecasting: Principles and Practice](#)
- 2) [Ad Click Prediction](#)
- 3) [Facebook Prophet](#)
- 4) [Forecasting with Neural Network Autoregression](#)
- 5) [Neural Network Playground](#)
- 6) [fast.ai—Making neural nets uncool again](#)
- 7) [Predictive Hacks](#)

5.7 Reproducibility

- 1) [The Reproducibility Crisis in ML-based Science](#)
- 2) [Leakage and the Reproducibility Crisis in ML-based Science](#)
- 3) [Data leakage causes reproducibility failures in ML-based science](#)

5.8 Health

- 1) [Covid-19 Data Science](#)
- 2) [Machine Learning For Healthcare](#)
- 3) [Population health analytics](#)
- 4) [Jameel Clinic](#)
- 5) [Medical AI Bootcamp](#)
- 6) [Shah Lab AI](#)
- 7) [Build a Recommender Engine for Medical Research Papers](#)
- 8) [RECONLearn](#)
- 9) [RECON](#)

5.9 Journals

- 1) [TOP PUB](#)
- 2) [Harvard Data Science Review](#)
- 3) [Machine Learning in Public Policy](#)

5.10 Newspaper Articles and Newsletters

- 1) [The Facebook Scandal - or how to predict psychological traits from Facebook likes.](#)
- 2) [How to solve a business problem using data](#)
- 3) [Health Data Science](#)
- 4) [13 Essential Newsletters for Data Scientists: Remastered](#)
- 5) [The economics of artificial intelligence](#)

6 Statistics

This page provides a comprehensive collection of resources on statistics. Included here is a list of links to books, lectures and courses, and tutorials focused on statistics.

6.0.1 Books

- 1) [Introduction to Modern Statistics](#)
- 2) [Causal Inference: What If \(Hernan MA\)](#)
- 3) [Bayes Rules!](#)
- 4) [Causal Inference: The Mixtape](#)
- 5) [Seeing Theory](#)
- 6) [Common statistical tests are linear models \(or: how to teach stats\)](#)
- 7) [Modern Statistics with R](#)
- 8) [Introduction to Probability for Data Science](#)
- 9) [Intro to Bayesian Business Analytics](#)
- 10) [Design and Analysis of Experiments and Observational Studies using R](#)
- 11) [Statistical rethinking with brms, ggplot2, and the tidyverse](#)
- 12) [Bayes Rules! An Introduction to Applied Bayesian Modeling](#)

6.0.2 Personal Websites and Blogs

- 1) [Statistical Thinking](#)

6.0.3 Lectures and Courses

- 1) [Statistics 110: Probability](#)
- 2) [Statistical Computing](#)
- 3) [BIOLOGY FOR LIFE](#)
- 4) [Program Evaluation for Public Service and Research Design](#)
- 5) [STA 363: Statistical Learning](#)
- 6) [STA 312: Linear Models](#)
- 7) [STA 212: Statistical Models](#)
- 8) [Statistics 431: Advanced Statistical Computing with R](#)
- 9) [GUiDe to STatistical Analysis in Microbial Ecology \(GUSTA ME\)](#)

- 10) [Regression Modelling in R](#)
- 11) [The Grammar of Experimental Design](#)

6.0.4 Tutorials

- 1) [Welcome to A/B Testing at Scale Tutorial](#)
- 2) [Causation and Correlation](#)
- 3) [An Introduction to Hierarchical Modeling](#)
- 4) [Tidy Statistics in R](#)
- 5) [Statistics and R](#)
- 6) [Tidy dataframes and statsExpressions](#)
- 7) [Power Analysis | Introduction in R](#)
- 8) [metrica: Prediction performance metrics](#)
- 9) [Bayes Explained 1](#)
- 10) [Bayes Explained 2](#)
- 11) [CausalImpact](#)

6.0.5 Troubleshooting

- 1) [Statistical Problems to Document and to Avoid](#)
- 2) [Datamethods](#)

6.0.6 Biostatistics

- 1) [Biostatistics for Biomedical Research](#)
- 2) [Tutorials in Biostatistics Papers](#)
- 3) [Biostatistics Series Papers](#)
- 4) [WCM Biostatistics Computing Club](#)
- 5) [Katherine \(Kat\) Hoffman Biostats](#)

7 Public Datasets

This page provides a comprehensive collection of resources on datasets that can be used for statistical analysis, machine learning, and data journalism. The page includes curated collections of free and open data sources, industry-specific data, and multiple public data sources.

7.1 Working with Data

- 1) [Data Literacy](#)
- 2) [Using Data Appropriately](#)
- 3) [Sharing Data](#)
- 4) [WTF CSV Data Organization](#)
- 5) [Data Cleaning](#)

7.2 Free and Open Data

- 1) [Open Africa Data](#)
- 2) [Data Africa](#)
- 3) [Our World in Data](#)
- 4) [California Data Sources](#)
- 5) [DataSF](#)
- 6) [NYC Open Data](#)
- 7) [The Humanitarian Data Exchange](#)
- 8) [Global Data Barometer](#)
- 9) [Microsoft Research Open Data](#)
- 10) [Awesome Public Datasets](#)
- 11) [Multiple Public Data Sources](#)
- 12) [Data Science Central Datasets](#)
- 13) [70 Amazing Free Data Sources You Should Know](#)
- 14) [Rebecca Barter - List of Public Data Repositories](#)

7.3 Machine Learning Training Data

- 1) [Datasets for Machine Learning and Deep Learning](#)
- 2) [Best Public Datasets for Machine Learning and Data Science](#)
- 3) [Machine Learning Food Datasets Collection](#)
- 4) [Datasets for Data Science and Machine Learning](#)
- 5) [R Datasets](#)
- 6) [Kaggle Datasets](#)
- 7) [Google Research Datasets](#)

7.4 Industry-Specific Data

- 1) [NHS R-community Datasets](#)
- 2) [Infectious Diseases Data Observatory](#)
- 3) [Inside Airbnb](#)
- 4) [School Shootings USA](#)
- 5) [Political Datasets](#)
- 6) [Eighty-Five Points Sports Data](#)
- 7) [Poverty and Inequality Platform API](#)
- 8) [Nasdaq Data Link](#)
- 9) [Analyze Survey Data for Free](#)
- 10) [Global Clinical Research Data](#)
- 11) [Google Community Mobility Reports](#)
- 12) [15 Open Datasets for Healthcare](#)
- 13) [Demographic and Health Surveys](#)
- 14) [FiveThirtyEight Data](#)
- 15) [Jo Hardin Datasets](#)
- 16) [StatsBomb Open Data](#)

8 Sports Data Science

This page provides a comprehensive collection of resources on sports data science. The page includes links to tutorials, datasets, journal articles, and personal blogs focused on sports data science.

8.1 Sports Data Science Blogs

8.1.1 Multiple Sports

- 1) [Sweep Sports Analytics](#)
- 2) [Jason Zivkovic](#)
- 3) [SPORT DATA SCIENCE](#)

8.1.2 Soccer

- 1) [SBNation Blogs](#)
- 2) [Sports Intelligence](#)
- 3) [11tegen11](#)
- 4) [Karun Singh](#)
- 5) [Mackay Analytics](#)
- 6) [Son of a Corner](#)
- 7) [Saints by Numbers](#)
- 8) [Achim Zeileis](#)
- 9) [Opta Analyst](#)
- 10) [The Come on Man](#)
- 11) [The Last Man Analytics](#)
- 12) [Analytics FC](#)
- 13) [Brendan Kent](#)
- 14) [WyScout](#)
- 15) [Biscuit Chaser FC](#)
- 16) [StatsBomb](#)
- 17) [EightyFivePoints](#)
- 18) [Owen Thompson](#)
- 19) [Tony El Habr](#)

- 20) [11v11](#)

8.1.3 Basketball Analytics

- 1) [Dribble analytics](#)

8.2 Sports Data Science Stories

- 1) [Armando Broja: A Case Study](#)
- 2) [Out of the Park](#)
- 3) [How Data \(and Some Breathtaking Soccer\) Brought Liverpool to the Cusp of Glory](#)
- 4) [How data analysis helps football clubs make better signings](#)
- 5) [LOAN PATHWAYS: THE ACADEMY MODEL](#)
- 6) [LOAN PATHWAYS: THE PORTFOLIO MODEL](#)
- 7) [LOAN PATHWAYS: THE MULTI-CLUB MODEL](#)

8.3 Sports Data Science Journals and Readings

- 1) [EightyFivePoints Reading List & Data](#)
- 2) [Sports analytics reading list](#)
- 3) [Sports analytics conference video archives](#)
- 4) [Journal of Sports Analytics](#)
- 5) [Journal of Quantitative Analysis in Sports](#)

8.4 Tools and Databases

8.4.1 Multiple Sports

- 1) [PySport](#)
- 2) [DTAI SPORTS ANALYTICS LAB](#)
- 3) [SportsDataVerse](#)
- 4) [Sports guideR](#)
- 5) [Sports Reference](#)
- 6) [SFU Sports Analytics Group Virtual Seminar](#)
- 7) [Brescia BDAS Lab](#)
- 8) [SPORT PERFORMANCE ANALYSIS](#)
- 9) [Bruin Sports Analytics](#)
- 10) [Spotrac](#)

- 11) [THE GREAT ANALYTICS RANKINGS](#)
- 12) [Sport Informatics and Analytics](#)

8.4.2 Soccer

8.4.2.1 Curated list of soccer resources

- 1) [Awesome Soccer Analytics](#)
- 2) [Edd Webster Football Analytics](#)
- 3) [Soccer Analytics Handbook](#)

8.4.2.2 Individual resources

- 1) [Soccermatics](#)
- 2) [MPL Soccer](#)
- 3) [American Soccer Analysis](#)
- 4) [Total Football Analysis](#)
- 5) [Learn Python & Data Science With Football](#)
- 6) [Learn R & Data Science With Football](#)
- 7) [Experimental 361](#)
- 8) [Barca Innovation Hub](#)
- 9) [Training Ground Guru](#)
- 10) [SportSciData](#)
- 11) [FourFourTwo](#)
- 12) [Twelve Football](#)
- 13) [Football Science](#)
- 14) [THE DATA CRUNCHING FOOTBALL EXPERT](#)
- 15) [worldfootballR](#)
- 16) [Languages and tools to learn for sports analytics](#)
- 17) [Coding for sports analytics](#)

8.4.2.3 Databases

- 1) [CIES Football Observatory](#)
- 2) [The Fjelstul English Football Database](#)
- 3) [The Fjelstul World Cup Database](#)
- 4) [StatsBomb Open Data](#)
- 5) [Football Data UK](#)
- 6) [SoFIFA](#)

8.4.2.4 American Football

- 1) [Big Data Bowl '21 Submission](#)
- 2) [Modeling Player Biomechanics](#)
- 3) [Fantasy Football Analytics](#)

8.4.2.5 Basketball

- 1) [NBASTuffer](#)

8.4.3 Tutorials

8.4.3.1 Soccer

- 1) [A Beginner's Introduction to Mixed Effects Models in Sport](#)
- 2) [How Our Club Soccer Predictions Work](#)
- 3) [Google Research Football with Manchester City F.C.](#)
- 4) [Data in Elite Sport](#)
- 5) [COMPUTER VISION IN SPORT](#)
- 6) [Getting Started with Free StatsBomb Event Data](#)
- 7) [Predicting Football Results With Statistical Modelling](#)
- 8) [Liverpool 20-21 EPL Performance](#)
- 9) [Explaining Expected Threat](#)
- 10) [Working with data in elite sport](#)
- 11) [Mathematical Modelling of Football](#)
- 12) [10 Steps to Get Started in Sports Analytics](#)
- 13) [Flank Attacks Visualisation](#)
- 14) [Doing More With StatsBomb Data in R](#)
- 15) [How sports scientists can use ggplot2 in R to make better visualisations](#)
- 16) [Visualizing the Premier League](#)
- 17) [Fitting your own football xG model](#)
- 18) [PassSonar](#)
- 19) [Adding Football Crests to Scatter Plots in R](#)

8.4.3.2 Basketball

- 1) [How to make NBA shots charts in R](#)
- 2) [Animating NBA Play by Play using R](#)

9 Social Media and Music Analysis

On this page, there is a comprehensive collection of resources focused on social media and music analysis. The page includes links to books, courses, and tutorials that can help with social media, music, and lyric analysis.

9.1 Social Media Analysis

9.1.1 Books

- 1) [Text Mining with R](#)
- 2) [Social Media Mining](#)
- 3) [Social Media Mining \(Slides\)](#)

9.1.2 Tutorials

- 1) [ANALYSING TWITTER DATA WITH R](#)
- 2) [ANALYSING TWITTER DATA: EXPLORING USER PROFILES AND RELATIONSHIPS](#)
- 3) [ANALYSING TWITTER DATA: EXPLORING TWEETS CONTENT](#)
- 4) [4 ways how to create a gradient word cloud in R](#)
- 5) [Analyzing Google Trends with R: Retrieve and plot with gtrendsR](#)
- 6) [Quantitative Analysis of Political Text](#)
- 7) [Classifying Tweets for Sentiment Analysis](#)
- 8) [Advancing Text Mining with R and quanteda](#)
- 9) [Using Twitter With Python And Tweepy](#)
- 10) [USING R TO ANALYZE AND VISUALIZE TWEETS FROM SACNAS](#)
- 11) [Twitter “Account Analysis” in R](#)
- 12) [Calculating \(Twitter\) Vocabulary Breadth of U.S. Presidential Candidates Using TTR](#)
- 13) [Twitter data analysis in R](#)
- 14) [Twitter’s Feelings About Programming Languages](#)
- 15) [Look who’s tweeting: 2020 Presidential Candidates](#)
- 16) [A TidyText Analysis With R](#)
- 17) [Using Twitter as a data source an overview of social media research tools \(2021\)](#)

9.1.3 Courses

- 1) [Text As Data](#)

9.1.4 Projects

- 1) [twitter-sentiment-analysis](#)
- 2) [Analytics Master's Research Project](#)

9.2 Music and Lyric Analysis

9.2.1 Basics

- 1) [Spotify R&D](#)
- 2) [spotifyr package](#)
- 3) [13 best websites to analyze your Spotify data](#)

9.2.2 Tutorials

- 1) [Spotify Songs: Tidy Tuesday](#)
- 2) [How to integrate Spotify and Genius API to easily crawl song lyrics with Python](#)
- 3) [Profiling Songs on Spotify Using Cluster Analysis](#)
- 4) [Sentiment Analysis of Most Streamed Spotify Artists](#)
- 5) [Lyric Analysis with NLP & Machine Learning with R](#)
- 6) [What's popping?: An Exploratory Analytics Project on What Makes Popular Music Popular](#)
- 7) [Analysis of my Spotify in R with spotifyr package](#)
- 8) [Game of Thrones: Exploratory and Sentiment Analysis](#)
- 9) [tayloR](#)
- 10) [Exploring the Spotify API with R](#)
- 11) [Learn tidytext with my new learnr course](#)
- 12) [Text mining with tidy data principles](#)
- 13) [Are Pop Lyrics Getting More Repetitive?](#)
- 14) [Spotify Songs: An analysis of the spotify dataset](#)
- 15) [Fitted Happier](#)
- 16) [Data Wrangling - Spotify Dataset Analysis](#)
- 17) [Understanding + classifying genres using Spotify audio features](#)
- 18) [What Makes a Hit](#)
- 19) [What Makes Songs Popular? It's All About 'You'](#)
- 20) [Can big data really predict what makes a song popular?](#)

- 21) [PCA and the #TidyTuesday best hip hop songs ever](#)
- 22) [Blue Christmas: A data-driven search for the most depressing Christmas song](#)
- 23) [Using Data to Find the Angriest Death Grips Song](#)
- 24) [Spotify songs popularity analysis](#)
- 25) [Nigerian Songs Spotify EDA](#)
- 26) [Sentiment Analysis of Spotify, Amazon Music, and Pandora user reviews](#)

10 Data Science Interview

On this page, there is a comprehensive collection of resources on preparing for data science interviews. There are general interview questions, language- and field-specific interview questions, and job descriptions.

10.1 Job Descriptions

- 1) [Role of a Data Science Manager](#)
- 2) [Six Steps To Take Before Pursuing Education To Get A Data Science Job](#)
- 3) [How do Data Professionals Spend their Time on Data Science Projects?](#)
- 4) [Things I Have Learned About Data Science](#)
- 5) [This Is What a Data Scientist Looks Like](#)
- 6) [How to land a Data Scientist job at your dream company — My journey to Airbnb](#)
- 7) [How to Become More Marketable as a Data Scientist](#)
- 8) [How to Build a Data Science Portfolio](#)

10.2 Curated collection of Data Science Resources

- 1) [Data Science Career Resources](#)
- 2) [The Big List of Data Science Interview Resources](#)

10.3 General Interview Questions

- 1) [101 important data science interview questions, answers, and key concepts](#)
- 2) [Interview Warmup](#)
- 3) [Amazon Data Science Case Question: Duplicate Products](#)
- 4) [10 Frequently Asked Interview Questions For Machine Learning In 2019](#)
- 5) [I Worked Through 500+ Data Science Interview Questions](#)
- 6) [16 Machine Learning Interview Questions and Answers for Entry-Level Applicants](#)
- 7) [A Guide for Machine Learning Technical Interviews \(FAANG Companies\)](#)
- 8) [Data Science Interviews](#)

10.4 Language- and Field-Specific Interview Questions

- 1) [How To Ace Data Science Interviews: SQL](#)
- 2) [How To Ace Data Science Interviews: Statistics](#)
- 3) [How To Ace Data Science Interviews: R & Python](#)
- 4) [Ace the SQL & Data Science Interview](#)
- 5) [Analyzing 89 Responses to a SQL Screener Question for a Senior Data Analyst Position](#)

Part III

Data Visualization and Storytelling

This page provides a comprehensive collection of resources on data visualization and storytelling. There is a curated collection of resources focused on general data visualization, and data visualization tools such as R/RStudio, Python, and Tableau. Additionally, there is a section focused on data journalism and storytelling.

A curated collection of resources

Visualization

10.4.0.0.1 * 1. [Data Visualization \(General\)](#)

10.4.0.0.2 * 2. [Data Visualization \(R\)](#)

10.4.0.0.3 * 3. [Data Visualization \(Python\)](#)

10.4.0.0.4 * 4. [Data Visualization \(Tableau\)](#)

Journalism and Storytelling

10.4.0.0.1 * 1. [Data Journalism and Storytelling](#)

11 Data Visualization (General)

On this page, you will find a list of general data visualization resources. First, there is a list of data visualization specialists and companies to follow. Additionally, there are practical introductions to data visualization, style guidelines, and books.

11.1 Data Visualization (Blogs)

A listing of data visualization specialists and companies to follow:

11.1.1 People

- 1) [Maarten Lambrechts](#)
- 2) [Ada Homolova](#)
- 3) [Lisa Charlotte Muth](#)
- 4) [Zev Ross](#)
- 5) [Edward Tufte](#)
- 6) [CÉDRIC SCHERER](#)
- 7) [Kieran Healy](#)
- 8) [Little Miss Data](#)
- 9) [SportsChord](#)

11.1.2 Companies

- 1) [Stamen](#)

11.2 Data Visualization (General)

11.2.1 Selecting the appropriate chart types

- 1) [Visual Vocabulary](#)
- 2) [The unspoken rules of visualisation](#)
- 3) [What Are Data Visualization Style Guidelines?](#)

- 4) [Data visualization](#)
- 5) [Why scientists need to be better at data visualization](#)
- 6) [Data Viz Project](#)
- 7) [A Tufte Handout Example](#)

11.2.2 Selecting the appropriate colors

- 1) [Color in Data Vis](#)
- 2) [Color Brewer](#)
- 3) [Your Friendly Guide to Colors in Data Visualisation](#)

11.2.3 Data Viz Challenges

- 1) [Black in Data Vis Challenge](#)
- 2) [VIZRISK CHALLENGE STEPS](#)

11.2.4 Books

- 1) [A Reader on Data Visualization](#)
- 2) [Data Visualization: A practical introduction](#)

12 Data Visualization (R)

This page contains a list of resources focused on data visualization in R. There are links to books, templates, tutorials, and courses. Additionally, for those interested in using coding for art, there is a section on "Data aRtistry."

12.1 Curated Links by other data viz specialists

- 1) [Awesome ggplot2](#)

12.2 Books

- 1) [An introduction to data visualization protocols for wet lab scientists](#)
- 2) [Cookbook for R](#)
- 3) [R Graphics Cookbook](#)
- 4) [Data Visualization with R](#)
- 5) [Data visualisation using R, for researchers who don't use R](#)
- 6) [Fundamentals of Data Visualization](#)
- 7) [ggplot2: Elegant Graphics for Data Analysis](#)
- 8) [R4DS Data Visualization](#)
- 9) [Interactive web-based data visualization with R, plotly, and shiny](#)
- 10) [BBC Visual and Data Journalism cookbook for R graphics](#)
- 11) [ggplot2](#)

12.3 Templates

- 1) [R CHARTS](#)
- 2) [The R Graph Gallery](#)
- 3) [ggplot2 extensions](#)
- 4) [A Scientist's Guide to R: Step 3.1 - data visualization with ggplot2+](#)
- 5) [ggplot tricks](#)
- 6) [Designing ggplots](#)
- 7) [BBC Style Plot Templates](#)

- 8) [a ggplot2 grammar guide](#)
- 9) [Tom's Cookbook for Better Viz](#)
- 10) [12 ggplot extensions for snazzier R graphics](#)
- 11) [Data Visualization Packages for R you should check out in 2022](#)

12.4 Courses

- 1) [Graphic Design with ggplot2](#)
- 2) [Data Visualization in R](#)
- 3) [Data Visualization](#)
- 4) [SDS 375: Data Visualization in R](#)

12.5 Tutorials and Notes

12.5.1 Basics

- 1) [Top 50 ggplot2 Visualizations - The Master List \(With Full R Code\)](#)
- 2) [The Complete ggplot2 Tutorial - Part1 | Introduction To ggplot2 \(Full R code\)](#)
- 3) [Best ggplot visualizations](#)
- 4) [A quick introduction to ggplot2](#)
- 5) [Getting started with data visualization in R using ggplot2](#)
- 6) [The Evolution of a ggplot \(Ep. 1\)](#)
- 7) [A ggplot2 Tutorial for Beautiful Plotting in R](#)
- 8) [R you ready to make charts?](#)
- 9) [The ggplot flibbook](#)
- 10) [Getting started with data visualization in R using ggplot2](#)
- 11) [ggplot2: Mastering the basics](#)
- 12) [Graphing with ggplot](#)
- 13) [10 Levels of ggplot2: From Basic to Beautiful](#)
- 14) [Practical ggplot2](#)
- 15) [Beautiful plotting in R: A ggplot2 cheatsheet](#)
- 16) [The Good, the Bad and the Ugly: how \(not\) to visualize data](#)
- 17) [Six things I always Google when using ggplot2](#)
- 18) [ggplot2tutor](#)

12.5.2 Troubleshooting

- 1) [Common R Mistakes in Data Viz](#)

12.5.3 Enhancing Plots (with Labels, Axes, Themes, Scales, and Other Options)

- 1) [Quick and easy ways to deal with long labels in ggplot2](#)
- 2) [Making text labels the same size as axis labels in ggplot2](#)
- 3) [Labelling area plots](#)
- 4) [Level Up Your Labels: Tips and Tricks for Annotating Plots](#)
- 5) [Increase & Decrease Number of Axis Ticks in R \(2 Examples\)](#)
- 6) [ggplot2 Theme Elements Demonstration](#)
- 7) [How to use multiple colour scales in ggplot with {ggnewscale}](#)
- 8) [Spruce up your ggplot2 visualizations with formatted text](#)
- 9) [Themes to Improve Your ggplot figures](#)
- 10) [How to Add Labels Directly in ggplot2](#)
- 11) [Generate your own ggplot theme gallery](#)
- 12) [ggpattern](#)
- 13) [{ggtext} for images as x-axis labels](#)
- 14) [GGPLOT THEMES GALLERY](#)
- 15) [ggplot theme tester](#)
- 16) [Introducing 'optout' - A package for optimizing graphics output](#)
- 17) [ggcute](#)
- 18) [Radial Patterns in ggplot2](#)
- 19) [Reordering and facetting for ggplot2](#)
- 20) [Accelerate your plots with ggforce](#)
- 21) [ggrapid](#)
- 22) [Patchwork R package goes nerd viral](#)
- 23) [ggthemes](#)
- 24) [htmlwidgets](#)
- 25) [ggchicklet](#)
- 26) [ggplot date helpers](#)
- 27) [12 Extensions to ggplot2 for More Powerful R Visualizations](#)
- 28) [Improving a Visualization](#)
- 29) [ggplot theme for publication ready Plots](#)
- 30) [WHY NOT DO THAT WITH R?](#)

12.5.4 Various graph types

- 1) [Creating flowcharts with {ggplot2}](#)
- 2) [Dendrograms in R, a lightweight approach](#)
- 3) [Racing Barchart with gganimate](#)
- 4) [Racing Barplots with {ggpattern}, {flagon} & {gganimate}](#)
- 5) [R base plotting without wrappers](#)
- 6) [How to basic: bar plots](#)
- 7) [Network Visualizations in R](#)

- 8) [Great Looking Tables](#)
- 9) [ggplot2 Bar Chart](#)
- 10) [How to build Animated Charts like Hans Rosling — doing it all in R](#)
- 11) [How to create Bar Race Animation Charts in R](#)
- 12) [Animating Your Data Visualizations Like a Boss Using R](#)
- 13) [barbarplots](#)
- 14) [Dynamic display of data with animated plots](#)
- 15) [Beginners Cookbook for Interactive Visualization in R with highcharter](#)

12.5.5 Data Viz Examples

- 1) [Can I do that? Inspiration from a Pudding data visualization](#)
- 2) [Visualizing the heterogeneity of single cell data from time-lapse imaging](#)
- 3) [the seven living generations in America](#)
- 4) [this old house](#)
- 5) [Visualizing real house price plots](#)
- 6) [State Unemployment Claims](#)
- 7) [A look into U.S. infectious diseases](#)
- 8) [NYT-style urban heat island maps](#)
- 9) [Ryo Viz Data Viz Gallery](#)
- 10) [Reproducing an Axios plot](#)
- 11) [Create a Visual CV](#)
- 12) [Visualize monthly precipitation anomalies](#)
- 13) [Inspiration from Flowing Data](#)
- 14) [Visualizing Time Series](#)
- 15) [Visualising high-dimensional data](#)
- 16) [Data Storytelling in R](#)
- 17) [William Chase - Data Viz Projects](#)

12.6 Data aRtistry

- 1) [Getting started in #rtistry](#)
- 2) [ART FROM CODE](#)
- 3) [Life Through A Mathematician's Eyes](#)
- 4) [aRt](#)
- 5) [Rtistry Gallery](#)
- 6) [Art by Claus O. Wilke](#)
- 7) [Unpredictable paintings](#)
- 8) [Mystery curves](#)
- 9) [Flow Fields](#)
- 10) [HOW TO WORK WITH FLOW FIELDS IN R](#)

13 Data Visualization (Python)

The Python Data Visualization page provides a list of books and tutorials for beginners. There are links to practical introductions to data visualization, tools and style guidelines.

13.1 Books

- 1) [The Python Graph Gallery](#)

13.2 Tutorials

- 1) [Visualizing Data with Pairs Plots in Python](#)
- 2) [How to Make a Scatter Plot in Python using Seaborn](#)
- 3) [Seaborn Heatmaps: 13 Ways to Customize Correlation Matrix Visualizations](#)
- 4) [How to make a matplotlib line chart](#)
- 5) [Python Plotting for Exploratory Data Analysis](#)
- 6) [Effectively Using Matplotlib](#)
- 7) [PyWaffle](#)
- 8) [Easily hiding items from the legend in matplotlib](#)
- 9) [The Next Level of Data Visualization in Python](#)
- 10) [Introducing Plotly Express](#)
- 11) [seaborn: statistical data visualization](#)
- 12) [PyGWalker](#)

14 Data Visualization (Tableau)

The Tableau Data Visualization page provides a list of books and tutorials for beginners. There are links to practical introductions to data visualization, tools and style guidelines.

14.1 Data Visualization (Tableau)

- 1) [A Beginner's Guide to Tableau Public](#)
- 2) [Tableau Resources](#)
- 3) [What Is Data Visualization? Definition, Examples, And Learning Resources](#)
- 4) [The 10 Best Data Visualization Blogs To Follow](#)
- 5) [How to ace the Tableau Desktop Qualified Associate exam!](#)
- 6) [Tableau Tips, Dashboards, and Training](#)
- 7) [7 tips and tricks from the dashboard experts](#)
- 8) [Data Visualization with Tableau](#)

15 Data Journalism Tools and Stories

This page provides a comprehensive collection of resources focused on data journalism tools and resources. On this page, you will find a list of personal websites and blogs that do some great work in data journalism. Additionally, there is a list of good examples of data stories to emulate.

15.1 Data Journalism

15.1.1 Blogs and Websites

A listing of data journalism practitioners and companies to follow:

- 1) [Seth Stephens-Davidowitz](#)
- 2) [Quant Editing](#)
- 3) [Amber Thomas](#)
- 4) [Data Journalism and Viz Accounts](#)
- 5) [Flowing Data Website](#)

15.1.2 Guides

- 1) [GitHub Data Journalism](#)
- 2) [A Beginner's Guide to Data Journalism: Let's Build a Story From Scratch](#)
- 3) [Quick Guide to Data Journalism](#)
- 4) [Data journalism: a guide for editors](#)
- 5) [Developing Data Journalists in the Developing World](#)

15.1.3 Writing

- 1) [Interrogating Data: A Science Writer's Guide to Data Journalism](#)
- 2) [The Beginning, The Middle, & The New Beginning](#)
- 3) [How data journalists create beautiful feature stories](#)
- 4) [A data journalist's guide to building a hypothesis](#)
- 5) [Flowing Data Story Formats](#)

- 6) [Telling Stories With Data: A Simple Math Guide](#)
- 7) [Data Biographies: Getting to Know Your Data](#)
- 8) [Putting data back into context](#)
- 9) [Dueling Data](#)

15.1.4 Tools

- 1) [Data Journalism Tools](#)
- 2) [Data Journalism and Visualization with Free Tools](#)
- 3) [Storytelling Tools](#)
- 4) [StoryBench: The future of machine learning in journalism](#)
- 5) [How ProPublica Analyzed Pardon Data](#)

15.1.5 Tutorials

- 1) [Mapping search data from Google Trends in R](#)
- 2) [Using R to Analyze Fargo Crime Data](#)

15.1.6 Organizations

- 1) [Global Investigative Journalism Network Data Journalism](#)
- 2) [ProPublica Data Institute 2018](#)
- 3) [ProPublica Data Institute 2019](#)
- 4) [SIGMA Awards](#)
- 5) [Knight Foundation: Misinformation](#)
- 6) [The Bureau of Investigative Journalism](#)
- 7) [Fake News Public Data Lab](#)
- 8) [AP DataKit](#)
- 9) [Data Journalism Grants](#)
- 10) [A community conference for data makers everywhere](#)
- 11) [Google News Initiative](#)

15.2 Data Stories (Longform and In-depth articles)

15.2.1 Story aggregators

- 1) [Global Investigative Journalism Network Top 10](#)
- 2) [Media Hack Collective](#)
- 3) [12 brilliant data journalism projects of 2021](#)

- 4) [Data Journalism: Some Favorite Examples](#)
- 5) [Examples of Award-winning Data Journalism](#)
- 6) [12 Best Data Journalism Examples: Storytelling with data](#)
- 7) [10 data journalism projects that made an impact in 2020: our ultimate COVID-19 roundup](#)

15.2.2 Top Data Story Examples

- 1) [10 years. 180 school shootings. 356 victims.](#)
- 2) [Blind Drunk](#)
- 3) [Uber Game](#)
- 4) [What Africa will look like in 100 years](#)
- 5) [ONE NATION, TRACKED](#)
- 6) [Life in the camps](#)
- 7) [Life in Hong Kong's shoebox housing](#)
- 8) [The Airbnb effect](#)
- 9) [NSA Files Decoded](#)
- 10) [The Deadliest Jobs in America](#)
- 11) [Congressional Tweets](#)
- 12) [The Kim Foxx Effect](#)
- 13) [What makes a titletown?](#)
- 14) [High street crisis deepens](#)
- 15) [The use of political ads on facebook in the run-up to the European Parliament elections 2019 in Austria](#)
- 16) [What the Democratic Presidential Candidates Care About, in 44,000 Tweets](#)
- 17) [How many high school stars make it in the NBA?](#)
- 18) [The Facebook Scandal](#)

Part IV

Generative AI and Large Language Models

This page contains a comprehensive collection of guides and resources to help you learn about generative AI and large language models (LLM).

16 Generative AI and Large Language Models (Part 1)

Part V

Teaching Resources

On this page, there is a curated collection of general teaching resources, links to RStudio and Carpentries certification resources, and content focused on medical, science, and technical writing.

A curated collection of resources

16.0.0.0.1 * 1. [General Teaching Resources](#)

16.0.0.0.2 * 2. [RStudio and Carpentries Certification](#)

16.0.0.0.3 * 3. [Medical, Science, and Technical Writing](#)

17 General Teaching Resources

The general teaching resources page provides resources focused on teaching STEM subjects, as well as links to teaching aids for more efficient teaching.

17.1 Teaching STEM Subjects

17.1.1 A curated collection of resources

- 1) [Teaching Tech Together](#)
- 2) [Maintaining a laboratory notebook](#)
- 3) [Subjects and Skills for Graduate Students in the Biological Sciences](#)
- 4) [Scientific Figure Design](#)

17.2 Teaching Aids

17.2.1 Typing Test

- 1) [Keyboard Test](#)
- 2) [Typing Test](#)

17.2.2 Simulations

- 1) [Physics Education and Training Simulations](#)
- 2) [Circuit simulation and schematics](#)
- 3) [Interactive STEM activities](#)
- 4) [GeoGebra for Teaching and Learning Math](#)
- 5) [GeoGebra Simulations](#)

17.2.3 Online Textbooks

- 1) [Free and flexible textbooks and resources](#)
- 2) [College Physics](#)

17.2.4 Design and Illustrations

- 1) [One Page Website](#)
- 2) [Illustrations](#)
- 3) [BioRender Learning Hub](#)
- 4) [Beautiful Free Stock Video Footage](#)

17.2.5 Miscellaneous Tools

- 1) [Whiteboard](#)
- 2) [Etherpad](#)
- 3) [Active Memory](#)

18 RStudio and Carpentries Certification

This page provides a list of tutorials, resources and study notes for RStudio and Carpentries certification.

18.1 RStudio Certification

18.1.1 Tutorials and Resources

- 1) [R4DS Book – Prep for Exam](#)
- 2) [POSIT/R Training Videos](#)
- 3) [POSIT Primers](#)
- 4) [RStudio Sample Exams](#)
- 5) [Miscellaneous Teaching Notes 1](#)
- 6) [Miscellaneous Teaching Notes 2](#)
- 7) [RStudio Instructor Training](#)

18.1.2 Study notes from RStudio Certified Trainers

- 1) [Laurie L Baker RStudio Instructor Training](#)
- 2) [Marly Comar Gotti Tidyverse Certification](#)
- 3) [Yuqi Liao: Getting Certified as an RStudio Instructor](#)
- 4) [Shel Kariuki: The RStudio Certification Process](#)
- 5) [Florenzia D’Andrea: Two examples of iteration with purrr - Class for the RStudio certification](#)
- 6) [Silvia Canelon RStudio Certification](#)
- 7) [Silvia Canelon RStudio Certification Training 1](#)
- 8) [Silvia Canelon RStudio Certification Training 2](#)
- 9) [Silvia Canelon R4DS Certification Training 1](#)
- 10) [Paloma Rojas Saunero RStudio Certification Training 1](#)
- 11) [Paloma Rojas Saunero RStudio Certification Exam Slides](#)
- 12) [Paloma Rojas Saunero RStudio Certification Exam Notes](#)

18.2 Carpentries Certification

- 1) [The Carpentries](#)
- 2) [Data Carpentry](#)
- 3) [Software Carpentry](#)
- 4) [Library Carpentry](#)

19 Medical, Science, and Technical Writing

For anyone interested in medical, science, and technical writing, this page offers a curated collection of resources. It provides practical introductions to medical, science, and technical writing, style guidelines, and challenges.

19.1 Grammar Basics

- 1) [Academic Phrasebank](#)
- 2) [Excelsior Online Writing Lab](#)
- 3) [Purdue Online Writing Lab](#)
- 4) [Quick and Dirty Tips](#)
- 5) [Inter-Biotech: Biomedical Language Editing](#)
- 6) [Features of academic writing](#)
- 7) [American Copy Editors Series](#)
- 8) [Go The Distance: Academic Writing](#)
- 9) [ACES Grammar Guide Quizzes](#)
- 10) [Introduction to Common Errors in English Usage](#)

19.2 Medical Writing

- 1) [Health Writer Hub: 30-day writing challenges](#)
- 2) [MedComms Networking](#)
- 3) [Debra Gordon](#)
- 4) [Armitage Medical](#)
- 5) [Martin Medical Writing](#)
- 6) [The Med Writers](#)

19.3 Science Writing

- 1) [A Note to Beginning Science Writers](#)
- 2) [Science Blogging 101](#)
- 3) [How to write a science blog](#)

- 4) [Carl Zimmer's advice for aspiring science writers](#)
- 5) [Six great pieces of science writing you may have missed this year](#)

19.4 Technical Writing

- 1) [Awesome Technical Writing](#)
- 2) [A beginner's guide to writing documentation](#)
- 3) [Technical Writing: Why and How?](#)
- 4) [Welcome to Google Season of Docs](#)
- 5) [Becoming a Technical Writer at Google](#)
- 6) [Advice for Technical Writing](#)
- 7) [The Society for Technical Communication](#)
- 8) [How to write a great technical blog post](#)

Part VI

Bio-X: All things Bio

On this page, I provide a curated collection of resources focused on all things related to biotechnology. The main topics that I focus on include open-source bio, biomechanics, and bioinformatics.

19.4.0.0.1 * 1. [Biomechanics and Sports Engineering](#)

19.4.0.0.2 * 2. [Open-Source Bio](#)

19.4.0.0.3 * 3. [Bioinformatics and Computational Biology](#)

20 Bioinformatics and Computational Biology

The Bioinformatics and Computational Biology page offers a curated collection of resources for bioinformatics and computational biology. The page includes personal and lab websites, tutorials, courses, books, and projects.

20.1 Personal and Lab Websites

- 1) [JT Leek](#)
- 2) [Karolis Koncevičius](#)
- 3) [Eytan Ruppin](#)
- 4) [Lior Pachter](#)
- 5) [Jean Fan](#)
- 6) [Liu Lab](#)
- 7) [Costa Lab](#)
- 8) [Ming Tang](#)
- 9) [James Zou](#)
- 10) [Karl Broman](#)
- 11) [Krishnaswamy Lab](#)
- 12) [Philip Compeau](#)
- 13) [ZHOU LAB](#)
- 14) [Childhood Cancer Data Lab](#)
- 15) [Diving into Genetics and Genomics](#)

20.2 Bioinformatics News Stories

- 1) [GENOMES ARISING](#)
- 2) [The next chapter for African genomics](#)

20.3 Bioinformatics Basics

20.3.1 Courses and Lectures

- 1) [Introduction to Genomics for Engineers](#)
- 2) [Introduction to Computational Biology](#)
- 3) [SDS 348, Computational Biology and Bioinformatics 2019](#)
- 4) [SDS 348, Computational Biology and Bioinformatics 2020](#)
- 5) [BMI 702 | Biomedical Artificial Intelligence](#)
- 6) [DIY Transcriptomics](#)
- 7) [BIG Bioinformatics](#)
- 8) [Intro to Biomedical Data Science](#)
- 9) [Introduction to Unix, Orchestra and RNA-Seq](#)
- 10) [Linux command line exercises for NGS data processing](#)
- 11) [Introduction to Unix](#)
- 12) [Engineering Biology Research Consortium](#)
- 13) [Genomic Data Science Community Network](#)
- 14) [Learn about Bioinformatics and Computational Tools for Biology](#)
- 15) [Introduction to R for Biologists](#)
- 16) [Graphs and Networks for Biology using R](#)
- 17) [HBC Training Program Basic Data Skills](#)

20.3.2 Books and Resources

- 1) [Welcome to a Little Book of R for Bioinformatics](#)
- 2) [Modern Statistics for Modern Biology](#)
- 3) [Learning Bioinformatics At Home](#)
- 4) [Important numbers in Cell Biology](#)
- 5) [The Biologist's Guide to Computing](#)
- 6) [Bioinformatics Starter Pack](#)
- 7) [Orchestrating Single-Cell Analysis with Bioconductor](#)
- 8) [Single-cell best practices](#)
- 9) [Biodatascience101](#)
- 10) [Introduction to biological circuit design](#)
- 11) [Biological Architecture Reading Roadmap](#)
- 12) [RNA-Seq Blog](#)
- 13) [7 Books for you to learn bioinformatics](#)
- 14) [Biomedical Data Science Book](#)
- 15) [CRISPRpedia](#)
- 16) [Data Science for the Biomedical Sciences](#)
- 17) [What is genomics?](#)
- 18) [COLAUTTI LAB Books and Resources](#)

- 19) [20 YouTube Channels Every Bioinformatician Should Follow](#)
- 20) [Data Analysis for the Life Sciences \(Free E-book\)](#)
- 21) [Practical Computing for Biologists](#)
- 22) [Bioinformatics Data Skills \(2015\)](#)
- 23) [The Supplementary Material Repository for Bioinformatics Data Skills](#)
- 24) [From cell line to command line](#)
- 25) [A Primer for Computational Biology](#)
- 26) [Computational Genomics with R](#)
- 27) [The Biostar Handbook. A bioinformatics e-book for beginners](#)
- 28) [Cleva Lab Biology](#)
- 29) [Folding@Home](#)
- 30) [Python for biologists](#)
- 31) [Modeling Life](#)
- 32) [Fundamentals of computational biology](#)
- 33) [Learn Bioinformatics in 100 hours](#)
- 34) [Collection of bioinformatics training materials](#)
- 35) [Bioinformatics Workbook](#)
- 36) [My opinionated selection of books/urls for bioinformatics/data science curriculum](#)
- 37) [Statistical Modeling of High Dimensional Counts](#)
- 38) [Genomics Boot Camp](#)
- 39) [Deep learning on computational biology and bioinformatics tutorial](#)

20.3.3 Tools

- 1) [R2 Genomics Analysis and Visualization Platform](#)
- 2) [Exploring Spatial Omics](#)
- 3) [BD2K-LINCS DATA COORDINATION AND INTEGRATION CENTER](#)
- 4) [The Top 15 Machine Learning Data Science Bioinformatics Open Source Projects](#)
- 5) [Biopython](#)
- 6) [gget](#)
- 7) [ROSALIND](#)
- 8) [Free and Open Source Tools for Bioinformatics and Molecular Biology](#)
- 9) [PiGx: Pipelines in Genomics](#)
- 10) [USC Libraries Bioinformatics](#)
- 11) [NCI Bioinformatics Training and Education Program](#)
- 12) [Harvard Chan Bioinformatics Core](#)
- 13) [Genomic Data Processing](#)
- 14) [Machine Learning in Genomics Workshop](#)

20.4 Bioinformatics Tutorials and Projects

20.4.1 Tutorials

- 1) [Oncology Bioinformatics project tutorial](#)
- 2) [Centre of Bioinformatics Research and Technology \(CBIRT\)](#)
- 3) [Galaxy Training](#)
- 4) [Childhood Cancer Data Lab Workshop Materials](#)
- 5) [Childhood Cancer Data Lab Research Resources](#)
- 6) [Crash Courses in Bioinformatics](#)
- 7) [Data Carpentry for Biologists](#)
- 8) [Illustrating Python via Examples from Bioinformatics](#)
- 9) [Illustrating Python via Bioinformatics Examples](#)
- 10) [EVOLUTION AND GENOMICS](#)
- 11) [Bioinformatics Introduction](#)
- 12) [Griffith Lab Tutorials and Courses](#)
- 13) [PCA in action](#)
- 14) [A Primer on Deep Learning in Genomics](#)
- 15) [Bioinformatics Tutorials](#)
- 16) [Free online training in bioinformatics and biostatistics](#)
- 17) [Bioinformatics Midterm Project](#)
- 18) [Introduction to bioinformatics](#)
- 19) [Explore the world of Bioinformatics with Machine Learning](#)

20.4.1.1 Tidyverse and R

- 1) [Bioinformatics in the Tidyverse](#)
- 2) [An Introduction to R through the tidyverse + Bioinformatics](#)
- 3) [R for Biology Data Science](#)
- 4) [Intro to R and RStudio for Genomics](#)
- 5) [The Bioconductor 2018 Workshop Compilation](#)
- 6) [Bioconductor Workshop 1 Jupyter Notebooks](#)
- 7) [MicroArray Project](#)

20.4.1.2 BioPython

- 1) [First Steps in Biopython](#)
- 2) [What is BioPython](#)
- 3) [Getting Started with Biopython](#)
- 4) [Biopython Tutorial](#)
- 5) [Practical Computing for Biologists](#)

- 6) [Biopython Tutorial and Cookbook](#)

20.4.2 Projects

- 1) [Analysing the HIV pandemic, Part 1: HIV in sub-Saharan Africa](#)
- 2) [Analysing the HIV pandemic, Part 2: Drug resistance testing](#)
- 3) [Analysing the HIV pandemic, Part 3: Genetic diversity](#)
- 4) [Analysing the HIV pandemic, Part 4: Classification of lab samples](#)
- 5) [Machine learning for biology part one](#)
- 6) [Machine learning for biology part two](#)
- 7) [Machine learning for biology part three](#)
- 8) [Machine learning for biology part four](#)

21 Biomechanics and Sports Engineering

The Biomechanics and Sports Engineering page provides a curated collection of resources. Included here is a list of books, tutorials, and personal blogs and websites.

21.1 Biomechanics

21.1.1 Curated collection of resources

- 1) [A curated repository of biomechanical resources](#)

21.1.2 Personal Blogs and Websites

- 1) [Stuart McErlain-Naylor Biomechanics Website](#)

21.1.3 Labs

21.1.4 Books

- 1) [Biomechanics of Movement](#)
- 2) [The Biomechanics of Solids and Fluids: The Physics of Life](#)

21.1.5 Tutorials and Lectures

- 1) [Biomechanics of Movement Classroom](#)
- 2) [Reinbolt Research Group](#)
- 3) [Notes on Scientific Computing for Biomechanics and Motor Control](#)
- 4) [Understanding Biomechanics and Body Movement](#)
- 5) [The Biomechanist](#)

21.1.6 Sports Engineering and Movement Science

- 1) [Becoming a Sports Engineer](#)
- 2) [Peak Performance Project](#)
- 3) [BTS Bioengineering](#)
- 4) [Sparta Science](#)

22 Biotechnology Startups

22.1 Resources for Biotechnology Startups

- 1) [biotickr](#)
- 2) [Investor Village Biotech and Med](#)
- 3) [Biotech2k Background Info](#)
- 4) [Alcimed Biotech Business Exploration](#)
- 5) [Benchmarks for growing health tech businesses](#)
- 6) [The best ways to follow biopharma news and learn about the industry \(in my opinion\)](#)
- 7) [Timmerman Report: Your edge in Biotech](#)
- 8) [How to start a biotech company](#)
- 9) [How to Build a Biotech](#)
- 10) [Scaling Biotech](#)
- 11) [Anatomy of a Biotech Business Development Deal](#)
- 12) [The case for young biotech founders](#)
- 13) [Data-driven perspectives on the business of biotech](#)
- 14) [The Bio Revolution](#)
- 15) [Synthetic Biology Is About to Disrupt Your Industry](#)
- 16) [Investing in IDRx](#)
- 17) [Why We're Investing in EQRx](#)
- 18) [Term Sheet Recommendations for Launching University Life Science Startups](#)
- 19) [How Do You Become a Bioentrepreneur?](#)
- 20) [How to Spin Your Scientific Research Out Of a University and Into a Startup](#)
- 21) [Early-Stage Biotech Value Creation: The Roles of Equity and Partnerships](#)
- 22) [How Biotech Startup Funding Will Change in the Next 10 Years](#)
- 23) [The Creation Of Biotech Startups: Evolution Not Revolution](#)
- 24) [Biotech New Venture Formation: Reflecting On A Decade's Changing Dynamics](#)
- 25) [Biotech researchers venture into the wild to start their own business](#)
- 26) [Platform-Disease Fit: Why You Need It and How to Find It](#)
- 27) [It's Time to Heal: 16 Trends Driving the Future of Bio and Healthcare](#)

23 Open-Source Bio

This page provides resources focused on open-source science, medicine, and technology. There are personal blogs and websites, tools, projects, and publications that highlight progress in this field.

23.1 Science, Medicine, and Technology

23.1.1 Personal Blogs and Websites

- 1) [Andre Maia Chagas Open-Source](#)
- 2) [Synthetic Biology for engineering plant growth](#)

23.1.2 Tools

- 1) [Open Source Tools for Computational Biology](#)

23.1.3 Labs and Groups

- 1) [DIY BIO - Listing of local groups](#)

23.1.4 Projects

- 1) [Open Labware](#)
- 2) [Open Neuroscience](#)
- 3) [Open Source BioArt](#)
- 4) [Open Source Bioreactor](#)
- 5) [Open Source Bioreactor Project](#)
- 6) [Farma: a Home Bioreactor for Pharmaceutical Drugs](#)
- 7) [Gaudi Labs Open Culture Technology](#)
- 8) [Raspberry Pi: a versatile tool for biological sciences](#)
- 9) [Open sourcing bioinstruments](#)
- 10) [Appropedia: Open Source Technology](#)

- 11) [IFIXIT Medical Device Repair](#)
- 12) [DFROBOT](#)
- 13) [Open-source robotic platform for experimental automation](#)
- 14) [Lab On The Cheap](#)
- 15) [DIYbio Lab](#)

23.1.5 Publications

- 1) [PloS Open Source Toolkit](#)
- 2) [Open Source Hardware Journal](#)
- 3) [Open Source Biomaker](#)
- 4) [TREND Africa Open Source Training](#)

23.1.6 mHealth

- 1) [Open Data Kit](#)
- 2) [Open mHealth](#)
- 3) [Mobile Sensor Data to Knowledge](#)

24 Biomedical Image Analysis and Microscopy

24.1 E-Books and Tutorials

- 1) [Bio-image Analysis Notebooks](#)
- 2) [Python BioImage Analysis Tutorial](#)
- 3) [BIAPy: BioImage Analysis with Python](#)
- 4) [Analyzing fluorescence microscopy images with ImageJ](#)
- 5) [NEUBIAS Academy](#)
- 6) [Grand Challenge: ML in Biomedical Imaging](#)
- 7) [Carpenter Lab Public Slides](#)
- 8) [Lecture - Applied Bioimage Analysis \(2020\)](#)
- 9) [Deep learning in microscopy collection](#)
- 10) [Awesome Cytodata](#)
- 11) [CellProfiler Practical: Counting of in situ PLA signals per cell](#)
- 12) [SciPy 2021 Bioimage Analysis Fundamentals in Python](#)
- 13) [Several image processing workflow examples](#)
- 14) [Recommended Readings and Viewings](#)
- 15) [BioImage Model Zoo](#)
- 16) [deepImageJ](#)
- 17) [Bristol: Image analysis courses](#)
- 18) [Bioconductor: Basics of image data and spatial patterns analysis in R](#)
- 19) [EASY IMAGE PROCESSING IN R USING THE MAGICK PACKAGE](#)
- 20) [Bio-Image analysis with Python and Napari - DIGS-BB Light Microscopy Course 2022](#)
- 21) [First Principles of Computer Vision](#)
- 22) [Data Science in Cell Imaging](#)
- 23) [The OpenAI Microscope](#)
- 24) [The Journal of Machine Learning for Biomedical Imaging](#)
- 25) [CELL TRACKING CHALLENGE](#)
- 26) [Bioimage Analysis Course: The Life Cycle of an Image Data Set](#)
- 27) [Statistics for Neuroscientists: Image Analysis](#)
- 28) [Image Segmentation with Python](#)

24.2 BioImaging Software

- 1) [A Hitchhiker's guide through the bio-image analysis software universe](#)
- 2) [ImageScience Software](#)
- 3) [DeepLabCut](#)
- 4) [DeepLabCut GUI](#)
- 5) [Discover, install, and share napari plugins](#)
- 6) [napari](#)
- 7) [Chan Zuckerberg napari](#)
- 8) [StarDist - Object Detection with Star-convex Shapes](#)
- 9) [cellfinder-napari](#)
- 10) [ilastik](#)
- 11) [YAPiC: Yet Another Pixel Classifier \(based on deep learning\)](#)
- 12) [CellPose](#)
- 13) [CellProfiler tutorials](#)
- 14) [Piximi App](#)
- 15) [Piximi GitHub](#)
- 16) [Icy](#)
- 17) [nucleAIzer](#)
- 18) [ScientiFig](#)
- 19) [uList](#)
- 20) [ImJoy](#)
- 21) [ZeroCostDL4Mic](#)

24.3 Microscopy

- 1) [The OpenFlexure Project](#)
- 2) [NanoJ-Fluidics: open-source fluid exchange in microscopy](#)
- 3) [Microscopy Series](#)
- 4) [The Power Of Observation](#)
- 5) [Royal Microscopy Society](#)

Part VII

Recommended Readings

The recommended readings page contains articles, stories, and books that I have read and enjoyed. There is a specific focus on STEM and another general section for everyone.

STEM students

24.3.0.0.1 * 1. [Recommended Readings](#)

For Everyone

24.3.0.0.1 * 1. [Recommended Readings](#)

25 Recommended Readings (STEM)

This page contains a list of readings that I would recommend for upper-level undergraduate and graduate students getting started in STEM research. It includes articles focused on best practices, important research papers, and science integrity.

25.1 Best Practices

- 1) [How to succeed in science: a concise guide for young biomedical scientists. Part I: taking the plunge.](#) Yewdell JW, Nat Rev Mol Cell Biol. (2008).
- 2) [How to succeed in science: a concise guide for young biomedical scientists. Part II: making discoveries.](#) Yewdell JW, Nat Rev Mol Cell Biol. (2008).
- 3) [The importance of stupidity in scientific research.](#) Schwartz MA, J Cell Sci. (2008).
- 4) [Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments \(Dunning-Kruger Effect\).](#) Kruger J and Dunning D, J Pers Soc Psychol. (1999).
- 5) [How to choose a good scientific problem.](#) Alon U, Mol Cell. (2009).
- 6) [PLOS "Ten Simple Rules" Collection provides a quick, concentrated guide for mastering some of the professional challenges research scientists face in their careers.](#) PloS.
- 7) [Resources for PhD Students](#) Guillaume Dalle.

25.2 Challenges

- 1) [The Emotional Toll of Graduate School.](#) Puri P, Sci Am. (2019).
- 2) [PhDs: the tortuous truth.](#) Woolston C, Nature (2019).
- 3) [Why doing a PhD is often a waste of time](#) The Economist, Medium (2016).
- 4) [There's an awful cost to getting a PhD that no one talks about](#) Walker J, Quartz (2015).
- 5) [Strategies for Supporting Engineering Student Mental Health](#) Wilson & Goldberg, CEE (2023).
- 6) [Prioritizing mental health and wellness in engineering](#) S. Wilson, ASEE (2022)

25.3 Transformative Research Papers in Biomedical Science

- 1) [A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity](#) Jinek et al., Science (2012) from the Doudna and Charpentier Lab
- 2) [Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors](#) Takahashi et al., Cell (2006) from the Yamanaka Lab
- 3) [Induction of pluripotent stem cells from adult human fibroblasts by defined factors](#) Takahashi et al., Cell (2006) from the Yamanaka Lab
- 4) [Matrix Elasticity Directs Stem Cell Lineage Specification](#) Engler et al., Cell (2006) from the Discher Lab

25.4 Statistics

- 1) [What exactly is ‘N’ in cell culture and animal experiments?](#) Lazic SE et al., PloS Biol. (2018).
- 2) [SuperPlots: Communicating reproducibility and variability in cell biology.](#) Lord SJ et al., J Cell Biol. (2020).
- 3) [Ten common statistical mistakes to watch out for when writing or reviewing a manuscript.](#) Makin TR et al. eLife (2019).
- 4) [Beyond bar and line graphs: time for a new data presentation paradigm.](#) Weissgerber TL et al. PloS Biol. (2015).
- 5) [Statistical relevance—relevant statistics, part I.](#) Klaus B EMBO J. (2015).
- 6) [Statistical relevance-relevant statistics, part II: presenting experimental data.](#) Klaus B EMBO J. (2015).
- 7) [Ten simple rules for getting started with statistics in graduate school](#) Zitomer RA et al. PloS Comput Biol. (2022).

25.5 Writing

- 1) [ACS Guide to Scholarly Communication.](#) ACS. (2020).
- 2) [The art of writing science.](#) Plaxco KW, Prot. Sci. (2010).
- 3) [Whitesides’ Group: Writing a Paper.](#) Whitesides GM, Adv. Mat. (2004).
- 4) [How to write a first-class paper.](#) Gewin V, Nature (2018).
- 5) [Writing readable prose: When planning a scientific manuscript, following a few simple rules has a large impact.](#) Bredan AS & van Roy F, EMBO Rep. (2006).
- 6) [Ten simple rules for structuring papers](#) Mensh B & Kording K, PLoS Comp Biol. (2017).
- 7) [Write accessibly](#) Editors, Nat Biomed Eng. (2021).

25.6 Science Integrity

- 1) [Science Integrity Digest](#) Bik E
- 2) [Brookes Lab](#) Brookes P. [Sample Post](#) focused on “Fake Data”
- 3) [PubPeer: The Online Journal Club](#)
- 4) [For Better Science](#) Schneider L
- 5) [Retraction Watch](#)
- 6) [Automatic detection of image manipulations in the biomedical literature](#) Bucci EM, Cell Death and Disease (2018).
- 7) [Learning to identify image manipulations in scientific publications](#) Mazaheri G et al., arXiv (2021).
- 8) [Publishing ethics in the era of paper mills](#) Hackett R and Kelly S, Biol Open (2020).

26 Recommended Readings (General)

This page contains a list of readings that I would recommend for anyone. It includes articles with topics that range from religion, mental health, classic books and short stories.

26.1 Interesting Articles and Speeches

- 1) How Will You Measure Your Life? Clayton Christensen (2010)
- 2) A World Split Apart. Alexander Solzhenitsyn (1978)
- 3) The Inner Ring. CS Lewis (1944)

26.2 My Favorite Short Stories

- 1) How Much Land Does a Man Require? Leo Tolstoy (1886)
- 2) The Americanization of Kambili. Tochi Eze (2022) [Link](#)
- 3) The Street Sweep. Meron Hadero (2018) [Link](#)

26.3 Classics

- 1) Night (Elie Wiesel)
- 2) Gilead (Marilynne Robinson)

26.4 Mental Health / Counseling / Peace of Mind

- 1) When People Are Big and God Is Small: Overcoming Peer Pressure, Codependency, and the Fear of Man (Edward T. Welch)
- 2) Seeing with New Eyes: Counseling and the Human Condition Through the Lens of Scripture (David A. Powlison)
- 3) Man's Search for Meaning (Viktor E. Frankl)
- 4) Absence of Mind: The Dispelling of Inwardness from the Modern Myth of the Self (Marilynne Robinson)

26.5 Religious

- 1) The Reason for God: Belief in an Age of Skepticism (Timothy J. Keller)
- 2) The Cost of Discipleship (Dietrich Bonhoeffer)
- 3) A Treatise of Earthly-Mindedness (Jeremiah Burroughs)
- 4) The Knowledge of the Holy (A.W. Tozer)
- 5) The Pursuit of God: The Human Thirst for the Divine (A.W. Tozer)
- 6) A Bunch of Everlastings; Or, Texts That Made History (F.W. Boreham)
- 7) A Packet of Surprises: The Best Essays and Sermons of F. W. Boreham (F.W. Boreham)
- 8) The Screwtape Letters (C.S. Lewis)
- 9) The Last Sin Eater (Francine Rivers)
- 10) Mere Christianity (C.S. Lewis)