

Table of Contents

Table of Contents	1
Terminals	2
Free Text editors & IDEs	2
Free Python Courses/Tutorials	3
Bootcamps	4
Math Problem solving with Python	4
Data Science Courses/Tutorials	4
Pandas Tutorials	4
Python Packages	5
R Packages	5
Data Science Books	5
Data Resources	6
Biology-focused Data Science Courses/Tutorials	9
Books/Texts/Websites	9
Biology-focused Books (Bioinformatics/Computational Biology)	10
Books to learn Data Visualization	13
Statistics books	13
Online Tutors	14
Hackathons/Competitions/Paired-Programming opportunities	14
Teaching materials for school kids	15
Conferences	15
Intermediate Learning	15
Free Cheatsheets	15
Free Videos	16
Tutors	16
Pay for Courses	16
Pay for Texts/Websites	16
Slack Channels to get help	17
Meetup Groups	17
To make a website	17
Mock Interviews	17
Tools	18
Scrapers	18

Terminals

1. iTerm2 (works on Macs with macOS 10.10 or newer, Check out the impressive features and screenshots. If you spend a lot of time in a terminal, then you'll appreciate all the little things that add up to a lot. It is free)
 - a. <https://www.iterm2.com/downloads.html>
2. tmux

Free Text editors & IDEs

1. IDE = **integrated development environment** = a piece of software that acts as text editor, debugger and compiler all in one
 - a. Note: you should choose an IDE that works with multiple languages
 - b. See popularity graph:
<https://insights.stackoverflow.com/survey/2018#technology-development-environments>
2. Vim
 - a. “Steeper learning curve than Emacs, but great to work with” (Josh C of www.meetup.com/codeselfstudy)
 - b. How to turn Vim into an IDE:
<https://arstechnic.ca.com/information-technology/2009/05/vim-made-easy-how-to-get-your-favorite-ide-features-in-vim/>
 - c. Instructions to set Vim up to work well with Python:
<https://realpython.com/vim-and-python-a-match-made-in-heaven/>
 - d. Why Vim: <https://coderoncode.com/tools/2017/04/16/vim-the-perfect-ide.html>
 - e. NeoVim:
 - i. Set up NeoVim on Mac for R: http://blog.jom.link/neovim_R_mac.html
 - f. VimR: <http://vimr.org/>
3. Emacs
 - a. “Easier learning curve, but less efficient than Vim” (Josh C of www.meetup.com/codeselfstudy)
4. Visual Studio IDE or Visual Studio Code (VSCode) <https://code.visualstudio.com/> (Microsoft)
 - a. Note: the IDE comes in 3 levels: Community, Professional, Enterprise. The latter two cost tons of [money](#)! The community version lacks features.
 - b. Python extension
<https://marketplace.visualstudio.com/items?itemName=ms-python.python>

- c. Recommended by FreeCodeCamp
<https://medium.freecodecamp.org/an-overview-of-visual-studio-code-for-front-end-developers-49a4aa0771fb>
- d. Recommendation:
<https://hackernoon.com/virtualstudio-code-the-editor-i-didnt-think-i-needed-16970c8356d5>
- 5. NetBeans <https://netbeans.org/> (Oracle)
- 6. Eclipse <http://www.eclipse.org/> (Made by IBM, owned by Eclipse Foundation, created in January 2004, an independent not-for-profit corporation)
- 7. Atom <https://ide.atom.io/> (Git/GitHub)
- 8. Sublime (a pop up window will ask you if you want to pay \$80, just close it & keep using for free)
 - a. Make Sublime a perfect IDE for R
<https://github.com/REditorSupport/sublime-ide-r>
- 9. Jupyter Notebooks - not as a text editor, but a notebook that can hold code

Free Python Courses/Tutorials

- 1. <http://exercism.io/> (Recommended by see github "seanfarley")
- 2. EdX (Recommended by Sean Farley)
 - a. MIT: Introduction to Computer Science and Programming Using Python (recommended by Josh C of www.meetup.com/codeselfstudy)
 - i. <https://www.edx.org/course/introduction-computer-science-mitx-6-00-1x-11>
 - ii. Josh says "The instructor Eric Grimson has a great way of presenting info"
- 3. For Beginners
 - a. <https://www.pythomn.org/about/gettingstarted/>
- 4. Interactive Python, Intro to Programming: (Intermediate)
 - a. <http://interactivepython.org/runestone/static/thinkcspy/index.html>
- 5. Interactive Python, Algorithms and Data Structures (Intermediate to Advanced)
 - a. <http://interactivepython.org/runestone/static/pythonds/index.html>
- 6. Rithm School Free Courses (web development, algorithms) (Intermediate to Advanced)
 - a. <https://www.rithmschool.com/courses#python>
- 7. PyBites Code Challenges (Intermediate to Advanced)
 - a. <https://codechalleng.es/challenges/?next=/>
- 8. Free Code Camp
 - a. <https://www.freecodecamp.org/>
- 9. Pandas - Python for data analysis
 - a. <https://pandas.pydata.org/>
 - b. <https://www.udemy.com/python-for-data-science-and-machine-learning-bootcamp/> (ES)

- c. Top 8 resources for learning data analysis with pandas
 - i. <https://www.dataschool.io/best-python-pandas-resources/>
- 10. Guru99.com
- 11. Real Python tutorials
 - a. <https://realpython.com/python-introduction/>
- 12. General Assembly (their HTML/CSS and JavaScript course is really good)
- 13. Codecademy
- 14. Parallelization (for help, ask Michelle Dong, studying CS at USF)
 - a. Python is used as a wrapper
 - b. (like Hadoop?)
- 15. Pluralsight
 - a. <https://www.pluralsight.com/>

Bootcamps

1. Lambda school, 9 hours per day, 30 weeks, full stack, remote
2. Metis
3. General Assembly

Math Problem solving with Python

1. Project Euler: <https://projecteuler.net/>
 - a. Project Euler is a series of challenging mathematical/computer programming problems that will require more than just mathematical insights to solve. Although mathematics will help you arrive at elegant and efficient methods, the use of a computer and programming skills will be required to solve most problems.

Data Science Courses/Tutorials

1. 9 free courses from <https://www.datacamp.com/pricing>
2. Software Carpentry for Data Science <https://software-carpentry.org> & <https://github.com/swcarpentry>
3. An online community for showcasing R & Python tutorials: <https://datascienceplus.com/>
4. Hadoop, Hive, Presto, Spark
5. Data Science masters (open source)
 - a. <https://github.com/datasciencemasters/go>

Pandas Tutorials

1. Brandon Rhodes - Pandas From The Ground Up - PyCon 2015 (Don't bother with the .zip file, just listen, Recommended by Chris Stehlik of Sunday's meetup)
 - a. <https://www.youtube.com/watch?v=5JnMutdy6Fw>

Python Packages

1. Presto
 - a. <https://pypi.org/project/presto-python-client/>
 - b. <https://prestodb.io/>

R Packages

1. Tidyverse: R packages for data science
 - a. <https://www.tidyverse.org/>

Data Science Books

1. Pandas for Everyone: Python Data Analysis (Addison-Wesley Data & Analytics Series), Dec 2017
 - a. <https://www.safaribooksonline.com/library/view/pandas-for-everyone/9780134547046/>
2. R books to learn Data Science (recommended by NYC Open statistical programming Meetup - all these meetings are live streamed here: <https://nyhackr.org/presentations.html>) (Note Drew Conway if the grandfather of Data Science) (Joseph J Allaire, the head of R Studio: https://en.wikipedia.org/wiki/Joseph_J._Allaire)
 - a. Introduction to R
 - i. "R for Everyone" by Jared Lander
 - b. Statistics books
 - i. Andrew Gelman's books
 1. <http://www.stat.columbia.edu/~gelman/books/>
 2. Andrew Gelman's Blog: <http://andrewgelman.com/>
 - ii. "The Elements of Statistical Learning: Data Mining, Inference, and Prediction" by Trevor Hastie, Robert Tibshirani, Jerome Friedman
 - iii. <https://web.stanford.edu/~hastie/Papers/ESLII.pdf>
 - iv. "An Introduction to Statistical Learning: with applications in R" by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani

1. <http://www-bcf.usc.edu/~gareth/ISL/>
- v. “Data Analysis Using Regression and Multilevel/Hierarchical Models” by Andrew Gelman & Jennifer Hill
 1. <http://www.stat.columbia.edu/~gelman/arm/>
- c. Machine Learning
 1. Linear Regression
 2. Penalized Regression
 3. “Applied Predictive Modeling” by Max Kuhn and Kjell Johnson
 - a. Will give you intro to Caret
 - b. https://vuquangnguyen2016.files.wordpress.com/2018/03/applied-predictive-modeling-max-kuhn-kjell-johnson_1518.pdf
 - c. <http://appliedpredictivemodeling.com/>
 4. “Deep Learning with R” by Joseph J Allaire
 - a. This book covers Tensorflow (<https://www.tensorflow.org/>) and Keras (<https://keras.rstudio.com/>)

Data Resources

1. Kaggle has a page linking to other data
2. Gapminder.org
3. More from The Data Incubator (in Oakland) see list below
 - a. <https://blog.thedataincubator.com/2014/10/data-sources-for-cool-data-science-projects-part-1/>
 - b. <https://blog.thedataincubator.com/2014/10/data-sources-for-cool-data-science-projects-part-2/>
4. Great data sources from Insight:
 - a. <https://docs.google.com/document/d/1kEg5oJj9lfCVO-Bf48MvYmCumPu3Thly3AmhHAOd55M/edit?usp=sharing>

Economic Data:

1. **Publicly Traded Market Data:** [Quandl](#) is an amazing source of finance data. [Google Finance](#) and [Yahoo Finance](#) are additional good sources of data. Corporate filings with the SEC are available on [Edgar](#).
2. **Housing Price Data:** You can use the [Trulia API](#) or the [Zillow API](#). In the UK, you can find [price paid in house sales](#) and historical [mean house price by region](#) (use [this tool](#) to translate between postcode and lat/long).

3. **Lending Data:** You can find [student loan defaults by university](#) and the complete collection of peer-to-peer loans from [Lending Club](#) and [Prosper](#), the two largest platforms in the space.
4. **Home Mortgage Data:** There is data made available by the [Home Mortgage Disclosure Act](#) and there's a lot of data from the [Federal Housing Finance Agency available here](#).

Content Data:

1. **Review Content:** You can get reviews of restaurants and physical venues from Foursquare and Yelp (see geodata). Amazon has a large repository of [Product Reviews](#). Beer reviews from Beer Advocate can be found [here](#). Rotten Tomatoes [Movie Reviews](#) are available from Kaggle.
2. **Web Content:** Looking for web content? Wikipedia provides [dumps of their articles](#). Common Crawl has a [large corpus of the internet available](#). ArXiv maintains all their data available via [Bulk Download from AWS S3](#). Want to know which URLs are malicious? There's a [dataset](#) for that. Music data is available from the [Million Songs Database](#). You can analyze the Q&A patterns on sites like [Stack Exchange \(including Stack Overflow\)](#).
3. **Media Data:** There's open annotated articles from the [New York Times](#), [Reuters Dataset](#), and [GDELT project](#) (a consolidation of many different news sources). Google Books has [published NGrams](#) for books going back to past 1800.
4. **Communications Data:** There's access to public messages of the [Apache Software Foundation](#) and communications [amongst former execs at Enron](#).

Government Data:

1. **Municipal Data:** Crime Data is available for [City of Chicago](#) and [Washington DC](#). Restaurant Inspection Data is available for [Chicago](#) and [New York City](#).
2. **Transportation Data:** [NYC Taxi Trips in 2013 are available courtesy of the Freedom of Information Act](#). There's bikesharing data from [NYC](#), [Washington DC](#), and [SF](#). There's also [Flight Delay Data from the FAA](#).
3. **Census Data:** [Japanese Census Data](#). US Census data from [2010](#), [2000](#), [1990](#). From census data, the government has also derived [time use data](#). [EU Census Data](#). Check out [popular male / female baby names going back to the 19th Century](#) from the Social Security Administration.
4. **World Bank:** They have a lot of data available [on their website](#).
5. **Election Data:** Political contribution data for the last few US elections can be downloaded from the FEC [here](#) and [here](#). Polling data is available from [Real Clear Politics](#).

6. **Food, Drugs, and Devices Data:** The USDA provides location-based information about the food environment in their [Food Atlas](#). The FDA also provides a number of high value [public datasets](#).

Data With a Cause:

1. **Environmental Data:** Data on [household energy usage is available](#) as well as [NASA Climate Data](#).
2. **Medical and Biological Data:** You can get anything from [anonymous medical records](#), to remote sensor reading [for individuals](#), to data on the Genomes of [1000 individuals](#).

Miscellaneous:

1. **Geo Data:** Try looking at these Yelp Datasets for [venues near major universities](#) and one for major cities in the [Southwest](#). The [Foursquare API](#) is another good source. Open Street Map has open [data on venues](#) as well.
2. **Twitter Data:** You can get access to [Twitter Data](#) used for sentiment analysis, [network Twitter Data](#), and [social Twitter data](#), on top of their [API](#).
3. **Games Data:** Datasets for games, including a large dataset of [Poker hands](#), dataset of [online Dominion Games](#), and datasets of [Chess Games](#) are available. Boardgamegeek.com also has a large [database](#) of games, prices, artists, etc.
4. **Web Usage Data:** Web usage data is a common dataset that companies look at to understand engagement. Available datasets include [anonymous usage data for MSNBC](#), [Amazon purchase history](#) (also anonymized), and [Wikipedia traffic](#).

Metasources: these are great sources for other web pages.

1. Stanford Network Data: <http://snap.stanford.edu/index.html>
2. Every year, the ACM holds a competition for machine learning called the KDD Cup. Their data is [available online](#).
3. UCI maintains [archives of data for machine learning](#).
4. [US Census Data](#).
5. Amazon is hosting [Public Datasets on s3](#).
6. Kaggle hosts machine-learning challenges and many of their datasets are [publicly available](#).
7. The cities of [Chicago](#), [New York](#), [Washington DC](#), and [SF](#) maintain public data warehouses.
8. Yahoo maintains a lot of data on its web properties [which can be obtained by writing them](#).

9. [BigML](#) is a blog that maintains a list of public datasets for the machine learning community.
10. [GroupLens Research](#) has collected and made available rating data sets from the MovieLens website.

Biology-focused Data Science Courses/Tutorials

1. Use Python for Research, with libraries such as Numpy and SciPy:
<https://www.edx.org/es/course/using-python-research-harvardx-ph526x-0> (Harvard professor, recommended by Mariana (Madu on Slack))
2. **Rosalind**, a platform for learning bioinformatics and programming through problem solving <http://rosalind.info/problems/locations/>
3. Biopython: Python Tools for Computational Molecular Biology <http://biopython.org>
4. Books, courses, & tutorials <https://pythonforbiologists.com/>
5. [DTU course Python and Unix for Bioinformaticians](#), last modified April 2018
6. [Biostar Handbook + courses](#)
 - a. Course: Python Programming in 100 hours » How to use Python for bioinformatics data analytics.
 - b. Course: Learn Bioinformatics in 100 hours » An Effective Introduction to Bioinformatics Data Analysis.
 - c. See [Development Timeline](#) - last updated in 2017

Books/Texts/Websites

1. Python tutorial (Recommended by Holberton)
 - a. See Index of Chapters here: <https://docs.python.org/3/tutorial/index.html>
2. **Real Python** <https://realpython.com/> (For Python beginners or intermediates, through new users with other programming experience)
3. **Automate the Boring Stuff** (recommended by Josh C of www.meetup.com/codeselfstudy, PyLadies has a donated book. Teaches how to work with csv files)
 - a. text here: <https://automatetheboringstuff.com/>
4. Front end webpage building (light)
 - a. <http://surge.sh/> (Recommended by Josh C of code self study Berkeley MeetUp)
5. **Python Playground** (recommended by a PyLadies mentor: they use some of these questions for job interviews at FB, dropbox, etc!)
6. Learn Python3 the Hard Way, by Zed A. Shaw (recommended against by who?)

7. **Python Crash Course:** Hands on Project Based introduction to programming, by Eric Matthes (recommended highly, the author is a high school teacher)
8. Cracking Cookies with Python (PyLadies has a donated book)
9. [Data Science and Analytics With Python](#), by Jesus Rogel-Salazar, 2017 (recommended by Michael Deutscher of Code Self Study)
10. Learn Python in 1 Day (Deena used this one to start, it took me weeks to work through, but I liked it & feel it was a fast way to learn the basics)
 - a. <http://www.learncodingfast.com/python>
 - b. <https://projecteuler.net/>
 - c. Author: jamie@learncodingfast.com.
11. Learning Python: Powerful Object-Oriented Programming, by Mark Lutz (O'Reilly)
12. Python in a Nutshell, by Alex Martelli (O'Reilly)
13. Python for Beginners: A Crash Course Guide to Learn Python in 1 week (coding, programming, web-programming, programmer)
14. Fluent Python - good for intermediate level (Recommended by Rahul's boss at Cruise)
15. Clean Architecture (2015) - lots of examples, symptoms
16. Design patterns (1994) - old, don't bother reading; [criticism](#)

Biology-focused Books (Bioinformatics/Computational Biology)

1. Python
 - a. Python for Bioinformatics, [Second Edition](#), by Sebastian Bassi (Chapman and Hall/CRC, 2017) \$64 on Kindle
 - i. updated throughout to Python 3 (note, 1st edition, Python2, was 2016)
 - ii. a solid introduction to programming with Python, making the book accessible for readers without previous programming experience
 - iii. Covers advanced topics, such as web applications with Bottle, databases (MySQL, SQLite and MongoDB), XML and Bokeh
 - iv. include NoSQL databases, the Anaconda Python distribution, graphical libraries like Bokeh
 - b. [Biostar Handbook + courses](#)
 - i. Course: Python Programming in 100 hours » How to use Python for bioinformatics data analytics.
 - ii. Course: Learn Bioinformatics in 100 hours » An Effective Introduction to Bioinformatics Data Analysis.
 - iii. See [Development Timeline](#) - last updated in 2017
 - c. Python For The Life Sciences (completed on 2017-02-07) \$35
 - i. <https://leanpub.com/pythonforthelifesciences>
 - ii. [TOC](#)
 - d. A Primer for Computational Biology, by Shawn T. O'Neil, (Only available in paper, Dec 2017)

- i. The text covers accessing and using remote servers via the command-line, writing programs and pipelines for data analysis, and provides useful vocabulary for interdisciplinary work. The book is broken into three parts: Unix/Linux, Python, R
 - ii. (recommended by Michael Deutscher of Code Self Study)
- e. Bioinformatics with Python Cookbook, by Tiago Antao (2015)
 - i. Learn how to use modern Python bioinformatics libraries and applications to do cutting-edge research in computational biology
 - ii. see [TOC & What you will learn here](#)
 - iii. also available in a version for R
- f. Python Programming for Biology: Bioinformatics and Beyond, by Tim J. Stevens and Wayne Boucher (2015)
 - i. Do you have a biological question that could be readily answered by computational techniques, but little experience in programming? Do you want to learn more about the core techniques used in computational biology and bioinformatics? Written in an accessible style, this guide provides a foundation for both newcomers to computer programming and those interested in learning more about computational biology. The chapters guide the reader through: a complete beginners' course to programming in Python, with an introduction to computing jargon; descriptions of core bioinformatics methods with working Python examples; scientific computing techniques, including image analysis, statistics and machine learning.
- g. Computing for Biologists: Python Programming and Principles, by by Ran Libeskind-Hadas and Eliot Bush (2014)
 - i. Computing is revolutionizing the practice of biology. This book, which assumes no prior computing experience, provides students with the tools to write their own Python programs and to understand fundamental concepts in computational biology and bioinformatics. Each major part of the book begins with a compelling biological question, followed by the algorithmic ideas and programming tools necessary to explore it: the origins of pathogenicity are examined using gene finding, the evolutionary history of sex determination systems is studied using sequence alignment, and the origin of modern humans is addressed using phylogenetic methods. In addition to providing general programming skills, this book explores the design of efficient algorithms, simulation, NP-hardness, and the maximum likelihood method, among other key concepts and methods. Easy-to-read and designed to equip students with the skills to write programs for solving a range of biological problems, the book is accompanied by numerous programming exercises, available at www.cs.hmc.edu/CFB.
 - ii. [book review](#)

- h. Illustrating Python via Examples from Bioinformatics:
<http://hplgit.github.io/bioinf-py/doc/pub/html/index.html>
 - i. Books, courses, & tutorials <https://pythonforbiologists.com/>
 - j. [An Introduction to Programming for Bioscientists: A Python-Based Primer](#), (PLoS Comp Bio, 2014)
 - k. Introduction to Python, Center for Computational Biology and Bioinformatics, University of Texas at Austin, http://sjspielman.org/bdib2016_python/ (Very beginner)
 - l. Bioinformatics Programming Using Python: Practical Programming for Biological Data (Animal Guide) by Mitchell L Model (O'Reilly, 2009) ([free pdf](#))
 - m. Python for Bioinformatics, Second Edition (Chapman & Hall/CRC Mathematical and Computational Biology) (2009)
 - n. Bioinformatics Programming in Python: A Practical Course for Beginners, by Ruediger-Marcus Flaig (2008)
2. R
- a. Bioinformatics Data Skills: Reproducible and Robust Research with Open Source Tools, by Vince Buffalo (O'Reilly, 2015)
 - i. Experience at UC Davis Genome Center
 - ii. [Summary](#): Unix, then R
 - b. R for Data Science: Import, Tidy, Transform, Visualize, and Model Data
 - c. R Cookbook: Proven Recipes for Data Analysis, Statistics, and Graphics (O'Reilly)
 - d. Statistics and Data Analysis for Microarrays using R and Bioconductor, Second Edition (Chapman & Hall/CRC Mathematical and Computational Biology)
 - e. RNA-seq Data Analysis: A practical Approach (Chapman & Hall/CRC Mathematical and Computational Biology) {demonstrates command-line tools, R, and other open source tools, such as the graphical Chipster software}
 - f. Data Computing: an introduction to wrangling and visualization with R, by Daniel Kaplan (Good Exercises at the back, some are bio related, well written)
3. Command line:
- a. Practical Computing for Biologists, by Steven Haddock and Casey Dunn (Sinauer, 2010)
4. Not sure which programming language or multiple languages
- a. A list of ~100 books (but there are not publication dates :(
 - i. <https://www.iscb.org/iscb-publications-bioinformatics-review/35-ISCB%20Publications/ISCB%20Publications/125-booklistaikchoontan>
 - b. A List of Jupyter Notebooks
 - i. <https://github.com/jupyter/jupyter/wiki/A-gallery-of-interesting-Jupyter-Notebooks>
 - c. A List of open source bioinformatics software
 - i. https://en.wikipedia.org/wiki/List_of_open-source_bioinformatics_software

- d. Bioinformatics: A Practical Handbook of Next Generation Sequencing and Its Applications (2017) {Linux, de novo genome assembly, metagenomics, not clear if they use a specific programming language}
- e. Bioinformatics and Functional Genomics, 3rd edition, 2013, [eTextbook \\$100](#)
- f. An Introduction to Bioinformatics Algorithms (Computational Molecular Biology) (MIT press, 2004) - doesn't use a specific programming language
- g. Bioinformatics for Beginners: Genes, Genomes, Molecular Evolution, Databases, and Analytical Tools (2014) {not clear which programming language}
- h. Understanding Bioinformatics (2007)
- i. Applied Bioinformatics: An introduction (2018, updated edition from 2003/8, Includes concrete examples with exercises and solutions), doesn't look good
- j. Deep Learning <http://www.deeplearningbook.org/> {a resource intended to help students and practitioners enter the field of machine learning in general and deep learning in particular. The online version of the book is now complete and will remain available online for free}
- k. Introducing Epigenetics: A Graphic Guide (2017) {a non-fiction comic book, no programming}
- l. Essential Bioinformatics (2006) {free online}
- m. Biodata Mining and Visualization (2010) {don't bother, more recent books are better}

Books to learn Data Visualization

1. A list maintained by <http://bibrainz.com/aof/2017/11/07/essential-list-of-data-visual-storytelling-resources/>
- 2.

Statistics books

1. [Discovering Statistics Using R](#), by Andy Field
 - a. good images, good humor
 - b. Recommended by Ian of Code Self-Study
- 2.

Data Viz Software Programs

1. Power BI (ranked #1 by Mico Yuk of Metis, due to its ease of use, flexibility, mobile capability, speed of innovation, end to end capabilities with the integration to the MSFT stack) this is by far the most promising and our top choice. It is also my favorite for authored visualizations.
2. Tableau (ranked #2 by Mico Yuk of Metis, due to its extended visual capabilities, mobile accessibility and quick adoption by users). It is however quite pricey and I'm waiting to see where they land on the more scientific end of the spectrum. They are making

investments. It is a stand alone tool though so there is no end to end. Def a toy for those who enjoy 'data discovery' nirvana! Great for data discovery, slow on AI.

3. Excel and Qlikview (ranked equally at #3 by Mico Yuk of Metis) Qlikview is basically one of the best out there for reporting type visualizations. I'm not sure today what their 'scientific' capabilities are but they are an end to end system and you can almost any visual done with some SQL code.

Data Viz Cheatsheets

1. A poster with 98 rules of the SUCCESS formula for IBCS® compliant reports and presentations <http://www.ibcs.com/product/ibcs-success-poster/>

Online Tutors

1. See how a script you run goes through the steps that are nested within different classes of your code. It is web-based:
 - a. <http://pythontutor.com/>

Hackathons/Competitions/Paired-Programming opportunities

1. Hackathons
 - a. DataKind
 - i. <https://www.meetup.com/DataKind-SF-Bay-Area/>
 - b. Girls in Tech
 - i. <http://girlsintech.org/programs/hacking-for-humanity/>
 - c. Lists of hackathons in SF/Bay Area
 - i. <https://www.eventbrite.com/d/ca--san-francisco/hackathon/>
 - ii. <https://www.hackathon.com/city/united-states/san-francisco-bay-area>
 - d. Science hackathons
 - i. <http://sf.sciencehackday.org/> (Oct 27-28, 2018)
 - e. Code 4 Good
 - i. <https://c4gw.glitch.me/>
2. Paired-Programming
 - a. Chingu: do a project with other people in a team of about 3 people. Aim: have an MVP in 2 months. Projects depend on
 - i. <https://chingu.io/apply> (you need some experience building stuff before joining)

- ii. They have a data science subgroup & a lot of resources
- 3. Kaggle
 - a. <https://www.kaggle.com/>
 - b. Kaggle is the place to do data science projects
 - i. It is a platform for predictive modelling and analytics competitions in which statisticians and data miners compete to produce the best models for predicting and describing the datasets uploaded by companies and users

Teaching materials for school kids

- 1. Zero Libraries
 - a. PyGame zero <https://pygame-zero.readthedocs.io/en/stable/>
 - b. Network zero
 - c. GPIO zero <https://www.raspberrypi.org/forums/viewtopic.php?t=122071>

Conferences

- 1. Open Data Science Conference
 - a. <https://odsc.com/california>

Intermediate Learning

- 1. Sara Packman - The Journey Over the Intermediate Gap - PyCon 2018
 - a. <https://www.youtube.com/watch?v=49Cllu1XkIE&feature=youtu.be>

Free Cheatsheets

- 1. Learn X in Y minutes
 - a. <https://learnxinyminutes.com/>
 - i. Python3, R, git, Statistical Computing with Python, tmux, vim, matlab, markdown, awk ...
 - b. Python3
 - i. <https://learnxinyminutes.com/docs/python3/>
 - ii. https://github.com/ehmatthes/pcc/tree/master/cheat_sheets

Free Videos

1. Indentation errors (Recommended by Holberton)
 - a. <https://www.youtube.com/watch?v=1QXOd2ZQs-Q>

Tutors

1. Python Tutor.com <http://pythontutor.com/>
 - a. It will tell you what every line of your code is doing
 - b. Also has messages with real people who can answer your questions

Pay for Courses

1. <https://www.udemy.com/topic/python/> (Recommended by many)
 - a. Tip - wait for courses to go on sale for \$9.99 or \$11.00. Sales happen ~1x every 2 weeks
 - b. Courses taken by peers
 - i. “Complete Python Bootcamp: Go From Zero To Hero” instructed by Jose Portilla (Christian Arrieta of Sunday Holberton 10-week group is taking this course) (recommended by Claudia)
2. <https://www.datacamp.com/#> (Personal choice by Ting)
 - a. Interactive exercises with video instructions. Great for beginners who would like to repeat the courses for refresher
 - b. Advance courses could be harder to learn independently
3. <https://www.dataquest.com/> (Recommend by Esther)
 - a. Data Analyst, Data Scientist and Data Engineering tracks

Pay for Texts/Websites

1. Programming Collective Intelligence (recommended by Josh C of www.meetup.com/codeselfstudy)
 - a. <http://shop.oreilly.com/product/9780596529321.do>
 - b. Building Smart Web 2.0 Applications “tap the power behind search rankings, product recommendations, social bookmarking, and online matchmaking? This fascinating book demonstrates how you can build Web 2.0 applications to mine the enormous amount of data created by people on the Internet.”
2. Front end website building (heavy)

- a. <https://www.netlify.com/> (Recommended by Josh C of Code Self Study)
- b. <https://www.digitalocean.com/community/tutorials/> (Recommended by Holberton)

Slack Channels to get help

1. <https://techbeacon.com/46-slack-groups-developers>
2. <https://codenewbie.typeform.com/to/uwsWIZ>
3. <https://devchat.devolio.net/>
4. Slackin.pyladies.com
 - a. allows you to join the pyladies slack channel

Meetup Groups

1. Local groups
 - a. Girls in Tech
 - b. Girl Develop It
 - c. PyLadies
 - d. Women in Machine Learning and Data Sci (WIMLDS)
2. Non-local groups that have free videos
 - a. Montreal <https://montrealpython.org/en/>
 - b. NYC (Data sci) <https://nyhackr.org/presentations.html>

To make a website

1. Django girls
 - a. <https://djangogirls.org/resources/>
2. Miguel Grinberg Flask tutorial
 - a. <https://blog.miguelgrinberg.com/post/the-flask-mega-tutorial-part-i-hello-world>

Mock Interviews

1. Pramp: <https://www.pramp.com/#/>

Tools

Scrapers

- [pdf scraper Tabula](#) (Don't be fooled by click bate: taboola)
-