

Training Resources & Finding Support

General

Resources for Python

Resources for R

GitHub resources

Code development

Troubleshooting

Community-based support

Manuals

Miscellaneous sources

Blogs

Podcasts

General

Please check for yourself the pricing structure of the courses below - it's often the case that the introductory courses are free, but you have to pay a subscription or a course fee for more advanced learning.

General Data Science courses

- [Coursera](#)
- [edX](#)
- [EMC](#)
- [Microsoft Professional Programme](#)

This GDS Academy course is for anyone interested in understanding what AI is, how it's being used in government now and how to get started with AI.

- [Introduction to Artificial Intelligence](#)

Udacity Data Science Courses - often delivered in collaboration with industry (like Google, Facebook etc) so are often very applicable to real-world situations

- <https://www.udacity.com/courses/data-science>

Intro post to David Venturi, who created his own data science masters

- <https://medium.com/@davidventuri/i-dropped-out-of-school-to-create-my-own-data-science-master-s-here-s-my-curriculum-1b400dcee412>

David Venturi's blog and website

- <https://medium.com/@davidventuri>
- <http://davidventuri.com/data-science-masters#scroll>

Data School - useful youtube videos and helpful links to other content

- <http://www.dataschool.io/>

Data Science Masters - includes reading suggestions and online training exercises

- <http://datasciencemasters.org/>

Coursera - machine learning course (foundation course)

- <https://www.coursera.org/learn/machine-learning>

Coursera - Andrew Ng's deep learning specialization (intermediate-advanced, requires commitment)

- <https://www.coursera.org/specializations/deep-learning>

Trello - a free useful project management tool to organise your tasks in an agile manner. There are other similar websites available.

- <https://trello.com>

Resources for Python

A comprehensive list of resources for learning Python, collated by Data Scientists in government

- <https://khub.net/documents/412856979/0/Python+learning+resources.pdf/0970e0c5-2146-c39d-17fd-ee36df68b624?t=1611920511450>

Coursera - python course

- <https://www.coursera.org/learn/python>

Open University - python course 'Learn to Code for Data Analysis'

- <https://www.futurelearn.com/courses/learn-to-code>

Codecademy - introduction to Python

- <https://www.codecademy.com/learn/python>

Datacamp - introduction to Python for Data Science

- <https://www.datacamp.com/courses/intro-to-python-for-data-science>

Pandas Cookbook

- <https://github.com/jvns/pandas-cookbook>

Read the docs

- <https://docs.python.org/3.6/>

Resources for R

A comprehensive list of resources for learning R, collated by Data Scientists in government (pdf attached)

- <https://khub.net/documents/412856979/0/R+learning+resources.pdf/77db35c8-17af-6f02-adf6-2e1d38a0d7be?t=1611920497935>

A collection of free R programming books

- <https://bookdown.org/>

O'Reilly book on data science in R

- <http://r4ds.had.co.nz/>

DataCamp - free introductory courses on R and Python, paid subscription for more advanced courses

- <https://www.datacamp.com/>

GitHub resources

Git is a version control system (VCS) for tracking changes in source code during development.

- Coordinate work among project contributors
- Track changes in any set of files

GitHub is a web-based hosting service for version control using Git, use it to:

- Version control and backup your project code
- Make collaboration with your mentor easier
- Revert and course-correct when things go wrong

Quickstart

- [Installing git](#)
- [Set up git](#)
- [Authenticating to GitHub / Generating a new SSH key and adding it to the ssh-agent](#)

[GitHub Tutorial GitHub Cheat Sheet](#)

Introduction to using Git

- <https://try.github.io/levels/1/challenges/1>

Code development

Integrated Development Environments (IDEs) I strongly recommend using Pycharm for Python code development. Very intuitive.

- <https://www.jetbrains.com/pycharm/>

Text editors

- [vim](#)
- [emacs](#)

Terminal app of choice

- <https://www.terminal2.com/>

Troubleshooting

Overall, the easiest thing to do is to either:

- Copy and paste any errors you may encounter, or
- Provide a short text description of the problem you want to solve (problem formulation becomes very important in programming/data science) into google and let it do its magic.

Community-based support

- [Stack Overflow](#) - question and answer site for professional and enthusiast programmers
- [Cross Validated](#) - question and answer site for people interested in statistics, machine learning, data analysis, data mining, and data visualization
- [Kaggle](#) - find and publish data sets, explore and build models, work with other data scientists and machine learning engineers, and enter competitions to solve data science challenges
- [GitHub](#) - search through issues and pull requests of relevant projects

Manuals

Refer to language or package/library specific API documentation. It will usually be hosted at <https://readthedocs.org/> or GitHub-based.

Quick reference, [man-style](#)

- `?function()`
- `man function`
- `function --help`

Miscellaneous sources

Blogs

- <https://hackernoon.com/>
- <https://medium.com/topic/technology>
- <https://machinelearningmastery.com/blog/>
- <https://www.kdnuggets.com/>

Podcasts

Some useful data science podcasts which help with conceptual understanding eg, explaining how algorithms work.

- [Data Skeptic](#)
- [The Talking Machines](#)
- [Not So Standard Deviations](#)
- [Partially Derivative](#)
- [Datacamp pod](#)