**Useful resources**[Help](https://class.coursera.org/bigdataschool-001/help/pages?url=https%3A%2F%2Fclass.coursera.org%2Fbigdataschool-001%2Fwiki%2FUseful_resources)

Here are some general resources that you may find useful for a further study.  This list is by no means complete.  We welcome additional suggestions through the appropriate discussion thread.

**Some other on-line classes and tutorials in this arena:**

* [Yaser Abu-Mostafa's Caltech class "Learning From Data"](https://work.caltech.edu/telecourse.html)
* [Andrew's Moore statistics and data mining tutorial](http://www.autonlab.org/tutorials/)
* [Bill Howe's "Introduction to Data Science" class on Coursera](https://www.coursera.org/course/datasci)
* [JHU's Data Sciences classes on Coursera](https://www.coursera.org/specialization/jhudatascience/1/courses)
* [Courses offered at IACS at Harvard](http://iacs.seas.harvard.edu/courses-websites)
* KDnuggets has a [good list](http://www.kdnuggets.com/education/online.html)

**General e-Science / Data Science resources:**

* ["The Fourth Paradigm"](http://research.microsoft.com/en-us/collaboration/fourthparadigm/)(free book download from Microsoft Research)
* Gregory Piatetsky-Shapiro's [KDNuggets](http://www.kdnuggets.com/)
* Andy Pryke's [The Data Mine](http://www.the-data-mine.com/)
* Fionn Murtagh's [Multivariate Data Analysis Software and Resources](http://www.classification-society.org/csna/mda-sw/)

**Books on Data Mining and Machine Learning:**

* "Data Mining - Concepts and Techniques", J. Han & M. Kamber, MK
* "Neural Networks for Pattern Recognition", C.M. Bishop, Oxford University Press
* ["Pattern Recognition and Machine Learning"](http://research.microsoft.com/en-us/um/people/cmbishop/PRML/index.htm), C.M. Bishop
* "The Elements of Statistical Learning", T. Hastie, R. Tibshirani J. Friedman, 2009, [free download](http://statweb.stanford.edu/~tibs/ElemStatLearn/)
* "Introduction to Information Retrieval", Christopher D. Manning, Prabhakar Raghavan & Hinrich Schütze, 2008, Cambridge University Press
* "Gaussian Processes for Machine Learning", Carl Edward Rasmussen and Christopher K. I. Williams, The MIT Press, 2006:[free download](http://www.gaussianprocess.org/gpml/)
* "[Tika in Action](http://manning.com/mattmann/)", Mattmann & Zitting, Manning Publications, 2011, 256 pages.

**Web Services and standalone applications:**

* [DAME](http://dame.dsf.unina.it/) (DAta Mining and Exploration)
* [Weka](http://www.cs.waikato.ac.nz/~ml/weka/) (Data mining software in Java)
* [Orange](http://orange.biolab.si/)(open source data analysis and visualization)
* VOStat (Virtual Observatory statistics web services): [Penn State version](http://astrostatistics.psu.edu:8080/vostat/), [VO India version](http://vo.iucaa.ernet.in/~voi/VOStat.html)

**Programming Languages and DM Libraries:**

* *R*: [Official Manuals](http://cran.r-project.org/manuals.html), [Advanced R programming](http://adv-r.had.co.nz/), [R-Inferno](http://www.burns-stat.com/documents/books/the-r-inferno/) (book download), [R-studio](http://www.rstudio.com/)
* *Matlab*: [Official Documentation](http://www.mathworks.com/help/matlab/), [SOM Toolbox](http://www.cis.hut.fi/somtoolbox/)[, Bayesian Net Toolbox for Matlab](https://code.google.com/p/bnt/), [Netlab](http://www.aston.ac.uk/eas/research/groups/ncrg/resources/netlab/)
* Other Libraries: [Sci-kit Learn (Python)](http://scikit-learn.org/stable/), [FANN](http://leenissen.dk/fann/wp/), [LibSVM](http://www.csie.ntu.edu.tw/~cjlin/libsvm/)