Open-Source Analytics Tools

Note: Some of these tools require registration in order to access.

1) Primary tools: Python and R are the primary open-source coding languages. Download the Anaconda Distribution - https://www.anaconda.com/products/distribution It has a huge toolkit and comes with Python and RStudio. With Anaconda, get Jupyter Notebook, which is a programming notebook that allows for R, Python and Julia. Within Python, the pandas, Numpy and SciKit Learn are the most used libraries for data science.

Note: Julia is an open-source version of MatLab

2) Scrapy: Python library for scraping - https://scrapy.org/download/

3) Orange: Easy Data Visualization software - https://orangedatamining.com 4) GGPlot2:

Data Visualization in R -- https://r-graph-gallery.com/ggplot2-package.html

5) Google tools: Google analytics, google sheets and Google document. TensorFlow, which is an open-source machine learning tool. The full library is not open source.

6) Latex and Beamer: Latex is an open-source document preparation software and beamer is a presentation tool using the Latex language. TexMaker is suggested IDE.

7) KNIME: Full data science workflow software

8) PostHog: This is an open-source analytics software similar to Tableau -cloud version avaiable

9) Open source Geographic tools –

- QGIS - https://www.qgis.org/en/site/
- Geoda - http://geodacenter.github.io/
- GRASS GIS - https://grass.osgeo.org