

Setting up a Project

Fundamentals of Data Science

Jeremy Teitelbaum

Key Software Tools

Anaconda Install for Python

- Download from <https://www.anaconda.com>

Anaconda includes:

- python
- jupyter: notebook working environment
- python libraries: ML, visualization, I/O and others
- conda package manager: for dealing with multiple versions of libraries
- anaconda navigator: a GUI gateway to anaconda tools
- lots of other stuff

Verify Anaconda

- Verify JupyterLab

From a command line

```
$ jupyter lab
```

or use anaconda navigator to launch jupyterlab.

- Verify python version

From a command line

```
$ python --version
```

or inside a jupyter notebook cell:

```
import sys
print(sys.version)
```

R and Rstudio Install for R

- R is an open source language for statistical computations.
- Rstudio is a working environment for the R language.
- R and Rstudio need to be installed separately.
- R is available at <https://cran.r-project.org>
- Rstudio is available at <https://posit.co/download/rstudio-desktop>

Verify R and Rstudio

For R, From a command line:

```
$ R
```

For Rstudio, use the icon/shortcut or from a command line:

```
$ rstudio
```

Setting up a Project

- Create a project directory
- Subdirectories
 - data for data files
 - docs for notes and documentation
 - others?
- Create a README.md file

For the directories: - the finder or File Manager - the command line

For the README file: - a text editor such as **nano** or notepad - jupyter or Rstudio (as we will see soon)