1. Essential Software

Fundamentals of Data Science

Jeremy Teitelbaum

Essential Software

- the anaconda distribution including python v3.10 or higher
- the vscode development environment
- the R language
- the Rstudio integrated development environment

Anaconda

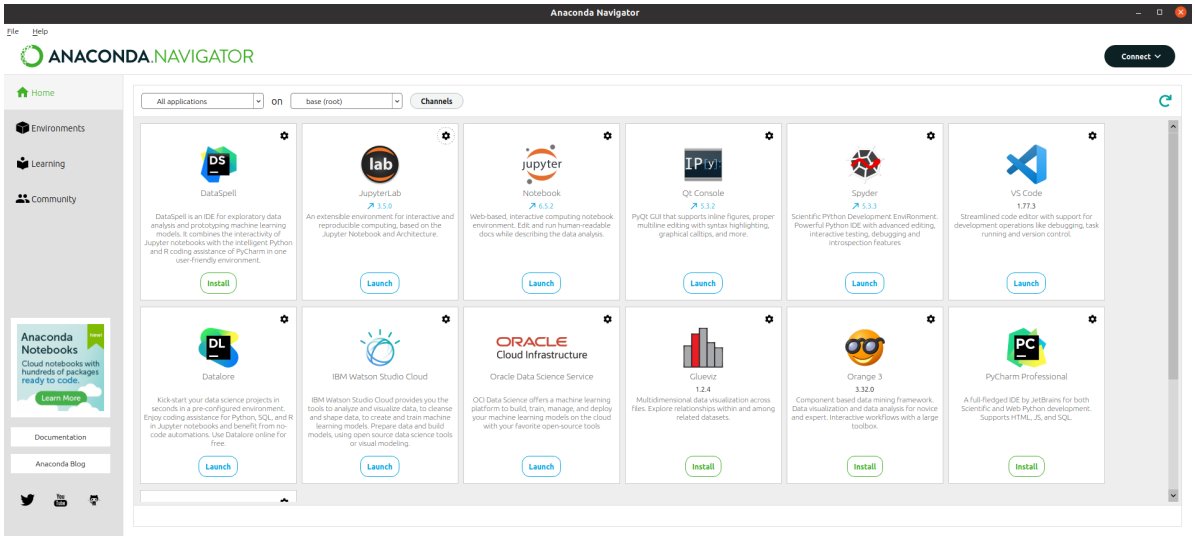
Anaconda is a data science software package that includes:

- An up-to-date version of the `python` language.
- A package manager (`conda`) for managing python libraries, and a collection of those libraries
- A GUI interface (`anaconda-navigator`) for launching many of the data science tools.
- You can install other tools (such as `R` and `Rstudio`) through `anaconda-navigator` but we won't go that route.

Anaconda Setup

1. Download the appropriate anaconda package for your operating system.
2. Run the installer.
3. Start anaconda-navigator to verify that it was installed.
4. Run an instance of `JupyterLab` to verify that it starts properly.

Anaconda Navigator



VSCode

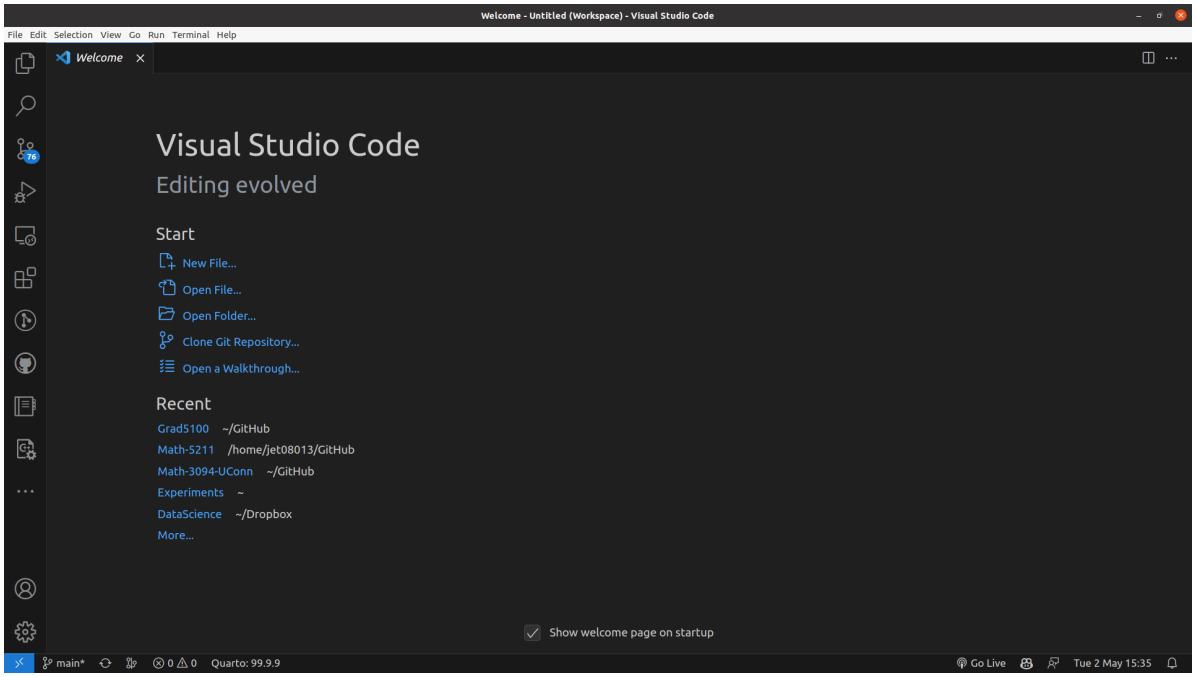
`VSCode`, or "Visual Studio Code", is a powerful code editor developed by Microsoft.

It offers many tools to help write and maintain code and text.

1. Download and install vscode from the home page.
2. Start `VSCode` to make sure that it works. Use the start menu (on Windows), the finder (on MacOS), or open a shell window and type:

```
$ code
```

VSCode



R

`R` is a programming language that is highly optimized for statistical modeling.

Install `R` by following the instructions on the home page.

R

Verify that `R` works by starting it:

- On MacOS, there is a graphical version of `R` called `R.app`
- The installation process on Windows creates a desktop shortcut to start `R`.
- On Linux, you start `R` from a desktop shortcut or the command line

```
$ R
```

Note that you quit `R` with

```
> q()
```

Rstudio

`Rstudio` is a development environment for `R` (and for other languages if you prefer it).

Install `Rstudio` from the home page.

Start `Rstudio` by double clicking a shortcut on Windows or Mac, or from the command line:

```
$ rstudio
```

Rstudio