

Carpintería de Datos

Prof. Dr. Germán Chaparro

Universidad ECCI



Python – Lo Bueno

- Actualizado continuamente
- Módulos para temas generales (Matplotlib, NumPy, SciPy)
- Módulos para temas especializados (AstroPy, SunPy,...)
- Inicialmente intuitivo
- Hay muchos *shells* (intérpretes) interactivos
- Fácil de combinar con el *shell* de Unix y con otros lenguajes
- Open source
- Usualmente su uso es de alto nivel

Python – Lo Malo

- Descentralizado
- Documentación “en la nube”
- Puede ser lento para manejar *arrays* (arreglos)
- Requiere uso indiscriminado de memoria
- Es interpretado, no compilado
- -> Optimizar no es trivial

Binder

Máquina virtual (nube)

+ Docker: Linux + Python + Jupyter
+ Requerimientos (librerías, etc.)

<https://binderhub.readthedocs.io/en/latest/overview.html>

Se tiene acceso a muchas herramientas
que tiene cualquier computador con
Linux que tenga conexión a internet de
alta velocidad



Turn a GitHub repo into a collection of
interactive notebooks

Have a repository full of Jupyter notebooks? With Binder, open those notebooks in an executable environment, making your code immediately reproducible by anyone, anywhere.

Build and launch a repository

GitHub repo or URL

Git branch, tag, or commit

Path to a notebook file (optional)

File ▾

launch ▾

Already built!

Launching

Build logs

show

Fuera de Binder (localmente)

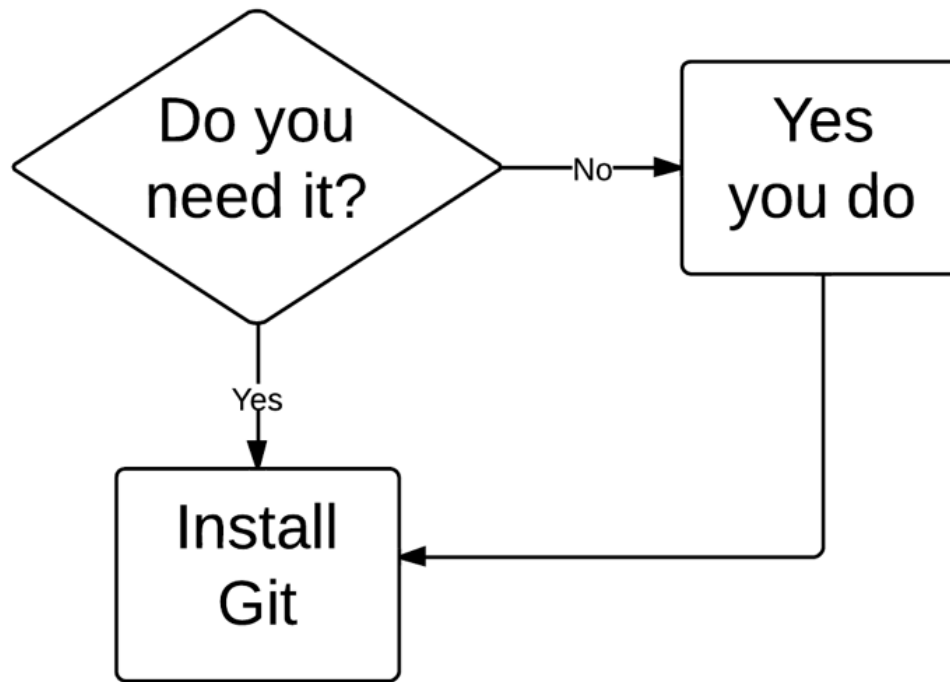
- OS: Windows, Linux, macOS
- Administrador de paquetes: Anaconda, pip
- Administrador de entornos: Anaconda
- Editor de texto: Nano, Sublime, emacs, vim... (Scripts -> .py)
- IDE: Spyder, PyCharm... (Scripts -> .py)
- Entorno de Notebooks: Jupyter (.ipynb)
- Repositorio en Línea con Control de Versión: Bitbucket, GitHub,...

Librerías para hoy

- Numpy
- Pandas
- Collections (por defecto en Python)
- Matplotlib

Control de Versión

Version Control Flowchart



Originally developed for large software projects with many developers.

Also useful for single user, e.g. to:

- Keep track of history and changes to files,
- Be able to revert to previous versions,
- Keep many different versions of code well organized,
- Easily archive exactly the version used for results in publications,
- Keep work in sync on multiple computers.