

# **INTRODUÇÃO À CIÊNCIA DE DADOS**

**Jupyter Notebook e o Google Colab**

# PYTHON (python.org)

- Nesta disciplina usaremos massivamente a linguagem Python.
- Python é uma linguagem de programação de alto nível, interpretada de script, com tipagem dinâmica e forte. Além disso é multiplataforma.
- Há muita conversa e às vezes competição em definir a melhor linguagem para DS (python, linguagem R, ou outras menos utilizadas).
- Saiba:
  - Tanto o Python como o R são ótimos, mas é preciso se especializar em uma delas.
  - Penso que o grande diferencial seja a quantidade de oferta de empregos e também a amplitude de possibilidades que o Python oferece em relação ao R.

# POR ONDE COMEÇAR?

- Já usa Python? Ótimo!
- Não usa ainda, **sugiro** instalar o Anaconda (<https://www.anaconda.com/products/individual>)
- O anaconda é um "kit" open source, disponível para Windows, Mac e Linux, que instala o Python e um conjunto de bibliotecas, pacotes e utilitários para iniciar o desenvolvimento tanto com Python como com R.
- O anaconda facilita muito o trabalho inicial de quem quer começar a ser um profissional de DS.

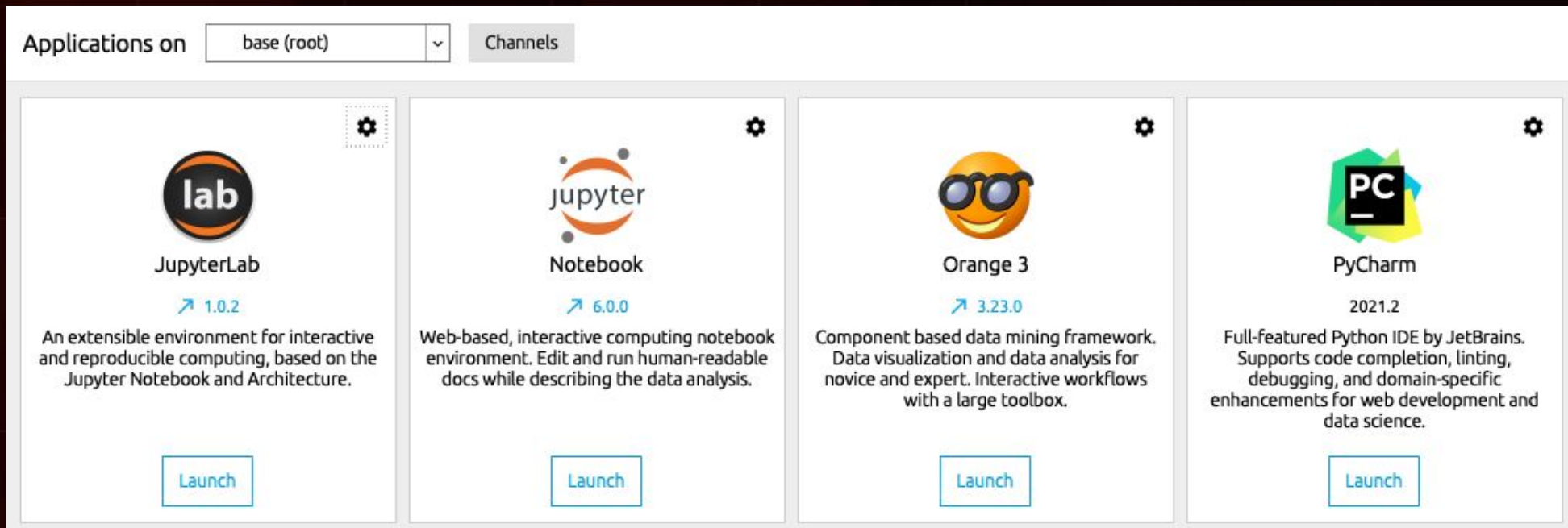
# ANACONDA! Muita coisa boa!!

The screenshot displays the Anaconda Navigator desktop application. The interface is divided into a left sidebar and a main content area. The sidebar contains navigation links for Home, Environments, Learning, and Community. The main content area shows a grid of application tiles under the heading 'Applications on base (root) Channels'. Each tile represents an application or environment, including its icon, name, version, a brief description, and a button to either 'Launch' or 'Install' it. The applications listed are:

Application	Version	Action
JupyterLab	1.0.2	Launch
Jupyter Notebook	6.0.0	Launch
Orange 3	3.23.0	Launch
PyCharm	2021.2	Launch
Qt Console	4.5.1	Launch
Spyder	3.3.6	Launch
VS Code	1.59.0	Launch
dioplas	0.4.1	Install
fsleyes	0.34.2	Install
Glueviz	1.0.0	Install
gnuradio	3.8.2.0	Install
gnuradio-core	3.8.2.0	Install
gnuradio-grc	3.8.2.0	Install
gnuradio-qtdgui	3.8.2.0	Install
gnuradio-uhd	3.8.2.0	Install
gnuradio-zeroomq	3.8.2.0	Install
gpl	1.3.0	Install
gqrx	2.13.5	Install
hyperspyui	1.1.2	Install
labelme	4.5.6	Install
mercurial-app	5.2	Install
opensesame	3.3.4	Install
psy-view	0.1.0	Install
psypilot-gui	1.3.0	Install
pyboat	0.8.22	Install
RStudio	1.1.456	Install
stradtize	0.1.3	Install
tlrps	1.5.1	Install
vaex	3.0.0	Install
veusz	3.2.1	Install
xonsh	0.6.0	Install

At the bottom of the sidebar, there are links to Documentation and Developer Blog, and social media icons for Twitter, YouTube, and GitHub.

# ANACONDA! Detalhe!



# IPYTHON ([ipython.org/](http://ipython.org/))

- IPython é um shell para Python (e também para outras linguagens) com muito mais funcionalidades do que o shell padrão do Python.
- Ele permite copiar código e rodar scripts de pequenos códigos.
- Tem uma particularidade de criar pequenos cadernos (notebooks) onde é possível ir registrando as atividades como se fosse um diário.

# JUPYTER NOTEBOOK ([jupyter.org](http://jupyter.org))

- Aplicativo open source para desenvolvimento.
- É uma evolução do IPython, entretanto seu ambiente de trabalho deixa de ser o shell (IPython) e passa a ser o Browser (navegador web).
- Excelente alternativa para desenvolvimento e prática de programação em Python em praticamente todas as fases de atividade do profissional de DS.

# JUPYTER LAB ([jupyter.org](https://jupyter.org))

- É a próxima geração do Jupyter Notebook.
- Open source, é uma renovação do Jupyter Notebook, que possibilita utilização de abas, instalação de extensões para os mais variados fins como: `#github`, `#toc` (índice), `#variableInspector` (inspetor de variáveis), `#spreadsheet` (visualização de planilhas) entre outros.
- Essa é uma ótima opção para desenvolvedores, para estudar DS. Iremos utilizar no Jupyter Lab nessa disciplina!



# JUPYTER LAB - DEMONSTRAÇÃO

- Faremos uma demonstração do uso do Jupyter Lab para o desenvolvimento em Python.
- Partiremos da ideia que o computador já tenha o Jupyter Lab instalado. A instalação do Anaconda facilita isso.
- Lembre-se que ter o Anaconda instalado é uma sugestão que seria muito importante, mas não é obrigatória, na sequência veremos outra opção similar para desenvolvimento.
- Vamos à demonstração!!



# GOOGLE COLAB

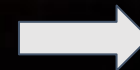
- A Google disponibiliza uma plataforma para desenvolvimento Python, muito similar ao Jupyter Notebook/Lab que o desenvolvedor instala em sua máquina pessoal.
- Uma customização do Jupyter! Mas lembre-se, para usar o Google Colab é necessário ter internet disponível.

# GOOGLE COLAB

- Algumas vantagens:
  - Não precisa ter o Anaconda/Python/Jupyter instalado.
  - Não precisa ficar gerenciando bibliotecas.
  - Acesso a GPUs e TPUs gratuitas (use se necessário).
  - Acesso de qualquer lugar (incluindo de outros dispositivos), tudo na nuvem.
  - Compartilhamento fácil.

# GOOGLE COLAB

- Não depende da máquina local (pode ser uma vantagem ou desvantagem dependendo da sua máquina local).
- Vamos usar muito aqui nessa disciplina!
- Temos desvantagens também:
  - Execução interrompida após algumas horas.
  - Bibliotecas não disponíveis precisam ser reinstaladas a cada execução.
- Demonstração!!



# FINALIZANDO

- Jupyter e Google Colab são ótimas alternativas para desenvolvedores de DS.
- Compreenda suas diferenças.
- Use muito, só fazendo muitas vezes é que se aprende!

# **INTRODUÇÃO À CIÊNCIA DE DADOS**

**Jupyter Notebook e o Google Colab**