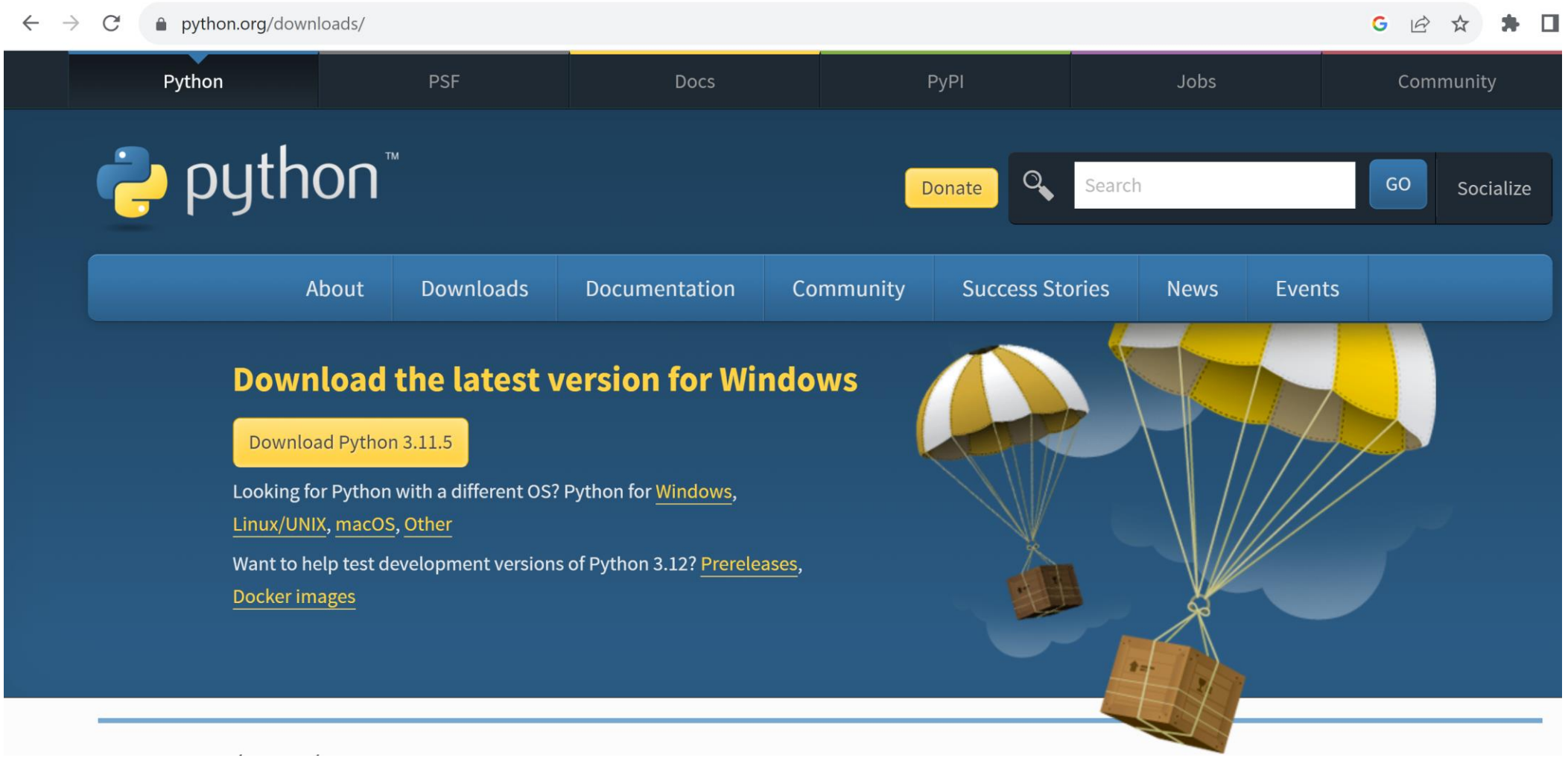
Installing Jupyter Notebook for Spark

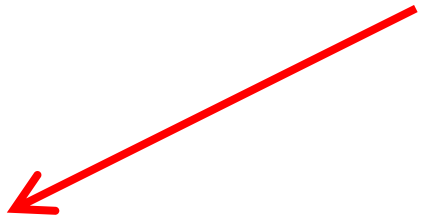
Arnaud.nauwynck@gmail.com

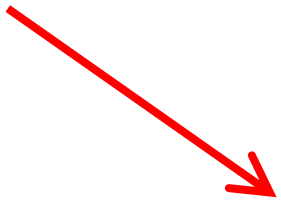
Step 1 : install python (simplest method, not using AnaConda)



Step 2 : install jupyter

Choose class Jupyter Notebook… OK





JupyterLab

Install JupyterLab with `pip`:

```
pip install jupyterlab
```

Note: If you install JupyterLab with conda or mamba, we recommend using the conda-forge channel.

Once installed, launch JupyterLab with:

```
jupyter lab
```

Jupyter Notebook

Install the classic Jupyter Notebook with:

```
pip install notebook
```

To run the notebook:

```
jupyter notebook
```

Step 3 : install jupyter spylon_kernel (for using Spark with Scala langage)

python –m spylon_kernel install --user

(or pip install spylon_kernel –user)

```
C:\Users\arnaud>python -m spylon_kernel install --user
0.00s - Debugger warning: It seems that frozen modules are being used, which may
0.00s - make the debugger miss breakpoints. Please pass -Xfrozen_modules=off
0.00s - to python to disable frozen modules.
0.00s - Note: Debugging will proceed. Set PYDEVD_DISABLE_FILE_VALIDATION=1 to disable this validation.
[InstallKernelSpec] Installed kernelspec spylon-kernel in C:\Users\arnaud\AppData\Roaming\jupyter\kernels\spylon-k
ernel
```

Star Jupyter Notebook

jupyter notebook

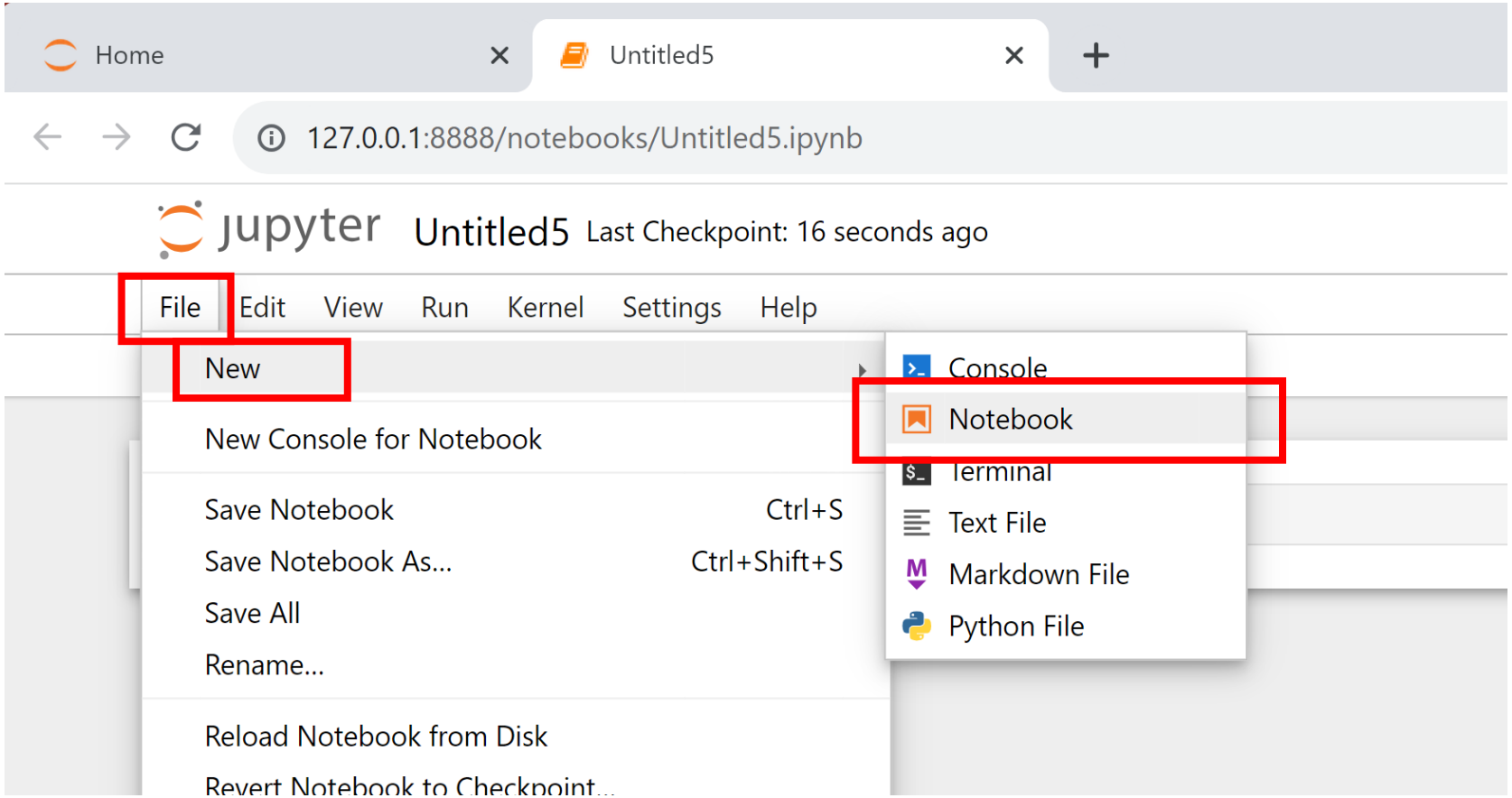
```
C:\Users\arnaud>jupyter notebook
[I 2023-09-06 23:09:00.357 ServerApp] Package notebook took 0.0000s to import
[I 2023-09-06 23:09:00.482 ServerApp] Package jupyter_lsp took 0.1187s to import
[W 2023-09-06 23:09:00.482 ServerApp] A `_jupyter_server_extension_points` function was not found in jupyter_lsp.
Instead, a `_jupyter_server_extension_paths` function was found and will be used for now. This function name will
be deprecated in future releases of Jupyter Server.
[I 2023-09-06 23:09:00.544 ServerApp] Package jupyter_server_terminals took 0.0598s to import
[I 2023-09-06 23:09:00.544 ServerApp] Package jupyterlab took 0.0000s to import
```

[[truncated logs …. Also contains errors??]]

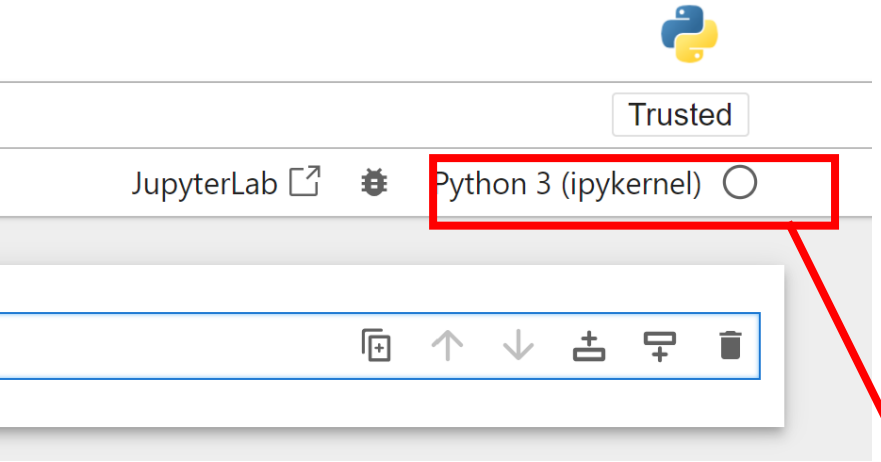
Jupyter Notebook ..

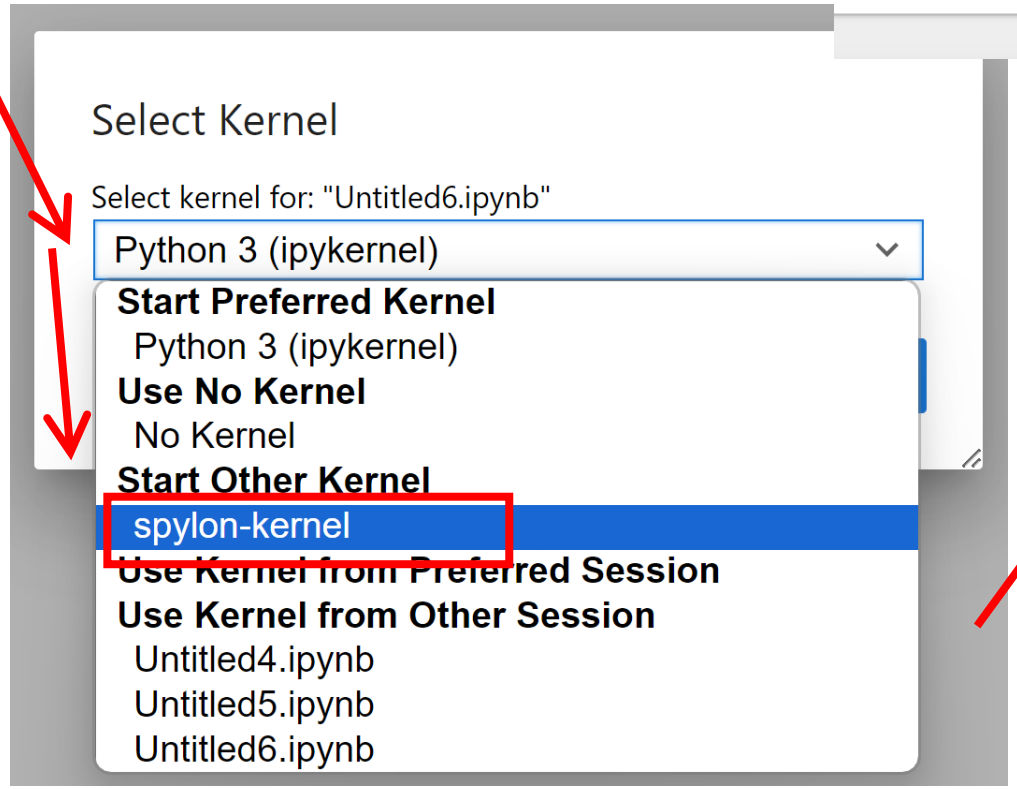
```
[I 2023-09-06 23:09:01.879 ServerApp] jupyterlab | extension was successfully loaded.
[I 2023-09-06 23:09:01.879 ServerApp] notebook | extension was successfully loaded.
[I 2023-09-06 23:09:01.895 ServerApp] Serving notebooks from local directory: C:\Users\arnaud
[I 2023-09-06 23:09:01.895 ServerApp] Jupyter Server 2.7.3 is running at:
[I 2023-09-06 23:09:01.895 ServerApp] http://localhost:8888/tree?token=a9195d7e950f42e4e26a7b93bbebf3664d1794b9495
9c5e3
[I 2023-09-06 23:09:01.911 ServerApp]     http://127.0.0.1:8888/tree?token=a9195d7e950f42e4e26a7b93bbebf3664d1794b
94959c5e3
[I 2023-09-06 23:09:01.911 ServerApp] Use Control-C to stop this server and shut down all kernels (twice to skip c
onfirmation).
[C 2023-09-06 23:09:01.974 ServerApp]

    To access the server, open this file in a browser:
        file:///C:/Users/arnaud/AppData/Roaming/jupyter/runtime/jpserver-636-open.html
    Or copy and paste one of these URLs:
        http://localhost:8888/tree?token=a9195d7e950f42e4e26a7b93bbebf3664d1794b94959c5e3
        http://127.0.0.1:8888/tree?token=a9195d7e950f42e4e26a7b93bbebf3664d1794b94959c5e3
```

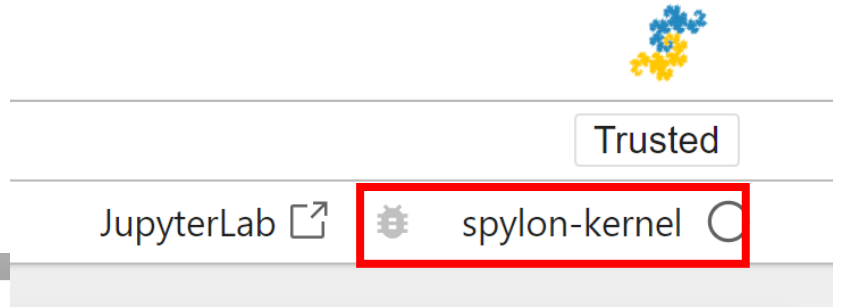
Creating a new notebook



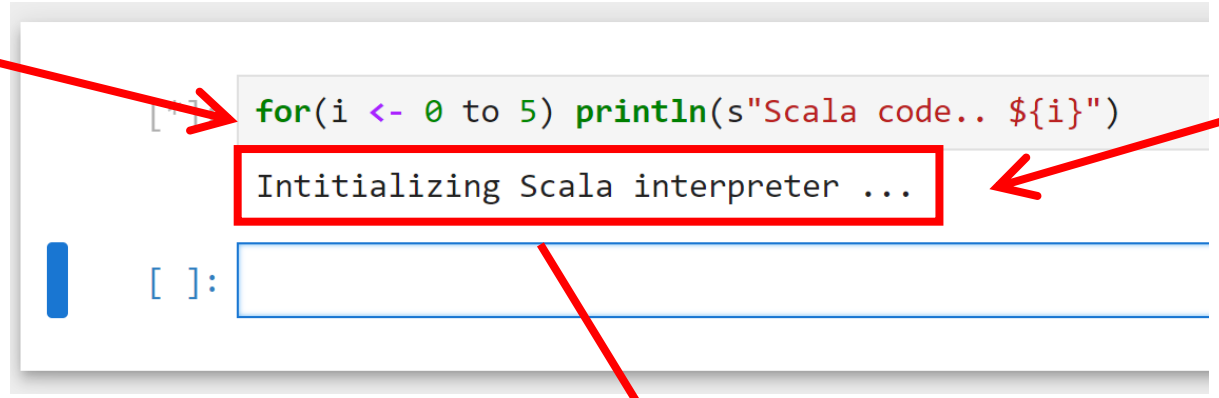
Step 3 … testing spylon

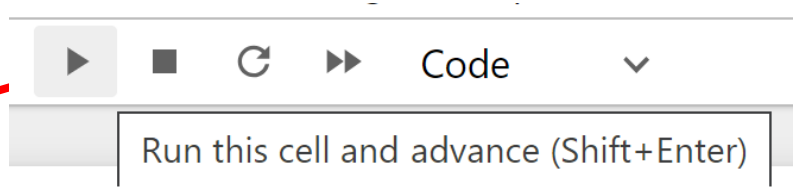






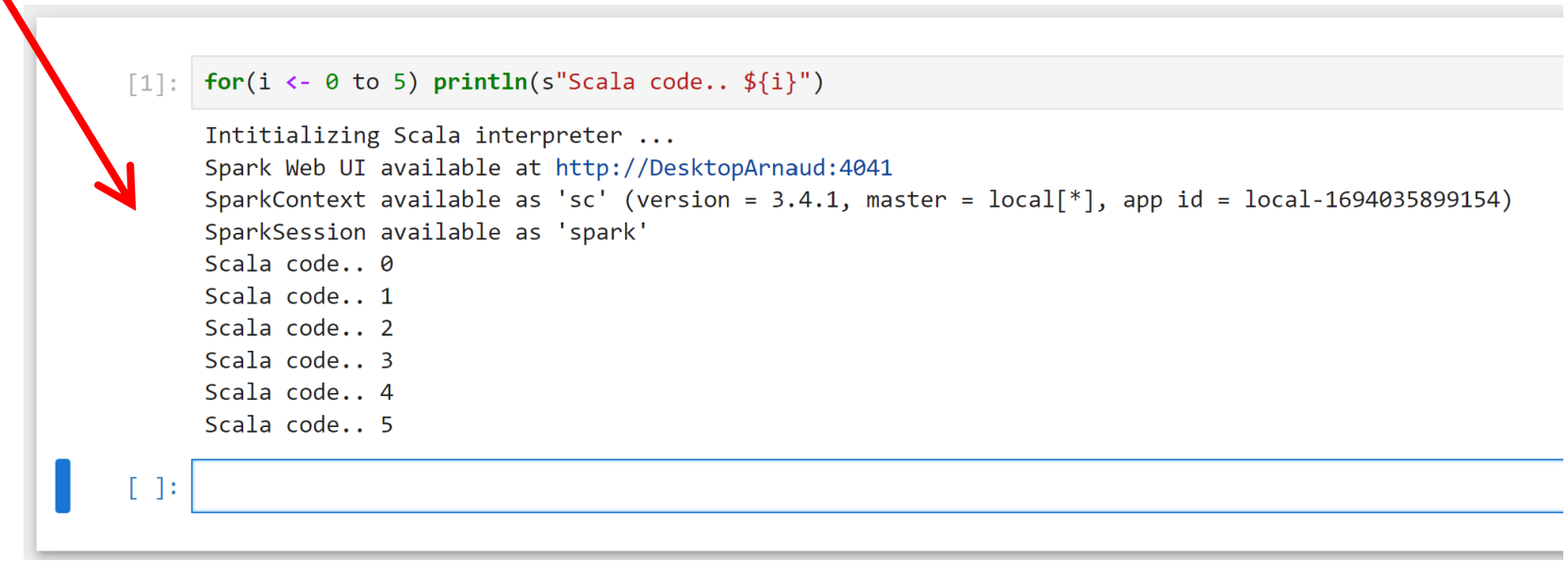
Step 3 … Testing Spylon .. Write Scala



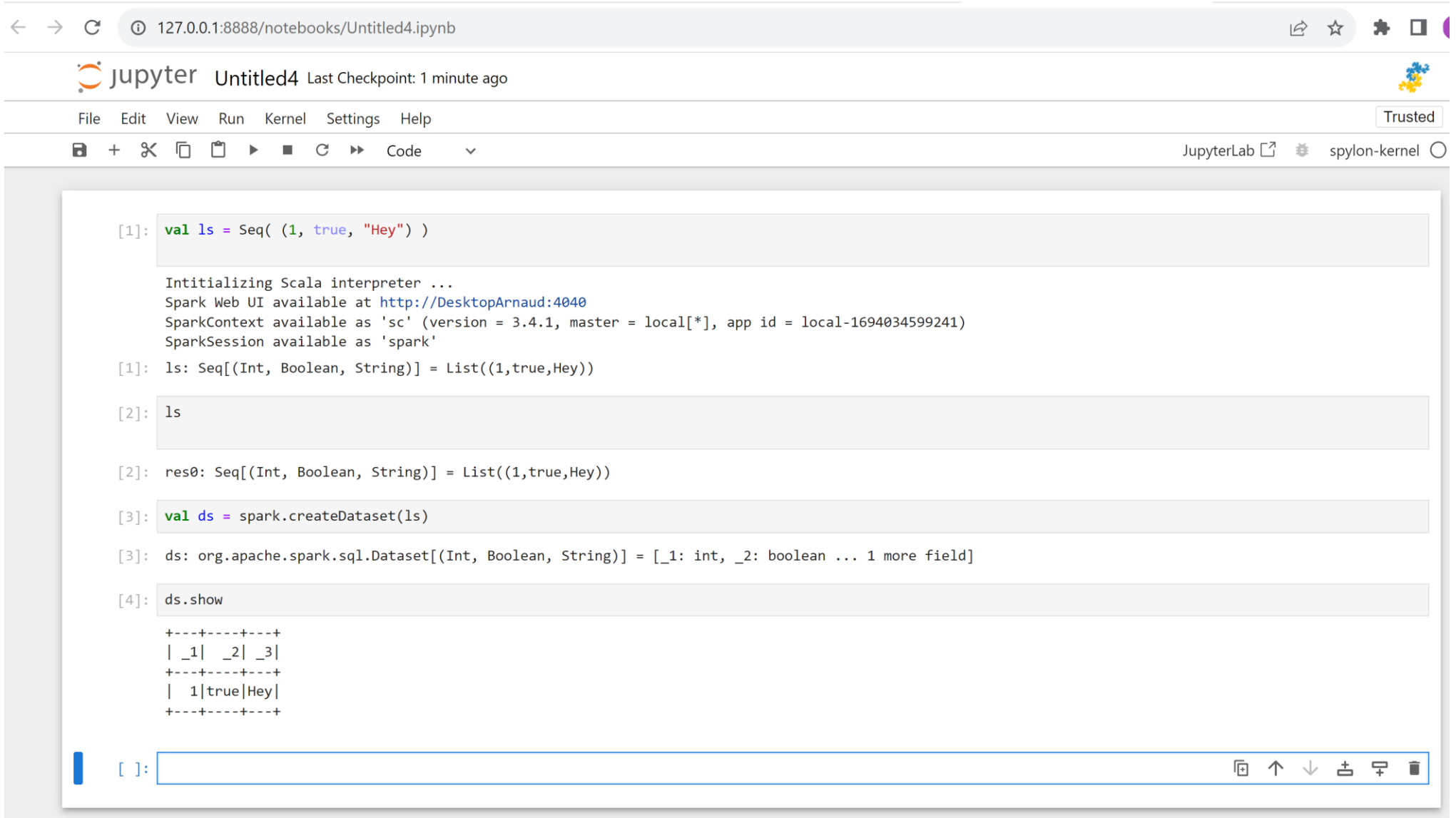


Type SCALA code
… Shift+Enter

Get Spark process
(Scala interpreter)



Step 3 … Testing Spylon .. Write Scala



Step 4 : install kernel for pyspark…
(for using PySpark with Python langage)