[One time setup] Install IPython

From the top left menu, Open a terminal: Applications => System Tools => Terminal

Type:

```
1   sudo easy_install ipython==1.2.1
```

Hit enter, administrator password is **cloudera**.

Launch pyspark with IPython

Every time you need to open the pyspark shell, open a terminal and type:

```
1   PYSPARK_DRIVER_PYTHON=ipython pyspark
```

Hit enter, after the startup logs, you should see the pyspark console:

Check version

To make sure that PySpark started correctly, print out the version by typing in the PySpark IPython terminal:

```
1   sc.version
2   
```

Verify that the output is:

```
1  u'1.3.0'
2  |
```

Python resources

This module requires a basic knowledge of Python to perform the programming assignments.

I recommend to take the interactive tutorial at: http://learnpython.org, the "Learn the basics" section should be enough.

Then if you want to learn more about Python, even if not necessary for this class, you can check the Coursera class:

https://www.coursera.org/learn/python

Mark as completed