

Data Science Tools and Models: Setting up Jupyter Notebooks

1 Download

Download the Python 3.6 version from: <https://www.anaconda.com/download> and install.

You may also install Python 3 www.python.org and Jupyter jupyter.org without installing Anaconda.

Install using the defaults.

2 Install packages

To integrate PostgreSQL and Jupyter (we will install Postgres in a later lab):

1. Run Anaconda with administrative privileges
2. Select “All” from the drop down at the top left of the main pane.
3. Search for psycopg2 and check the box next to it. For more information on this package, see <https://pypi.python.org/pypi/psycopg2>
4. Click Apply
5. Allow any dependencies to be installed as well.

Also install the ipython-sql package using one of the following commands:

```
pip install ipython-sql
python3.6 -m pip install ipython-sql
```

For more information on ipython-sql, see pypi.python.org/pypi/ipython-sql.

3 Set up Jupyter

1. Choose a directory that will contain the course notebooks, such as:

```
/Users/rbm2/Documents/School/Rice/DSToolsAndModels/notebooks
```

2. Create a configuration file

```
jupyter notebook --generate-config
```

3. Find your configuration file (the previous step should tell you where it is). Typically it is your home directory, in a new folder named .jupyter. The configuration file will be named jupyter_notebook_config.py.
4. Edit the configuration file to point to the default notebook directory.

(a) Find the row that says: `#c.NotebookApp.notebook_dir = ''`

(b) Uncomment it and include the path to your notebook folder

```
c.NotebookApp.notebook_dir = '/Users/rbm2/Documents/School/Rice/DSToolsAndModels/n
```