

Jupyter Notebooks on the Internet

There are thousands of interesting Jupyter Notebooks available on the internet for you to learn from. One of the best sources is: <https://github.com/jupyter/jupyter/wiki>

It is important to notice that you can download such notebooks to your local computer or import them to a cloud based notebook tool so that you can rerun, modify, and apply what's explained in the notebook.

Very often, Jupyter Notebooks are already shared in a rendered view. This means that you can look at them as if they were running locally on your machine. But sometimes, folks only share a link to the Jupyter file (which you can make out by the *.ipynb extension). In this case, you can pick the URL to that file and paste it to the NB-Viewer => <https://nbviewer.jupyter.org/>

The list of Jupyter Notebooks provides you with a huge collection of materials to explore. Therefore, it might be useful to give you some pointers to interesting notebooks. You have covered some examples with data in the labs. Let's highlight some useful data that further explores data science. In addition, as we have covered different tasks in data science, we will also provide a sample notebook for each task.

First, you start with exploratory data analysis, for which this notebook is highly recommended: https://nbviewer.jupyter.org/github/Tanu-N-Prabhu/Python/blob/master/Exploratory_data_Analysis.ipynb

For data integration/cleansing at a smaller scale, the python library `pandas` is often used. For this task, you can have a look at this notebook: <https://towardsdatascience.com/data-cleaning-with-python-using-pandas-library-c6f4a68ea8eb>

If you want to know more about clustering, have a look at this notebook: https://nbviewer.jupyter.org/github/temporacr/tutorial_ml_gkbionics/blob/master/2%20-%20KMeans.ipynb

And finally, if you want an in-depth notebook on the `iris` dataset, have a look at this: <https://www.kaggle.com/lalitharajesh/iris-dataset-exploratory-data-analysis>

Author(s)

Romeo Kienzler

