

## 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 follow along what's explained in the notebook.

Very often jupyter notebook 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 just grab the URL to that file and past it to the NB-Viewer => <https://nbviewer.jupyter.org/>

The list above gives you a very nice start with a huge collection of materials to explore. Therefore it's maybe more useful to give you some pointers to interesting notebooks. As we have covered some toy examples with toy data in the labs, let me just point to some work which uses these data and goes further down the road of data science. In addition, as we've covered the different tasks in data science we'll also provide an exemplar notebook for each of those.

First you start with exploratory data analysis, so this notebook is highly recommended to have a look at:

[https://nbviewer.jupyter.org/github/Tanu-N-Prabhu/Python/blob/master/Exploratory\\_data\\_Analysis.ipynb](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. Please have a look at this notebook:

<https://towardsdatascience.com/data-cleaning-with-python-using-pandas-library-c6f4a68ea8eb>

If you want to already experience what clustering is, have a look at this:

[https://nbviewer.jupyter.org/github/temporaer/tutorial\\_ml\\_gkbionics/blob/master/2%20-%20KMeans.ipynb](https://nbviewer.jupyter.org/github/temporaer/tutorial_ml_gkbionics/blob/master/2%20-%20KMeans.ipynb)

And finally, if you want to go for a more in-depth notebook on the *iris* dataset have a look here: <https://www.kaggle.com/lalitharajesh/iris-dataset-exploratory-data-analysis>

## Author(s)

Romeo Kienzler

## Change log

Date	Version	Changed by	Change Description
2020-10-21	2.0	Malika Singla	Migarted to GitLab