GESIS Workshop

Big Data: Introduction to Data Science with Python

Preparing for the Course

1. Preparing Python basics:

Thoroughly go through the six tutorial notebooks we have prepared for you to learn (or repeat) the basics of Python in Jupyter Notebooks. Start with the first, `00b_preparation.ipynb`. Do the **exercises** within, i.e. actually try to solve them yourself (if you fail, don't worry, getting started writing some code is key - or even simply copy-pasting). A basic understanding of these notebooks is very helpful as preparation for the course. We have put up a version that runs completely on GESIS Servers for preparation (if you are not already on it right now).

2. Local installation for the actual workshop:

Please bring your own laptop. We strongly recommend that you install Anaconda (including Python, which you need anyway). It can be downloaded here:

https://www.continuum.io/downloads

Anaconda is an open data science platform powered by Python. It comes with many code libraries / packages for Python already installed (and it manages those packages for you). It also comes equipped with Jupyter Notebooks. We will be working with the newest Python version (currently 3.7), and we will use Jupyter Notebooks for the teaching.
You can launch Jupyter notebooks either directly (on Windows through the icon in the start menu; on Linux by typing `jupyter notebook` in the terminal) or through the Anaconda Navigator. (If you want, you can also try out Jupyter Lab or Spyder, but all teaching materials are optimized for Jupyter Notebooks)

A good video describing the general installation of Anaconda:
https://www.youtube.com/watch?v=YJC6IdI3hWk
(we recommend using the option to add it to your PATH - you will be asked in the installation process - on your machine if you have not previously installed Python on it)

And starting/using Jupyter:
https://www.youtube.com/watch?v=HW29067qVWk

Troubleshooting resources can be found at:
https://docs.continuum.io/anaconda/install
http://jupyter-notebook-beginner-guide.readthedocs.io/en/latest/