Rail Data Science

0.1 Introduction to Python and Jupyter

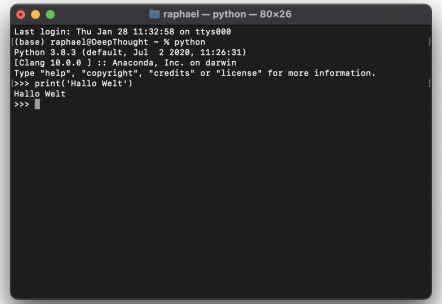
Prof. Dr. Raphael Pfaff

23. Mai 2023

Fachhochschule Aachen

Python

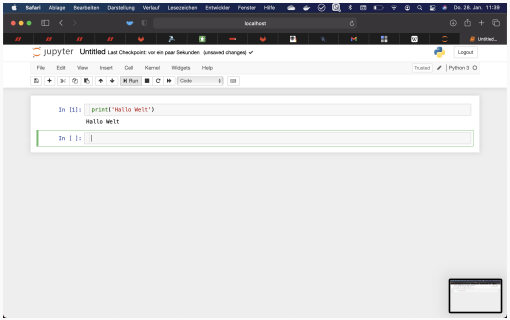
- Interpreted high-level language
- Focus on readibility
- Program structure by indentation - no braces
- Dynamically typed - no variable definition required
- Highly popular
  - Provides numerous powerful packages
  - Trained staff and freelancers available



Python in Terminal

Jupyter

- Interactive Notebook
- Can contain Text, images, code
  - Multiple languages, e.g. R, Julia, Python
- Powerful combination of code, results and explanations - in one document
- Notable extensions:
  - JupyterHub: Multi-User server
  - JupyterLab: Enriched user interface, more formatting options



Basic Jupyter Notebook

Installation

- (Mostly) required: Anaconda[1]
  - Preselected python packages for data analysis
  - Package manager `conda`: conveniently add packages from console
  - Environment manager
  - Download at `https://www.anaconda.com/products/individual`
  - Follow installation instruction
- Convenient: Github desktop app
  - Version control of code
  - Clone course repository (and others...)
  - Download at `https://desktop.github.com`https://desktop.github.com
  - Follow installation instruction

[1] In case an installation is impossible, google Colab or Binder online-Solutions can be arranged.

Special packages individual branches

- Base packages: anaconda
- Using `conda package manager`
- Data analysis:
  - Plotly: `conda install -c conda-forge plotly`
- Simulation:
  - PyControl: `conda install -c conda-forge control`
- Computer Vision:
  - OpenCV: `pip install opencv-python`
  - Tesseract: `pip install tesseract pytesseract`