Data Visualization with Python

Generate beautiful plots with pandas and matplotlib.

Dr. Michael Horn, Fachreferat Naturwissenschaften, UB Bern
Coffee & Bit(e)s Fall 2020
This Lecture

Content

- why data visualization?
- why data visualization with python and python tools?
- what are these tools?
- presentation of a jupyter notebook
- conclusion
Why Data Visualization (DV) ?
Understand data and communicate information

Understanding of data
➢ DV makes data perceptible by humans.
➢ DV highlights patterns in data (correlations, outliers, etc.).
➢ DV promotes scientific ideation.

Communication of information
➢ DV highlights the essence of a survey (by putting all information in a nutshell).
➢ DV is esthetical.
➢ DV is fundamental for storytelling.

Data visualization is essential in a scientific context.
Why DV with Python and Python Tools? Extraordinary open source tools

➢ free of cost, non-proprietary

➢ very flexible and versatile
  o can be integrated in various workflows
  o have a good interoperability (among each other and with other types of software)

➢ well-established and reliable
  o have a large user community, are maintained and refined by a large community of independent developers
  o have a good documentation, are transparent
What are these Tools?

Python, jupyter, pandas, matplotlib

➢ python:
  o programming (scripting) language
  o easy to learn
  o very powerful, flexible and versatile

➢ jupyter notebook:
  o browser based application
  o integrates code (e.g. python), code output (e.g. plots) and documentation

➢ pandas:
  o python library for data processing
  o fast, powerful, flexible, easy to use
  o integrates matplotlib

➢ matplotlib:
  o python library for data visualization
  o enables DV in matlab style
  o is platform independent and very robust
Data Visualization

Presentation of a jupyter notebook

Jupyter notebook on GitHub:

https://github.com/ubnpl/pytools/blob/master/data_visualization/
Data_Visualization_with_Python_CL_HS2020.ipynb
Conclusions
Take home message

- data visualization is fundamental for understanding data and communicating information.
- Python, Jupyter, Pandas and Matplotlib are extraordinary open source tools that can be used for data visualization.
- The University Library constructs a digital toolbox which offers exemplary use of such tools.
Useful literature:


Useful links:

➢ python: https://www.python.org/

➢ jupyter: https://jupyter.org/

➢ pandas: https://pandas.pydata.org

➢ matplotlib: https://matplotlib.org
Thank you
for your attention

Dr. Michael Horn, Fachreferat Naturwissenschaften, UB Bern
Coffee & Bit(e)s Fall 2020

@ www.unibe.ch/ub/sciencelibrary ➔ see «Coffee & Bit(e)s» for lecture notes