Interactive data visualisation in Python

A Workshop

Instructor: Thomas Bury[1]

[1]Dept. Physiology, McGill University

McGill
UNIVERSITY

QLS SQV
QUANTITATIVE LIFE SCIENCES
SCIENCES QUANTITATIVES DU VIVANT

Software checklist

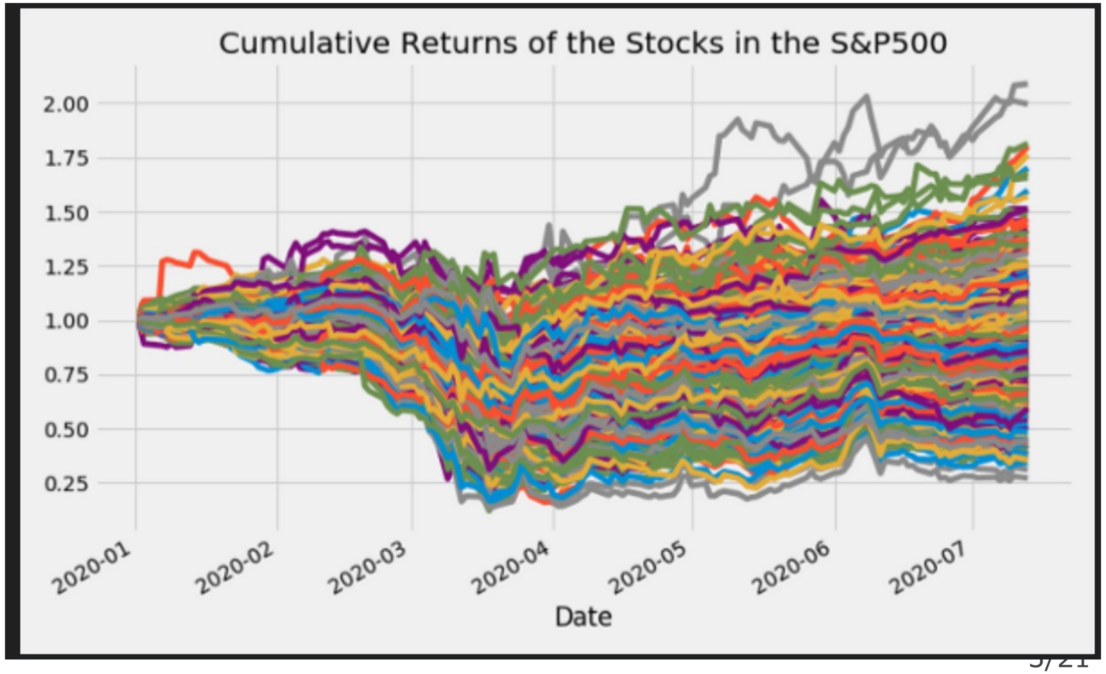
- Installed the necessary software:
  - Python 3
  - JupyterLab
  - numpy, pandas, plotly, dash, jupyter-dash

- Forked and cloned the github repository
- Ran the Jupyter notebook `test.ipynb` successfully.

Why bother with *interactive* visualisations?

- Big data: lots of variables (dimensions), lots of entries.
- Can be difficult adequately represent it on a 2D static plot
- Having interactivity allows us to
  1. Navigate through different sections of data
  2. Vary the scale of the axes
  3. View all variable values with a tooltip
  4. Compute statistics on the fly

Icentia CardioSTAT



Dashboard demonstration

Learning objectives

By the end of this workshop, you will know how to

- Import public datasets into Python
- Use basic pandas functionality for manipulating data frames.
- Create basic interactive plots (e.g. scatter, histogram, box) as html files
- Visualise higher dimensions using grid, 3D and heat plots
- Create a slider to view a fourth dimension

Agenda

2.35pm ▪ Introduction

2.45pm ▪ Notebook #1 – basic Plotly functionality
- Demonstration (me) [15 mins]
- Participation (you) [15 mins]

3.15pm ▪ Notebook #2 – visualisation with higher dimensions [30 mins]

3.45pm ▪ Break [5 mins]

3.50pm ▪ Notebook #3 - sliders and buttons [45 mins]

4.20pm ▪ Closing remarks

4.25pm ▪ Fin

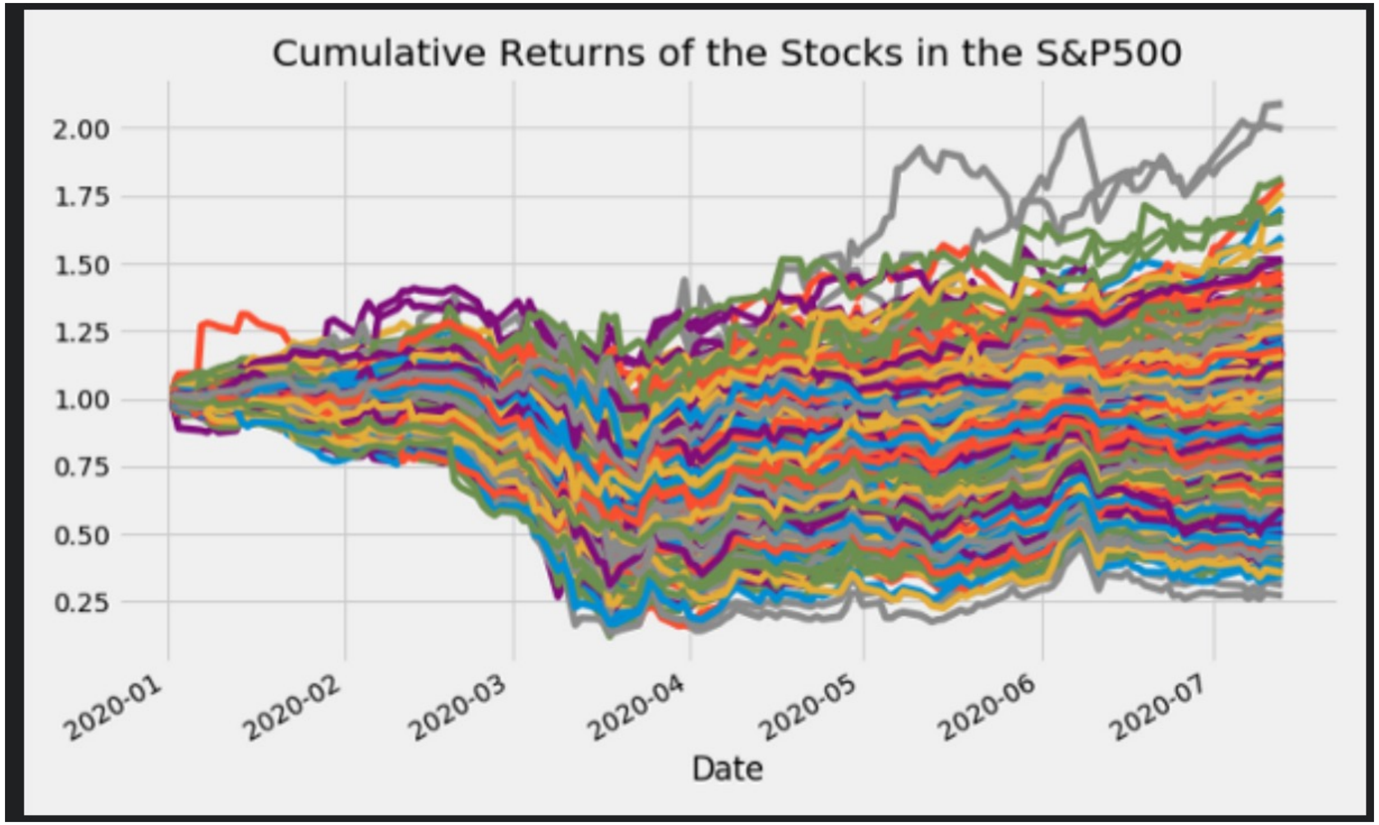
Notebook #1

Plotly fundamentals

Visualising additional dimensions

What attributes can you add to a plot to view a higher dimension?

Brainstorm!



Possible approaches

- Static visualisation methods:
  - Colour of data points
  - Size of datapoints
  - Grid of plots
  - Contour plots / Heat maps
- Interactive visualisation methods
  - Navigation through 3D plots
  - Hover data
  - Buttons, sliders and drop-down boxes

Possible approaches

- Static visualisation methods:
  - Colour of data points
  - Size of datapoints
  - Grid of plots
  - Contour plots / Heat maps
- Interactive visualisation methods
  - Navigation through 3D plots
  - Hover data
  - Buttons, sliders and drop-down boxes

Notebook #2

Visualising higher dimensions

Notebook #3

Sliders and buttons

Resources for Dash app deployment

- Read through the Plotly tutorial on deployment
  https://dash.plotly.com/deployment

- You can deploy several applications with Heroku for free (easiest option, though memory restrictions apply)
  https://devcenter.heroku.com/articles/getting-started-with-python

- Alternatively, your institution may provide cloud services where you can host Dash apps. E.g with Compute Canada.
  https://www.computecanada.ca/research-portal/national-services/compute-canada-cloud/

  "Users of the cloud service should ideally have at least an intermediate technical knowledge in systems management"

Resources for further Plotly/Dash tutorials

- Youtube channel ‘Charming data’
  https://www.youtube.com/channel/UCqBFsuAz41sqWcFjZkqmJqQ
- Plotly tutorial for beginners: Kaggle
  https://www.kaggle.com/kanncaa1/plotly-tutorial-for-beginners

Tips for fast improvement

- Design and create your own visualisations and look up the documentation as required.
  https://plotly.com/python/
- Get ideas from other Dash apps (e.g. Dash app gallery https://dash-gallery.plotly.host/Portal/), get the source code, and experiment with it.