Hands-on: retrieve from data source

1. Think about your project: what data would be interesting / useful for later?

2. Choose a source of data that you are interested in, e.g.
   - 2.1 (Yahoo)
   - 2.2 Quandl
   - 2.3 Intraday FX
   - 2.4 ...?

3. In `lib_dataretrieve.py`, add a function that retrieves data and outputs either a file or a Pandas DF.

4. Launch download

5. While data is being downloaded, write a data parsing/cleaning/formatting function.
6. Apply the function to downloaded data

Project: in your report

Minimum content

- Rolling computations
- Out-of-sample results (prediction)

Bonus:

- Reach a computing impossibility (time, memory)
- Identify the cause of the problem (computer, Operating System, your programming abilities)
- Explain how you would exploit more computing power by making computing time budgeting estimates with X computers, N cores, etc

Your project: structure

In essence, what to do with today's work

```
for t0 in range(0,len(mydata)-T):

    # 1. build predictors
    #  either portfolios
    #  or market states
    #  or performance measures of strats
    #  (optional layer) use machine learning
    #
    # 2. apply to t_oos in range(t0+T,t0+T+dT)
```

Your project: discussion time / start now

Mandatory elements:

1. Data acquisition/cleaning

2. Rolling calibration windows

3. Out-of-sample performance of something financially meaningful
   - 3.1 portfolio optimisation: compare vanilla Markowitz with
     - 3.1.1 RMT-clean correlation matrices
     - 3.1.2 shrinkage
     - 3.1.3 an alternate cost function (e.g. Rényi entropy)
   - 3.2 strategies of some kind
     - 3.2.1 clustering of days
     - 3.2.2 off-the-mill trading strategies
     - 3.2.3 alternate data (Twitter), Google Trends, ...

Your project: which strategies?

The aim is to an another systematic computation layer

- Portfolio building methods
- Define market states ($+-+$, clustering of days, ...)
  - Compute conditional performance measures for all assets
  - Rank, invest
  - Machine learning (next two weeks)
- [quantopian.com]: implement and test on-line strategies. Clone existing strategies as well.