Reading Group

Why change?

- Data sets are getting bigger
- New technologies
- Cost-effectiveness

Goals

- New tools, improved methods
- *Easier* to manage, explore "small" data
- *Possible* to process "big" data

Goals

- Learn new tools
- Learn to write code that's fast *where it counts*

What makes a language good?

- Productivity
  - Easy to write good code
  - Easy to read and modify
  - Most code already written
- Extensibility
- Ubiquity
- Speed?

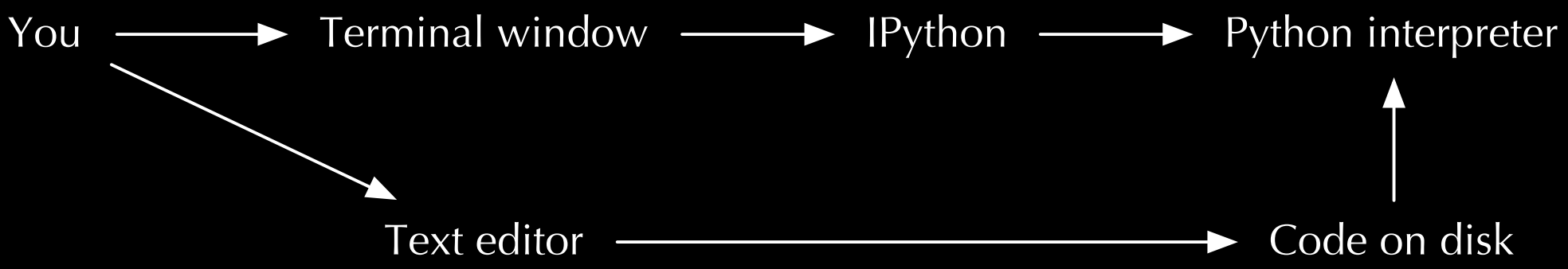
Why Python?

- All those things
- The way forward
- Python skills → MATLAB, etc.

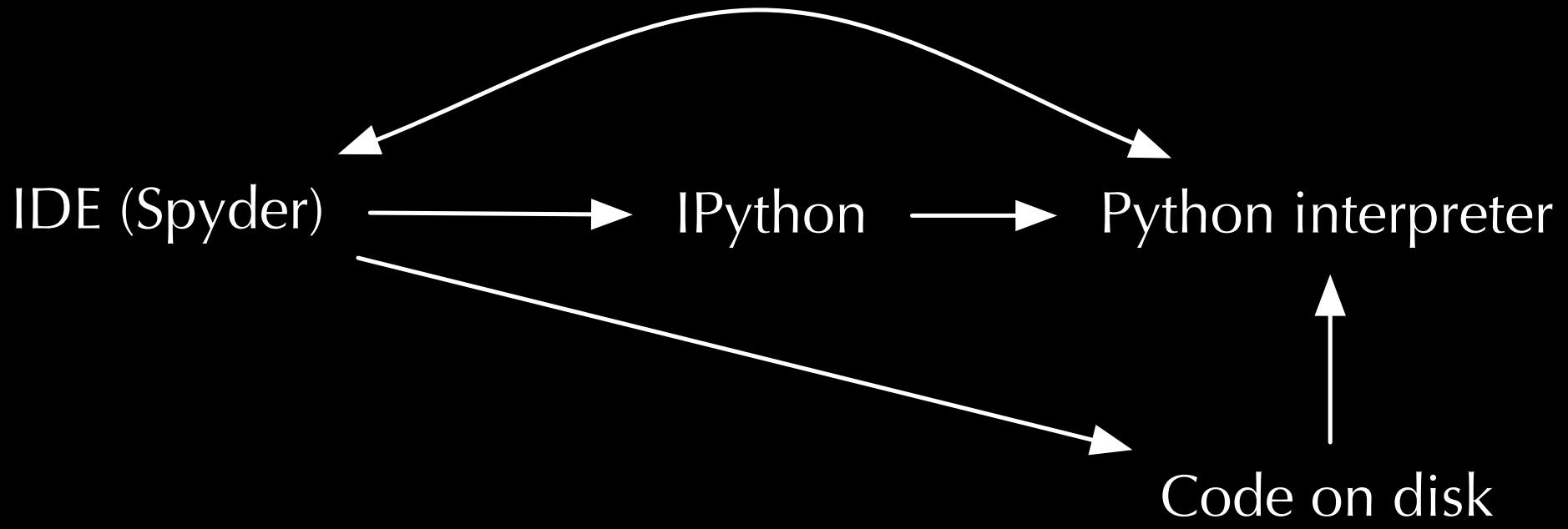
Today

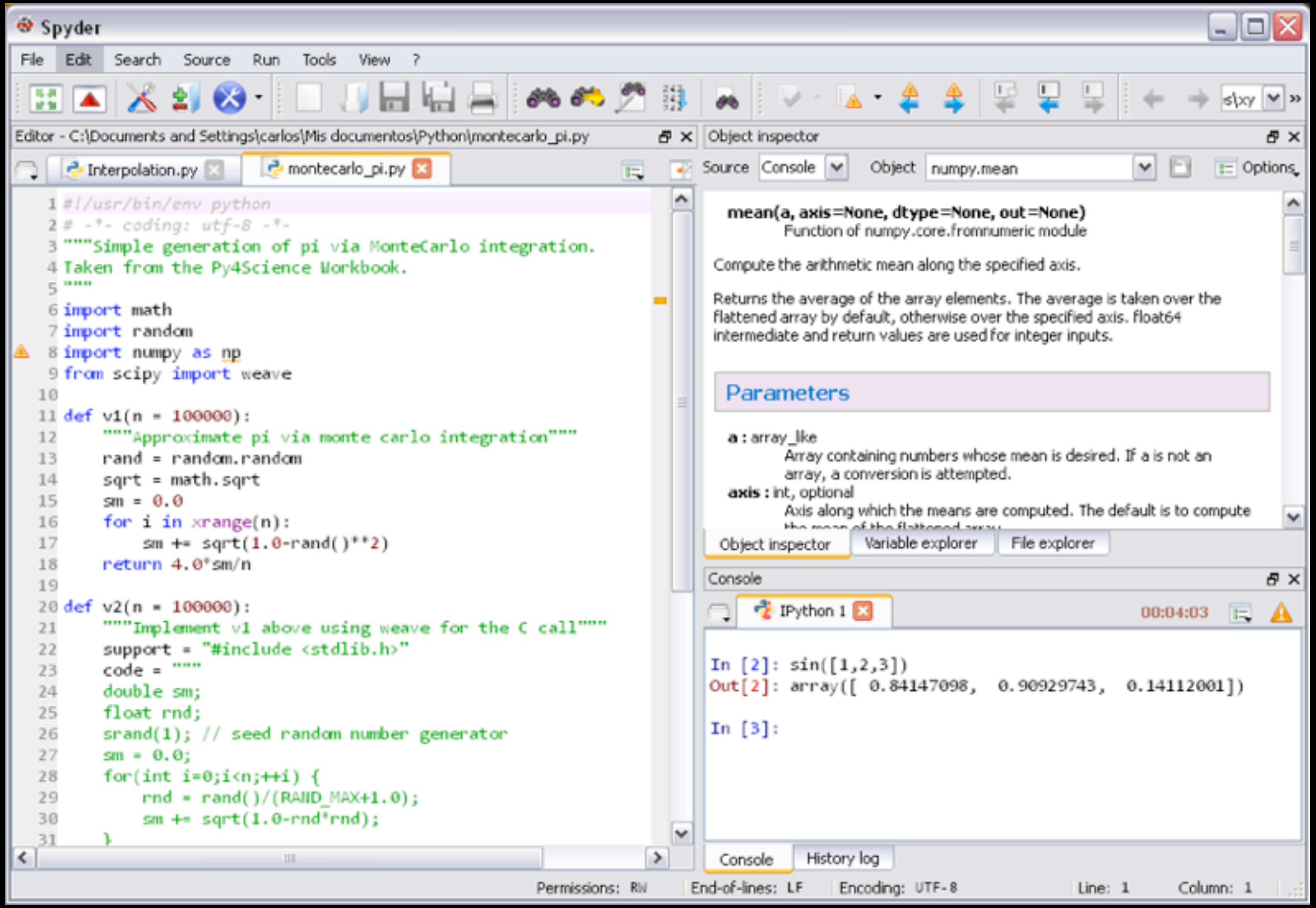
- Survey what Python is
- Plan future meetings

How do I use it?



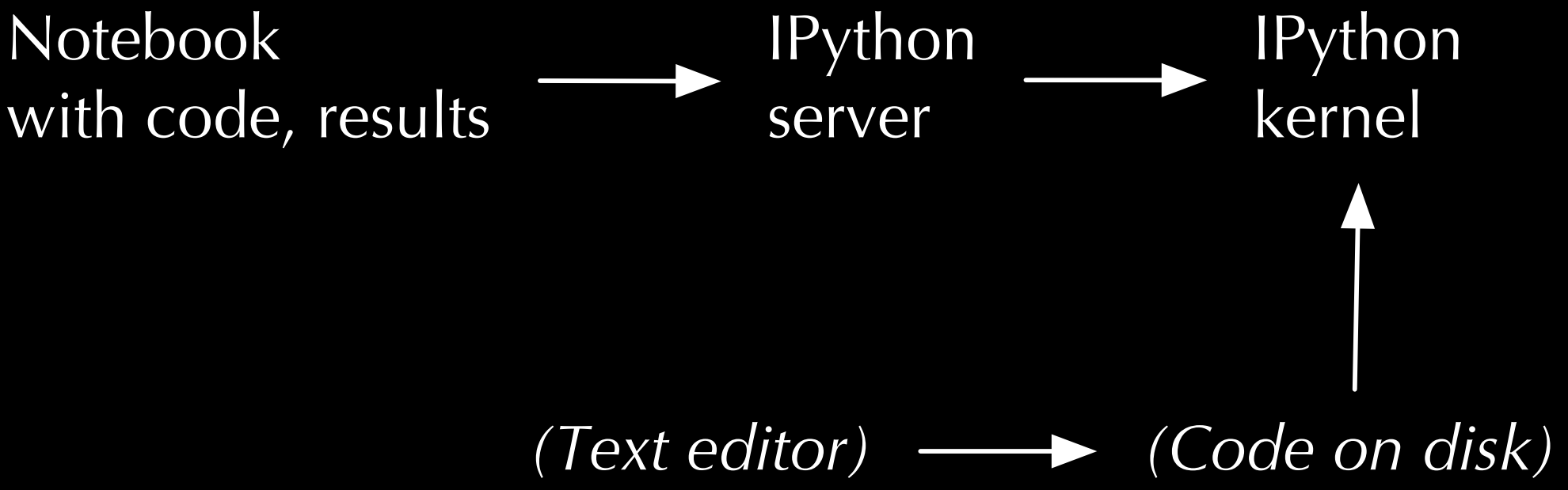
How do I use it?



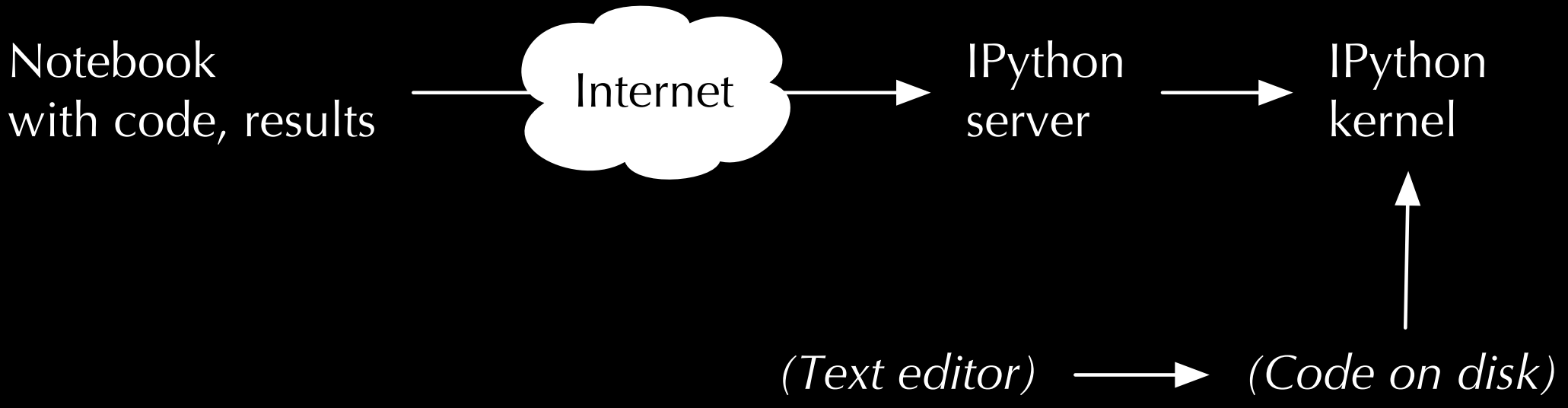

Spyder
File Edit Search Source Run Tools View ?
Editor - C:\Documents and Settings\carlos\Mis documentos\Python\montecarlo_pi.py
Interpolation.py
montecarlo_pi.py
#!/usr/bin/env python
-*- coding: utf-8 -*-
"""Simple generation of pi via MonteCarlo integration.
Taken from the Py4Science Workbook.
"""
import math
import random
import numpy as np
from scipy import weave
def v1(n = 100000):
"""Approximate pi via monte carlo integration"""
rand = random.random
sqrt = math.sqrt
sm = 0.0
for i in xrange(n):
sm += sqrt(1.0-rand()**2)
return 4.0*sm/n
def v2(n = 100000):
"""Implement v1 above using weave for the C call"""
support = "#include <stdlib.h>"
code = """
double sm;
float rnd;
srand(1); // seed random number generator
sm = 0.0;
for(int i=0;i<n;++i) {
rnd = rand()/(RAND_MAX+1.0);
sm += sqrt(1.0-rnd*rnd);
}
Object inspector
Source Console Object numpy.mean Options
mean(a, axis=None, dtype=None, out=None)
Function of numpy.core.fromnumeric module
Compute the arithmetic mean along the specified axis.
Returns the average of the array elements. The average is taken over the flattened array by default, otherwise over the specified axis. float64 intermediate and return values are used for integer inputs.
Parameters
a : array_like
Array containing numbers whose mean is desired. If a is not an array, a conversion is attempted.
axis : int, optional
Axis along which the means are computed. The default is to compute
Object inspector
Variable explorer
File explorer
Console
IPython 1
00:04:03
In [2]: sin([1,2,3])
Out[2]: array([0.84147098, 0.90929743, 0.14112001])
In [3]:
Console
History log
Permissions: RW
End-of-lines: LF
Encoding: UTF-8
Line: 1
Column: 1

Notebook demo

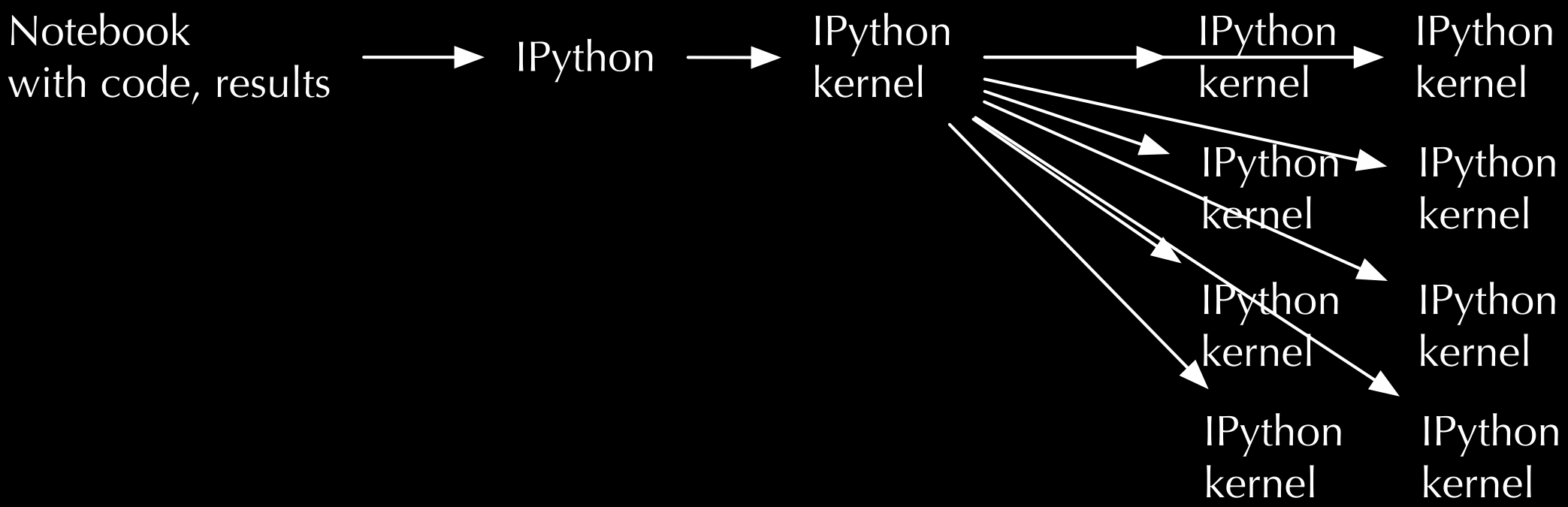
How do I use it?



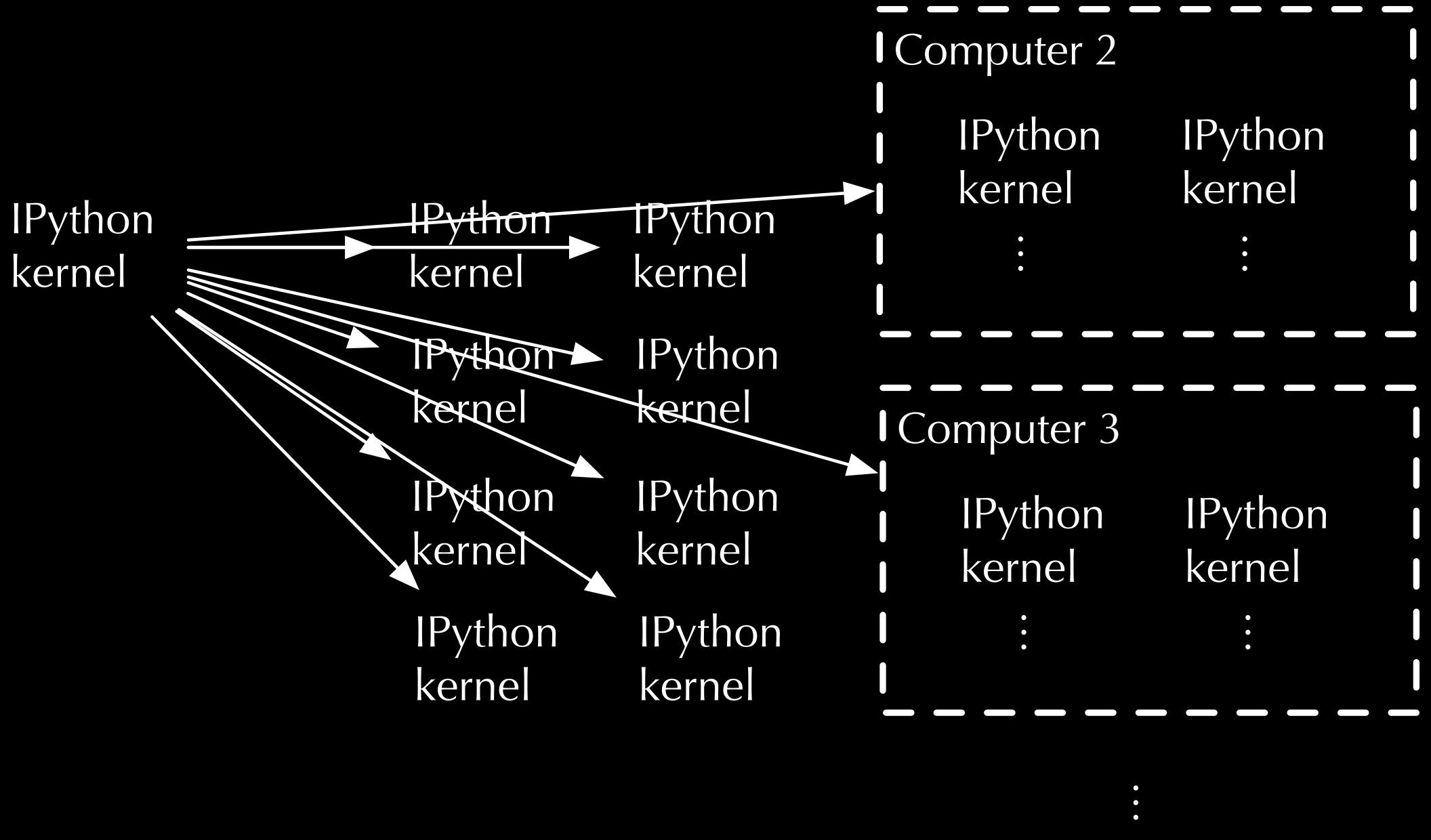
How do I use it?



How do I use it?



How do I use it?



… etc.

- Scientific
  - scikit-image, OpenCV, MayaVi
  - pandas, statsmodels, scikit-learn, sympy
- Big data
  - pytables, disco, PiCloud
- High-performance
  - cython, pyCUDA
- Scientific GUIs: Enthought Tool Suite

…and more

- Send email, serve webpages, etc.
- Access Google Docs
- Write an IDE
- Run YouTube, Yelp, reddit, …

Get yourself started!

- Install Python (EPD Free)
- “`ipython notebook`”
- Python Tutorial (v2.7)
- My example notebooks

Future meetings

Mix and match

- Array math, plotting
- Image processing
- Sharing data w/ Matlab, IDL, DataGraph, etc.
- Version control and the UNIX shell
- Parallel processing
- Working with particle tracks (pandas)
- Curve fitting & statistics
- cython
  (and when to use it)
- 3D visualization (MayaVi)
- **Anything you get excited about!**