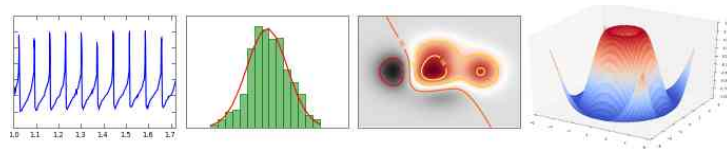


Matplotlib

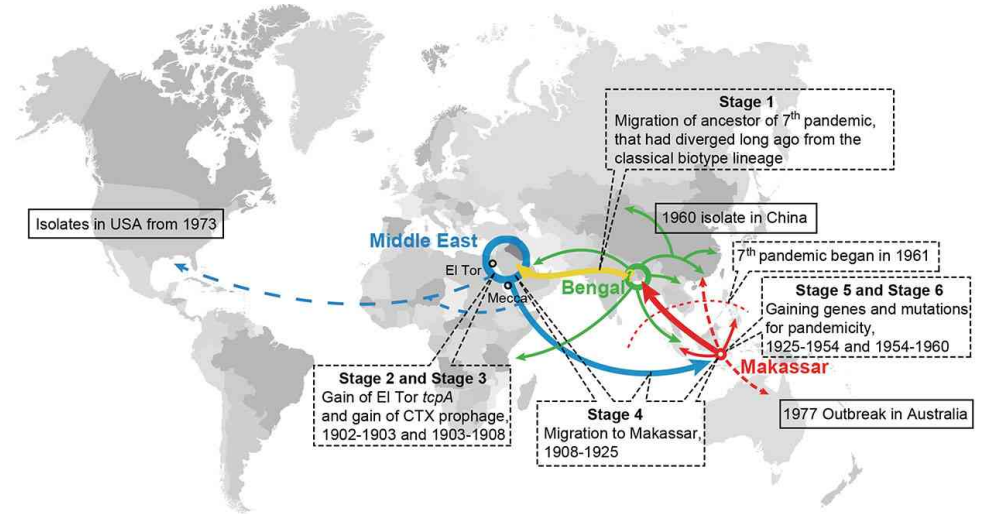
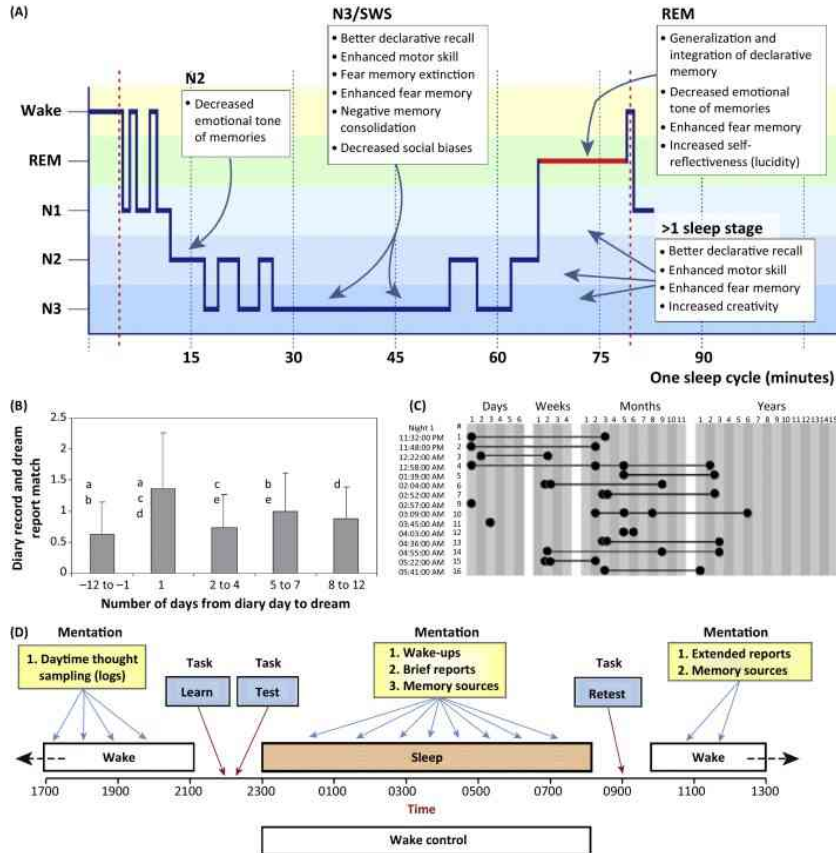
Let's do something useful for a change

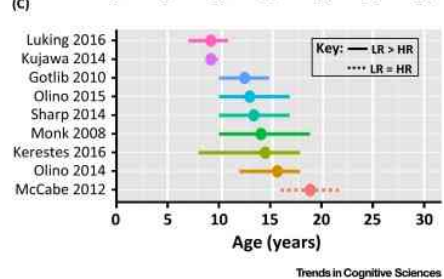
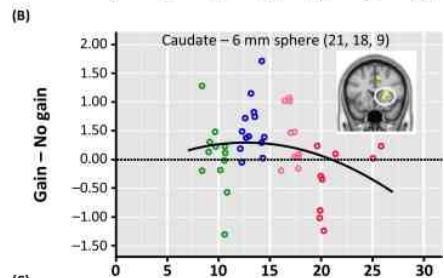
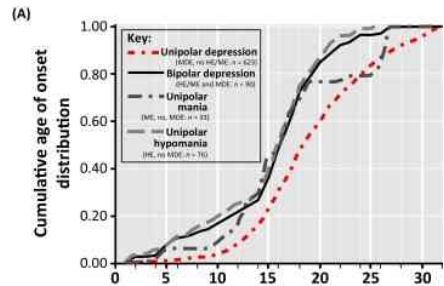


Matplotlib

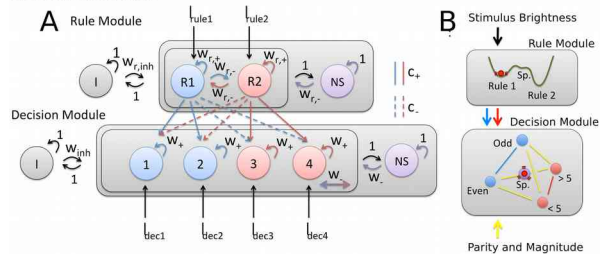
- Programming plots instead of making them in excel, SPSS,...
 - Write a program that constructs a plot step by step
- Why?
 - Flexibility: scientific plotting has moved beyond single line graphs
 - Exploration and on-line visualization
 - Data might change
 - Your ideas might change
 - Consistency: define a style and use in all plots
 - Quality – there is no substitute for a good plot

Flexibility

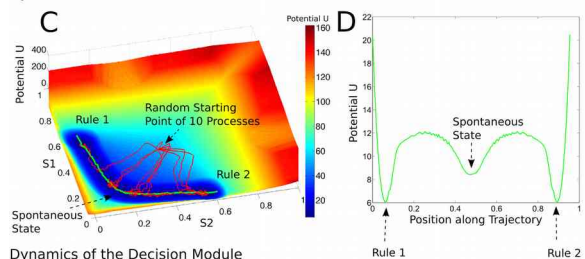




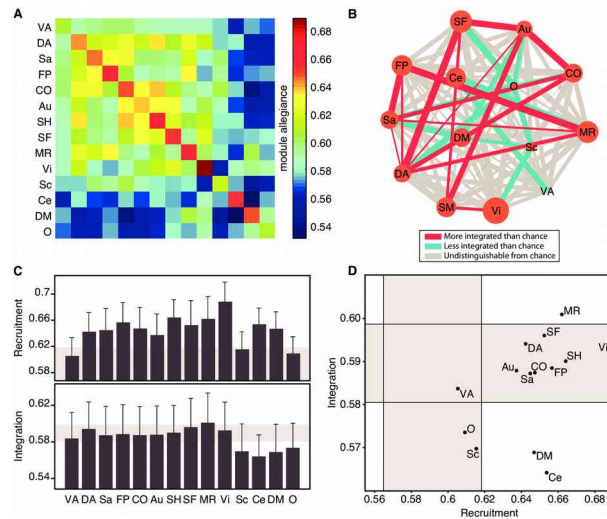
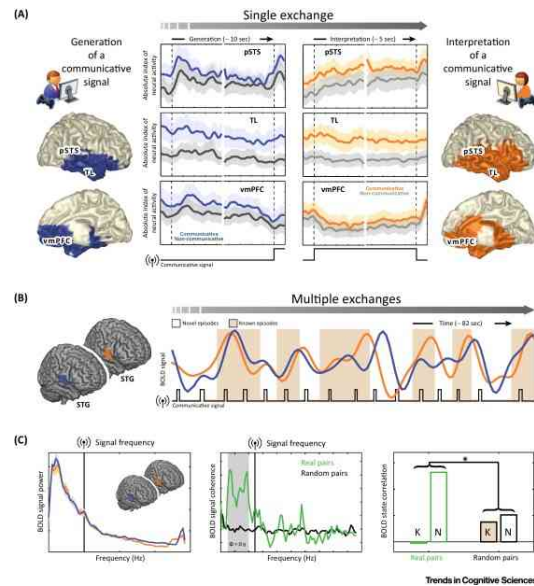
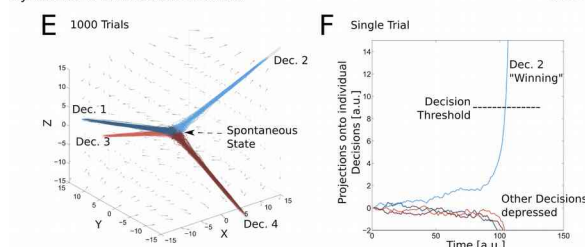
General Architecture



Dynamics of the Rule Module



Dynamics of the Decision Module



But it's not just science journals

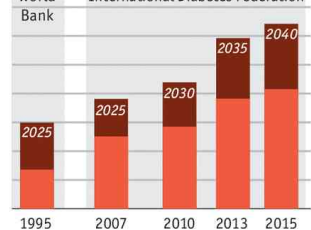
A growing problem

Diabetes prevalence*

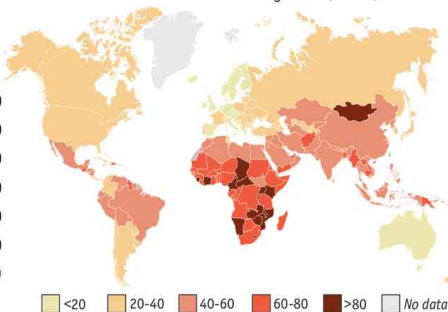
Worldwide population with the condition, m

Estimate Projected increase, by year
By organisation

World Bank International Diabetes Federation

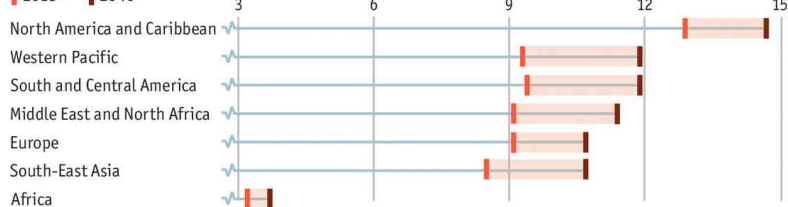


Share of diabetes deaths before the age of 60, 2015, %



Regional prevalence, as % of adult population

2015† 2040‡



Sources: World Bank; IDF Diabetes Atlas, 2016

*Aged 20- to -79-year-olds

†Estimate

‡Forecast

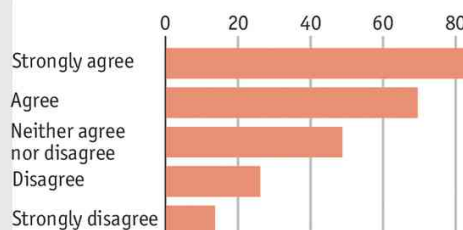
Economist.com

It's me not EU

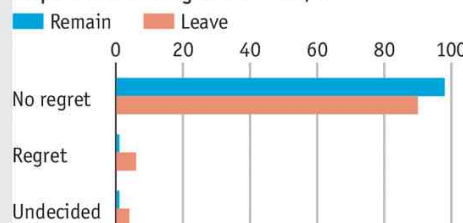
Polling before and after European Union referendum*

I'd rather put my trust in the wisdom of ordinary people than in the opinion of experts

Respondents who also voted to leave EU, %



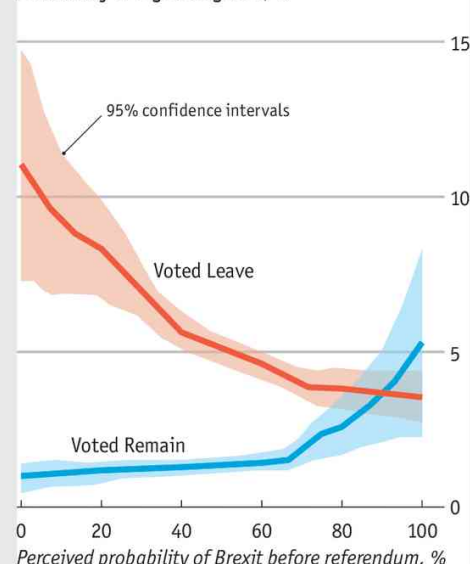
Respondents who regret their vote, %



Source: British Election Study

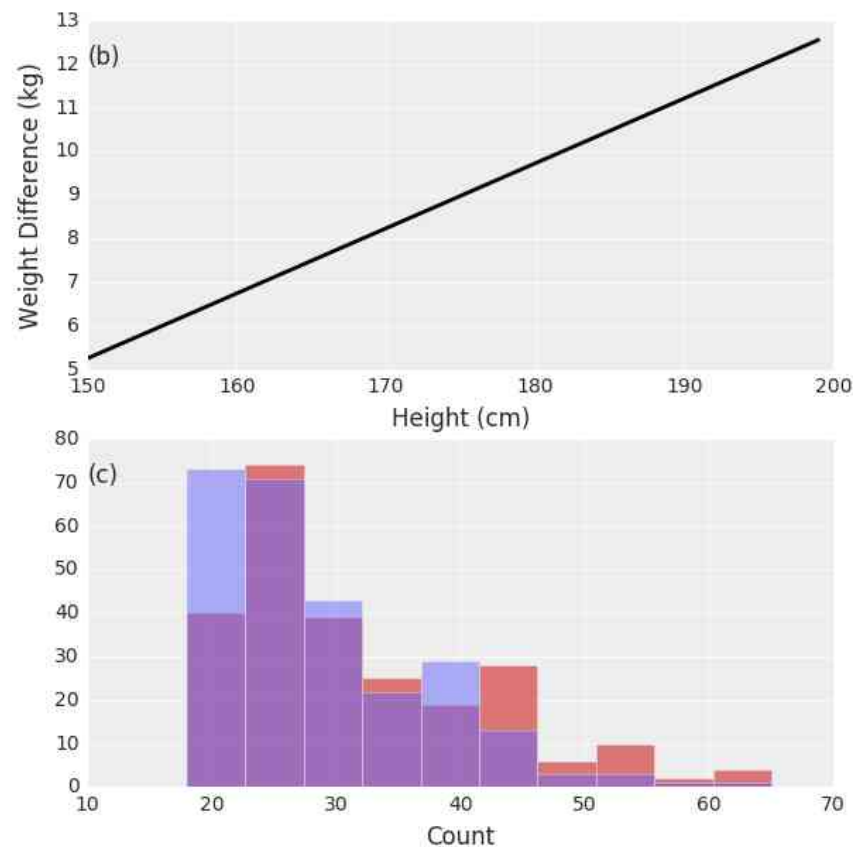
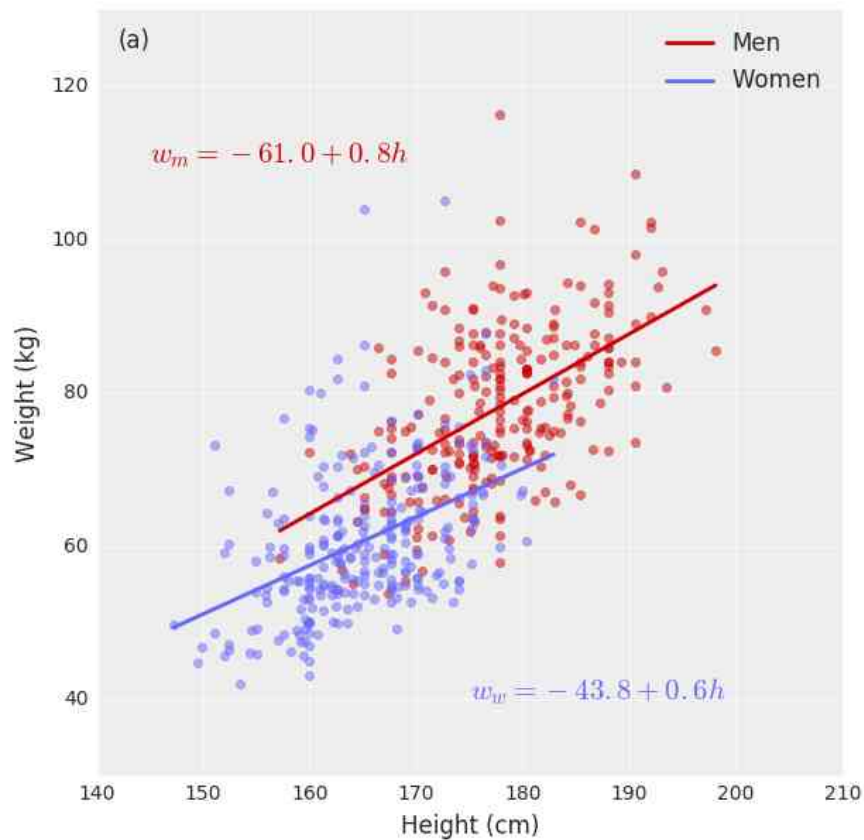
Economist.com

Probability of regretting vote, %



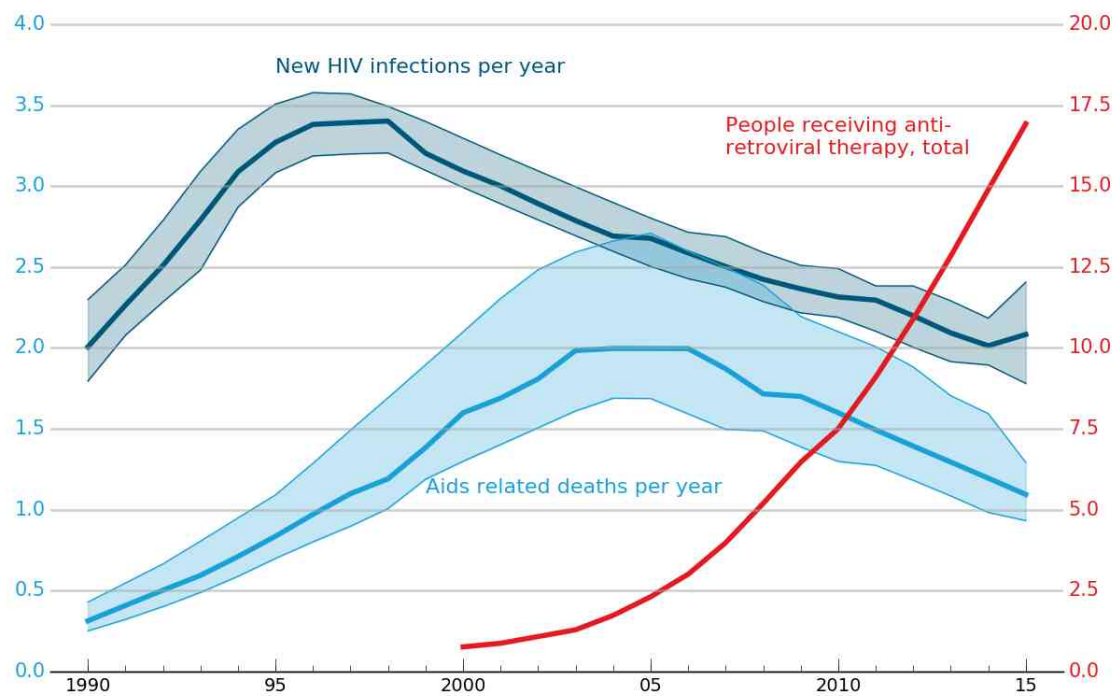
*Fieldwork: 28,069 respondents between April and July 2016

Example 1



Example 2

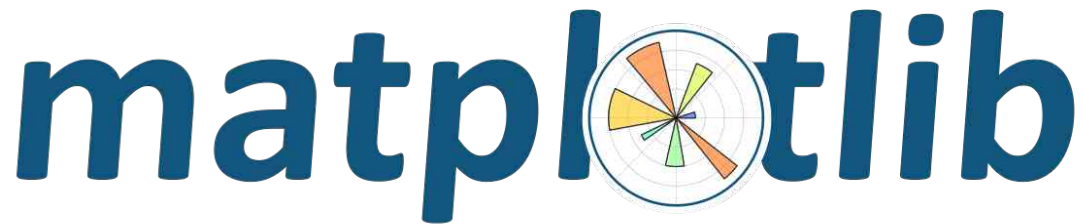
Keeping the pressure up



Typical work flow

- Enter or Read in the data
- Plot it quick and dirty
- Spend **a lot of** time refining, tuning and polishing the plot
 - Change aspect ratio, choose colors, line styles, add titles, labels, annotations
- Think you're done
- Refine some more
- Enjoy the work of art you've created
- Mock excel and spss graphs

Matplotlib



- Pyplot is the (sub)module that contains most plotting commands
- Documentation:
 - Go to the matplotlib website
 - Go to pyplot: <https://tinyurl.com/y8xslmv6>
 - Go over the **scatter** command

`scatter`

Make a scatter plot of x vs y. Marker size is scaled by s and marker color is mapped to c. Parameters

——— x, y : array_like, shape (n,) Input data s : scalar or array_like, shape (n,), optional size in points².

Matplotlib

matplotlib

```
matplotlib.pyplot.scatter(x, y, s=None, c=None, marker=None, cmap=None, norm=None, vmin=None, vmax=None, alpha=None,  
linewidths=None, verts=None, edgecolors=None, hold=None, data=None, **kwargs)
```

●
X and Y are required
arguments

●
All the others are
optional and their default
value is listed

```
matplotlib.pyplot.scatter(x, y, s=None, c=None, marker=None, cmap=None, norm=None, vmin=None, vmax=None, alpha=None, linewidths=None, verts=None, edgecolors=None, hold=None, data=None, **kwargs) ¶
```

Make a scatter plot of x vs y

Marker size is scaled by s and marker color is mapped to c

Parameters:

x, y : array_like, shape (n,)

Input data

Input data

s : scalar or array_like, shape (n,), optional

size in points². Default is rcParams['lines.markersize'] ** 2.

Size of the points

c : color, sequence, or sequence of color, optional, default: 'b'

c can be a single color format string, or a sequence of color specifications of length N, or a sequence of N numbers to be mapped to colors using the cmap and norm specified via kwargs (see below). Note that c should not be a single numeric RGB or RGBA sequence because that is indistinguishable from an array of values to be colormapped. c can be a 2-D array in which the rows are RGB or RGBA, however, including the case of a single row to specify the same color for all points.

Color of the points

marker : MarkerStyle, optional, default: 'o'

See [markers](#) for more information on the different styles of markers scatter supports. marker can be either an instance of the class or the text shorthand for a particular marker.

Markers used [See link]

cmap : Colormap, optional, default: None

A Colormap instance or registered name. cmap is only used if c is an array of floats. If None, defaults to rc.image.cmap.

norm : Normalize, optional, default: None

A Normalize instance is used to scale luminance data to 0, 1. norm is only used if c is an array of floats. If None, use the default normalize().

vmin, vmax : scalar, optional, default: None

vmin and vmax are used in conjunction with norm to normalize luminance data. If either are None, the min and max of the color array is used. Note if you pass a norm instance, your settings for vmin and vmax will be ignored.

alpha : scalar, optional, default: None

The alpha blending value, between 0 (transparent) and 1 (opaque)

linewidths : scalar or array_like, optional, default: None

If None, defaults to (lines.linewidth),

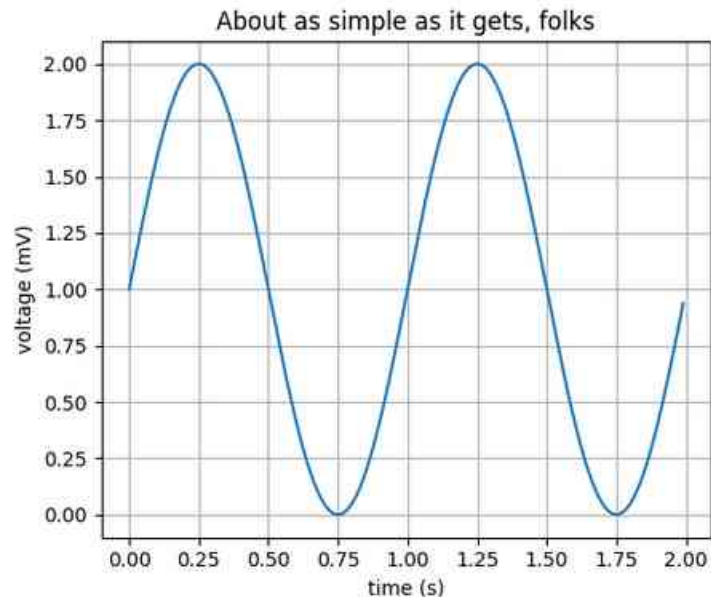
Matplotlib



- Gallery: good source of examples
 - Go to the matplotlib website
 - Go to the gallery:
 - `simple_plot`

pylab_examples example code: simple_plot.py

([Source code](#), [png](#), [pdf](#))



```
import matplotlib.pyplot as plt
import numpy as np

t = np.arange(0.0, 2.0, 0.01)
s = 1 + np.sin(2*np.pi*t)
plt.plot(t, s)

plt.xlabel('time (s)')
plt.ylabel('voltage (mV)')
plt.title('About as simple as it gets, folks')
plt.grid(True)
plt.savefig("test.png")
plt.show()
```

Keywords: python, matplotlib, pylab, example, codex (see [Search examples](#))

seaborn

- <https://seaborn.pydata.org/index.html>
- Built on top of matplotlib

Other plotting packages: bokeh

- New plot types
- Interactivity
- Large data

Texas Unemployment, 2009

