Scientific Plots in Practice

Analysis and Visualization of Big Data

Franziska Peter and Josep Perelló

UNIVERSITAT DE BARCELONA

Contents - Creating and editing graphics

Session II

Topic 1: Scientific Plots in Practice. Static plots for scientific publication. Python plotting libraries.

Sessions VI, XI, XII

Interactive plots for web apps. Displaying Geographical Data. Dashboards.

UNIVERSITAT DE BARCELONA

Some references

1. https://matplotlib.org/, https://seaborn.pydata.org/

2. Dynamical Systems with Applications using Python, Stephen Lynch, https://doi-org.sire.ub.edu/10.1007/978-3-319-78145-7, Springer Nature 2018

3. Rougier, Droettboom, Bourne (2014) Ten Simple Rules for Better Figures. PLOS Comp. Biology 10(9): e1003833. https://doi.org/10.1371/journal.pcbi.1003833

4. Python and Matplotlib Essentials for Scientists and Engineers, Matt A Wood, http://dx.doi.org/10.1088/978-1-6270-5620-5, IOP Publishing 2015

5. Essential Python for the Physicist, Giovanni Moruzzi, https://doi-org.sire.ub.edu/10.1007/978-3-030-45027-4, Springer Nature Switzerland AG 2020

6. Introduction to Scientific Programming with Python, Joakim Sundnes, https://doi-org.sire.ub.edu/10.1007/978-3-030-50356-7, Simula SpringerBriefs on Computing 2020

*. https://matplotlib.org/cheatsheets/cheatsheets.pdf

UNIVERSITAT DE BARCELONA

Evaluation

Gradual and incremental set of tasks (in class and through Campus Virtual)

Task 1: Data Management Plan Forensics, in group (Tues 9, JPerelló): 10%

Task 2: Sharing code in Github, individual (Wed 10, FPeter): 10%

Task 3: Write an abstract (Mon 15, JPerelló): 10%

Task 4: Create a dashboard (Thu 18, FPeter): 30%

Task 5: Oral presentation, in group (Fri 19, JPerelló + FPeter): 40%

To set a group between 2 and 4. You will work together during the course.

UNIVERSITAT DE BARCELONA

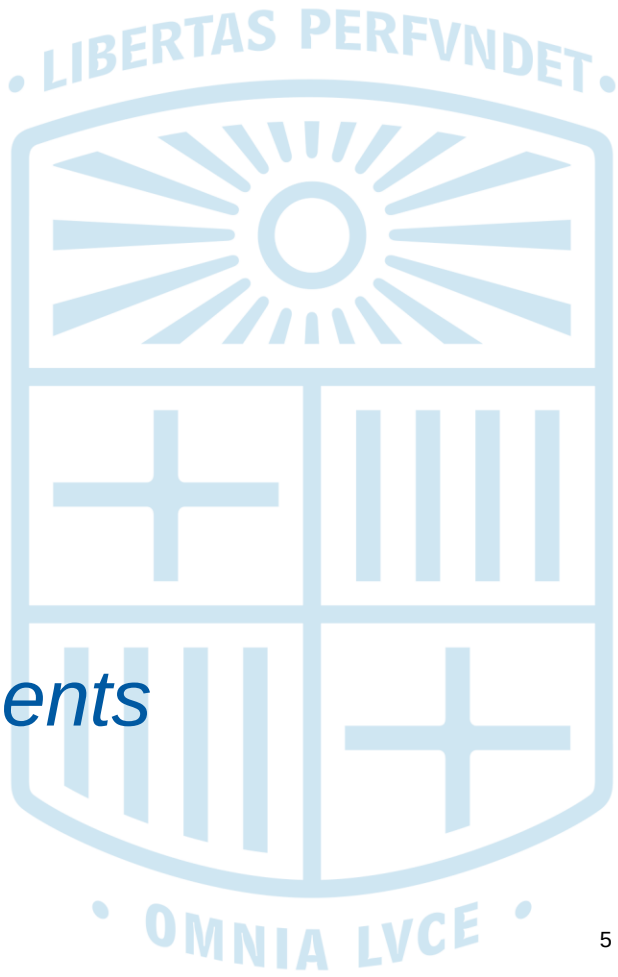
Outline

1. Python plotting libraries

2. DIY: 6 plots in 30min

3. Dive into Matplotlib

4. Scientific Journals' Requirements



UNIVERSITAT DE BARCELONA

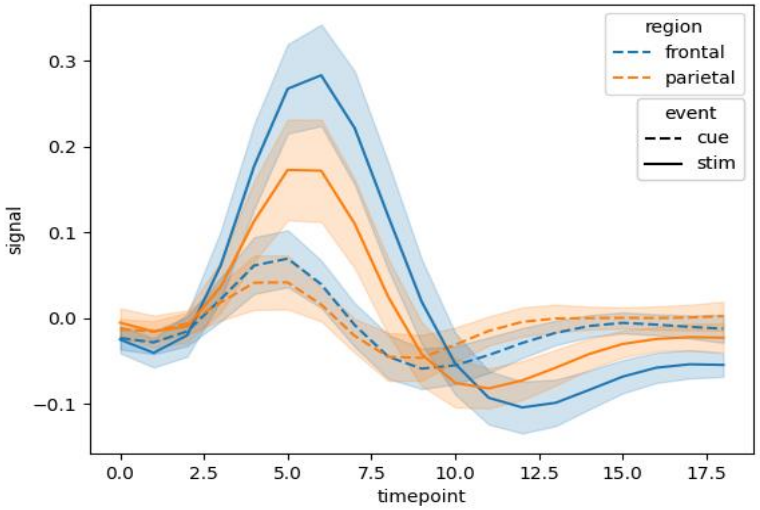
Python plotting libraries

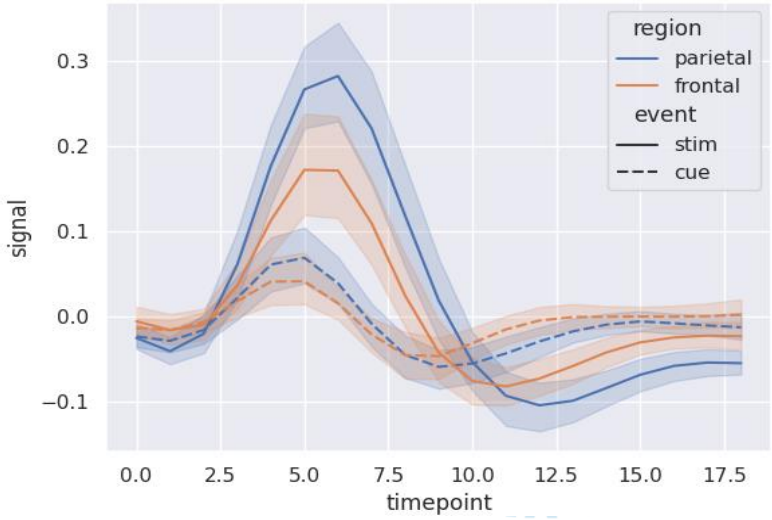
	Matplotlib (mpl)	Seaborn
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl
cons	- complex syntax - not interactive	- gaining full control is harder

UNIVERSITAT DE BARCELONA

Python plotting libraries

	Matplotlib (mpl)	Seaborn
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl
cons	- complex syntax - not interactive	- gaining full control is harder

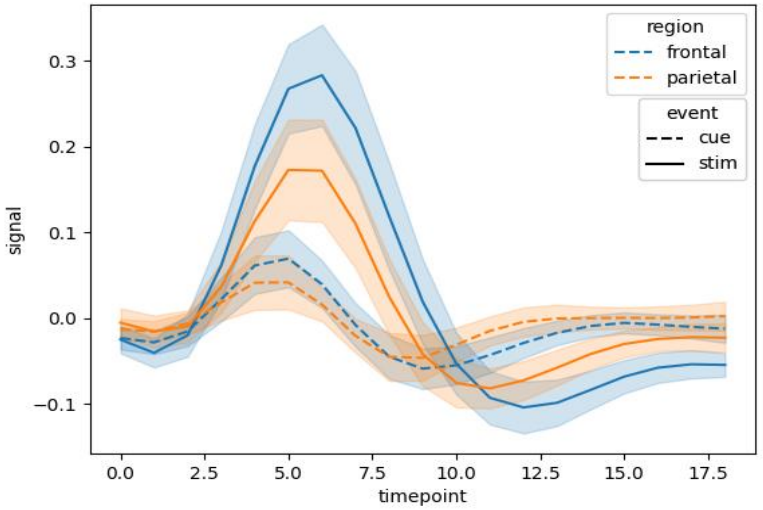




UNIVERSITAT DE BARCELONA

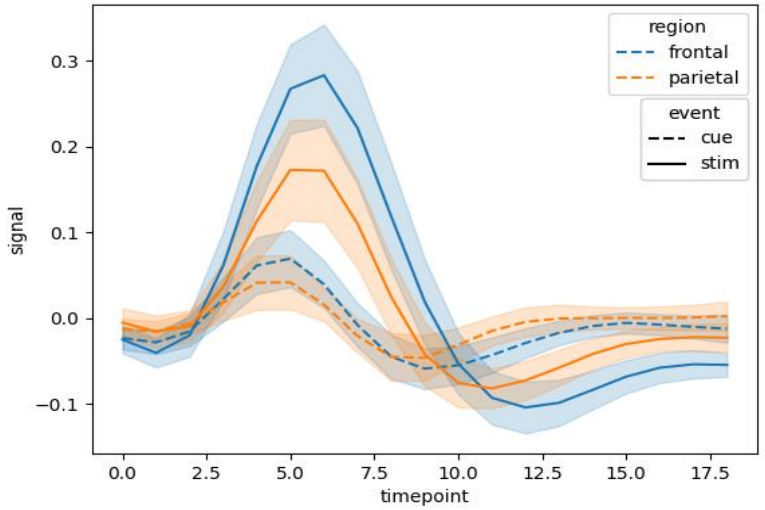
Matplotlib vs Seaborn

```
import seaborn as sns
import matplotlib.pyplot as plt
from scipy.stats import sem

# Load an example dataset with long-form data
fmri = sns.load_dataset("fmri") # returns panda dataframe

# sort data in chronological order
fmri_mpl = fmri.sort_values('timepoint')

# create canvas
fig, ax =plt.subplots()

# chose color and linestyle for the different curves
colors = {"parietal":"C0", "frontal":"C1"}
line_styles = {"stim":"solid", "cue":"dashed"}

# plot curves
for event in set(fmri_mpl["event"]):
    # filter for one event
    fmri_mpl_ev = fmri_mpl[fmri_mpl["event"]==event]
    for region in set(fmri_mpl_ev["region"]):
        # filter for one region
        fmri_mpl_re = fmri_mpl_ev[fmri_mpl_ev["region"]==region]
        # calculate mean and 95% confidence interval with scipy stats
        df_mean = fmri_mpl_re.groupby('timepoint').signal.mean()
        df_se = fmri_mpl_re.groupby('timepoint').signal.apply(sem).mul(1.96)
        # plot mean
        plt.plot(df_mean, # time stamps and mean are in one DataFrame
                 color = colors[region], # color according to region
                 ls = line_styles[event]), # linestyle according to event
                 #label = region)
        # plot confidence interval
        plt.fill_between(df_mean.index, df_mean - df_se, df_mean + df_se, color=colors[region], alpha = 0.2)
```



UNIVERSITAT DE BARCELONA

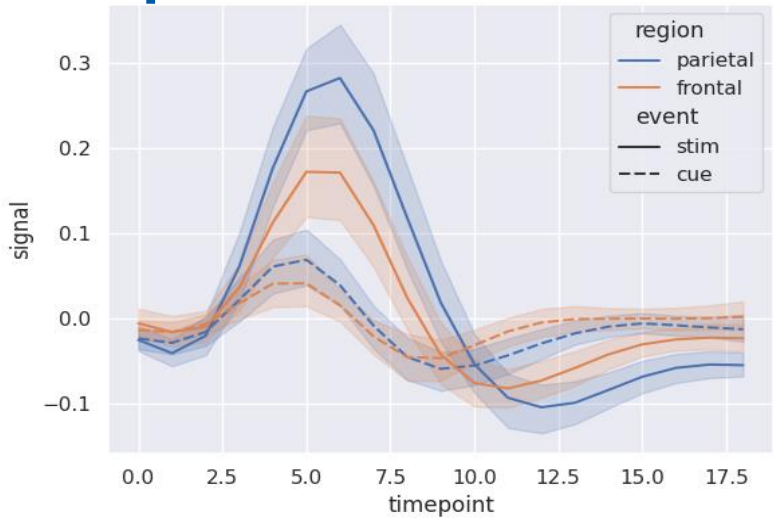
Matplotlib vs Seaborn




```
# axes labels
ax.set_xlabel('timepoint')
ax.set_ylabel("signal")

# dummy lines for legend
dummy_lines = []
for event in set(fmri_mpl["event"]):
    dummy_lines.append(ax.plot([],[], c="black", ls = line_styles[event])[0])
lines = ax.get_lines()
# legend regions
legend1 = plt.legend([lines[i] for i in [0,1]], ["frontal", "parietal"],
                     title = "region",
                     loc="upper right")
# legend events
legend2 = plt.legend([dummy_lines[i] for i in [0,1]], ["cue", "stim"],
                     title = "event",
                     loc="center right",
                     bbox_to_anchor=(1,0.7))

ax.add_artist(legend1)
ax.add_artist(legend2)

# show in API
plt.show()
```

UNIVERSITAT DE BARCELONA

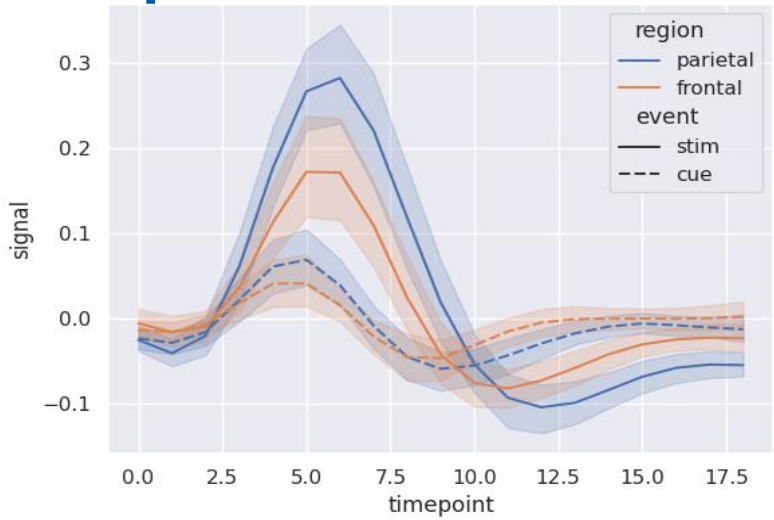
Seaborn vs Matplotlib

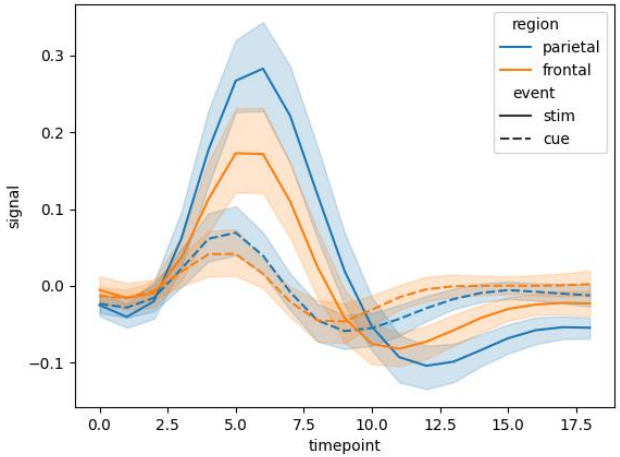


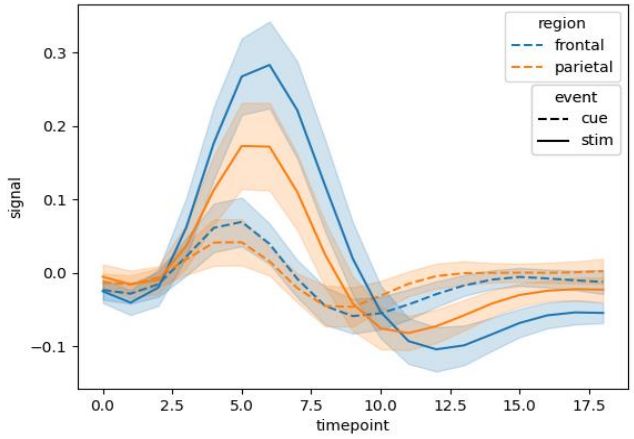

```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set_theme(style="darkgrid")

# Load an example dataset with long-form data
fmri = sns.load_dataset("fmri")

# Plot the responses for different events and regions
sns.lineplot(x="timepoint", y="signal",
             hue="region", style="event",
             data=fmri)
plt.show()
```

UNIVERSITAT DE BARCELONA

Seaborn vs Matplotlib





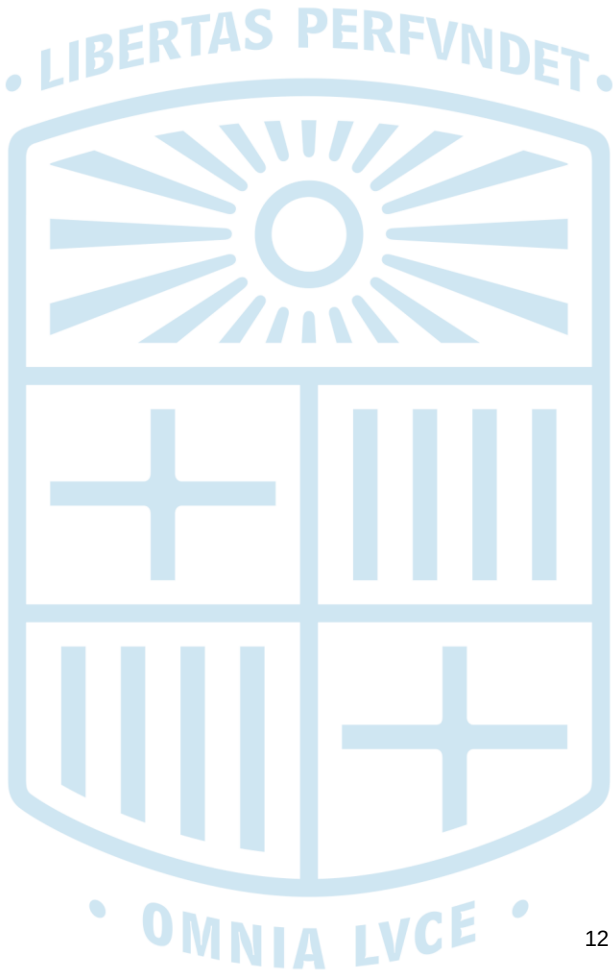



```
import seaborn as sns
import matplotlib.pyplot as plt

sns.set_theme(style="darkgrid")

# Load an example dataset with long-form data
fmri = sns.load_dataset("fmri")

# Plot the responses for different events and regions
sns.lineplot(x="timepoint", y="signal",
             hue="region", style="event",
             data=fmri)
plt.show()
```

UNIVERSITAT DE BARCELONA

Python plotting libraries

	Matplotlib (mpl)	Seaborn
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl
cons	- complex syntax - not interactive	- gaining full control is harder



UNIVERSITAT DE BARCELONA

Python plotting libraries

	Matplotlib (mpl)	Seaborn	Plotly	Bokeh	Altair
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes	for web apps (essentially JavaScript)	for web apps (output = html)	for web apps
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl	- interactive plots - quick plotting with plotly.express - 3D - >community than bokeh	- interactive plots	- focus on statistics
cons	- complex syntax - not interactive	- gaining full control is harder	- might not stay FOSS for ever - different syntax than mpl (but similar concepts)	- only interactive plots - different syntax than mpl	- no 3D - not fully customizable - different syntax than mpl

UNIVERSITAT DE BARCELONA

Python plotting libraries

	Matplotlib (mpl)	Seaborn	Plotly	Bokeh	Altair
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes	for web apps (essentially JavaScript)	for web apps (output = html)	for web apps
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl	- interactive plots - quick plotting with plotly.express - 3D - >community than bokeh	- interactive plots	- focus on statistics
cons	- complex syntax - not interactive	- gaining full control is harder	- might not stay FOSS for ever - different syntax than mpl (but similar concepts)	- only interactive plots - different syntax than mpl	- no 3D - not fully customizable - different syntax than mpl

THURSDAY: interactive visualizations

Python plotting libraries

	Matplotlib (mpl)	Seaborn	Plotly	Bokeh	Altair	pandas
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes	for web apps (essentially JavaScript)	for web apps (output = html)	for web apps	essentially mpl
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl	- interactive plots - quick plotting with plotly.express - 3D - >community than bokeh	- interactive plots	- focus on statistics	- quick plot of DataFrames - geopandas
cons	- complex syntax - not interactive	- gaining full control is harder	- might not stay FOSS for ever - different syntax than mpl (but similar concepts)	- only interactive plots - different syntax than mpl	- no 3D - not fully customizable - different syntax than mpl	- very basic

Python plotting libraries

	Matplotlib (mpl)	Seaborn	Plotly	Bokeh	Altair	pandas
	oldest python plotting library (gold standard)	built on top of mpl, contains ready-made themes	for web apps (essentially JavaScript)	for web apps (output = html)	for web apps	essentially mpl
pros	- complete control - almost anything is possible - 3D - animations	- simple syntax - if necessary, still full control - more appealing than mpl	- interactive plots - quick plotting with plotly.express - 3D - >community than bokeh	- interactive plots	- focus on statistics	- quick plot of DataFrames - geopandas
cons	- complex syntax - not interactive	- gaining full control is harder	- might not stay FOSS for ever - different syntax than mpl (but similar concepts)	- only interactive plots - different syntax than mpl	- no 3D - not fully customizable - different syntax than mpl	- very basic
gallery	matplotlib.org/stable/gallery	seaborn.pydata.org/examples	plotly.com/python	docs.bokeh.org/en/latest/docs/gallery.html	altair-viz.github.io/gallery/	pandas.pydata.org/pandas-docs/stable/user_guide/visualization.html

UNIVERSITAT DE BARCELONA

Python plotting libraries

See also:

- Mayavi (3D plots)

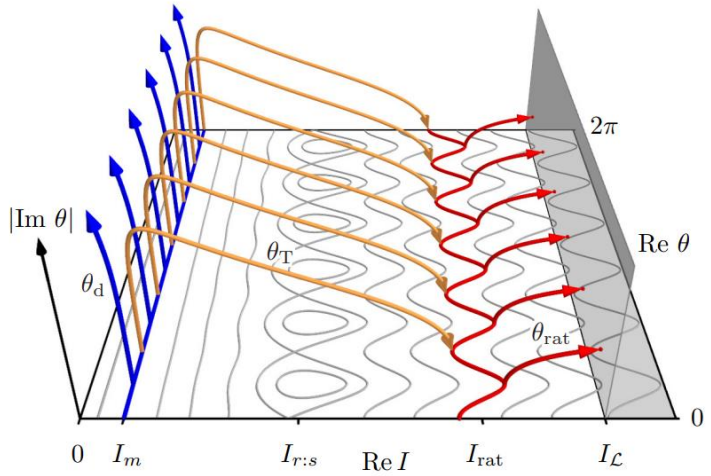


FIG. 2. (color online) Phase space of $\mathcal{H}_{r:s}$ (thin [gray] lines) and leaky region $\mathcal{L}$ (shaded area). Real tori an complex paths (thick lines and arrows) are labeled in the figure.

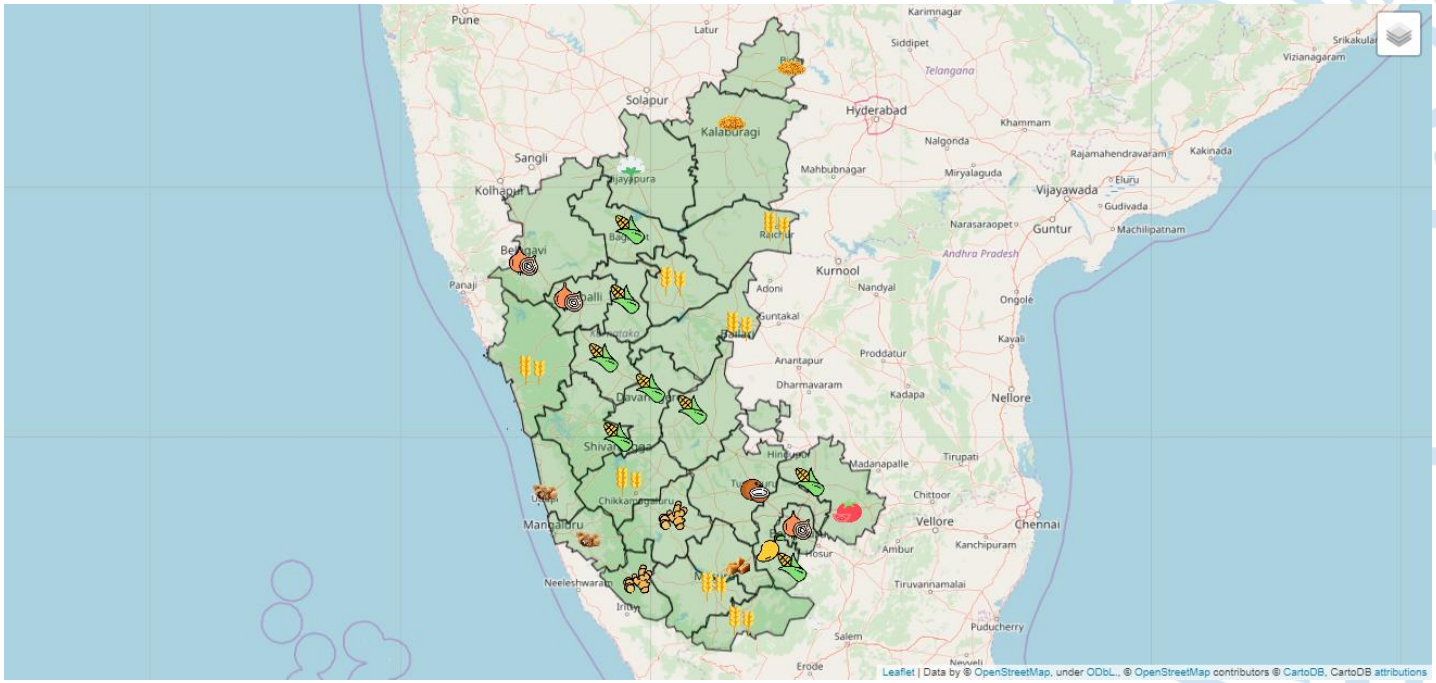
https://arxiv.org/pdf/1609.09276.pdf

UNIVERSITAT DE BARCELONA

Python plotting libraries

See also:

- Mayavi (3D plots)
- Folium (geodata mostly)

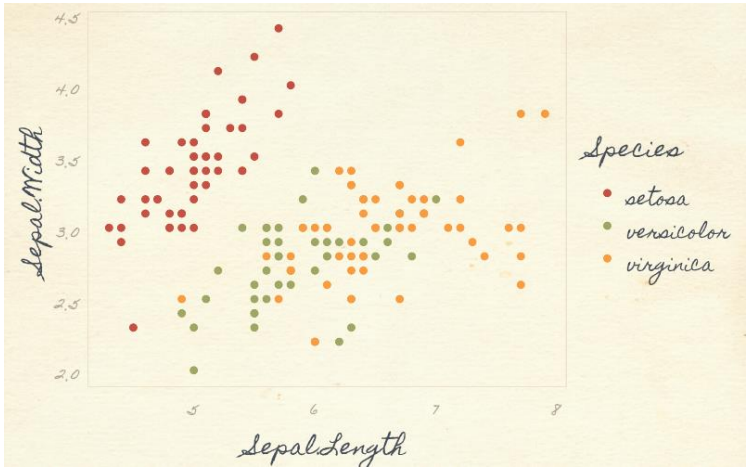


www.nagarajbhat.com/post/folium-visualization

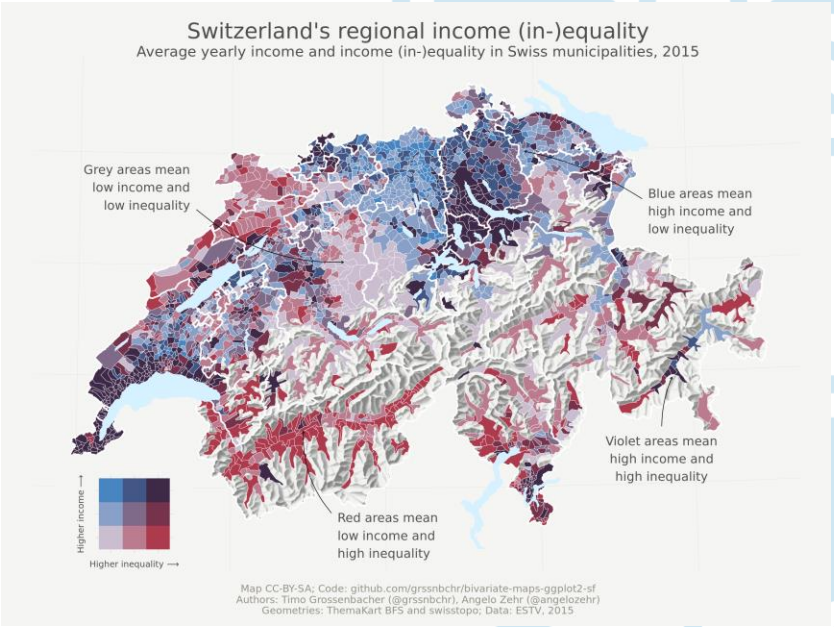
Python plotting libraries

See also:

- Mayavi (3D plots)
- Folium (geodata mostly)
- ggplot2 (fun themes and beautiful visualizations)



www.garrickadenbuie.com/project/ggpomological



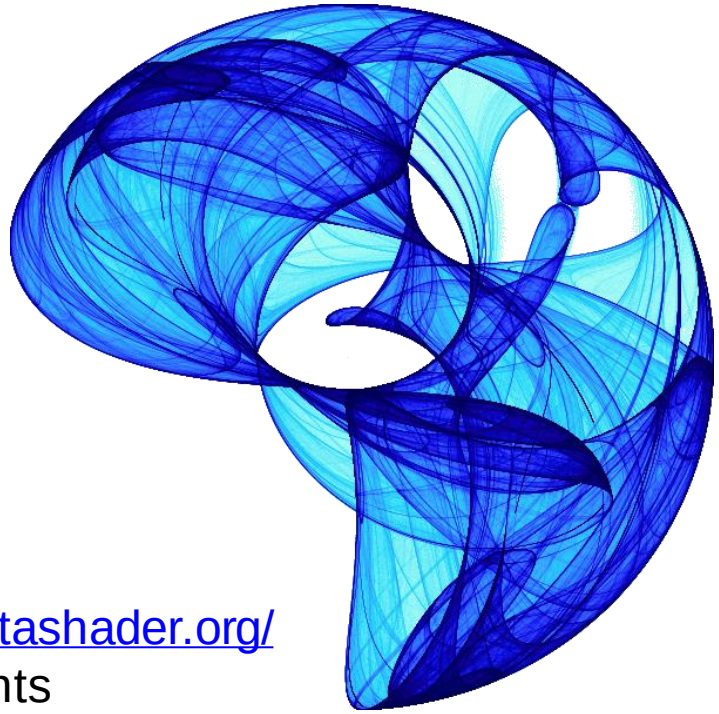
timogrossenbacher.ch/2019/04/bivariate-maps-with-ggplot2-and-sf

UNIVERSITAT DE BARCELONA

Python plotting libraries

See also:

- Mayavi (3D plots)
- Folium (geodata mostly)
- ggplot2 (fun themes and beautiful visualizations)
- Datashader (Big Data, plots: bokeh or holoviews)



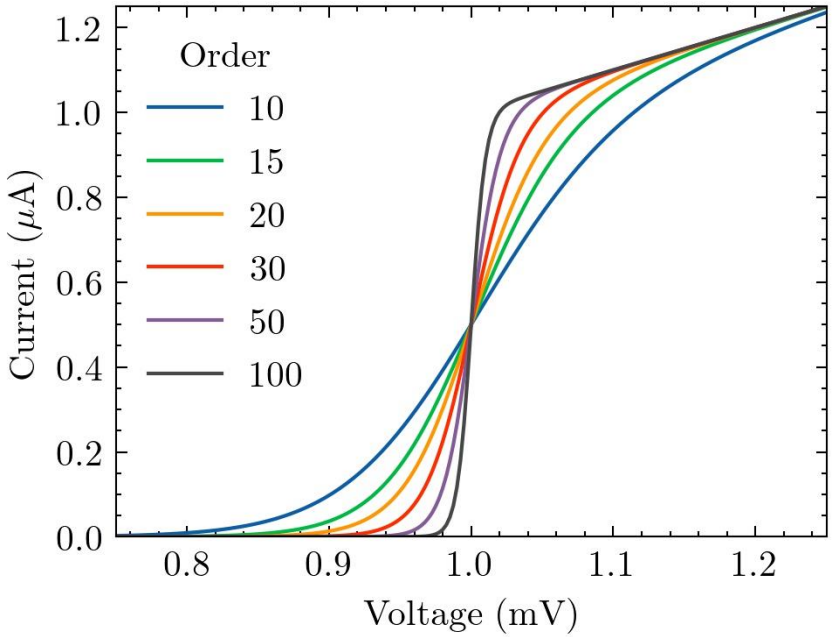
https://datashader.org/
$2 \cdot 10^8$ points

UNIVERSITAT DE BARCELONA

Python plotting libraries

See also:

- Mayavi (3D plots)
- Folium (geodata mostly)
- ggplot2 (fun themes and beautiful visualizations)
- Datashader (Big Data, plots: bokeh or holoviews)
- SciencePlots (adapted to several journals)

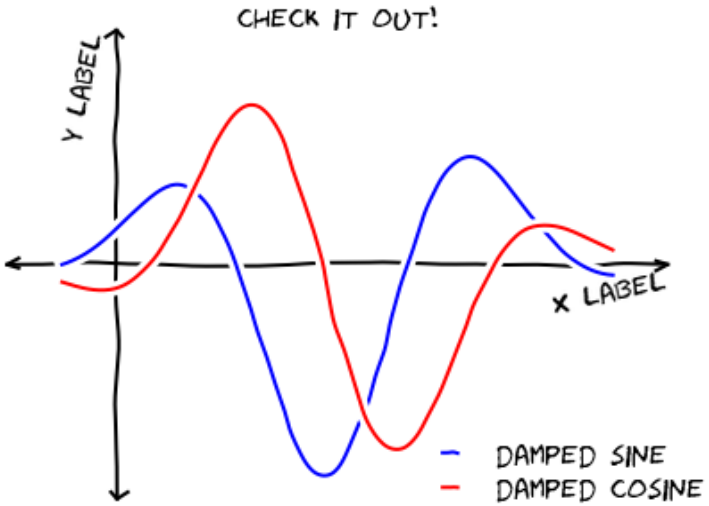


UNIVERSITAT DE BARCELONA

Python plotting libraries

See also:

- Mayavi (3D plots)
- Folium (geodata mostly)
- ggplot2 (fun themes and beautiful visualizations)
- Datashader (Big Data, plots: bokeh or holoviews)
- SciencePlots (adapted to several journals)
- Xkcd style (developer stage)



UNIVERSITAT DE BARCELONA

Python plotting libraries

NOW YOU! Try them out on your own:

- pick one plot from each gallery
- copy the source code to a .py file on your pc
- try to make code run (e.g. by installing the necessary packages*)

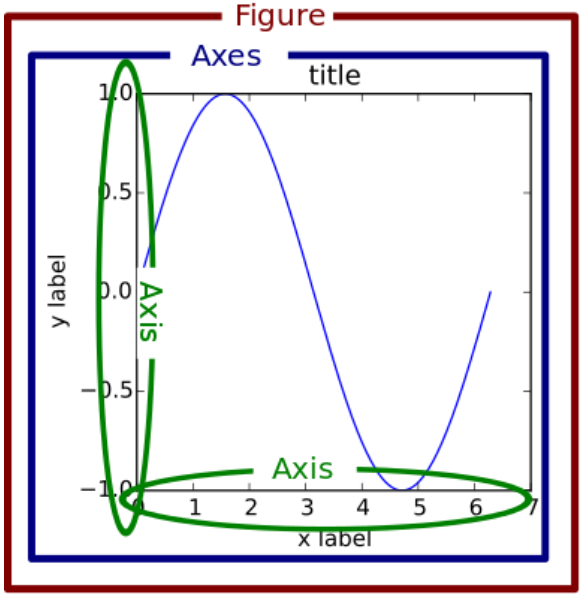
* for bokeh please install version 2.2.2 (we'll need that later)
pip install 'bokeh==2.2.2'

	Matplotlib (mpl)	Seaborn	Plotly	Bokeh	Altair	pandas
gallery	matplotlib.org/stable/gallery	seaborn.pydata.org/examples	plotly.com/python	docs.bokeh.org/en/latest/docs/gallery.html#standalone-examples	altair-viz.github.io/gallery/	pandas.pydata.org/pandas-docs/stable/user_guide/visualization.html
To actually plot sth, add lines:		from matplotlib import pyplot as plt plt.show()			<chart>.show()	from matplotlib import pyplot as plt plt.show()

UNIVERSITAT DE BARCELONA

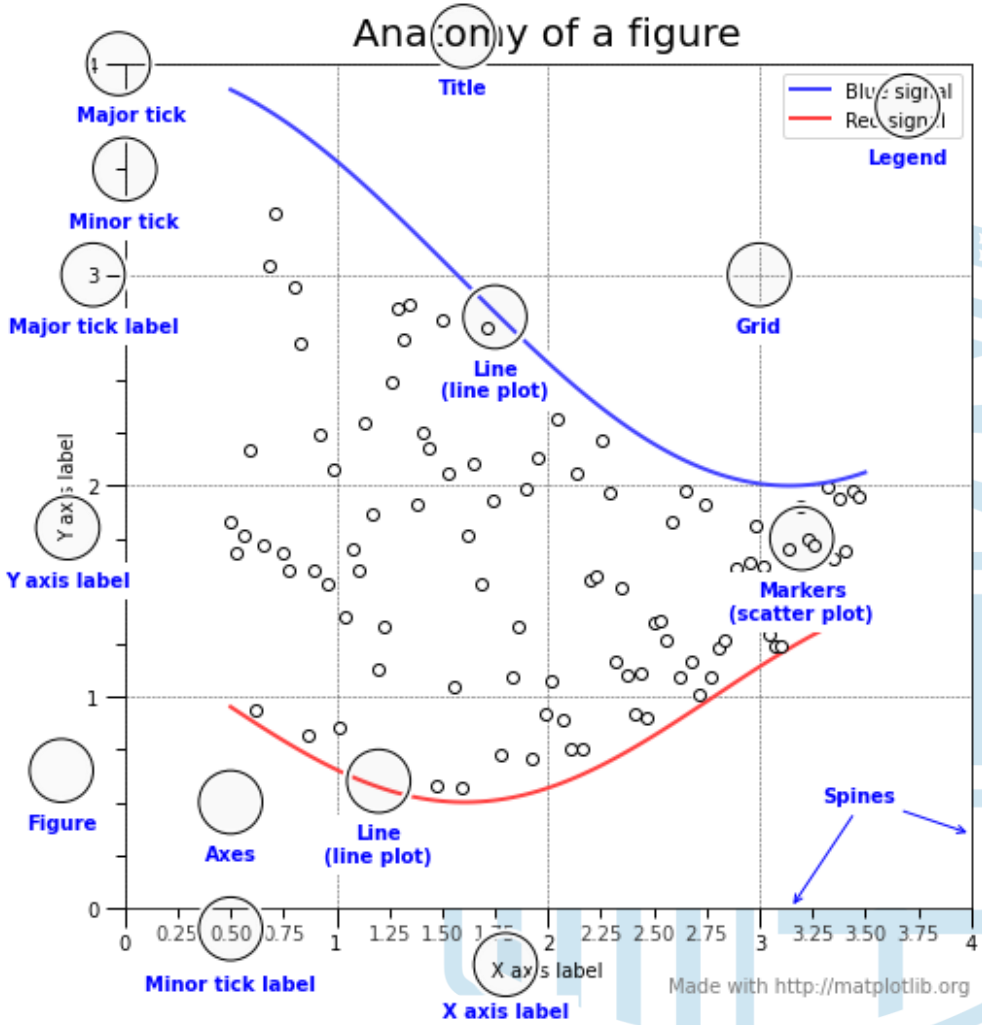
matplotlib

```
from matplotlib import pyplot as plt
```



Matplotlib Object Hierarchy

```
f, ax = plt.subplots()
```



[Anatomy of a figure]

UNIVERSITAT DE BARCELONA

We have a canvas, now we need sth we can plot, i.e: data!

get it from:

- Formulas (explicit or implicit)
- Numerical simulations
- Own measurements
- Open databases

--> jupyter notebooks on
https://github.com/Chaotique/Master_Visualizations_2021

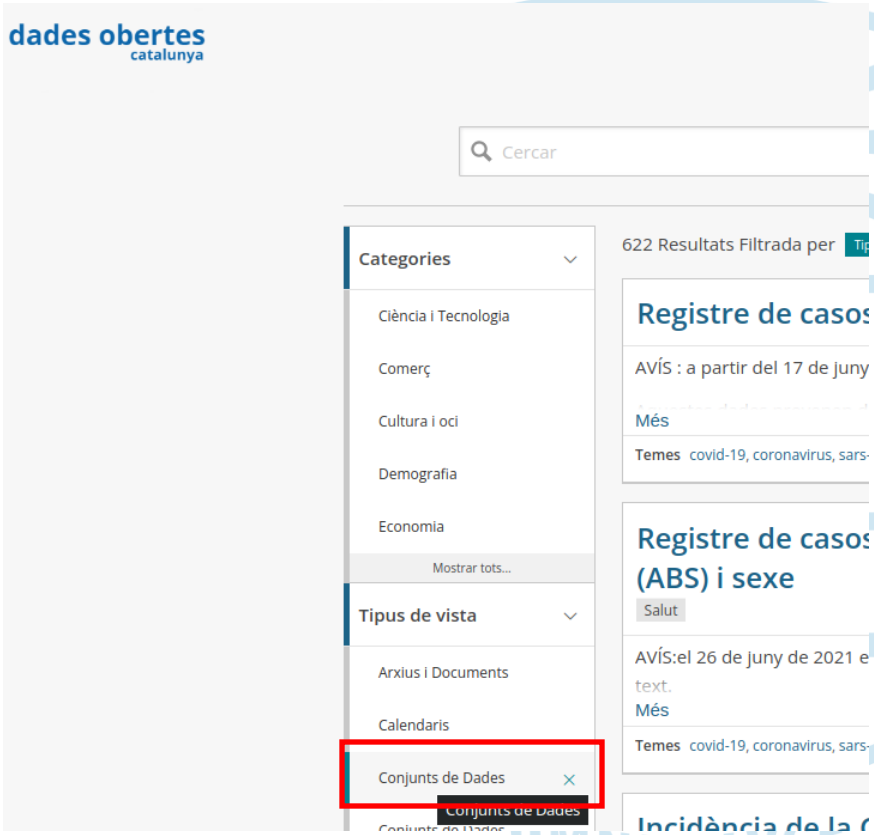
UNIVERSITAT DE BARCELONA

Homework until Thursday

Search on analisi.transparenciacatalunya.cat/browse?limitTo=datasets for datasets that might be interesting to visualize (on a dashboard)

- **What data do they "offer" exactly?**
  (e.g. columns of the table)
- **Does it contain both geographical and temporal data?**
  (would be nice)

On **Thursday** (session VI) we decide for one or two topics/ datasets together, for which we create a simple **dashboard** in sessions VI, XI, and XII.



UNIVERSITAT DE BARCELONA

Scientific Journals Requirements

PHYSICAL REVIEW E
covering statistical, nonlinear, biological, and soft matter physics
Highlights Recent Accepted Collections Authors

Information for Authors
- Manuscript Preparation
  - Content Guidelines
  - Technical Format and Style Guidelines
- Manuscript Submission and Resubmission
  - Submission

PUBLISH
PLOS ONE

Style and Format
Manuscript Organization
Submission
Information
at Submission
for Specific Study
eligible for APC

Submission Guidelines

Related information for authors
- PLOS Writing Center
- Submission system
- Journal scope and publication criteria
- Getting started guide
- Guidelines for revisions
- Publication fees
- APC Support

Style and Format

Science
Current Issue First release papers Archive
HOME > SCIENCE > *SCIENCE*: INFORMATION FOR AUTHORS

***Science*: Information for authors**

Science is a weekly, peer-reviewed journal that publishes significant original sci
reviews and analyses of current research and science policy. We seek to publish
influential in their fields or across fields and that will substantially advance sci

nature
Explore content About the journal Publish with us Subscribe
nature > for authors > formatting guide

For Authors
Editorial criteria and processes
Formatting guide
Presubmission enquiries
Initial submission

Formatting guide

This guide describes how to prepare contribu
this in full if you have not previously submitte
that, before submission, you familiarize your
the journal, either in print or online, particular

IEEE Author Center
Resources & Tools for Authors
HOME NEW AUTHORS JOURNAL AUTHORS

Magazine Authors
Get Started with IEEE Magazines
Create Your IEEE Magazine Article
The IEEE Magazines Publication Process
Article Submission Requirements
Authoring Tools and Templates
IEEE Article Templates

Search this section

Home » Create Your IEEE Magazine Article » Article Submission Requir

Article Submission Requirements

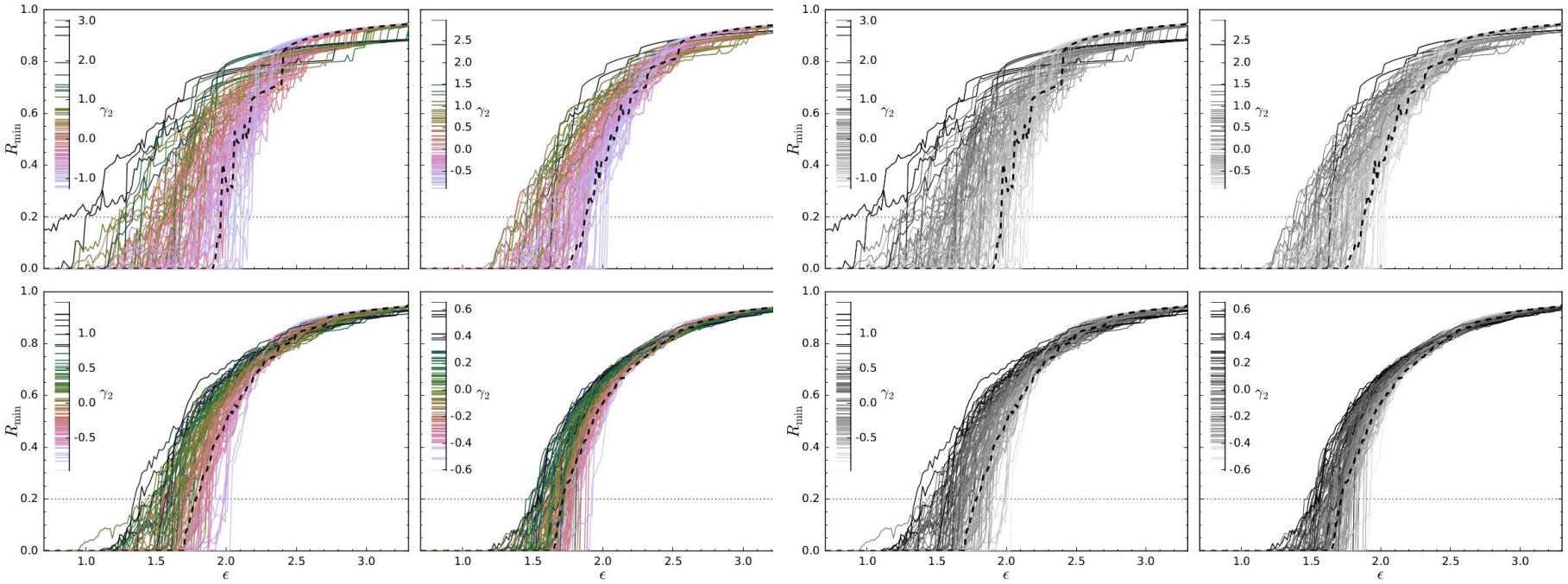
Speed up the publishing process by making sure all your files a
article submission requirements, including graphics and multi

Supply the following during article submission:
- Source file in TXT, DOC, DOCX, or TeX format
- Source files for any graphics or multimedia items, labeled se

UNIVERSITAT DE BARCELONA

Scientific Journals Requirements

- Plot should be 'readable' in printed version (cubehelix palettes)



FPs dissertation

UNIVERSITAT DE BARCELONA

Scientific Journals Requirements

- Plot should be 'readable' in printed version (cubehelix palettes)

OPINION

Obama's Divided Nation

ama presides over
ica more divided
any time in 50
ry that was riven
racial lines gath-
2008 to elect its
ident. That presi-
four years dividing
the basis of eco-
The campaign re-
vealed no evi-
dence that Mr.
Obama will close
the chasm he has
created between
his voters and
those he attacked
and vilified.

It may be true
that Mitt Romney
failed to respond

R.I.
Conn.
Del.
D.C.

Obama Romney

Source: AP

drawn attention to what hap-

Obama spokestotum replied:

problem with pols,
verbally facile as Mr.
that in crunch time,
reverts to No. 1. Exi
that 9% of the elect
who to vote for jus
Tuesday; and amo
42% said Mr. Oban
Sandy response—t
tie photo-op—wa
factor. Of those,
voted for Mr. Ob
Mr. Christie is o
politico who is
Yes, Republic
across two pres
that there are
how crudely
issue like ill
Blowing up th
if you thoug
day's results

In other words.

UNIVERSITAT DE BARCELONA

Scientific Journals Requirements

- Plot should be 'readable' in printed version (cubehelix palettes)
- Plot should fit into a column of the paper
- Limits to number of figures
- Concise end descriptive caption

UNIVERSITAT DE BARCELONA

Scientific Journals Requirements

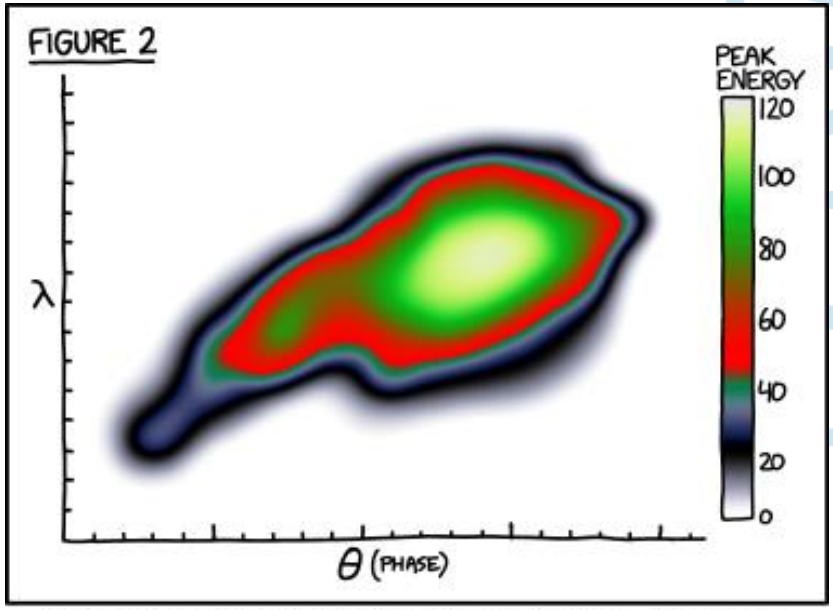
- Plot should be 'readable' in printed version (cubehelix palettes)
- Plot should fit into a column of the paper
- Limits to number of figures
- Concise end descriptive caption

Other Aspects to consider

- Think of colorblind people
- Figure fontsize = fonsize of article
- Limit axes to range of data (exception bar diagrams)
- consider using log, polar etc
- Clear not cluttered, consistent, self-explanatory, and not misleading
- Labels, specifications in the caption (reproducibility)
- As complex as necessary, as simple as possible

UNIVERSITAT DE BARCELONA

Appendix



EVERY YEAR, DISGRUNTLED SCIENTISTS COMPETE FOR THE PAINBOW AWARD FOR WORST COLOR SCALE.

https://xkcd.com/2537/

Stateful vs stateless approach

stateful

```
import matplotlib.pyplot as plt

plt.figure()
plt.plot([0, 1, 2, 3])

# get current figure and current axes
plt.gcf().gca().set_xlabel("tide")
```

stateless (OOP)

```
import matplotlib.pyplot as plt

fig = plt.figure()
ax  = fig.add_subplot(111)
li  = ax.plot([0,1], [1,0])
```