

# R/Python

## 需要理由嘛？

*DAMA 的自私传销;-)*

Zoom.Quiet



# #免责声明



一切资料来自网络互动挖掘  
一切想法来自日常学习工作  
一切体悟来自各种沟通交流  
一切知识来自社区分享印证  
一切经验来自个人失败体验



#PPT去死去死



幻灯

Presentation



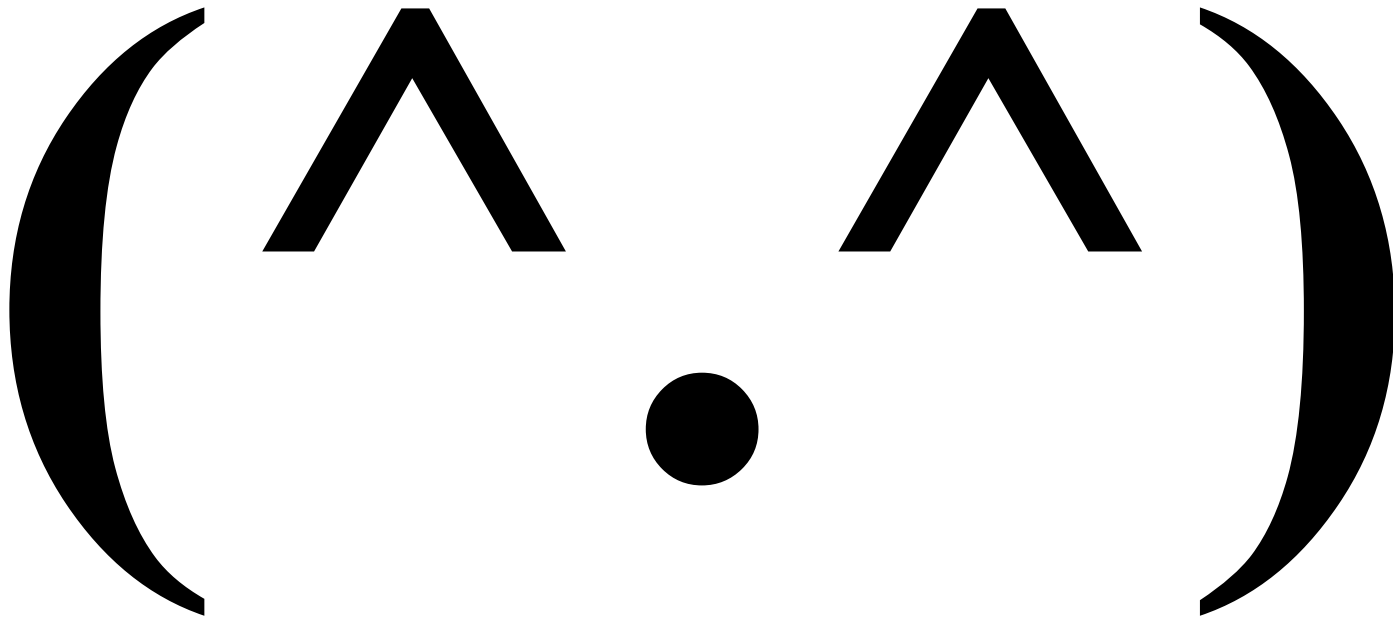
# #授权声明



录音



#是也乎



#风格提醒



高橋流

*Takahashi*



#高橋流



大



# #高橋流





# #高橋流



# 快



#高橋流



很 快



#高橋流



非常快



#所以



听



#DAMA



大媽



# #Dama dama 黠鹿(tiānlù)



#DAMA



15+



#DAMA 自证:发现 gmail...



2000





#DAMA: 2003~



**EZUG**



#DAMA: 2004~



# 啄木鸟

Python技术社区



#DAMA: 2005~

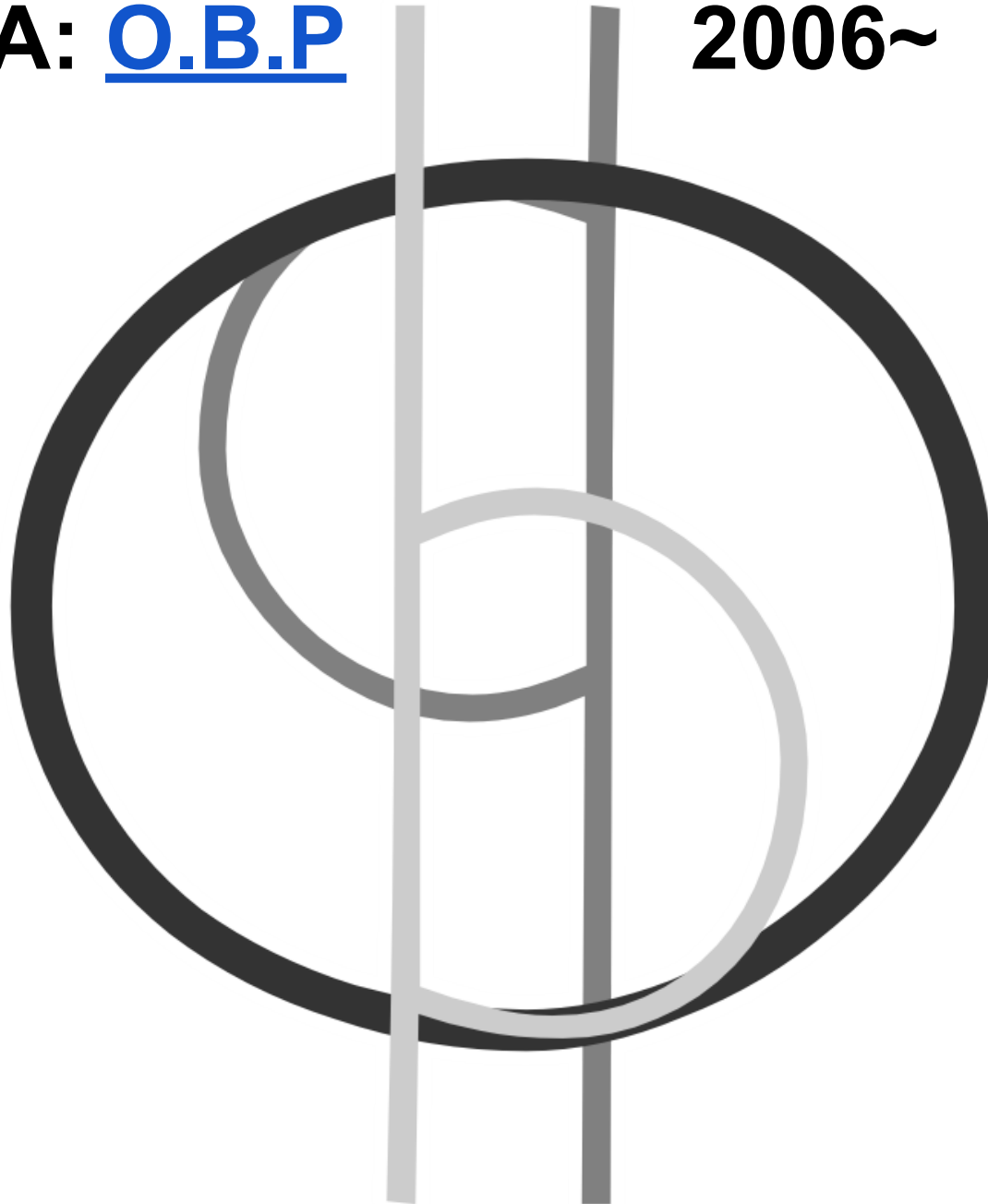


CPyUG



#DAMA: O.B.P

2006~



#DAMA: 2006~ Zeuux



哲思



#DAMA: 2007~



**E****E****U****G**



#DAMA: 2008~



# 教育大发现

*sociallearnlab*



#DAMA: 2010~



# 开源卫士

*code.ijinshan*





#DAMA: 2011~



# OpenResty

Nginx with Lua



# #DAMA: 2013~



gdg livin zhuhai life ;-)



Google

Developers

珠海谷歌开发者社区



# #感谢...



什么是....



R



R是....



# 游乐场



**R vs ...**



**SAS**

**SPSS**



R vs ..



Julia



什么是....



Py





# Py vs R



DataCamp  
Learn data analysis for free,  
*interactively*

# DATA SCIENCE WARS



VS.

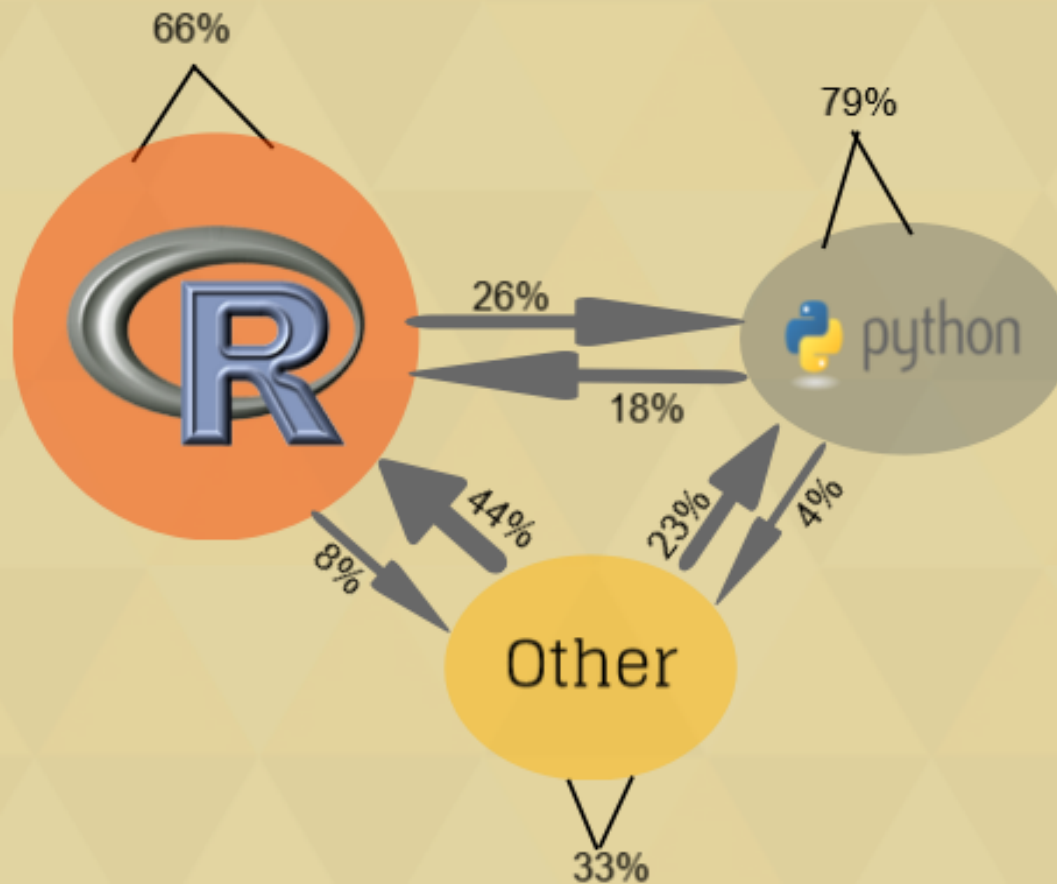


# Py vs



## Switching Between R and Python?

Number of people switching between R and Python in 2013 \*



\*Percentages on the arrows are relative to the base



Py+R=



#但是...



IP [y]:



IP[y]: 的伙伴们

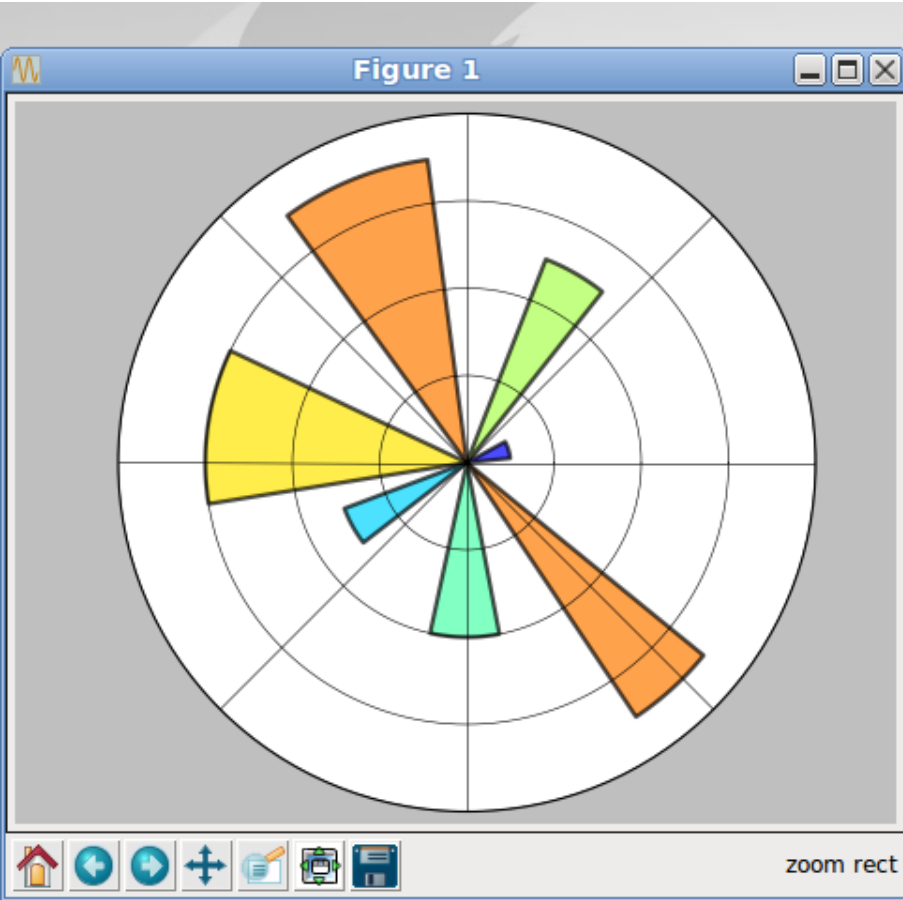


# Pandas



# numpy





```

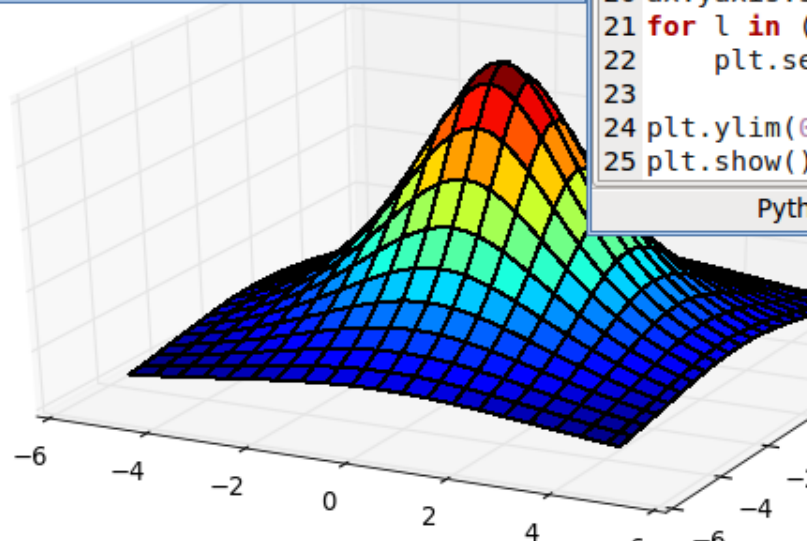
matplotlib_logo.py (~) - gedit
Datei Bearbeiten Ansicht Suchen Werkzeuge Dokumente Hilfe

Öffnen Speichern Rückgängig

matplotlib_logo.py x
1 import matplotlib as mpl
2 import matplotlib.pyplot as plt
3 from math import *
4
5 t, w, r = zip((0.1, 0.4, 1), (0.9, 0.3, 5),
6 (1.7, 0.5, 7), (2.7, 0.6, 6), (3.5, 0.3, 3),
7 (4.5, 0.4, 4), (5.3, 0.3, 7))
8
9 ax = plt.subplot(111, polar=True)
10 bars = ax.bar(t, r, width=w, bottom=0.0, lw=2)
11
12 for r, bar in zip(r, bars):
13     bar.set_facecolor(plt.cm.jet(r / 9.))
14     bar.set_alpha(0.7)
15
16 tic = mpl.ticker
17 no_labels = mpl.ticker.NullFormatter()
18 ax.yaxis.set_major_locator(tic.MultipleLocator(2))
19 ax.xaxis.set_major_formatter(no_labels)
20 ax.yaxis.set_major_formatter(no_labels)
21 for l in ('xgridlines', 'ygridlines'):
22     plt.setp(plt.getp(ax, l), 'linestyle', '-')
23
24 plt.ylim(0, 8)
25 plt.show()

Python Tabulatorbreite: 4 Z. 11, Sp. 1 EINF

```



```

Terminal
Datei Bearbeiten Ansicht Suchen Terminal Hilfe

user@pc ~ $ python matplotlib_logo.py
user@pc ~ $

```

# pandas pivot\_table explained



pandas 透视表...

	Account	Name	Rep	Manager	Product	Quantity	Price	Status
0	714466	Trantow-Barrows	Craig Booker	Debra Henley	CPU	1	30000	presented
1	714466	Trantow-Barrows	Craig Booker	Debra Henley	Software	1	10000	presented
2	714466	Trantow-Barrows	Craig Booker	Debra Henley	Maintenance	2	5000	pending
3	737550	Fritsch, Russel and Anderson	Craig Booker	Debra Henley	CPU	1	35000	declined
4	146832	Kiehn-Spinka	Daniel Hilton	Debra Henley	CPU	2	65000	won

```
pd.pivot_table(df,  
index=["Manager", "Status"],  
columns=["Product"],  
aggfunc=[np.sum],  
values=["Price"],  
fill_value=0,  
margins=True,  
dropna=True)
```

Can also use a dictionary:  
aggfunc={"Quantity":len,  
"Price":[np.sum,np.mean]}

		sum					
		Price					
		Product	CPU	Maintenance	Monitor	Software	All
Manager	Status						
Debra Henley	declined	70000	0	0	0	0	70000
	pending	40000	10000	0	0	0	50000
	presented	30000	0	0	0	20000	50000
	won	65000	0	0	0	0	65000
Fred Anderson	declined	65000	0	0	0	0	65000
	pending	0	5000	0	0	0	5000
	presented	30000	0	5000	10000	0	45000
	won	165000	7000	0	0	0	172000
All		465000	22000	5000	30000	0	522000



IP[y]: + ...



**+graphviz**

**+D3.js**





IP[y]: 3.\*



IP[y]: 的角色



# 强化Py外壳 Jupyter内核之一 并发Python



Jupyter 的使命



REPL 协议支持

Notebook

JupyterHub

... 语言无关的!



Jupyter 已能



# Julia

# Ruby ...



# 150507 正式联姻



DropBox, GitLab, Google Drive....



# 150507 正式联姻



c. [US] <https://github.com/OpenBookProjects/ipynb/blob/master/XKCD-style/XKCD-demos.ipynb>

## XKCD 样图表

### 为什么!?

因为喜欢哪!

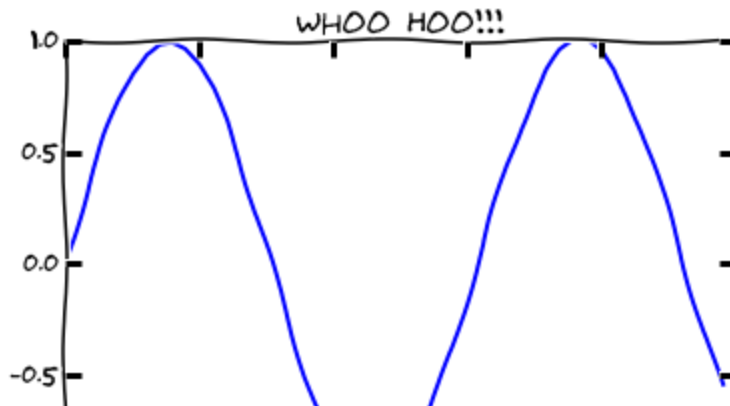
```
In [56]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [67]: %pylab inline
```

Populating the interactive namespace from numpy and matplotlib

```
In [68]: plt.xkcd() # Yes...
plt.plot(sin(linspace(0, 10)))
plt.title('Whoo Hoo!!!')
```

```
Out[68]: <matplotlib.text.Text at 0x11181cd90>
```





# #总之

- + R 和 Py 都是好的
- + IP[y]: 值得折腾
- + Jupyter 必须关注
- + Pandas... 很好玩



IP[y]: + Jupyter



样例





#是也乎



# Q&A

[zoomquiet+dama@gmail.com](mailto:zoomquiet+dama@gmail.com)



*Life's pathetic, go Pythonic !*



# Happy Hacking

