

数据可视化基础

讲师：萨缪尔
Samuel



萨缪尔老师

网易数据分析教研负责人

前盛大游戏战略规划总监、前腾讯游戏商业智能中心Leader

上海交通大学高金硕士、《哈佛管理导师》外部导师

- 知乎大V：「萨缪尔」主要聚集与商业分析、行业与战略分析、Python数据分析等
- BAT互联网巨头商业洞察分析 & 咨询公司战略咨询背景
- 擅长行业趋势研究和战略管理咨询工具，为20多家上市公司提供战略发展决策建议

课程

亮点

1

Python可视化库比较

2

常见图表类型及语法

3

数据可视化的探索过程

4

案例

Python可视化库

1

Python可视化库比较及选择

	Matplotlib	Seaborn	Bokeh	Plotly	Pyecharts
基本图表	●	●	●	●	●
地图			●	●	●
3D	●		○	●	●
多子图	●		●	●	●
动画	○		○	●	●
交互控件			●	●	

常见图表类型及语法



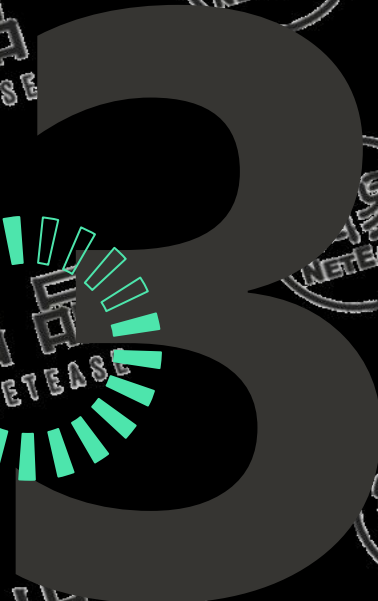
不同图形语法的代码示例

图形语法	散点图	统计直方图	箱形图
matplotlib	<pre>plt.scatter(df['SOD'], df['tau'], c='black', s=15, marker='o')</pre>	<pre>plt.hist(df['SOD'], 30, density= False, facecolor='w',edgecolor="k")</pre>	<pre>labels=np.unique(df['Class']) all_data = [df[df['Class']==label]['SOD'] for label in labels] plt.boxplot(all_data, widths=0.6, notch= False,labels=labels)</pre>
Seaborn	<pre>sns.relplot(x="SOD", y="tau", data=df,color='k')</pre>	<pre>sns.distplot(df['SOD'], kde=False, bins=30, hist_kws=dict(edgecolor="k", facecolor="w",linewidth=1,alpha=1))</pre>	<pre>sns.boxplot(x="Class", y="SOD", data= df, width =0.6,palette=['w'])</pre>
plotnine	<pre>(ggplot(df, aes(x='SOD',y='tau')) + geom_point())</pre>	<pre>(ggplot(df, aes(x='SOD')) + geom_ histogram(bins=30,colour="black",fil l="white"))</pre>	<pre>(ggplot(df, aes(x='Class',y='SOD'))+ geom_boxplot(show_legend=False))</pre>

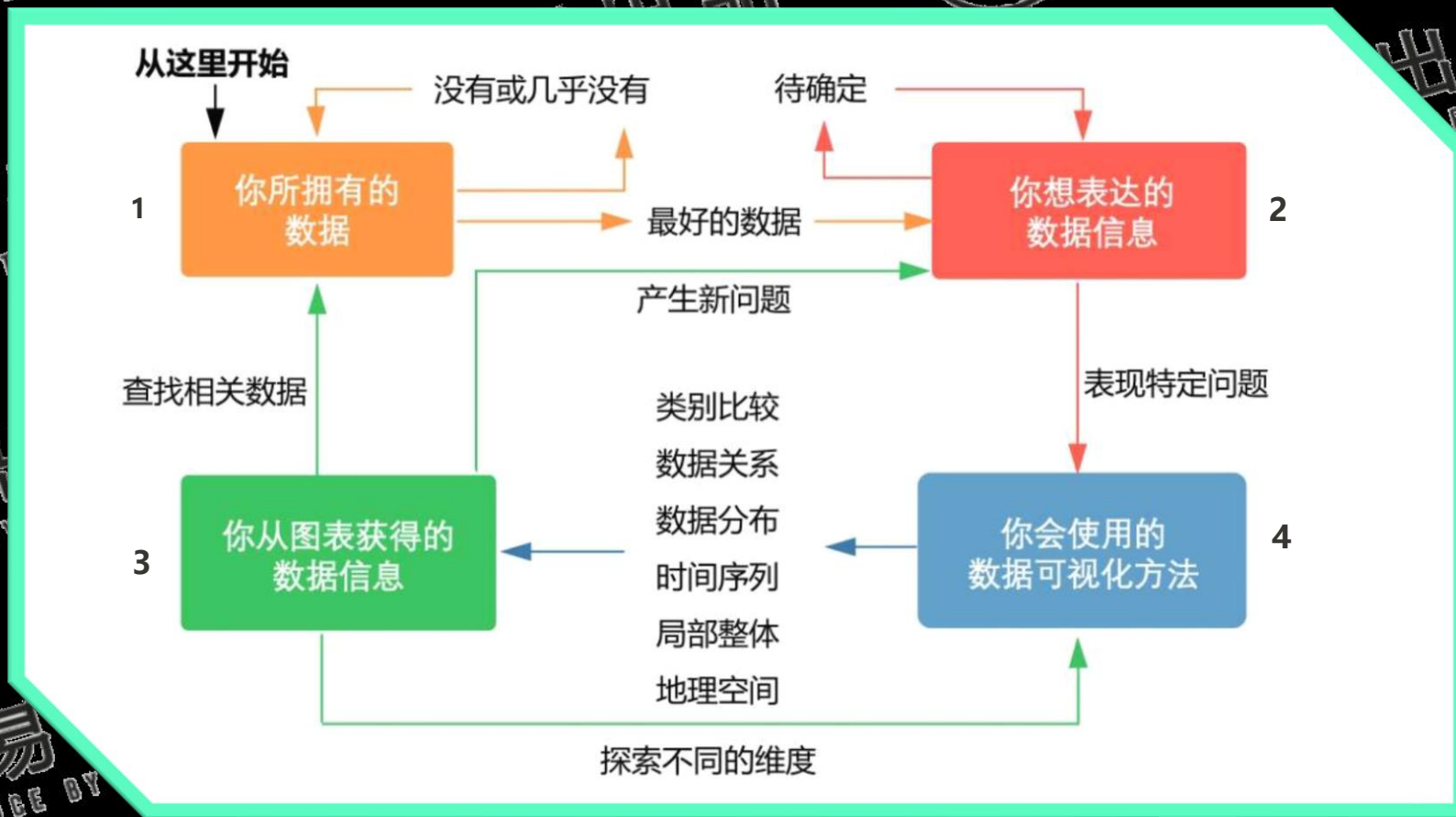
Matplotlib图表主要元素调整的函数

ID	函数	核心参数说明	功能
1	figure()	figsize (图表尺寸)、dpi (分辨率)	设置图表的大小与分辨率
2	title()	str (图名)、fontdict (文本格式, 包括字体大小、类型等)	设置标题
3	xlabel()、ylabel()	xlabel (X 轴名) 或 ylabel (Y 轴名)	设置 X 轴和 Y 轴的标题
4	axis()、xlim()、ylim()	xmin、xmax 或 ymin、ymax	设置 X 轴和 Y 轴的范围
5	xticks()、yticks()	ticks (刻度数值)、labels (刻度名称)、fontdict	设置 X 轴和 Y 轴刻度
6	grid()	b (有无网格线)、which (主/次网格线)、axis (X 轴和 Y 轴网格线)、color、linestyle、linewidth、alpha (透明度)	设置 X 轴和 Y 轴的主要和次要网格线
7	legend()	loc (位置)、edgecolor、facecolor、fontsize	控制图例显示

数据可视化的探索过程



数据可视化的探索过程



案例



注意：设置中文字体，防止作图时出现乱码

```
plt.rcParams['font.sans-serif'] = ['SimHei']
```

```
# Matplotlib中设置字体-黑体，解决Matplotlib中文乱码问题
```

```
plt.rcParams['axes.unicode_minus'] = False
```

```
# 解决Matplotlib坐标轴负号 '-' 显示为方块的问题
```

```
sns.set(font='SimHei')
```

```
# Seaborn中设置字体-黑体，解决Seaborn中文乱码问题
```


案例：简单的Matplotlib作图



课程

总结

1

Python常用的可视化库比较：

Matplotlib、Seaborn、Bokeh、Plotly、Pyecharts

2

常见的图表类型以及使用Python的实现：

散点图、统计直方图、箱形图

3

Matplotlib图表美化的主要函数：

figure、title、xlabel、ylabel、grid。。。

4

图表设置中文字体的方法