

# 数据关系型图表

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

数据关系型图表的种类



2

散点图



3

气泡图



# 数据关系型图表种类



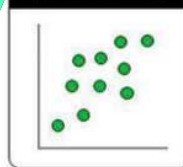
# 数据关系型图表

数值关系  
型

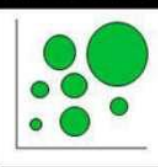
层次关系  
型

网络关系  
型

散点图  
(scatter chart)



气泡图  
(bubble chart)



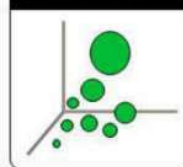
散点曲线图  
(connected scatterplot)



曲线图  
(line chart)



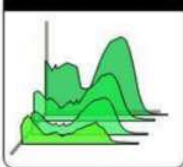
三维气泡图  
(3D bubble chart)



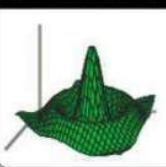
等高线图  
(contour chart)



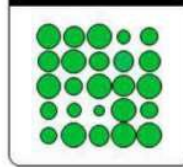
瀑布图  
(waterfall chart)



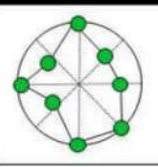
曲面图  
(surface plot)



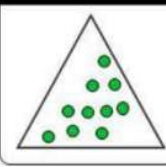
相关系数图  
(correlation matrix)



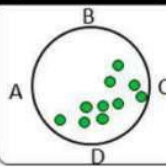
雷达图  
(radar chart)



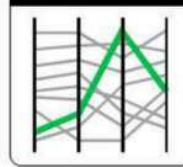
三元相图  
(ternary chart)



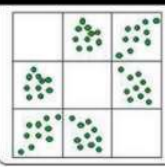
径向坐标图  
(radviz chart)



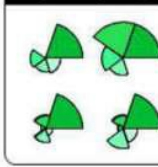
平行坐标系  
(parallel coordination)



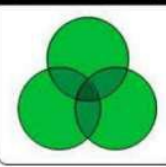
矩阵散点图  
(scatter matrix)



星形图  
(star plot)



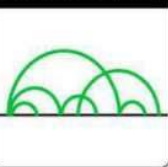
韦恩图  
(venn diagram)



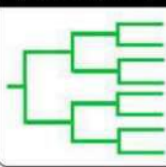
和弦图  
(chord diagram)



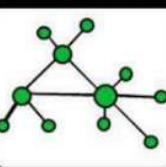
弧长链接图  
(arc-connection chart)



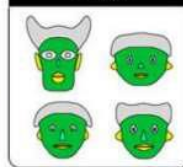
树形图  
(dendrogram)



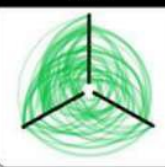
网络图  
(network chart)



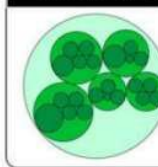
切尔夫夫脸谱图  
Isotype



蜂巢图  
Hive



圆填充图  
Circle Packing



# 散点图





# 散点图

## #1、导入可视化库，并设置中文字体

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

## #2、数据导入，3作图数据

```
df=pd.read_csv('/Users/samuelzhan/Downloads/1960-2020年全球各国GDP（清洗版）.csv')
df.set_index('序列')
df=df.loc[:,['序列','国家/地区','所在洲','Year','GDP(十亿美元)']]
df['Year']=pd.to_datetime(df['Year'])
df1=df[((df['国家/地区']=='美国')|(df['国家/地区']=='中国')|(df['国家/地区']=='日本'))&(df['Year']>'1999-01-01')]
```

## #4、选择可视化的图表类型并实现

```
plt.scatter(df1["Year"],df1['GDP(十亿美元)'],s=15,marker='o')
```

## #5、图例注释等装饰

```
plt.title('2000-2020年中美日GDP增长趋势（单位：十亿美元）')
plt.xlabel('年份')
plt.ylabel('GDP')
plt.figure()
plt.show()
```

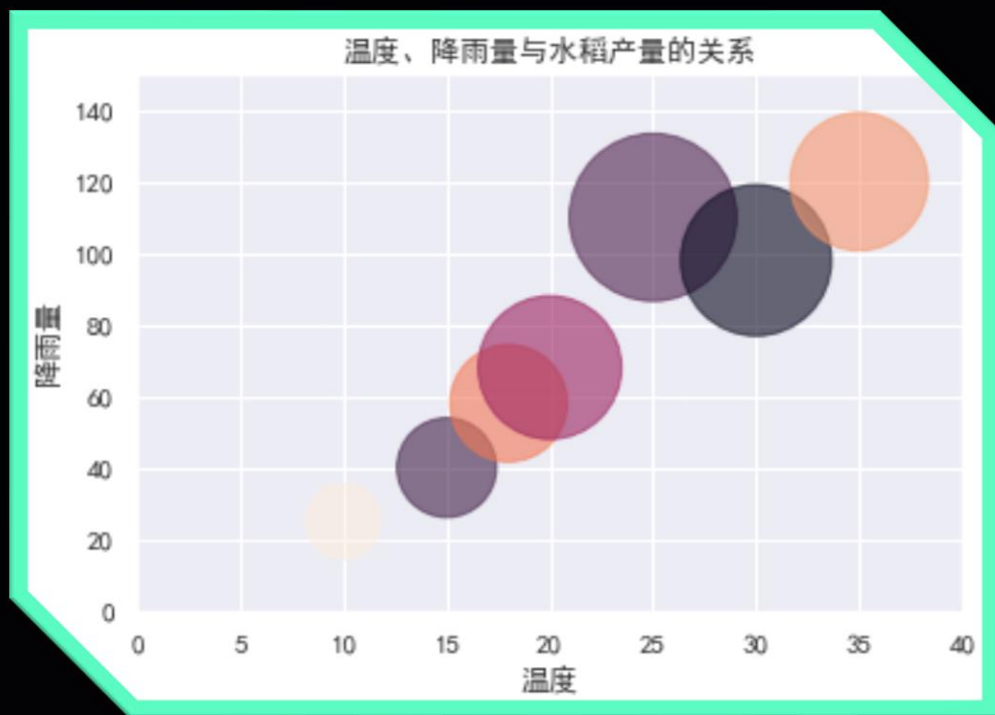


# 气泡图





# 气泡图



## #1、导入可视化库，并设置中文字体

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import numpy as np
```

```
plt.rcParams['font.sans-serif'] = ['SimHei']
plt.rcParams['axes.unicode_minus'] = False
sns.set(font='SimHei')
```

## #2、数据导入，3作图数据

```
production = [925, 1625, 2250, 3375, 4607, 3750, 3125]
tem = [10, 15, 18, 20, 25, 30, 35]
rain = [25, 40, 58, 68, 110, 98, 120]
```

## #4、选择可视化的图表类型并实现

```
colors = np.random.rand(len(tem)) # 颜色数组
size = production
plt.scatter(tem, rain, s=size, c=colors, alpha=0.6) # 画散点图, alpha=0.6 表示不透明度为 0.6
```

## #5、图例注释等装饰

```
plt.ylim([0, 150]) # 纵坐标轴范围
plt.xlim([0, 40]) # 横坐标轴范围
plt.title('温度、降雨量与水稻产量的关系')
plt.xlabel('温度') # 横坐标轴标题
plt.ylabel('降雨量') # 纵坐标轴标题
plt.show()
```

# 课程总结



1

数据关系型图表的种类：

数值关系型、层次关系型、网络关系型

2

散点图的制作

3

气泡图的制作