



绘制散点图

主讲：孙 睿





绘制散点图

- 散点图 (scatter diagram) 又称为散点分布图，由若干数据点组成。主要用于判断两变量之间是否存在某种关联，或者总结数据点的分布模式。
- 每个数据点有位置特征：横坐标和纵坐标。利用坐标点（散点）的分布形态反映特征间的统计关系。





绘制散点图

>>> 基础语法

```
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, data=None )
```



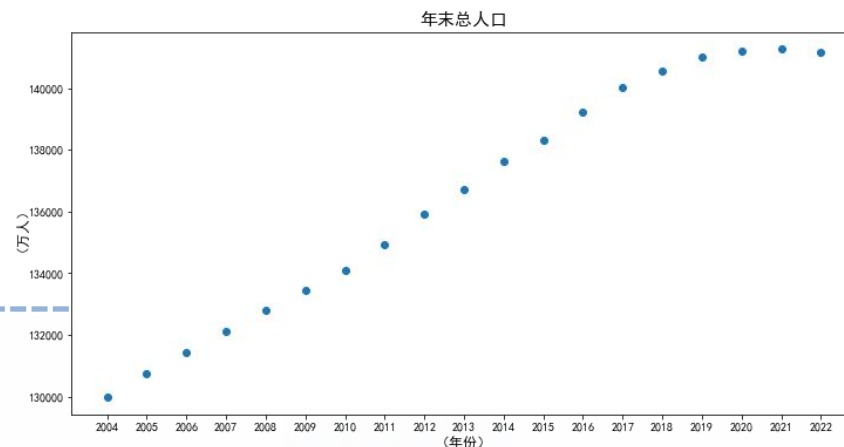
参数名	说明
x, y	表示数据点的位置 (x, y)。必备参数。x, y都不能为空。x, y支持多维数组，但是x, y的维度必须一致，多维数组会被展平 (flatten) 为一维数组。
s	表示数据点标记的大小。可选参数。默认值为rcParams['lines.markersize'] ** 2。
c	表示数据点的颜色。 可选参数。默认是蓝色 'b'，表示的是标记的颜色，或者可以是一个表示颜色的字符，或者是一个长度为n的表示颜色的序列等等如 'b' =blue, 'y' =yellow, 'k' =black等。
marker	表示数据点标记的样式，默认为圆形 'o'：rcParams["scatter.marker"] ('o')
alpha	表示标记的透明度。取值范围为[0, 1]，0为透明，1为不透明。默认值为None。
linewidths	表示数据点的描边（边缘线）宽度。浮点数或类数组结构。默认值为rcParams["lines.linewidth"] (1.5)。
edgecolors	表示数据点的描边（边缘线）颜色。取值为{'face', 'none', None}、颜色、颜色序列。默认值为rcParams["scatter.edgecolors"] ('face')。'face'：标记边缘颜色与标记颜色相同'none'：没有边线。
vmin, vmax	表示亮度的最小值和最大值。当norm存在的时候忽略。



绘制散点图

>>> 绘制散点图展示近20年人口增长趋势

```
df=pd.read_excel('人口数据.xlsx','总人口',index_col=0)
fig=plt.figure(figsize=(12,6))
plt.scatter(df.index,df['年末总人口(万人)'])
plt.xticks(df.index)
plt.xlabel(' (年份) ',fontsize=12)
plt.ylabel(' (万人) ',fontsize=12)
plt.title('年末总人口',fontsize=15)
plt.show()
```

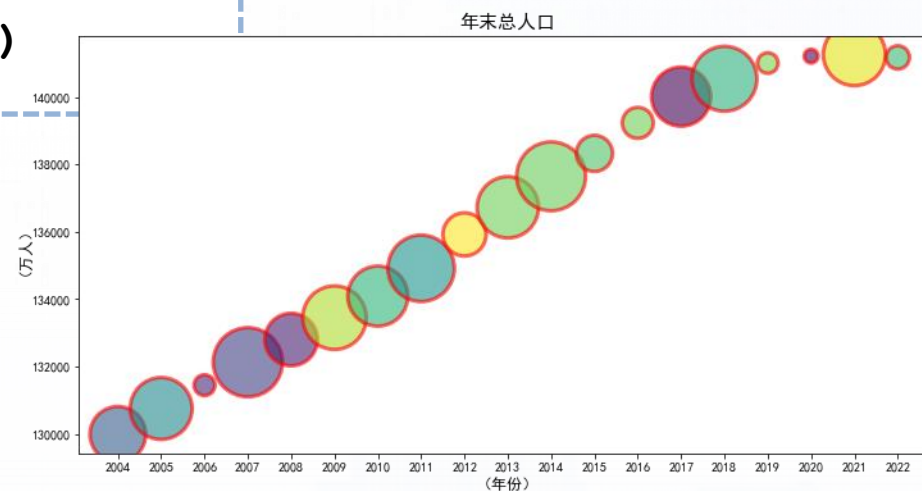




绘制散点图

>>> 绘制气泡图展示近年来人口增长趋势

```
plt.scatter(df.index, df['年末总人口(万人)'],  
            s=np.random.randint(50, 3000, 19),  
            c=np.random.rand(19), alpha=0.6,  
            linewidths=3, edgecolors='r')
```



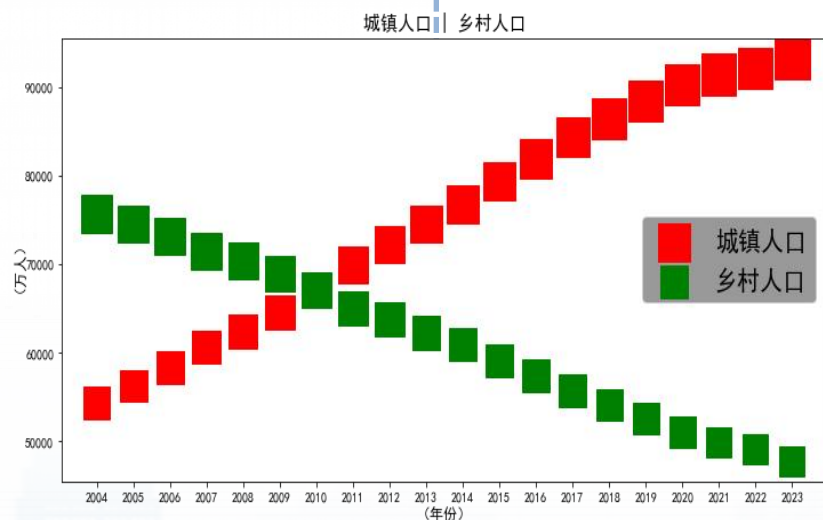


绘制散点图

>>> 绘制散点图比较“城镇人口”和“乡村人口”数据变化

```
plt.scatter(df.index, df['城镇人口(万人)'], color='r',  
            marker='s', s=df['城镇人口(万人)']/100 )
```

```
plt.scatter(df.index, df['乡村人口(万人)'], color='g',  
            marker='s', s=df['乡村人口(万人)']/100 )
```

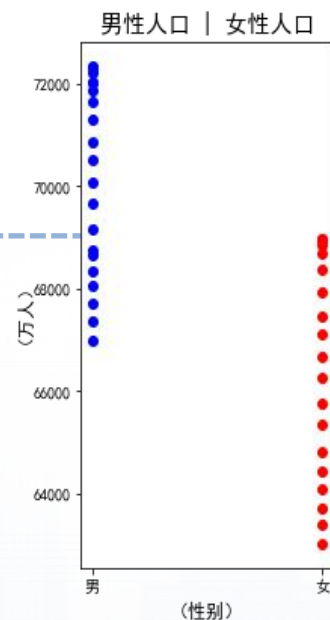




绘制散点图

>>> 绘制散点图比较“男性人口”和“女性人口”数据

```
df=pd.read_excel('人口数据.xlsx','总人口',index_col=0)
fig=plt.figure(figsize=(3,6))
plt.scatter(np.array(['男'] * 20),df['男性人口(万人)'],c='blue')
plt.scatter(np.array(['女'] * 20),df['女性人口(万人)'],c='red')
plt.xlabel(' (性别) ',fontsize=12)
plt.ylabel(' (万人) ',fontsize=12)
plt.title('男性人口 | 女性人口',fontsize=15)
plt.show()
```





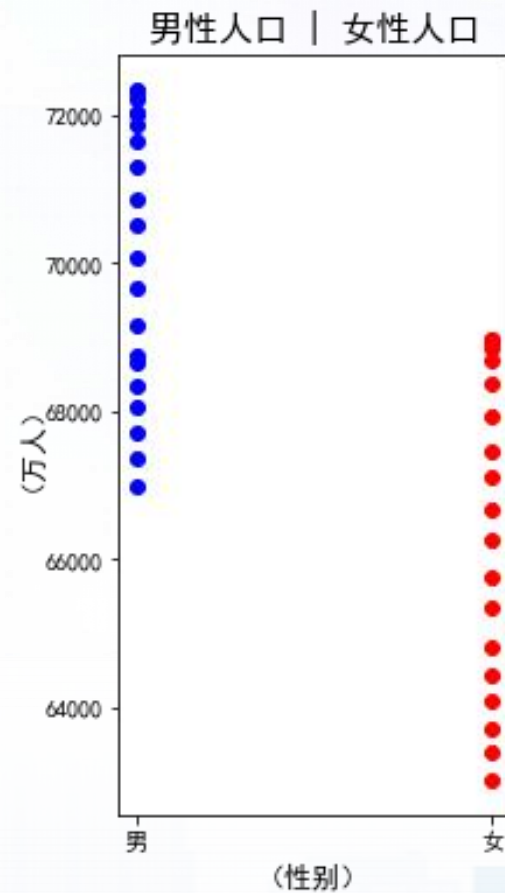
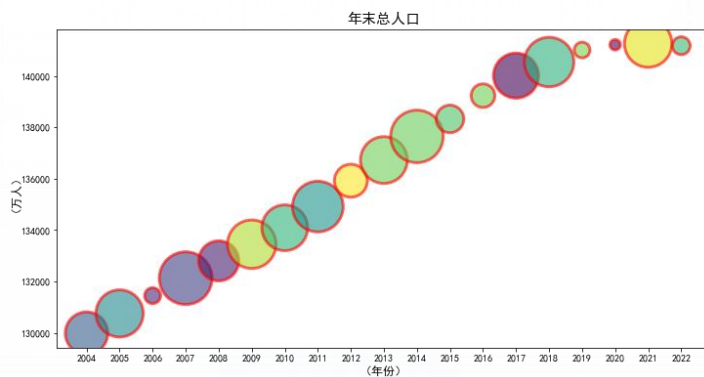
绘制散点图

>>> 总结



matplotlib.pyplot.scatter ()

s	数据点标记的大小
c	数据点标记的颜色
marker	数据点标记的样式
alpha	数据点标记的透明度





【matplotlib绘图库】

绘制散点图

主讲：孙 睿

