

matplotlib 模块绘图

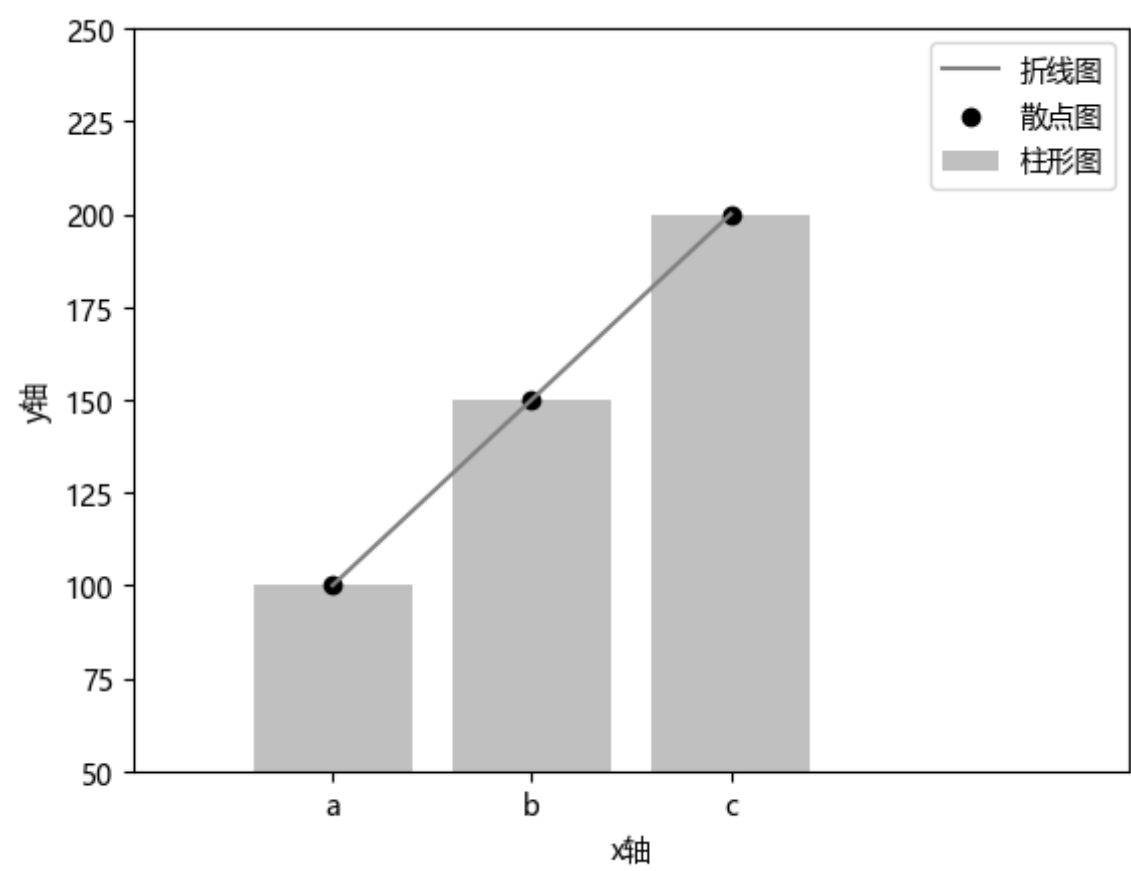
班级_____学号_____姓名_____

matplotlib 是使用最广泛的 python 第三方统计图表绘图库，利用它能够绘制各种各样的统计图表。考试和教材中主要考查的是 matplotlib 的子库 pyplotlib，其考查的函数如表 1 所示。

表 1 pyplot 子库常用的函数

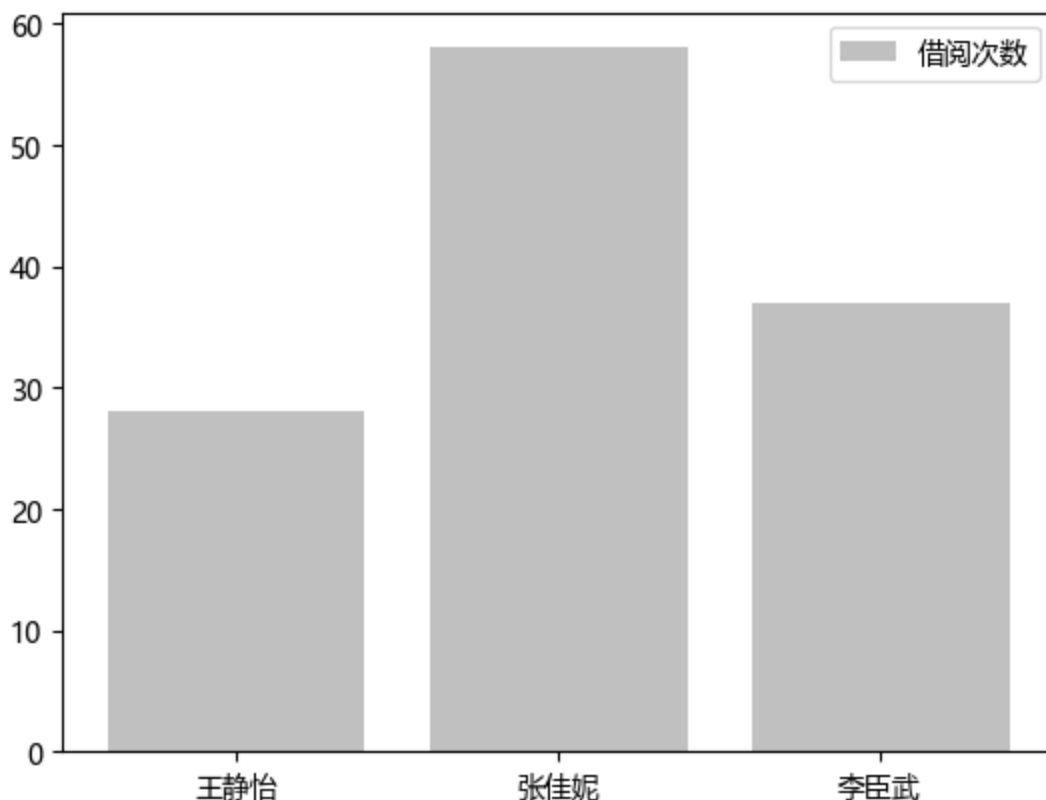
函数	说明
plot	绘制线型图
bar	柱形图（垂直柱型）
barh	条形图（水平柱型）
scatter	散点图
title	设置图表标题
xlim()、ylim()	设置 X 轴、Y 轴的取值范围
xlabel()、ylabel()	设置 x 轴标签、y 轴标签
legend()	显示图例
show()	显示创建的绘图图表

例 1 柱形图、折线图、散点图极简示例



```
import matplotlib.pyplot as plt #导入matplotlib的pyplot子库
x = ['a', 'b', 'c'] # 绘图点的x轴坐标
y = [100, 150, 200] # 绘图点的y轴坐标
ser = pd.Series([100, 150, 200], index=['a', 'b', 'c'])
# bar、plot、scatter的第1个参数是绘图点的x轴坐标，第2个参数是绘图点的y轴
# 坐标。label是显示的图例。color是绘制图形的颜色。
plt.bar(x, y, label="柱形图", color='silver')
plt.plot(x, y, label="折线图", color='gray')
plt.scatter(ser.index, ser.values, label="散点图", color="black")
plt.legend() # 显示图例
plt.ylim(50, 250) # 设置y轴坐标范围
plt.xlim(-1, 4) # 显示x轴坐标范围
plt.xlabel("x轴") # 显示x轴标签
plt.ylabel("y轴") # 显示y轴标签
plt.show() # 显示图表
```

例2 学生图书借阅次数柱形图



```
import pandas as pd
import matplotlib.pyplot as plt
data = {"姓名": ["王静怡", "张佳妮", "李臣武"],
```

```
"性别": ["女", "女", "男"],
```

```
"借阅次数": [28, 58, 37]}
```

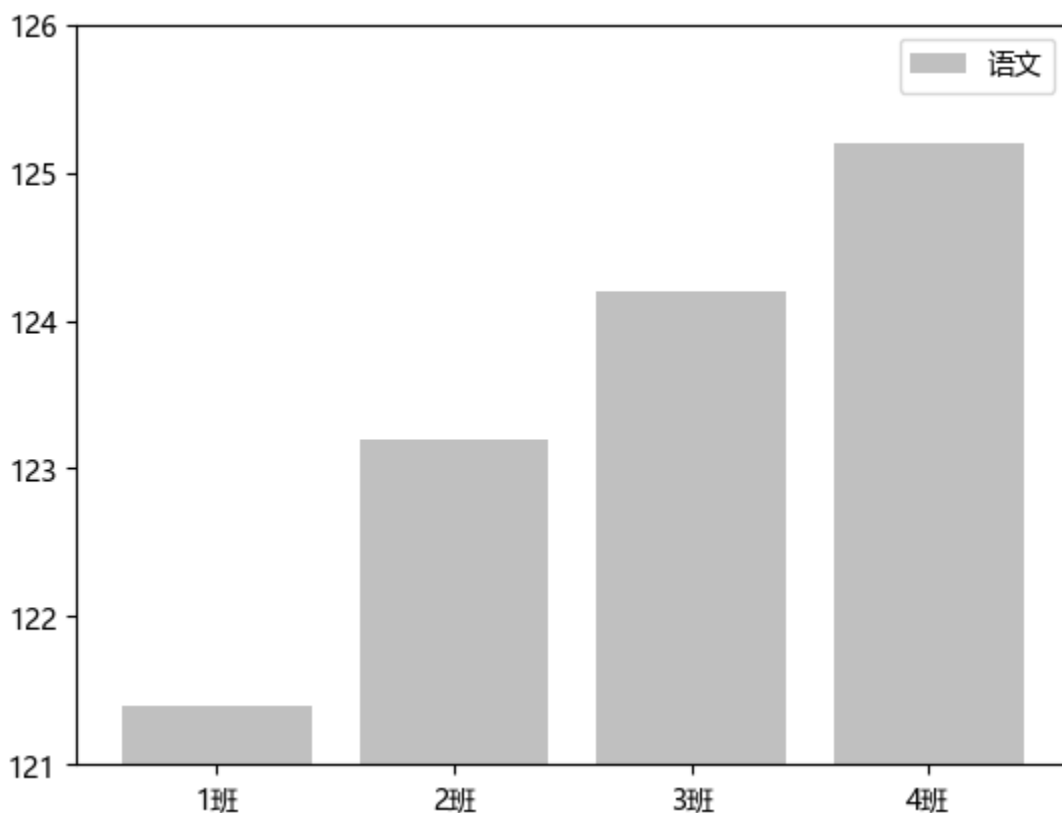
```
df = pd.DataFrame(data, columns=["姓名", "性别", "借阅次数"])
```

```
plt.bar(df["姓名"], df["借阅次数"], label="借阅次数", color='silver')
```

```
plt.legend()
```

```
plt.show()
```

例 3 考试各班语文平均分对比柱形图



```
df = pd.read_excel("登分表.xlsx")
```

```
g = df.groupby("班级", as_index=False)
```

```
df2 = g.mean()
```

```
plt.bar(df2["班级"], df2["语文"], label="语文", color='silver')
```

```
plt.ylim(121, 126)
```

```
plt.legend()
```

```
plt.show()
```

加框处代码，`as_index=False`，所以生成的 `df2` 保留了班级列，其它数值列进行了求平均数的操作。考试中有两种常见的等价代码：

代码 1

```
g = df.groupby("班级", as_index=True)
```

```
df2 = g.mean()
```

```
plt.bar(df2.index, df2["语文"], label="语文", color='silver')
```

代码 1 中, `as_index=True`, 生成的 `df2` 的班级存放在 `index` 中, 所有传入的 `x` 用 `df2.index`。

代码 2

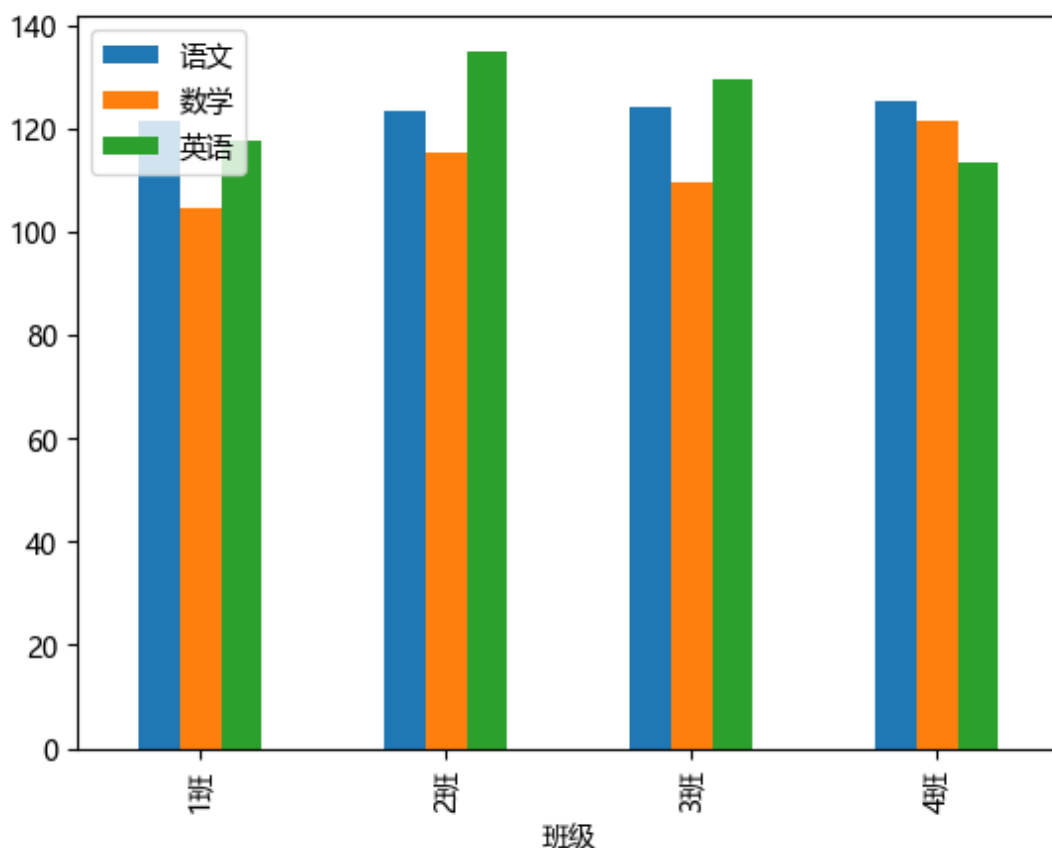
```
g = df.groupby("班级", as_index=True)
```

```
s = g["语文"].mean()
```

```
plt.bar(s.index, s.values, label="语文", color='silver')
```

代码 2 中, `as_index=True`, 只取了分组语文列求平均值, 求得的 `s` 是一个 `series`, `s` 的 `index` 是班级的 `values` 是语文的平均分。

例 4 使用 `DataFrame` 上的 `plot` 函数进行绘图



```
df = pd.read_excel("登分表.xlsx")
```

```
g = df.groupby("班级", as_index=True)
```

```
df2 = g[["语文", "数学", "英语"]].mean()
```

```
df2.plot(kind="bar") # plot 函数作图的 x 是 index, y 是所有的列
```

```
# kind 参数的取值: bar-柱形图; barh-条形图; line-折线图; scatter-散点图。
```