

可视化进阶 seaborn

2024



01

seaborn简介

02

图表的常用设置

03

数据分组统计

04

数据位移

05

数据转换

06

数据合并

07

数据导出

08

日期数据处理

09

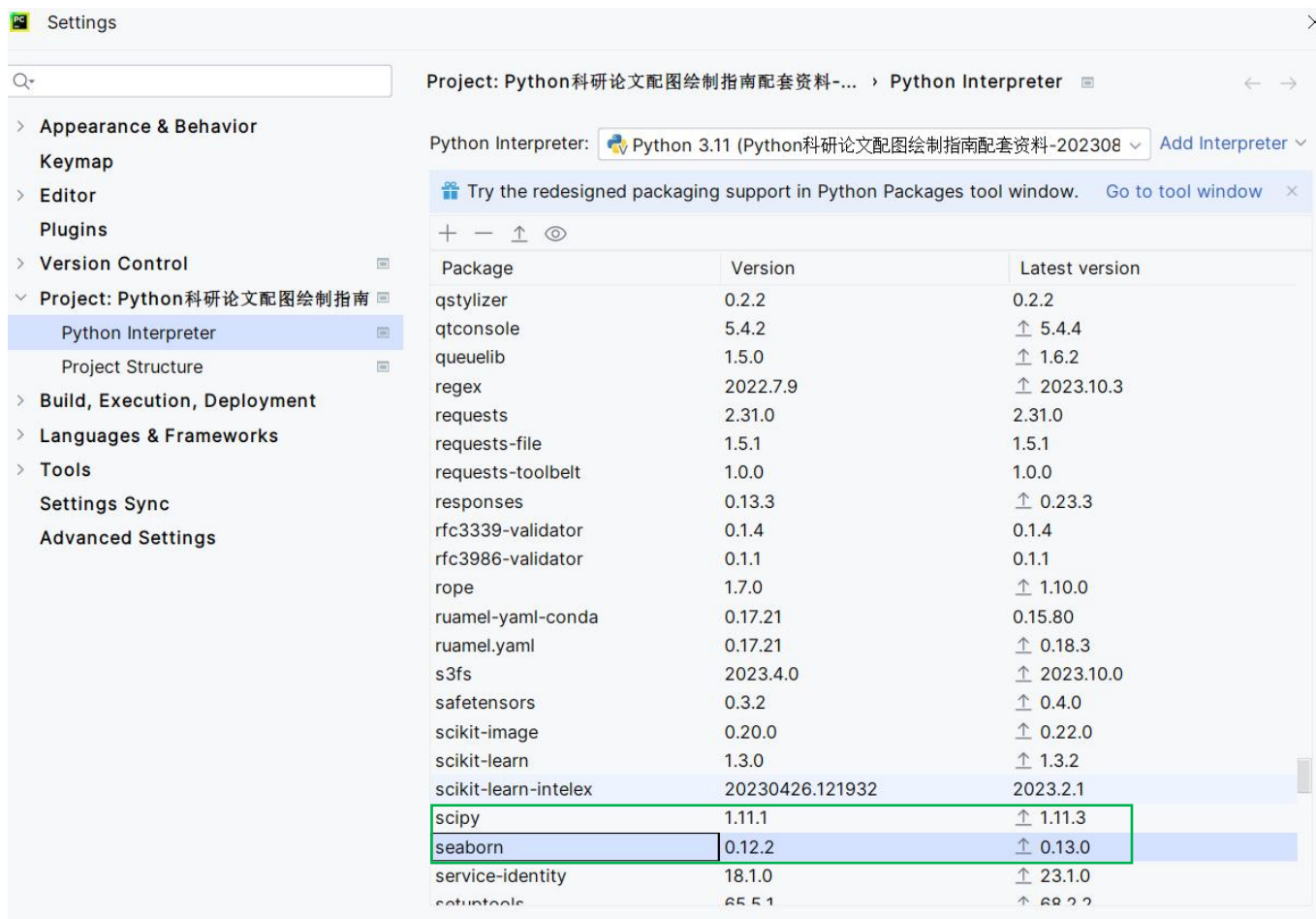
时间序列

01

seaborn简介



- seaborn是基于matplotlib的python可视化库。
 - 安装seaborn: `pip install seaborn`
 - seaborn 依赖于 scipy, 因此, 也需要安装scipy





● seaborn初接触

代码

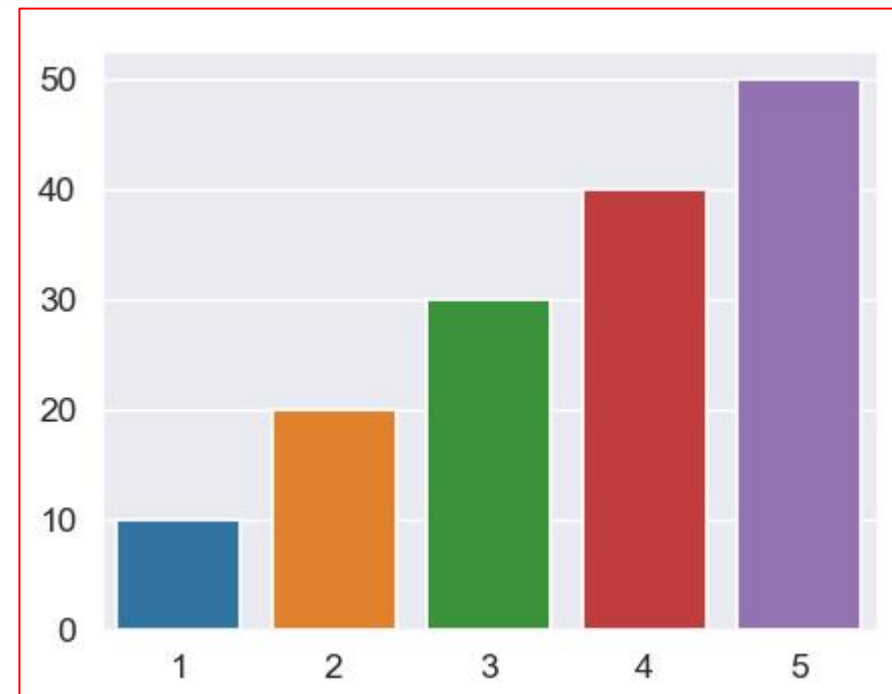
matplotlib库是seaborn的依赖库，基本都需要导入

{darkgrid, whitegrid, dark, white, ticks}

语法与matplotlib略有不同

```
import seaborn as sns
import matplotlib.pyplot as plt
plt.figure(figsize=(4, 3))
x=[1, 2, 3, 4, 5]
y=[10, 20, 30, 40, 50]
#plt.bar(x, y)
#plt.show()
sns.set_style('darkgrid')
sns.barplot(x=x, y=y)
plt.show()
```

输出

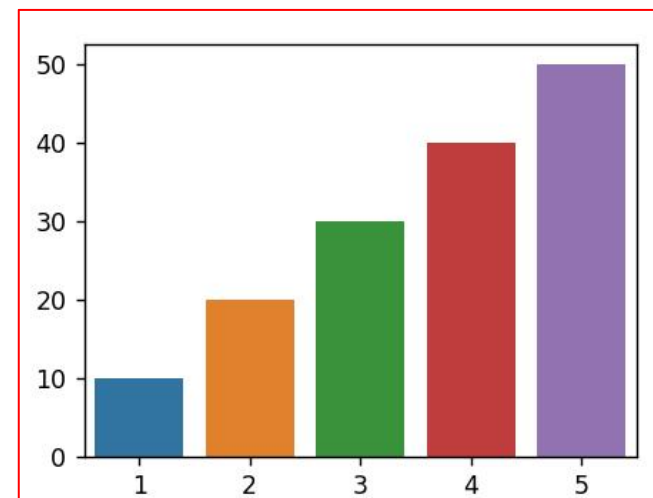
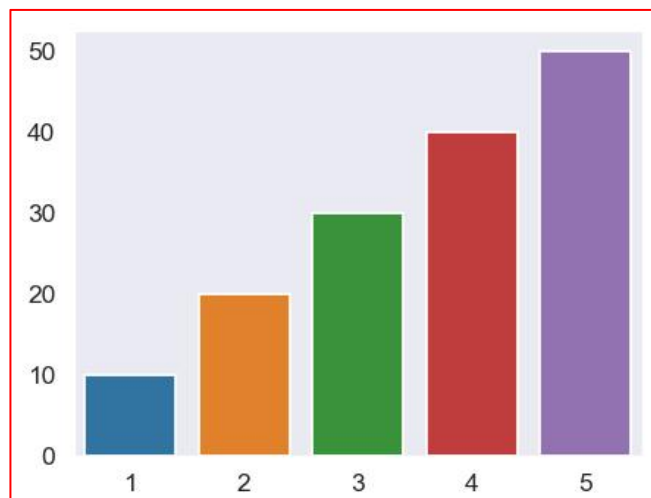
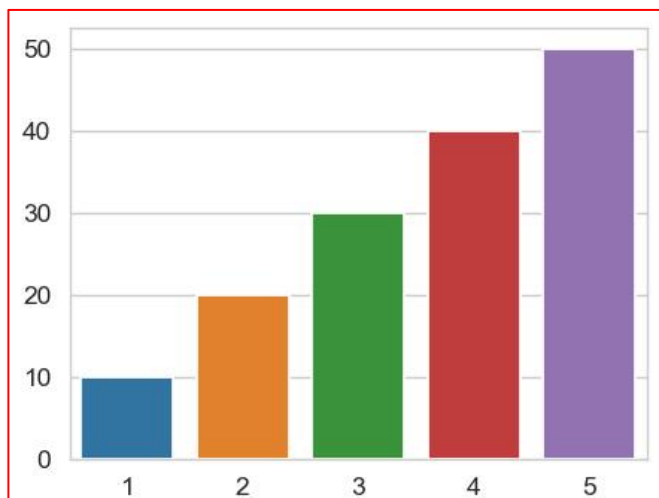
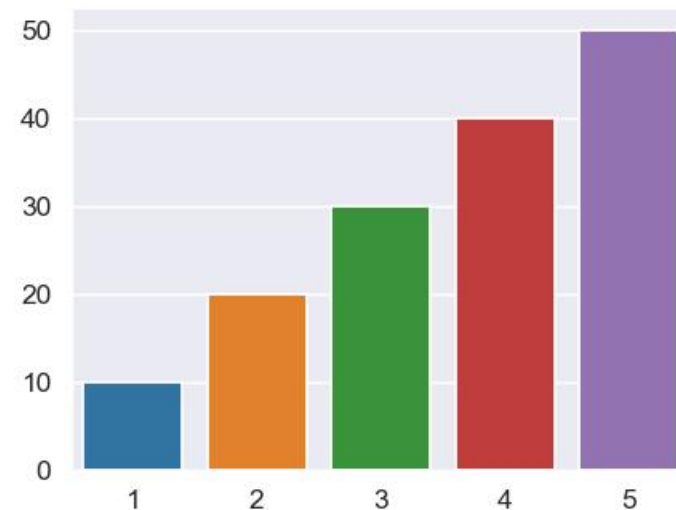


02

seaborn 绘图



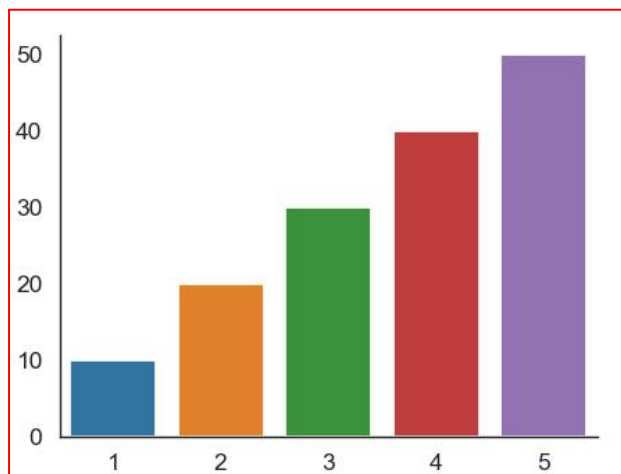
- seaborn图表基本设置
- 背景风格
 - `sns.axes_style` `sns.set_style`
 - `darkgrid,`
 - `whitegrid,`
 - `dark,`
 - `white,`
 - `ticks}`



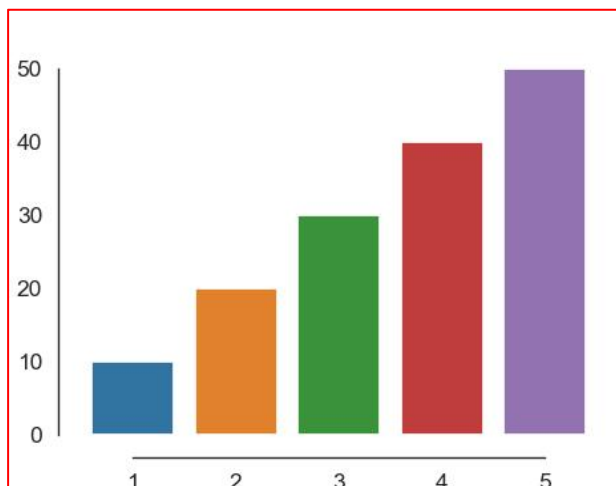


- 边框设置
 - `sns.despine(fig=None, ax=None, top= True, right=Ture,left=True,top=True,bottom=Flase,offset=None)`

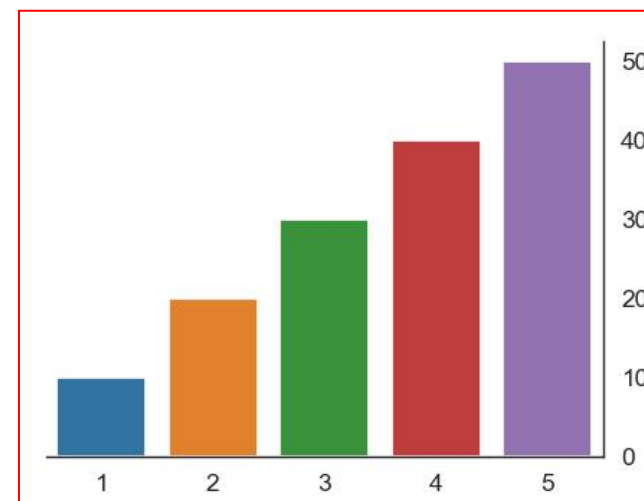
```
sns. despine()
```



```
sns. despine(offset=10, trim=True)
```



```
sns. despine(left=True, right=False)
```





2.seaborn绘图

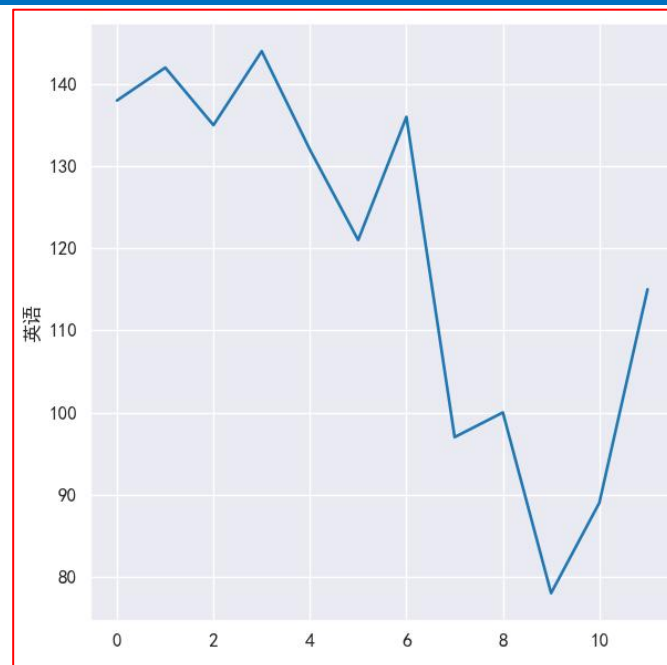
Python 数据分析: pandas

- 折线绘制
- `sns.relplot()`

代码

```
sns.set_style('darkgrid')
plt.rcParams['font.sans-serif']=['SimHei'] #解决中文乱码
df1=pd.read_excel('data.xls') #导入Excel文件
#绘制折线图
sns.relplot(x=range(0,12,1),y="英语",kind="line",data=df1)
plt.show()# 显示
```

输出

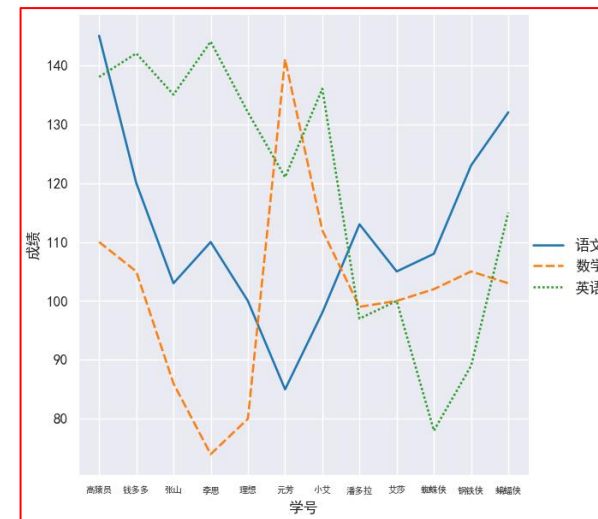
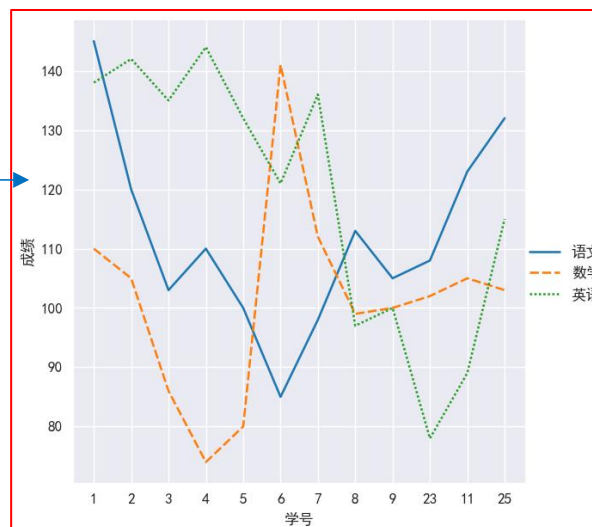


尝试

```
sns.relplot(data=df1.iloc[:,2:],kind='line')
```

观察excel表格, 把学号或者姓名作为x坐标 ticks

学号	姓名	语文	数学	英语
1	高猿员	145	110	138
2	钱多多	120	105	142
3	张山	103	86	135
4	李思	110	74	144
5	理想	100	80	132
6	元芳	85	141	121
7	小艾	98	112	136
8	潘多拉	113	99	97
9	艾莎	105	100	100
23	蜘蛛侠	108	102	78
11	钢铁侠	123	105	89
25	蝙蝠侠	132	103	115



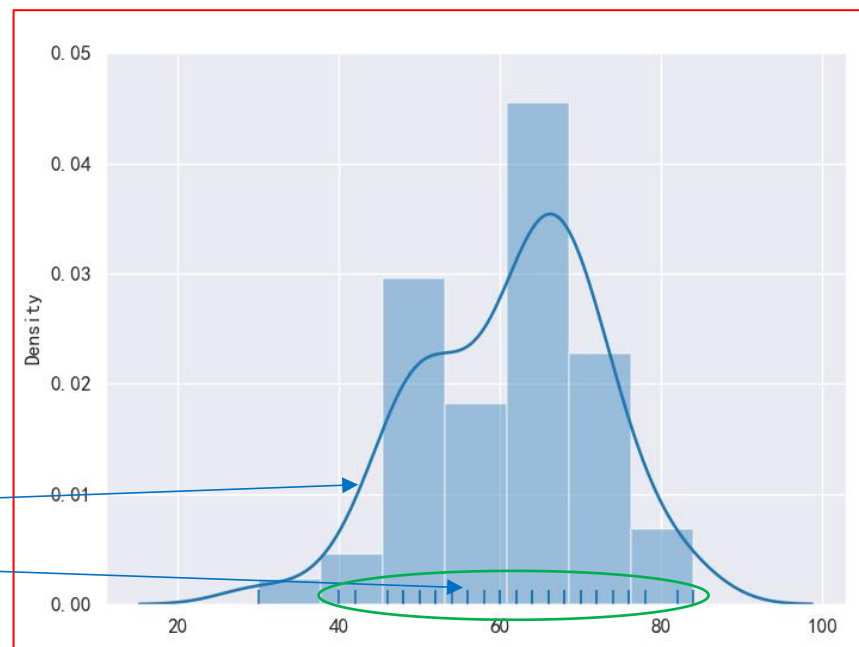


- 直方图绘制
 - `sns.distplot()`

代码

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_style('darkgrid')
plt.rcParams['font.sans-serif']=['SimHei'] #解决中文乱码
df1=pd.read_excel('data2.xls')
data=df1[['得分']]
#sns.distplot(data,bins = 3, hist = True, rug=True,kde= True)
sns.distplot(data,rug=True,kde=True) #直方图, 显示观测的小细条
plt.show() # 显示
```

输出



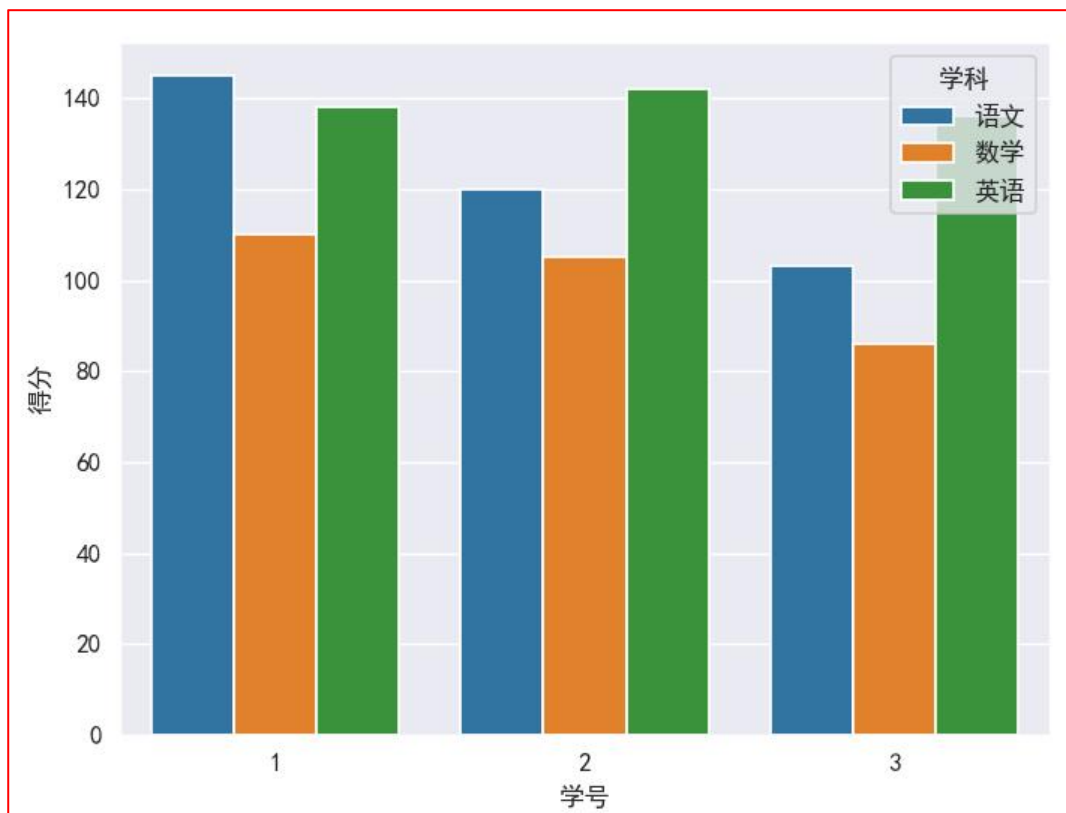


- 直方图绘制
- `sns.barplot()`

代码

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_style('darkgrid')
plt.rcParams['font.sans-serif']=['SimHei'] #解决中文乱码
df1=pd.read_excel('data.xls',sheet_name='sheet2')
sns.barplot(x='学号',y='得分',hue='学科',data=df1)
plt.show() # 显示
```

输出





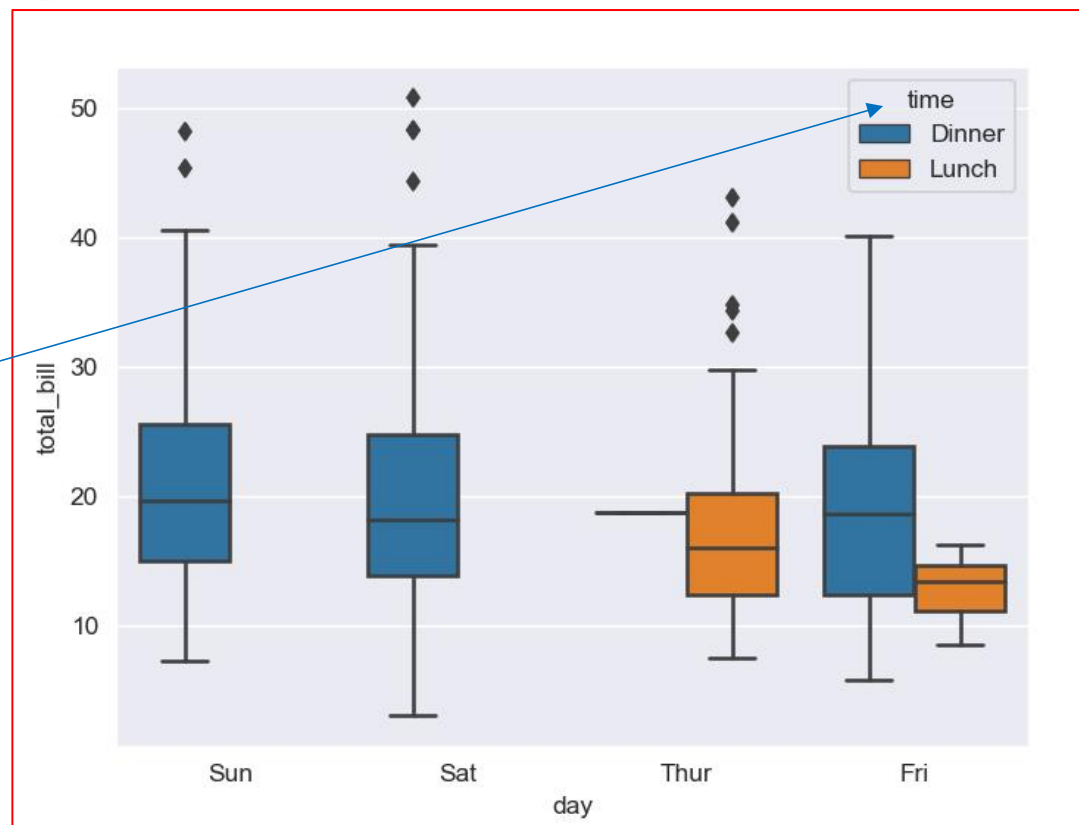
- 箱型绘制
- `sns. boxplot()`

代码

```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
sns.set_style('darkgrid')
#读取数据集tips
tips=pd.read_csv('tips.csv')
#绘制箱形图
sns.boxplot(x='day',y='total_bill',hue='time',data=tips)
plt.show() # 显示图表
```

hue代表分类字段

输出



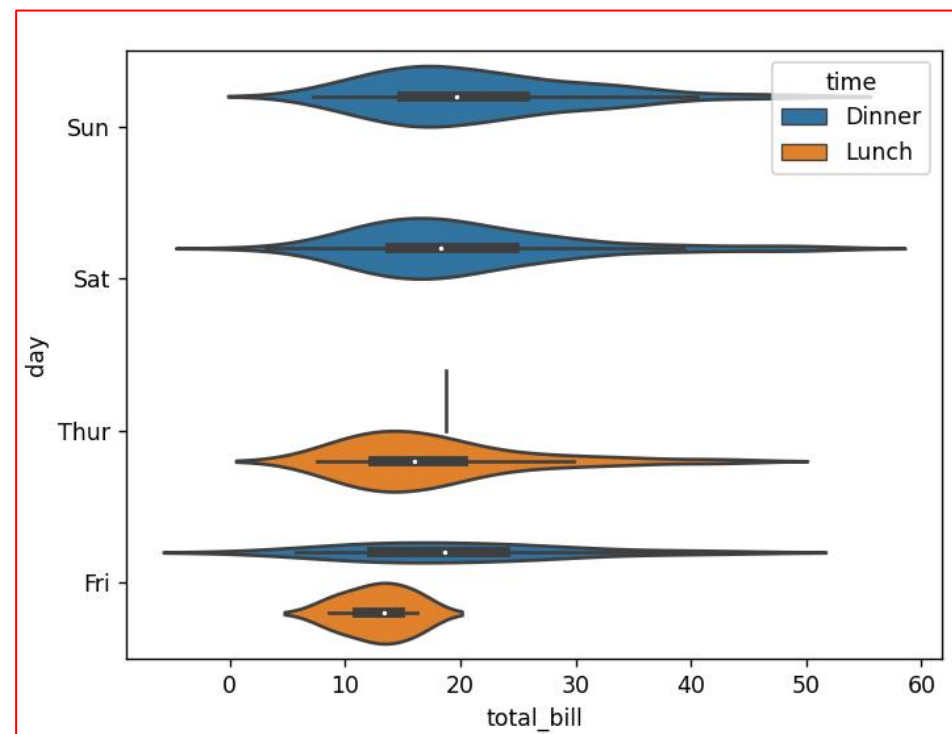


- 小提琴图绘制
- `sns.violinplot()`

代码

```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
#读取数据集tips
tips = pd.read_csv('tips.csv')
sns.violinplot(x='total_bill', y='day', hue='time', data=tips)
plt.show()
```

输出





- 线性回归模型绘制
- `sns.lmplot()`

代码

```
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
sns.set_style('darkgrid')
#读取数据集tips
tips=pd.read_csv('tips.csv')
#绘制回归模型, 描述线性关系
sns.lmplot(x='total_bill', y='tip', data=tips)
plt.show() # 显示
```

	A	B	C	D	E	
1	total_bill	tip	sex	smoker	day	time
2	16.99	1.01	Female	No	Sun	Dinner
3	10.34	1.66	Male	No	Sun	Dinner
4	21.01	3.5	Male	No	Sun	Dinner
5	23.68	3.31	Male	No	Sun	Dinner
6	24.59	3.61	Female	No	Sun	Dinner
7	25.29	4.71	Male	No	Sun	Dinner
8	8.77	2	Male	No	Sun	Dinner
9	26.88	3.12	Male	No	Sun	Dinner
10	15.04	1.96	Male	No	Sun	Dinner
11	14.78	3.23	Male	No	Sun	Dinner

输出

