

# 实训五 Matplotlib数据可视化基础

本实训针对一组关于全球星巴克门店的统计数据，分析了在不同国家和地区以及中国不同城市的星巴克门店的数量。

## 导入模块

```
In [ ]: import pandas as pd
import matplotlib.pyplot as plt
plt.rcParams['font.sans-serif']=['SimHei']
```

## 获取数据

```
In [ ]: data=pd.read_csv('directory.csv')
data.head()
```

```
Out[ ]:
```

	Brand	Store Number	Store Name	Ownership Type	Street Address	City	State/Province	Country	Postcod	
0	Starbucks	47370-257954	Meritxell, 96	Licensed	Av. Meritxell, 96	Andorra la Vella		7	AD	AD50
1	Starbucks	22331-212325	Ajman Drive Thru	Licensed	1 Street 69, Al Jarf	Ajman		AJ	AE	Na
2	Starbucks	47089-256771	Dana Mall	Licensed	Sheikh Khalifa Bin Zayed St.	Ajman		AJ	AE	Na
3	Starbucks	22126-218024	Twofour 54	Licensed	Al Salam Street	Abu Dhabi		AZ	AE	Na
4	Starbucks	17127-178586	Al Ain Tower	Licensed	Khalidiya Area, Abu Dhabi Island	Abu Dhabi		AZ	AE	Na

### (1) 查看星巴克旗下有哪些品牌。

如果我们只关心星巴克咖啡门店，  
则只需获取星巴克中Brand的数据集，并查看全世界一共有多少家星巴克门店。

```
In [ ]: print("星巴克旗下品牌",data["Brand"].unique())
print("星巴克旗下各品牌数量")
print(data["Brand"].value_counts())
```

```
星巴克旗下品牌 ['Starbucks' 'Teavana' 'Evolution Fresh' 'Coffee House Holdings']
星巴克旗下各品牌数量
Starbucks          25249
Teavana            348
Evolution Fresh    2
Coffee House Holdings 1
Name: Brand, dtype: int64
```

## (2) 查看全世界一共有多少国家和地区开设了星巴克门店

显示门店数量排名前10和后10的国家和地区。

```
In [ ]: print("共有",data["Country"].unique().shape[0],"个国家开设星巴克店")
        print("共有",data["State/Province"].unique().shape[0],"个州/省开设星巴克店")
        print("共有",data["City"].unique().shape[0],"个城市开设星巴克店")
```

```
共有 73 个国家开设星巴克店
共有 338 个州/省开设星巴克店
共有 5470 个城市开设星巴克店
```

```
In [ ]: print("门店数量前10的国家")
        print(data["Country"].value_counts().sort_values()[-10:])
        print("门店数量后10的国家")
        print(data["Country"].value_counts().sort_values()[ :10])
```

门店数量前10的国家

```
PH      298
TR      326
TW      394
MX      579
GB      901
KR      993
JP     1237
CA     1468
CN     2734
US    13608
```

Name: Country, dtype: int64

门店数量后10的国家

```
AD      1
LU      2
MC      2
AW      3
CW      3
SK      3
ZA      3
TT      3
AZ      4
BO      4
```

Name: Country, dtype: int64

```
In [ ]: print("门店数量前10的州/省")
        print(data["State/Province"].value_counts().sort_values()[-10:])
        print("门店数量后10的州/省")
        print(data["State/Province"].value_counts().sort_values()[ :10])
```

门店数量前10的州/省

AZ	536
IL	575
31	580
NY	645
FL	694
11	706
WA	757
ENG	787
TX	1042
CA	2821

Name: State/Province, dtype: int64

门店数量后10的州/省

SA	1
70	1
65	1
DJ	1
K	1
JUN	1
57	1
VAN	1
TU	1
YO	1

Name: State/Province, dtype: int64

```
In [ ]: print("门店数量前10的城市")
print(data["City"].value_counts().sort_values()[-10:])
print("门店数量后10的城市")
print(data["City"].value_counts().sort_values()[:10])
```

门店数量前10的城市

Seattle	156
Las Vegas	156
Mexico City	180
Chicago	180
Toronto	192
London	216
New York	232
北京市	234
Seoul	243
上海市	542

Name: City, dtype: int64

门店数量后10的城市

Midrand	1
Greenvale	1
Glenville	1
Monticello	1
Glenmont	1
Covington	1
Colville	1
Geneseo	1
DuPont	1
Fresh Meadows	1

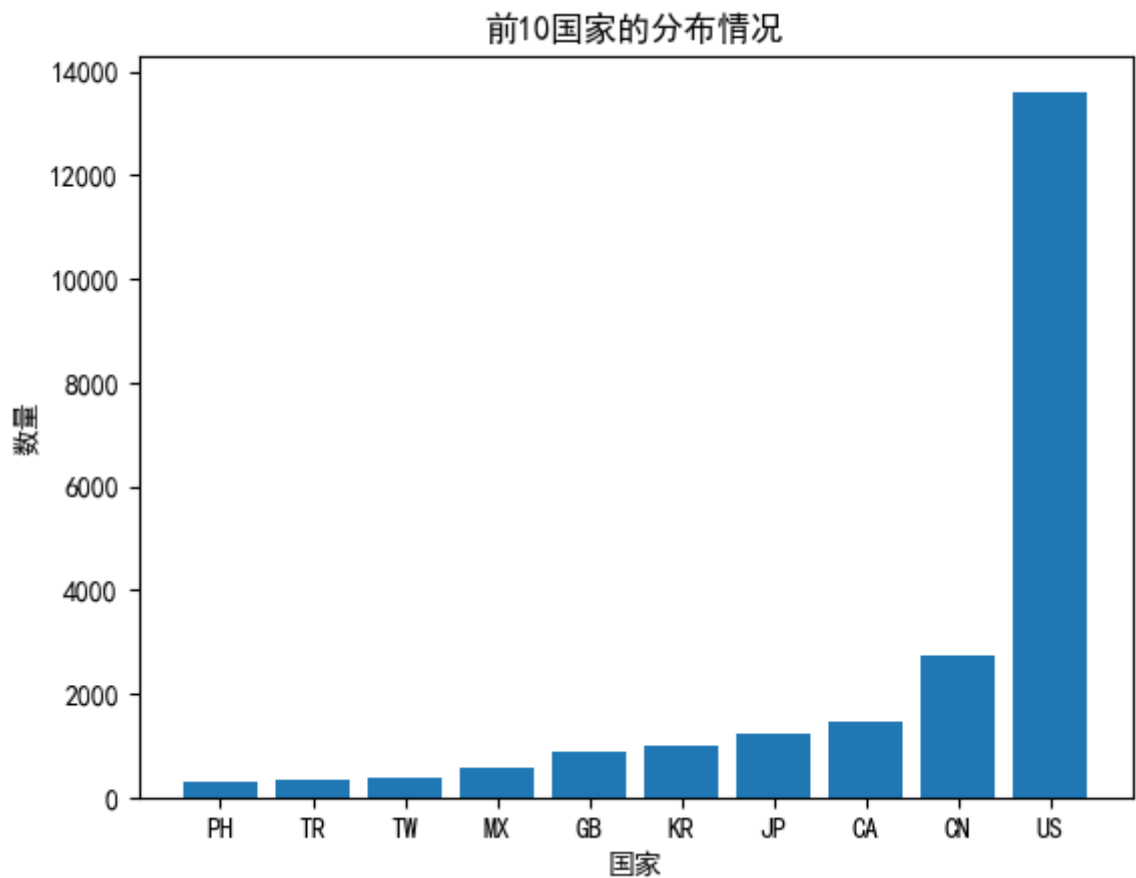
Name: City, dtype: int64

### (3) 用柱状图绘制排名前10的国家的分布情况。

```
In [ ]: from numpy import arange

country=data["Country"].value_counts().sort_values()[-10:]
```

```
plt.bar(country.index.values, country)
plt.title("前10国家的分布情况")
plt.xlabel("国家")
plt.ylabel("数量")
plt.show()
```



(4) 显示拥有星巴克门店数量排名前10 的国家和地区。

```
In [ ]: print("门店数量前10的国家")
print(data["Country"].value_counts().sort_values()[-10:])

print("门店数量前10的州/省")
print(data["State/Province"].value_counts().sort_values()[-10:])
```

门店数量前10的国家

PH	298
TR	326
TW	394
MX	579
GB	901
KR	993
JP	1237
CA	1468
CN	2734
US	13608

Name: Country, dtype: int64

门店数量前10的州/省

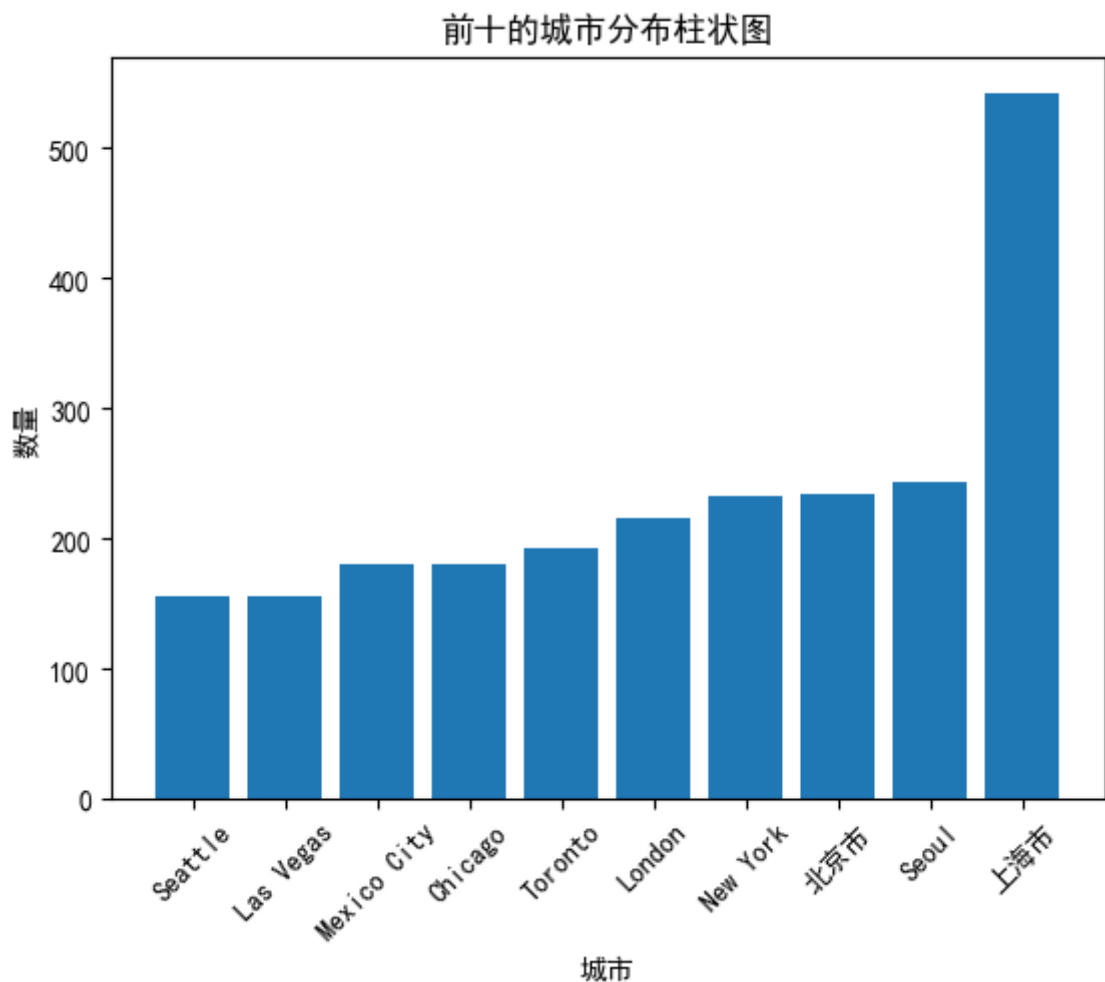
AZ	536
IL	575
31	580
NY	645
FL	694
11	706
WA	757
ENG	787
TX	1042
CA	2821

Name: State/Province, dtype: int64

## (5) 绘制星巴克门店数前10的城市分布柱状图

```
In [ ]: city=data["City"].value_counts().sort_values()[-10:]
plt.bar(city.index.values,city)
plt.xticks(rotation=45)
plt.title("前十的城市分布柱状图")
plt.xlabel("城市")
plt.ylabel("数量")
```

Out[ ]: Text(0, 0.5, '数量')



(6) 按照星巴克门店在中国的分布情况，统计排名前10 的城市。

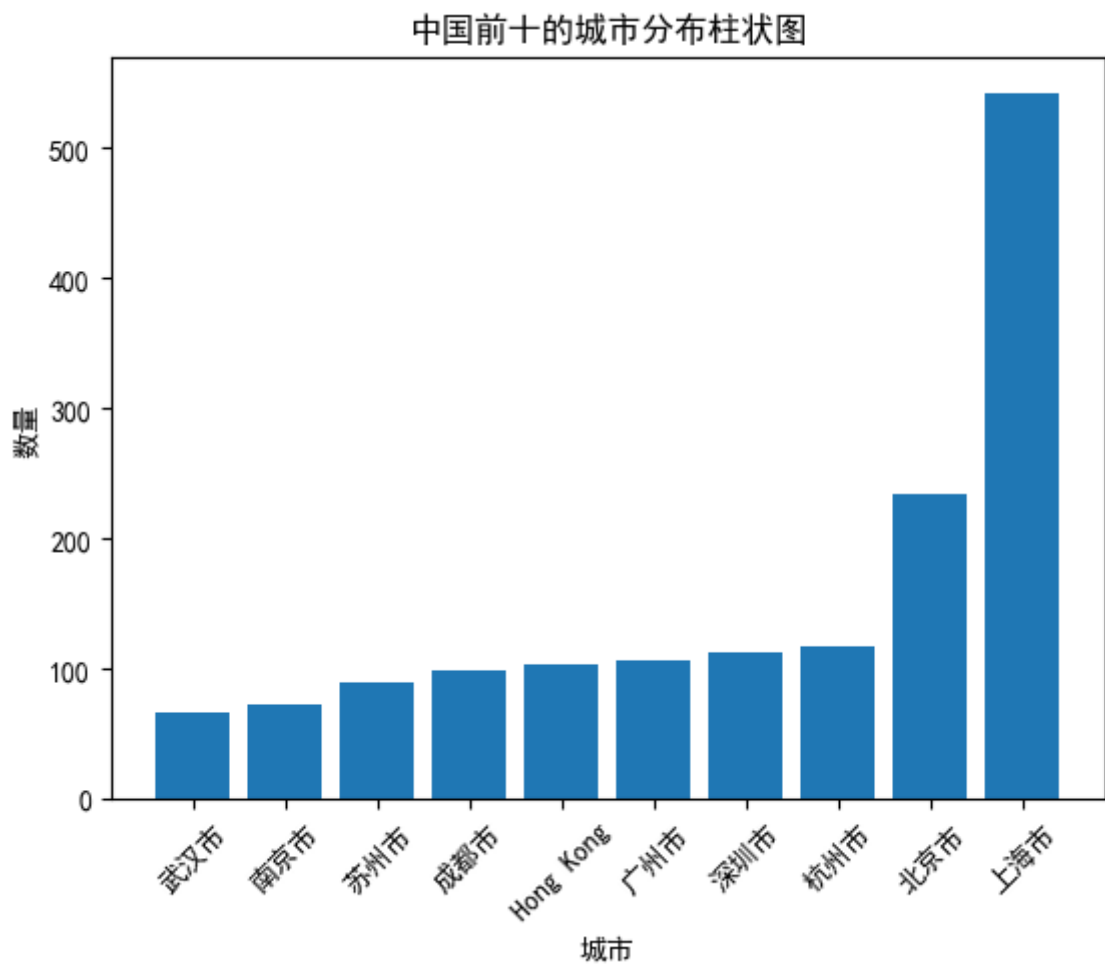
```
In [ ]: city=data[data["Country"]=="CN"]["City"].value_counts().sort_values()[-10:]
print(city)
```

```
武汉市          67
南京市          73
苏州市          90
成都市          98
Hong Kong      104
广州市          106
深圳市          113
杭州市          117
北京市          234
上海市          542
Name: City, dtype: int64
```

(8) 绘制中国排名前十的城市

```
In [ ]: city=data[data["Country"]=="CN"]["City"].value_counts().sort_values()[-10:]
plt.bar(city.index.values,city)
plt.xticks(rotation=45)
plt.title("中国前十的城市分布柱状图")
plt.xlabel("城市")
plt.ylabel("数量")
```

```
Out[ ]: Text(0, 0.5, '数量')
```



(8) 用饼状图显示星巴克门店的经营方式有哪几种。

```
In [ ]: type=data["Ownership Type"].value_counts()
plt.pie(type,labels=type.index.values,autopct="%.1f%%")
plt.title("星巴克门店经营方式")
```

```
Out[ ]: Text(0.5, 1.0, '星巴克门店经营方式')
```

星巴克门店经营方式

