山东大学计算机科学与技术学院

可视化技术课程实验报告

学号：202300130239	姓名：陆云	班级：23 数据
实验题目：数据质量实践		
实验学时：2	实验日期：2025. 9. 19	
实验目标：		

实验步骤：

1. 删去最后两行：

```
[1]: import pandas as pd
     from pandas import DataFrame
     import numpy as np
     import matplotlib.pyplot as plt

     import chardet

     # 先检测文件编码
     with open("Pokemon.csv", 'rb') as f:
         result = chardet.detect(f.read())  # 检测编码
         encoding = result['encoding']
         print(f"检测到的编码: {encoding}")

     data = pd.read_csv("Pokemon.csv", encoding="MacRoman")
     data
```

检测到的编码: MacRoman

[1]:

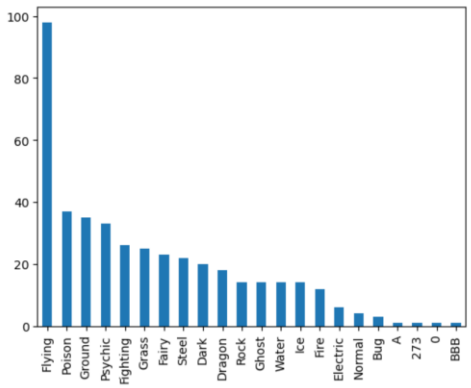
	#	Name	Type 1	Type 2	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary
0	1	Bulbasaur	Grass	Poison	318	45	49	49	65	65	45	1	FALSE
1	2	Ivysaur	Grass	Poison	405	60	62	63	80	80	60	1	FALSE
2	3	Venusaur	Grass	Poison	525	80	82	83	100	100	80	1	FALSE
3	3	VenusaurMega Venusaur	Grass	Poison	625	80	100	123	122	120	80	1	FALSE
4	4	Charmander	Fire	NaN	309	39	52	43	60	50	65	1	FALSE
...	...	...	...	...	...	...	...	...	...	...	...	...	...
805	721	Volcanion	Fire	Water	600	80	110	120	130	90	70	6	TRUE
806	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined
807	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined	undefined
808	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
809	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

810 rows × 13 columns

2. 清空 type2 存在异常的数值

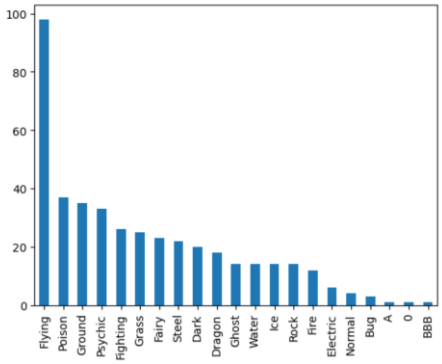
```
[3]: # • type2存在异常的数值取值，可清空
     data["Type 2"].value_counts().plot(kind='bar')
```

[3]: <Axes: >



[4]:
```
cut2 = data[data["Type 2"] == "273"]
data = data[data["Type 2"] != "273"]
data["Type 2"].value_counts().plot(kind='bar')
```

[4]: <Axes: >



3. 删除重复行

[5]:
```
# 筛选出数据框 data 中重复行的操作
data[data.duplicated()]

# # 删除data中的重复行
# data = data.drop_duplicates()
# data
```

[5]:

	#	Name	Type 1	Type 2	Total	HP	Attack	Defense	Sp. Atk	Sp. Def	Speed	Generation	Legendary
15	11	Metapod	Bug	NaN	205	50	20	55	25	25	30	1	FALSE
23	17	Pidgeotto	Normal	Flying	349	63	60	55	50	50	71	1	FALSE
185	168	Ariados	Bug	Poison	390	70	90	70	60	60	40	2	FALSE
186	168	Ariados	Bug	Poison	390	70	90	70	60	60	40	2	FALSE
187	168	Ariados	Bug	Poison	390	70	90	70	60	60	40	2	FALSE

结果图片：