

DataFrameExercises

```
import pandas as pd
import numpy as np
```

案例1:學生成績分析

場景:某班級的學生成績資料如下,請完成以下任務:

1. 計算每位學生的總分和平均分。
2. 找出數學成績高於90分或英語成績高於85分的學生。
3. 按總分從高到低排序,並輸出前3名學生。

```
import pandas as pd \
data = {\
'姓名': ['張三', '李四', '王五', '趙六', '錢七'],\
'數學': [85, 92, 78, 88, 95],\
'英語': [90, 88, 85, 92, 80],\
'物理': [75, 80, 88, 85, 90]\
}
```

```
data = {\
    '姓名': ['張三', '李四', '王五', '趙六', '錢七'],\
    '數學': [85, 92, 78, 88, 95],\
    '英語': [90, 88, 85, 92, 80],\
    '物理': [75, 80, 88, 85, 90]\
}
```

```
scores = pd.DataFrame(data)
scores
```

	姓名	數學	英語	物理
0	張三	85	90	75
1	李四	92	88	80

	姓名	數學	英語	物理
2	王五	78	85	88
3	趙六	88	92	85
4	錢七	95	80	90

```
scores["Total"]("數學",%20"英語",%20"物理").sum(axis=1)
scores
```

	姓名	數學	英語	物理	Total
0	張三	85	90	75	250
1	李四	92	88	80	260
2	王五	78	85	88	251
3	趙六	88	92	85	265
4	錢七	95	80	90	265

```
scores["Average"]("數學",%20"英語",%20"物理").mean(axis=1)
scores
```

	姓名	數學	英語	物理	Total	Average
0	張三	85	90	75	250	83.333333
1	李四	92	88	80	260	86.666667
2	王五	78	85	88	251	83.666667
3	趙六	88	92	85	265	88.333333
4	錢七	95	80	90	265	88.333333

```
scores[(scores["數學"] > 90) | (scores["英語"] > 85)]
```

	姓名	數學	英語	物理	Total	Average
0	張三	85	90	75	250	83.333333
1	李四	92	88	80	260	86.666667
3	趙六	88	92	85	265	88.333333
4	錢七	95	80	90	265	88.333333

```
scores.loc[(scores["數學"] > 90) | (scores["英語"] > 85), "姓名"]
```

姓名 数学 英语 物理 Total Average

0 张三 85 90 75 250 83.333333

1 李四 92 88 80 260 86.666667

3 赵六 88 92 85 265 88.333333

4 钱七 95 80 90 265 88.333333

```
scores.sort_values("Total", ascending=False)
```

	姓名	數學	英語	物理	Total	Average
4	錢七	95	80	90	265	88.333333
3	趙六	88	92	85	265	88.333333
1	李四	92	88	80	260	86.666667
2	王五	78	85	88	251	83.666667
0	張三	85	90	75	250	83.333333

```
scores.sort_values("Total", ascending=False).head(3)
```

	姓名	數學	英語	物理	Total	Average
4	錢七	95	80	90	265	88.333333
3	趙六	88	92	85	265	88.333333
1	李四	92	88	80	260	86.666667

```
scores.nlargest(3, columns="Total")
```

	姓名	數學	英語	物理	Total	Average
3	趙六	88	92	85	265	88.333333
4	錢七	95	80	90	265	88.333333
1	李四	92	88	80	260	86.666667

案例2:銷售資料分析

場景:某公司銷售資料如下,請完成以下任務:

1. 計算每種產品的總銷售額(銷售額單價×銷量)。

2. 找出銷售額最高的產品。
3. 按銷售額從高到低排序,並輸出所有產品資訊。

```
import pandas as pd
```

```
data = {
    '產品名稱': ['A', 'B', 'C', 'D'],
    '單價': [100, 150, 20, 120],
    '銷量': [50, 30, 20, 40]
}
df = pd.DataFrame(data)
```

```
data = {
    '產品名稱': ['A', 'B', 'C', 'D'],
    '單價': [100, 150, 20, 120],
    '銷量': [50, 30, 20, 40],
}
df = pd.DataFrame(data)
df
```

	產品名稱	單價	銷量
0	A	100	50
1	B	150	30
2	C	20	20
3	D	120	40

```
df["Total sales price"] = df["單價"] * df["銷量"]
df
```

	產品名稱	單價	銷量	Total sales	Total sales price
0	A	100	50	5000	5000
1	B	150	30	4500	4500
2	C	20	20	400	400
3	D	120	40	4800	4800

```
df["Total sales price"].max()
df.loc[df["Total sales price"] == df["Total sales price"].max(), "產品名稱"]
```

0 A

Name: 產品名稱, dtype: object

```
df.sort_values("Total sales price", ascending=False)
```

	產品名稱	單價	銷量	Total sales	Total sales price
0	A	100	50	5000	5000
3	D	120	40	4800	4800
1	B	150	30	4500	4500
2	C	20	20	400	400

案例3:電商使用者行為分析

場景:某電商平臺的使用者行為資料如下,請完成以下任務:

1. 計算每位使用者的總消費金額(消費金額= 商品單價 購買數量)
2. 找出消費金額最高的使用者,並輸出其所有資訊
3. 計算所有使用者的平均消費金額(保留2位小數)
4. 統計電子產品的總購買數量

```
import pandas as pd
data = {
    '使用者ID': [101, 102, 103, 104, 105],
    '使用者名稱': ['Alice', 'Bob', 'Charlie', 'David', 'Eve'],
    '商品類別': ['電子產品', '服飾', '電子產品', '家居', '服飾'],
    '商品單價': [1200, 300, 889, 150, 200],
    '購買數量': [1, 3, 2, 5, 4]
}
df = pd.DataFrame(data)
```