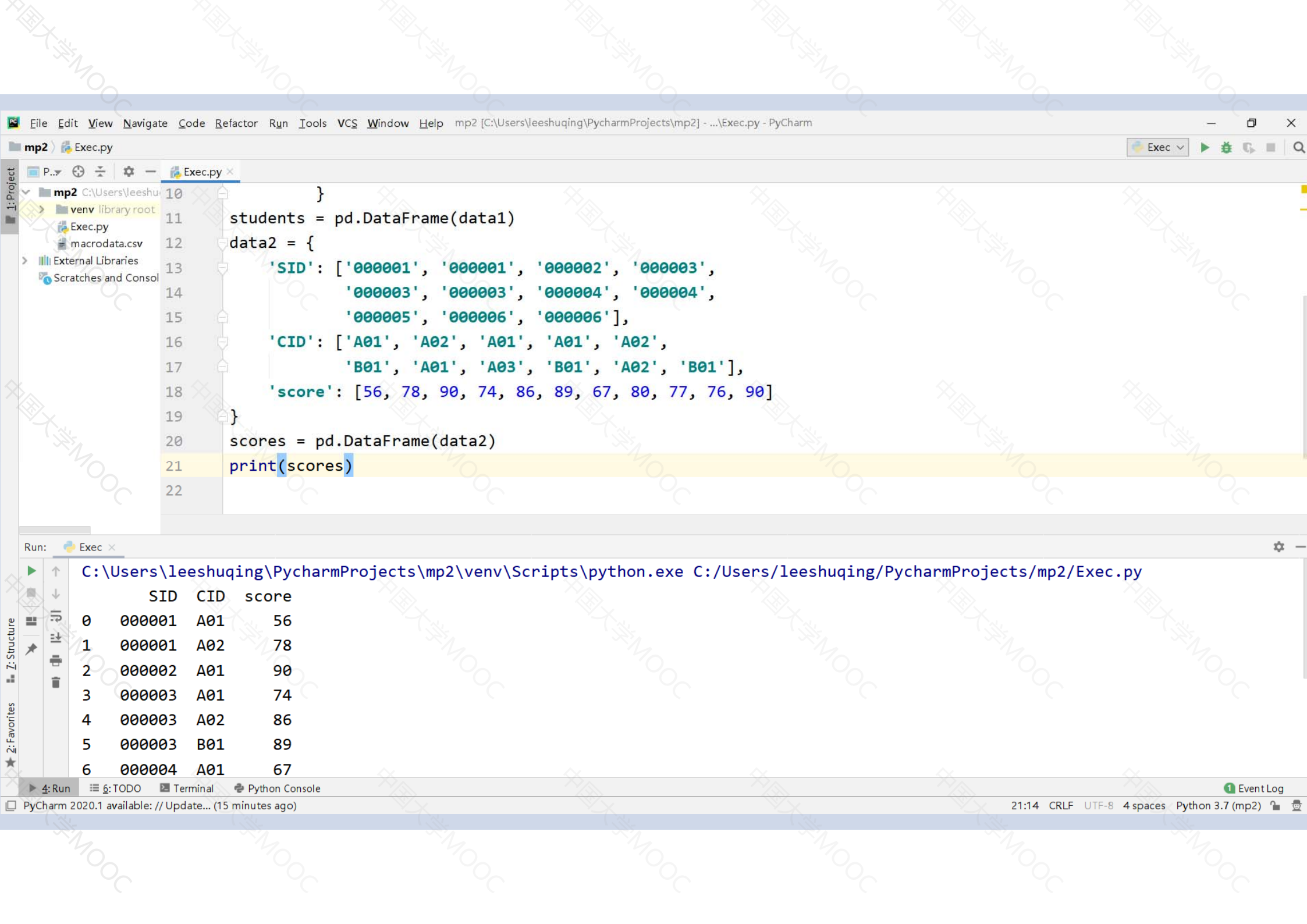


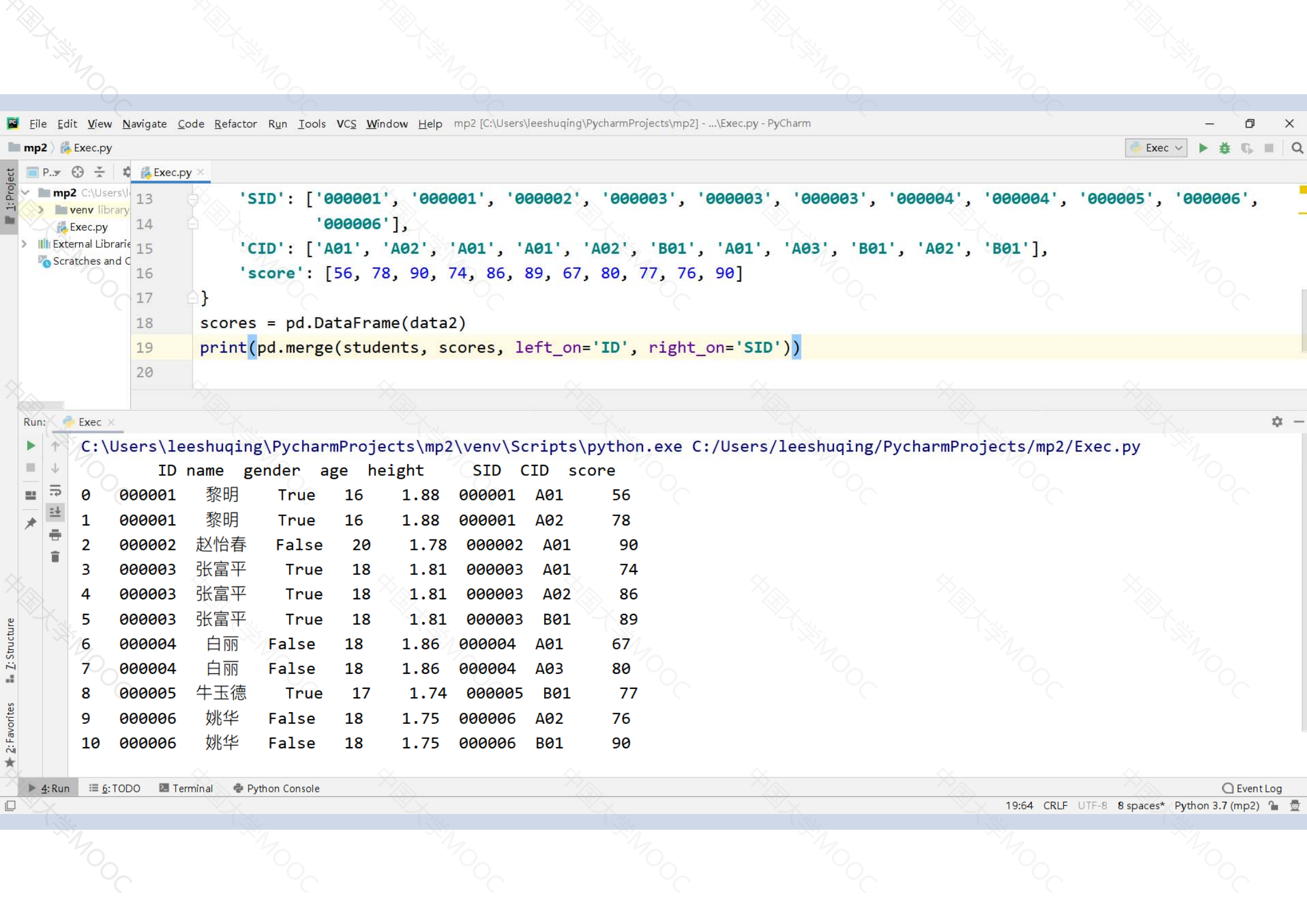
# Python大数据分析

## 七、数据的连接

# 成绩数据

SID	CID	Score
000001	A01	56
000001	A02	78
000002	A01	90
000003	A01	74
000003	A02	86
000003	B01	89
000004	A01	67
000004	A03	80
000005	B01	77
000006	A02	76
000006	B01	90

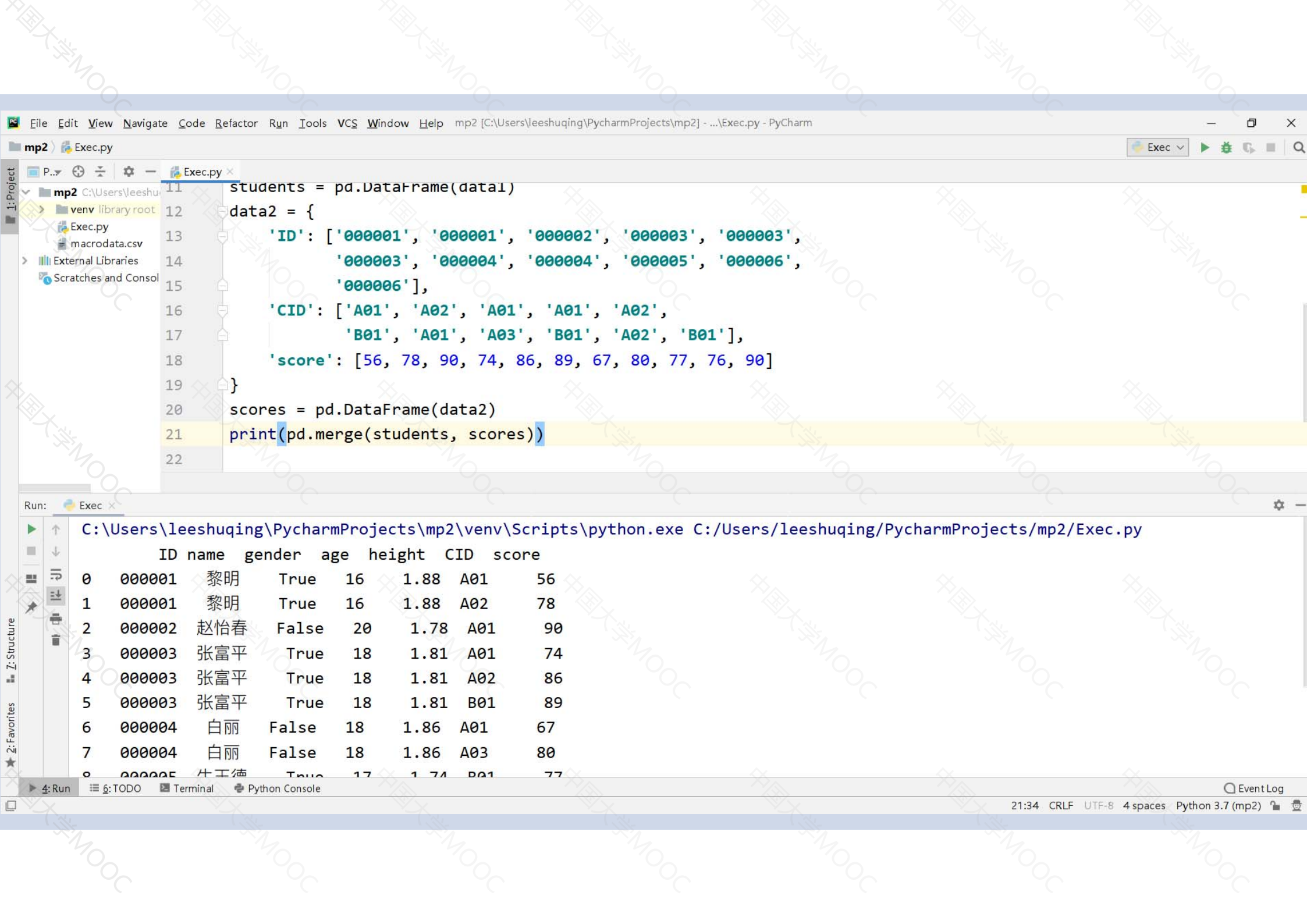




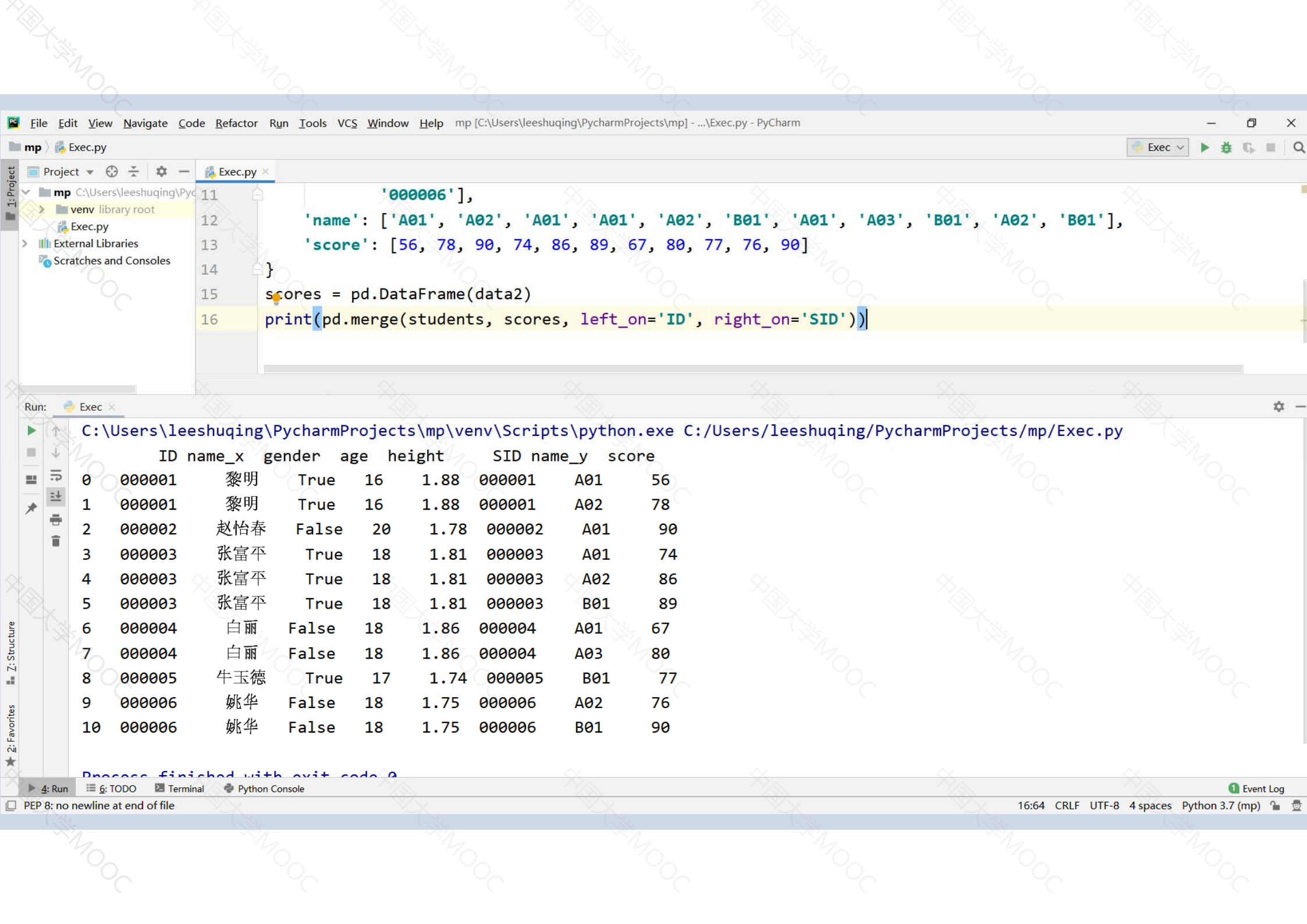
```
13 'SID': ['000001', '000001', '000002', '000003', '000003', '000003', '000004', '000004', '000005', '000006',
14       '000006'],
15 'CID': ['A01', 'A02', 'A01', 'A01', 'A02', 'B01', 'A01', 'A03', 'B01', 'A02', 'B01'],
16 'score': [56, 78, 90, 74, 86, 89, 67, 80, 77, 76, 90]
17 }
18 scores = pd.DataFrame(data2)
19 print(pd.merge(students, scores, left_on='ID', right_on='SID'))
20
```

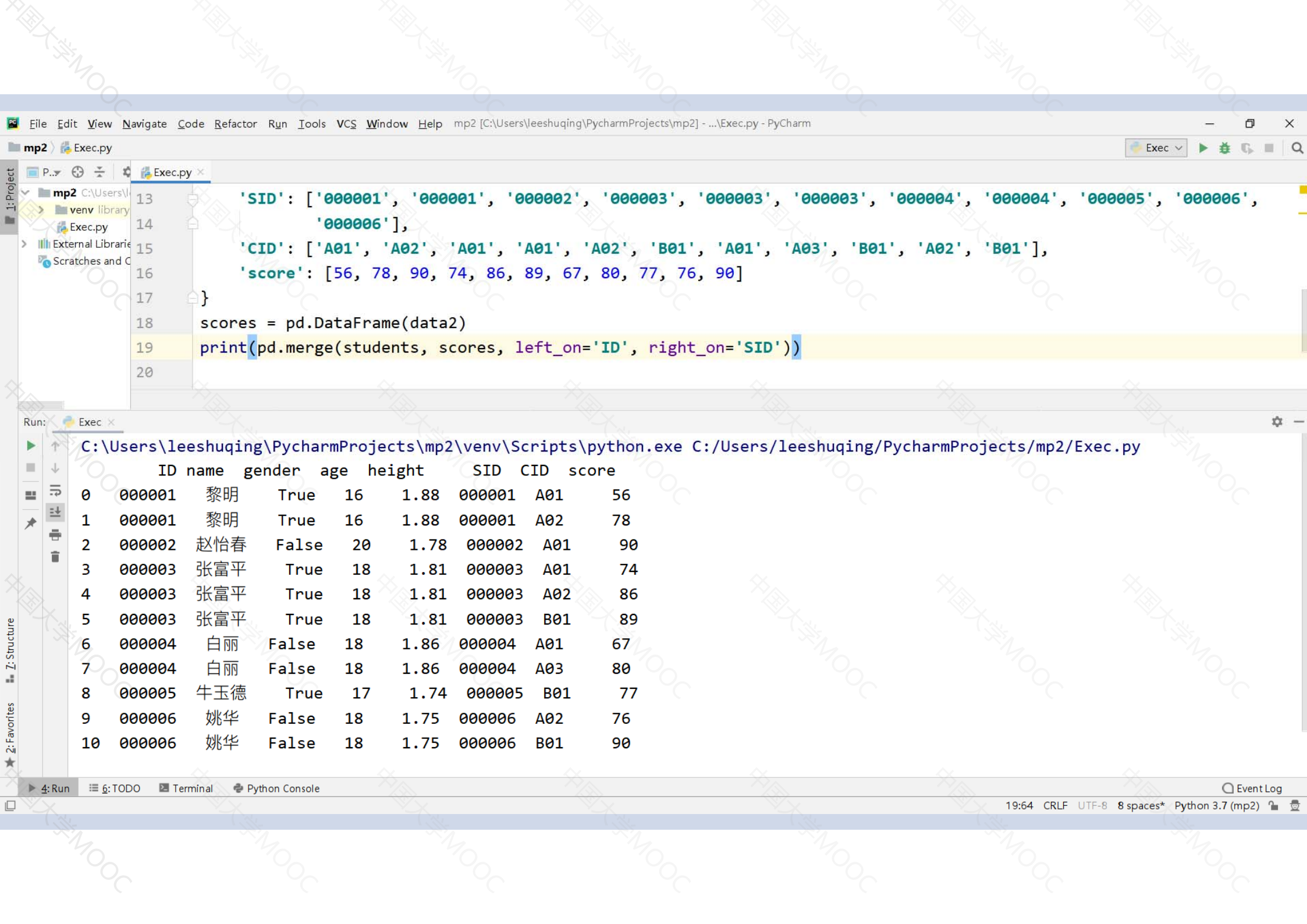
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py

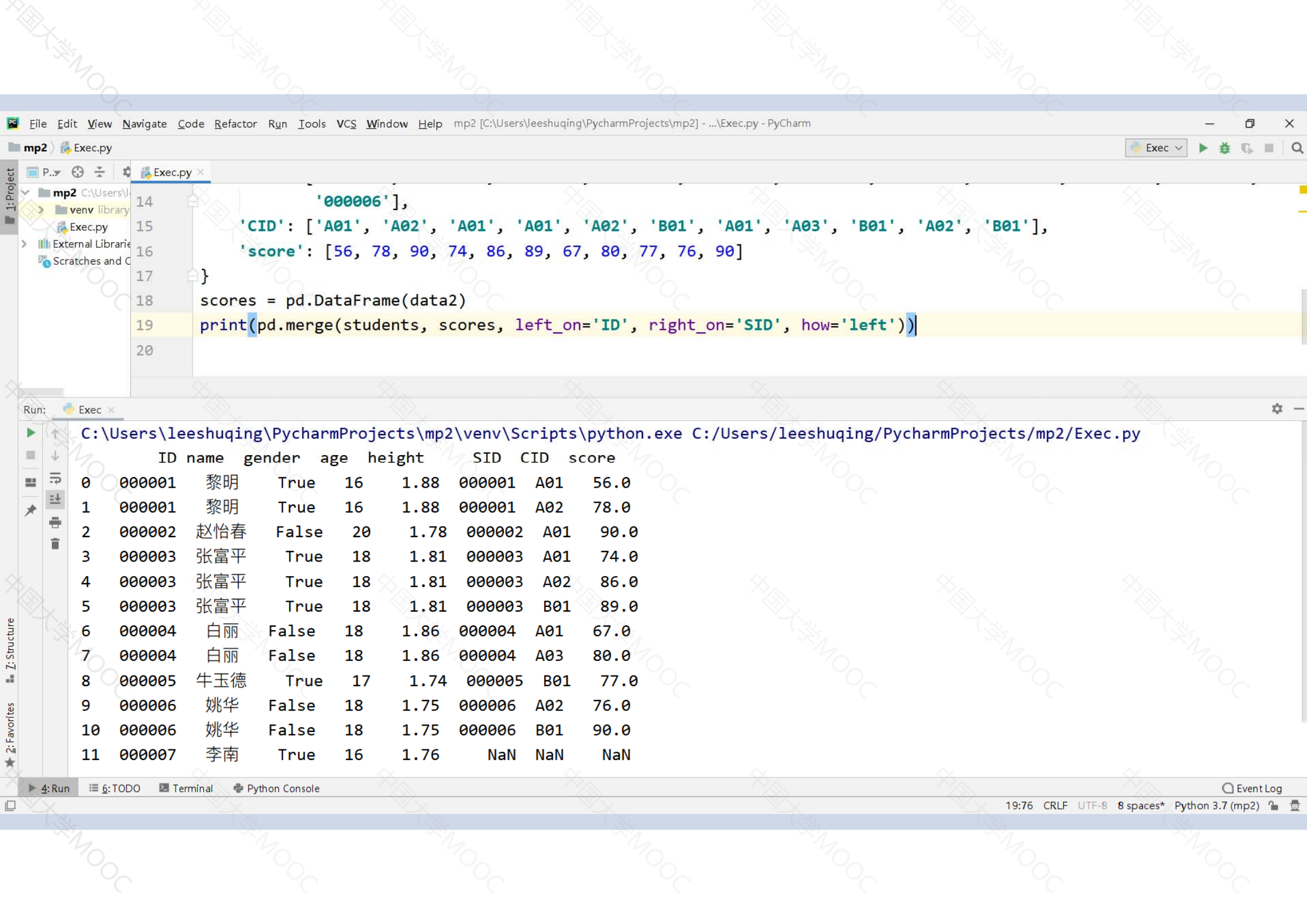
	ID	name	gender	age	height	SID	CID	score
0	000001	黎明	True	16	1.88	000001	A01	56
1	000001	黎明	True	16	1.88	000001	A02	78
2	000002	赵怡春	False	20	1.78	000002	A01	90
3	000003	张富平	True	18	1.81	000003	A01	74
4	000003	张富平	True	18	1.81	000003	A02	86
5	000003	张富平	True	18	1.81	000003	B01	89
6	000004	白丽	False	18	1.86	000004	A01	67
7	000004	白丽	False	18	1.86	000004	A03	80
8	000005	牛玉德	True	17	1.74	000005	B01	77
9	000006	姚华	False	18	1.75	000006	A02	76
10	000006	姚华	False	18	1.75	000006	B01	90













```
11 students = pd.DataFrame(data1)
12 data2 = {
13     'SID': ['000001', '000001', '000002', '000003',
14            '000003', '000003', '000004', '000004',
15            '000005', '000006', '000006'],
16     'CID': ['A01', 'A02', 'A01', 'A01', 'A02',
17            'B01', 'A01', 'A03', 'B01', 'A02', 'B01'],
18     'score': [56, 78, 90, 74, 86, 89, 67, 80, 77, 76, 90]
19 }
20 scores = pd.DataFrame(data2)
21 frame = pd.merge(students, scores, left_on='ID', right_on='SID', how='left')
22 print(frame[pd.isna(frame['score']) == True]['name'])
23
```

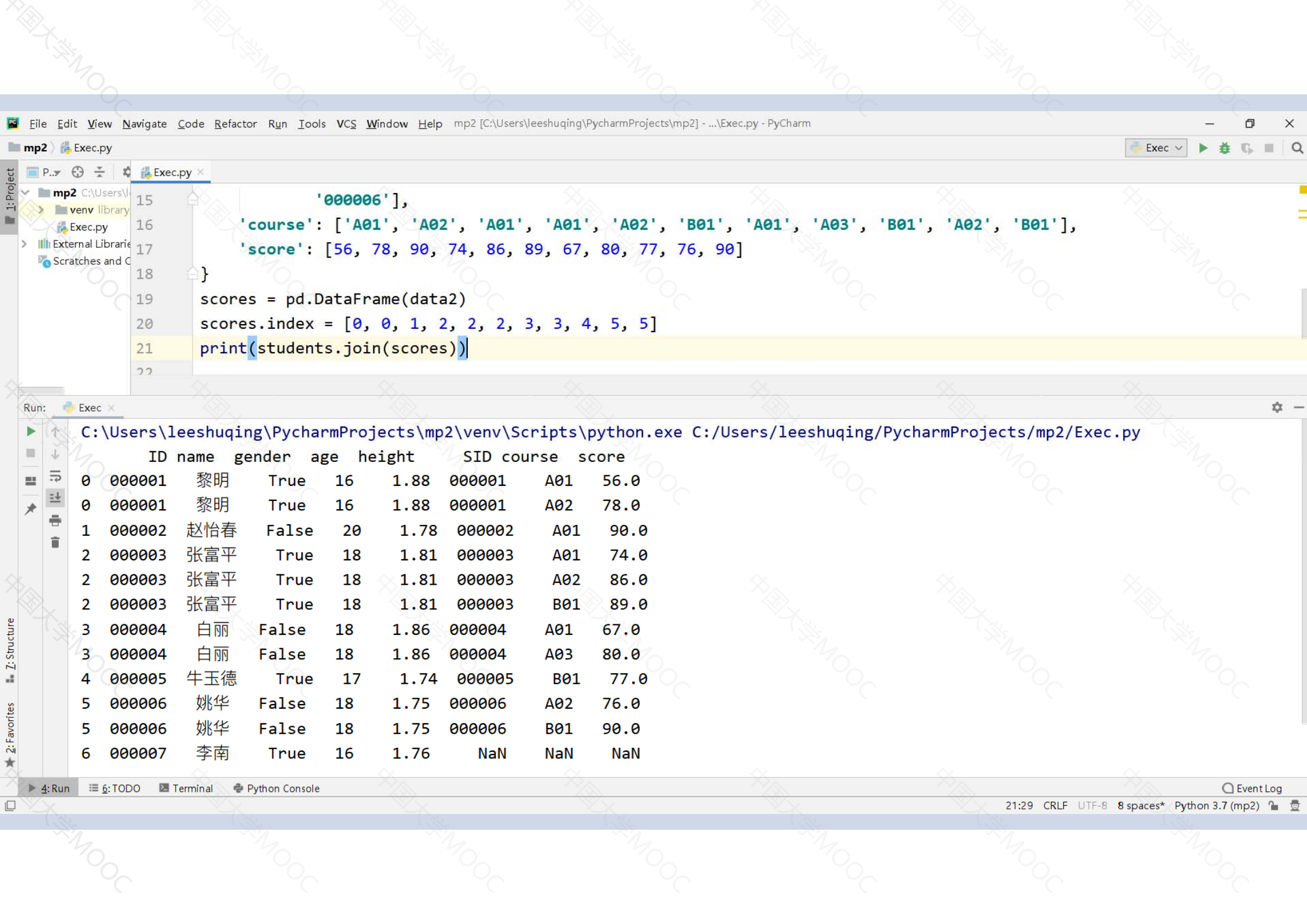
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py

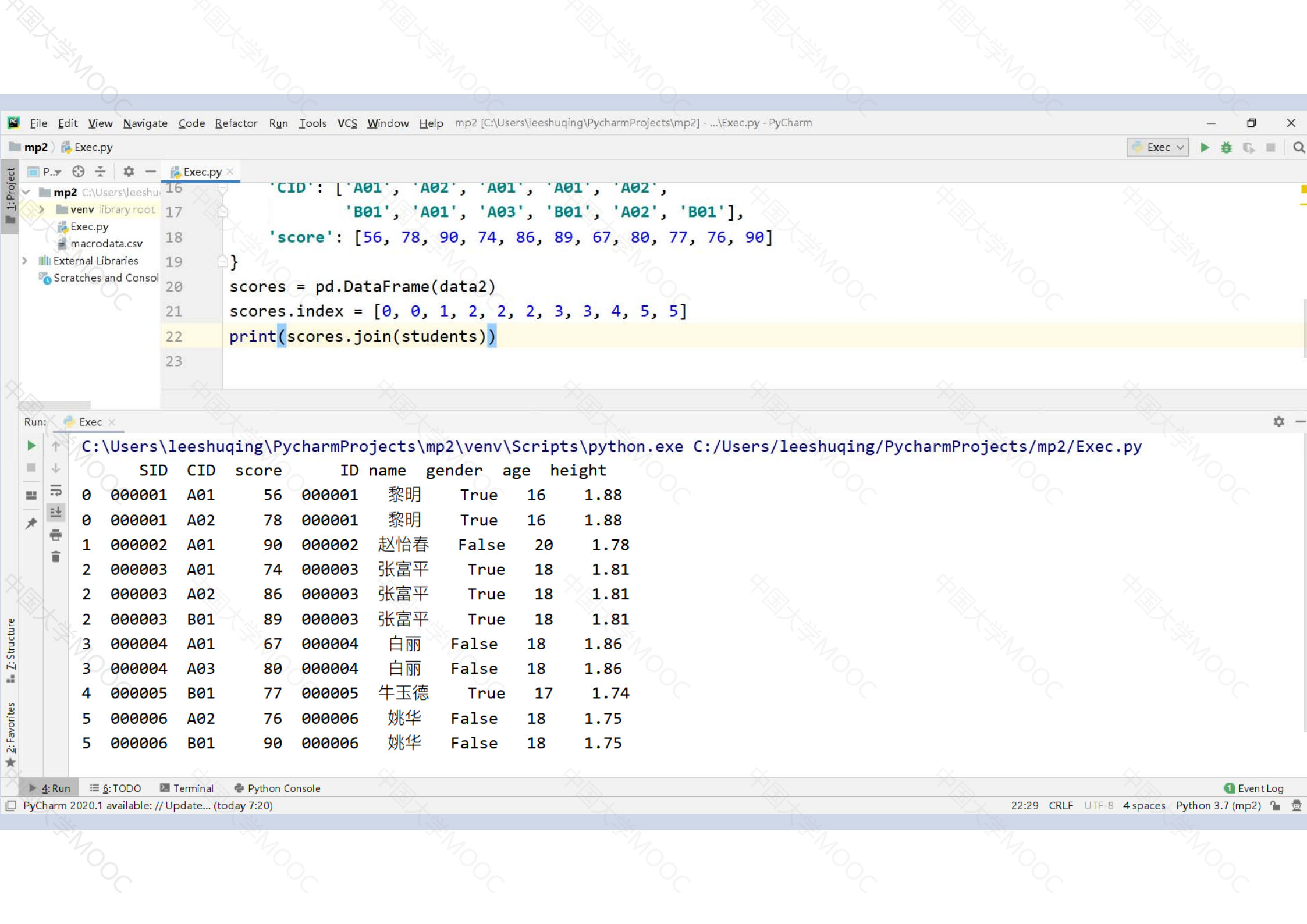
11 李南

Name: name, dtype: object

Process finished with exit code 0







```
16 'CID': ['A01', 'A02', 'A01', 'A01', 'A02',  
17        'B01', 'A01', 'A03', 'B01', 'A02', 'B01'],  
18 'score': [56, 78, 90, 74, 86, 89, 67, 80, 77, 76, 90]  
19 }  
20 scores = pd.DataFrame(data2)  
21 scores.index = [0, 0, 1, 2, 2, 2, 3, 3, 4, 5, 5]  
22 print(scores.join(students))  
23
```

C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py

	SID	CID	score	ID	name	gender	age	height
0	000001	A01	56	000001	黎明	True	16	1.88
0	000001	A02	78	000001	黎明	True	16	1.88
1	000002	A01	90	000002	赵怡春	False	20	1.78
2	000003	A01	74	000003	张富平	True	18	1.81
2	000003	A02	86	000003	张富平	True	18	1.81
2	000003	B01	89	000003	张富平	True	18	1.81
3	000004	A01	67	000004	白丽	False	18	1.86
3	000004	A03	80	000004	白丽	False	18	1.86
4	000005	B01	77	000005	牛玉德	True	17	1.74
5	000006	A02	76	000006	姚华	False	18	1.75
5	000006	B01	90	000006	姚华	False	18	1.75

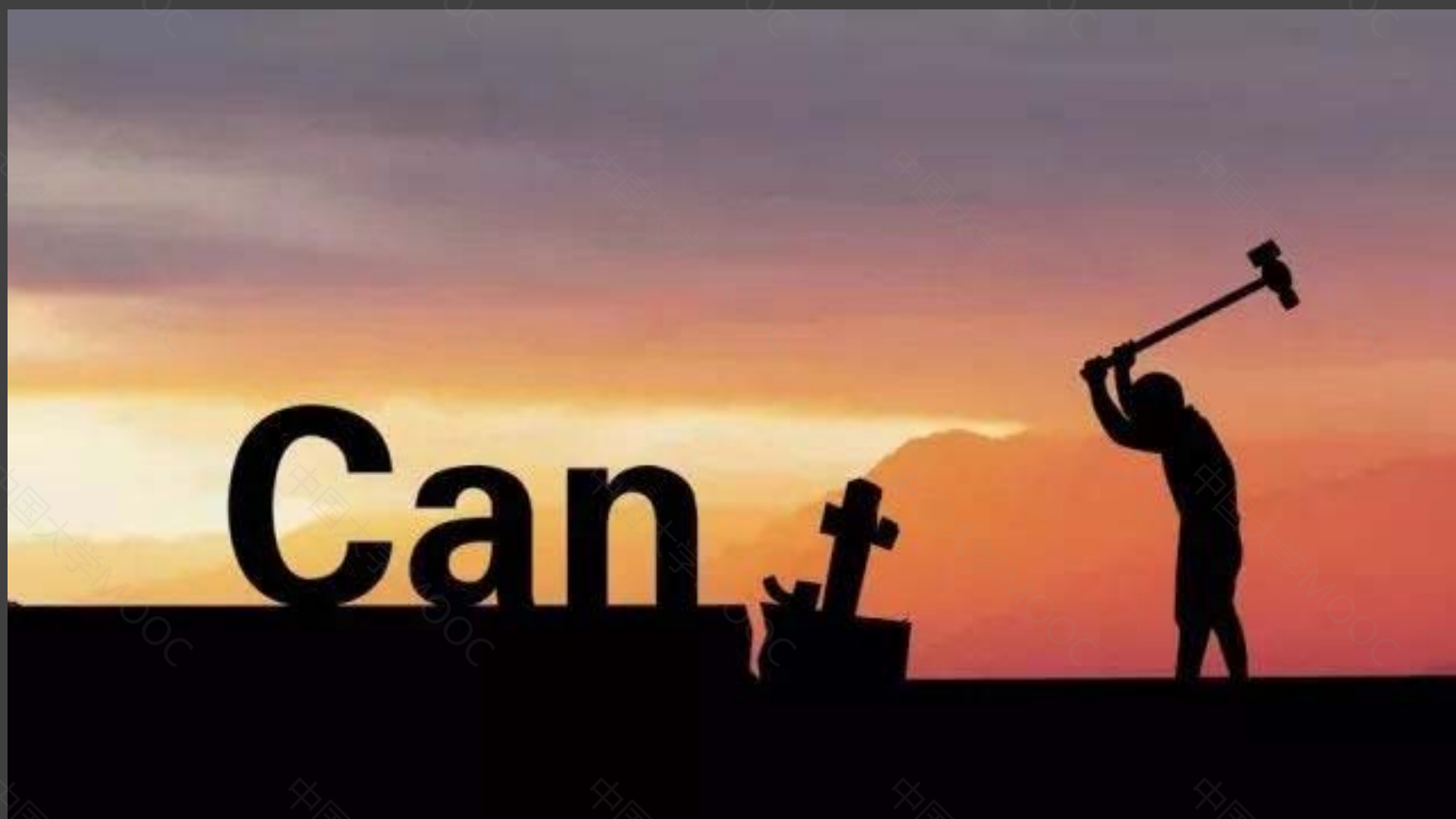








只有多练习：)



**一次不学多，下次再学**