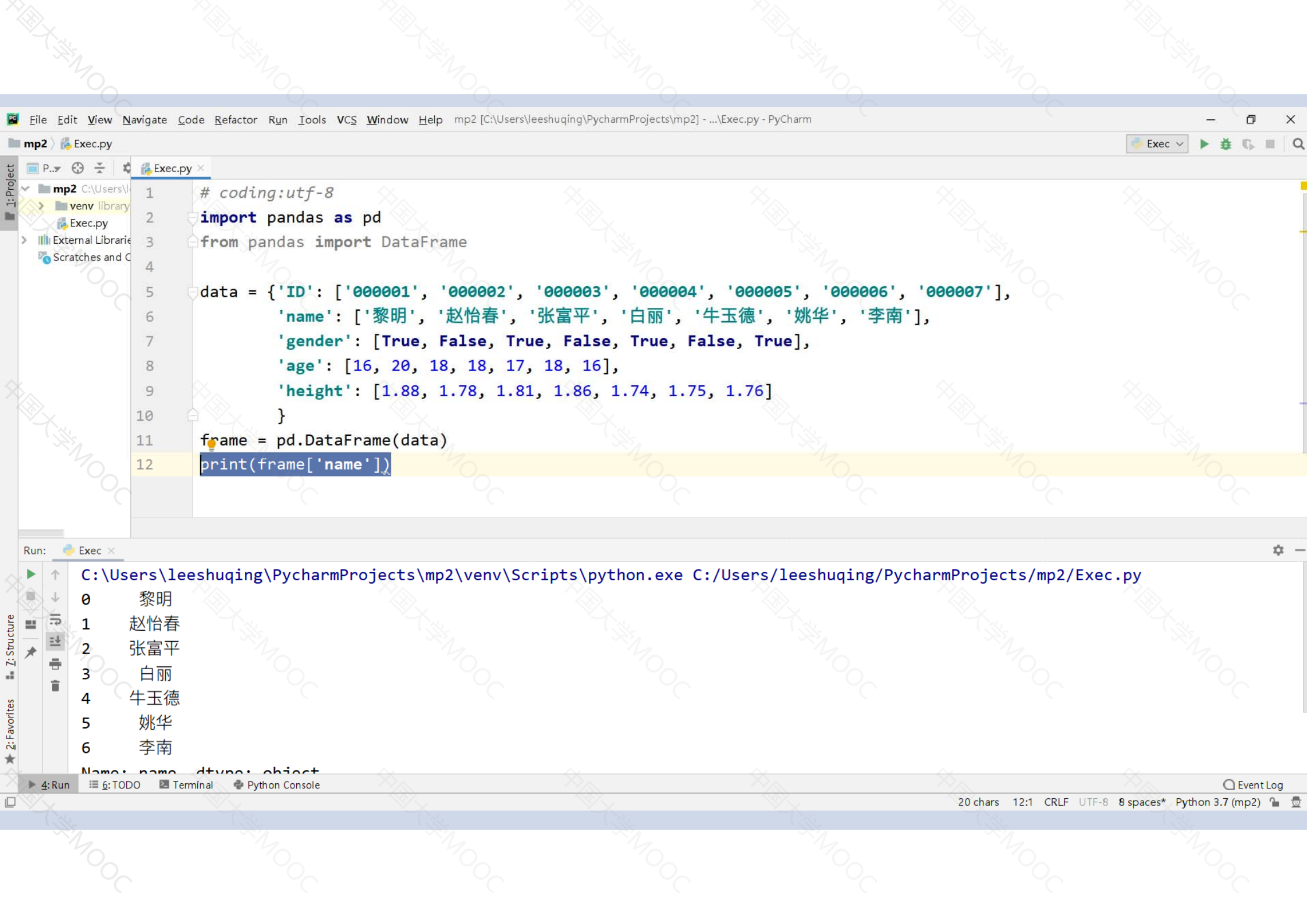


Python大数据分析

二、数据的选择



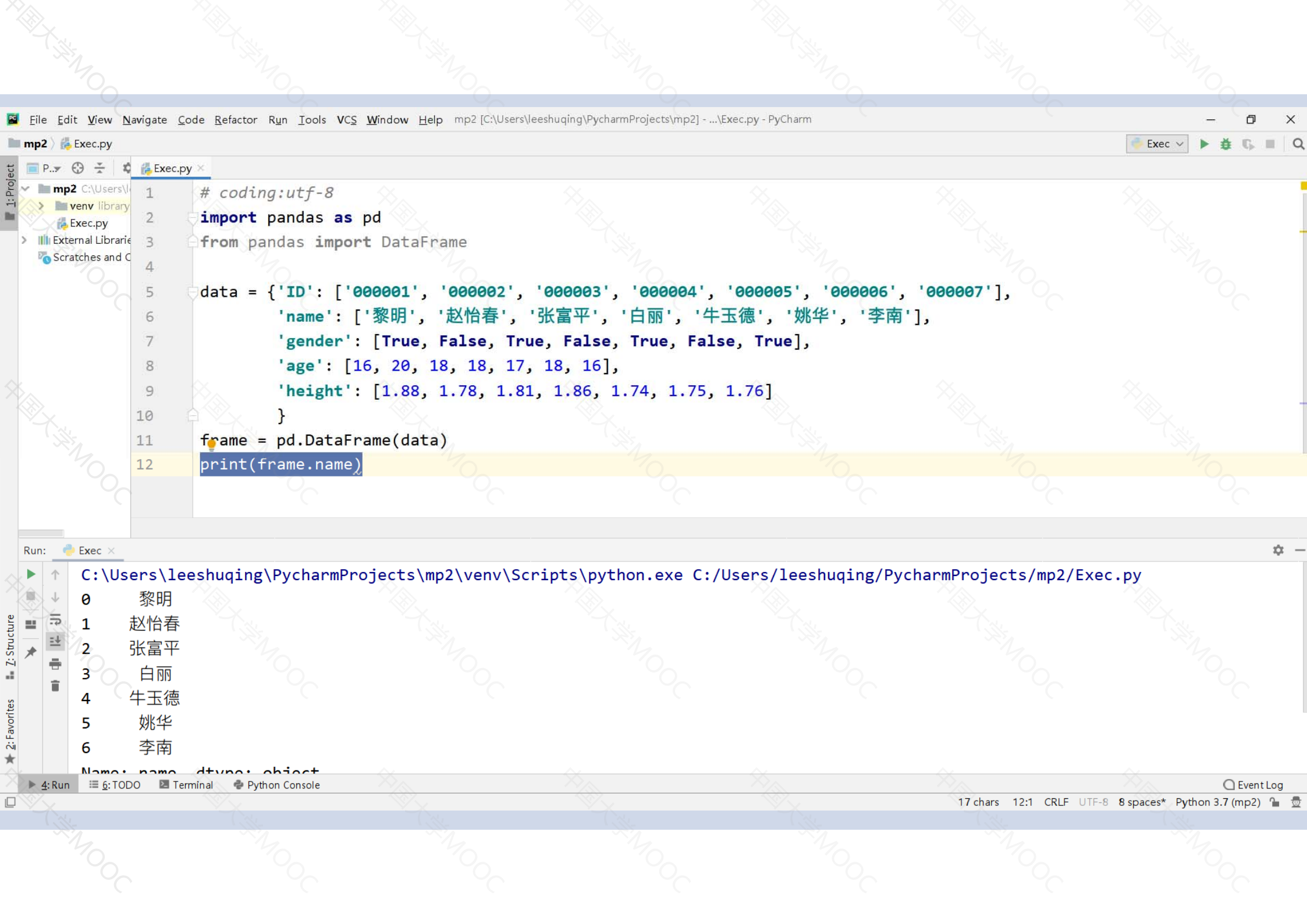
```
# coding:utf-8
import pandas as pd
from pandas import DataFrame

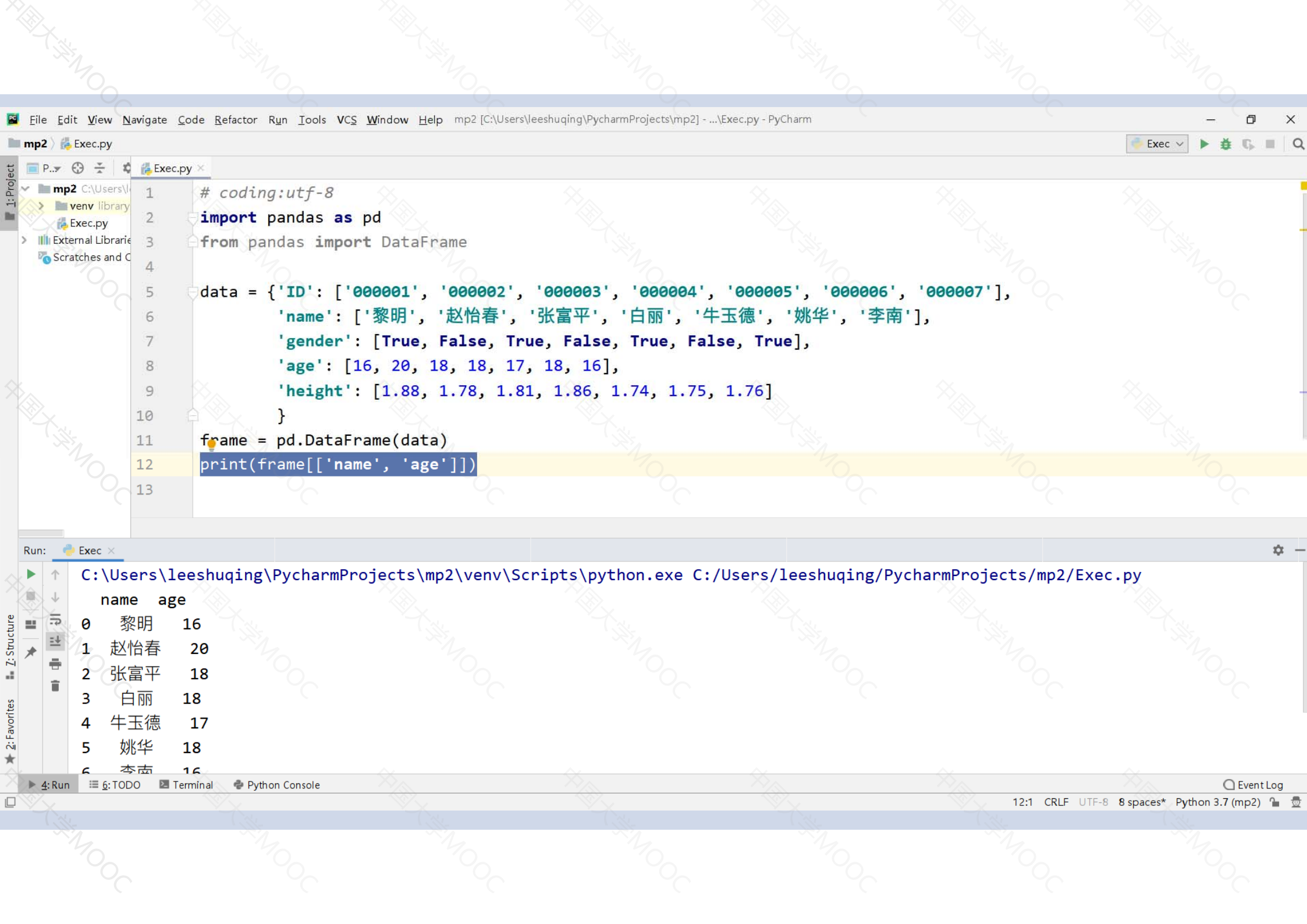
data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
        'gender': [True, False, True, False, True, False, True],
        'age': [16, 20, 18, 18, 17, 18, 16],
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}

frame = pd.DataFrame(data)
print(frame['name'])
```

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

```
0 黎明
1 赵怡春
2 张富平
3 白丽
4 牛玉德
5 姚华
6 李南
Name: name, dtype: object
```

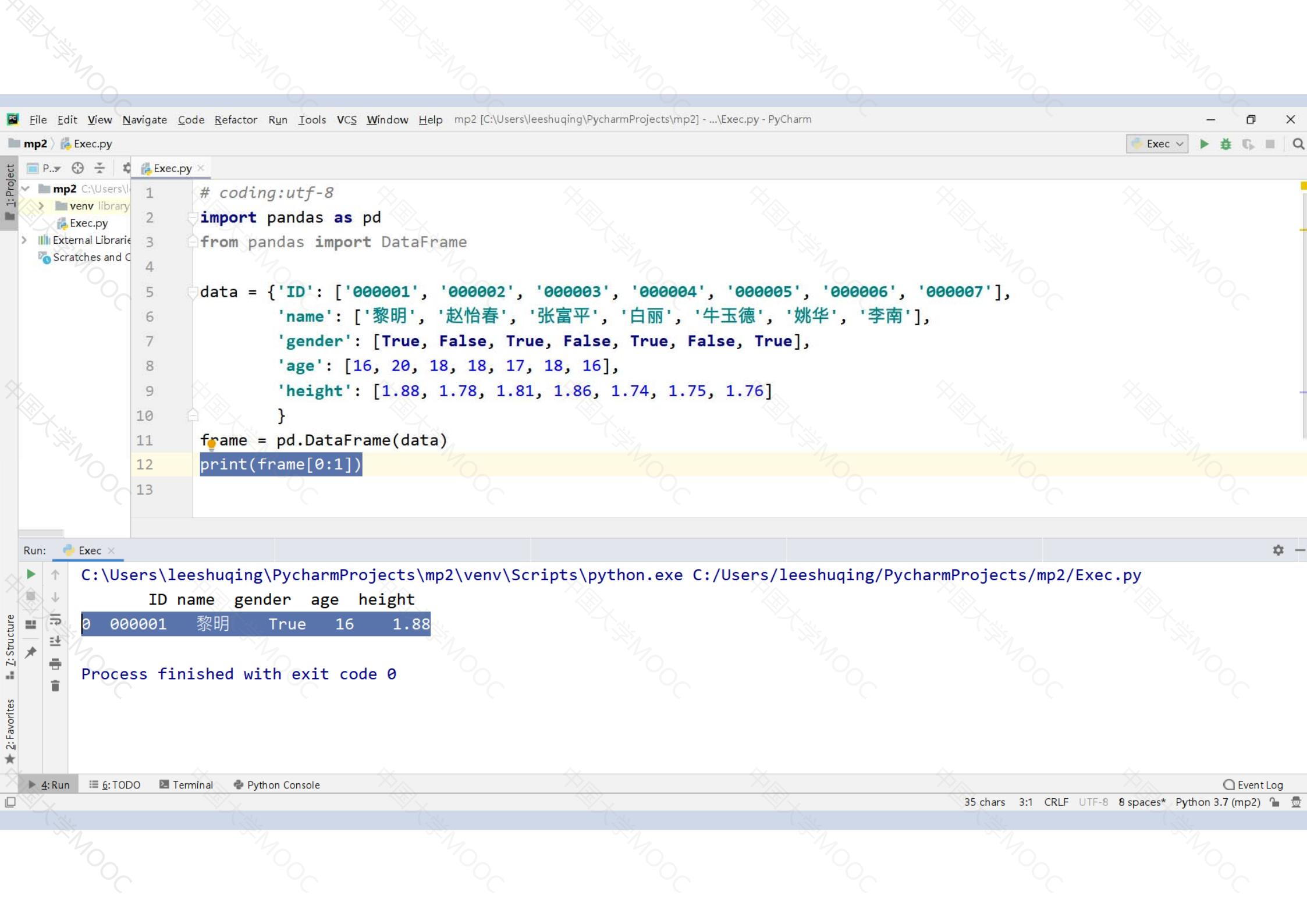


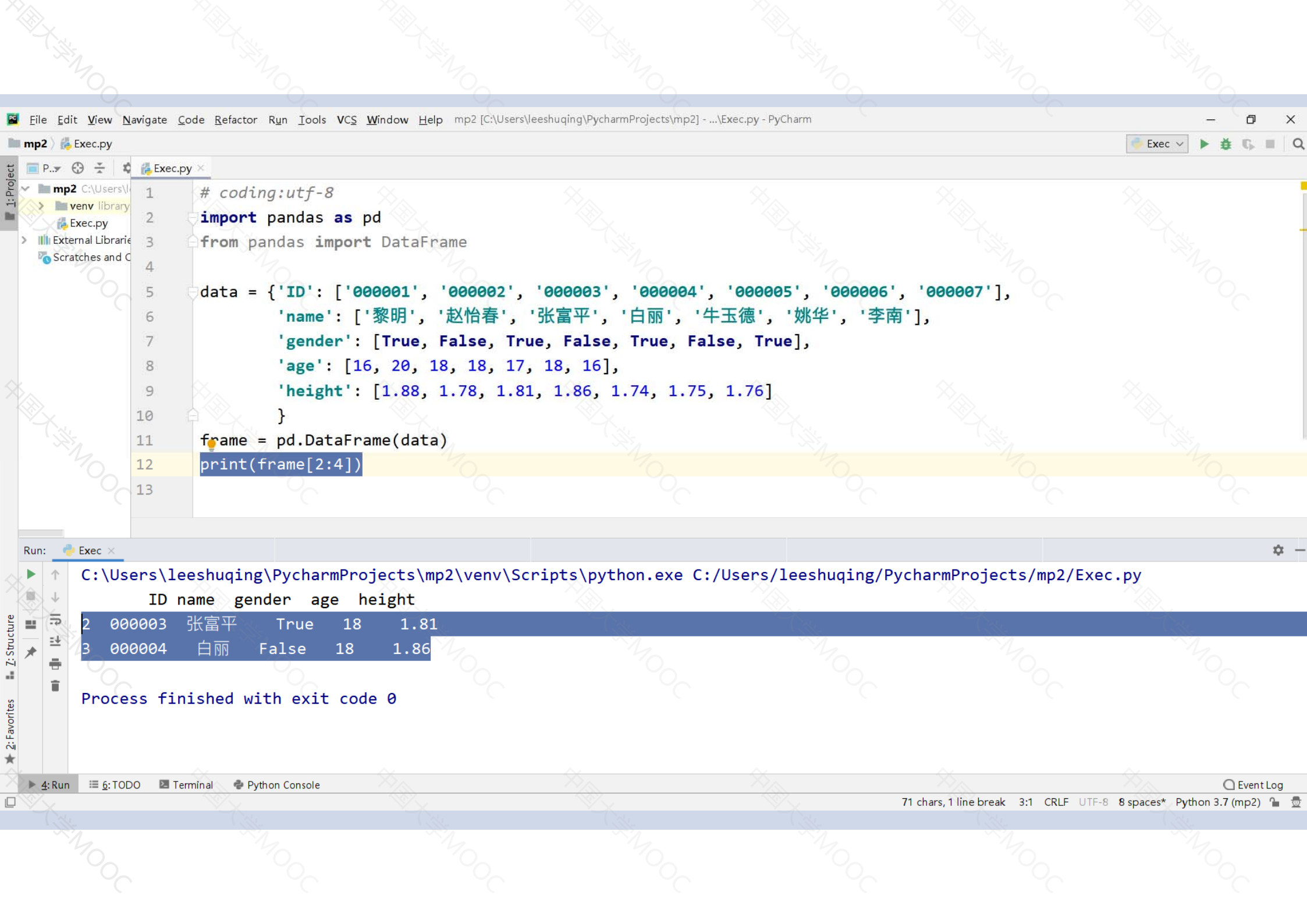


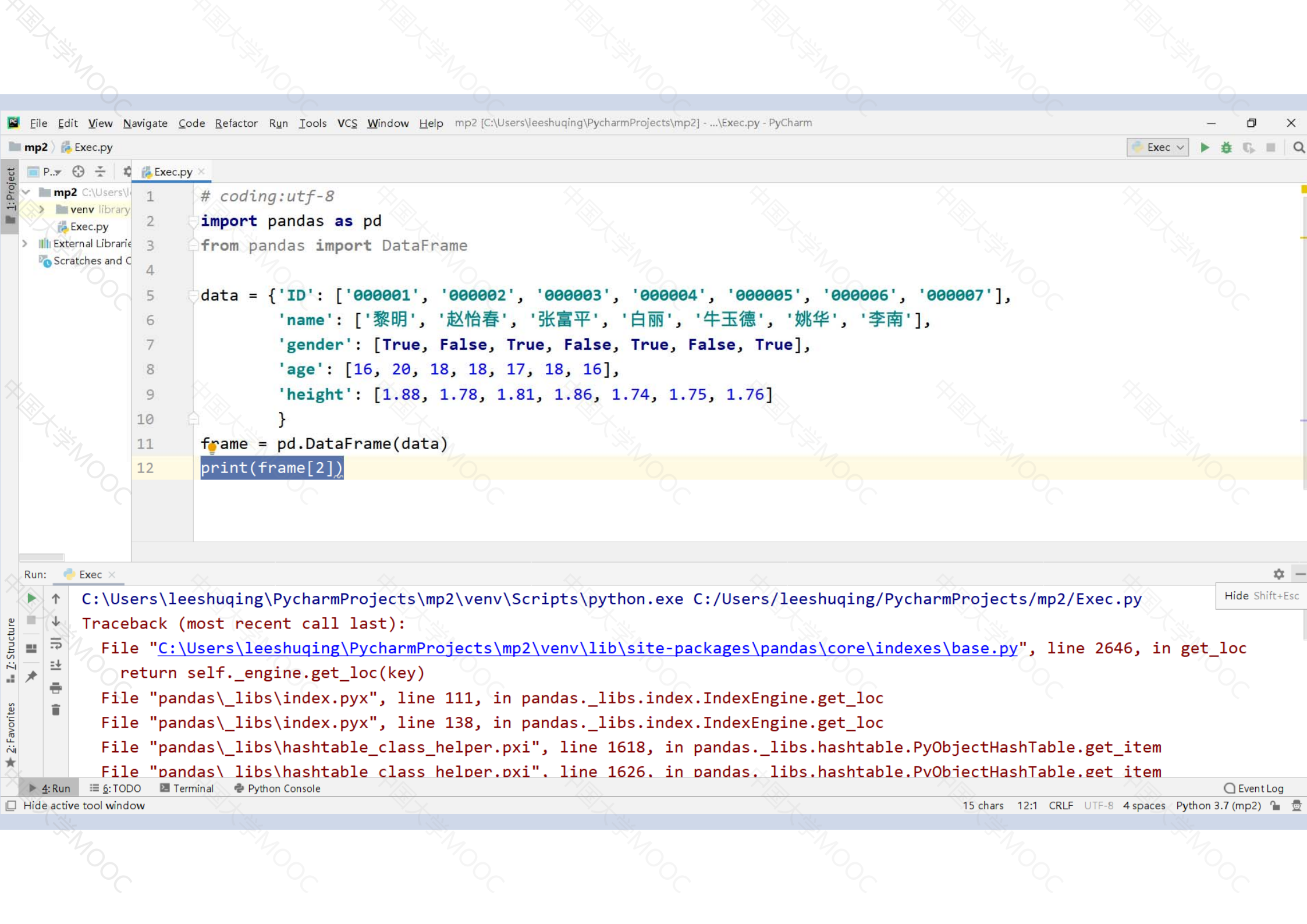
```
1 # coding:utf-8
2 import pandas as pd
3 from pandas import DataFrame
4
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7         'gender': [True, False, True, False, True, False, True],
8         'age': [16, 20, 18, 18, 17, 18, 16],
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}
10
11 frame = pd.DataFrame(data)
12 print(frame[['name', 'age']])
13
```

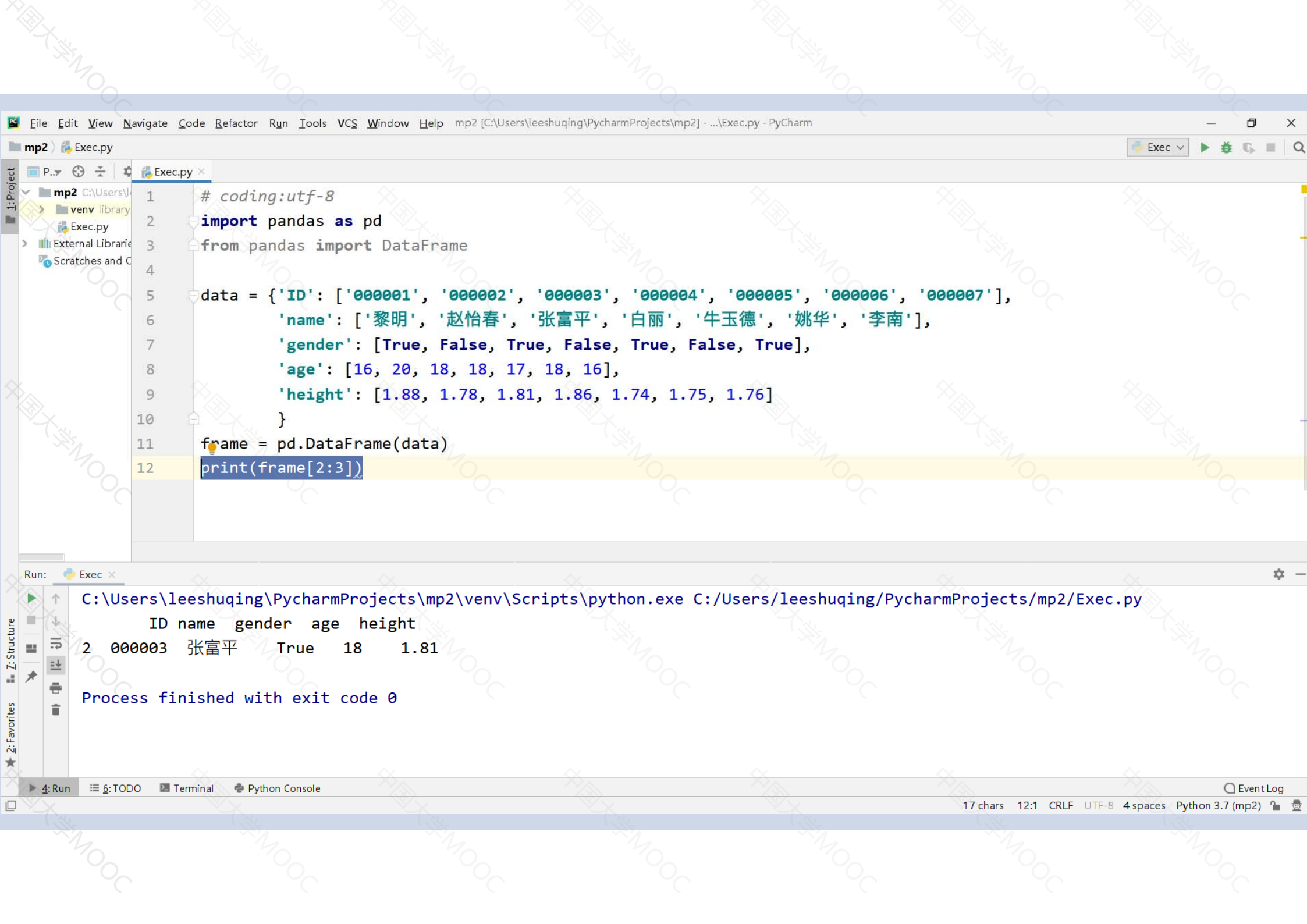
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py

	name	age
0	黎明	16
1	赵怡春	20
2	张富平	18
3	白丽	18
4	牛玉德	17
5	姚华	18
6	李南	16









```
1 # coding:utf-8
2 import pandas as pd
3 from pandas import DataFrame
4
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7         'gender': [True, False, True, False, True, False, True],
8         'age': [16, 20, 18, 18, 17, 18, 16],
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}
10
11 frame = pd.DataFrame(data)
12 print(frame[2:3])
```

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

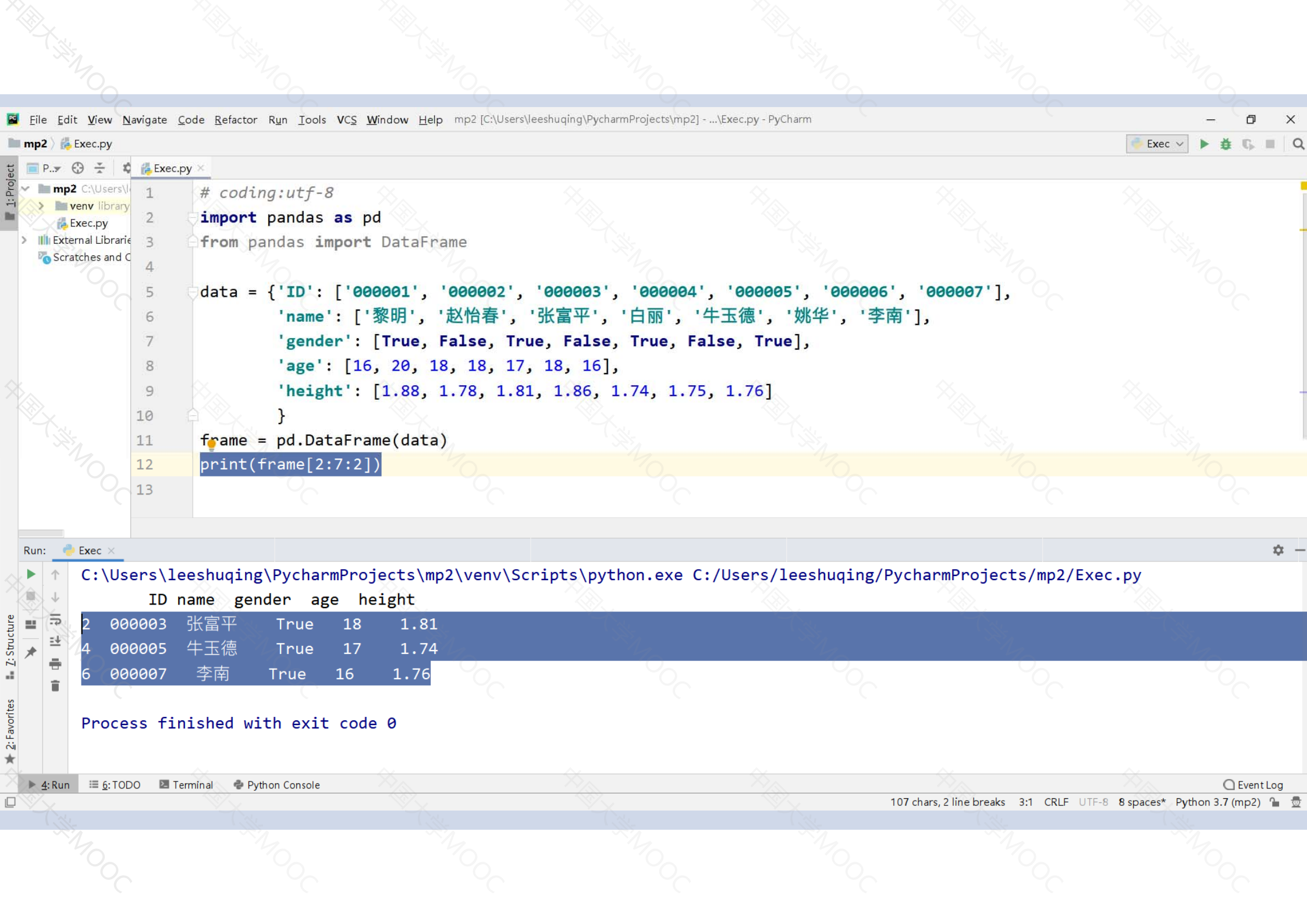
	ID	name	gender	age	height
2	000003	张富平	True	18	1.81

Process finished with exit code 0

4: Run | 6: TODO | Terminal | Python Console

Event Log

17 chars 12:1 CRLF UTF-8 4 spaces Python 3.7 (mp2)

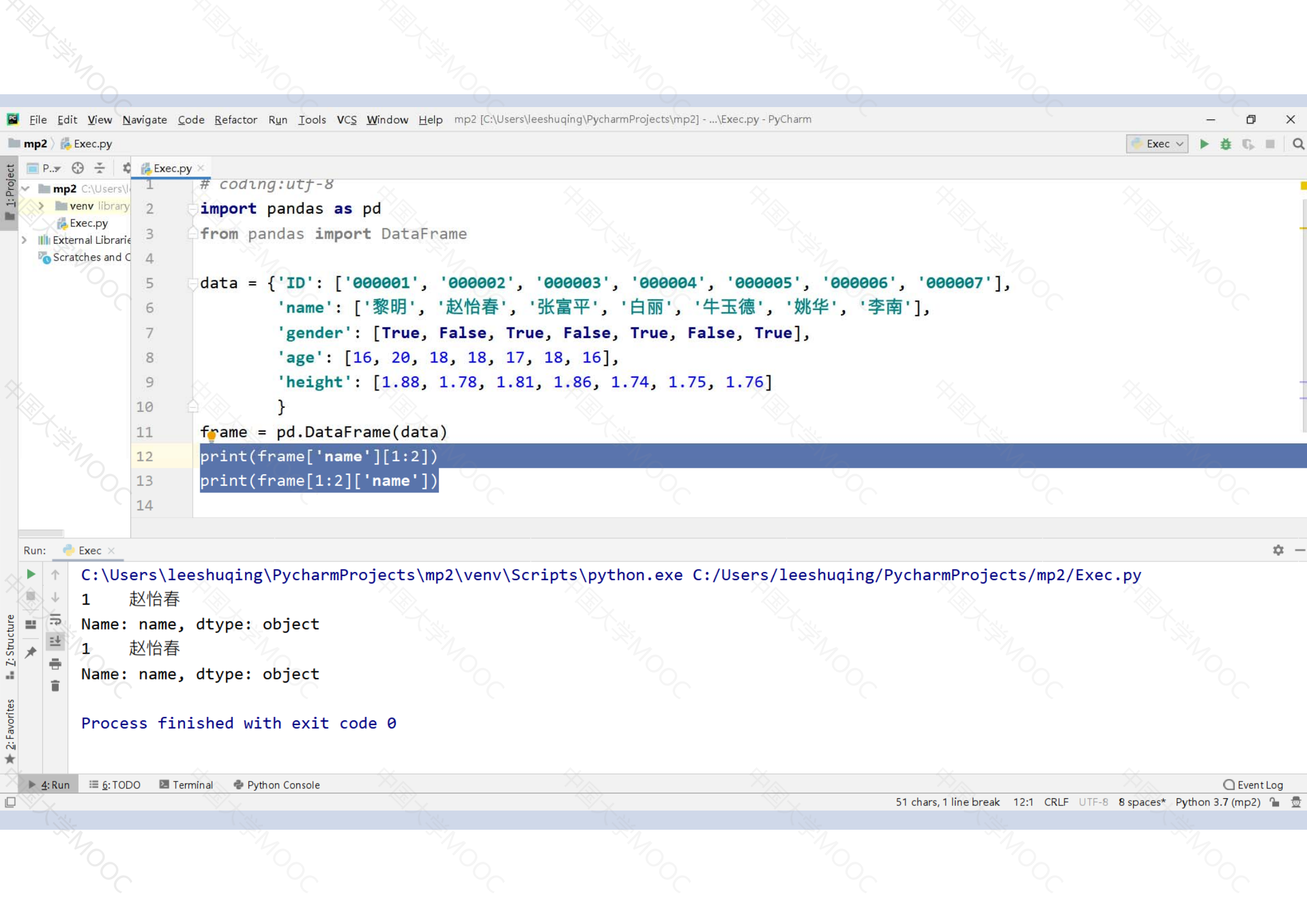


```
1 # coding:utf-8
2 import pandas as pd
3 from pandas import DataFrame
4
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7         'gender': [True, False, True, False, True, False, True],
8         'age': [16, 20, 18, 18, 17, 18, 16],
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}
10
11 frame = pd.DataFrame(data)
12 print(frame[2:7:2])
13
```

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

	ID	name	gender	age	height
2	000003	张富平	True	18	1.81
4	000005	牛玉德	True	17	1.74
6	000007	李南	True	16	1.76

Process finished with exit code 0



```
# coding:utf-8
```

```
import pandas as pd
```

```
from pandas import DataFrame
```

```
data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],  
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],  
        'gender': [True, False, True, False, True, False, True],  
        'age': [16, 20, 18, 18, 17, 18, 16],  
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]  
    }
```

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

```
print(frame['name'][1:2])
```

```
print(frame[1:2]['name'])
```

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

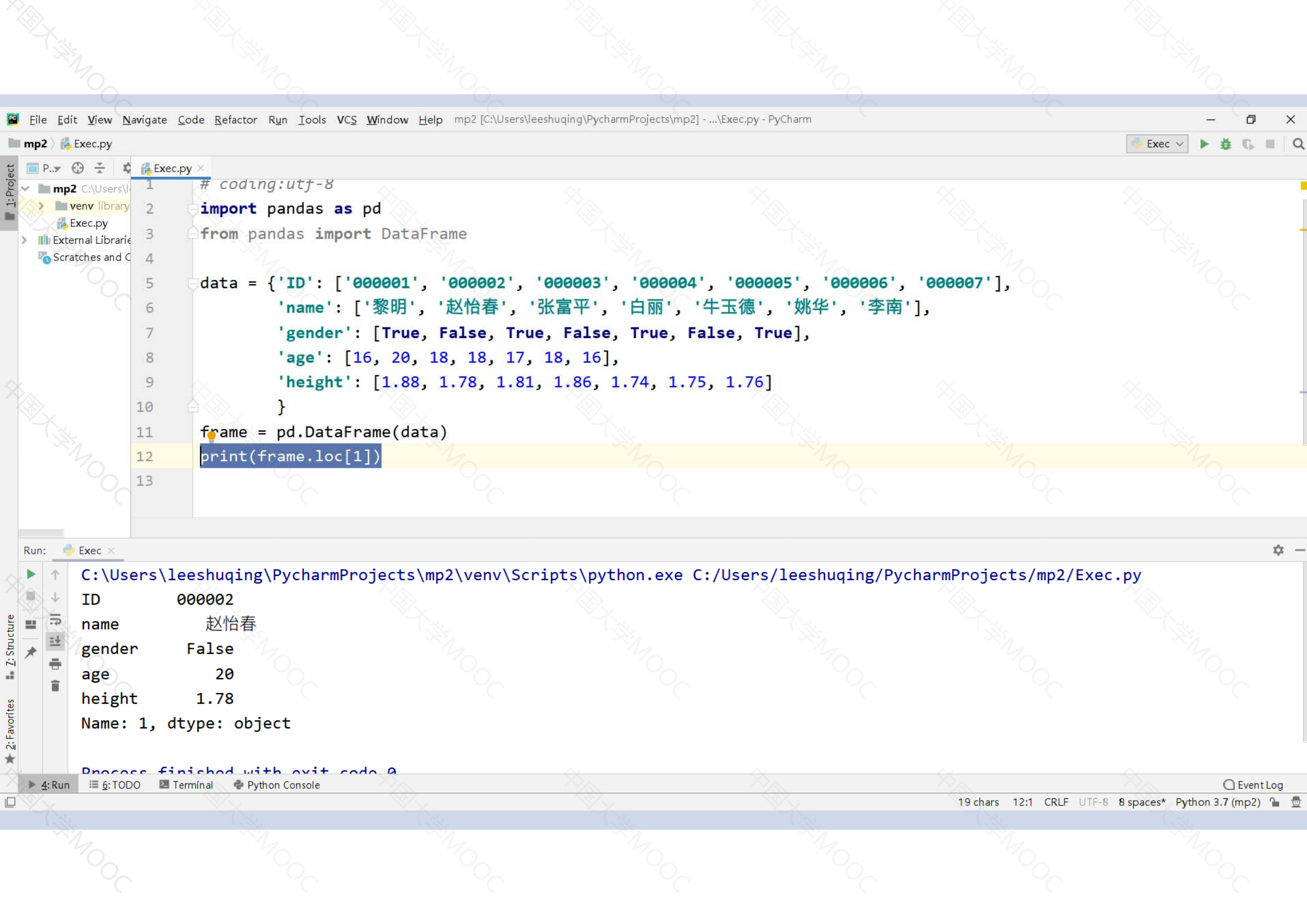
1 赵怡春

Name: name, dtype: object

1 赵怡春

Name: name, dtype: object

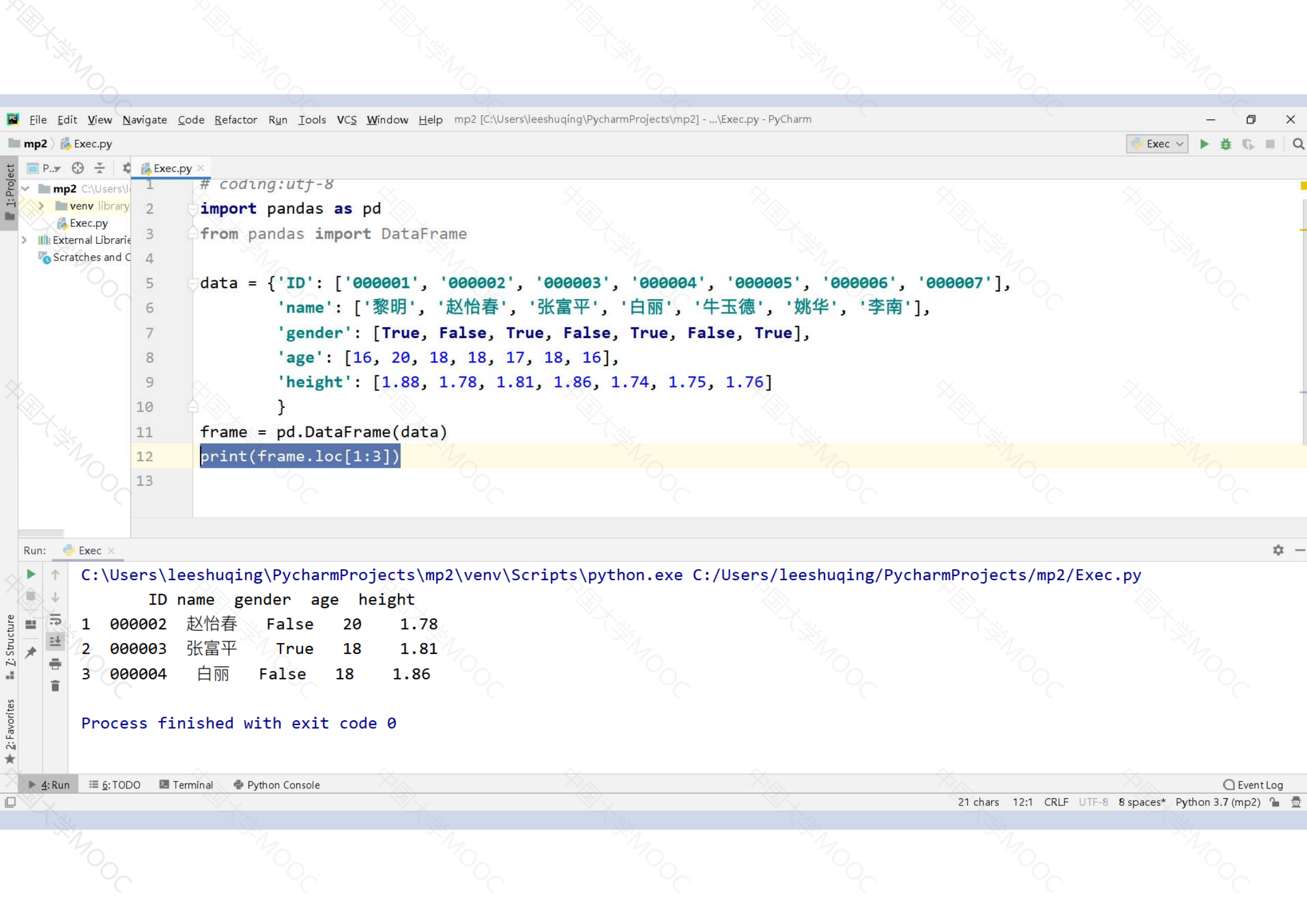
Process finished with exit code 0

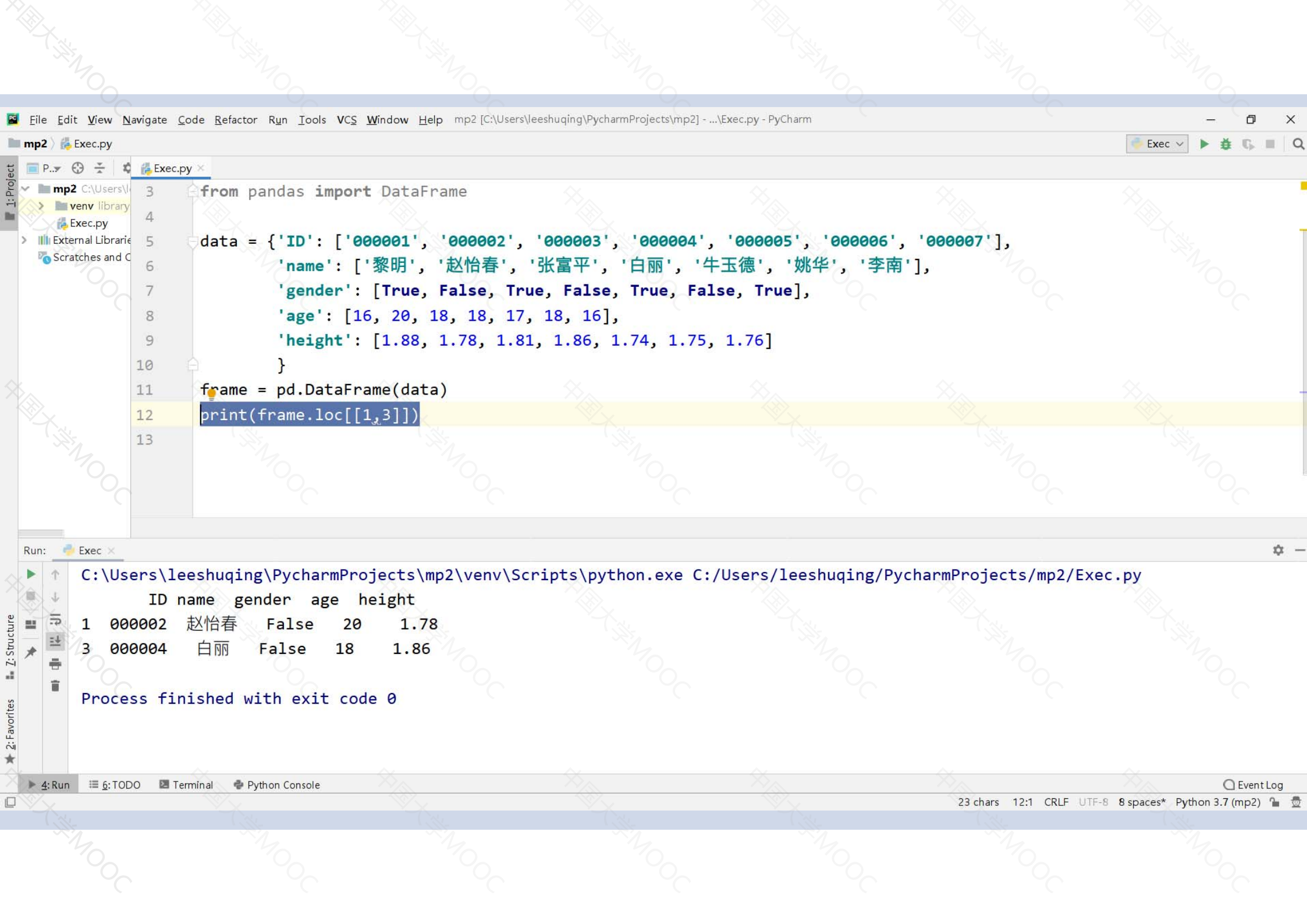


```
1 # coding:utf-8
2 import pandas as pd
3 from pandas import DataFrame
4
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7         'gender': [True, False, True, False, True, False, True],
8         'age': [16, 20, 18, 18, 17, 18, 16],
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
10        }
11 frame = pd.DataFrame(data)
12 print(frame.loc[1])
13
```

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
ID          000002
name        赵怡春
gender      False
age         20
height      1.78
Name: 1, dtype: object

Process finished with exit code 0
```

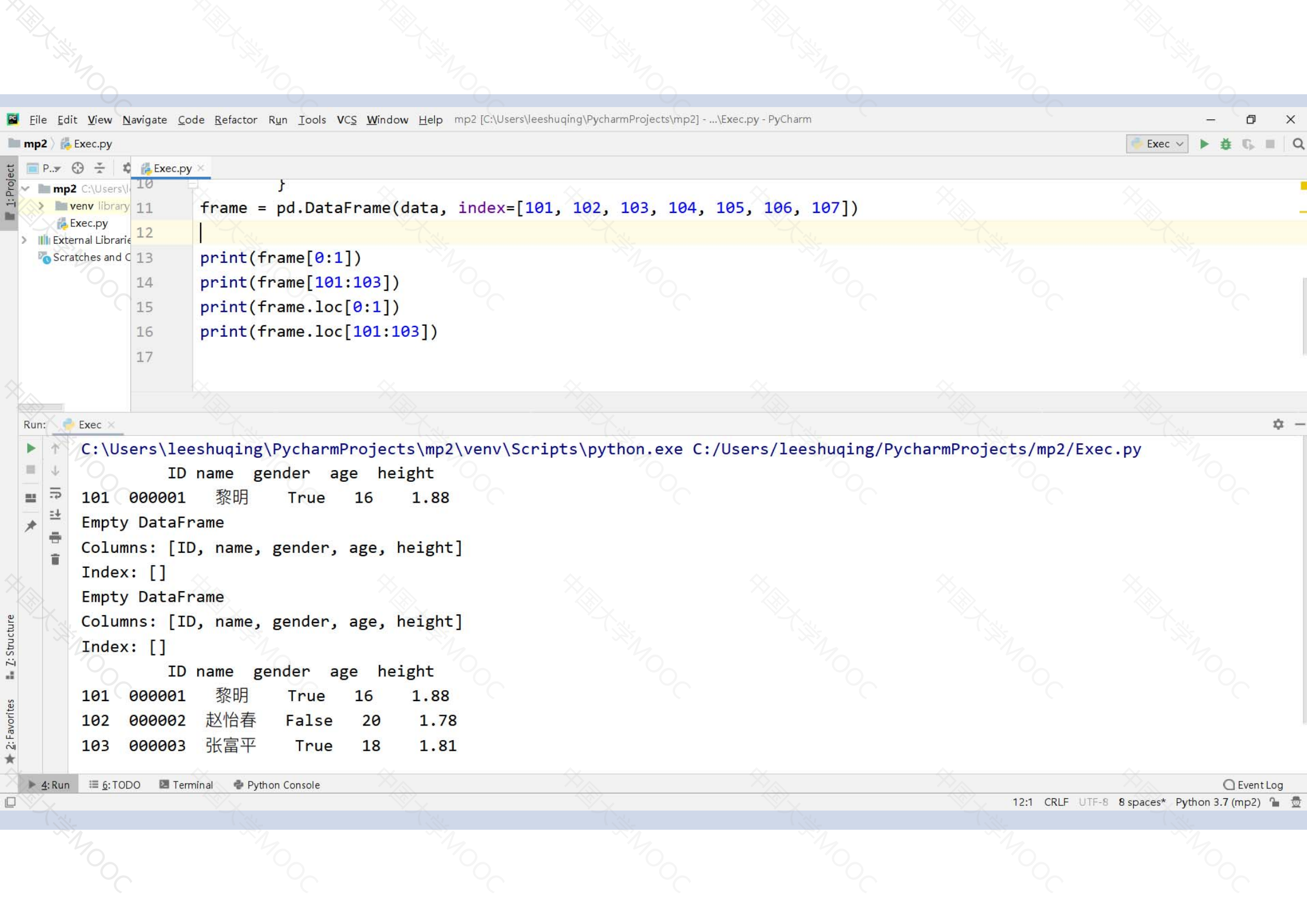


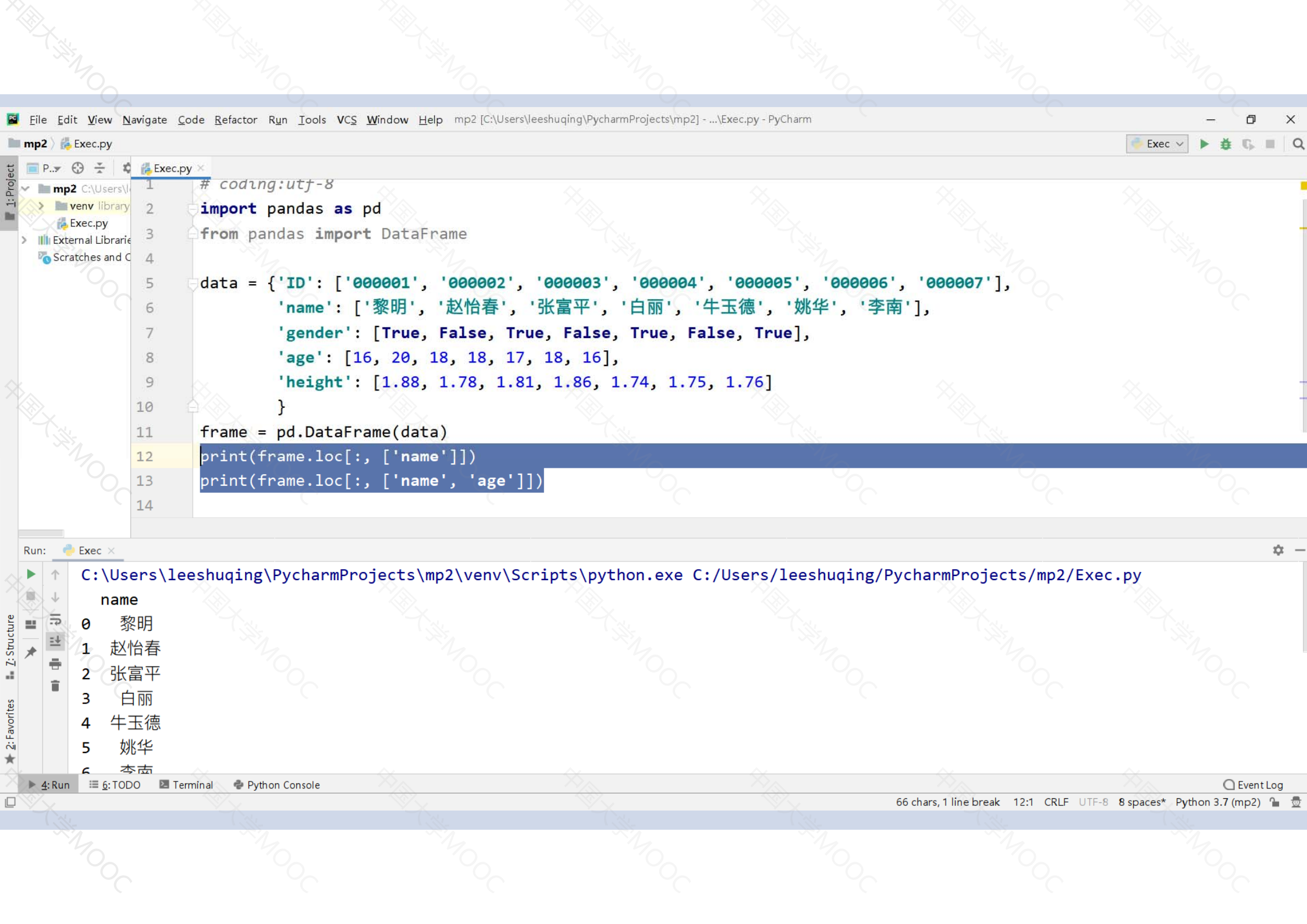


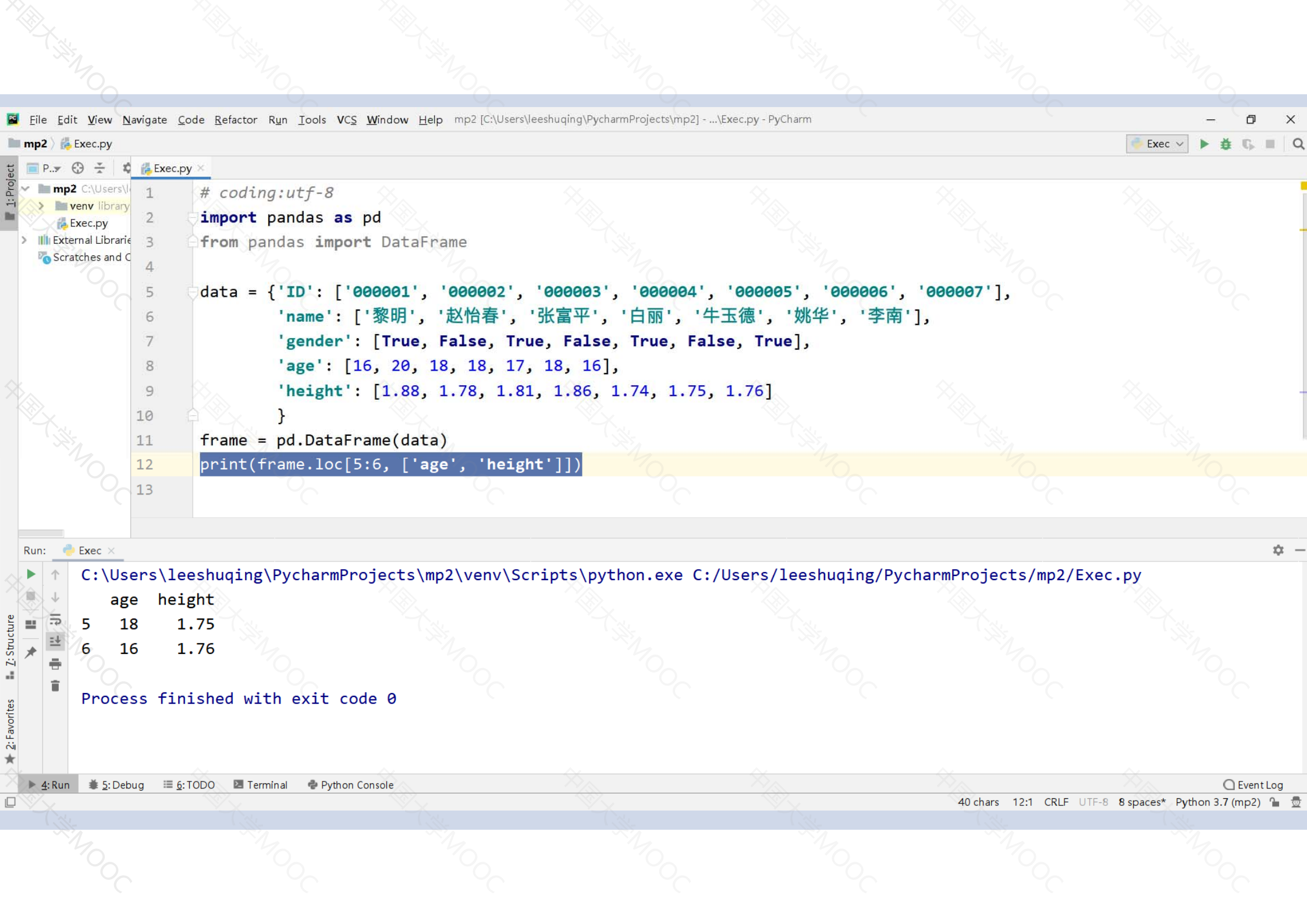
```
3 from pandas import DataFrame
4
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7         'gender': [True, False, True, False, True, False, True],
8         'age': [16, 20, 18, 18, 17, 18, 16],
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
10      }
11 frame = pd.DataFrame(data)
12 print(frame.loc[[1,3]])
13
```

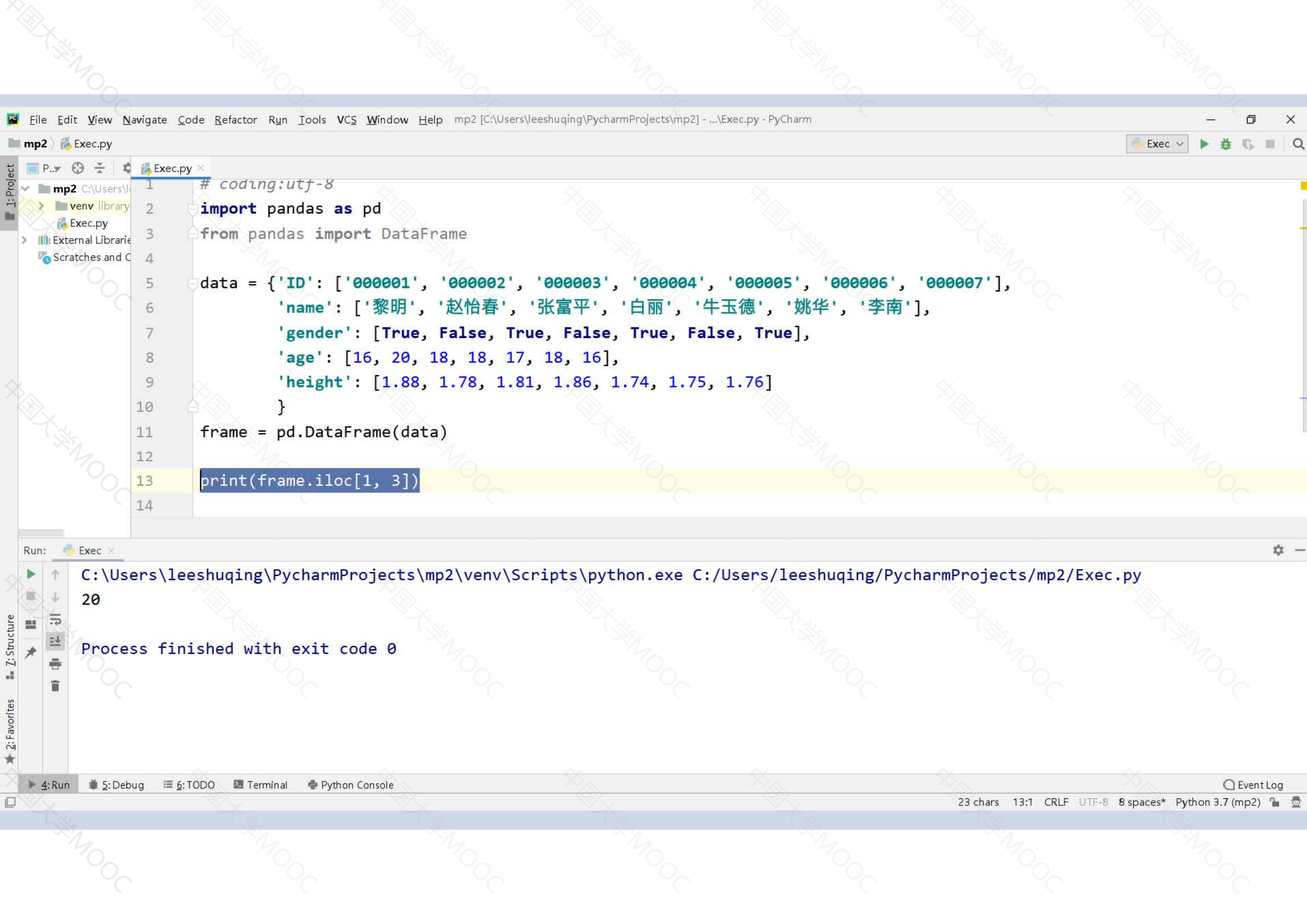
```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
   ID name  gender  age  height
1  000002  赵怡春  False   20    1.78
3  000004   白丽  False   18    1.86

Process finished with exit code 0
```







```
# coding:utf-8
```

```
import pandas as pd
```

```
from pandas import DataFrame
```

```
data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],  
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],  
        'gender': [True, False, True, False, True, False, True],  
        'age': [16, 20, 18, 18, 17, 18, 16],  
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]  
    }
```

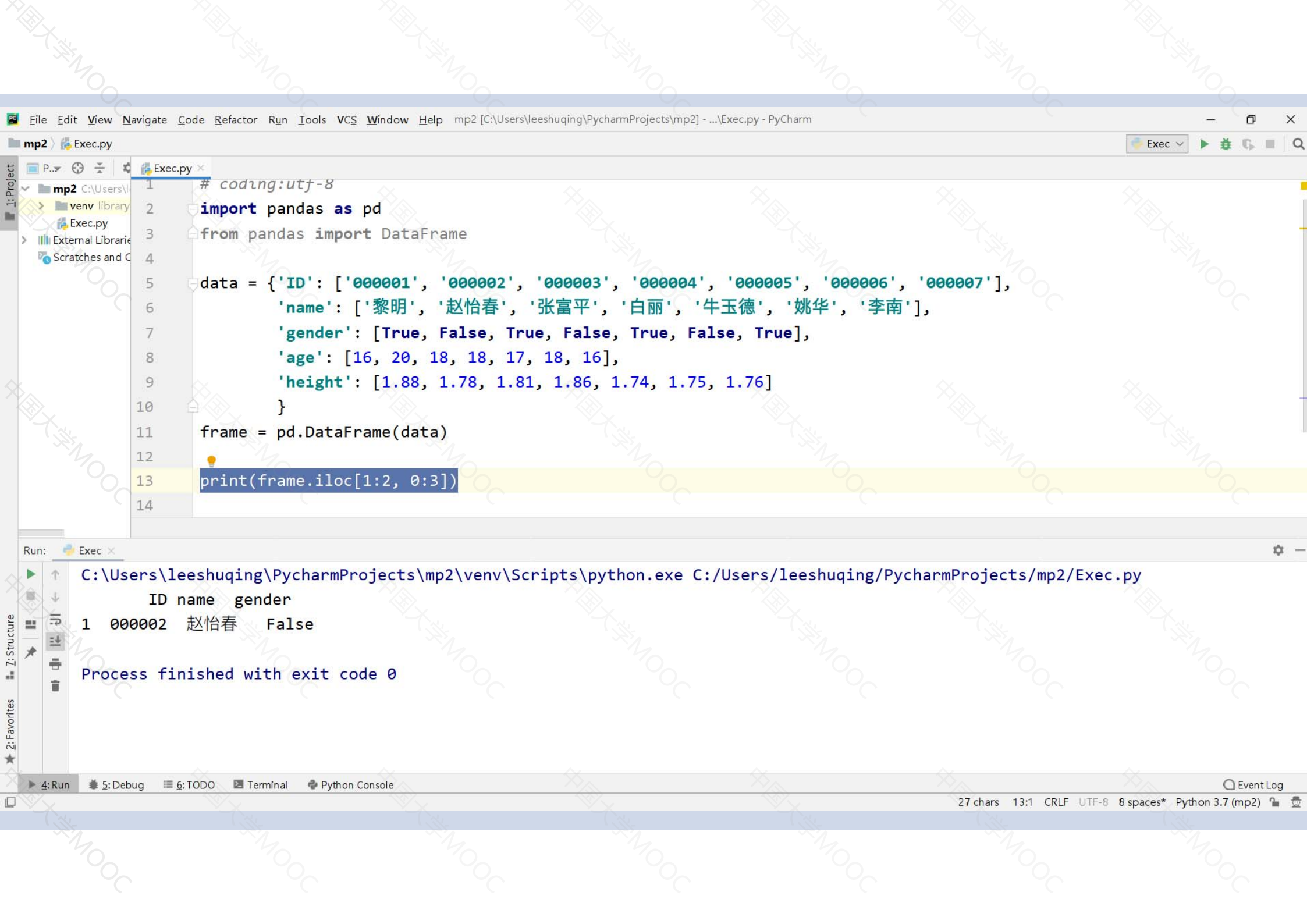
```
frame = pd.DataFrame(data)
```

```
print(frame.iloc[1, 3])
```

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

```
20
```

```
Process finished with exit code 0
```

```
# coding:utf-8
```

```
import pandas as pd
```

```
from pandas import DataFrame
```

```
data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],  
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],  
        'gender': [True, False, True, False, True, False, True],  
        'age': [16, 20, 18, 18, 17, 18, 16],  
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]  
    }
```

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

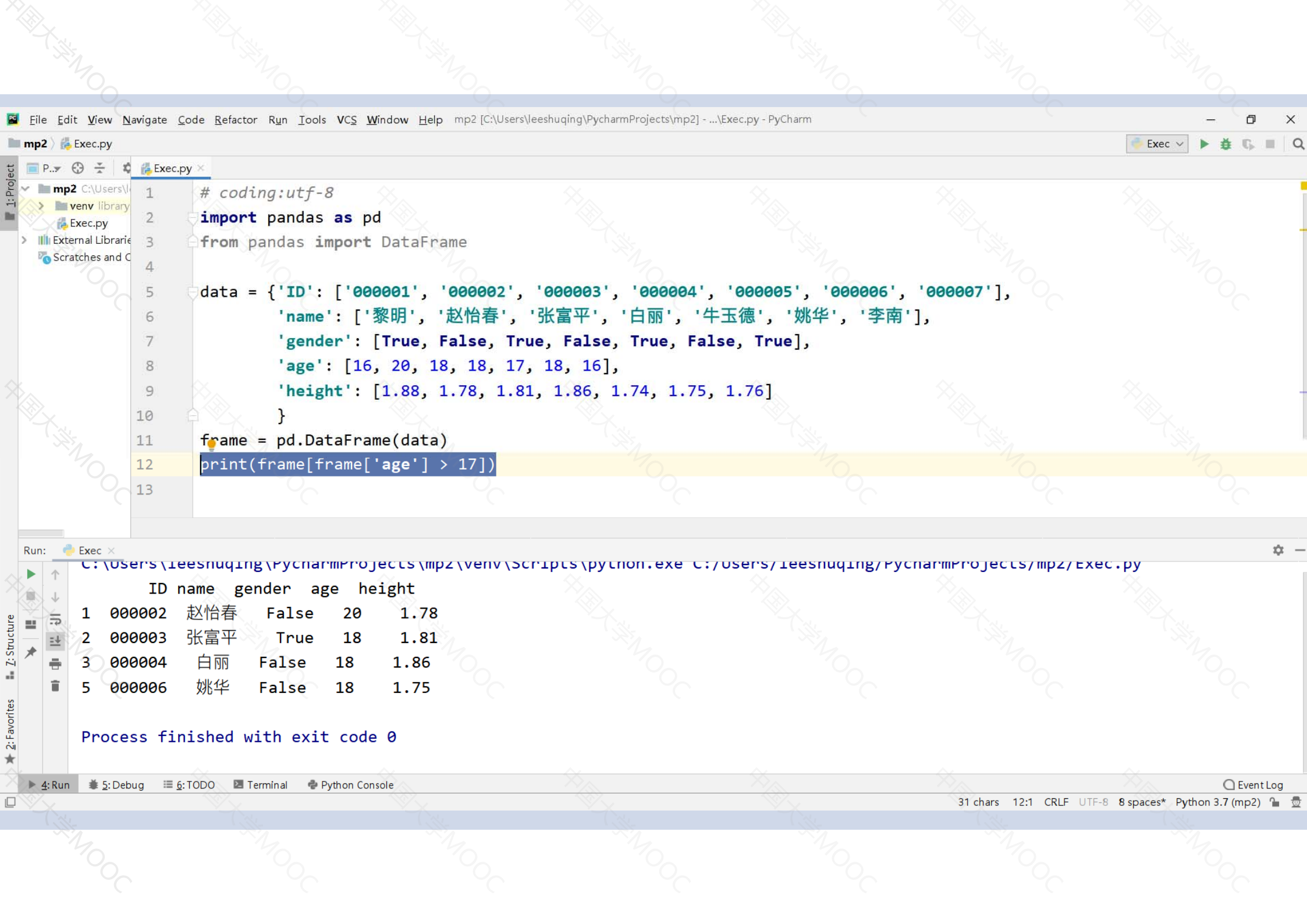
```
print(frame.iloc[1:2, 0:3])
```

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

```
ID name gender
```

```
1 000002 赵怡春 False
```

Process finished with exit code 0



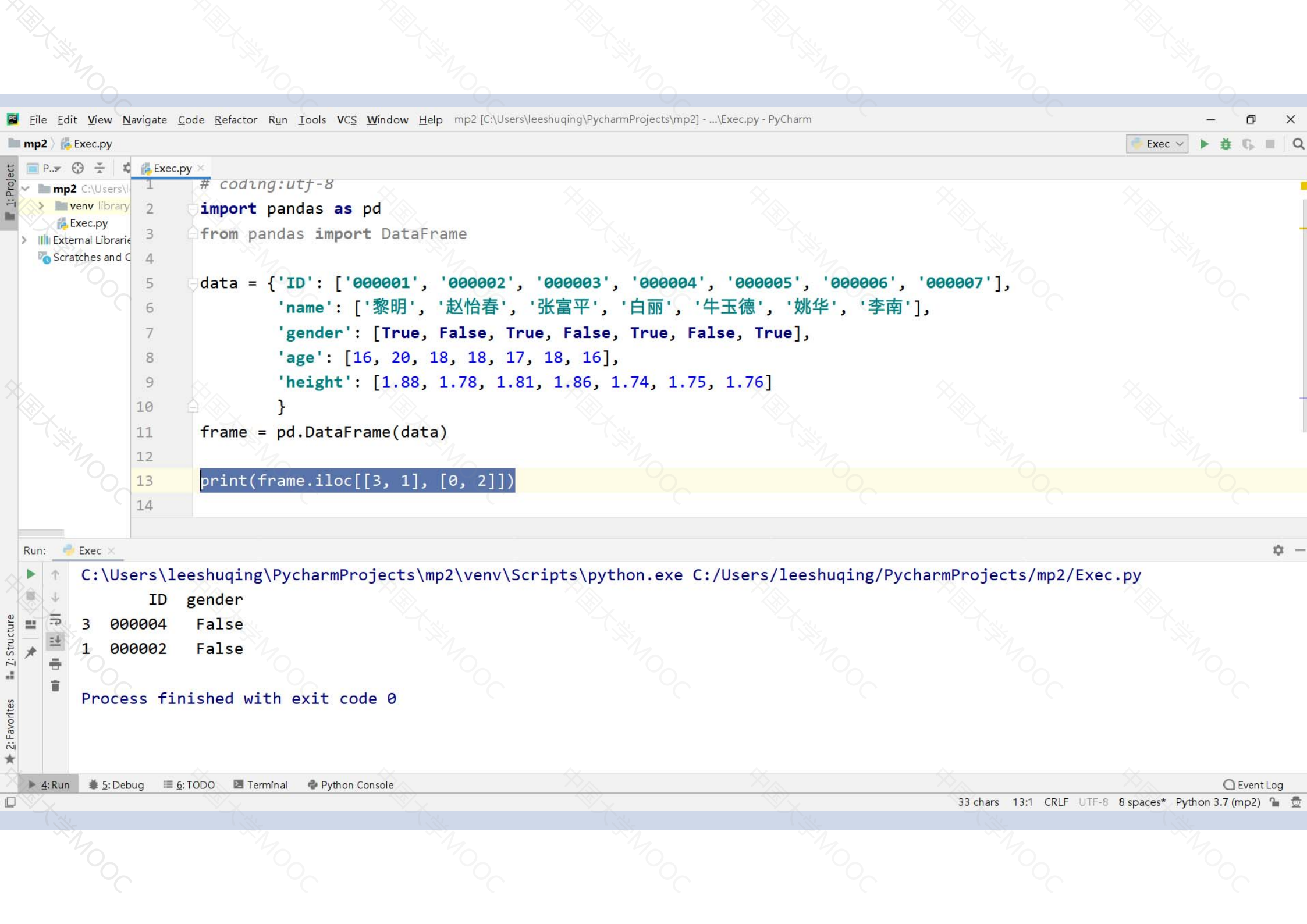
```
# coding:utf-8
import pandas as pd
from pandas import DataFrame

data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
        'gender': [True, False, True, False, True, False, True],
        'age': [16, 20, 18, 18, 17, 18, 16],
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}

frame = pd.DataFrame(data)
print(frame[frame['age'] > 17])
```

	ID	name	gender	age	height
1	000002	赵怡春	False	20	1.78
2	000003	张富平	True	18	1.81
3	000004	白丽	False	18	1.86
5	000006	姚华	False	18	1.75

Process finished with exit code 0



```

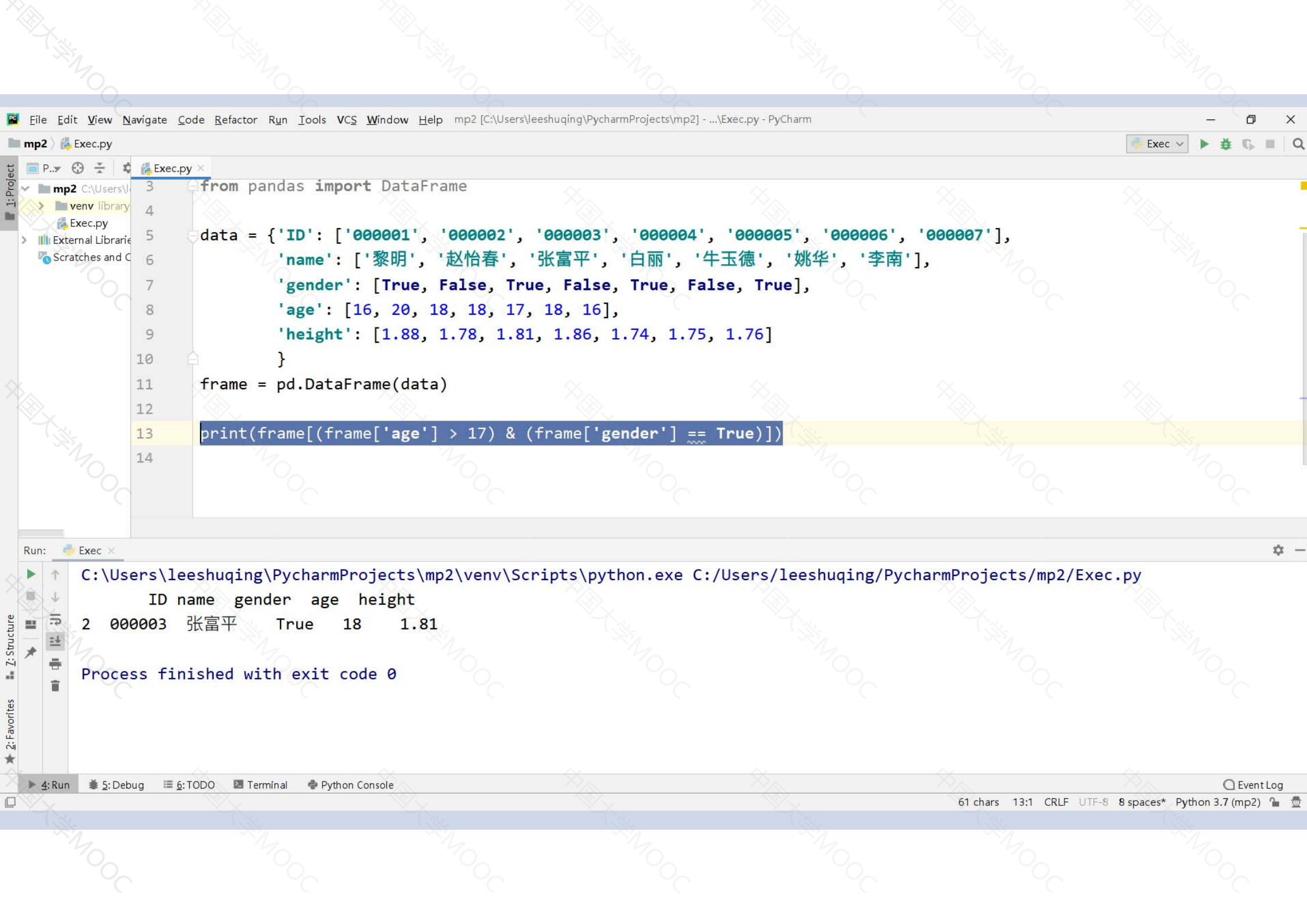
3 import numpy as np
4 from pandas import DataFrame
5
6 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
7         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
8         'gender': [True, False, True, False, True, False, True],
9         'age': [16, 20, 18, 18, 17, 18, 16],
10        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
11        }
12 frame = pd.DataFrame(data)
13 print(frame[(frame['age'] > 17) & (frame['height'] > 1.80)])
14

```

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

	ID	name	gender	age	height
2	000003	张富平	True	18	1.81
3	000004	白丽	False	18	1.86

Process finished with exit code 0



```
3 from pandas import DataFrame
```

```
4  
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],  
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],  
7         'gender': [True, False, True, False, True, False, True],  
8         'age': [16, 20, 18, 18, 17, 18, 16],  
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]  
10      }
```

```
11 frame = pd.DataFrame(data)
```

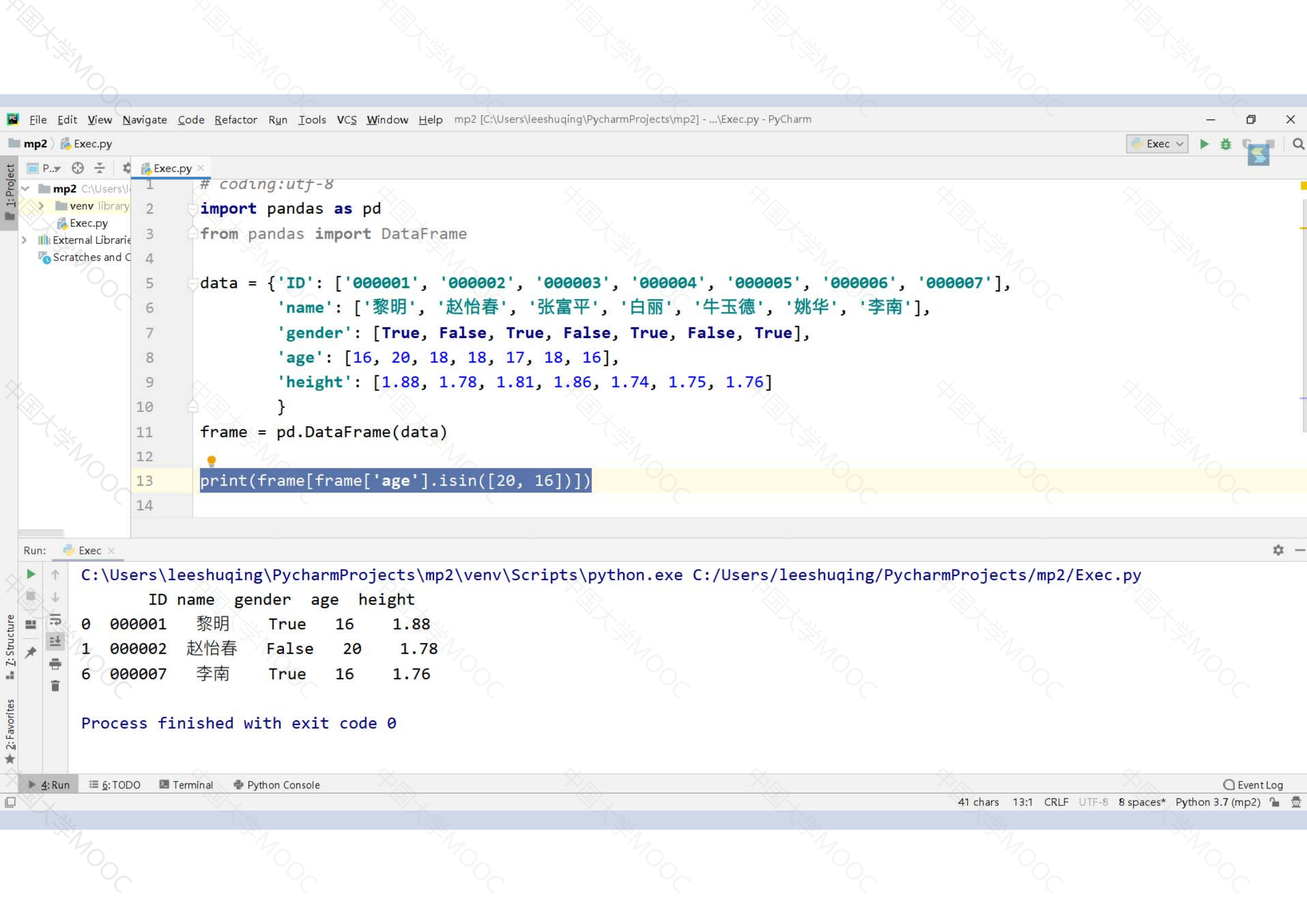
```
12  
13 print(frame[(frame['age'] > 17) & (frame['gender'] == True)])  
14
```

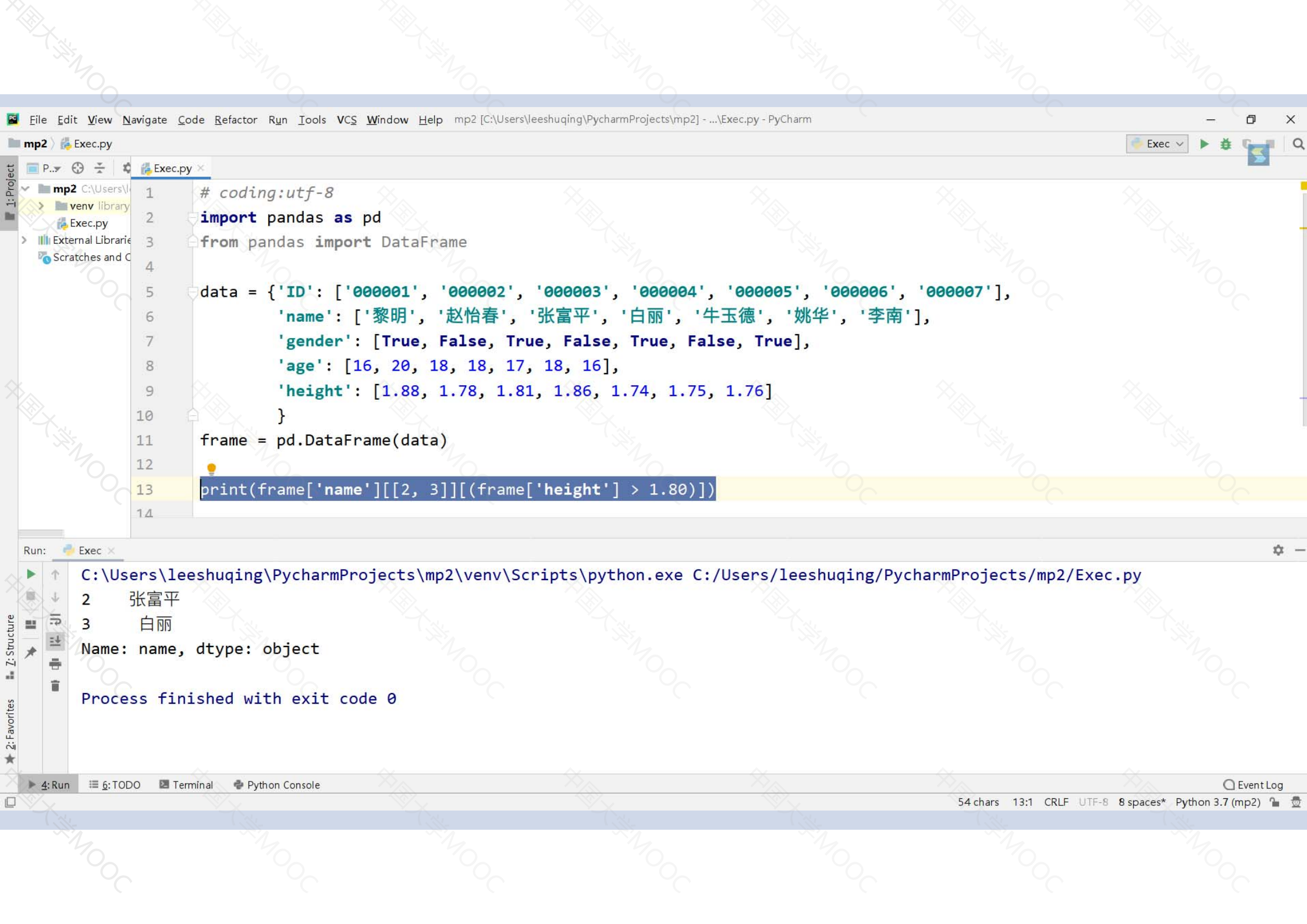
```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
```

```
ID name gender age height
```

```
2 000003 张富平 True 18 1.81
```

```
Process finished with exit code 0
```



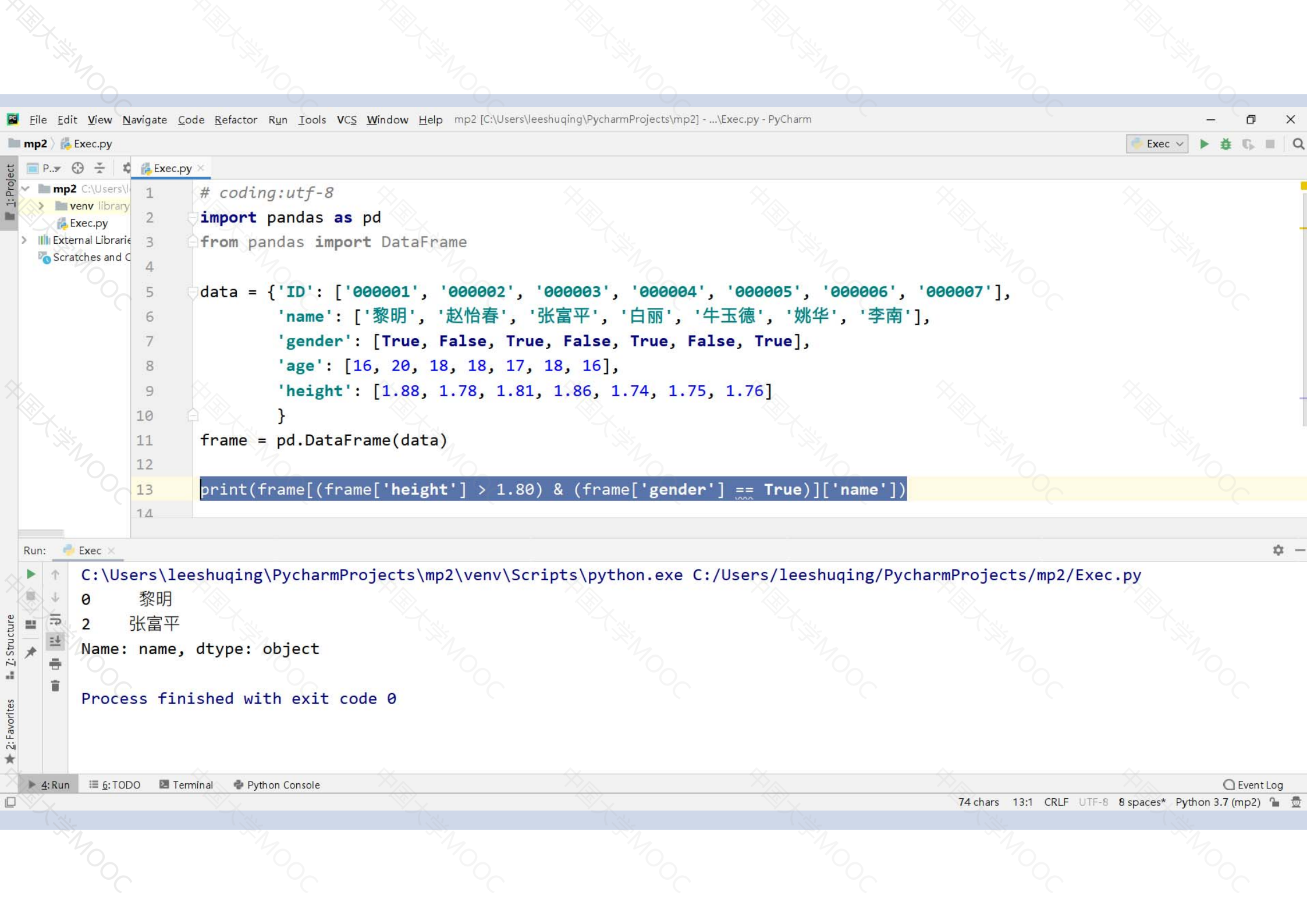
```
1 # coding:utf-8
2 import pandas as pd
3 from pandas import DataFrame
4
5 data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6         'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7         'gender': [True, False, True, False, True, False, True],
8         'age': [16, 20, 18, 18, 17, 18, 16],
9         'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}
10
11 frame = pd.DataFrame(data)
12
13 print(frame['name'][(frame['height'] > 1.80)])
```

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
2 张富平
3 白丽
Name: name, dtype: object

Process finished with exit code 0
```

练习

- 查询身高大于1.80的男生姓名



```
# coding:utf-8
import pandas as pd
from pandas import DataFrame

data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
        'gender': [True, False, True, False, True, False, True],
        'age': [16, 20, 18, 18, 17, 18, 16],
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]}

frame = pd.DataFrame(data)

print(frame[(frame['height'] > 1.80) & (frame['gender'] == True)][ 'name'])
```

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

0 黎明

2 张富平

Name: name, dtype: object

Process finished with exit code 0

一次不学多，下次再学