Python大数据分析

四、数据的更改

File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm

mp2 > Exec.py   Exec ▼ ▶ ✿ ⟳ ■ 🔍
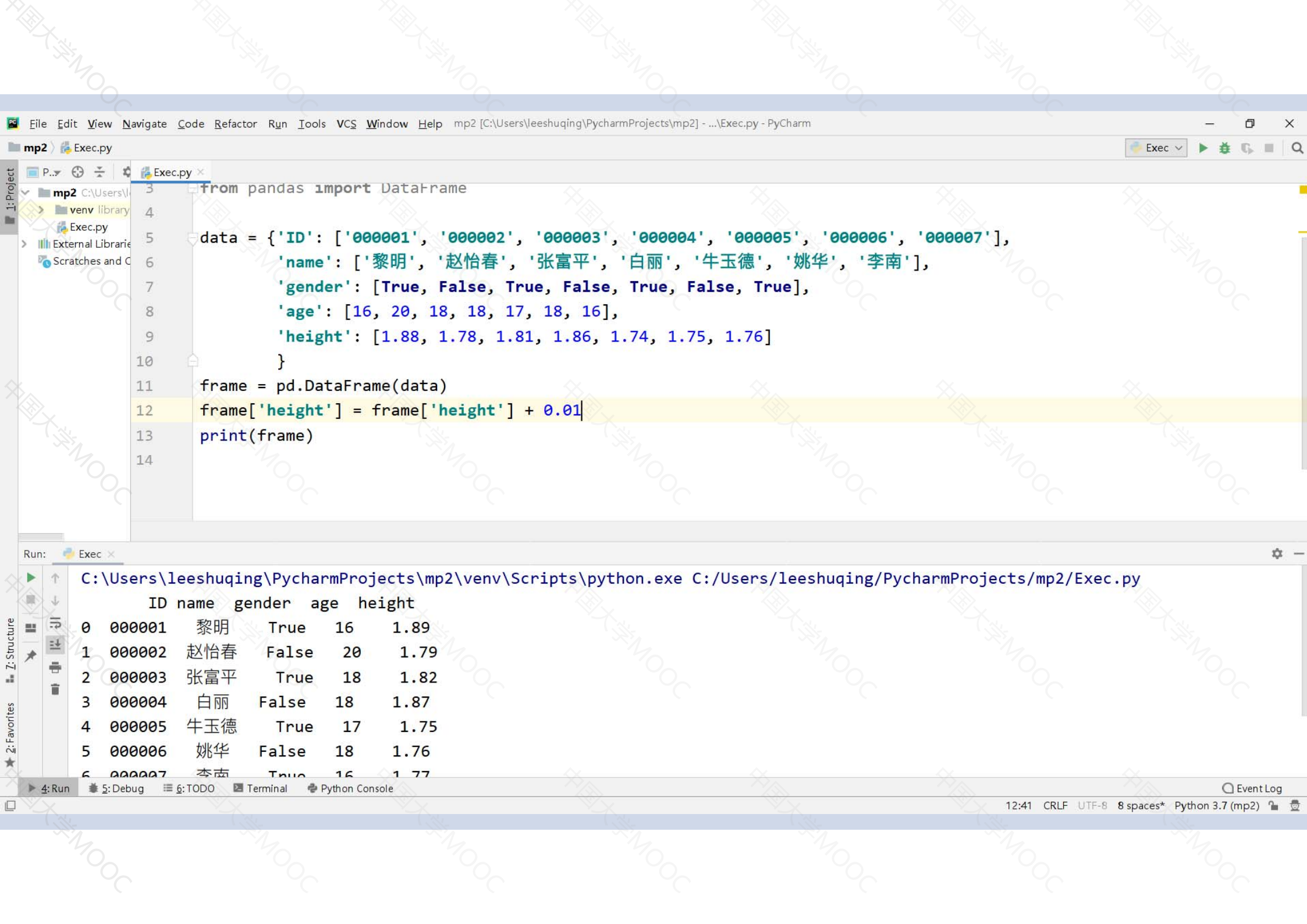
P..▼ ⊕ ↓↑ ✿   Exec.py ×

```python
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   frame['height'] = frame['height'] + 0.01
13   print(frame)
14
```
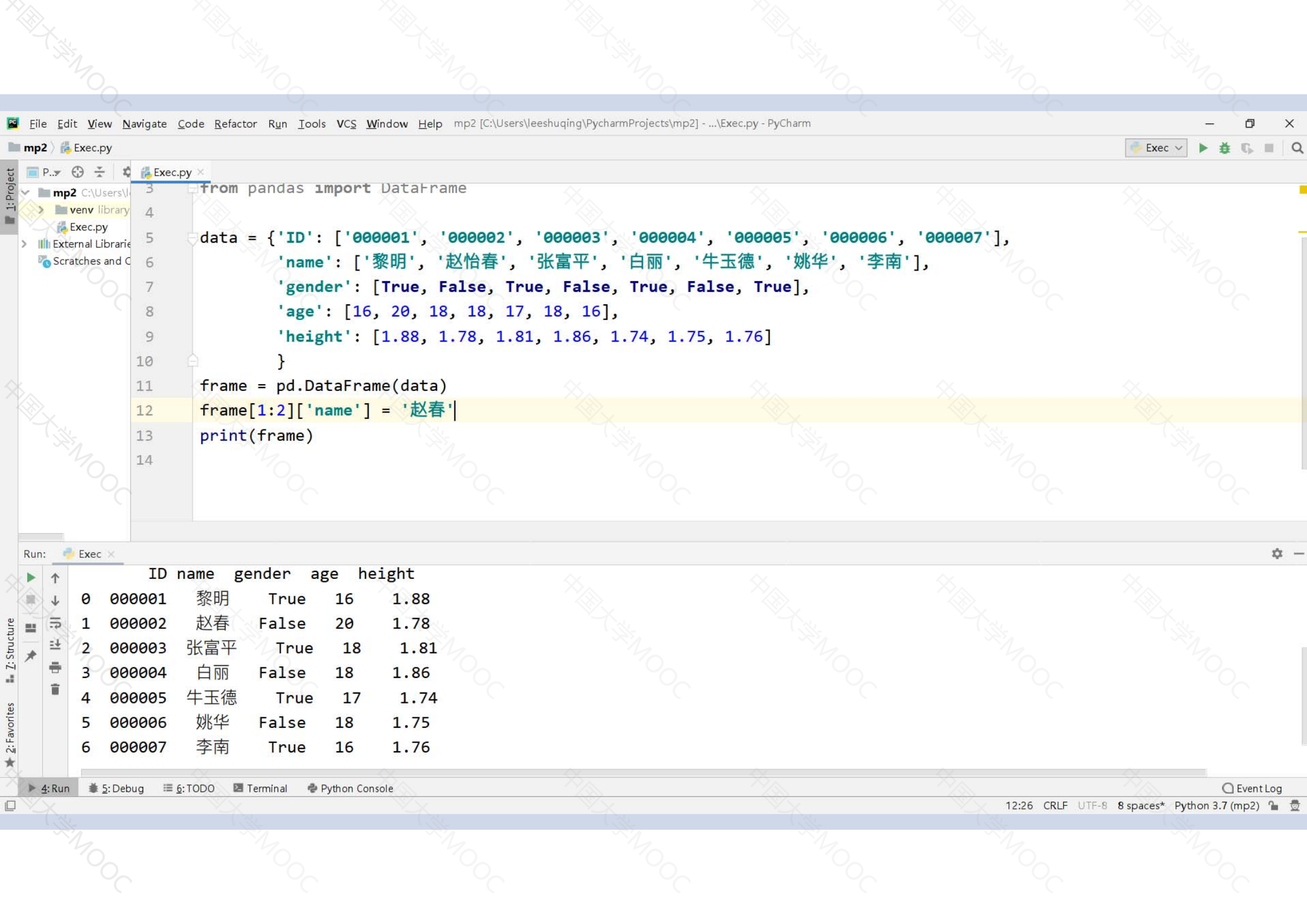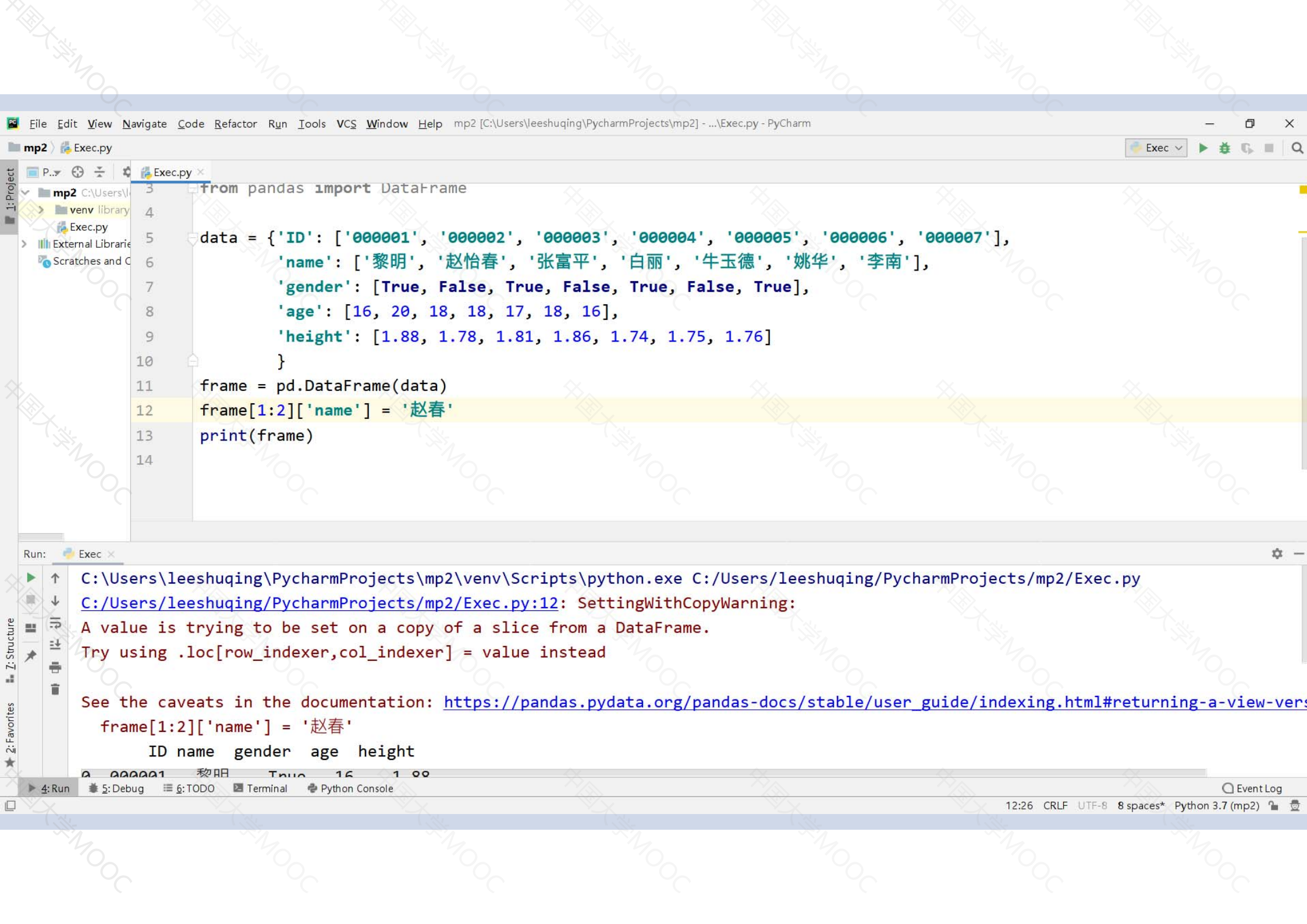
mp2 C:\Users\l
venv library
Exec.py
External Librarie
Scratches and C

Run: 🐍 Exec ×

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
         ID name  gender  age  height
0  000001   黎明    True   16    1.89
1  000002  赵怡春   False   20    1.79
2  000003  张富平    True   18    1.82
3  000004   白丽   False   18    1.87
4  000005  牛玉德    True   17    1.75
5  000006   姚华   False   18    1.76
6  000007   李南    True   16    1.77
```

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

Run: Exec
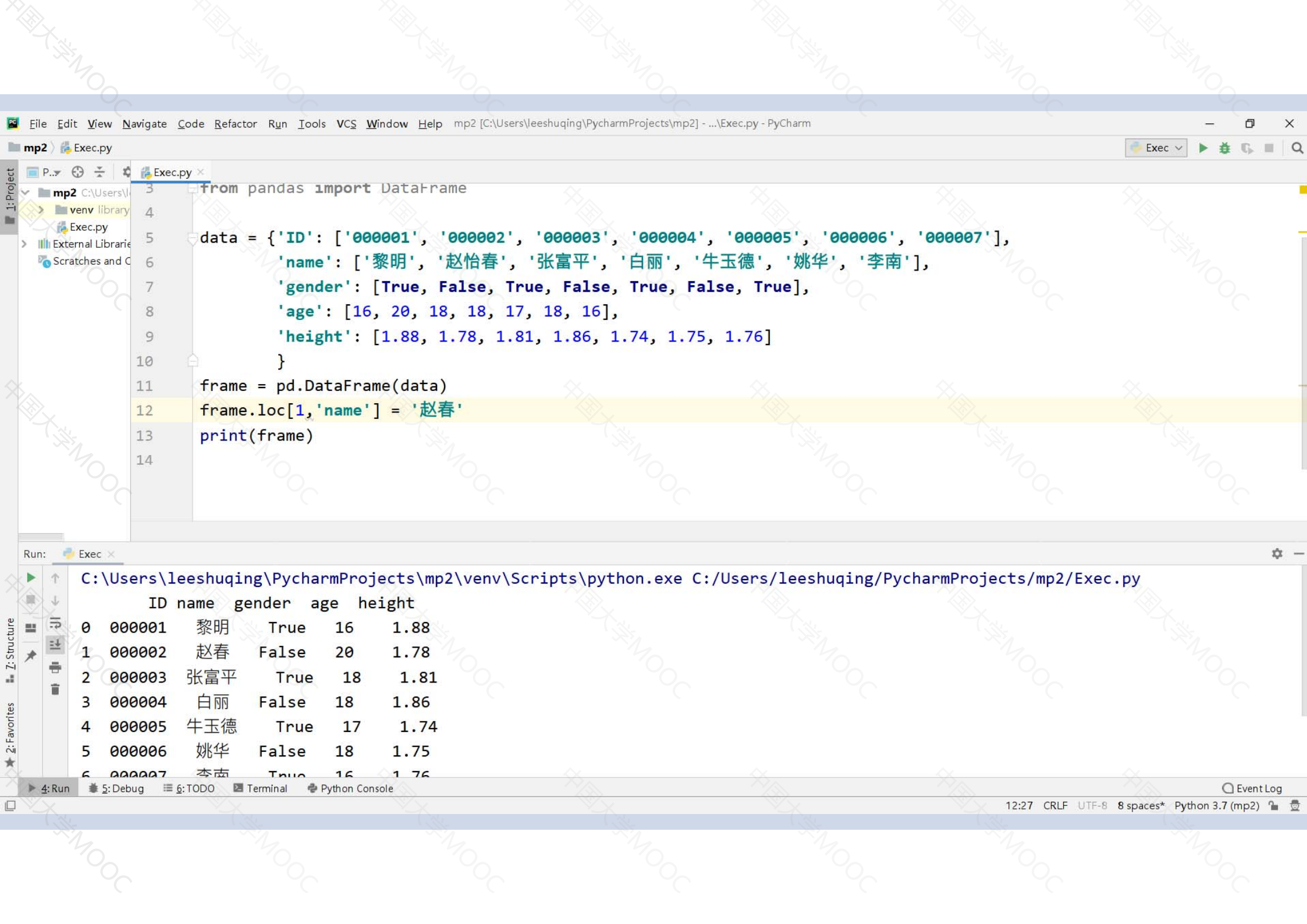
```
       ID name  gender  age  height
0  000001  黎明    True   16    1.88
1  000002  赵春   False   20    1.78
2  000003  张富平   True   18    1.81
3  000004  白丽   False   18    1.86
4  000005  牛玉德   True   17    1.74
5  000006  姚华   False   18    1.75
6  000007  李南    True   16    1.76
```

PC File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm   — ☐ ✕

mp2 ⟩ Exec.py   Exec ▾ ▶ 🐞 ⏵ ☐ 🔍

```
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      frame[1:2]['name'] = '赵春'
13      print(frame)
14
```

Run: 🐍 Exec ✕

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py:12: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-vers
  frame[1:2]['name'] = '赵春'
       ID name  gender  age   height
0  000001   黎明    True   16    1.88
```
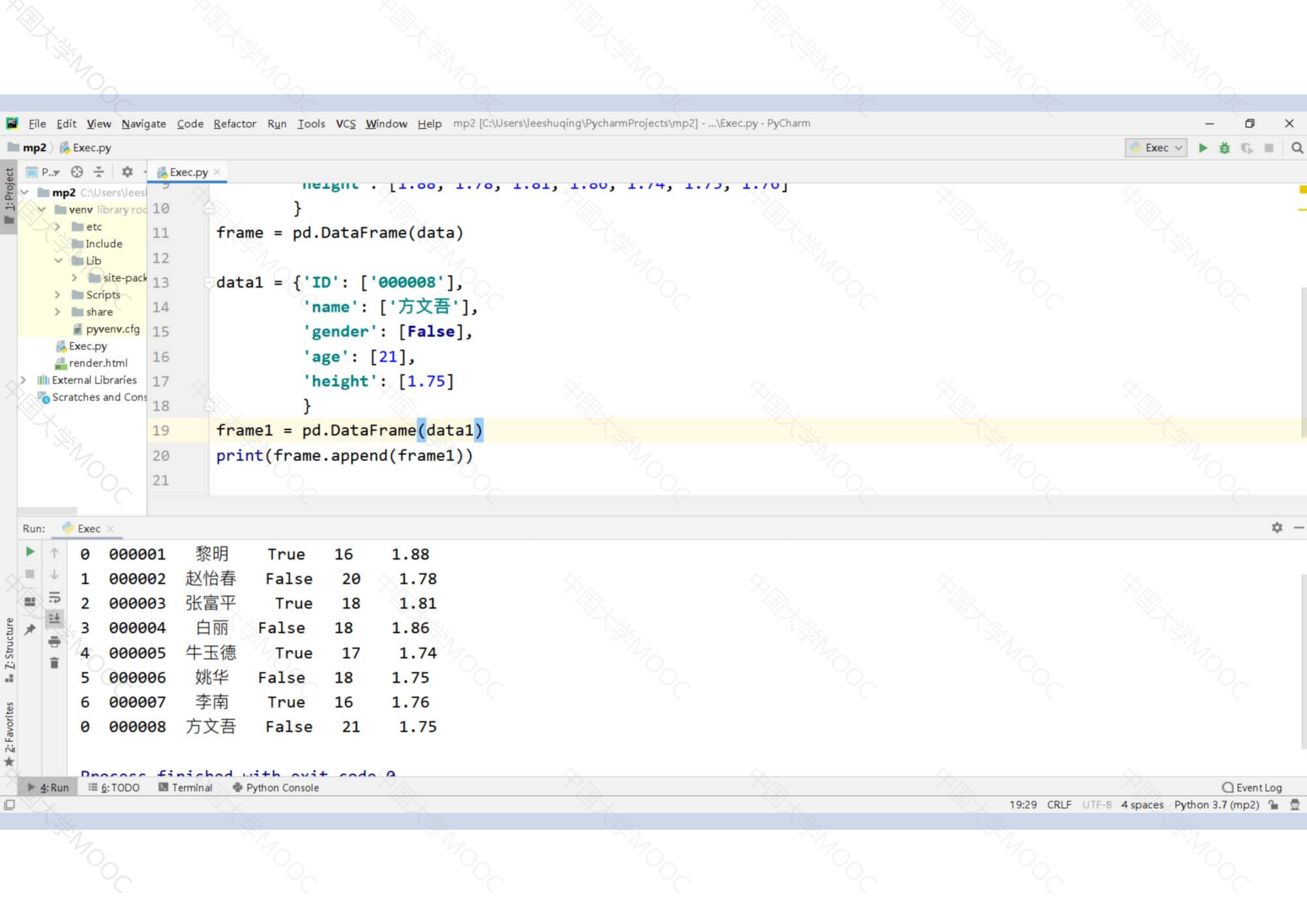
File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm

mp2 ▸ Exec.py   Exec ▾ ▶ ✿ ⟳ ■ ☌

```python
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)
frame.loc[1,'name'] = '赵春'
print(frame)
```

Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
       ID name  gender  age  height
0  000001   黎明    True   16    1.88
1  000002   赵春   False   20    1.78
2  000003  张富平    True   18    1.81
3  000004   白丽   False   18    1.86
4  000005  牛玉德    True   17    1.74
5  000006   姚华   False   18    1.75
6  000007   李南    True   16    1.76
```
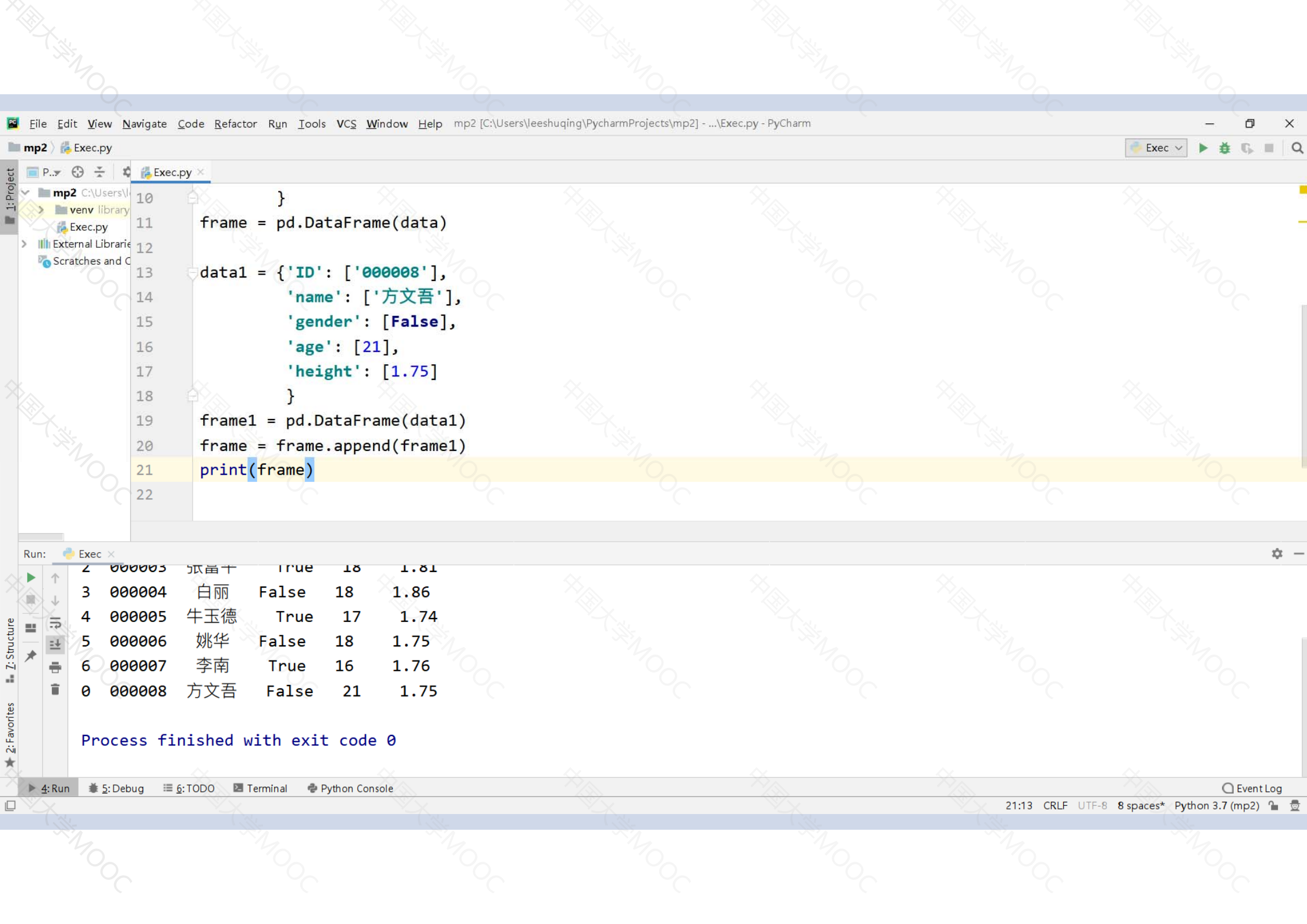
File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm   —   □   ✕

mp2 ⟩ Exec.py   Exec ∨   ▶ ⚐ ⟳ ■   Q

```
 9              height : [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
10          }
11     frame = pd.DataFrame(data)
12
13     data1 = {'ID': ['000008'],
14              'name': ['方文吾'],
15              'gender': [False],
16              'age': [21],
17              'height': [1.75]
18          }
19     frame1 = pd.DataFrame(data1)
20     print(frame.append(frame1))
21
```

Run: Exec ✕

```
0  000001   黎明     True    16     1.88
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
4  000005   牛玉德    True    17     1.74
5  000006   姚华     False    18     1.75
6  000007   李南     True    16     1.76
0  000008   方文吾   False    21     1.75


Process finished with exit code 0
```
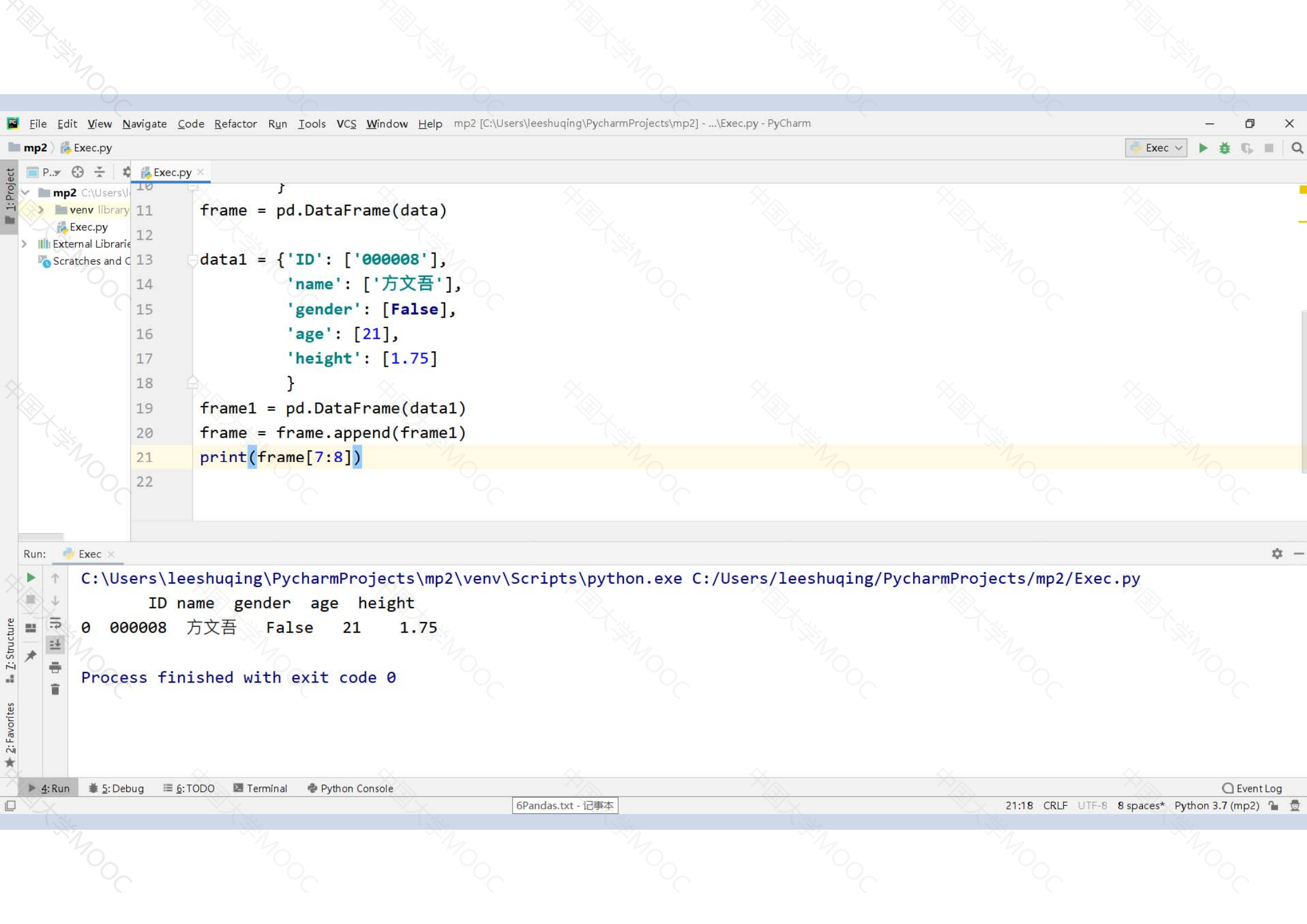
```python
          }
frame = pd.DataFrame(data)

data1 = {'ID': ['000008'],
         'name': ['方文吾'],
         'gender': [False],
         'age': [21],
         'height': [1.75]
         }
frame1 = pd.DataFrame(data1)
frame = frame.append(frame1)
print(frame)
```

```
2  000003  张富平   True   18   1.81
3  000004  白丽    False  18   1.86
4  000005  牛玉德   True   17   1.74
5  000006  姚华    False  18   1.75
6  000007  李南    True   16   1.76
0  000008  方文吾   False  21   1.75


Process finished with exit code 0
```

File Edit View Navigate Code Refactor Run Tools VCS Window Help mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm

mp2 > Exec.py   Exec ▼ ▶ ✿ ✿ ✿ ✿   Q

```python
10        }
11    frame = pd.DataFrame(data)
12
13    data1 = {'ID': ['000008'],
14             'name': ['方文吾'],
15             'gender': [False],
16             'age': [21],
17             'height': [1.75]
18             }
19    frame1 = pd.DataFrame(data1)
20    frame = frame.append(frame1)
21    print(frame[7:8])
22
```
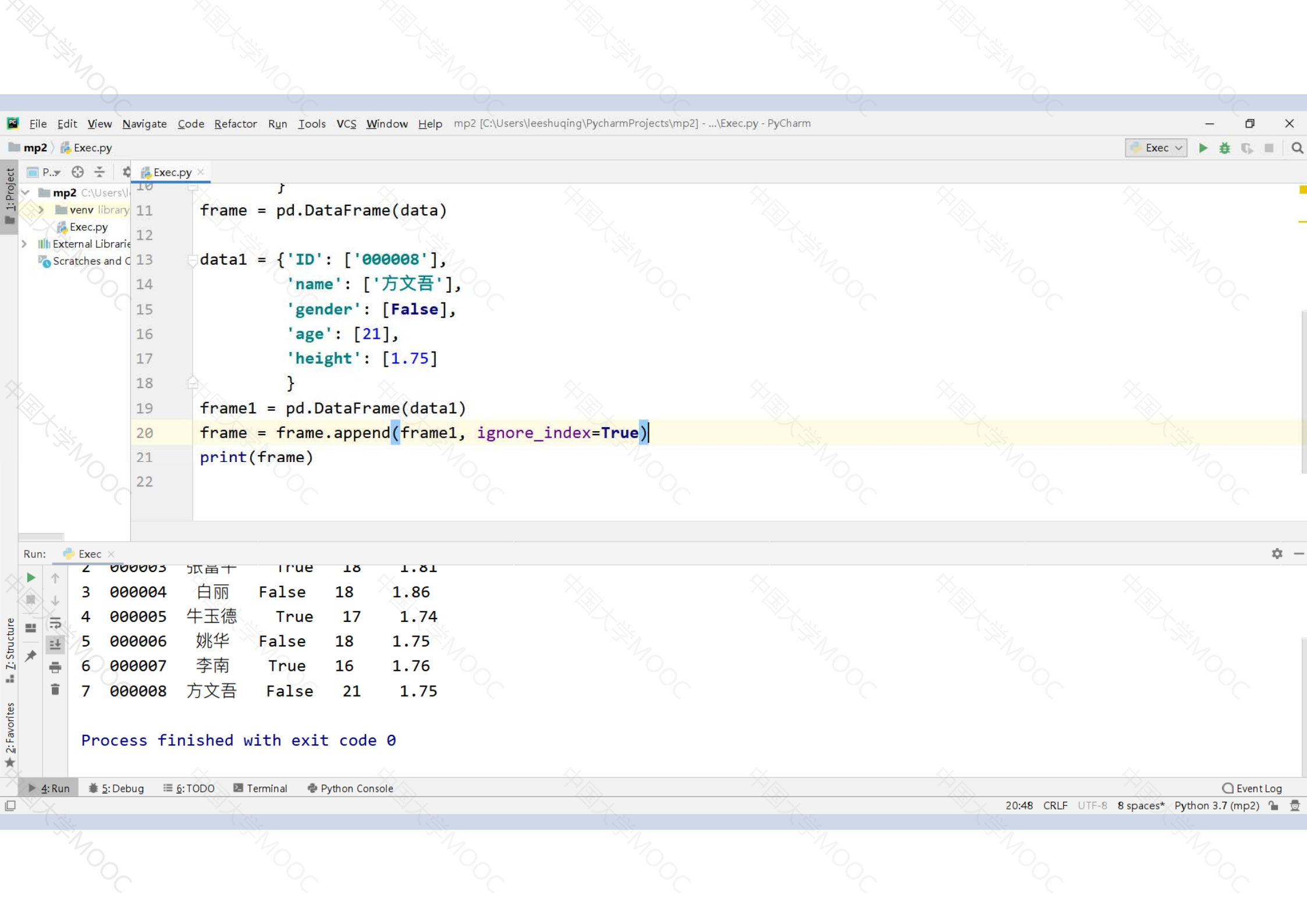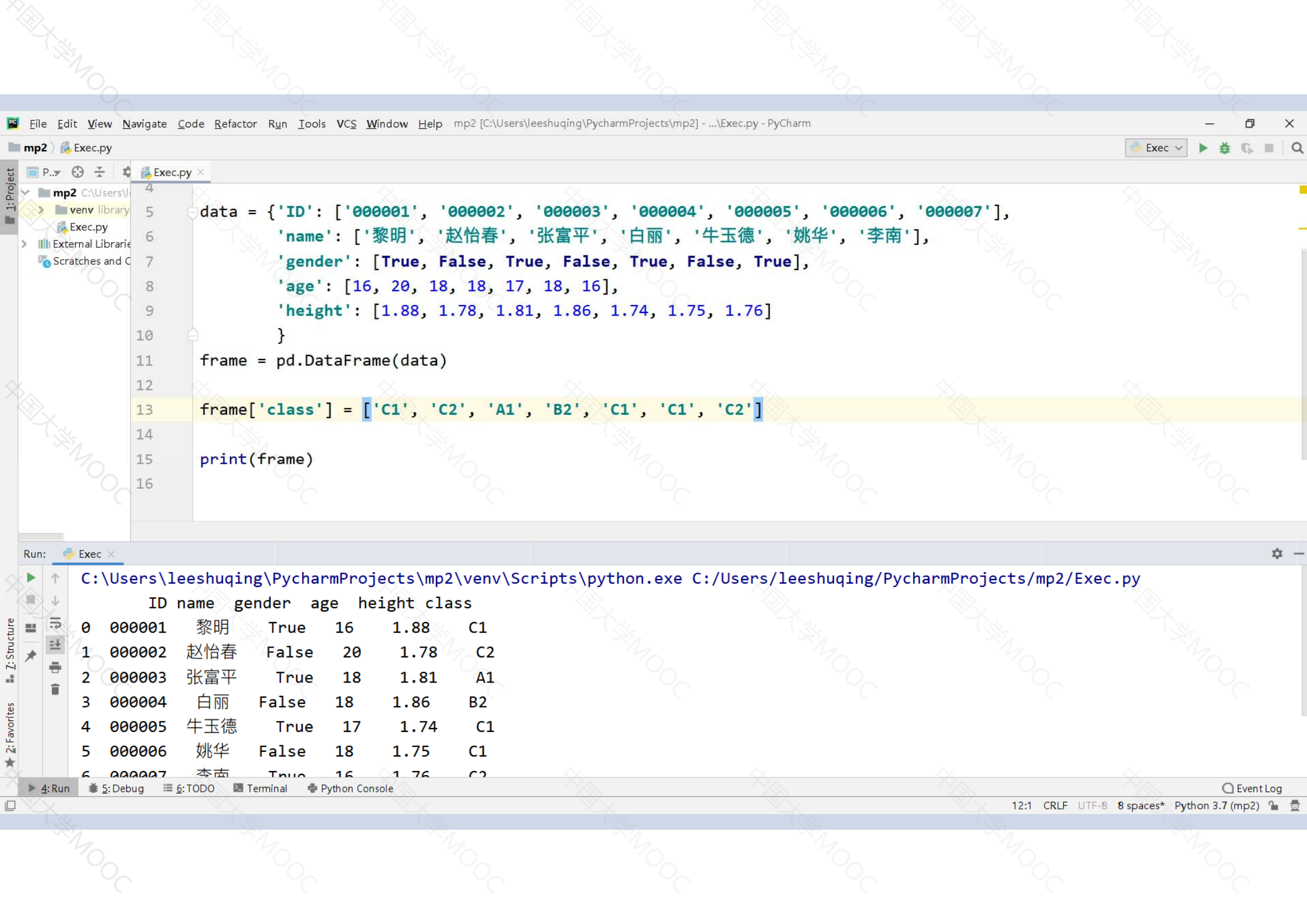
Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
       ID name  gender  age  height
0  000008  方文吾   False   21    1.75

Process finished with exit code 0
```

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

data1 = {'ID': ['000008'],
         'name': ['方文吾'],
         'gender': [False],
         'age': [21],
         'height': [1.75]
         }
frame1 = pd.DataFrame(data1)
frame = frame.append(frame1, ignore_index=True)
print(frame)
```

```
2  000003  张富千   True    18    1.81
3  000004  白丽    False   18    1.86
4  000005  牛玉德   True    17    1.74
5  000006  姚华    False   18    1.75
6  000007  李南    True    16    1.76
7  000008  方文吾   False   21    1.75


Process finished with exit code 0
```

File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm

mp2 › Exec.py   Exec ⌄
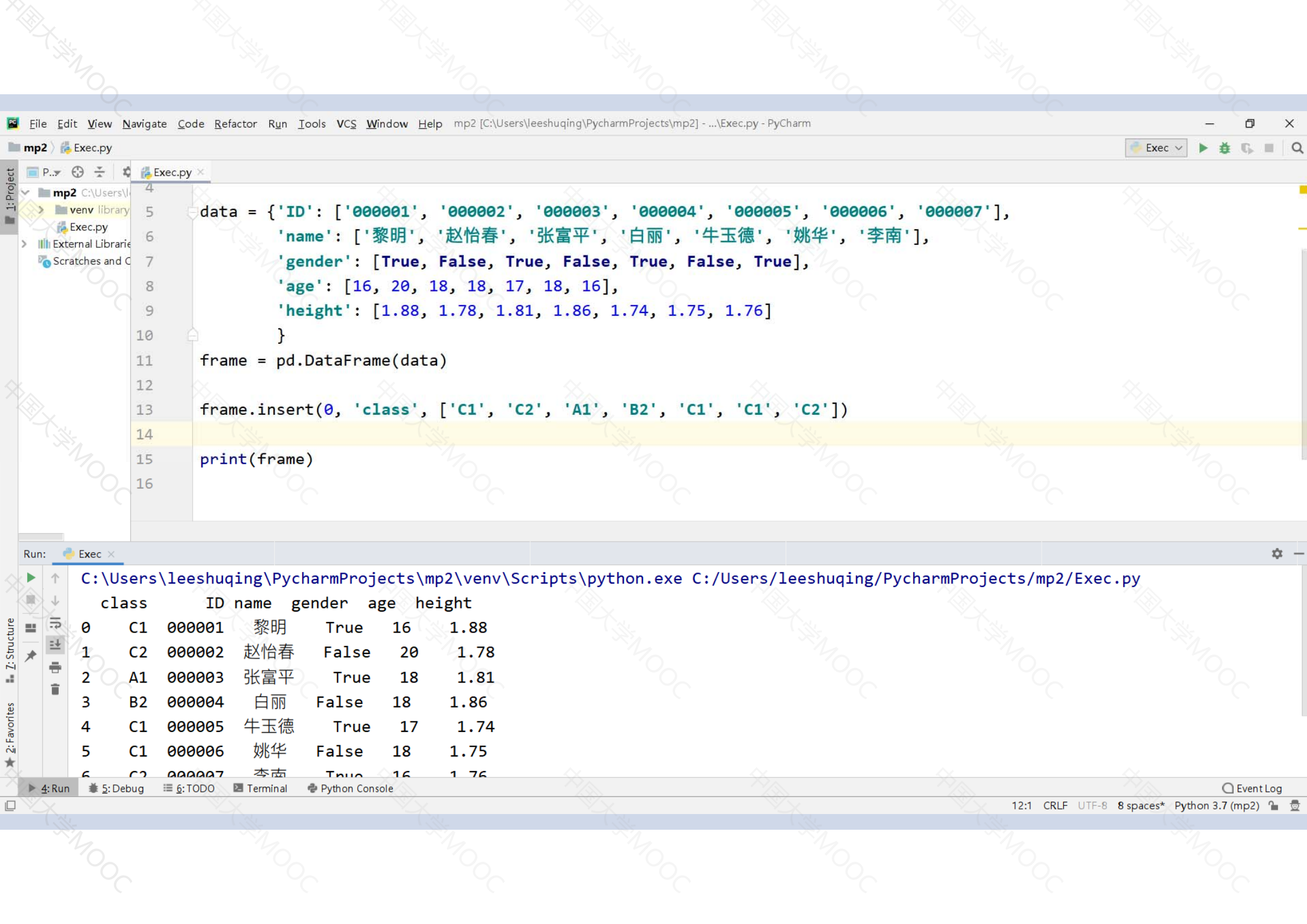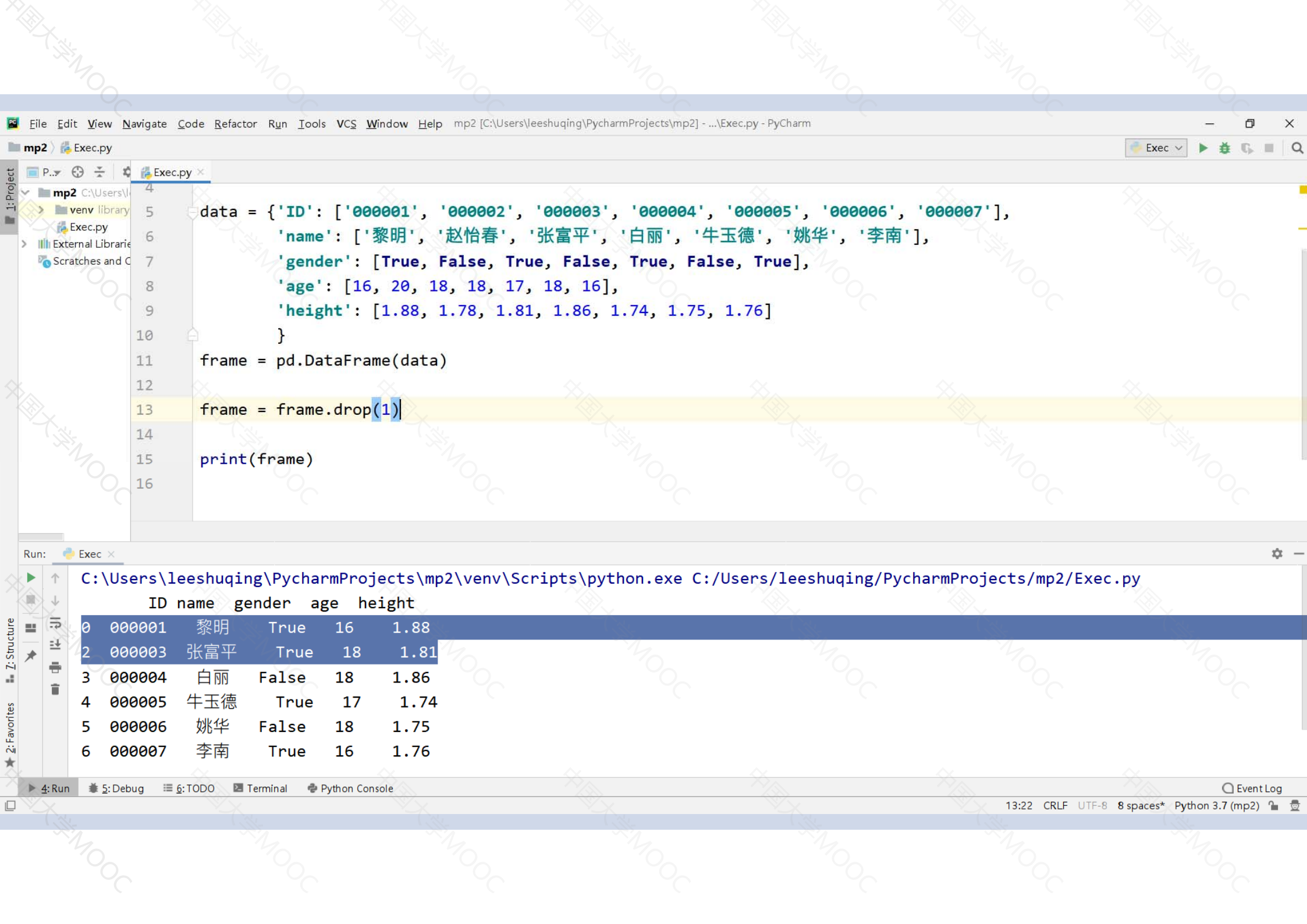
```python
 4
 5    data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
 6            'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
 7            'gender': [True, False, True, False, True, False, True],
 8            'age': [16, 20, 18, 18, 17, 18, 16],
 9            'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
10            }
11    frame = pd.DataFrame(data)
12
13    frame['class'] = ['C1', 'C2', 'A1', 'B2', 'C1', 'C1', 'C2']
14
15    print(frame)
16
```

Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
        ID name  gender  age  height class
0   000001   黎明    True   16    1.88    C1
1   000002  赵怡春   False   20    1.78    C2
2   000003  张富平    True   18    1.81    A1
3   000004   白丽   False   18    1.86    B2
4   000005  牛玉德    True   17    1.74    C1
5   000006   姚华   False   18    1.75    C1
6   000007   李南    True   16    1.76    C2
```
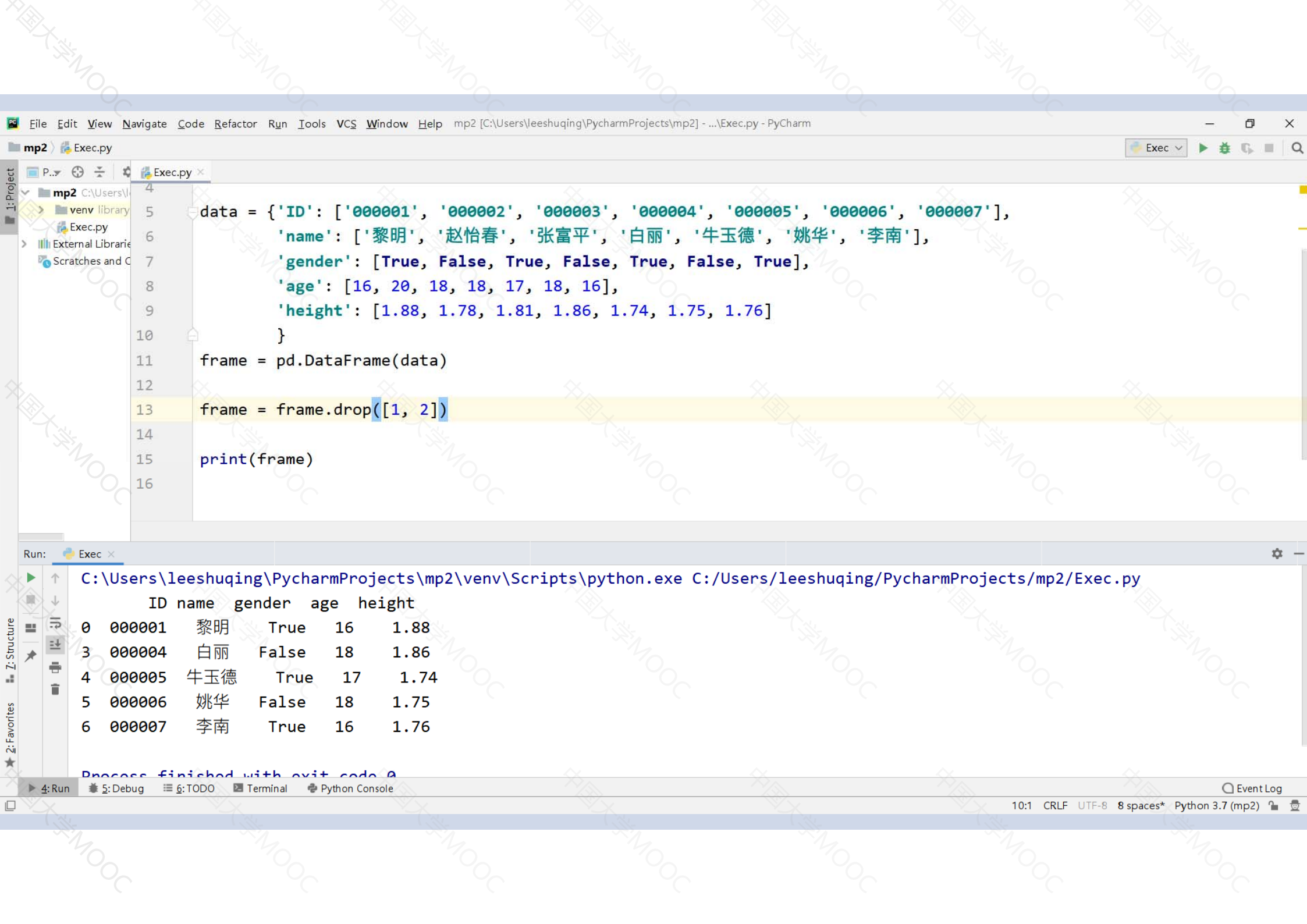
```python
4
5    data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
6            'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
7            'gender': [True, False, True, False, True, False, True],
8            'age': [16, 20, 18, 18, 17, 18, 16],
9            'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
10           }
11   frame = pd.DataFrame(data)
12
13   frame.insert(0, 'class', ['C1', 'C2', 'A1', 'B2', 'C1', 'C1', 'C2'])
14
15   print(frame)
16
```

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
   class      ID name  gender  age  height
0     C1  000001   黎明    True   16    1.88
1     C2  000002  赵怡春   False   20    1.78
2     A1  000003  张富平    True   18    1.81
3     B2  000004   白丽   False   18    1.86
4     C1  000005  牛玉德    True   17    1.74
5     C1  000006   姚华   False   18    1.75
6     C2  000007   李南    True   16    1.76
```

mp2 > Exec.py   Exec

Exec.py

```python
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)

frame = frame.drop(1)

print(frame)
```

Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
       ID name  gender  age  height
0  000001   黎明    True   16    1.88
2  000003  张富平    True   18    1.81
3  000004   白丽   False   18    1.86
4  000005  牛玉德    True   17    1.74
5  000006   姚华   False   18    1.75
6  000007   李南    True   16    1.76
```

4: Run   5: Debug   6: TODO   Terminal   Python Console   Event Log

13:22   CRLF   UTF-8   8 spaces*   Python 3.7 (mp2)

File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm   — ☐ ✕

mp2 Exec.py   Exec ▾ ▶ 🐞 🔁 ■ 🔍

```python
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)

frame = frame.drop([1, 2])

print(frame)
```
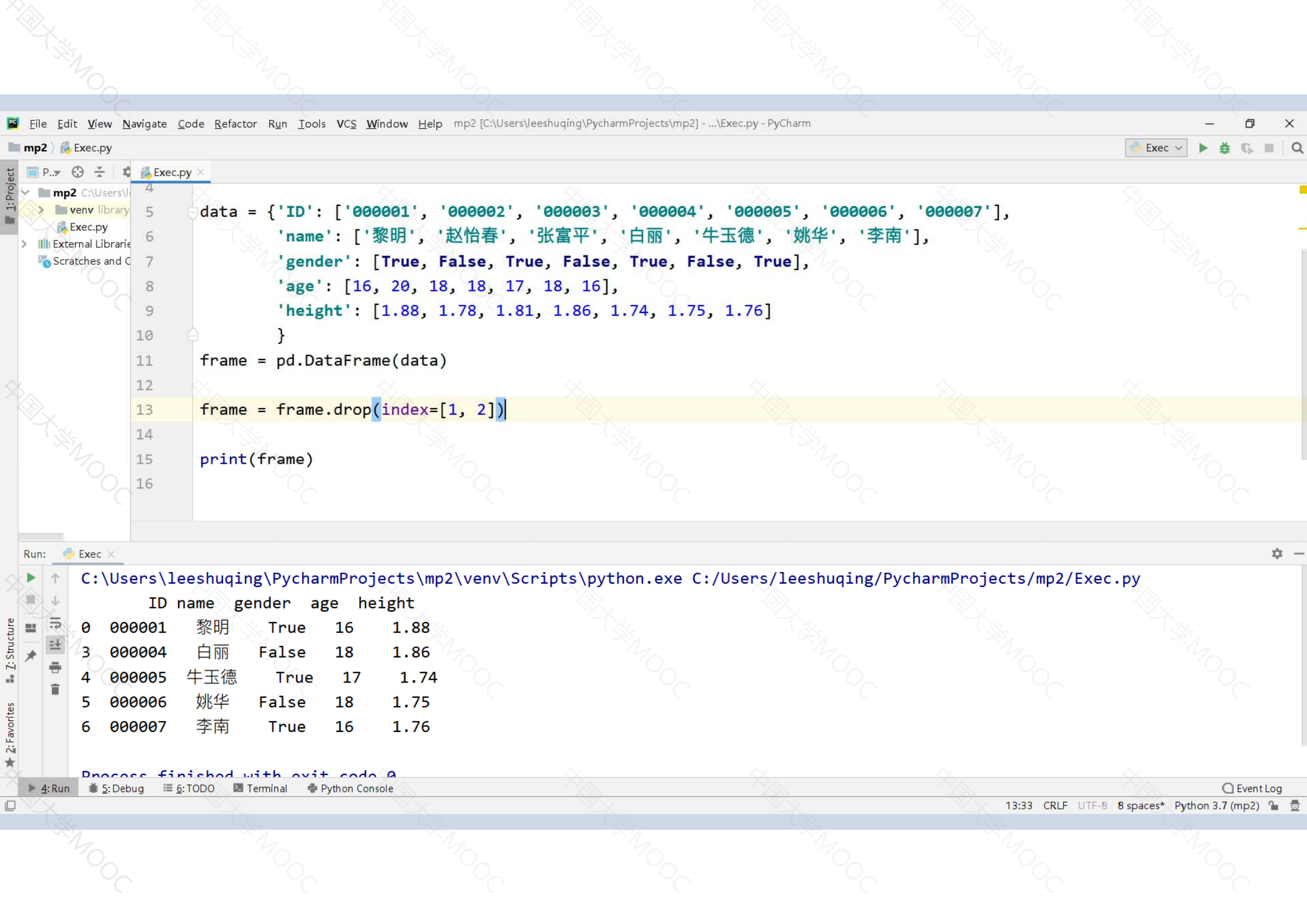
Run: 🐍 Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
       ID  name   gender   age   height
0  000001   黎明     True    16     1.88
3  000004   白丽    False    18     1.86
4  000005  牛玉德    True    17     1.74
5  000006   姚华    False    18     1.75
6  000007   李南     True    16     1.76

Process finished with exit code 0
```
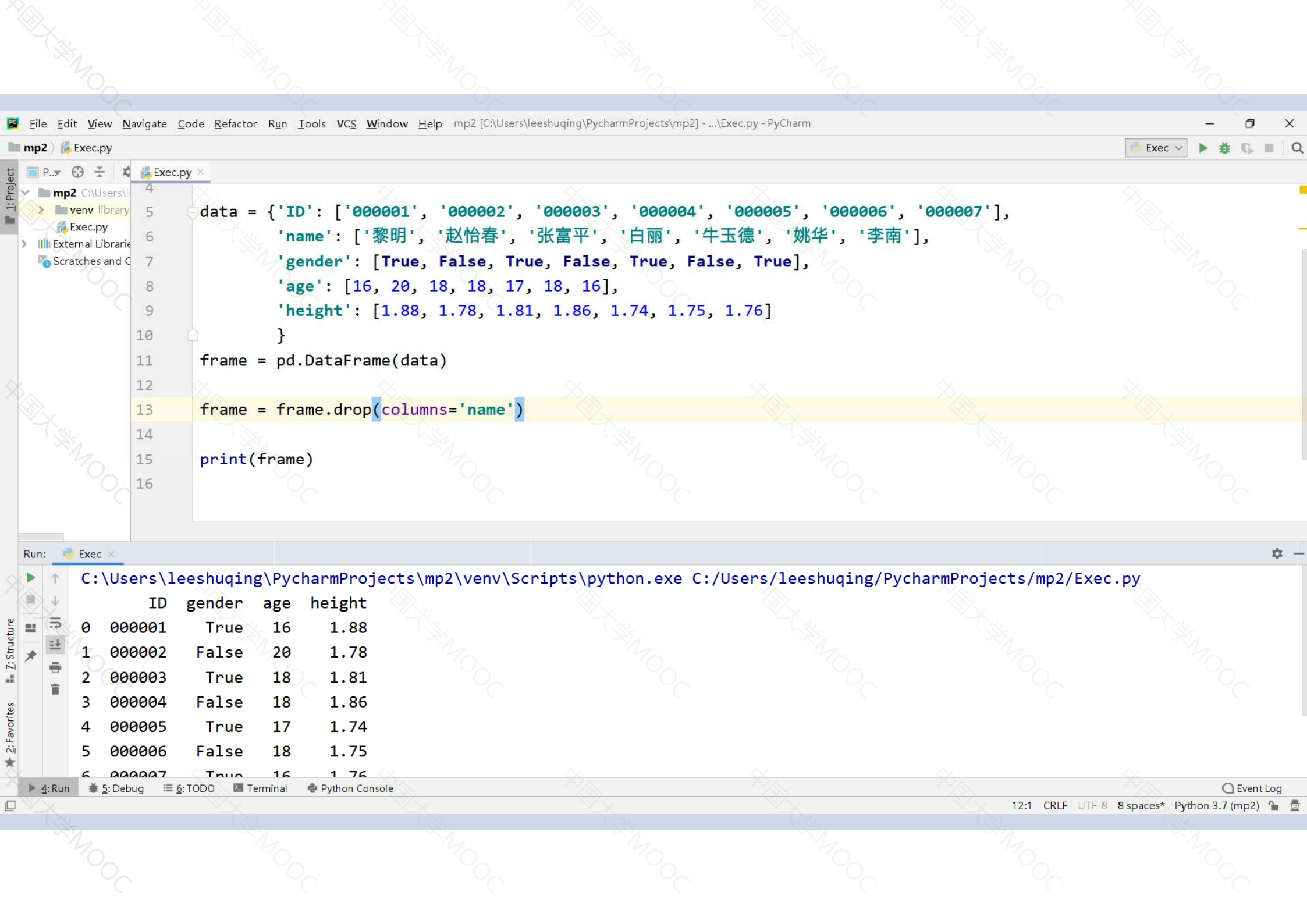
```python
data = {'ID': ['000001', '000002', '000003', '000004', '000005', '000006', '000007'],
        'name': ['黎明', '赵怡春', '张富平', '白丽', '牛玉德', '姚华', '李南'],
        'gender': [True, False, True, False, True, False, True],
        'age': [16, 20, 18, 18, 17, 18, 16],
        'height': [1.88, 1.78, 1.81, 1.86, 1.74, 1.75, 1.76]
        }
frame = pd.DataFrame(data)

frame = frame.drop(index=[1, 2])

print(frame)
```

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
       ID name  gender  age  height
0  000001   黎明    True   16    1.88
3  000004   白丽   False   18    1.86
4  000005  牛玉德    True   17    1.74
5  000006   姚华   False   18    1.75
6  000007   李南    True   16    1.76


Process finished with exit code 0
```
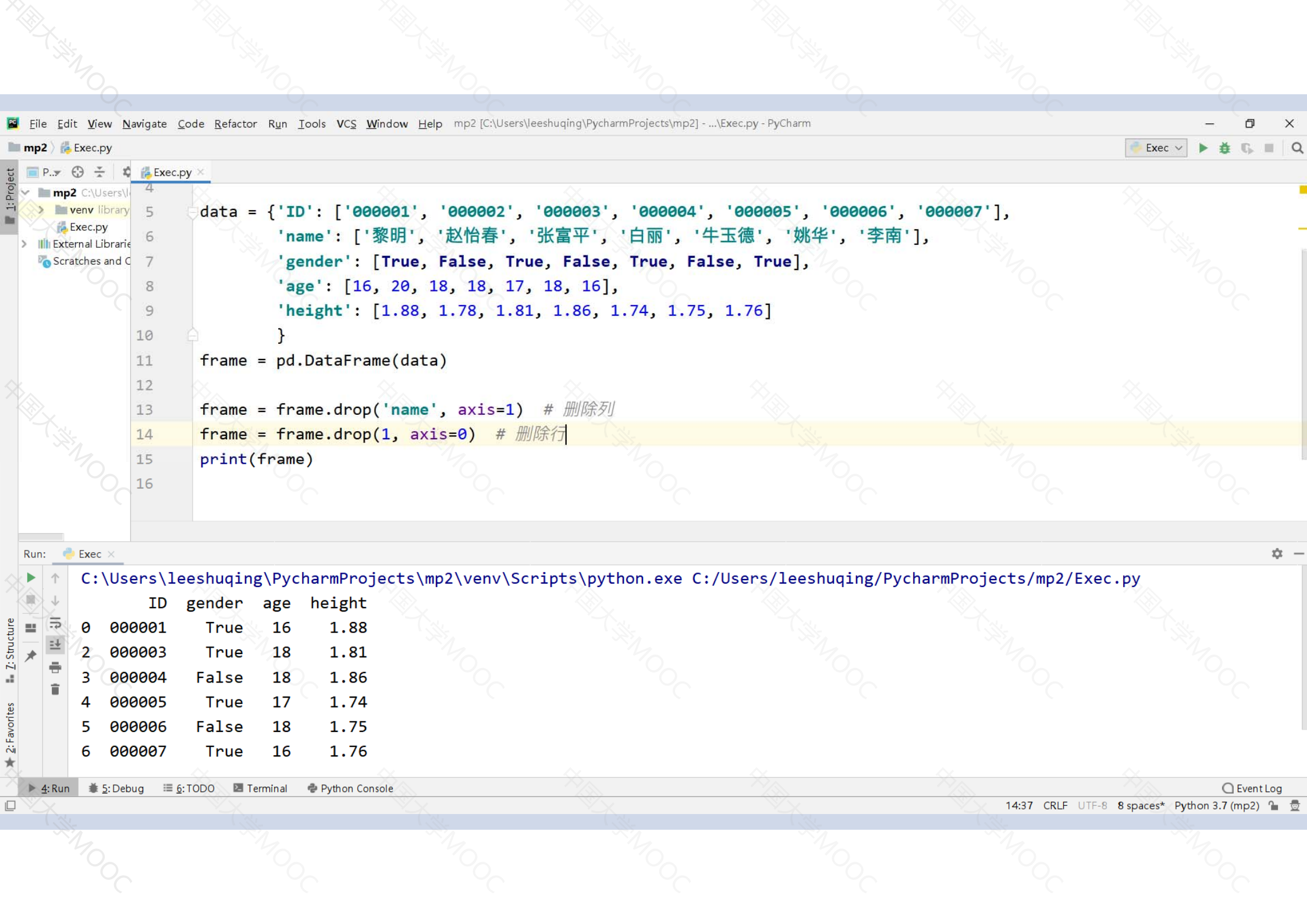
File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm

mp2 Exec.py   Exec   🔍

```python
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   frame = frame.drop(columns='name')
14
15   print(frame)
16
```

Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
        ID  gender  age  height
0   000001    True   16    1.88
1   000002   False   20    1.78
2   000003    True   18    1.81
3   000004   False   18    1.86
4   000005    True   17    1.74
5   000006   False   18    1.75
6   000007    True   16    1.76
```

mp2   Exec.py   Exec
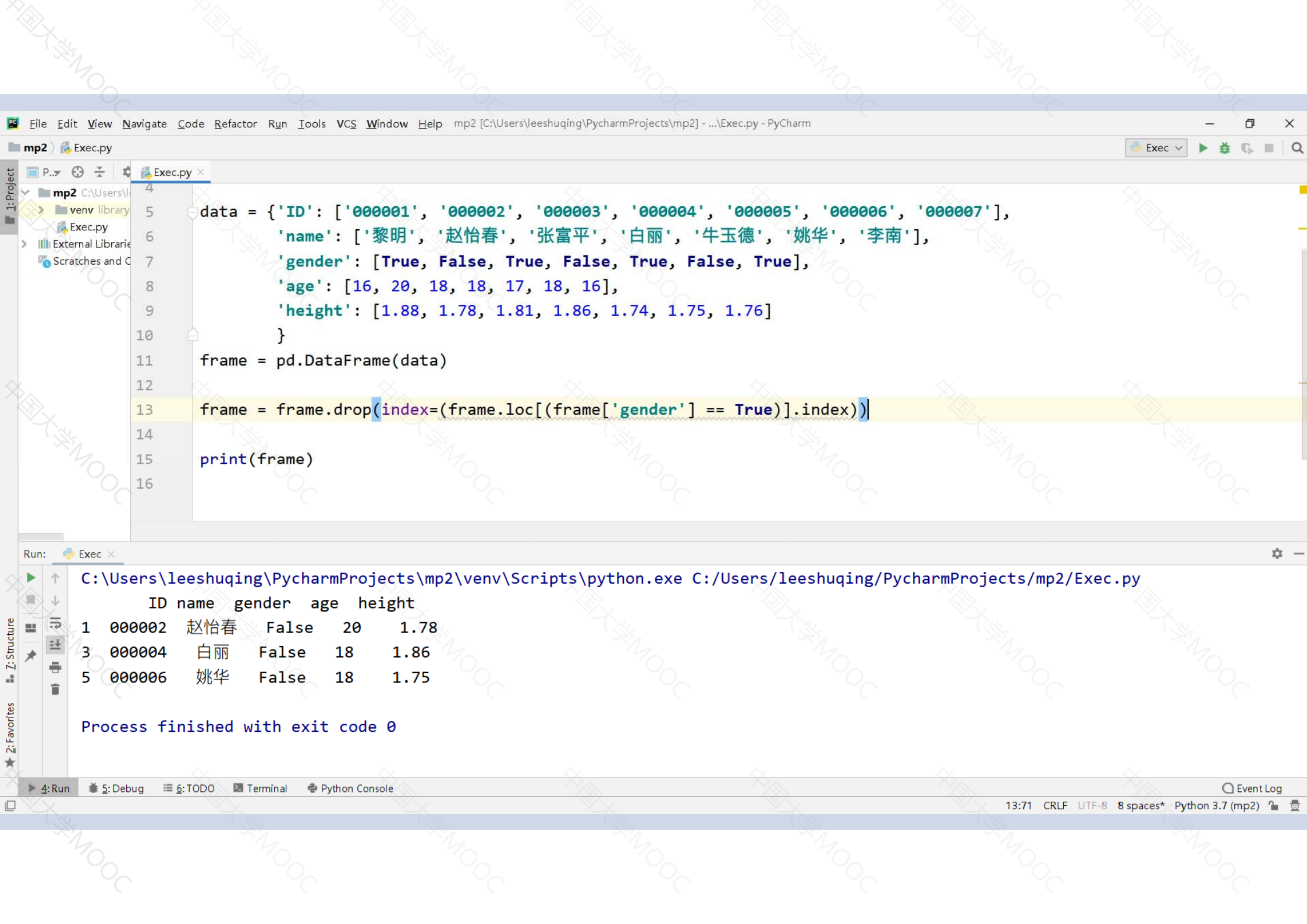
P..   Exec.py

```python
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    frame = frame.drop('name', axis=1)  # 删除列
14    frame = frame.drop(1, axis=0)  # 删除行
15    print(frame)
16
```

Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
       ID  gender  age  height
0  000001    True   16    1.88
2  000003    True   18    1.81
3  000004   False   18    1.86
4  000005    True   17    1.74
5  000006   False   18    1.75
6  000007    True   16    1.76
```

4: Run   5: Debug   6: TODO   Terminal   Python Console   Event Log

14:37   CRLF   UTF-8   8 spaces*   Python 3.7 (mp2)

File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm
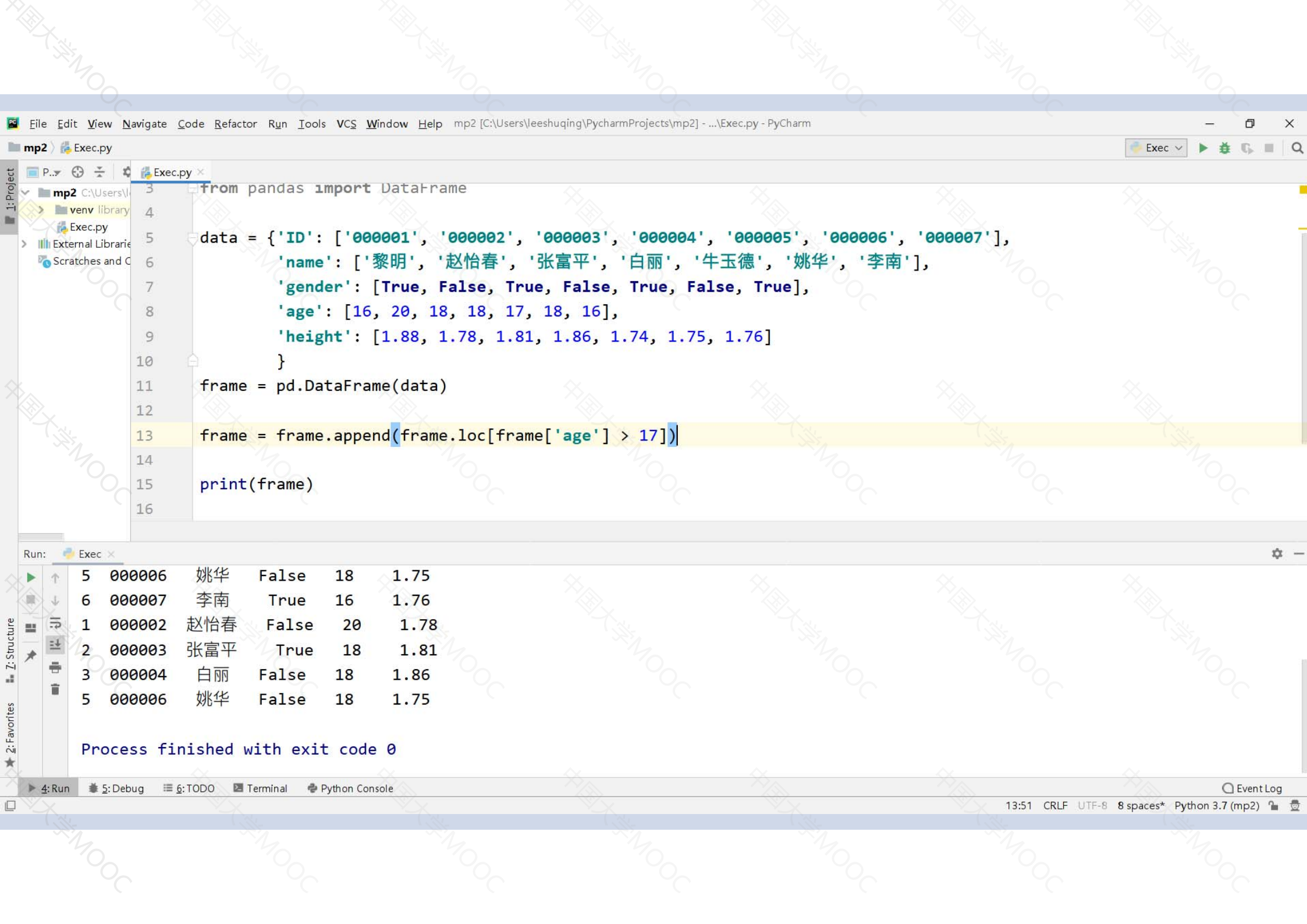
mp2   Exec.py   Exec

```python
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)

frame = frame.drop(index=(frame.loc[(frame['gender'] == True)].index))

print(frame)
```

Run: Exec

```
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
5  000006    姚华   False   18    1.75


Process finished with exit code 0
```

4: Run   5: Debug   6: TODO   Terminal   Python Console   Event Log

13:71   CRLF   UTF-8   8 spaces*   Python 3.7 (mp2)

mp2 > Exec.py

Exec

```python
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)


frame = frame.drop(index=(frame.loc[(frame['gender'] == True)].index))


print(frame.index)
print(frame.columns)
```

Run: Exec

```
C:\Users\leeshuqing\PycharmProjects\mp2\venv\Scripts\python.exe C:/Users/leeshuqing/PycharmProjects/mp2/Exec.py
Int64Index([1, 3, 5], dtype='int64')
Index(['ID', 'name', 'gender', 'age', 'height'], dtype='object')

Process finished with exit code 0
```

▶ 4: Run   5: Debug   6: TODO   Terminal   Python Console   Event Log

14:1   CRLF   UTF-8   8 spaces*   Python 3.7 (mp2)

File Edit View Navigate Code Refactor Run Tools VCS Window Help   mp2 [C:\Users\leeshuqing\PycharmProjects\mp2] - ...\Exec.py - PyCharm

mp2 ⟩ Exec.py   Exec

Exec.py ×

```python
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   frame = frame.append(frame.loc[frame['age'] > 17])
14
15   print(frame)
16
```
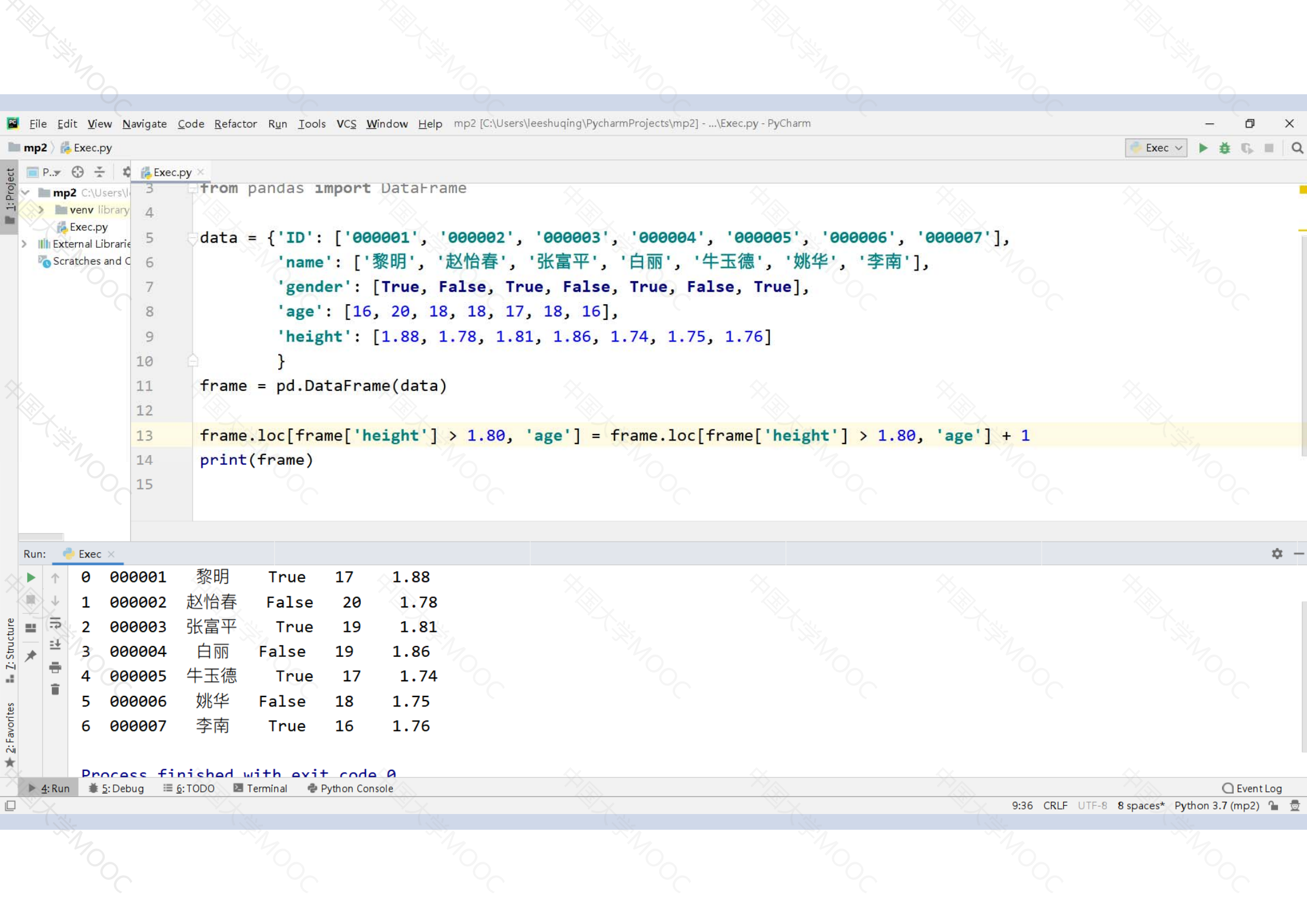
Run: Exec ×

```
5   000006    姚华     False    18    1.75
6   000007    李南     True     16    1.76
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
```

mp2   Exec.py   Exec

P..   Exec.py

mp2 C:\Users\l
venv library
Exec.py
External Librarie
Scratches and C

```python
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)


frame.loc[frame['height'] > 1.80, 'age'] = frame.loc[frame['height'] > 1.80, 'age'] + 1
print(frame)
```

Run: Exec

```
0  000001   黎明     True    17    1.88
1  000002   赵怡春   False    20    1.78
2  000003   张富平    True    19    1.81
3  000004   白丽     False    19    1.86
4  000005   牛玉德    True    17    1.74
5  000006   姚华     False    18    1.75
6  000007   李南      True    16    1.76


Process finished with exit code 0
```

4: Run   5: Debug   6: TODO   Terminal   Python Console   Event Log

9:36   CRLF   UTF-8   8 spaces*   Python 3.7 (mp2)

一次不学多，下次再学