```
1 import pandas as pd
2 import numpy as np
```

[문제1] 아래 그림과 같이 데이터프레임을 만드시오 .

	name	age	job
0	Jone	15	student
1	Jenny	30	developer
2	Nate	30	teacher

```
1 pr1 = {'name':['Jone','Jenny','Nate'],'age':[15,30,30],'job':['student','developer','teacher']}
2 pr1 = pd.DataFrame(pr1)
3 pr1
```

	name	age	job
0	Jone	15	student
1	Jenny	30	developer
2	Nate	30	teacher

[문제2] age(나이)가 20보다 작은 데이터를 표시하여라.

```
1 pr1.query('age<20')
2
```

	name	age	job
0	Jone	15	student

job_list.csv 파일을 로드하여 데이터프레임을 생성 한 뒤 아래 문제를 풀이하시오.

[문제3] 위 데이터프레임에 컬럼명을 넣어라. name, job, age

```
1 from google.colab import files
2 uploaded = files.upload()
```

파일 선택 joblist.csv

```
1 jl = pd.read_csv('joblist.csv',header=None)
2 jl.columns = (['name','job','gender','age'])
3 jl
```

	name	job	gender	age
0	John	teacher	male	40.0
1	Nate	teacher	male	35.0
2	Fred	teacher	male	37.0
3	Abraham	student	male	10.0
4	Brian	student	male	15.0
5	Janny	developer	female	16.0
6	Nate	teacher	male	32.0
7	Obrian	dentist	male	50.0
8	Yuna	teacher	female	38.0
9	Rob	lawyer	male	31.0
10	Brian	student	female	NaN
11	Matt	student	male	17.0
12	Wendy	banker	female	33.0
13	Edward	teacher	male	NaN
14	Chris	banker	female	NaN
15	Philip	lawyer	male	48.0
16	Janny	developer	female	16.0

[문제4] 데이터가 중복인 행을 찾아 삭제하시오.

```
1 jl.drop_duplicates(['name'])
2 jl
```

	name	job	gender	age
0	John	teacher	male	40.0
1	Nate	teacher	male	35.0
2	Fred	teacher	male	37.0
3	Abraham	student	male	10.0
4	Brian	student	male	15.0
5	Janny	developer	female	16.0
6	Nate	teacher	male	32.0
7	Obrian	dentist	male	50.0
8	Yuna	teacher	female	38.0
9	Rob	lawyer	male	31.0

[문제5] 나이가 없는 값(NaN)을 0으로 채워라.

11	Matt	student	male	17.0

```
1 jl['age'] = jl['age'].fillna(0)
2 jl
```

	name	job	gender	age
0	John	teacher	male	40.0
1	Nate	teacher	male	35.0
2	Fred	teacher	male	37.0
3	Abraham	student	male	10.0
4	Brian	student	male	15.0
5	Janny	developer	female	16.0
6	Nate	teacher	male	32.0
7	Obrian	dentist	male	50.0
8	Yuna	teacher	female	38.0
9	Rob	lawyer	male	31.0
10	Brian	student	female	0.0
11	Matt	student	male	17.0
12	Wendy	banker	female	33.0
13	Edward	teacher	male	0.0
14	Chris	banker	female	0.0
15	Philip	lawyer	male	48.0
16	Janny	developer	female	16.0

[문제6] major_list.csv파일을 로드하여 데이터프레임을 만들고 count열을 추가하여 직업별 데이터수를 카운트한 값을 표시하여라.

```
1 from google.colab import files
2 uploaded = files.upload()
```

파일 선택 major_list.csv

- **major_list.csv**(application/vnd.ms-excel) - 304 bytes, last modified: 2021. 5. 13. - 100% done

```
Saving major_list.csv to major_list (2).csv
```

```
1 mj = pd.read_csv('major_list.csv')
2 mj = pd.DataFrame(mj)
3 mj['major'].value_counts()
```

```
Compurter Science    4
Psychology           3
Economics            3
Physics              1
Name: major, dtype: int64
```

```
1 mgg = mj.groupby('major')
2 for name,group in mgg:
3   print(name+":"+str(len(group)))
4   print(group)
5   print()
```

```
Compurter Science:4
       name              major     sex
0      John  Compurter Science    male
1      Nate  Compurter Science    male
6  Jeniffer  Compurter Science  female
7    Edward  Compurter Science    male

Economics:3
    name      major     sex
4  Janny  Economics  female
5   Yuna  Economics  female
9  Wendy  Economics  female

Physics:1
      name    major   sex
2  Abraham  Physics  male

Psychology:3
     name       major     sex
3   Brian  Psychology    male
8    Zara  Psychology  female
10   Sera  Psychology  female
```

✓ 0초 오후 3:55에 완료됨