

In [1]:

```
import pandas
```

In [2]:

```
one=pandas.DataFrame({  
    'emp_id':[1,2,3,4,5],  
    'dept_name':['duvet','bidet','footwear','clothing','electronics'],  
    'aisle':[1,2,3,4,5]})
```

In [3]:

```
first=pandas.DataFrame(one)
```

In [4]:

```
two=pandas.DataFrame({  
    'emp_id':[6,7,8,9,10],  
    'dept_name':['grocery','toys','laundry','frozen','stationery'],  
    'aisle':[6,2,2,9,10]})
```

In [5]:

```
second=pandas.DataFrame(two)
```

In [6]:

```
three=pandas.read_csv('supermarket.csv')  
pandas.isnull(three).any()
```

Out[6]:

```
emp_id      False  
dept_name    False  
aisle        True  
dtype: bool
```

In [7]:

```
three.head()
```

Out[7]:

	emp_id	dept_name	aisle
0	6	grocery	6.0
1	7	toys	2.0
2	8	laundry	2.0
3	9	frozen	NaN
4	10	stationary	10.0

In [8]:

```
three.dropna()
```

Out[8]:

	emp_id	dept_name	aisle
0	6	grocery	6.0
1	7	toys	2.0
2	8	laundry	2.0
4	10	stationary	10.0

In [9]:

```
three.groupby('aisle').dept_name.value_counts()
```

Out[9]:

aisle	dept_name	
2.0	laundry	1
	toys	1
6.0	grocery	1
10.0	stationary	1

Name: dept\_name, dtype: int64

In [10]:

```
two.aisle.unique()
```

Out[10]:

```
array([ 6,  2,  9, 10])
```

In [11]:

```
len(two)
```

Out[11]:

```
5
```

In [12]:

```
two[two['aisle']>3]
```

Out[12]:

	emp_id	dept_name	aisle
0	6	grocery	6
3	9	frozen	9
4	10	stationery	10

In [13]:

```
two[two['aisle']==2]
```

Out[13]:

	emp_id	dept_name	aisle
1	7	toys	2
2	8	laundry	2

In [14]:

```
two[(two['aisle']==2) & (two['dept_name']=='toys')]
```

Out[14]:

	emp_id	dept_name	aisle
1	7	toys	2

In [15]:

```
(two[(two['aisle']==2) & (two['dept_name']=='toys')]).describe()
```

Out[15]:

	emp_id	aisle
count	1.0	1.0
mean	7.0	2.0
std	NaN	NaN
min	7.0	2.0
25%	7.0	2.0
50%	7.0	2.0
75%	7.0	2.0
max	7.0	2.0

In [16]:

```
two[two.aisle.isin([2,7,9])]
```

Out[16]:

	emp_id	dept_name	aisle
1	7	toys	2
2	8	laundry	2
3	9	frozen	9

In [17]:

```
(two.pivot(index='emp_id',columns='aisle',values='dept_name')).tail()
```

Out[17]:

	aisle	2	6	9	10
emp_id					
6	NaN	grocery	NaN	NaN	NaN
7	toys	NaN	NaN	NaN	NaN
8	laundry	NaN	NaN	NaN	NaN
9	NaN	NaN	frozen	NaN	NaN
10	NaN	NaN	NaN	stationery	NaN

In [18]:

```
(two.pivot(index='emp_id',columns='aisle',values='dept_name')).tail(3)
```

Out[18]:

	aisle	2	6	9	10
emp_id					
8	laundry	NaN	NaN	NaN	NaN
9	NaN	NaN	frozen	NaN	NaN
10	NaN	NaN	NaN	stationery	NaN

In [19]:

```
(two.pivot(index='emp_id',columns='aisle',values='dept_name')).tail(3).shift(1)
```

Out[19]:

	aisle	2	6	9	10
emp_id					
8	NaN	NaN	NaN	NaN	NaN
9	laundry	NaN	NaN	NaN	NaN
10	NaN	NaN	frozen	NaN	NaN

In [20]:

```
(two.melt(id_vars=['emp_id'],value_name='dept_name')).tail()
```

Out[20]:

	emp_id	variable	dept_name
5	6	aisle	6
6	7	aisle	2
7	8	aisle	2
8	9	aisle	9
9	10	aisle	10

In [21]:

```
melt1=two.melt(id_vars=['emp_id'],value_name='dept_name')
melt2=two.melt(id_vars=['emp_id'],value_name='aisle')
pandas.merge(melt1,melt2,on=['emp_id']).tail(1)
```

Out[21]:

	emp_id	variable_x	dept_name	variable_y	aisle
19	10	aisle	10	aisle	10

In [22]:

```
from functools import reduce
base=two[['aisle','dept_name','emp_id']]
feature=[base]+[melt1,melt2]
abt=reduce(lambda left,right: pandas.merge(left,right,on=['emp_id']),[melt1,melt2])
abt.tail(1)
```

Out[22]:

	emp_id	variable_x	dept_name	variable_y	aisle
19	10	aisle	10	aisle	10

In [23]:

```
pandas.concat([one,two])
```

Out[23]:

	emp_id	dept_name	aisle
0	1	duvet	1
1	2	bidet	2
2	3	footwear	3
3	4	clothing	4
4	5	electronics	5
0	6	grocery	6
1	7	toys	2
2	8	laundry	2
3	9	frozen	9
4	10	stationery	10

In [26]:

```
res=pandas.concat([one,two])  
res.to_csv('res.csv')
```

In [ ]: