In [1]:
```
import numpy as np
import pandas as pd

p1 = 'PONNURI ANIRUDDHA (RA2112704010015) - train.csv'
data = pd.read_csv(p1)
```

In [2]:
```
import pandas as pd
pd.set_option('display.max_rows', None)

orders_data = pd.DataFrame({
'ord_no':[70001,70009,70002,70004,70007,70005,70008,70010,70003,70012,70011,70013],
'purch_amt':[150.5,270.65,65.26,110.5,948.5,2400.6,5760,1983.43,2480.4,250.45, 75.29,3045.6],
'ord_date': ['2022-10-05','2022-09-10','2022-10-05','2022-08-17','2022-09-10','2022-07-27','2022-09-10','2022-10-10','2022-10-10','2022-06-
'customer_id':[5005,5001,5002,5009,5005,5007,5002,5004,5009,5008,5003,5002],
'salesman_id': [5002,5005,5001,5003,5002,5001,5001,5006,5003,5002,5007,5001]})
print("Original Orders DataFrame:")
print(orders_data)
print("\nGroup by two columns and count by each row:")
result = orders_data.groupby(['salesman_id','customer_id']).size().reset_index().groupby(['salesman_id','customer_id'])[[0]].max()
print(result)
```

```
Original Orders DataFrame:
    ord_no  purch_amt    ord_date  customer_id  salesman_id
0    70001     150.50  2022-10-05         5005         5002
1    70009     270.65  2022-09-10         5001         5005
2    70002      65.26  2022-10-05         5002         5001
3    70004     110.50  2022-08-17         5009         5003
4    70007     948.50  2022-09-10         5005         5002
5    70005    2400.60  2022-07-27         5007         5001
6    70008    5760.00  2022-09-10         5002         5001
7    70010    1983.43  2022-10-10         5004         5006
8    70003    2480.40  2022-10-10         5009         5003
9    70012     250.45  2022-06-27         5008         5002
10   70011      75.29  2022-08-17         5003         5007
11   70013    3045.60  2022-04-25         5002         5001

Group by two columns and count by each row:
                          0
salesman_id customer_id
5001        5002          3
            5007          1
5002        5005          2
            5008          1
5003        5009          2
5005        5001          1
5006        5004          1
5007        5003          1
```

In [3]:

```python
import pandas as pd
pd.set_option('display.max_rows', None)
#pd.set_option('display.max_columns', None)
df = pd.DataFrame({
'ord_no':[70001,70009,70002,70004,70007,70005,70008,70010,70003,70012,70011,70013],
'purch_amt':[150.50,270.65,65.26,110.50,948.50,2400.60,5760.00,1983.43,2480.40,250.45, 75.29,3045.60],
'ord_date': ['05-10-2022','09-10-2022','05-10-2022','08-17-2022','10-09-2022','07-27-2022','10-09-2022','10-10-2022','10-10-2022','06-17-20
'customer_id':[5001,5001,5005,5001,5005,5001,5005,5001,5005,5001,3005,5005],
'salesman_id': [5002,5005,5001,5003,5002,5001,5001,5006,5003,5002,5007,5001]})

print("Original Orders DataFrame:")
print(df)
df['ord_date']= pd.to_datetime(df['ord_date'])
print("\nMonth wise purchase amount:")
result_m = df.set_index('ord_date').groupby(pd.Grouper(freq='M')).agg({'purch_amt':sum})
print(result_m)
result_y = df.set_index('ord_date').groupby(pd.Grouper(freq='Y')).agg({'purch_amt':sum})
print("\nYEAR wise purchase amount:")
```

```
print(result_y)
#result = df.set_index('salesman_id').groupby(pd.Grouper()).agg({'purch_amt':sum})
```

```
Original Orders DataFrame:
    ord_no  purch_amt    ord_date  customer_id  salesman_id
0    70001     150.50  05-10-2022         5001         5002
1    70009     270.65  09-10-2022         5001         5005
2    70002      65.26  05-10-2022         5005         5001
3    70004     110.50  08-17-2022         5001         5003
4    70007     948.50  10-09-2022         5005         5002
5    70005    2400.60  07-27-2022         5001         5001
6    70008    5760.00  10-09-2022         5005         5001
7    70010    1983.43  10-10-2022         5001         5006
8    70003    2480.40  10-10-2022         5005         5003
9    70012     250.45  06-17-2022         5001         5002
10   70011      75.29  07-08-2022         3005         5007
11   70013    3045.60  04-25-2022         5005         5001

Month wise purchase amount:
            purch_amt
ord_date
2022-04-30    3045.60
2022-05-31     215.76
2022-06-30     250.45
2022-07-31    2475.89
2022-08-31     110.50
2022-09-30     270.65
2022-10-31   11172.33

YEAR wise purchase amount:
            purch_amt
ord_date
2022-12-31   17541.18
```

In [4]:

```
import pandas as pd
df = pd.DataFrame( {'id' : [1, 2, 1, 1, 2, 1, 2],'type' : [10, 15, 11, 20, 21, 12, 14],
                    'book' : ['C','C++','Java','C','C++','Java','C++']})

print("Original DataFrame:")
print(df)
result = df.groupby(['id', 'type', 'book']).size().unstack(fill_value=0)
print("\nResult:")
print(result)
```

```
Original DataFrame:
   id  type  book
0   1    10     C
1   2    15   C++
2   1    11  Java
3   1    20     C
4   2    21   C++
5   1    12  Java
6   2    14   C++

Result:
book     C  C++  Java
id type
1  10    1    0     0
   11    0    0     1
   12    0    0     1
   20    1    0     0
2  14    0    1     0
   15    0    1     0
   21    0    1     0
```

In [ ]: