EE4211 Project - Question 3

Group name: WeUsOurs (Group No.26)

Group members: Liu Tianshu, Nie Hongtuo, Pan Jie, Zhang Chenxi

The whole code takes ~3 mins to finish running.

Please read our annotations while running the code:)

In Question 1, we have found out the top 5 correlated meters for each meterID. We are interesting about this results and want to explore more on this matter.

Here, in Question 3, we propose to use two of the most popular clustering algorithms: K-Means and Gaussian mixture model(GMM) to conduct the cluster for each meter.

The clustering results will be compared with the top 5 correlated results that we obtained in Question1. We want to explore:

1,If all the top 5 correlated meters are in the same cluster (For K-Means & GMM)?

2,If some of the top 5 correlated meters are in the same cluster (For K-Means & GMM)?

3,If none of the top 5 correlated meters are in the same cluster (For K-Means & GMM)?

4,Any top 5 correlated meters exist in both K-Means cluster and GMM cluster?

5,What is the matching rate between 5 correlated meters and the clustering results (For K-Means & GMM)?

We will present detailed information in our presentation as well.

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.interpolate import interp1d
import seaborn as sns
```

In [2]:
```
df = pd.read_csv('dataport-export_gas_oct2015-mar2016.csv')
df
grouped = df.groupby(['dataid'])
len(grouped)
keys = pd.Series(df.dataid.values)
keys = keys.unique()
keys = np.sort(keys)
```

In [3]:
```
def rectify(df):
    grouped = df.groupby(['dataid'])
    for key,group in grouped:
        val = group.meter_value.values
        diff = val[1:] - val[:-1]
        defect_time = np.where(diff < 0)[0] + 1
        if len(defect_time):
            for i in range(1,len(val)):
                if val[i] > val[-1]:
                    val[i] = val[i - 1]
                elif val[i] < val[i - 1]:
                    val[i] = val[i - 1]
            df.loc[df['dataid'] == key, 'meter_value'] = val
    return df
```

In [4]:
```
df = rectify(df)
rec_df = df
```

In [5]:
```
def local2utc(time):
    return pd.Timestamp(time)
df.localminute = df.apply(lambda r: local2utc(r.localminute), axis=1)

def select_data(df, month) -> pd.DataFrame:
    column_names = df.columns.values
    values = df.values
    values = filter(lambda x: x[0].month == month, values)
    values = pd.DataFrame(values)
    for i in range(len(column_names)):
        values = values.rename(columns={i:column_names[i]})
    return values
```

In [6]:

```
def hourly_range_month(df, start_month, end_month):
    month_dict = {1:31,2:28,3:31,4:30,5:31,6:30,7:31,8:31,9:30,10:31,11:30,12:
31}
    groups = df.groupby(['dataid'], sort=['localminute'])
    processed_group = {}
    for key in groups.groups.keys():
        new_group = []
        current_group = groups.get_group(key).values
        # init_first_point
        first_point = current_group[0]
        first_point[0] = first_point[0].replace(month = start_month, day = 1,
hour = 0, minute = 0, second = 0)
        index_current_group = 0
        tmp = []
        # replace all time on the hour
        for row in current_group:
            row[0] = row[0].replace(minute = 0, second = 0)
            tmp.append(row)
        current_group = tmp
        # interpolate data 24 hour for 30 day
        for month in range(start_month, end_month + 1):
            month = month % 12 if month % 12 != 0 else 12
            for day in range(1, month_dict[month] + 1):
                for hour in range(0, 24):
                    # make sure its the last data of the current time
                    while (True):
                        if (index_current_group >= len(current_group)):
                            break
                        current_row = current_group[index_current_group]
                        time = current_row[0]
                        if (time.month == month and time.day == day and time.h
our == hour):
                            index_current_group += 1
                        else:
                            break
                    # assign data
                    current_row = current_group[index_current_group - 1]
                    time = current_row[0]
                    date = pd.Timestamp(year = time.year, month = time.month,
day = time.day, hour = time.hour, minute = 0, second = 0)
                    current_row = [date, current_row[1], current_row[2]]
                    if (time.month == month and time.day == day and time.hour
== hour):
                        if (len(new_group) > 0 and new_group[-1][0].month == m
onth and new_group[-1][0].day == day and new_group[-1][0].hour == hour):
                            new_group[-1] = current_row
                        else:
                            new_group.append(current_row)
                    else:
                        date = pd.Timestamp(year = time.year, month = month, d
ay = day, hour = hour, minute = 0, second = 0)
                        row = [date,new_group[-1][1],new_group[-1][2]]
                        new_group.append(row)
        processed_group[key] = new_group
    return processed_group
```

In [7]:
```
hourly_data_raw = hourly_range_month(df,10,15)

def hourly_data_matrix(hourly_data_raw, keys):
    hourly_data = np.zeros((157, 4368), dtype=int)
    for i in range(157):
        key = keys[i]
        monthly_value = hourly_data_raw[key]
        monthly_value = list(zip(*monthly_value))[2]
        monthly_value = np.array(monthly_value)
        hourly_data[i] = monthly_value

    return hourly_data

hourly_data = hourly_data_matrix(hourly_data_raw, keys)
hourly_data.shape
hourly_corr = np.corrcoef(hourly_data)
hourly_corr.shape
import datetime
df = pd.read_csv('dataport-export_gas_oct2015-mar2016.csv')
def rectify(df):

    grouped = df.groupby(['dataid'])
    for key,group in grouped:
        val = group.meter_value.values
        diff = val[1:] - val[:-1]
        defect_time = np.where(diff < 0)[0] + 1
        if len(defect_time):
            for i in range(1,len(val)):
                if val[i] > val[-1]:
                    val[i] = val[i - 1]
                elif val[i] < val[i - 1]:
                    val[i] = val[i - 1]
            df.loc[df['dataid'] == key, 'meter_value'] = val
    return df
```

First,similar as Q1.3, we find out the top 5 relative families for each MeterID.

In [8]:

```
Top5 = []
for i in range(157):
    order = np.argsort(hourly_corr[i,:])
    #print('The top 5 relative families of consumer No.', keys[i], 'are:\t', k
eys[np.flip(order)[1:6]])
    top5 = [keys[i],keys[np.flip(order)[1:6]]]
    Top5.append(top5)
    #print(Top5)
    print('The top 5 relative families of consumer No.', keys[i], 'are:\t', ke
ys[np.flip(order)[1:6]])
```

```
The top 5 relative families of consumer No. 35 are:      [2449 1697 5810 7287
5403]
The top 5 relative families of consumer No. 44 are:      [1697 1283   35 2449
484]
The top 5 relative families of consumer No. 77 are:      [9849 1697   35   44
2449]
The top 5 relative families of consumer No. 94 are:      [2575 5892 9134 5129
1507]
The top 5 relative families of consumer No. 114 are:     [1556 4514 2094 3893
1619]
The top 5 relative families of consumer No. 187 are:     [9766 4296 8086 1507
2018]
The top 5 relative families of consumer No. 222 are:     [7965   44 1283 2965
77]
The top 5 relative families of consumer No. 252 are:     [3577 5395 4732 1556
5814]
The top 5 relative families of consumer No. 370 are:     [7117 1589 3527 6836
1714]
The top 5 relative families of consumer No. 483 are:     [3367 4193 5439 8829
7017]
The top 5 relative families of consumer No. 484 are:     [5484 5403 5892 5810
9982]
The top 5 relative families of consumer No. 661 are:     [7794 5814 3723  739
4514]
The top 5 relative families of consumer No. 739 are:     [ 661 5814 7794 3723
9295]
The top 5 relative families of consumer No. 744 are:     [7900 8829 3544 4998
3527]
The top 5 relative families of consumer No. 871 are:     [8890 3577 9052 5131
2072]
The top 5 relative families of consumer No. 1042 are:    [7030 9982 4296 9766
5193]
The top 5 relative families of consumer No. 1086 are:    [5972 4356 9295 9121
3723]
The top 5 relative families of consumer No. 1103 are:    [3310 8386 6830 9631
7016]
The top 5 relative families of consumer No. 1185 are:    [5972 4356 5395 3918
1086]
The top 5 relative families of consumer No. 1283 are:    [1697  484   44 9982
5484]
The top 5 relative families of consumer No. 1403 are:    [9600 5317 3036 2755
8244]
The top 5 relative families of consumer No. 1415 are:    [ 744 8086 2018 2470
3527]
The top 5 relative families of consumer No. 1507 are:    [5636 6910 2094 3778
5129]
The top 5 relative families of consumer No. 1556 are:    [4514 2094 4732  114
4031]
The top 5 relative families of consumer No. 1589 are:    [3527  370 7117 1714
5439]
The top 5 relative families of consumer No. 1619 are:    [1556 4031  114 2072
3577]
The top 5 relative families of consumer No. 1697 are:    [2449   35 5810  484
5403]
The top 5 relative families of consumer No. 1714 are:    [5439 3527 1589 8386
7117]
The top 5 relative families of consumer No. 1718 are:    [1801 1791  114 2094
```

4514]
The top 5 relative families of consumer No. 1790 are: [5814 3134 3723 661
739]
The top 5 relative families of consumer No. 1791 are: [1718 114 1801 2094
1556]
The top 5 relative families of consumer No. 1792 are: [1619 114 4031 2034
2072]
The top 5 relative families of consumer No. 1800 are: [5193 1042 6830 2470
3849]
The top 5 relative families of consumer No. 1801 are: [5636 1718 1507 8086
9729]
The top 5 relative families of consumer No. 2018 are: [8086 7741 9766 4296
1801]
The top 5 relative families of consumer No. 2034 are: [4031 8829 9729 3544
6412]
The top 5 relative families of consumer No. 2072 are: [4031 1619 3577 2034
4732]
The top 5 relative families of consumer No. 2094 are: [1556 1507 4514 5636
114]
The top 5 relative families of consumer No. 2129 are: [661 7794 3723 9121
9134]
The top 5 relative families of consumer No. 2233 are: [3039 7919 4421 7460
9956]
The top 5 relative families of consumer No. 2335 are: [7989 9134 6910 4732
5636]
The top 5 relative families of consumer No. 2378 are: [7017 2335 2470 7989
1589]
The top 5 relative families of consumer No. 2449 are: [5810   35 1697 484
5484]
The top 5 relative families of consumer No. 2461 are: [483 4767 4193 7016
3367]
The top 5 relative families of consumer No. 2470 are: [3527 7017 5439 3849
8086]
The top 5 relative families of consumer No. 2575 are: [94 7989 5892 1507
3778]
The top 5 relative families of consumer No. 2638 are: [4352 9474 4421 7460
7919]
The top 5 relative families of consumer No. 2645 are: [9278 8155 6578 1185
7682]
The top 5 relative families of consumer No. 2755 are: [3036 8244 9600 9160
2946]
The top 5 relative families of consumer No. 2814 are: [6673 9474 4352 6101
2638]
The top 5 relative families of consumer No. 2818 are: [3367 5439 7017 8829
483]
The top 5 relative families of consumer No. 2945 are: [9052 3577 3544 8829
1556]
The top 5 relative families of consumer No. 2946 are: [8244 3036 9160 2755
7566]
The top 5 relative families of consumer No. 2965 are: [1415 6505 2378 5193
5658]
The top 5 relative families of consumer No. 2980 are: [9631 6836 1714 3527
5193]
The top 5 relative families of consumer No. 3036 are: [9600 2755 8244 9160
5317]
The top 5 relative families of consumer No. 3039 are: [2233 7919 4421 7460
2638]

```
The top 5 relative families of consumer No. 3134 are:    [9639 1790 1086 3723
739]
The top 5 relative families of consumer No. 3310 are:    [9631 1714 2461 1103
7117]
The top 5 relative families of consumer No. 3367 are:    [ 483 5439 7017 8829
4193]
The top 5 relative families of consumer No. 3527 are:    [1589 2470 5439 7017
1714]
The top 5 relative families of consumer No. 3544 are:    [8829 2034 4031 9729
5636]
The top 5 relative families of consumer No. 3577 are:    [9052 8890 4732 1556
4031]
The top 5 relative families of consumer No. 3635 are:    [ 661 5814 7794 8890
9134]
The top 5 relative families of consumer No. 3723 are:    [ 661 9121 7794 5814
4514]
The top 5 relative families of consumer No. 3778 are:    [1507 5636 6910 2094
5892]
The top 5 relative families of consumer No. 3849 are:    [2470 7017 5439 3367
3527]
The top 5 relative families of consumer No. 3893 are:    [7794 4514 1556 5814
661]
The top 5 relative families of consumer No. 3918 are:    [5395  252 8890 3577
5814]
The top 5 relative families of consumer No. 4029 are:    [7030 8156 3849 4296
1042]
The top 5 relative families of consumer No. 4031 are:    [2034 1556 2072 9729
1619]
The top 5 relative families of consumer No. 4193 are:    [ 483 3367 7900 5439
8829]
The top 5 relative families of consumer No. 4228 are:    [4514 7794 2094 1556
3893]
The top 5 relative families of consumer No. 4296 are:    [9766 8086 7741 1801
2018]
The top 5 relative families of consumer No. 4352 are:    [2638 9474 4421 7460
7919]
The top 5 relative families of consumer No. 4356 are:    [5972 1086 9121  661
3723]
The top 5 relative families of consumer No. 4373 are:    [4732 1556 2094 6910
1507]
The top 5 relative families of consumer No. 4421 are:    [7460 7919 2638 9474
4352]
The top 5 relative families of consumer No. 4447 are:    [2645 8967 6578 9278
8155]
The top 5 relative families of consumer No. 4514 are:    [7794 1556 3893 4228
114]
The top 5 relative families of consumer No. 4671 are:    [6863 7016 2461 3310
4767]
The top 5 relative families of consumer No. 4732 are:    [1556 9052 3577 4373
6910]
The top 5 relative families of consumer No. 4767 are:    [ 483 2072 2461 4031
2034]
The top 5 relative families of consumer No. 4874 are:    [2814  222 6673 7965
9474]
The top 5 relative families of consumer No. 4998 are:    [7900 8829 5439  483
3544]
The top 5 relative families of consumer No. 5129 are:    [1507 6910 4373 4732
```

2094]
The top 5 relative families of consumer No. 5131 are: [9052 8890 8829 2335
3544]
The top 5 relative families of consumer No. 5193 are: [2470 3527 3849 1415
1589]
The top 5 relative families of consumer No. 5275 are: [871 2072 2034 1619
7429]
The top 5 relative families of consumer No. 5317 are: [9600 3036 2755 8244
9160]
The top 5 relative families of consumer No. 5395 are: [252 3918 5814 661
8890]
The top 5 relative families of consumer No. 5403 are: [9982 484 5810 9766
5892]
The top 5 relative families of consumer No. 5439 are: [3367 7900 483 3527
7017]
The top 5 relative families of consumer No. 5484 are: [484 2575 5892 2449
5810]
The top 5 relative families of consumer No. 5545 are: [7739 9620 6830 1103
5317]
The top 5 relative families of consumer No. 5636 are: [1507 1801 6910 2094
3778]
The top 5 relative families of consumer No. 5658 are: [6505 7989 744 7900
2335]
The top 5 relative families of consumer No. 5785 are: [9639 3134 1086 5972
8155]
The top 5 relative families of consumer No. 5810 are: [2449 7287 9121 5403
35]
The top 5 relative families of consumer No. 5814 are: [661 7794 3893 3723
252]
The top 5 relative families of consumer No. 5892 are: [1507 3778 6910 2094
5129]
The top 5 relative families of consumer No. 5972 are: [4356 1086 5395 5814
661]
The top 5 relative families of consumer No. 6101 are: [2946 9160 8244 7566
2755]
The top 5 relative families of consumer No. 6412 are: [2034 8829 3544 3367
7429]
The top 5 relative families of consumer No. 6505 are: [1589 370 7117 2378
744]
The top 5 relative families of consumer No. 6578 are: [9278 8155 5785 7682
8967]
The top 5 relative families of consumer No. 6673 are: [2814 5545 6101 9474
4352]
The top 5 relative families of consumer No. 6685 are: [6505 2378 744 1589
370]
The top 5 relative families of consumer No. 6830 are: [8386 2980 9631 1714
6836]
The top 5 relative families of consumer No. 6836 are: [370 7117 1589 3527
1714]
The top 5 relative families of consumer No. 6863 are: [9620 7739 3310 9631
1103]
The top 5 relative families of consumer No. 6910 are: [1507 5636 4732 4373
3778]
The top 5 relative families of consumer No. 7016 are: [2461 1714 483 8386
4767]
The top 5 relative families of consumer No. 7017 are: [3367 5439 2470 3527
483]

```
The top 5 relative families of consumer No. 7030 are:    [8156 4296 9766 1042
8086]
The top 5 relative families of consumer No. 7117 are:    [ 370 1589 3527 1714
6836]
The top 5 relative families of consumer No. 7287 are:    [9295 5810 9121 1086
7682]
The top 5 relative families of consumer No. 7429 are:    [8829 2034 3544 4031
6412]
The top 5 relative families of consumer No. 7460 are:    [4421 7919 2638 9474
4352]
The top 5 relative families of consumer No. 7566 are:    [8244 2946 3036 2755
9160]
The top 5 relative families of consumer No. 7674 are:    [1801 9729 8086 3544
2018]
The top 5 relative families of consumer No. 7682 are:    [7287 1086 9121 9295
35]
The top 5 relative families of consumer No. 7739 are:    [6830 1103 3310 9631
6836]
The top 5 relative families of consumer No. 7741 are:    [8086 1801 2018 4296
9766]
The top 5 relative families of consumer No. 7794 are:    [4514 3893  661 4228
3723]
The top 5 relative families of consumer No. 7900 are:    [ 744 5439 4998 8829
3527]
The top 5 relative families of consumer No. 7919 are:    [4421 7460 2638 9474
4352]
The top 5 relative families of consumer No. 7965 are:    [2965 1283   44 1697
5484]
The top 5 relative families of consumer No. 7989 are:    [2335 5636 2018 2575
1801]
The top 5 relative families of consumer No. 8059 are:    [4031 3577 2034 1619
8890]
The top 5 relative families of consumer No. 8084 are:    [7674 8086 1791 1801
1718]
The top 5 relative families of consumer No. 8086 are:    [2018 1801 7741 4296
9766]
The top 5 relative families of consumer No. 8155 are:    [5785 1185 9278 5972
9639]
The top 5 relative families of consumer No. 8156 are:    [7030 4296 7741 9766
1801]
The top 5 relative families of consumer No. 8244 are:    [3036 2946 9160 2755
9600]
The top 5 relative families of consumer No. 8386 are:    [1714 3527 5439 2470
3849]
The top 5 relative families of consumer No. 8467 are:    [  77 9849 1283 2233
7965]
The top 5 relative families of consumer No. 8703 are:    [3310 9631 6836  370
1103]
The top 5 relative families of consumer No. 8829 are:    [3544 2034 4031 9729
7429]
The top 5 relative families of consumer No. 8890 are:    [9052 3577 4732  252
1556]
The top 5 relative families of consumer No. 8967 are:    [6578 4447 9278 8155
2645]
The top 5 relative families of consumer No. 9052 are:    [3577 8890 4732 1556
4031]
The top 5 relative families of consumer No. 9121 are:    [3723 9295 7794  661
```

```
4356]
The top 5 relative families of consumer No. 9134 are:    [4373  661 2335 4732
6910]
The top 5 relative families of consumer No. 9160 are:    [8244 3036 2755 2946
9600]
The top 5 relative families of consumer No. 9278 are:    [8155 5785 9639 1185
1086]
The top 5 relative families of consumer No. 9295 are:    [9121  739 1086 3723
4356]
The top 5 relative families of consumer No. 9474 are:    [2638 4352 4421 7460
7919]
The top 5 relative families of consumer No. 9600 are:    [3036 5317 2755 8244
9160]
The top 5 relative families of consumer No. 9620 are:    [6863 7739 5545 3310
6836]
The top 5 relative families of consumer No. 9631 are:    [2980 1714 3310 6836
2461]
The top 5 relative families of consumer No. 9639 are:    [1086 3134 5972 1790
5785]
The top 5 relative families of consumer No. 9729 are:    [4031 2034 3544 1801
8829]
The top 5 relative families of consumer No. 9766 are:    [4296 8086 2018 1801
187]
The top 5 relative families of consumer No. 9849 are:    [  77   35 1697 2449
44]
The top 5 relative families of consumer No. 9956 are:    [7429 2945 7900  744
5131]
The top 5 relative families of consumer No. 9982 are:    [5403  484 4296 1042
9766]
```

For a better performance, here we use the hourly useage for each meter.

In [9]:
```
month = 10
```

In [10]:
```
from sklearn.cluster import KMeans
from sklearn.mixture import GaussianMixture
if month >= 10:
    start_hour = (month - 10)*728
    end_hour = start_hour + 728 - 1
if month <10:
    start_hour = (month + 12 - 10)*728
    end_hour = start_hour + 728 - 1


hourly_data = hourly_data[:,start_hour:end_hour]

meter_num,hour_num = hourly_data.shape
hourly_usage = np.zeros(shape=(meter_num,hour_num))

for i in range(157):
    hourly_usage[i] = hourly_data[i] - hourly_data[i][0]
```

Let start from K-means. By default, we set the number of clusters as 4.

In [11]:

```
Clusters = 4

kmeans = KMeans(n_clusters=Clusters)
kmeans.fit(hourly_usage)
cluster = kmeans.predict(hourly_usage)
cluster = np.c_[cluster,keys]
Clus = []

for i in range(Clusters):
    idx = np.where(cluster[:,0] == i)
    l = cluster[idx[0]][:,1]
    Clus.append(list(l))

print(Clus)
YK = []
Not_match_num = 0
for i in range(len(Top5)):
    print('For meted id '+str(Top5[i][0])+':')
    Idx =  np.where(cluster[:,1] == Top5[i][0])
    label = cluster[Idx,0]
    No = []
    yk = []
    for j in Top5[i][1]:
        if j not in Clus[label[0][0]]:
            No.append(j)
            Not_match_num = Not_match_num + 1
        else:
            yk.append(j)
    YK.append(yk)
    if No == []:
        print('All its top 5 correlated meters are in the same clusters')
    else:
        print('The meter(s) is(are) of top 5 correlated but not the same clust
er:')
        print(No)
    print('------------------------')


print('##################')
print('##################')
print('##################')
print(str(round(100-100*Not_match_num/(157*5),3))+
      '% of the K-means result can match with top 5 correlation')
```

```
[[35, 44, 77, 94, 114, 252, 483, 484, 661, 739, 744, 1185, 1556, 1589, 1619, 1
714, 1791, 1792, 1800, 1801, 2034, 2072, 2094, 2233, 2335, 2378, 2470, 2575, 2
638, 2645, 2980, 3039, 3134, 3367, 3527, 3577, 3778, 3849, 3893, 4031, 4193, 4
296, 4352, 4356, 4373, 4421, 4447, 4732, 4767, 4998, 5129, 5131, 5275, 5403, 5
439, 5636, 5785, 5814, 6412, 6578, 6830, 6836, 6863, 7017, 7429, 7741, 7900, 7
965, 8084, 8155, 8156, 8467, 8703, 8829, 8890, 9052, 9134, 9278, 9631, 9639, 9
729, 9849, 9956, 9982], [222], [187, 1042, 1086, 1283, 1415, 1507, 1697, 1718,
1790, 2018, 2129, 2449, 3723, 4029, 4514, 5193, 5484, 5810, 5892, 5972, 6910,
7030, 7287, 7460, 7674, 7682, 7794, 7919, 8086, 9121, 9295, 9474, 9766], [370,
871, 1103, 1403, 2461, 2755, 2814, 2818, 2945, 2946, 2965, 3036, 3310, 3544, 3
635, 3918, 4228, 4671, 4874, 5317, 5395, 5545, 5658, 6101, 6505, 6673, 6685, 7
016, 7117, 7566, 7739, 7989, 8059, 8244, 8386, 8967, 9160, 9600, 9620]]
For meted id 35:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2449, 1697, 5810, 7287]
------------------------
For meted id 44:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1697, 1283, 2449]
------------------------
For meted id 77:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1697, 2449]
------------------------
For meted id 94:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5892, 1507]
------------------------
For meted id 114:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4514]
------------------------
For meted id 187:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4296]
------------------------
For meted id 222:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7965, 44, 1283, 2965, 77]
------------------------
For meted id 252:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5395]
------------------------
For meted id 370:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1589, 3527, 6836, 1714]
------------------------
For meted id 483:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 484:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5484, 5892, 5810]
------------------------
For meted id 661:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

[7794, 3723, 4514]

For meted id 739:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 3723, 9295]

For meted id 744:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]

For meted id 871:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8890, 3577, 9052, 5131, 2072]

For meted id 1042:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9982, 4296]

For meted id 1086:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4356]

For meted id 1103:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830, 9631]

For meted id 1185:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5972, 5395, 3918, 1086]

For meted id 1283:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[484, 44, 9982]

For meted id 1403:
All its top 5 correlated meters are in the same clusters

For meted id 1415:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[744, 2470, 3527]

For meted id 1507:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5636, 2094, 3778, 5129]

For meted id 1556:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4514]

For meted id 1589:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[370, 7117]

For meted id 1619:
All its top 5 correlated meters are in the same clusters

For meted id 1697:

```
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35, 484, 5403]
------------------------
For meted id 1714:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8386, 7117]
------------------------
For meted id 1718:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 1791, 114, 2094]
------------------------
For meted id 1790:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5814, 3134, 661, 739]
------------------------
For meted id 1791:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1718]
------------------------
For meted id 1792:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 1800:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5193, 1042]
------------------------
For meted id 1801:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1718, 1507, 8086]
------------------------
For meted id 2018:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7741, 4296, 1801]
------------------------
For meted id 2034:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 2072:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 2094:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1507, 4514]
------------------------
For meted id 2129:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 9134]
------------------------
For meted id 2233:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7919, 7460]
------------------------
For meted id 2335:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7989, 6910]
------------------------
```

```
For meted id 2378:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7989]
------------------------
For meted id 2449:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35, 484]
------------------------
For meted id 2461:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[483, 4767, 4193, 3367]
------------------------
For meted id 2470:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8086]
------------------------
For meted id 2575:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7989, 5892, 1507]
------------------------
For meted id 2638:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 7460, 7919]
------------------------
For meted id 2645:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7682]
------------------------
For meted id 2755:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 2814:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 4352, 2638]
------------------------
For meted id 2818:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3367, 5439, 7017, 8829, 483]
------------------------
For meted id 2945:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9052, 3577, 8829, 1556]
------------------------
For meted id 2946:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 2965:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1415, 2378, 5193]
------------------------
For meted id 2980:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5193]
------------------------
For meted id 3036:
All its top 5 correlated meters are in the same clusters
------------------------
```

```
For meted id 3039:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7919, 7460]
------------------------
For meted id 3134:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1790, 1086, 3723]
------------------------
For meted id 3310:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9631, 1714]
------------------------
For meted id 3367:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3527:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3544:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8829, 2034, 4031, 9729, 5636]
------------------------
For meted id 3577:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3635:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 5814, 7794, 8890, 9134]
------------------------
For meted id 3723:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 5814]
------------------------
For meted id 3778:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1507, 6910, 5892]
------------------------
For meted id 3849:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3893:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 4514]
------------------------
For meted id 3918:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[252, 8890, 3577, 5814]
------------------------
For meted id 4029:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8156, 3849, 4296]
------------------------
For meted id 4031:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 4193:
All its top 5 correlated meters are in the same clusters
```

```
------------------------
For meted id 4228:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4514, 7794, 2094, 1556, 3893]
------------------------
For meted id 4296:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9766, 8086, 2018]
------------------------
For meted id 4352:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 7460, 7919]
------------------------
For meted id 4356:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5972, 1086, 9121, 3723]
------------------------
For meted id 4373:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6910, 1507]
------------------------
For meted id 4421:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7460, 7919, 9474]
------------------------
For meted id 4447:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8967]
------------------------
For meted id 4514:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1556, 3893, 4228, 114]
------------------------
For meted id 4671:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6863, 4767]
------------------------
For meted id 4732:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6910]
------------------------
For meted id 4767:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2461]
------------------------
For meted id 4874:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[222, 7965, 9474]
------------------------
For meted id 4998:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 5129:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1507, 6910]
------------------------
```

```
For meted id 5131:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 5193:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2470, 3527, 3849, 1589]
------------------------
For meted id 5275:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[871]
------------------------
For meted id 5317:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 5395:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[252, 5814, 661, 8890]
------------------------
For meted id 5403:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5810, 9766, 5892]
------------------------
For meted id 5439:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 5484:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[484, 2575]
------------------------
For meted id 5545:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830]
------------------------
For meted id 5636:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1507, 6910]
------------------------
For meted id 5658:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[744, 7900, 2335]
------------------------
For meted id 5785:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1086, 5972]
------------------------
For meted id 5810:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5403, 35]
------------------------
For meted id 5814:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 3723]
------------------------
For meted id 5892:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3778, 2094, 5129]
```

```
------------------------
For meted id 5972:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4356, 5395, 5814, 661]
------------------------
For meted id 6101:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 6412:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 6505:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1589, 2378, 744]
------------------------
For meted id 6578:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7682, 8967]
------------------------
For meted id 6673:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 4352]
------------------------
For meted id 6685:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2378, 744, 1589]
------------------------
For meted id 6830:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8386]
------------------------
For meted id 6836:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[370, 7117]
------------------------
For meted id 6863:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9620, 7739, 3310, 1103]
------------------------
For meted id 6910:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5636, 4732, 4373, 3778]
------------------------
For meted id 7016:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1714, 483, 4767]
------------------------
For meted id 7017:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 7030:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8156, 4296]
------------------------
For meted id 7117:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

```
[1589, 3527, 1714, 6836]
------------------------
For meted id 7287:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 7429:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 7460:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4421, 2638, 4352]
------------------------
For meted id 7566:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 7674:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 9729, 3544]
------------------------
For meted id 7682:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35]
------------------------
For meted id 7739:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830, 9631, 6836]
------------------------
For meted id 7741:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8086, 2018, 9766]
------------------------
For meted id 7794:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3893, 661, 4228]
------------------------
For meted id 7900:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 7919:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4421, 2638, 4352]
------------------------
For meted id 7965:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2965, 1283, 1697, 5484]
------------------------
For meted id 7989:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2335, 5636, 2018, 2575, 1801]
------------------------
For meted id 8059:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4031, 3577, 2034, 1619, 8890]
------------------------
For meted id 8084:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

[7674, 8086, 1718]

For meted id 8086:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 7741, 4296]

For meted id 8155:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5972]

For meted id 8156:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7030, 9766]

For meted id 8244:
All its top 5 correlated meters are in the same clusters

For meted id 8386:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1714, 3527, 5439, 2470, 3849]

For meted id 8467:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1283]

For meted id 8703:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3310, 370, 1103]

For meted id 8829:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]

For meted id 8890:
All its top 5 correlated meters are in the same clusters

For meted id 8967:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6578, 4447, 9278, 8155, 2645]

For meted id 9052:
All its top 5 correlated meters are in the same clusters

For meted id 9121:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 4356]

For meted id 9134:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6910]

For meted id 9160:
All its top 5 correlated meters are in the same clusters

For meted id 9278:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1086]

```
------------------------
For meted id 9295:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[739, 4356]
------------------------
For meted id 9474:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2638, 4352, 4421]
------------------------
For meted id 9600:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 9620:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6863, 6836]
------------------------
For meted id 9631:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3310, 2461]
------------------------
For meted id 9639:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1086, 5972, 1790]
------------------------
For meted id 9729:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 9766:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4296, 1801]
------------------------
For meted id 9849:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1697, 2449]
------------------------
For meted id 9956:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2945]
------------------------
For meted id 9982:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1042, 9766]
------------------------
##################
##################
##################
57.325% of the K-means result can match with top 5 correlation
```

From the result we can tell that: 1,For most of the meters, some of its top 5 correlated meters are in the K-Means cluster. 2,For several meters,all of its top 5 correlated meters are in the K-Means cluster. 3,For only a small amount of meters, none of its top 5 correlated meters are in the K-Means cluster.

Now let's try another method, GMM

In [12]:

```
GMM = GaussianMixture(n_components=Clusters,covariance_type='tied',reg_covar=
0.1)
GMM.fit(hourly_usage)
cluster = GMM.predict(hourly_usage)
cluster = np.c_[cluster,keys]
Clus = []

for i in range(Clusters):
    idx = np.where(cluster[:,0] == i)
    l = cluster[idx[0]][:,1]
    Clus.append(list(l))

print(Clus)
YG = []
Not_match_num = 0
for i in range(len(Top5)):
    print('For meted id '+str(Top5[i][0])+':')
    Idx =  np.where(cluster[:,1] == Top5[i][0])
    label = cluster[Idx,0]
    No = []
    yg = []
    for j in Top5[i][1]:
        if j not in Clus[label[0][0]]:
            No.append(j)
            Not_match_num = Not_match_num + 1
        else:
            yg.append(j)
    YG.append(yg)
    if No == []:
        print('All its top 5 correlated meters are in the same clusters')
    else:
        print('The meter(s) is(are) of top 5 correlated but not the same clust
er:')
        print(No)
    print('------------------------')


print('##################')
print('##################')
print('##################')
print(str(round(100-100*Not_match_num/(157*5),3))+
      '% of the GMM result can match with top 5 correlation')
```

```
[[187, 484, 1042, 1086, 1283, 1415, 1507, 1697, 1718, 1790, 1800, 2018, 2094,
2129, 2449, 3723, 4029, 4296, 4356, 4514, 5193, 5403, 5484, 5636, 5810, 5892,
5972, 6910, 7030, 7287, 7460, 7674, 7682, 7794, 7919, 8086, 9121, 9295, 9474,
9766, 9982], [222], [35, 44, 77, 94, 114, 252, 483, 661, 739, 744, 1185, 1556,
1589, 1619, 1714, 1791, 1792, 1801, 2034, 2072, 2233, 2335, 2378, 2461, 2470,
2575, 2638, 2645, 2980, 3039, 3134, 3367, 3527, 3577, 3778, 3849, 3893, 4031,
4193, 4228, 4352, 4373, 4421, 4447, 4732, 4767, 4998, 5129, 5131, 5275, 5395,
5439, 5785, 5814, 6412, 6578, 6830, 6836, 6863, 7017, 7117, 7429, 7741, 7900,
7965, 8084, 8155, 8156, 8467, 8703, 8829, 8890, 9052, 9134, 9278, 9631, 9639,
9729, 9849, 9956], [370, 871, 1103, 1403, 2755, 2814, 2818, 2945, 2946, 2965,
3036, 3310, 3544, 3635, 3918, 4671, 4874, 5317, 5545, 5658, 6101, 6505, 6673,
6685, 7016, 7566, 7739, 7989, 8059, 8244, 8386, 8967, 9160, 9600, 9620]]
For meted id 35:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2449, 1697, 5810, 7287, 5403]
------------------------
For meted id 44:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1697, 1283, 2449, 484]
------------------------
For meted id 77:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1697, 2449]
------------------------
For meted id 94:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5892, 1507]
------------------------
For meted id 114:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4514, 2094]
------------------------
For meted id 187:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 222:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7965, 44, 1283, 2965, 77]
------------------------
For meted id 252:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 370:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7117, 1589, 3527, 6836, 1714]
------------------------
For meted id 483:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 484:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 661:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 3723, 4514]
------------------------
For meted id 739:
```

```
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 3723, 9295]
------------------------
For meted id 744:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 871:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8890, 3577, 9052, 5131, 2072]
------------------------
For meted id 1042:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 1086:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 1103:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830, 9631]
------------------------
For meted id 1185:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5972, 4356, 3918, 1086]
------------------------
For meted id 1283:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[44]
------------------------
For meted id 1403:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 1415:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[744, 2470, 3527]
------------------------
For meted id 1507:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3778, 5129]
------------------------
For meted id 1556:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4514, 2094]
------------------------
For meted id 1589:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[370]
------------------------
For meted id 1619:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 1697:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35]
------------------------
For meted id 1714:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

[8386]

For meted id 1718:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 1791, 114]

For meted id 1790:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5814, 3134, 661, 739]

For meted id 1791:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1718, 2094]

For meted id 1792:
All its top 5 correlated meters are in the same clusters

For meted id 1800:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830, 2470, 3849]

For meted id 1801:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5636, 1718, 1507, 8086]

For meted id 2018:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7741, 1801]

For meted id 2034:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]

For meted id 2072:
All its top 5 correlated meters are in the same clusters

For meted id 2094:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1556, 114]

For meted id 2129:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 9134]

For meted id 2233:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7919, 7460]

For meted id 2335:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7989, 6910, 5636]

For meted id 2378:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7989]

For meted id 2449:

```
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35]
------------------------
For meted id 2461:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7016]
------------------------
For meted id 2470:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8086]
------------------------
For meted id 2575:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7989, 5892, 1507]
------------------------
For meted id 2638:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 7460, 7919]
------------------------
For meted id 2645:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7682]
------------------------
For meted id 2755:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 2814:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 4352, 2638]
------------------------
For meted id 2818:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3367, 5439, 7017, 8829, 483]
------------------------
For meted id 2945:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9052, 3577, 8829, 1556]
------------------------
For meted id 2946:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 2965:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1415, 2378, 5193]
------------------------
For meted id 2980:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5193]
------------------------
For meted id 3036:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3039:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7919, 7460]
------------------------
For meted id 3134:
```

```
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1790, 1086, 3723]
------------------------
For meted id 3310:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9631, 1714, 2461, 7117]
------------------------
For meted id 3367:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3527:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3544:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8829, 2034, 4031, 9729, 5636]
------------------------
For meted id 3577:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3635:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 5814, 7794, 8890, 9134]
------------------------
For meted id 3723:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661, 5814]
------------------------
For meted id 3778:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1507, 5636, 6910, 2094, 5892]
------------------------
For meted id 3849:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 3893:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 4514]
------------------------
For meted id 3918:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5395, 252, 8890, 3577, 5814]
------------------------
For meted id 4029:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8156, 3849]
------------------------
For meted id 4031:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 4193:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 4228:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4514, 7794, 2094]
------------------------
```

```
For meted id 4296:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7741, 1801]
------------------------
For meted id 4352:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 7460, 7919]
------------------------
For meted id 4356:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661]
------------------------
For meted id 4373:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2094, 6910, 1507]
------------------------
For meted id 4421:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7460, 7919, 9474]
------------------------
For meted id 4447:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8967]
------------------------
For meted id 4514:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1556, 3893, 4228, 114]
------------------------
For meted id 4671:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6863, 2461, 4767]
------------------------
For meted id 4732:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6910]
------------------------
For meted id 4767:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 4874:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[222, 7965, 9474]
------------------------
For meted id 4998:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 5129:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1507, 6910, 2094]
------------------------
For meted id 5131:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 5193:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

[2470, 3527, 3849, 1589]

For meted id 5275:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[871]

For meted id 5317:
All its top 5 correlated meters are in the same clusters

For meted id 5395:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3918]

For meted id 5403:
All its top 5 correlated meters are in the same clusters

For meted id 5439:
All its top 5 correlated meters are in the same clusters

For meted id 5484:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2575]

For meted id 5545:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830]

For meted id 5636:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 3778]

For meted id 5658:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[744, 7900, 2335]

For meted id 5785:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1086, 5972]

For meted id 5810:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35]

For meted id 5814:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7794, 3723]

For meted id 5892:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3778, 5129]

For meted id 5972:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5395, 5814, 661]

For meted id 6101:
All its top 5 correlated meters are in the same clusters

```
------------------------
For meted id 6412:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 6505:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1589, 7117, 2378, 744]
------------------------
For meted id 6578:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7682, 8967]
------------------------
For meted id 6673:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9474, 4352]
------------------------
For meted id 6685:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2378, 744, 1589]
------------------------
For meted id 6830:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8386]
------------------------
For meted id 6836:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[370]
------------------------
For meted id 6863:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[9620, 7739, 3310, 1103]
------------------------
For meted id 6910:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4732, 4373, 3778]
------------------------
For meted id 7016:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2461, 1714, 483, 4767]
------------------------
For meted id 7017:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 7030:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8156]
------------------------
For meted id 7117:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[370]
------------------------
For meted id 7287:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 7429:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

[3544]

For meted id 7460:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4421, 2638, 4352]

For meted id 7566:
All its top 5 correlated meters are in the same clusters

For meted id 7674:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 9729, 3544]

For meted id 7682:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[35]

For meted id 7739:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6830, 9631, 6836]

For meted id 7741:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[8086, 2018, 4296, 9766]

For meted id 7794:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3893, 661, 4228]

For meted id 7900:
All its top 5 correlated meters are in the same clusters

For meted id 7919:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4421, 2638, 4352]

For meted id 7965:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2965, 1283, 1697, 5484]

For meted id 7989:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2335, 5636, 2018, 2575, 1801]

For meted id 8059:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[4031, 3577, 2034, 1619, 8890]

For meted id 8084:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7674, 8086, 1718]

For meted id 8086:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801, 7741]

For meted id 8155:

```
The meter(s) is(are) of top 5 correlated but not the same cluster:
[5972]
------------------------
For meted id 8156:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[7030, 4296, 9766]
------------------------
For meted id 8244:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 8386:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1714, 3527, 5439, 2470, 3849]
------------------------
For meted id 8467:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1283]
------------------------
For meted id 8703:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3310, 370, 1103]
------------------------
For meted id 8829:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 8890:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 8967:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6578, 4447, 9278, 8155, 2645]
------------------------
For meted id 9052:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 9121:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[661]
------------------------
For meted id 9134:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6910]
------------------------
For meted id 9160:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 9278:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1086]
------------------------
For meted id 9295:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[739]
------------------------
For meted id 9474:
The meter(s) is(are) of top 5 correlated but not the same cluster:
```

```
[2638, 4352, 4421]
------------------------
For meted id 9600:
All its top 5 correlated meters are in the same clusters
------------------------
For meted id 9620:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[6863, 6836]
------------------------
For meted id 9631:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3310]
------------------------
For meted id 9639:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1086, 5972, 1790]
------------------------
For meted id 9729:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[3544]
------------------------
For meted id 9766:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1801]
------------------------
For meted id 9849:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[1697, 2449]
------------------------
For meted id 9956:
The meter(s) is(are) of top 5 correlated but not the same cluster:
[2945]
------------------------
For meted id 9982:
All its top 5 correlated meters are in the same clusters
------------------------
##################
##################
##################
61.401% of the GMM result can match with top 5 correlation
```

From the result we can tell that: 1,For most of the meters, some of its top 5 correlated meters are in the GMM cluster. 2,For several meters,all of its top 5 correlated meters are in the GMM cluster. 3,For only a small amount of meters, none of its top 5 correlated meters are in the GMM cluster.

Now let's see which top 5 correlated meter(s) is(are) in the both K-means cluster and GMM cluster.

In [13]:

```
The_same = []
for i in range(len(YG)):
    l = []
    l.append(cluster[i,1])
    if YG[i] != []:
        if YK[i]!= []:
            the_same = [x for x in YG[i] if x in YK[i]]
            if the_same != []:
                print('Based on correlation,K-means and GMM')
                print('MeterID:'+ str(cluster[i,1]))
                print('and')
                print(the_same)
                print('are in the both K-Means and GMM clusters')
                print(' ')
                l.append(the_same)
    The_same.append(l)

for i in range(157):
    if len(The_same[i]) == 1:
        The_same[i].append([])


TOP5 = []
ALL = []
for i in range(157):
    TOP5.append(Top5[i][1])
    ALL.append(The_same[i][1])

all_match = 0
for i in range(157):
    all_match = all_match + len(ALL[i])


print(str(round(100*Not_match_num/(157*5),3))+
      '% result of these 3 methods can match up')
```

Based on correlation,K-means and GMM
MeterID:44
and
[35]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:77
and
[9849, 35, 44]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:94
and
[2575, 9134, 5129]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:114
and
[1556, 3893, 1619]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:187
and
[9766, 8086, 1507, 2018]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:252
and
[3577, 4732, 1556, 5814]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:483
and
[3367, 4193, 5439, 8829, 7017]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:484
and
[5403, 9982]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:661
and
[5814, 739]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:739
and

[661, 5814]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:744
and
[7900, 8829, 4998, 3527]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1042
and
[7030, 9766, 5193]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1086
and
[5972, 9295, 9121, 3723]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1103
and
[3310, 8386, 7016]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1283
and
[1697, 5484]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1403
and
[9600, 5317, 3036, 2755, 8244]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1415
and
[8086, 2018]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1507
and
[6910]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1556
and
[4732, 114, 4031]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1589
and
[3527, 1714, 5439]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1619
and
[1556, 4031, 114, 2072, 3577]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1697
and
[2449, 5810]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1714
and
[5439, 3527, 1589]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1718
and
[4514]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1790
and
[3723]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1791
and
[114, 1801, 1556]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1792
and
[1619, 114, 4031, 2034, 2072]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:1801
and
[9729]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2018
and

[8086, 9766]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2034
and
[4031, 8829, 9729, 6412]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2072
and
[4031, 1619, 3577, 2034, 4732]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2094
and
[5636]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2129
and
[7794, 3723, 9121]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2233
and
[3039, 4421, 9956]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2335
and
[9134, 4732]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2378
and
[7017, 2335, 2470, 1589]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2449
and
[5810, 1697, 5484]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2470
and
[3527, 7017, 5439, 3849]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2575
and
[94, 3778]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2638
and
[4352, 4421]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2645
and
[9278, 8155, 6578, 1185]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2755
and
[3036, 8244, 9600, 9160, 2946]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2814
and
[6673, 6101]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2945
and
[3544]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2946
and
[8244, 3036, 9160, 2755, 7566]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2965
and
[6505, 5658]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:2980
and
[9631, 6836, 1714, 3527]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3036
and

[9600, 2755, 8244, 9160, 5317]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3039
and
[2233, 4421, 2638]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3134
and
[9639, 739]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3310
and
[1103]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3367
and
[483, 5439, 7017, 8829, 4193]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3527
and
[1589, 2470, 5439, 7017, 1714]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3577
and
[9052, 8890, 4732, 1556, 4031]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3723
and
[9121, 7794, 4514]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3849
and
[2470, 7017, 5439, 3367, 3527]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:3893
and
[1556, 5814, 661]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4029
and
[7030, 1042]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4031
and
[2034, 1556, 2072, 9729, 1619]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4193
and
[483, 3367, 7900, 5439, 8829]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4352
and
[2638, 4421]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4373
and
[4732, 1556]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4421
and
[2638, 4352]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4447
and
[2645, 6578, 9278, 8155]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4514
and
[7794]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4671
and
[7016, 3310]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4732
and

[1556, 9052, 3577, 4373]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4767
and
[483, 2072, 4031, 2034]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4874
and
[2814, 6673]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:4998
and
[7900, 8829, 5439, 483]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5129
and
[4373, 4732]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5131
and
[9052, 8890, 8829, 2335]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5193
and
[1415]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5275
and
[2072, 2034, 1619, 7429]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5317
and
[9600, 3036, 2755, 8244, 9160]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5403
and
[9982, 484]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5439
and
[3367, 7900, 483, 3527, 7017]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5484
and
[5892, 2449, 5810]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5545
and
[7739, 9620, 1103, 5317]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5636
and
[2094]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5658
and
[6505, 7989]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5785
and
[9639, 3134, 8155]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5810
and
[2449, 7287, 9121]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5814
and
[661, 3893, 252]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5892
and
[1507, 6910]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:5972
and

[1086]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6101
and
[2946, 9160, 8244, 7566, 2755]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6412
and
[2034, 8829, 3367, 7429]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6505
and
[370]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6578
and
[9278, 8155, 5785]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6673
and
[2814, 5545, 6101]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6685
and
[6505, 370]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6830
and
[2980, 9631, 1714, 6836]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6836
and
[1589, 3527, 1714]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6863
and
[9631]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:6910
and
[1507]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7016
and
[8386]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7017
and
[3367, 5439, 2470, 3527, 483]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7030
and
[9766, 1042, 8086]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7287
and
[9295, 5810, 9121, 1086, 7682]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7429
and
[8829, 2034, 4031, 6412]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7460
and
[7919, 9474]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7566
and
[8244, 2946, 3036, 2755, 9160]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7674
and
[8086, 2018]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7682
and

[7287, 1086, 9121, 9295]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7739
and
[1103, 3310]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7741
and
[1801]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7794
and
[4514, 3723]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7900
and
[744, 5439, 4998, 8829, 3527]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7919
and
[7460, 9474]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:7965
and
[44]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8084
and
[1791, 1801]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8086
and
[2018, 9766]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8155
and
[5785, 1185, 9278, 9639]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8156
and
[7741, 1801]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8244
and
[3036, 2946, 9160, 2755, 9600]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8467
and
[77, 9849, 2233, 7965]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8703
and
[9631, 6836]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8829
and
[2034, 4031, 9729, 7429]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:8890
and
[9052, 3577, 4732, 252, 1556]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9052
and
[3577, 8890, 4732, 1556, 4031]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9121
and
[3723, 9295, 7794]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9134
and
[4373, 661, 2335, 4732]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9160
and

[8244, 3036, 2755, 2946, 9600]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9278
and
[8155, 5785, 9639, 1185]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9295
and
[9121, 1086, 3723]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9474
and
[7460, 7919]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9600
and
[3036, 5317, 2755, 8244, 9160]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9620
and
[7739, 5545, 3310]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9631
and
[2980, 1714, 6836]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9639
and
[3134, 5785]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9729
and
[4031, 2034, 1801, 8829]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9766
and
[8086, 2018, 187]
are in the both K-Means and GMM clusters

```
Based on correlation,K-means and GMM
MeterID:9849
and
[77, 35, 44]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9956
and
[7429, 7900, 744, 5131]
are in the both K-Means and GMM clusters

Based on correlation,K-means and GMM
MeterID:9982
and
[5403, 484, 4296]
are in the both K-Means and GMM clusters

38.599% result of these 3 methods can match up
```

We can tell that even we restrict the measurements, the meter can still match a lot.You can also see the summary table below

In [14]:
```
meter_close = {'meterid':keys,'TOP5':TOP5,'K-means':YK,'GMM':YG,'all':ALL}
meter_close = pd.DataFrame(meter_close)

print(meter_close)
```

```
     meterid                              TOP5                   K-means  \
0         35  [2449, 1697, 5810, 7287, 5403]                    [5403]
1         44     [1697, 1283, 35, 2449, 484]                 [35, 484]
2         77      [9849, 1697, 35, 44, 2449]             [9849, 35, 44]
3         94  [2575, 5892, 9134, 5129, 1507]         [2575, 9134, 5129]
4        114  [1556, 4514, 2094, 3893, 1619]  [1556, 2094, 3893, 1619]
..       ...                             ...                       ...
152     9729  [4031, 2034, 3544, 1801, 8829]  [4031, 2034, 1801, 8829]
153     9766   [4296, 8086, 2018, 1801, 187]         [8086, 2018, 187]
154     9849        [77, 35, 1697, 2449, 44]              [77, 35, 44]
155     9956   [7429, 2945, 7900, 744, 5131]   [7429, 7900, 744, 5131]
156     9982   [5403, 484, 4296, 1042, 9766]         [5403, 484, 4296]

                              GMM                      all
0                              []                       []
1                            [35]                     [35]
2                [9849, 35, 44]           [9849, 35, 44]
3             [2575, 9134, 5129]        [2575, 9134, 5129]
4             [1556, 3893, 1619]        [1556, 3893, 1619]
..                           ...                      ...
152     [4031, 2034, 1801, 8829]  [4031, 2034, 1801, 8829]
153      [4296, 8086, 2018, 187]         [8086, 2018, 187]
154                  [77, 35, 44]              [77, 35, 44]
155      [7429, 7900, 744, 5131]   [7429, 7900, 744, 5131]
156  [5403, 484, 4296, 1042, 9766]         [5403, 484, 4296]

[157 rows x 5 columns]
```

In [ ]: