import the required libraries

In [1]:

```
import numpy as np
import pandas as pd
```

read the dataset and get the stats

In [2]:

```
d_csv=pd.read_csv('5_d.csv')
d_csv.head(10)
d_csv.describe()
```

Out[2]:

	y	pred
count	157200.000000	157200.000000
mean	66.562087	66.483715
std	63.714792	62.340840
min	0.000000	-6.000000
25%	11.000000	12.000000
50%	52.000000	52.000000
75%	104.000000	103.000000
max	440.000000	421.000000

1)compute the mean squared error for each point

In [3]:

```
from tqdm import tqdm
for i in tqdm(range(0,len(d_csv))):

    d_csv['pred'].loc[i]=((d_csv.y[i])-(d_csv.pred[i]))*((d_csv.y[i])-(d_csv.pred[i]))
    #d_csv['pred'].loc[i]**2
avg=(d_csv['pred'].sum())/len(d_csv)
print('MSE :',avg)
```

```
100%|██████████████████████████████████████████████████████████████████
███| 157200/157200 [00:48<00:00, 3233.35it/s]

MSE : 177.16569974554707
```

In [4]:

```
d_csv.head(10)
```

Out[4]:

	y	pred
0	101.0	1.0
1	120.0	400.0
2	131.0	324.0
3	164.0	1521.0
4	154.0	4.0
5	133.0	400.0
6	148.0	81.0
7	172.0	729.0
8	153.0	81.0
9	162.0	64.0

2)compute mean absolute percentage error

In [6]:

```
d_csv=pd.read_csv('5_d.csv')
d_csv['pred']=abs(d_csv['pred'])
d_csv.describe()
for i in range(0,len(d_csv)):
    d_csv['pred'].loc[i]=abs((d_csv.y[i])-(d_csv.pred[i]))
d_csv.head(10)
#modified_MAPE=mod(sum of errors)/sum of actual y's
errors=d_csv['pred'].sum()
classlabels=d_csv['y'].sum()

MAPE=(errors/classlabels)*100
print('MAPE :',MAPE)
```

```
MAPE : 12.911896142421892
```

3)compute r-squared error

In [10]:

```
from tqdm import tqdm
d_csv=pd.read_csv('5_d.csv')
y_mean=d_csv['y'].sum()/len(d_csv)
ss_tot=0
ss_res=0
def ss_total():
    for i in tqdm(range(0,len(d_csv))):
        global ss_tot
        ss_tot+=(d_csv['y'].loc[i]-y_mean)*(d_csv['y'].loc[i]-y_mean)
    return ss_tot

def ss_residual():
    global ss_res
    for i in tqdm(range(0,len(d_csv))):
        ss_res+=((d_csv.y[i])-(d_csv.pred[i]))*((d_csv.y[i])-(d_csv.pred[i]))
    return ss_res

ss_total()
ss_residual()
r_squared=1-(ss_res/ss_tot)
print('sstotal',ss_tot)

print('ssresidual',ss_res)

print('R_squared error is:',r_squared)
```

```
100%|████████████████████████████████████████████████████████████████
██| 157200/157200 [00:10<00:00, 14874.75it/s]
100%|████████████████████████████████████████████████████████████████
██| 157200/157200 [00:09<00:00, 16545.95it/s]

sstotal 638161080.035662
ssresidual 27850448.0
R_squared error is: 0.9563582786990964
```