In [5]:

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

In [6]:

```
from sklearn.cluster import KMeans
from sklearn import preprocessing # for standardization
```

In [7]:

```
df=pd.read_csv('ai_mid5.csv', header=0)
```

In [10]:

```
df.head()
```

Out[10]:

	age	balance
0	30	1787
1	33	4789
2	35	1350
3	30	1476
4	59	0

In [11]:

```
n_cls = 5 # number of class
km = KMeans(n_clusters=n_cls, random_state=20)
```

In [12]:

```
# 標準化してからk-meansを実行
# k-means clustering with standardized data
cls = km.fit_predict(preprocessing.scale(df))
```

In [13]:

```
pd.value_counts(pd.Series(cls), ).sort_index()
```

Out[13]:

```
0    1988
1     822
2     329
3      48
4    1334
dtype: int64
```

Ans.

1988, 822, 329, 48, 1334