```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.cluster import KMeans

loan=pd.read_csv('loan_data.csv')

loan.head()
```


```
{"summary":"{\n  \"name\": \"loan\",\n  \"rows\": 45000,\n
\"fields\": [\n    {\n      \"column\": \"person_age\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
6.045108211348622,\n        \"min\": 20.0,\n        \"max\": 144.0,\n
\"num_unique_values\": 60,\n        \"samples\": [\n          22.0,\n
26.0,\n          53.0\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"person_gender\",\n      \"properties\": {\n        \"dtype\":
\"category\",\n        \"num_unique_values\": 2,\n        \"samples\":
[\n          \"male\",\n          \"female\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"person_education\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
\"num_unique_values\": 5,\n        \"samples\": [\n          \"High
School\",\n          \"Doctorate\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"person_income\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
80422.49863189556,\n        \"min\": 8000.0,\n        \"max\":
7200766.0,\n        \"num_unique_values\": 33989,\n
\"samples\": [\n          48967.0,\n          31001.0\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"person_emp_exp\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
6,\n        \"min\": 0,\n        \"max\": 125,\n
\"num_unique_values\": 63,\n        \"samples\": [\n          93,\n
76\n        ],\n        \"semantic_type\": \"\",\n
\"description\": \"\"\n      }\n    },\n    {\n      \"column\":
\"person_home_ownership\",\n      \"properties\": {\n
\"dtype\": \"category\",\n        \"num_unique_values\": 4,\n
\"samples\": [\n          \"OWN\",\n          \"OTHER\"\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"loan_amnt\",\n
\"properties\": {\n        \"dtype\": \"number\",\n        \"std\":
6314.8866905411405,\n        \"min\": 500.0,\n        \"max\":
35000.0,\n        \"num_unique_values\": 4483,\n        \"samples\":
[\n          5800.0,\n          28338.0\n        ],\n
\"semantic_type\": \"\",\n        \"description\": \"\"\n      }\
n    },\n    {\n      \"column\": \"loan_intent\",\n
\"properties\": {\n        \"dtype\": \"category\",\n
```

\"num_unique_values\": 6,\n   \"samples\": [\n
\"PERSONAL\",\n   \"EDUCATION\"\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"loan_int_rate\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
2.9788082802254734,\n   \"min\": 5.42,\n   \"max\": 20.0,\n
\"num_unique_values\": 1302,\n   \"samples\": [\n   15.0,\
n   13.45\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"loan_percent_income\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 0.08721230801403355,\n   \"min\":
0.0,\n   \"max\": 0.66,\n   \"num_unique_values\": 64,\n
\"samples\": [\n   0.45,\n   0.54\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"cb_person_cred_hist_length\",\n
\"properties\": {\n   \"dtype\": \"number\",\n   \"std\":
3.8797018451620433,\n   \"min\": 2.0,\n   \"max\": 30.0,\n
\"num_unique_values\": 29,\n   \"samples\": [\n   24.0,\n
25.0\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"credit_score\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 50,\n   \"min\": 390,\n
\"max\": 850,\n   \"num_unique_values\": 340,\n
\"samples\": [\n   492,\n   484\n   ],\n
\"semantic_type\": \"\",\n   \"description\": \"\"\n   }\
n   },\n   {\n   \"column\": \"previous_loan_defaults_on_file\",\
n   \"properties\": {\n   \"dtype\": \"category\",\n
\"num_unique_values\": 2,\n   \"samples\": [\n   \"Yes\",\
n   \"No\"\n   ],\n   \"semantic_type\": \"\",\n
\"description\": \"\"\n   }\n   },\n   {\n   \"column\":
\"loan_status\",\n   \"properties\": {\n   \"dtype\":
\"number\",\n   \"std\": 0,\n   \"min\": 0,\n
\"max\": 1,\n   \"num_unique_values\": 2,\n   \"samples\":
[\n   0,\n   1\n   ],\n   \"semantic_type\":
\"\",\n   \"description\": \"\"\n   }\n   }\n]\
n}","type":"dataframe","variable_name":"loan"}

```
loan.isnull().sum()
```

person_age	0
person_gender	0
person_education	0
person_income	0
person_emp_exp	0
person_home_ownership	0
loan_amnt	0
loan_intent	0
loan_int_rate	0
loan_percent_income	0
cb_person_cred_hist_length	0

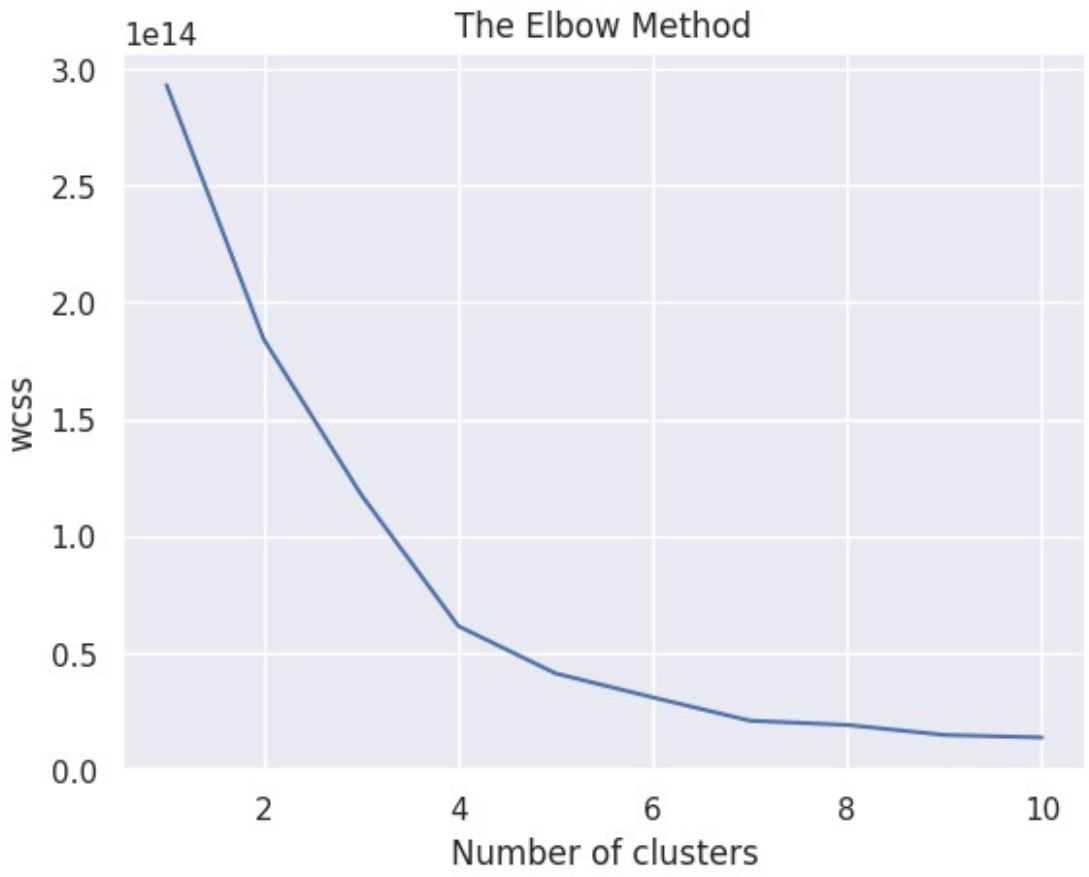
```
credit_score                      0
previous_loan_defaults_on_file    0
loan_status                       0
dtype: int64
```

```
x=loan.iloc[:,[3,6,11]].values
print(x)
```

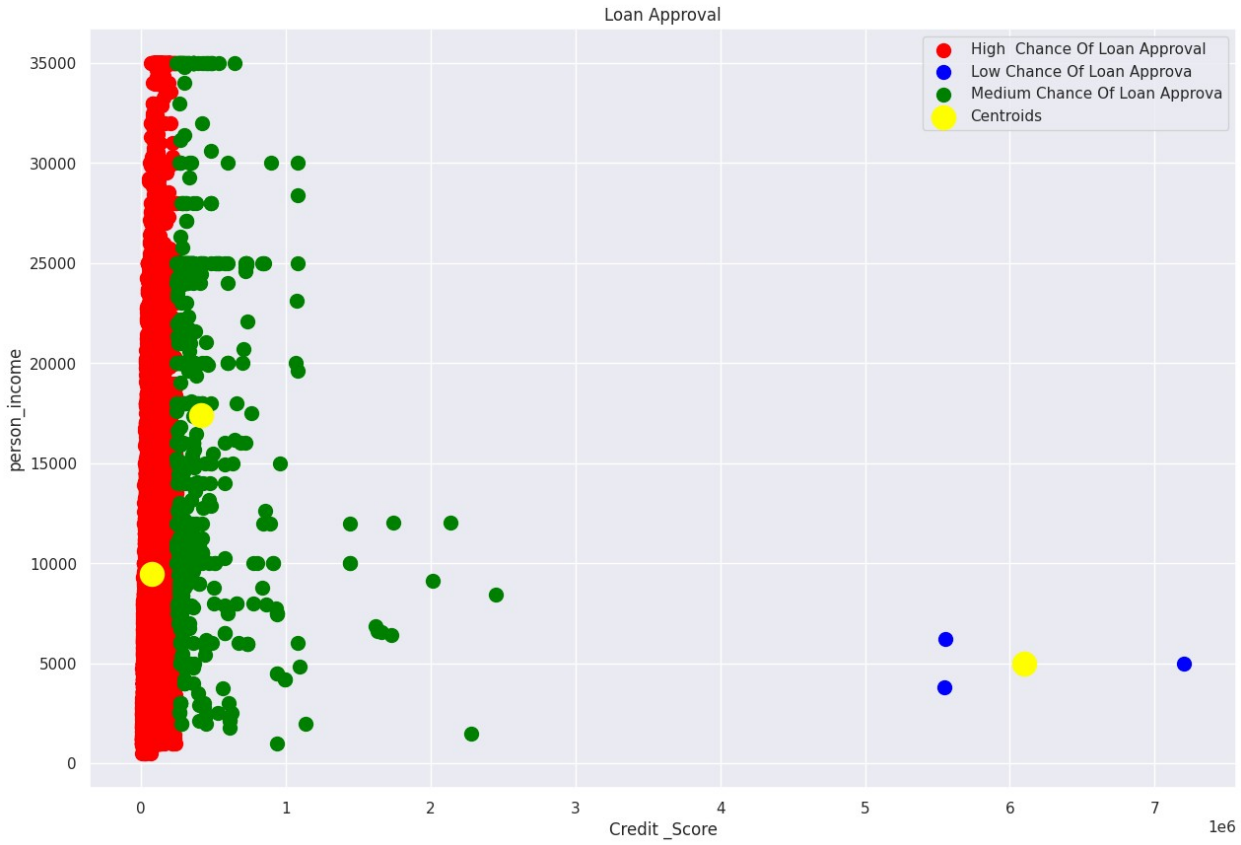
```
[[71948. 35000.   561.]
 [12282.  1000.   504.]
 [12438.  5500.   635.]
 ...
 [56942.  2771.   668.]
 [33164. 12000.   604.]
 [51609.  6665.   628.]]
```

```
wcss=[]
for i in range(1,11):
    kmeans=KMeans(n_clusters=i,init='k-means++',random_state=0)
    kmeans.fit(x)
    wcss.append(kmeans.inertia_)

sns.set()
plt.plot(range(1,11),wcss)
plt.title('The Elbow Method')
plt.xlabel('Number of clusters')
plt.ylabel('wcss')
plt.show()
```




```
kmeans=KMeans(n_clusters=3,init='k-means++',random_state=0)
y_kmeans=kmeans.fit_predict(x)
plt.figure(figsize=(15,10))
plt.scatter(x[y_kmeans==0,0],x[y_kmeans==0,1],s=100,c='red',label='Hig
h  Chance Of Loan Approval')
plt.scatter(x[y_kmeans==1,0],x[y_kmeans==1,1],s=100,c='blue',label='Lo
w Chance Of Loan Approva')
plt.scatter(x[y_kmeans==2,0],x[y_kmeans==2,1],s=100,c='green',label='M
edium Chance Of Loan Approva')
plt.scatter(kmeans.cluster_centers_[:,0],kmeans.cluster_centers_[:,1],
s=300,c='yellow',label='Centroids')
plt.title('Loan Approval')
plt.xlabel('Credit _Score')
plt.ylabel('person_income')
plt.legend()
plt.show()
```


Loan Approval
High Chance Of Loan Approval
Low Chance Of Loan Approva
Medium Chance Of Loan Approva
Centroids
person_income
Credit _Score
1e6
0
5000
10000
15000
20000
25000
30000
35000
0
1
2
3
4
5
6
7