In [1]:

```
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

In [3]:

```
df=pd.read_csv(r"C:\Users\MY HOME\Desktop\Income.csv")
df
```

Out[3]:

	Gender	Age	Income($)
0	Male	19	15
1	Male	21	15
2	Female	20	16
3	Female	23	16
4	Female	31	17
...	...	...	...
195	Female	35	120
196	Female	45	126
197	Male	32	126
198	Male	32	137
199	Male	30	137

200 rows × 3 columns

In [4]:

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 3 columns):
 #   Column     Non-Null Count  Dtype 
---  ------     --------------  ----- 
 0   Gender     200 non-null    object
 1   Age        200 non-null    int64 
 2   Income($)  200 non-null    int64 
dtypes: int64(2), object(1)
memory usage: 4.8+ KB
```

In [5]:

```
1 df.describe()
```

Out[5]:

	Age	Income($)
count	200.000000	200.000000
mean	38.850000	60.560000
std	13.969007	26.264721
min	18.000000	15.000000
25%	28.750000	41.500000
50%	36.000000	61.500000
75%	49.000000	78.000000
max	70.000000	137.000000

In [6]:

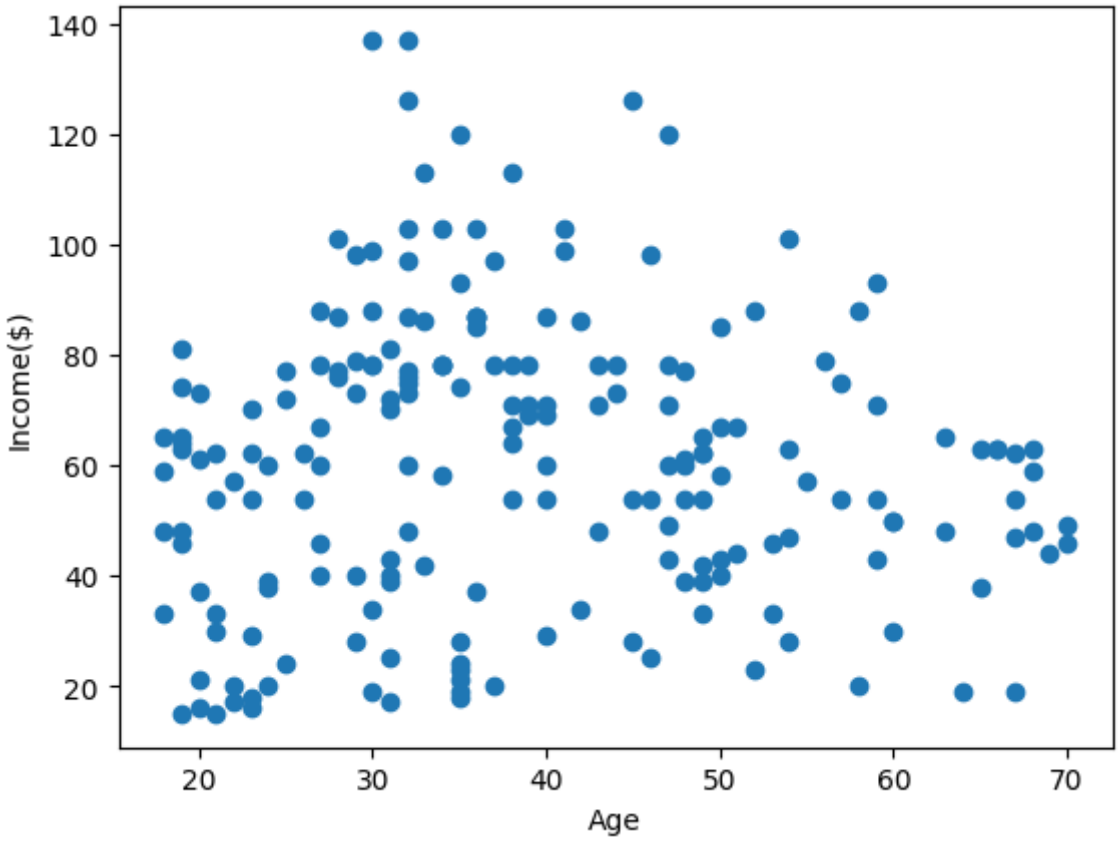
```
1 df.isna().any()
```

Out[6]:

```
Gender       False
Age          False
Income($)    False
dtype: bool
```

In [17]:

```
plt.scatter(df["Age"],df["Income($)"])
plt.xlabel("Age")
plt.ylabel("Income($)")
plt.show()
```



In [18]:

```
from sklearn.cluster import KMeans
```

In [19]:

```
km=KMeans()
km
```

Out[19]:

KMeans()

In [25]:

```
y_predicted=km.fit_predict(df[["Age","Income($)"]])
y_predicted
```

```
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
```

Out[25]:

```
array([0, 0, 0, 0, 0, 0, 0, 0, 2, 0, 2, 0, 2, 0, 0, 0, 0, 0, 2, 0, 0, 0,
       2, 0, 2, 0, 2, 0, 0, 0, 2, 0, 2, 0, 2, 0, 2, 0, 0, 0, 2, 0, 2, 0,
       2, 0, 2, 0, 0, 0, 2, 0, 0, 3, 2, 2, 3, 3, 6, 3, 3, 6, 3, 3, 3, 6,
       3, 3, 6, 6, 3, 3, 3, 3, 3, 6, 3, 7, 6, 3, 3, 6, 3, 3, 6, 3, 3, 6,
       6, 3, 3, 6, 7, 7, 6, 6, 7, 6, 7, 6, 6, 7, 3, 6, 7, 6, 3, 7, 3, 3,
       3, 6, 7, 6, 6, 6, 3, 7, 7, 7, 6, 7, 7, 7, 6, 1, 7, 7, 7, 7, 7, 7,
       1, 1, 1, 1, 7, 1, 1, 1, 7, 1, 1, 1, 1, 1, 7, 1, 1, 1, 7, 1, 7, 1,
       7, 1, 1, 1, 1, 1, 7, 1, 1, 1, 7, 1, 1, 1, 1, 1, 1, 1, 1, 1, 5, 1,
       5, 1, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 4, 4, 4, 4,
       4, 4])
```

In [26]:

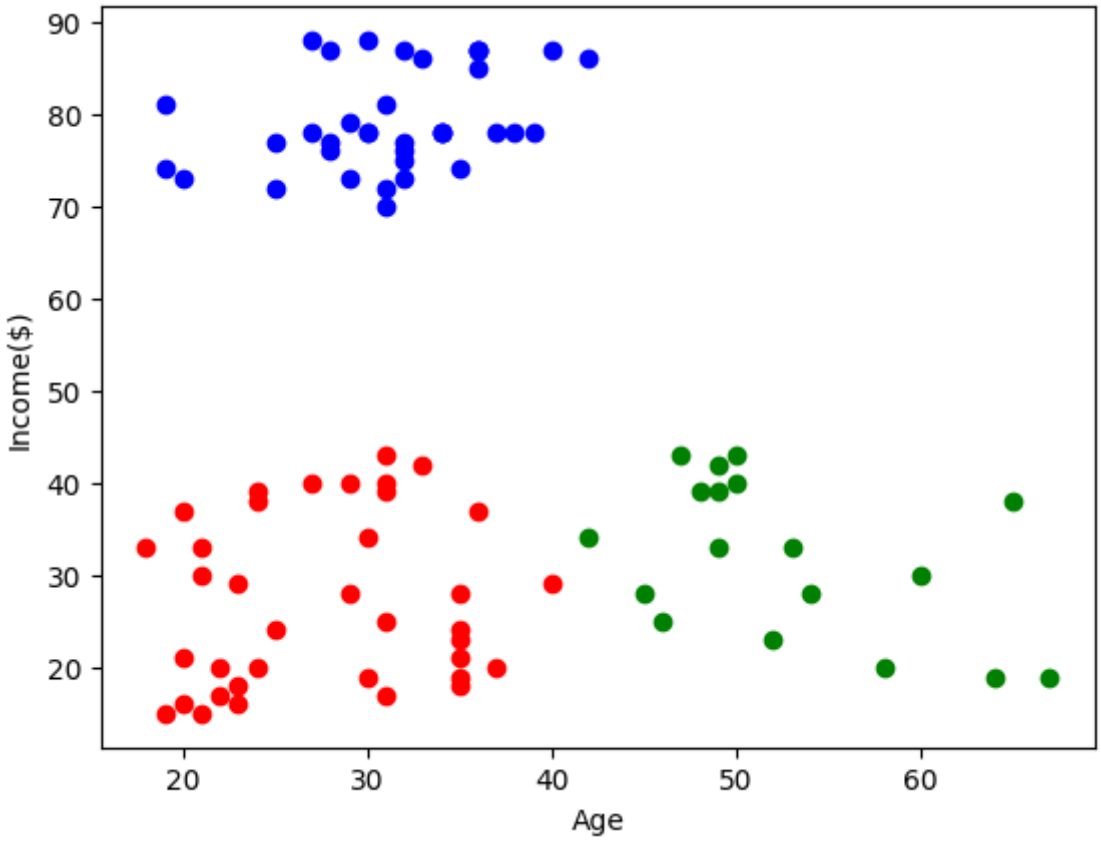
```
df["cluster"]=y_predicted
df.head()
```

Out[26]:

	Gender	Age	Income($)	cluster
0	Male	19	15	0
1	Male	21	15	0
2	Female	20	16	0
3	Female	23	16	0
4	Female	31	17	0

In [34]:

```
df1=df[df.cluster==0]
df2=df[df.cluster==1]
df3=df[df.cluster==2]
plt.scatter(df1["Age"],df1["Income($)"],color="red")
plt.scatter(df2["Age"],df2["Income($)"],color="blue")
plt.scatter(df3["Age"],df3["Income($)"],color="green")
plt.xlabel("Age")
plt.ylabel("Income($)")
plt.show()
```



In [36]:

```
from sklearn.preprocessing import MinMaxScaler
```

In [38]:

```
scaler=MinMaxScaler()
scaler.fit(df[["Income($)"]])
d=scaler.transform(df[["Income($)"]])
df.head()
```

Out[38]:

	Gender	Age	Income($)	cluster
0	Male	19	15	0
1	Male	21	15	0
2	Female	20	16	0
3	Female	23	16	0
4	Female	31	17	0

In [39]:

```
scaler.fit(df[["Age"]])
d=scaler.transform(df[["Age"]])
df.head()
```

Out[39]:

	Gender	Age	Income($)	cluster
0	Male	19	15	0
1	Male	21	15	0
2	Female	20	16	0
3	Female	23	16	0
4	Female	31	17	0

In [40]:

```
km.cluster_centers_
```

Out[40]:

```
array([[ 27.72972973,  27.21621622],
       [ 31.27027027,  79.37837838],
       [ 52.66666667,  32.        ],
       [ 59.1       ,  53.        ],
       [ 36.83333333, 127.66666667],
       [ 39.72222222,  99.44444444],
       [ 23.8       ,  57.88      ],
       [ 46.03448276,  68.93103448]])
```

In [54]:

```
km=KMeans()
km.fit_predict(df[["Age","Income($)"]])
```

```
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
```

Out[54]:

```
array([1, 1, 1, 1, 1, 1, 1, 1, 6, 1, 6, 1, 6, 1, 1, 1, 1, 1, 6, 1, 1, 1,
       6, 1, 6, 1, 6, 1, 1, 1, 6, 1, 6, 1, 6, 1, 6, 1, 1, 1, 6, 1, 6, 1,
       6, 1, 6, 1, 1, 1, 6, 1, 1, 6, 6, 6, 6, 5, 3, 6, 5, 3, 5, 6, 5, 3,
       6, 5, 3, 3, 5, 6, 5, 5, 5, 3, 0, 0, 3, 0, 5, 0, 5, 0, 3, 0, 5, 3,
       3, 0, 5, 3, 0, 0, 3, 3, 0, 3, 0, 3, 3, 0, 5, 3, 0, 3, 5, 0, 5, 5,
       5, 3, 0, 3, 3, 3, 5, 0, 0, 0, 3, 0, 0, 0, 3, 2, 0, 0, 0, 0, 0, 0,
       2, 2, 2, 2, 0, 2, 2, 2, 0, 2, 2, 2, 2, 2, 0, 2, 2, 2, 2, 2, 2, 2,
       0, 2, 2, 2, 2, 2, 0, 2, 2, 2, 7, 2, 7, 2, 2, 2, 7, 2, 2, 2, 7, 2,
       7, 2, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 7, 4, 4, 4, 4, 4, 4,
       4, 4])
```

In [55]:

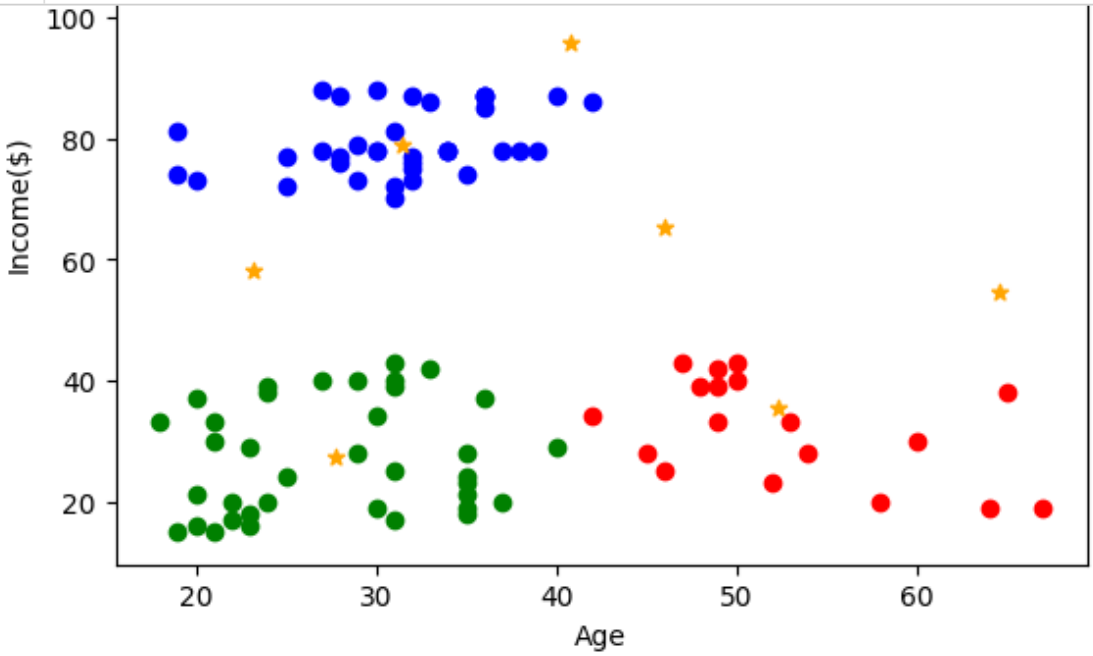
```
df["New cluster"]=y_predicted
df.head()
```

Out[55]:

	Gender	Age	Income($)	cluster	New cluster
0	Male	19	15	0	0
1	Male	21	15	0	0
2	Female	20	16	0	0
3	Female	23	16	0	0
4	Female	31	17	0	0

In [59]:

```
df1=df[df["New cluster"]==0]
df2=df[df["New cluster"]==1]
df3=df[df["New cluster"]==2]
plt.scatter(df1["Age"],df1["Income($)"],color="green")
plt.scatter(df2["Age"],df2["Income($)"],color="blue")
plt.scatter(df3["Age"],df3["Income($)"],color="red")
plt.scatter(km.cluster_centers_[:,0],km.cluster_centers_[:,1],color="orange",marker=
plt.xlabel("Age")
plt.ylabel("Income($)")
```



In [62]:

```
k_rng=range(1,20)
sse=[]
for k in k_rng:
    km=KMeans(n_clusters=k)
    km.fit(df[["Age","Income($)"]])
    sse.append(km.inertia_)
```

```
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
```

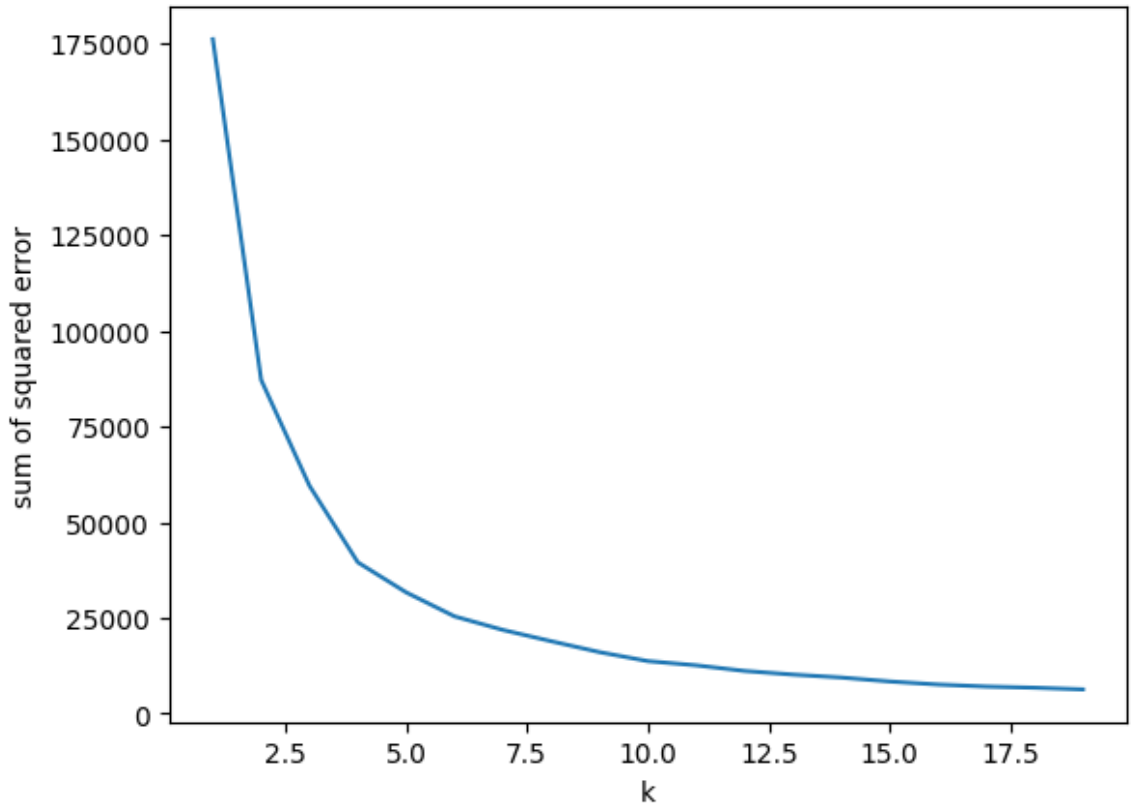
```
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
C:\Users\MY HOME\AppData\Local\Programs\Python\Python311\Lib\site-packages
\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_in
it` will change from 10 to 'auto' in 1.4. Set the value of `n_init` explic
itly to suppress the warning
  warnings.warn(
```

In [63]:

```
plt.plot(k_rng,sse)
plt.xlabel("k")
plt.ylabel("sum of squared error")
```

Out[63]:

```
Text(0, 0.5, 'sum of squared error')
```



In [ ]:

```

```