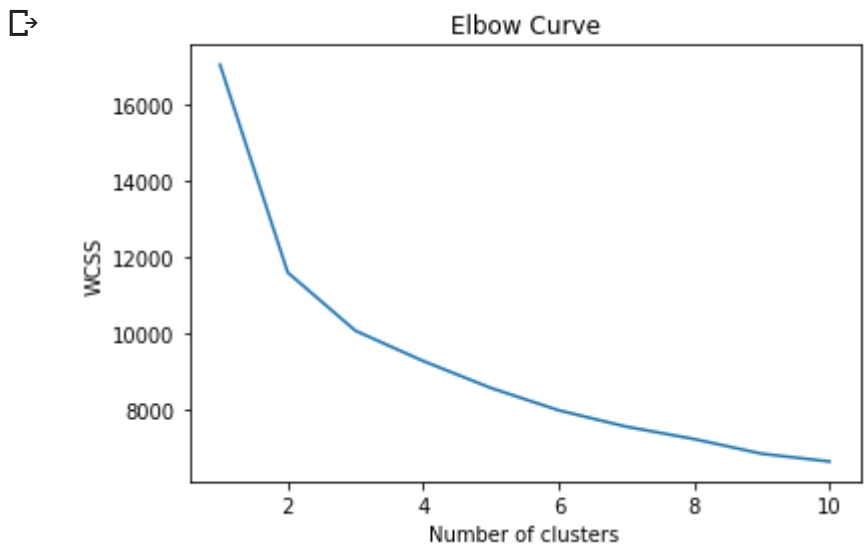
```
# import necessary libraries
import pandas as pd
import numpy as np
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler
from sklearn.datasets import load_breast_cancer


# load data
data = load_breast_cancer()


# preprocess data
scaler = StandardScaler()
scaled_data = scaler.fit_transform(data.data)


wcss = []
for i in range(1, 11):
  kmeans = KMeans(n_clusters=i, init='k-means++', max_iter=300, n_init=10, random_s
  kmeans.fit(scaled_data)
  wcss.append(kmeans.inertia_)


# plot elbow curve to find optimal number of clusters
import matplotlib.pyplot as plt
plt.plot(range(1, 11), wcss)
plt.title('Elbow Curve')
plt.xlabel('Number of clusters')
plt.ylabel('WCSS')
plt.show()
```




```
# use optimal number of clusters to perform KMeans clustering
kmeans = KMeans(n_clusters=2, init='k-means++', max_iter=300, n_init=10, random_sta
kmeans.fit(scaled_data)
```

KMeans

```
# identify anomalies
distances = kmeans.transform(scaled_data)
anomaly_threshold = np.percentile(np.min(distances, axis=1), 95)
anomaly_indices = np.where(np.min(distances, axis=1) > anomaly_threshold)


# print indices of anomalies
print(anomaly_indices)
```

```
(array([  3,   9,  12,  68,  71,  78, 108, 122, 146, 151, 152, 176, 180,
       190, 192, 212, 213, 265, 288, 290, 314, 318, 352, 368, 376, 379,
       461, 504, 505]),)
```