Q12

In [ ]:
```
from sklearn.datasets import load_iris
from sklearn.cluster import KMeans
from sklearn.decomposition import PCA
import numpy as np
import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
```

In [ ]:
```
data = load_iris()
X = data["data"]
y = data["target"]
```

In [ ]:
```
p = PCA(n_components=2)
X = pd.DataFrame(p.fit_transform(X), columns=["X0" , "X1"])
X
```

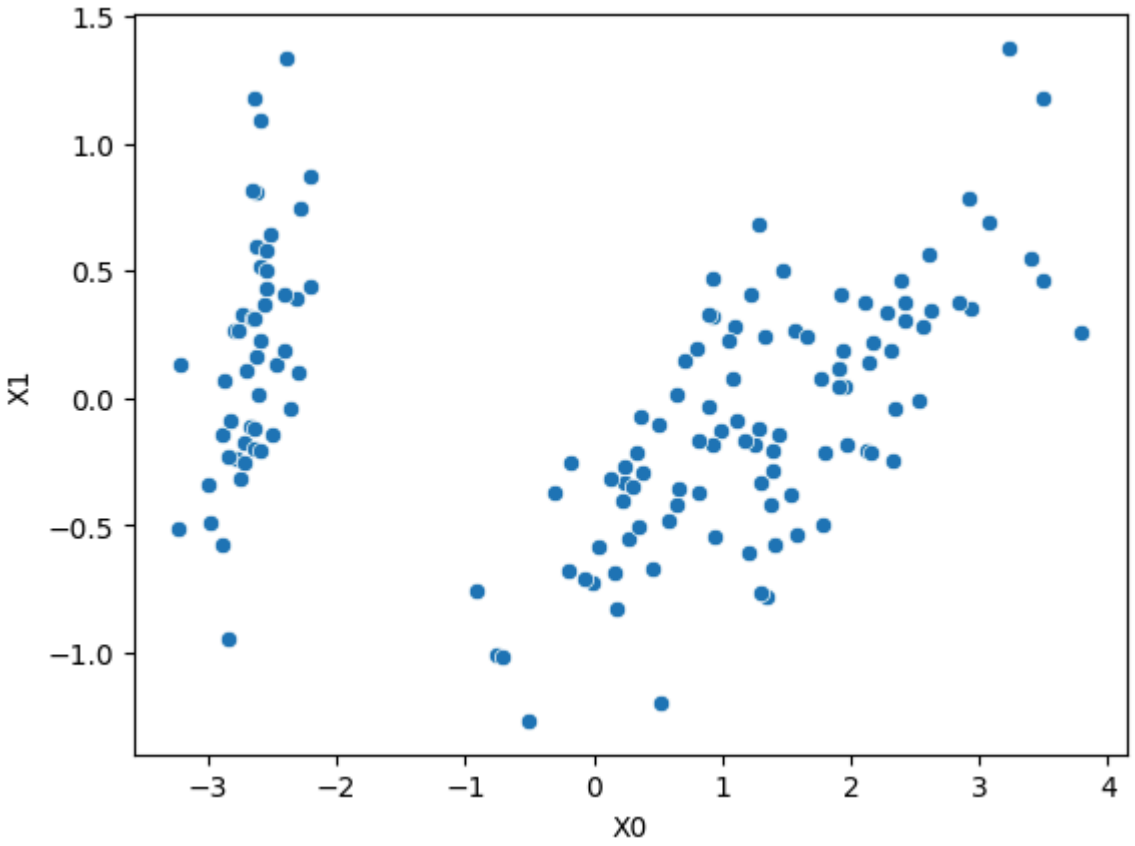
Out[ ]:

	X0	X1
0	-2.684126	0.319397
1	-2.714142	-0.177001
2	-2.888991	-0.144949
3	-2.745343	-0.318299
4	-2.728717	0.326755
...	...	...
145	1.944110	0.187532
146	1.527167	-0.375317
147	1.764346	0.078859
148	1.900942	0.116628
149	1.390189	-0.282661

150 rows × 2 columns

In [ ]:
```
sns.scatterplot(data = X, x="X0" , y="X1")
```

Out[ ]:
```
<Axes: xlabel='X0', ylabel='X1'>
```

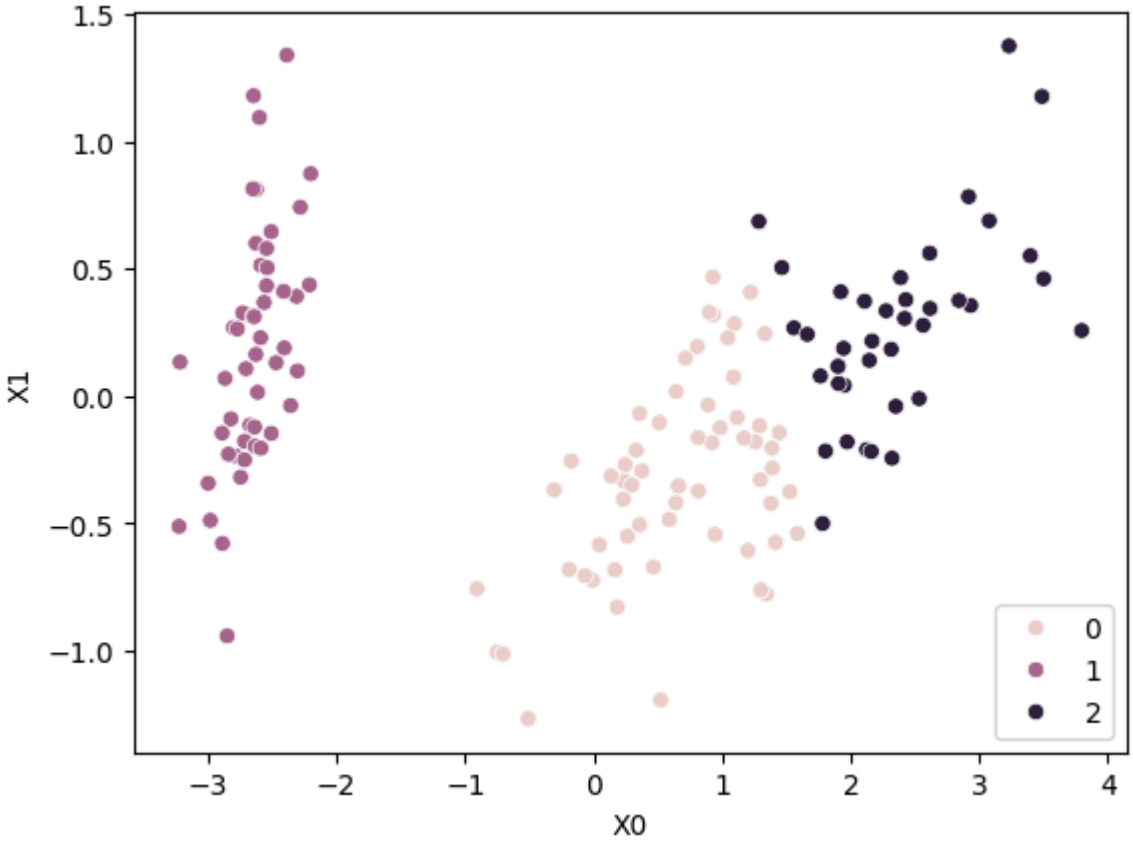


In [ ]:
```
kmeans = KMeans(n_clusters=3 ,random_state = 0, n_init='auto')
kmeans.fit(X)
```

Out[ ]:
```
KMeans
KMeans(n_clusters=3, random_state=0)
```

In [ ]:
```
sns.scatterplot(data = X, x="X0" , y="X1", hue = kmeans.labels_)
```

Out[ ]:
```
<Axes: xlabel='X0', ylabel='X1'>
```


1.5
1.0
0.5
0.0
−0.5
−1.0
X1
−3
−2
−1
0
1
2
3
4
X0
0
1
2