In [1]:

```
from sklearn.datasets import make_blobs
import pandas as pd
```

In [2]:

```
dataset, classes = make_blobs(n_samples=300, n_features=2, centers=5, cluster_std=0.
df = pd.DataFrame(dataset, columns=['X', 'Y'])
df.head(2)
```

Out[2]:

| | X | Y |
|---|---|---|
| **0** | 1.569719 | -0.838530 |
| **1** | -3.750696 | -4.419262 |

In [3]:

```
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=5, init='k-means++', random_state=0).fit(df)
```

In [4]:

```
print(kmeans.n_iter_) #total number of iterations to convergence
```

```
3
```

In [5]:

```
print(kmeans.cluster_centers_) #cluster centers
```

```
[[-5.83451299  2.30779804]
 [-1.35495664 -9.43666895]
 [-4.06321324 -4.81843763]
 [ 1.00220205 -1.26008793]
 [-1.56907217 -3.44508562]]
```

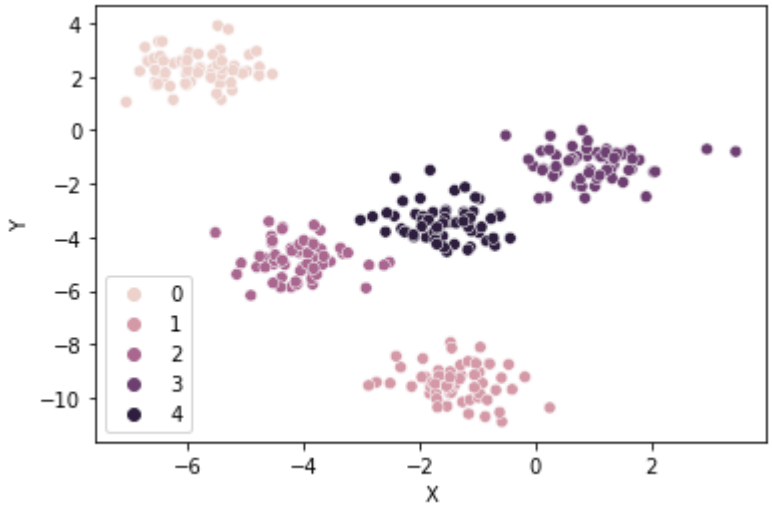
In [6]:

```
print(kmeans.inertia_) #defines how well the dataset is clustered
```

```
225.49694297568254
```

In [7]:

```
import seaborn as sns
import matplotlib.pyplot as plt
sns.scatterplot(data=df, x="X", y="Y", hue=kmeans.labels_)
plt.show()
```




In [ ]:

```
```