In[1]:

```
import matplotlib.pyplot as plt

import seaborn as sns

%matplotlib inline
```

For this example, we are going to create our data and then apply the algorithm to it.

In[2]:

```
from sklearn.datasets import make_blobs

data = make_blobs(n_samples = 500, n_features = 2,centers = 6, cluster_std = 1.8, random_state = 101)
```

In[3]:

```
plt.scatter(data[0][:,0],data[0][:,1],c=data[1],cmap='rainbow')
```

In[4]:

```
from sklearn.cluster import KMeans

model = KMeans(n_clusters=6)

model.fit(data[0])
```

In[5]:

```
f, (ax1, ax2) = plt.subplots(1, 2, sharey=True, figsize=(10,6))


ax1.set_title('K Means')

ax1.scatter(data[0][:,0], data[0][:,1], c=kmeans.labels_, cmap='rainbow')


ax2.set_title("Original")

ax2.scatter(data[0][:,0], data[0][:,1], c=data[1], cmap='rainbow')
```

In[6]:

```
SSE = []
list_k = list(range(1, 10))

for k in list_k:
    km = KMeans(n_clusters=k)
    km.fit(data[0])
    SSE.append(km.inertia_)
```

In[7]:

```
plt.figure(figsize=(6, 6))
plt.plot(list_k, sse, '-o')
plt.xlabel(r'Number of clusters *K*')
plt.ylabel('Sum of squared distance');
```