mpg

June 21, 2024

[6]:
```
# Import necessary libraries
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler

# Load dataset
mpg = sns.load_dataset('mpg')

# Selecting relevant features for clustering
X = mpg[['mpg', 'acceleration']]  # Using 'mpg' (miles per gallon) and␣
 ↪'acceleration' features for clustering

# Scale the data
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

# Perform KMeans clustering
kmeans = KMeans(n_clusters=3, random_state=42)
kmeans.fit(X_scaled)

# Predict cluster labels
labels = kmeans.labels_

# Visualize the clusters
plt.figure(figsize=(8, 6))
sns.scatterplot(x=X.iloc[:, 0], y=X.iloc[:, 1], hue=labels, palette='viridis',␣
 ↪legend='full')
plt.scatter(kmeans.cluster_centers_[:, 0], kmeans.cluster_centers_[:, 1],␣
 ↪marker='s', color='red', s=100, label='Centroids')
plt.title('KMeans Clustering')
plt.xlabel('Miles per Gallon (mpg)')
plt.ylabel('Acceleration')
plt.legend()
plt.show()
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870:
FutureWarning: The default value of `n_init` will change from 10 to 'auto' in
```

```
1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
```