Exercise 7

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import AgglomerativeClustering
from sklearn.preprocessing import StandardScaler

dataset = pd.read_csv(r"C:\Users\sridharan\Downloads\data.csv", encoding='latin1')

dataset['Quantity'] = pd.to_numeric(dataset['Quantity'], errors='coerce')
dataset['UnitPrice'] = pd.to_numeric(dataset['UnitPrice'], errors='coerce')

dataset = dataset.dropna(subset=['Quantity', 'UnitPrice'])

X = dataset[['Quantity', 'UnitPrice']]

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

clustering = AgglomerativeClustering(n_clusters=4, linkage='complete')
clustering.fit(X_scaled)

plt.figure(figsize=(8, 6))
plt.scatter(X_scaled[:,0], X_scaled[:,1], c=clustering.labels_, cmap='viridis', s=50)
plt.title('CURE Algorithm Clustering')
plt.xlabel('Standardized Quantity')
plt.ylabel('Standardized Unit Price')
```

```
plt.colorbar()

plt.show()
```

...