

ALGORITHMS USED FOR BOTH THE CREATION OF THE CLUSTER CSV FILE AND PLOTTING THE GRAPH.

1. Creating the cluster csv file.

```
from sklearn.cluster import KMeans

# coming up with the columns used for clustering
duka = delivery[['customer_number', 'duka_latitude', 'duka_longitude']].drop_duplicates().reset_index(drop=True)

# warehouse Longitudes and Latitudes
warehouse = centroids[['territory', 'warehouse_latitude', 'warehouse_longitude']]

duka_coords = duka[['duka_latitude', 'duka_longitude']].values

#implementing the KMeans clustering model
num_clusters = warehouse.shape[0]
kmeans = KMeans(n_clusters=num_clusters, init = 'k-means++', random_state=42)
duka['cluster'] = kmeans.fit_predict(duka_coords)

warehouse_coords = warehouse[['warehouse_latitude', 'warehouse_longitude']].values
warehouse['cluster'] = range(num_clusters)

#joining the two clusters
duka = duka.merge(warehouse, left_on='cluster', right_on='cluster', how = 'left', suffixes=('_duka', '_warehouse'))

duka.head()
```

2. Plotting the graph.

```
# Plot the clusters on a map
plt.figure(figsize=(12, 8))
for cluster_id in range(num_clusters):
    duka_cluster = duka[duka['cluster'] == cluster_id]
    plt.scatter(duka_cluster['duka_latitude'], duka_cluster['duka_longitude'], label=f'Cluster {cluster_id + 1}')

# Plot warehouse Locations
plt.scatter(warehouse['warehouse_latitude'], warehouse['warehouse_longitude'], color='red', marker='x', s=100, label='Warehouse')

plt.xlabel('Latitude')
plt.ylabel('Longitude')
plt.legend()
plt.title('duka Clusters Around Fulfillment Centers')
plt.show()
```

CLUSTER SCREENSHOT.

