TASK 3 (CLUSTERING)

```
# Import necessary libraries
import pandas as pd
from sklearn.preprocessing import LabelEncoder, MinMaxScaler
from sklearn.metrics.pairwise import cosine_similarity
from datetime import datetime


# Load datasets
customers = pd.read_csv("Customers.csv")
products = pd.read_csv("Products.csv")
transactions = pd.read_csv("Transactions.csv")


# Merge datasets
merged_data = transactions.merge(customers, on="CustomerID").merge(products, on="ProductID")
merged_data.rename(columns={'Price_y': 'Price'}, inplace=True)


# Convert TransactionDate to datetime
merged_data['TransactionDate'] = pd.to_datetime(merged_data['TransactionDate'])



# Current date for recency calculation
current_date = datetime.now()


# Aggregate features for customer profiles
customer_profiles = merged_data.groupby('CustomerID').agg({
    'Region': 'first',
    'Price': 'mean',
    'Quantity': 'sum',
    'TotalValue': 'sum',
    'TransactionDate': lambda x: (current_date - x.max()).days,
    'Category': 'nunique'
}).reset_index()



# Rename TransactionDate to Recency
customer_profiles.rename(columns={'TransactionDate': 'Recency'}, inplace=True)


# ----> Encode 'Region' column before scaling <----
le_region = LabelEncoder()
customer_profiles['Region'] = le_region.fit_transform(customer_profiles['Region'])


# Normalize features
scaler = MinMaxScaler()
normalized_features = scaler.fit_transform(customer_profiles.iloc[:, 1:]) # Exclude CustomerID


# Calculate cosine similarity
similarity_matrix = cosine_similarity(normalized_features)
similarity_df = pd.DataFrame(
    similarity_matrix,
    index=customer_profiles['CustomerID'],
    columns=customer_profiles['CustomerID']
)
```

SMALL CHANGES (CORRECT CODE)

```
# Import necessary libraries
import pandas as pd
from sklearn.preprocessing import LabelEncoder, MinMaxScaler
from sklearn.cluster import KMeans
from sklearn.metrics import davies_bouldin_score
import matplotlib.pyplot as plt
import seaborn as sns
```

Start coding or generate with AI.

Start coding or generate with AI.

Start coding or generate with AI.

Start coding or generate with AI.

Start coding or generate with AI.

Start coding or generate with AI.

Start coding or generate with AI.