PHASE 4

To build the market basket insights project by:

- Performing association analysis
- Generating insights.

Creating a Market Basket Insights project involves two key steps: performing association analysis and generating insights. Here's a brief overview of each step:

Performing Association Analysis:

- Start by gathering transaction data, which includes a list of items purchased in each transaction.
- Utilize data mining techniques like Apriori or FP-growth to identify frequent itemsets and association rules. These rules reveal patterns of items that tend to be purchased together.
- Set a minimum support and confidence threshold to filter out less relevant associations.
- Calculate lift or other relevant metrics to prioritize and refine the rules.
- Visualize the associations using tools like scatter plots or network diagrams.

Generating Insights:

- Analyze the association rules to gain insights into customer behavior and purchasing patterns.
- Identify cross-selling opportunities by finding items frequently bought together.
- Understand which items are commonly bought alone, indicating potential bundling opportunities.
- Explore the impact of promotions or discounts on item associations.
- Create recommendations for optimizing product placement, marketing strategies, and pricing

CODE & OUTPUT:

```
Import numpy as np

Import pandas as pd

Import matplotlib.pyplot as plt

From mlxtend.frequent_patterns import apriori

From mlxtend.frequent_patterns import association_rules

Dataset = pd.read_excel('C:\\ELCOT\\downloads\\Assignment-1_Data.xlsx')


# Split the 'Itemname' column into individual items

Items_df = transaction_data['Itemname'].str.split(', ', expand=True)


# Concatenate the original DataFrame with the new items DataFrame

Transaction_data = pd.concat([transaction_data, items_df], axis=1)


# Drop the original 'Itemname' column

Transaction_data = transaction_data.drop('Itemname', axis=1)


# Display the resulting DataFrame

Print(transaction_data.head())

# Convert items to boolean columns
```

Df_encoded = pd.get_dummies(transaction_data, prefix='', prefix_sep='').groupby(level=0, axis=1).max()

 # Save the transaction data to a CSV file

Df_encoded.to_csv('transaction_data_encoded.csv', index=False)

 # Load transaction data into a DataFrame

Df_encoded = pd.read_csv('transaction_data_encoded.csv')

 # Association Rule Mining

Frequent_itemsets = apriori(df_encoded, min_support=0.007, use_colnames=True)

Rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=1)

 # Display information of the rules

Rules.head(100)

	Antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
0	(10 COLOUR SPACEBOY PEN)	(LUNCH BAG APPLE DESIGN)	0.024070	0.061269	0.010941	0.454545	7.418831	0.009466	1.721007	0.886547
1	(LUNCH BAG APPLE DESIGN)	(10 COLOUR SPACEBOY PEN)	0.061269	0.024070	0.010941	0.178571	7.418831	0.009466	1.188089	0.921678

2 (PLASTERS IN TIN SPACEBOY) (10 COLOUR SPACEBOY PEN) 0.107221 0.024070 0.008753 0.081633 3.391466 0.006172 1.062679 0.789828

3 (10 COLOUR SPACEBOY PEN) (PLASTERS IN TIN SPACEBOY) 0.024070 0.107221 0.008753 0.363636 3.391466 0.006172 1.402938 0.722534

4 (10 COLOUR SPACEBOY PEN) (ROUND SNACK BOXES SET OF4 WOODLAND) 0.024070 0.245077 0.008753 0.363636 1.483766 0.002854 1.186308 0.334081

… … … … … … … … … … …

95 (36 PENCILS TUBE WOODLAND) (JUMBO BAG APPLES) 0.026258 0.061269 0.008753 0.333333 5.440476 0.007144 1.408096 0.838202

96 (36 PENCILS TUBE WOODLAND) (JUMBO BAG PINK POLKADOT) 0.026258 0.035011 0.008753 0.333333 9.520833 0.007833 1.447484 0.919101

97 (JUMBO BAG PINK POLKADOT) (36 PENCILS TUBE WOODLAND) 0.035011 0.026258 0.008753 0.250000 9.520833 0.007833 1.298322 0.927438

98 (36 PENCILS TUBE WOODLAND) (JUMBO BAG RED RETROSPOT) 0.026258 0.078775 0.008753 0.333333 4.231481 0.006684 1.381838 0.784270

99 (JUMBO BAG RED RETROSPOT) (36 PENCILS TUBE WOODLAND)

Rules = rules.sort_values(by='lift', ascending = False)

Rules

	Antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction	zhangs_metric
58982	(SPACEBOY BIRTHDAY CARD, PINK VINTAGE SPOT BEA…	(SPACEBOY CHILDRENS CUP, RED VINTAGE SPOT BEAK…	0.008753	0.008753	0.008753	1.000000	114.250000	0.008676	inf	1.000000
57294	(PLASTERS IN TIN WOODLAND ANIMALS, JUMBO BAG P…	(JUMBO BAG WOODLAND ANIMALS, PLASTERS IN TIN V…	0.008753	0.008753	0.008753	1.000000	114.250000	0.008676	inf	1.000000
2611	(MONSTERS STENCIL CRAFT)	(HAPPY STENCIL CRAFT)	0.008753	0.008753	0.008753	1.000000	114.250000	0.008676	inf	1.000000
56376	(CHILDRENS CUTLERY SPACEBOY, LUNCH BAG PINK PO…	(SKULL LUNCH BOX WITH CUTLERY, LUNCH BAG SPACE…	0.008753	0.008753	0.008753	1.000000	114.250000	0.008676	inf	1.000000
66328	(ROUND SNACK BOXES SET OF4 WOODLAND, CARD PSYC…	(REGENCY CAKESTAND 3 TIER, RED HARMONICA IN B…	0.008753	0.008753	0.008753	1.000000	114.250000	0.008676	inf	1.000000
…	…	…	…	…	…	…	…	…	…	…
3769	(MEMO BOARD COTTAGE DESIGN)	(ROUND SNACK BOXES SET OF 4 FRUITS)	0.054705	0.157549	0.008753	0.160000	1.015556	0.000134	1.002918	0.016204

5552 (REGENCY CAKESTAND 3 TIER) (STRAWBERRY LUNCH BOX WITH CUTLERY) 0.137856 0.078775 0.010941 0.079365 1.007496 0.000081 1.000641 0.008629

5553 (STRAWBERRY LUNCH BOX WITH CUTLERY) (REGENCY CAKESTAND 3 TIER) 0.078775 0.137856 0.010941 0.138889 1.007496 0.000081 1.001200 0.008076

3594 (REGENCY CAKESTAND 3 TIER) (LUNCH BAG WOODLAND) 0.137856 0.078775 0.010941 0.079365 1.007496 0.000081 1.000641 0.008629

3595 (LUNCH BAG WOODLAND) (REGENCY CAKESTAND 3 TIER)