MARKET BASKET INSIGHTS
Phase 4: Development part 2

Introduction:

Market basket analysis is a data mining technique used to discover associations between products that are frequently purchased together. It's a valuable tool for retailers and businesses to gain insights into customer behaviour and improve various aspects of their operations. Here are some insights you can generate from performing association analysis using market basket data. Developing market basket insights involves several key parts and processes. Market basket analysis is a data mining technique used to discover associations between products or items that are frequently purchased together. Here are the main components of developing market basket insights Market basket analysis can provide valuable insights for businesses to enhance their sales, customer experience, and inventory management, ultimately leading to increased revenue and customer satisfaction. Association analysis or learning of the association rule is defined as a process of searching for hidden association or pattern in a large dataset.

1. Frequent item set:

Identify the most common item sets or combinations of products that customers tend to buy together. This can help you optimize product placement, create effective product bundles, and cross-sell related items.

2. Association Rules:

Discover association rules, such as "if a customer buys X, they are likely to buy Y." This information can guide your marketing strategies and product recommendations, both in-store and online.

3. Lift Analysis:

Calculate the lift for different item pairs. Lift measures the strength of association between items. A lift greater than 1 indicates a positive association, while a lift less than 1 indicates that the items are less likely to be bought together. This helps you understand which products are truly associated and which are not.

4. Support, Confidence, and Lift Metrics:

Use support, confidence, and lift metrics to assess the significance of association rules. High support means that the rule is applicable to a significant portion of transactions, high confidence indicates a strong correlation between items, and high lift suggests a substantial increase in the likelihood of buying both items together.

5. Market Basket Size:

Analyse the size of market baskets, i.e., the number of items purchased in a single transaction. This can help you tailor promotions and discounts to encourage customers to buy more items per visit.

6. Customer Segmentation:

Segment customers based on their purchase history and preferences. For example, you can group customers who buy similar items and create targeted marketing campaigns for each segment.

7. Inventory Management:

Optimize inventory by stocking items that are frequently bought together, thus reducing out-of-stock situations and ensuring you have the right products in stock.

8. Strategies:

Adjust pricing strategies based on associations. For example, you might offer discounts on associated items or increase prices on stand-alone products to encourage customers to buy bundles.

9. Cross-Selling and Recommendations:

Implement personalized product recommendations on your website or in-store displays, suggesting items that are commonly purchased with the products customers are viewing or purchasing. Market Basket Analysis is a technique that assists us in determining which sports items are purchased together according to the association rules, primarily with the aim of identifying cross- selling opportunities.

10. Seasonal Trends and Special Occasions:

Identify any seasonal or special occasion associations. For example, which items tend to be bought together during holidays, events, or specific times of the year.

11. A/B Testing:

Conduct A/B testing to measure the impact of different association-based strategies on sales, such as the introduction of new product bundles or changing in-store layouts.

12. Basket Analysis for E-commerce:

For online retailers, use market basket insights to improve website design, navigation, and search functionalities to make it easier for customers to find and add associated items to their carts

13. Customer Loyalty Programs:

Develop loyalty programs based on frequent item sets, rewarding customers for purchasing associated products, which can increase customer retention.

14. Loss Leader Strategies:

Identify potential "loss leaders," which are products sold at a loss to attract customers, with the expectation of selling them other, more profitable items.

Coding:

```python
import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from mlxtend.frequent_patterns import apriori

from mlxtend.frequent_patterns import association_rules

sns.set(style="darkgrid", color_codes=True)

pd.set_option('display.max_columns', 75)

import pandas as pd

linkcode

df = pd.read_excel("Assignment-1_Data.csv",header=None)

df.info()

df.head()

df.describe()

color = plt.cm.rainbow(np.linspace(0, 1, 40))

df[0].value_counts().head(40).plot.bar(color = color, figsize=(13,5))

plt.title('frequency of most popular items', fontsize = 20)

plt.xticks(rotation = 90 )

plt.grid()

plt.show()

import networkx as nx

df['things'] = 'Things'

things= df.truncate(before = -1, after = 15)

things = nx.from_pandas_edgelist(things, source = 'things', target = 0, edge_attr = True)

import warnings

warnings.filterwarnings('ignore')

plt.rcParams['figure.figsize'] = (13, 13)

pos = nx.spring_layout(things)

color = plt.cm.Set1(np.linspace(0, 15, 1))

nx.draw_networkx_nodes(things, pos, node_size = 15000, node_color = color)

nx.draw_networkx_edges(things, pos, width = 3, alpha = 0.6, edge_color = 'black')
```

```
nx.draw_networkx_labels(things, pos, font_size = 20, font_family = 'sans-serif')

plt.axis('off')

plt.grid()

plt.title('Top 15 First Choices', fontsize = 20)

plt.show()

transactions[:1]

from itertools import permutations

flattened = [item for transaction in transactions for item in transaction]

items = list(set(flattened))

linkcode

print('# of items:',len(items))

print(list(items))

if 'KINGS CHOICE BISCUIT TIN' in items: items.remove('KINGS CHOICE BISCUIT TIN')

print(list(items))

rules = list(permutations(items, 2))

print('# of rules:',len(rules))

print(rules[:5])
```

Output:

class 'pandas.core.frame.DataFrame'>

RangeIndex: 522065 entries, 0 to 522064

Data columns (total 7 columns):

#	Column	Non-Null Count	Dtype
0	0	522065 non-null	object
1	1	520610 non-null	object
2	2	522065 non-null	object
3	3	522065 non-null	object
4	4	522065 non-null	object
5	5	388024 non-null	object
6	6	522065 non-null	object

dtypes: object(7)

memory usage: 27.9+ MB

	0	1	2	3	4	5	6
0	BillNo	Itemname	Quantity	Date	Price	CustomerID	Country
1	536365	WHITE HANGING HEART T-LIGHT HOLDER	6	2010-12-01 08:26:00	2.55	17850	United Kingdom
2	536365	WHITE METAL LANTERN	6	2010-12-01 08:26:00	3.39	17850	United Kingdom
3	536365	CREAM CUPID HEARTS COAT HANGER	8	2010-12-01 08:26:00	2.75	17850	United Kingdom
4	536365	KNITTED UNION FLAG HOT WATER BOTTLE	6	2010-12-01 08:26:00	3.39	17850	United Kingdom

	0	1	2	3	4	5	6
count	522065	520610	522065	522065	522065.00	388024	522065
unique	21664	4186	691	19642	1286.00	4298	31
top	573585	WHITE HANGING HEART T-LIGHT HOLDER	1	2011-10-31 14:41:00	1.25	17841	United Kingdom
freq	1114	2269	147848	1114	48778.00	7847	487622

frequency of most popular items

Top 15 First Choices
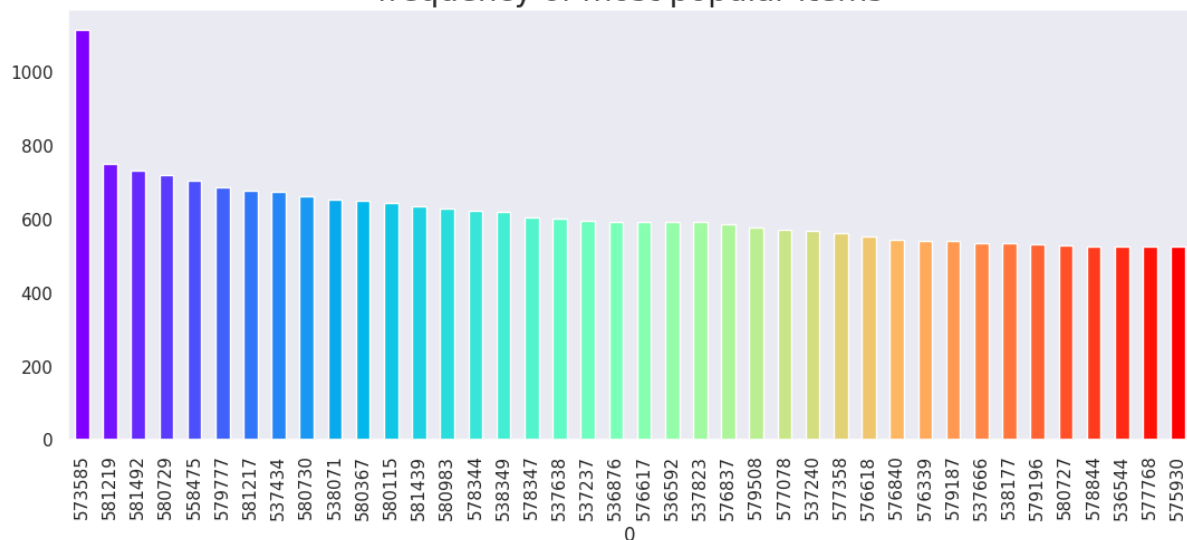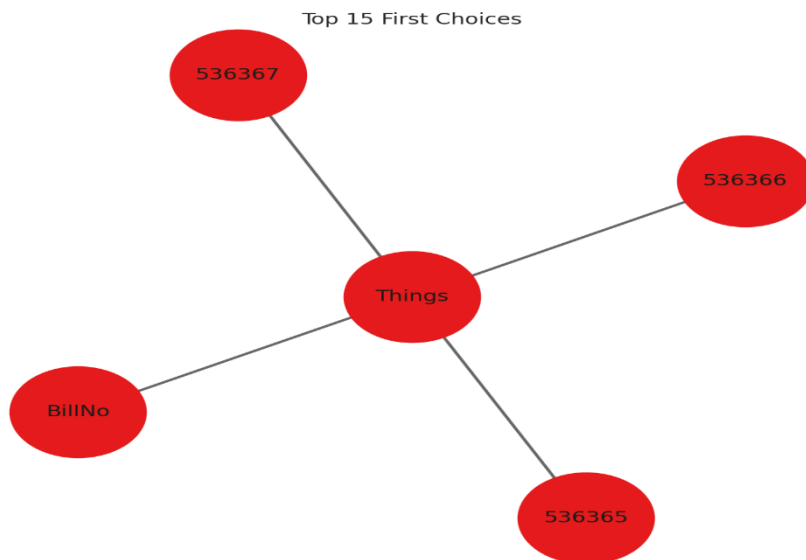
# of rules: 1021110

[('5 STRAND GLASS NECKLACE CRYSTAL', '2010-12-01 13:17:00'), ('5 STRAND GLASS NECKLACE CRYSTAL', '2010-12-01 12:43:00'), ('5 STRAND GLASS NECKLACE CRYSTAL', 'CAKE PLATE LOVEBIRD WHITE'), ('5 STRAND GLASS NECKLACE CRYSTAL', 'TRADITIONAL WOODEN CATCH CUP GAME'), ('5 STRAND GLASS NECKLACE CRYSTAL', 'RED RETROSPOT OVEN GLOVE')]

# of items: 1012

['KINGS CHOICE BISCUIT TIN', '5 STRAND GLASS NECKLACE CRYSTAL', '2010-12-01 13:17:00', '2010-12-01 12:43:00', 'CAKE PLATE LOVEBIRD WHITE', 'TRADITIONAL WOODEN CATCH CUP GAME', 'RED RETROSPOT OVEN GLOVE', '30', '0.55', 'WOOD S/3 CABINET ANT WHITE FINISH', 'SET OF 6 FUNKY BEAKERS', 'LADIES & GENTLEMEN METAL SIGN', 'WHITE WOOD GARDEN PLANT LADDER', '536502', 'ALARM CLOCK BAKELIKE GREEN', '2010-12-01 09:34:00', '17905', '536380', 'SET OF 3 NOTEBOOKS IN PARCEL', 'LIGHT GARLAND BUTTERFILES PINK', 'AGED GLASS SILVER T-LIGHT HOLDER', '536463', 'BOX OF VINTAGE ALPHABET BLOCKS', 'GLASS CLOCHE SMALL', '2010-12-01 09:59:00', 'BIRD HOUSE HOT WATER BOTTLE', '60 CAKE CASES VINTAGE CHRISTMAS', 'RIBBON REEL MAKING SNOWMEN', '13767', 'VINTAGE SNAKES & LADDERS', 'PACK OF 60 MUSHROOM CAKE CASES', 'JUMBO SHOPPER VINTAGE RED PAISLEY', 'VINTAGE HEADS AND TAILS CARD GAME', '5', 'ENGLISH ROSE SPIRIT LEVEL', '2010-12-01 12:22:00', '12748', 'CERAMIC STRAWBERRY CAKE MONEY BANK', '536372', 'FELTCRAFT PRINCESS CHARLOTTE DOLL', 'PACK OF 12 LONDON TISSUES', 'STRAWBERRY BATH SPONGE', 'COSY HOUR GIANT TUBE MATCHES', '536500', 'CHOCOLATE 3 WICK MORRIS BOX CANDLE', 'IVORY GIANT GARDEN THERMOMETER', 'SMALL GLASS HEART TRINKET POT', 'TRADITIONAL CHRISTMAS RIBBONS', 'RUSTIC SEVENTEEN DRAWER SIDEBOARD', 'LUNCH BAG SPACEBOY DESIGN', '2010-12-01 11:33:00', 'CHARLIE + LOLA RED HOT WATER BOTTLE', '3.25', 'RECORD FRAME 7" SINGLE SIZE', '100', 'FELTCRAFT DOLL MOLLY', 'EDWARDIAN PARASOL NATURAL', '2010-12-01 13:23:00', 'PLASTERS IN TIN VINTAGE PAISLEY', '15513', '13408', 'SET OF 2 WOODEN MARKET CRATES', '2010-12-01 12:23:00', '144', 'METAL 4 HOOK HANGER FRENCH CHATEAU', '2010-12-01 10:19:00', '3.39', 'WHITE LOVEBIRD LANTERN', '2010-12-01 11:22:00', 'TOADSTOOL MONEY BOX', 'SNOWFLAKE PORTABLE TABLE LIGHT', 'SAVE THE PLANET MUG', 'RED HARMONICA IN BOX', 'RETROSPOT CIGAR BOX MATCHES', '15291', 'MIRRORED DISCO BALL', '2010-12-01 12:48:00', 'GUMBALL MONOCHROME COAT RACK', 'PAPER CHAIN KIT RETROSPOT', 'CIRCUS PARADE LUNCH BOX', 'GINGHAM HEART DOORSTOP RED', '17181', 'FOLDING UMBRELLA CREAM POLKADOT', '4 PURPLE FLOCK DINNER CANDLES', 'RETROSPOT CHILDRENS APRON', 'RETROSPOT RED WASHING UP GLOVES', 'VICTORIAN SEWING BOX MEDIUM',

'536385', '64', 'SILVER HANGING T-LIGHT HOLDER', 'HERB MARKER PARSLEY', 'PACK OF 12 RED RETROSPOT TISSUES', 'SKULLS WRITING SET', 'METAL SIGN EMPIRE TEA', '2010-12-01 11:21:00', '14307', 'CARD BILLBOARD FONT', 'ENAMEL BREAD BIN CREAM', '2010-12-01 12:36:00', 'PINK PAISLEY SQUARE TISSUE BOX', 'VINTAGE BILLBOARD LOVE/HATE MUG', '7', '2010-12-01 12:49:00', 'RAINY LADIES BIRTHDAY CARD', 'HAND WARMER BABUSHKA DESIGN', '536397', 'JUMBO BAG STRAWBERRY', 'MIRRORED WALL ART FOXY', 'HAND OVER THE CHOCOLATE   SIGN', 'JAM MAKING SET PRINTED', 'RED RETROSPOT MUG', 'PARTY INVITES FOOTBALL', 'TOY TIDY DOLLY GIRL DESIGN', 'WHITE SPOT RED CERAMIC DRAWER KNOB', 'RECIPE BOX WITH METAL HEART', 'Things', '2010-12-01 11:58:00', 'TOTE BAG I LOVE LONDON', '36 PENCILS TUBE RED RETROSPOT', '7.95', '72 SWEETHEART FAIRY CAKE CASES', 'TRADITIONAL WOODEN SKIPPING ROPE', '2010-12-01 09:37:00', '23', 'BLACK CANDELABRA T-LIGHT HOLDER', 'FELTCRAFT HAIRBAND PINK AND PURPLE', 'ROTATING SILVER ANGELS T-LIGHT HLDR', 'FELTCRAFT CUSHION RABBIT', 'AIRLINE BAG VINTAGE JET SET WHITE', 'APPLE BATH SPONGE', 'MINI JIGSAW CIRCUS PARADE', 'CHRISTMAS DECOUPAGE CANDLE', '536401', 'SET OF 6 T-LIGHTS SANTA', 'PLASTERS IN TIN WOODLAND ANIMALS', 'LUNCH BOX WITH CUTLERY RETROSPOT', 'RETROSPOT GIANT TUBE MATCHES', 'OFFICE MUG WARMER POLKADOT', 'SET/4 MODERN VINTAGE COTTON NAPKINS', 'BillNo', '56', 'GIN & TONIC DIET GREETING CARD', '17920', 'Germany', 'DOORMAT BLACK FLOCK', 'CHRISTMAS HANGING STAR WITH BELL', '2010-12-01 12:51:00', 'SET/10 BLUE POLKADOT PARTY CANDLES', 'RETROSPOT TEA SET CERAMIC 11 PC', '3.95', '15983', 'SANDWICH BATH SPONGE', 'TRIPLE PHOTO FRAME CORNICE', 'PACK OF 72 SKULL CAKE CASES', 'ALARM CLOCK BAKELIKE RED', 'PACK 3 FIRE ENGINE/CAR PATCHES', 'DIAMANTE HAIR GRIP PACK/2 MONTANA', '80', 'FELTCRAFT CHRISTMAS']

Conclusion:

Association Analysis or Market Basket Analysis is a useful tool for organizations who wants to better understand the relationships between the topics that their members discuss, which can help them to devise efficient call handling strategies. In market basket analysis, association rules are used to predict the likelihood of products being purchased together. Association rules count the frequency of items that occur together, seeking to find associations that occur far more often than expected. To put it simply, market basket analysis helps the retailers know about the products frequently bought together so as to keep those items always available in their inventory. The source from which these patterns are found is the vast amount of data that is continually collected and stored. Is one of the most important methods used to uncover relationships between items. It looks for combinations of items that frequently occur together in transactions.