**MARKET BASKET INSIGHTS:**

**Unveiling customer behavior through association analysis**

**INTRODUCTION:**

Analysis of customer purchase patterns to identify associations and trends among products bought together.

**PLAN OF ACTIONS**

**STEP -1:**

The collection of data from the customers using IOT devices.

The data is used to estimate the customer’s product preferences.

Code for collection of data is given below,

**import** pandas **as** pd

df = pd.read\_csv('Groceries\_dataset.csv')

df.head()

**STEP-2:**

The data collected is processed by removing the irrelevant information .

Those data are converted into suitable format.

**STEP-3:**

Creating a complete dataset by merging structured and unstructured data while employing data integration tools for accuracy and consistency.

Code for grouping data is given below,

df[‘single\_transaction’] = df[‘Member\_number’].astype(str)+’\_’+df[‘Date’].astype(str)

df.head()

df2 = pd.crosstab(df['single\_transaction'], df['itemDescription'])

df2.head()

**def** **encode**(item\_freq):

res = 0

**if** item\_freq > 0:

res = 1

**return** res

basket\_input = df2.applymap(encode)

**STEP-4:**

The Apriori algorithm is the technique used for performing market basket analysis.

APRIORI ALGORITHM:

It is used for **association rule mining**, which is a rule-based process used to identify correlations between items purchased by users.

The code for concluding the mostly brought products using above algorithm is as follows,

**from** mlxtend.frequent\_patterns **import** apriori

**from** mlxtend.frequent\_patterns **import** association\_rules

frequent\_itemsets = apriori(basket\_input, min\_support=0.001, use\_colnames=True)

rules = association\_rules(frequent\_itemsets, metric="lift")

rules.head()

**STEP-5:**

Apply filters to select the most interesting and relevant rules based on confidence, lift and support.

**Support(item 1)= transactions comprising the item 1/total transactions**

**Confidence(item 1 and item 2)= transactions comprising item 1 and**

**item 2/ transaction comprising item 1**

**lift=confidence(item 1 and item 2)/support (item 1)**

To get the most frequent item combinations in the entire dataset, let’s sort the dataset by support, confidence, and lift:

rules.sort\_values(["support", "confidence","lift"],axis = 0, ascending = False).head(8)

**STEP-6:**

By using matplotlib library in python ,the results are visualized using charts, graphs or other visualization tools.

This type of visualization is used to communicate to finding the products effectively.

**STEP-7:**

The regular update of data is required for market basket analysis and re-run the analysis to adapt to changing customer behavior.

**CONCLUSION:**

Market based insights provide valuable information on the purchasing behavior of customers and helping the businesses to enhance sales and customer satisfaction.