International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

Association Rule Mining using FP-Growth Algorithm in Market Basket Analysis

Iyad S. Alkronz[1] and Samy S. Abu-Naser[2]

Department of information technology,
Faculty of Engineering and Information Technology
Al-Azhar University, Gaza, Palestine
Email: eyadalkronz@gmail.com [1], abunaser@alazhar.edu.pd[2]

***Abstract**: This study was conducted in order to make a Market Basket Analysis by using Association Rules mining through FP-Growth Algorithm. The data used in the study are transactions of customers in a UK based online retail outfit. Data were analyzed in the R Studio program and Spark using a data set containing 541910 transactions and 4070 different products. FP-Growth Algorithm was tried which are Association Rules algorithms, The best rule accordingly a customer who buys Regency Tea Plate Pink and Regency Tea Plate Roses also gets Regency Tea Plate Green with confidence 0.94 and the lift ratio value 64.3 Similarly, also other rules were interpreted in this study. As a result, product placement in the supermarket can be made according to these rules. Thus, sales of these products will increase and revenue will increase directly.*

Keywords: Market Basket Analysis, Association Rule Mining, FP-Growth Algorithm, Data mining, R, Spark

1 Introduction

The increasing volume of data and the growing importance of retail analytics made it easy for retailers to know their customers better. With a large amount of data, analytics has become more important to make decisions. The data can help retailers to understand customer behavior, plan and promote products, increase sales, improve customer experience. There are many algorithms and techniques used in retail that help uncover better insights and predict future events. One of the keys and widely used techniques in retail is Market Basket Analysis. This paper talks about Market Basket Analysis using R and spark through FP-Growth Algorithm and highlights the importance of such techniques in retail to boost sales.

2 Literature Review

In this section, we have concentrated on presenting different areas where data mining algorithms are used.

2.1 A Survey on Association Rule Mining in Market Basket Analysis

Data mining refers to extracting knowledge from large amount of data. Market basket analysis is a data mining technique to discover associations between datasets. Association rule mining identifies relationship between a large set of data items. When large quantity of data is constantly obtained and stored in databases, several industries are becoming concerned in mining association rules from their databases. For example, the detection of interesting association relationships between large quantities of business transaction data can help in catalog design, cross-marketing and various business decision making processes [8].

2.2 Market basket analysis using apriori algorithm to find consumer patterns in buying goods through transaction data (case study of Mizan computer retail stores)

Mizan Computer Shop is a shop that is engaged in the trading sector, especially in the field of selling computers and supporting accessories. Growing and increasing number of business actors in the computer sector, can makes the players challenged to be able to create unique differentiation and clear positioning. So, that consumers can differentiate from their competitors. Competitive and dynamic market conditions make every company should always observe competition in their business environment. Retail stores need to use all of available resources including data. Data processing is expected to be able to provide information that can be used to support marketing strategies. One of the data processing methods that are often used in marketing strategies is the use of data mining techniques i.e Market Basket Analysis using a priori algorithm [3].

2.3 Application of market–basket analysis on healthcare

Data analysis plays a vital role in the present era as it helps us to understand the patterns by exploring it in meaningful ways. Market—basket is one of the main methods used to find frequently occurring items in a transactional database and many researchers use the Apriori algorithm for this purpose[4].

3 Methodology

There are two major sections:

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

- Data preparation
- Exploratory Data Analysis and Data Visualization
- Rules generation by applying FP-Growth Algorithm.

An R program has been developed and implemented in R Studio environment. The following packages were imported:

- dplyr: A Grammar of Data Manipulation
- sparklyr: R interface for Apache Spark
- ggplot2: Data visualization

3.1 Data Pre-Processing

The data used in the study are transactions of customers in a UK based online retail outfit and it taken from UCI ML Repository, it contains 8 column fields and 541910 transactions and 4070 different products. This ".csv" is read using read.csv("file_name") function and stored as a data frame.

```
df <- read.csv("online_retail_II.csv")
head(df)

##   Invoice StockCode                     Description Quantity    InvoiceDate
## 1  536365    85123A  WHITE HANGING HEART T-LIGHT HOLDER        6 12/1/2010 8:26
## 2  536365     71053                 WHITE METAL LANTERN        6 12/1/2010 8:26
## 3  536365    84406B      CREAM CUPID HEARTS COAT HANGER        8 12/1/2010 8:26
## 4  536365    84029G KNITTED UNION FLAG HOT WATER BOTTLE        6 12/1/2010 8:26
## 5  536365    84029E      RED WOOLLY HOTTIE WHITE HEART.        6 12/1/2010 8:26
##   Price Customer.ID        Country
## 1  2.55       17850 United Kingdom
## 2  3.39       17850 United Kingdom
## 3  2.75       17850 United Kingdom
## 4  3.39       17850 United Kingdom
## 5  3.39       17850 United Kingdom

#To get Number of transactions
count(df)

##        n
## 1 541910

# to get number of poducts
df%>% distinct(StockCode) %>% summarize(count = n())

##   count
## 1  4070
```

Before applying algorithms, the dataset has to be prepared by doing the following data preparation steps:

1. Dropping all Null values.
2. Dropping all the duplicate items, if one item duplicates in one transaction then just keep one item
3. Dropping all transactions with just one item.
4. Convert data to transaction structure (Transactional Data) because of this type of data structure can't be used as input to the algorithm.

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

	items
1	c("85123A", "71053", "84406B", "84029G", "84029E", "22752", "21730")
2	c("22633", "22632")
3	c("84879", "22745", "22748", "22749", "22310", "84969", "22623", "22622", "21754", "21755", "21777", "48187")
4	c("22960", "22913", "22912", "22914")
5	c("22728", "22727", "22726", "21724", "21883", "10002", "21791", "21035", "22326", "22629", "22659", "22631", "22661", "21731", "22900", "21913", "22540", "22544", "22492"...
6	c("22632", "22633")
7	c("85123A", "71053", "84406B", "20679", "37370", "21871", "21071", "21068", "82483", "82486", "82482", "82494L", "84029G", "84029E", "22752", "21730")
8	c("85123A", "71053", "84406B", "20679", "37370", "21871", "21071", "21068", "82483", "82486", "82482", "82494L", "84029G", "84029E", "22752", "21730")
9	c("22114", "21733")
10	c("22632", "22633")

Figure 1: Prepared Data

After previous preparation steps, we have

3.2 Exploratory Data Analysis and Data Visualization
In this section, we use visualization and transformation to explore data in a systematic way.

- **Most frequent items, top (10)**

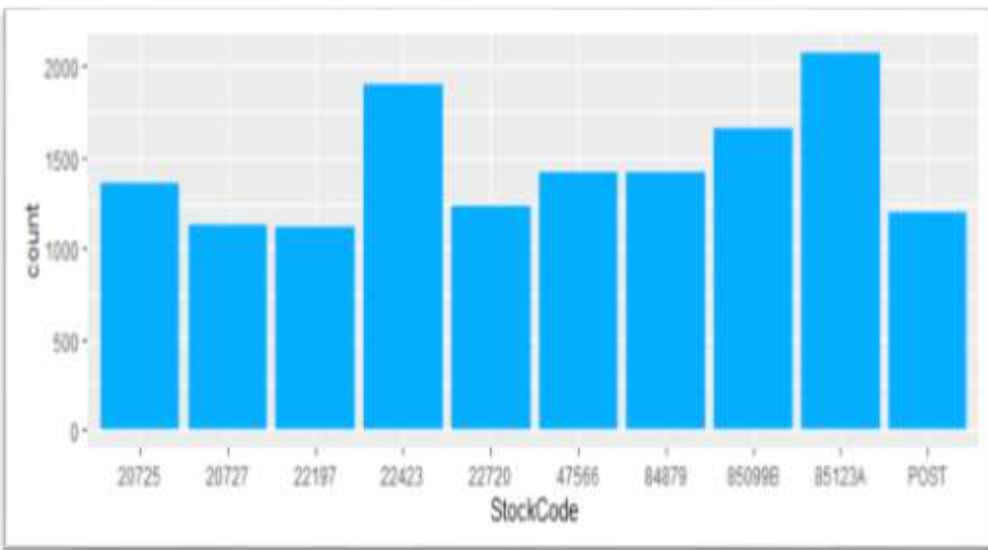

Figure 2: Most frequent items top (10)

From graph we see that most frequently item is that item with StockCode equal 85123A and its name is WHITE HANGING HEART T-LIGHT HOLDER

- **Number of items in each transaction**

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

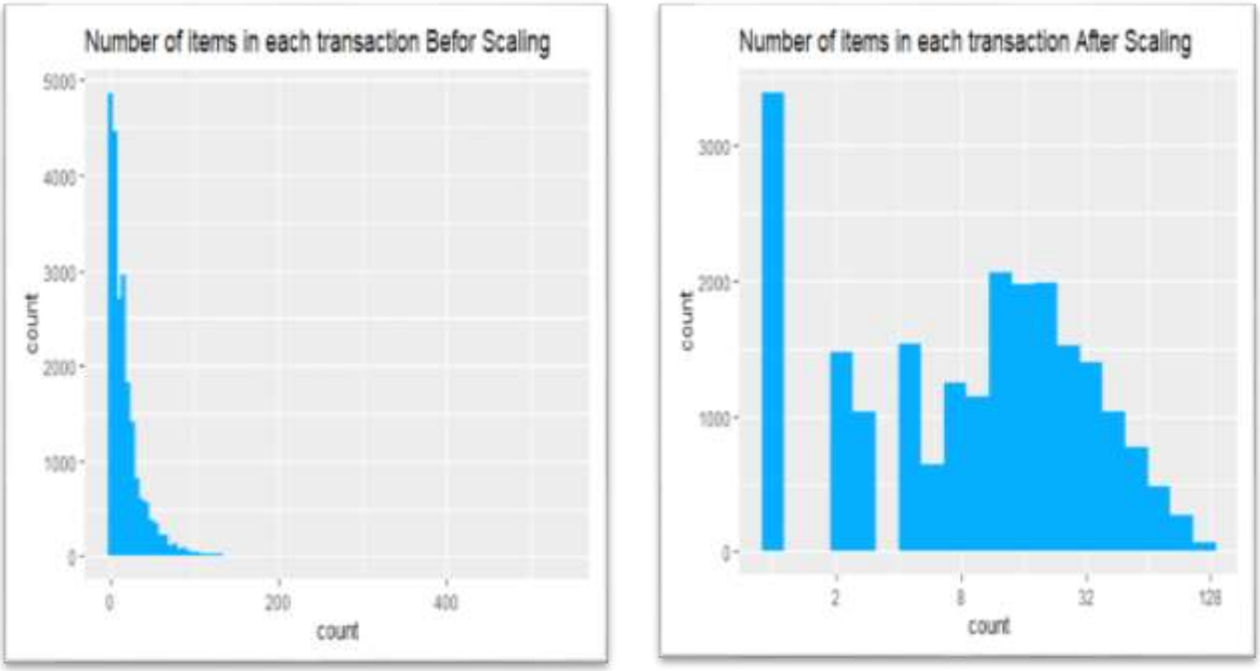

Figure 3: Number of Items in each transaction

From graph we can insight that a lot of transactions has only one item, and there are a lot of transactions have item count between (8,32) items.

3.3 Market Basket Analysis

Market basket analysis is also known as association analysis is an unsupervised machine learning method, is one of the key techniques used by large retailers to uncover associations between items. It works by looking for combinations of items that occur together frequently in transactions. To put it another way, it allows retailers to identify relationships between the items that people buy and help identify customer behavior and pattern and also Optimize marketing campaigns and strategies.

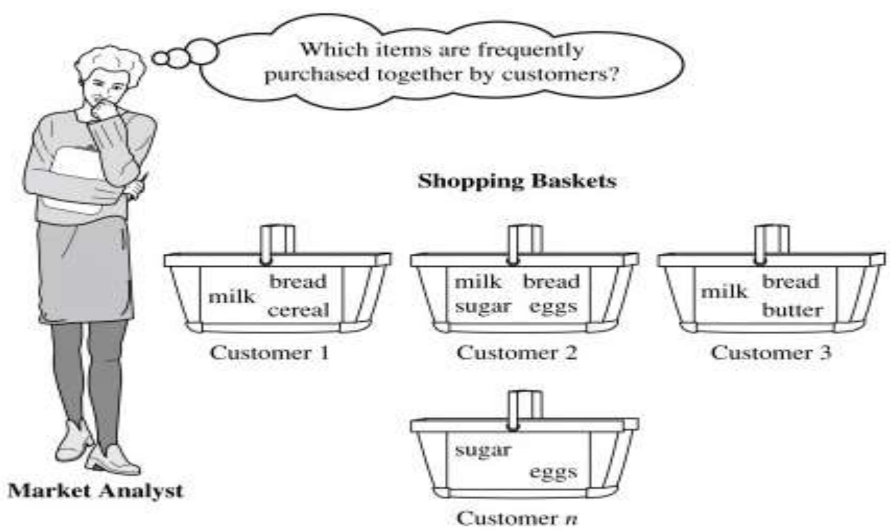

Figure 4: Overall picture of market basket analysis

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

Market basket analysis gives us insight into the goods by telling us which products tend to be purchased together and which are most enable to purchase. FP-growth algorithm is an efficient algorithm for mining frequent patterns. It does not need to produce the candidate sets and that is quite time consuming. It scans database only twice and frequent item set is mining by using of FP tree. In this paper, R and spark is used to find association rules.

3.4 Association Rule Mining

Association Rule Mining is a rule-based machine learning method to find associations and relationships between large sets of items, this rule also shows how frequently an item occurs in the itemset based on the occurrences of other items in a transaction, Association rules are widely used to analyze basket or transaction data to discover strong rules based on the interestingness and frequency of occurrences and it can be understood as the “if this, then that” rule. For example, if a user buys coffee and sugar, then he/she is likely to buy milk. This rule could be written as: If {A} Then {B}, Here, if part of the rule is known as antecedent and THEN part of the rule is known as consequent. {A} part is the condition and {B} part is considered as the result. These rules are applied to hundreds and thousands of records to obtain closer and accurate results. But it is not considered significantly accurate if applied to a small set of data.

3.5 SUPPORT, CONFIDENCE, LIFT

The association rule has primarily three measures to decide the degree of confidence, these are: Support, Confidence and Lift.

Support: This is one of the important measures to determine how frequently an itemset occurs in the transaction as a percentage of all transactions, Support is the number of transactions that include both {A} and {B} parts as a percentage of the total number of transactions.

$$support = \frac{(A+B)}{Total}$$

Confidence: This rule is the ratio of the number of transactions that include items in {A} and {B} to the number of transactions that include items in {A}. It can be understood as to how often items in B appear in transactions that contain A only. It is a conditional probability.

$$confidence = \frac{(A+B)}{A}$$

Lift: This third measure, lift or lift ratio is the ratio of confidence to expected confidence. Greater lift value tells how strong the association is. It shows us the rate of confidence that B will be purchased given that A was purchased.

$$lift = \frac{\frac{(A+B)}{A}}{\frac{B}{Total}}$$

3.6 Implementation

After Pre-processing the data, now data is ready for modeling, we use Spark through R by using sparklyr (R interface for Apache Spark), the sparklyr library provide ml_fpgrowth
A parallel FP-growth algorithm to mine frequent item sets, with two parameters

- min_support: Minimal support level of the frequent pattern.
- min_confidence: Minimal confidence for generating Association Rule

```
# A parallel FP-growth algorithm to mine frequent itemsets.
#min_support Minimal support level of the frequent pattern.
#min_confidence Minimal confidence for generating Association Rule
fp_model <- items_tbl %>%
  ml_fpgrowth(min_support = 0.01, min_confidence = 0.6)
```

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

After model created successfully, we can call two methods and pass the model as parameter

- The first one is ml_freq_itemsets to get frequency for items
- The second one is ml_association_rules to get generated rules with support, confidence, lift for each one.

	items	freq
1	list("85123A")	1993
2	list("22423")	1812
3	list("85099B")	1624
4	list("84879")	1382
5	list("47566")	1381
6	list("20725")	1322
7	list("22720")	1195
8	list("23203")	1092
9	list("POST")	1077
10	list("20727")	1067

Figure 5: item sets frequency

As we show in [figure 6] there are a lot of confidant association rule was generated by the model

For example, the first rule indicates that item [23171] will be purchased given that item [23170] and item [23172] were purchased with rate of confidence **94%.**

Model has generated 151 different rules, these rules allow retailers to identify relationships between the items that people buy and help identify customer behavior and pattern and also Optimize marketing campaigns and strategies

	antecedent	consequent	confidence	lift	support
1	list("23172", "23170")	list("23171")	0.9400000	64.345401	0.01002346
2	list("23172", "23171")	list("23170")	0.9126214	50.943828	0.01002346
3	list("22698", "22699", "22423")	list("22697")	0.8892989	22.911662	0.01284922
4	list("23172")	list("23171")	0.8841202	60.520284	0.01098315
5	list("22698", "22699")	list("22697")	0.8802661	22.678943	0.02116656
6	list("22698", "22697", "22423")	list("22699")	0.8795620	19.996443	0.01284922
7	list("22698", "22423")	list("22697")	0.8589342	22.129353	0.01460866
8	list("23172")	list("23170")	0.8583691	47.915389	0.01066325
9	list("22746")	list("22748")	0.8503937	45.701961	0.01151631
10	list("22698", "22423")	list("22699")	0.8495298	19.313673	0.01444871

Figure 6: Generated association rules top (10)

International Journal of Academic Engineering Research (IJAER)
ISSN: 2643-9085
Vol. 5 Issue 11, November - 2021, Pages: 5-13

3.7 EXPERIMENTAL RESULT

List Of All generated association rules:

	antecedent	consequent	confidence	lift	support
1	list("22384", "20728", "20727")	list("20725")	0.7394636	10.491210	0.01029004
2	list("23254")	list("23256")	0.7711268	43.695630	0.01167626
3	list("22384", "22382")	list("20728")	0.6081081	11.579366	0.01199616
4	list("22384", "22382")	list("20725")	0.6837838	9.701247	0.01348902
5	list("22384", "22382")	list("20727")	0.6108108	10.736989	0.01204948
6	list("21094")	list("21086")	0.7261538	48.126295	0.01258264
7	list("21094")	list("21080")	0.7046154	18.029695	0.01220943
8	list("20728", "20727", "20725")	list("22384")	0.7338403	14.518891	0.01029004
9	list("21231")	list("21232")	0.6379310	19.022312	0.01380891
10	list("22384", "850998")	list("20725")	0.6773050	9.609328	0.01018341
11	list("82581")	list("82580")	0.7394137	35.835770	0.01210279
12	list("20728", "20725")	list("22384")	0.6051502	11.972782	0.01503519
13	list("47590B")	list("47590A")	0.7071823	37.258179	0.01364897
14	list("23199", "23202")	list("23203")	0.6608392	11.350457	0.01007678
15	list("22698", "22423")	list("22697")	0.8589342	22.129353	0.01460866
16	list("22698", "22423")	list("22699")	0.8495298	19.313673	0.01444871
17	list("22699")	list("22697")	0.6751515	17.394426	0.02969716
18	list("20728", "22382")	list("20725")	0.6000000	8.512557	0.01263596
19	list("20728", "22382")	list("22383")	0.6101266	10.785612	0.01284922
20	list("22142")	list("22144")	0.6335404	28.091450	0.01087652
21	list("23199", "23203")	list("23200")	0.6018519	23.815049	0.01039667
22	list("22617")	list("22138")	0.7188940	15.534074	0.01663468
23	list("22662", "20725")	list("22383")	0.6569579	11.613481	0.01082320
24	list("22662", "20725")	list("22382")	0.6925566	13.002595	0.01140968
25	list("22384", "23209")	list("20725")	0.6891026	9.776708	0.01146300
26	list("22728")	list("22727")	0.6458333	13.489143	0.02148646
27	list("23294")	list("23293")	0.6915254	27.714211	0.01087652
28	list("23294")	list("23295")	0.6610169	34.060533	0.01039667
29	list("23200", "23203")	list("23199")	0.7169118	18.294418	0.01039667
30	list("22745")	list("22746")	0.6509434	48.067301	0.01103647
31	list("22745")	list("22748")	0.8018868	43.095097	0.01359565
32	list("22698", "22697", "22423")	list("22699")	0.8795620	19.996443	0.01284922
33	list("23343")	list("23344")	0.6382429	20.289633	0.01316912
34	list("21136")	list("84879")	0.7252125	9.842319	0.01364897
35	list("22748")	list("22746")	0.6189112	45.701961	0.01151631
36	list("22748")	list("22745")	0.7306590	43.095097	0.01359565
37	list("23206", "22382")	list("20725")	0.6314103	8.958193	0.01050331
38	list("22662", "22383")	list("20725")	0.6744186	9.568378	0.01082320
39	list("22662", "22383")	list("22382")	0.6810631	12.786807	0.01092984
40	list("23200", "23199")	list("23203")	0.6170886	10.599005	0.01039667
41	list("84997C")	list("84997D")	0.7508197	32.979798	0.01220943
42	list("22630")	list("22629")	0.6816720	17.982335	0.02260610
43	list("21086")	list("21080")	0.6784452	17.360053	0.01023672
44	list("21086")	list("21094")	0.8339223	48.126295	0.01258264
45	list("23172", "23171")	list("23170")	0.9126214	50.943828	0.01002346
46	list("23206", "22384")	list("20725")	0.6712329	9.523180	0.01044999
47	list("21931", "22411")	list("850998")	0.7401575	8.548272	0.01002346
48	list("20726", "23206")	list("20725")	0.6500000	9.221936	0.01039667
49	list("22699", "22423")	list("22697")	0.7314815	18.845696	0.01684794
50	list("22699", "22423")	list("22698")	0.6273148	20.321100	0.01444871
51	list("22384", "20727")	list("20725")	0.6616702	9.387509	0.01647473
52	list("22662", "22382")	list("20725")	0.6011236	8.528498	0.01140968
53	list("23206", "20727")	list("22383")	0.6312500	11.159025	0.01076989
54	list("23206", "20727")	list("20725")	0.6375000	9.044592	0.01087652
55	list("20728", "22383")	list("20725")	0.6083151	8.630528	0.01482192
56	list("22730", "22726")	list("22727")	0.7826087	16.345889	0.01151631
57	list("22698")	list("22697")	0.8134715	20.958065	0.02511196
58	list("22698")	list("22699")	0.7789292	17.708601	0.02404564
59	list("22730", "22727")	list("22726")	0.6792453	15.984849	0.01151631
60	list("22804")	list("85123A")	0.7331731	6.899847	0.01626146
61	list("23171")	list("23172")	0.7518248	60.520284	0.01098315
62	list("23171")	list("23170")	0.8467153	47.264859	0.01236938
63	list("85099F", "22386")	list("850998")	0.7925170	9.152986	0.01242269
64	list("22730")	list("22727")	0.6666667	13.924276	0.01695457
65	list("23200")	list("23199")	0.6666667	17.012245	0.01684794
66	list("22411", "22386")	list("850998")	0.7661290	8.848224	0.01013009
67	list("22382", "20727")	list("22383")	0.6014320	10.631912	0.01343570
68	list("22382", "20727")	list("20725")	0.6133652	8.702176	0.01370228
69	list("23172")	list("23171")	0.8841202	60.520284	0.01098315
70	list("23172")	list("23170")	0.8583691	47.915389	0.01066325
71	list("22698", "22699", "22423")	list("22697")	0.8892989	22.911662	0.01284922
72	list("22698", "22697")	list("22699")	0.8428875	19.162664	0.02116656
73	list("20726", "22384")	list("20725")	0.7430341	10.541866	0.01279591
74	list("20726", "22384")	list("20728")	0.6160991	11.731527	0.01060994
75	list("20726", "22384")	list("22383")	0.6130031	10.836462	0.01055662
76	list("20726", "22384")	list("20727")	0.6222910	10.938791	0.01071657
77	list("20726", "22384")	list("22382")	0.6284830	11.799626	0.01082320
78	list("23301")	list("23300")	0.6083551	17.828606	0.02484538
79	list("23209", "850998")	list("23203")	0.7898833	13.566896	0.01082320
80	list("22356")	list("20724")	0.6468254	16.089996	0.01738110
81	list("23295")	list("23293")	0.6785714	27.195055	0.01316912
82	list("21933")	list("21932")	0.6982456	40.798426	0.01060994
83	list("22729")	list("22726")	0.6021505	14.170559	0.01194284
84	list("22729")	list("22727")	0.6827957	14.261154	0.01354233
85	list("22569")	list("22570")	0.6847458	39.275509	0.01076989
86	list("22382", "22383")	list("20725")	0.6053215	8.588056	0.01455534
87	list("23170")	list("23171")	0.6904762	47.264859	0.01236938
88	list("22728", "22727")	list("22726")	0.6650124	15.649903	0.01428876
89	list("20728", "20727")	list("22384")	0.6141176	12.150201	0.01391555
90	list("20728", "20727")	list("20725")	0.6188235	8.779617	0.01402218
91	list("22697", "22699")	list("22698")	0.7127469	23.088567	0.02116656
92	list("23172", "23170")	list("23171")	0.9400000	64.345401	0.01002346
93	list("47590A")	list("47590B")	0.7191011	37.258179	0.01364897
94	list("22627")	list("22624")	0.6000000	18.034615	0.01007678
95	list("85099F")	list("850998")	0.6336336	7.318000	0.02249947
96	list("22578")	list("22577")	0.7356322	29.672080	0.01706121

	antecedent	consequent	confidence	lift	support
97	list("22579")	list("22577")	0.7256318	29.268709	0.01071657
98	list("22579")	list("22578")	0.8194946	35.334346	0.01210279
99	list("22577")	list("22578")	0.6881720	29.672080	0.01706121
100	list("20726", "20728")	list("20725")	0.6368564	9.035460	0.01252932
101	list("22725")	list("22726")	0.6374622	15.001558	0.01124973
102	list("22725")	list("22727")	0.6948640	14.513218	0.01226274
103	list("23174")	list("23173")	0.7007042	36.204982	0.01060994
104	list("23174")	list("23175")	0.7323944	49.059960	0.01108978
105	list("22698", "22697", "...	list("22423")	0.6070529	6.283601	0.01284922
106	list("23209", "20727")	list("20725")	0.6197605	8.792910	0.01103647
107	list("23175")	list("23174")	0.7428571	49.059960	0.01108978
108	list("23175")	list("23173")	0.7000000	36.168595	0.01044999
109	list("22697", "22423")	list("22699")	0.8250653	18.757484	0.01684794
110	list("22697", "22423")	list("22698")	0.7154047	23.174664	0.01460866
111	list("22726")	list("22727")	0.6725220	14.046572	0.02857752
112	list("21931", "22386")	list("85099B")	0.7852113	9.068610	0.01188953
113	list("23296")	list("23293")	0.6718266	26.924744	0.01156963
114	list("23209", "22383")	list("20725")	0.6404834	9.086919	0.01130305
115	list("23209", "22383")	list("23206")	0.6042296	12.937135	0.01066325
116	list("20726", "22382")	list("20725")	0.6356968	9.019009	0.01386223
117	list("23206", "22383")	list("20725")	0.6077348	8.622295	0.01172958
118	list("21733")	list("85123A")	0.6695652	6.301237	0.02463212
119	list("84596F")	list("84596B")	0.7820069	40.517464	0.01204948
120	list("22386")	list("85099B")	0.6288309	7.262532	0.02953721

	antecedent	consequent	confidence	lift	support
128	list("22728", "22726")	list("22727")	0.7790698	16.271974	0.01428876
129	list("23171", "23170")	list("23172")	0.8103448	65.231020	0.01002346
130	list("23256")	list("23254")	0.6616314	43.695630	0.01167626
131	list("23202", "23209")	list("23203")	0.8306452	14.267015	0.01098315
132	list("22384", "20728")	list("20725")	0.6527778	9.261346	0.01503519
133	list("22384", "20728")	list("20727")	0.6041667	10.620197	0.01391555
134	list("22570")	list("22569")	0.6177370	39.275509	0.01076989
135	list("22570")	list("22568")	0.6207951	27.526319	0.01082320
136	list("22386", "23203")	list("85099B")	0.7255639	8.379727	0.01029004
137	list("22750")	list("22749")	0.6279683	29.372005	0.01268927
138	list("21668")	list("21669")	0.6268222	32.566971	0.01146300
139	list("22384", "22383")	list("20725")	0.7004950	9.938340	0.01508851
140	list("22384", "22383")	list("20727")	0.6039604	10.616571	0.01300917
141	list("22697")	list("22699")	0.7651099	17.394426	0.02969716
142	list("22697")	list("22698")	0.6469780	20.958065	0.02511196
143	list("22384", "20728", "...	list("20727")	0.6843972	12.030509	0.01029004
144	list("22697", "22699", "...	list("22698")	0.7626582	24.705385	0.01284922
145	list("84596B")	list("84596F")	0.6243094	40.517464	0.01204948
146	list("22383", "20727")	list("20725")	0.6038544	8.567241	0.01503519
147	list("20726", "20727")	list("20725")	0.6590909	9.350915	0.01236938
148	list("22384", "20727", "...	list("20728")	0.6245955	11.893312	0.01029004
149	list("22910")	list("22086")	0.6467236	12.390142	0.02420559
150	list("22746")	list("22748")	0.8503937	45.701961	0.01151631
151	list("22746")	list("22745")	0.8149606	48.067301	0.01103647

Conclusion

This study was conducted in order to make a Market Basket Analysis by using Association Rules mining through FP-Growth Algorithm. Data were analyzed in the R Studio program and Spark using a data set containing 541910 transactions and 4070 different products. The best rule accordingly a customer who buys Regency Tea Plate Pink and Regency Tea Plate Roses also gets Regency Tea Plate Green with confidence 0.94 and the lift ratio value 64.3 As a result, product placement in the supermarket can be made according to these rules. Thus, sales of these products will increase and revenue will increase directly.

REFERENCES

[1] Jiawei Han and Micheline Kamber ,Data Mining: Concepts and Techniques, 2nd ,March 2006,Chapter .

[2] Journal, IRJET. "IRJET- Market Basket Analysis Using Data Mining Techniques." IRJET, 2021.

[3] M Qisman et al 2021 J. Phys.: Conf. Ser. 1722 012020

[4] Rao, A.B., Kiran, J.S. & G, P. Application of market–basket analysis on healthcare. Int J Syst Assur Eng Manag (2021). https://doi.org/10.1007/s13198-021-01298-2

[5] S. K. Solanki and J. T. Patel, "A Survey on Association Rule Mining," 2015 Fifth International Conference on Advanced Computing & Communication Technologies, 2015, pp. 212-216, doi: 10.1109/ACCT.2015.69.

[6] Chen, D., Sain, S. & Guo, K. Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. J Database Mark Cust Strategy Manag 19, 197–208 (2012).

[7] Karthiyayini, R.. "Affinity Analysis and Association Rule Mining using Apriori Algorithm in Market Basket Analysis." (2016).

[8] R. Agrawal, R. Srikant, "Fast algorithms for mining association Rules", Proceedings of the 20th VLDB Conference, Santiago, Chile, 1994, pp. 478–499.