Advances in Economics and Business Management (AEBM)
Print ISSN: 2394-1545; Online ISSN: 2394-1553; Volume 2, Number 3; January-March, 2015 pp. 241-245

http://www.krishisanskriti.org/aebm.html

Analytical Prescription for Unorganized Retail in India

Shubhradeep Nandi

Indian Institute of Management Bangalore
E-mail: shunandi@gmail.com

Abstract—*This is a paper to evaluate if analytical prescriptions can help the unorganized retail sector in India to optimize its cost and efficiently utilize its resources. Every month enormous amount of natural resources are consumed and converted to an irreversible form i.e. a consumer consumable product. Post production they are shipped to various retail outlets. Many of these shipped items are either recycled or disposed because there is a gap between 'items in demand' and 'items stocked by the retailers'. Non-packaged perishable food (NPPFP) items like potato, onion, tomato are lesser durable, if it has no taker(s) within its shelf life it ends up in rot dumps. We made a diligent analytical effort to bridge this gap using statistical analysis. This work acts as a stepping stone to develop customizable statistical models that can be efficiently applied across any geography and under multiple correlated factors. We attempted to determine how an unorganized retailer in a particular locality can intelligently decide on a) the set of products to shelf, b) the brands to choose, c) SKU to stock, and, d) when to replenish.*

1. INTRODUCTION

Most of the retail sector in India is unorganized. It is dominated by a large number of small retailers. Local Kirana shops, owner-managed general stores, chemists, footwear shops, apparel shops, pan and beedi shops, hand-cart hawkers, pavement vendors, etc. together make up the so- called "unorganized retail" or traditional retail.

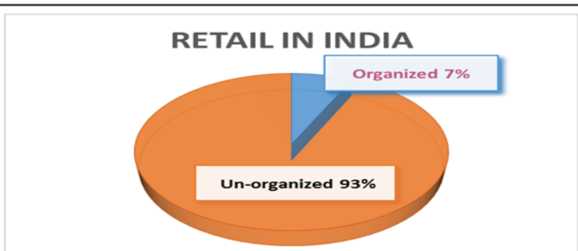

Figure 1. Retail distribution

The retail sector contributes to over 20 percent of Indian GDP [1]. Out of this, the Organized Retail(henceforth referred as OR) sector accounts only for about 7 percent share, and the remaining 93% is contributed by the Un-Organized Retail (henceforth referred as UOR) sector which is mostly a family owned business in India.

2. OPPORTUNITIES AND CHALLENGES

Opportunity
• Favourable demography dynamics.
• Proportionate increase in earning and spending.
• Better transportation and infrastructure coming up.
• Availability of cheaper technology.
Challenges
• Lack of best practice in inventory management and supply chain management.
• Lack of process standardization.
• Lack of knowledge, skills and training.
• Stiff competition from organized retail sector.
• Lack of data based analytic acumen.

3. PROBLEM STATEMENT

The major problems that an un-organized retailer would like to get resolved are mentioned below:-

1. What are the associated set of products that my customer need?
2. Why certain products expire or rot in my shelf? Is there a way to reduce this wastage yet have sufficient SKU to meet the demand?
3. Why certain brand have no takers, they give me good margin and I can make a fortune if people take it? With which brand should I replace them?
4. When should I replenish a product so that I use minimum shelf space, have minimum wastage, and, achieve maximum profit by meeting the demand?

In this project we have analytically derived an approach to solve these questions.

4. DERIVING THE CATALOGUE OF ASSOCIATED PRODUCTS

Our market survey showed that 87% of consumers were always delighted to shop all their needs under one roof. This is also one of the major reason they remained loyal and gave repeat business to a retailer. Hence stocking the right catalogue of products that meets your consumer's need is of utmost importance. We investigated the daily sale basket conceived from the approach of market basket analysis [3], of an UOR and derived the approach to deduce a catalogue of associated products that a retailer needs to stock.

In our experiment we primarily focused on two categories

a) Non-Packaged Perishable Food Products. (NPPFP)

b) Packaged Perishable Food Products. (PPFP)

A single transaction for this experiment is an accumulation of items sold in a single sale day.

The Data set consists of a combined 89 transactions of two unorganized retail stores. We collected 2 months of consecutive sales data from a UOR store in JP Nagar, and one month sales data from a store in Madiwala, Bangalore.

We mined the accumulated daily sales data by evaluating multiple association algorithms and finally resorting to Apriori Algorithm [2].

Our execution was split in to 2 steps based on the location from which this data is collected. As mentioned above we have collected data from two localities in Bangalore:-

a) JP Nagar(12.9120° N, 77.5930° E)

b) Madiwala(12.9200° N, 77.6200° E)

Execution 1:- Our dataset has 59 transactions collected from an UOR in JP Nagar. We mined the data using with parameters (minsup=10%, minconf=80%, lift ratio >=1)

Generated correlations are as follows:-

Experiment Inferences of execution 1

1. {Snacks} and {Beverage} together comprises of 10% of the total transactions and in 60% of the purchases where consumer buy {Snacks} they also opted for {Beverage}.

2. {Onion} and {Potatoes} together comprises of 14% of the total transactions and in 62% of the purchases where consumer buy {Onions}/ {Potatoes} they also opted for {Potatoes}/ {Onions}.

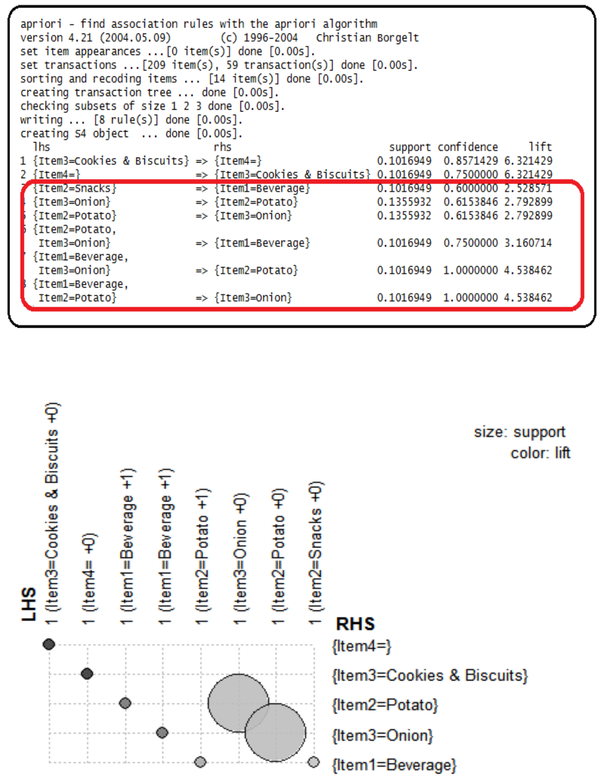

Figure 2: Experiment results of execution 1

3. {Cookies and Biscuits} alone compromises about 10% of the transactions and in 86% of the purchases consumers have chosen {Cookies and Biscuits} alone and not as a combo.

Execution 2:- Our dataset has 30 transactions collected from an UOR in Madiwala. We mined the data using parameters (minsup =10%, minconf=80%, lift ratio >=1)

Generated correlations are as follows:-

Experiment Inferences of execution 2

1. {Snacks} and {Beverage} together comprises of 10% of the total transactions and in almost 100% of the purchases where consumer buy {Snack} they also opted for {Beverage}.
2. {Onion} and {Potato} together comprises of 23% of the total transactions and in 100% of the purchases where consumer buy {Onions}/ {Potatoes} they also opted for {Potatoes}/ {Onions}.
3. {Bread} and {Milk} alone compromises about 11% of the transactions and in 90% of the purchases where consumer buy {Bread}/ {Milk} they also opted for {Milk}/ {Bread}.

```
apriori - find association rules with the apriori algorithm
version 4.21 (2004.05.09)        (c) 1996-2004   Christian Borgelt
set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[88 item(s), 30 transaction(s)] done [0.00s].
sorting and recoding items ... [20 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 done [0.00s].
writing ... [16 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
   lhs                 rhs               support   confidence     lift
1  {Item5=Tomato} => {Item4=Onion}     0.1000000          1.0 4.285714
2  {Item3=Bread}  => {Item1=Milk}      0.1000000          1.0 2.727273
3  {Item2=Snacks} => {Item1=Beverage} 0.1000000          1.0 2.500000
4  {Item4=Tomato} => {Item3=Onion}     0.1333333          0.8 3.428571
5  {Item4=Tomato} => {Item2=Potato}    0.1333333          0.8 3.428571
6  {Item4=Tomato} => {Item1=Beverage} 0.1333333          0.8 2.000000
7  {Item2=Bread}  => {Item1=Milk}      0.1333333          0.8 2.181818
8  {Item4=Ginger} => {Item3=Garlic}    0.1333333          0.8 4.000000
9  {Item3=Onion}  => {Item2=Potato}    0.2333333          1.0 4.285714
10 {Item2=Potato} => {Item3=Onion}     0.2333333          1.0 4.285714
11 {Item3=Onion,
    Item4=Tomato} => {Item2=Potato}    0.1333333          1.0 4.285714
```

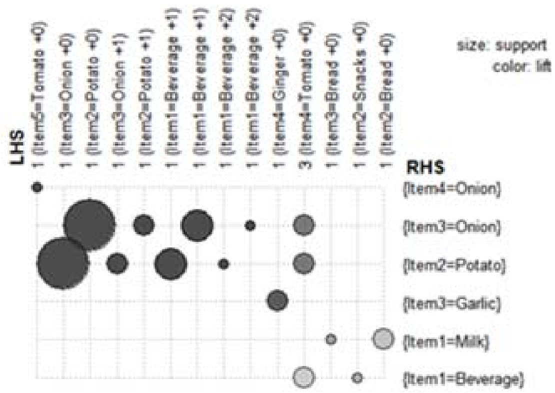

Figure 3: Experiment results of execution 2

Combining the dataset of execution 1 and execution 2 we also derived the accumulated proportional results. The accumulated proportional results even though diluted the individual inferences but we were able to derive a concrete correlation pattern of the consumer purchase.

Composite experiment inferences

1. Consider the combo {lhs = {Potato, Beverage} and rhs= Onion}, {lhs = {Onion, Beverage}, rhs= potato} each comprises of 12% of the total transactions and almost 100% of the purchases where consumer buy {Potato, Beverage}/{Onion, Beverage} they also opted for{Onion}/{potato} but

a) On a choice of {lhs = {Potato, Onion}, rhs= {Beverage}} only 70 % of the purchases were done by consumers

b) And, with {lhs=potato, rhs=onion} and {lhs=onion, rhs=potato} 75% of the purchases were done by the consumer.

The combination of {Potato, Onion} with unavailability of beverage will still attract 75-80% of the purchases

2. Cookies and Biscuits alone contributed to 6% of the total transactions and in 80% of the purchases done have gone alone.

```
apriori - find association rules with the apriori algorithm
version 4.21 (2004.05.09)        (c) 1996-2004   Christian Borgelt
set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[274 item(s), 89 transaction(s)] done [0.00s].
sorting and recoding items ... [24 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 done [0.00s].
writing ... [10 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
   lhs                              rhs                              support    confidence lift
1  {Item3=Cookies & Biscuits} => {Item4=}                          0.06741573 0.8571429 9.535714
2  {Item4=}                    => {Item3=Cookies & Biscuits} 0.06741573 0.7500000 9.535714
3  {Item4=Garlic}              => {Item3=Onion}                   0.05617978 0.7142857 3.178571
4  {Item4=Garlic}              => {Item1=Beverage}                0.05617978 0.7142857 2.445055
5  {Item4=Onion}               => {Item3=Potato}                  0.08988764 0.8000000 4.746667
6  {Item2=Potato}              => {Item3=Onion}                   0.16853933 0.7500000 3.337500
7  {Item3=Onion}               => {Item2=Potato}                  0.16853933 0.7500000 3.337500
8  {Item2=Potato,
    Item3=Onion}               => {Item1=Beverage}                0.12359551 0.7333333 2.510256
9  {Item1=Beverage,
    Item2=Potato}              => {Item3=Onion}                   0.12359551 1.0000000 4.450000
10 {Item1=Beverage,
    Item3=Onion}               => {Item2=Potato}                  0.12359551 1.0000000 4.450000
```

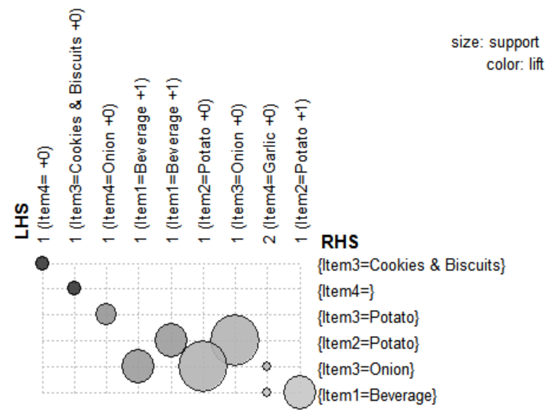

Figure 4: Experiment results of execution 3

3. {lhs=Garlic and rhs=Onion}, {lhs=Garlic and rhs=Beverage} individually contributed to more than 5% of the total transactions and contributed to the 70% of total sales.

Hence we found that there is a concrete correlation between {Onion, Potato}, {Milk, Bread}, and, {Snack, Beverage} as each of these items are primarily bought in pairs whereas {cookies and biscuits} sell alone.

5. PRESCRIBING THE PROFIT MAXIMIZING SKU

The Profit Maximizing SKU for a particular item is where Sales(S) meet Demand (D) coupled with minimum decay (d).

In order to simulate this condition we collected 31 day demand and sales data for Non-Packaged Perishable Food Product – Potato from an UOR in JP Nagar.

An UOR may not be able to meet the demand for an available item primarily due to 2 conditions:-

a) When Demand (D) > available SKU(S)
b) When available SKU(S) > = Demand(S) but Sales(S) is a combination of Decayed (d) + Sellable(s) items and S – d < D or s < D
Hence from the above two conditions we can derive that
Profit Maximizing daily SKU or Prescribed SKU = Ceil {Daily Demand + Daily Decay}

Execution and Results

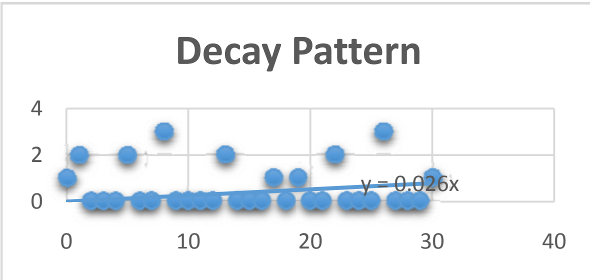

Figure 5: Decay time correlation

STEP 1. I derived the decay time correlation

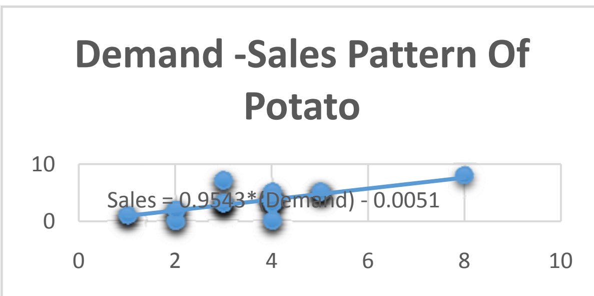

Figure 6: Demand sales correlation

Step 2. I derived the Current Demand Sales correlation

Step 3. Derived the Profit Maximizing SKU using decay time correlation and Demand Sales correlation

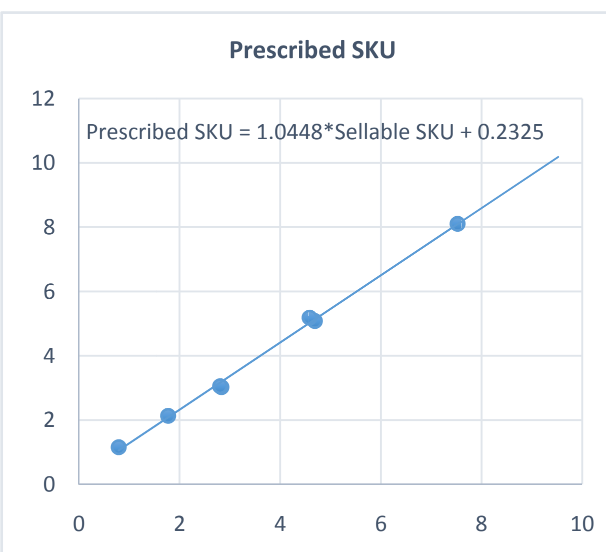

Figure 7: Prescribed SKU

Hence for this UOR considering the product type, the demand for the product, the decay and the storage facility available the prescribed daily SKU came out as *1.0448*Sellable SKU + 0.2325*

6. PRESCRIBING THE BRAND

The choice of brand depends primarily on three dimensions and these dimensions are detailed below:-

1. Historical Sale Data- This includes the quantitative historical value of the many SKU's of a product that were sold during a particular period. Historical Sales(h) = Consumer Demand(D) - Substitution SKU(s) - Rejection Number (r)
2. Consumer Preference - This is the index derived from consumer basket observations. 7 Day observation data was collected from JP Nagar and Madiwala.
3. Retailer Margin - The retail margin equals the difference between the price that a retailer pays for an item and the price at which he/she sells the item to a consumer. The retail margin helps a retailer to determine which item should be sold and how it should be priced.

Now based on the above three dimensions I have designed the BrandPrescriptionModel algorithm in statistical computing and graphics programming language R which consumes structured data, processes it and provides a point based output for the various competing brands involved in this process.

We ran two scenarios to validate our findings

Scenario 1:- In a competitive market how an UOR can decide which brand to stock.

Product Type: - PPFP

Item: - Potato Chips and Snacks

Primary Data collected as part of our research:-

Brand Name	Consumer Brand Preference	Historical Sales	Retailer Margin
Bingo	34	553	15
Lays	21	373	11
Haldiram	18	273	20

Scenario 1:- In a competitive market how do you decide which brand to stock.

Based on the scenario chosen a strategist assigns appropriate weights to the involved factors as input to the algorithm.

Results from the algorithm:-

```
> source('D:/IIM/Retailer/BrandPrescribtionModel.R')
          [,1]      [,2] [,3] [,4]
[1,] 0.4657534 0.4612177    0    0
[2,] 0.2876712 0.3110926    0    0
[3,] 0.2465753 0.2276897    0    0
          [,1]      [,2] [,3] [,4]
[1,] 0.4657534 0.4612177 0.15 0.33
[2,] 0.2876712 0.3110926 0.11 0.33
[3,] 0.2465753 0.2276897 0.20 0.33
  Serial Number BrandName PreferencePoint
1             1     Bingo        43.32709
2             2      Lays        27.45884
3             3  Haldiram        23.81407
```

Figure 8: Automated output from algorithm in a competitive market

The algorithm clearly states that in a highly competitive market Bingo with a higher consumer preference is supposedly the best choice and our prescribed one.

Scenario 2:- In a monopolistic market based how do you decide which brand to stock.

Based on the scenario chosen a strategist assigns appropriate weights to the involved factors as input to the algorithm.

The algorithm nearly states that in a monopolistic market Haldiram with a very low consumer preference and high retailer margin is supposedly the best choice and our prescribed one.

Thus using the above algorithm, support of our primary data and inputs from strategists we can prescribe the appropriate brands to retailer.

7. PRESCRIBING THE CYCLE TIME TO REPLENISH (CT2R)

An optimum replenishment cycle is one that has -

- Least amount of variation or dispersion in total demand, and,
- Least wastage/decay of product.

If Cycle1 has amount of variation = $\sigma 1$ and amount of wastage = d1 and Cycle 2 amount of variation = $\sigma 2$ and amount of wastage = d2 and,

$\sigma 1 + d1 > \sigma 2 + d2$

Then, optimum replenishment cycle is Cycle 2.

In order to derive the optimum replenishment cycle for potato we took 31 day consecutive demand of Potato from an UOR and derived the decay-time relationship which is 0.0264x x=1 to n where n= Replenishment cycle period.

For NPPF products the 3 most followed replenishment cycles were chosen. 3-7-15 day implementation plan was designed and presented [4].

We executed accordingly and found the below results:-

3 Day Replenishment

Variance (σ) 2.790858
Decay (d) 0.1584
σ + d 2.94

7 Day Replenishment

Variance(σ) 2.160246899
Decay(d) 0.7392
σ + d 2.89

15 Day Replenishment

Variance(σ) 4.949747468
Decay(d) 3.168
σ + d 8.11

We found that a 7 day replenishment cycle is best as it had minimum variance and optimum decay. Also retailer needs to keep a maximum of 28 Kilos of Potato, a reduction in storage space from 50 Kilos. So we prescribe a reduction of CT2R from 15 to 7 days.

8. CONCLUSION AND FUTURE SCOPE

Unorganized retail sector is the largest channel through which products are traded in India. More than 90% of the retail sector in India is unorganized as of date. In this project we took a statistical approach to derive solution for the four basic problems faced by this channel. The outcome is a 'win-win' scenario for the producer, retailer and consumer based on their positioning.

The retailer plays the pivotal role because of his/her positioning, (s)he is the primary conveyer of the customer needs to the producer. We have derived an approach to help the retailer decide on 'the product and its brand' (s)he needs to sell, and, how many SKU of that products (s)he needs to stock so that (s)he meets maximum demand yet has optimum wastage.

My project is in its most humble form and researchers can take a cue from this or enthusiasts can fund me to carry on future research in order to create the gigantic statistical models that can mutate and evolve over time using the genetic algorithm to prescribe better. The evolution if appropriately done will reach a stage where the models shall prescribe the producer what to produce, and, prescribe the consumer what to buy.

9. ACKNOWLEDGEMENTS

This work was done under the valuable guidance of Assistant Professor Shankar Venkatagiri from IIM Bangalore.

REFERENCES

[1] Prof. Kalpana Singh., "*Retail Sector in India: Present Scenario, Emerging Opportunities and Challenges*", e-ISSN: 2278-487X, p-ISSN: 2319-7668. Volume 16, Issue 4. Ver. I (Apr. 2014), PP 72-81.

[2] Rakesh Agrawal, and Ramakrishnan Srikant, "*Fast Algorithms for Mining Association Rules*", in Proceedings VLDB Conference Santiago, Chile, 1994.

[3] Bill Qualls, *Introduction to Market Basket Analysis*, First Analytics, Raleigh, NC.

[4] Juan-Carlos Méndez García., "*Reduction of the Cycle Time to Replenishment*", in Management Consulting, solution for decision makers.