Big Data Retail Analysis and Product Distribution (BREAD) Model for Sales Prediction

** Riktesh Srivastava*

Abstract

Retailing is one of the most promising and significant commercial sectors of the world. But within retail, there is scale difference in the way it operates at both, businesses and store level. People believed that marketing is an art, however, advent of big data analysis added a scientific flavor to marketing. Retail companies are now using big data and analytics for every stage- identifying the products with predicting drifts, evaluating customers purchase behavior, forecasting demand trends of each product, thereby segmenting and targeting customers accurately. Companies are using algorithms and models for storing and using customer data for sales prediction. However, they are still facing difficulty in correctly mapping products with customers. In this study, a new technique called BREAD (Big Data Retail Analytics and Product Distribution) model was developed for product distribution for retailers. As an experiment, the model was used for product distribution of ABC Stores (name changed, as requested).The algorithm takes product details from each store unconnectedly (10 in the case) and maps it with demand forecasting and product visibility. After evaluating the two results, the algorithm further assesses the price for each product category (termed as price optimization) and devise strategies accordingly.

Keywords: Demand forecasting, FCM, k-means clustering algorithm, price optimization, product visibility.

I. INTRODUCTION

Almost all retail outlets (within the same retail segment) offer like products, incorporate similar IT tools and infrastructure and use almost similar business models. According to Statista [1], worldwide retail sector is estimated to grow at 3.4% every year, with multi-national retailers dominating the business. They operate in a competitive environment, through a range of hypermarkets, supermarkets, and convenience stores. In a state of such penetrating race, analytics can be a foremost differentiator, which can help retailers to make knowledgeable verdicts, confidently influence sales, and gain competitive advantage. Lavalle[2] found that the top performing companies in retail sector are three times more effective than those without analytics, making analytics a sole competitive differentiator. Data analytics lashes the passage from merchant-driven business models to digital models, where every decision is cognizant by data, which increases output and productivity by 5-6% [3]. However, Big Data Analysis also poses challenges of proper data management, as poor data management and analysis cost upto 35% of businesses operating revenue [4]. Conversely, proper and accurate data management increases the sale by 73% [5]. IBM forecasted 107% growth in use of analytics for retail sector from 30% in 2010 to 62% in 2012 [6], which it surpassed beyond anticipation. The report also illustrated that retailers were taking business-driven decisions and adopting reasonable methodology for big data analytics. Reports from EKN Research supported by SAS EKNResearch specified that 2 out of 5 retailers lag behind competitors in terms of their analytics maturity. The competition in retail sector is getting even tougher as non-retail players (Amazon.com, netgrocer.com, flipkart.com to name a few) are also in the fray and they trust big data and retail analytics profoundly [7].

Table I illustrates the market share of two types of retailers, traditional retail stores, and e-tailers operating in UAE, their market shares, and use of analytics by them

Surprisingly, e-tailers' market share in UAE accounted for just 1% of total retailing value sales, which is considerably a lower share in comparison to the international average of 15-20% [8]. However, the use of analytics is much higher (almost 100%)as compared to traditional retailers (≈10%).

Manuscript received November 10, 2017; revised December 10, 2017; accepted December 11, 2017. Date of publication January 6, 2018.
* R.Srivastava is with Skyline University College, University City of Sharjah, Sharjah, UAE -1797 (email: rsrivastava@skylineuniversity.ac.ae)
DOI: 10.17010/ijcs/2018/v3/i1/121853

TABLE I.

RETAIL INDUSTRY AND USE OF ANALYTICS

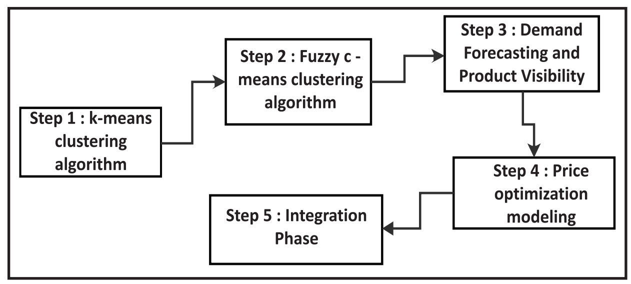
Types of Retail	Market Share	Use of Retail Analytics
Online Retaile (Euromonitor, 2016)	1%	100%
Grocery Store (Euromonitor, 2016)	31%	5%
Non-Grocery Stores (Electronics, Fashion, G&D (Euromonitor, 2016)	63%	5%
Miscellaneous	5%	0%

Retail analytics includes four different types of analysis i.e. price optimization, attract new customers, demand forecasting, and product consumer trends [9]. Levya [10] steered a research on price optimization for retail sector and concluded that optimal pricing is not a static problem. The study also described that retailers should react swiftly to changes in sales patterns. Fader [11] conducted a study pertaining to ways of attracting new customers based on three aspects, recency, frequency, and monetary value called RFM analysis. Fildes [12] established the study on demand forecasting as a crucial aspect of retail analytics. He analyzed demand forecasting for retail sector, and concluded that demand forecast adjustment is the only practical way for most organizations to improve their disaggregated sales forecasts. Kolyshkina [13] piloted a study for optimal utilization of retail analytics by focusing on key stages of analytics process. The study included the three aspects mentioned earlier and introduced an additional factor called predicting consumer trends. Adams [14] even enhanced the study by collecting data regarding which products constitute a customer's order, thereby analyzing ways of attracting customers, demand forecasting and predicting consumer trends. This study is often referred to as Shopping Basket Analysis (SBA).

The retail analysis piloted earlier were largely motivated on data puddle for research and was guided individually, thus limiting its scope. To overcome the limitations an integrated approach was required. The proposed BREAD model encompasses price optimization and demand forecasting from data collected from day-to-day operations of ABC stores as an integrated approach, and thus comprehends big data analysis. The model uses clustering, price optimization, demand forecasting, and product visibility based on the information collected from ABC stores. It then segments the data into different product types, thereby analyzing product wise sales in all outlets. Since, BREAD model is a big data analysis approach, the amount of data collected is also humongous.

BREAD model uses five steps, as mentioned in Fig. 1 to give recommendations regarding product wise sales status (store level and overall status), price optimization (min range and max range for each product type) and product visibility of all products types. Product visibility is an additional feature in the algorithm, owing to study conducted by Tyco [15] that incorrect inventory distortion costs retailers $800 billion a year.

Fig.1. Steps of BREAD model



As mentioned in Fig. 1, there are two levels of clustering (Step 1 and Step 2) analysis for BREAD model. The first level of clustering divides the entire dataset into n clusters (which are actually the number of outlets). The second level of clustering takes these n clusters into c clusters (for different product types). These c clusters are analyzed separately for demand forecasting, product visibility, and price optimization.

It thus,empirically evaluates the quality of clusters to ensure there is little or no deviation from the cluster centroids. Sum of Squared error (SSE) is used for the same, with standard deviation σ evaluated as

$$\sigma = \sqrt{\frac{1}{N-1} \Sigma^{N}_{i=1} (x_i - c)^2} \tag{1}$$

where $x_i = (1,2,\ldots,N)$ is an element in the cluster with N objects and c is the center of cluster.

$$SSE = \Sigma^{k}_{i=1} \Sigma_{x \in ci} (x_i - c)^2 \tag{2}$$

where, k is the number of cluster and c_i is the center of i^{th} cluster.

Rest of the paper is divided as follows. Section II describes steps for BREAD model. Section III mentions the BREAD strategies, and section IV illustrates the implementation of the model with four different outcomes. section V gives outcomes and recommendations accordingly.

II. BREAD MODEL STEPS

This section explains the five steps of the proposed BREAD model. These steps are executed in sequence to get the final outcome.

Step 1: Divide the entire dataset into k clusters

The first step is to distribute the entire dataset into k clusters in order to ease the analysis. K-means clustering algorithm is one of the naivest unsupervised clustering algorithm, where the clusters are placed at a distance from each other based on association with the centroid [16]. The algorithm continues to generate the clusters until no more clusters are possible.

If $X = \{x_1, x_2 \ldots\ldots\ldots x_n\}$ are set of data points and $V=\{v_1, v_2, \ldots\ldots\ldots v_n\}$ are the set of centers, then p clusters are generated as:

$$J(V) = \sum_{i=1}^{p}\sum_{j=1}^{p_i}(\|x_i - v_j\|)^2 \quad (3)$$

where,

$\|x_i - v_j\|$ is the Euclidean distance between x_i and v_j.

p_i is the number of data points in i^{th} cluster, and,

p is the number of cluster centers.

The clusters are recalculated as:

$$vi = \left(\frac{1}{p_i}\right)\sum_{j=1}^{p_i} x_i \quad (4)$$

Step 2: Clustering of substructures using Fuzzy c-means clustering algorithm (FCM)

Bezdek [17] introduced FCM, which generates fuzzy partitions of already created clusters called substructures. Once the first level of clustering is performed, the substructures based on different product types for each retail are created. Thus, within these substructures, different product types are identified. FCM is applied to n clusters generated from step 1 as:

(a) Randomly, p clusters are selected with $V = \{v_1, v_2, v_3, \ldots\ldots\ldots\ldots, v_n,\}$ to be the set of centres. The fuzzy membership for p clusters are evaluated as :

$$\mu_{ij} = 1/\Sigma^{p}_{k=1}\,(d_{ij}/d_{ik})^{(2/m-1)} \quad (5)$$

(b) Fuzzy centres v_j are calculated as:

$$v_j = (\Sigma^{p}_{k=1}(\mu_{ij})^{m}x_i)\,/\,(\Sigma^{n}_{i=1}(\mu_{ij})^{m}),\ \forall j = 1, 2, \ldots\ldots\ldots p \quad (6)$$

(c) Objective function is to minimize

$$J(U, V) = \Sigma^{n}_{i=1}\,\Sigma^{n}_{j=1}(\mu_{ij})^{m}\,\|x_i - v_j\|^2 \quad (7)$$

where,

n is the number of data points,

v_j represents the j^{th} cluster centre,

m is the fuzziness index $m \in [1, \infty]$,

p is the number of cluster centre,

μ_{ij} represents association of i^{th} data to j^{th} cluster,

d_{ij} is the Euclidean distance between i^{th} data and j^{th} cluster,

d_{ik} is the Euclidean distance between i^{th} data and k^{th} cluster.

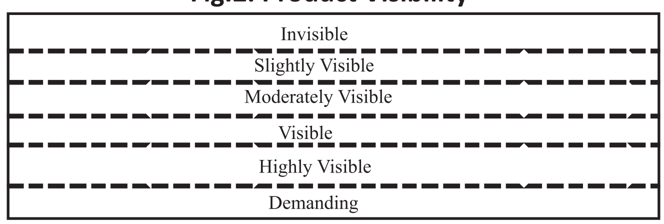
Step 3: Demand Forecasting and Product visibility (IV)

A key requirement for the execution of BREAD model is the ability to identify the status of demand and make predictions accordingly. The demand df_i for product type i depends on the number of sales in the past. If φ denotes all products that store sells for a given product type and $S(I)$ denotes the set of all product types s.t., $s \in S(i)$, then unique product type is represented as $(i, s) \in \varphi * S(i)$ and is denoted as is. Demand forecasting is thus represented as:

$$df_i = \Sigma_{s\in S(I)}\, df_{is} \quad (8)$$

It was also observed that demand for a particular product varies depending on either seasonal requirements or placement of different product types i in the shelf as represent in Fig. 2:

Fig.2. Product Visibility



Product Visibility (PV) is the concept used to evaluate the score of all products type i and is computed as:

$$PV_i = \frac{\Sigma^{6}_{i=1} i * \gamma_{6,i} * df_i}{\Sigma^{6}_{i=1} i * \gamma_{6,i} * df_i} \quad (9)$$

As specified in (9), df_i denotes the demand of different product type placed in six locations of shelf. As the dataset is imbalanced, and assuming that products placed at location 6 (or demanding) position are the most sellable product, the ratio $\gamma_{6,i}$ can be evaluated as:

$$\gamma_{6,i} = \frac{|Pv_6|}{|\Sigma^{5}_{i=1} PV_5|} \quad (10)$$

As given in (10), numerator is the product sold for the demanding PV position and the denominator for the other five PV positions.

Based on (9) and (10), the product visibility (PV) status range is given as follows:

TABLE II.
PV SCORE STATUS RANGE

PV Status Range		
Value	**Low**	**High**
Invisible	0	20
Slightly Visible	21	40
Moderately Visible	41	60
Visible	61	80
Highly Visible	81	100
Demanding	101	

Source:Author's calculation

Step 4: Price Optimization Modelling

All the outlets do not have the ability to buy inventory based on expected demand, rather, their purchasing decision is usually based on the requirements of customers. Price optimization modelling thus considers different products in outlets and the consolidated cost for each product type.

Let P be the set of products with P = $|p|$ and C is the possible price for each product types s.t. $C = |c|$. As an example for ABC stores, for product category soft drink (Low, Fat, and Regular), the possible set of prices for outlet 1 from dataset are:

c_1={32.89,47.50,52.93,55.66,57.39,85.58,107.09,115.08,141.48,142.51,143,151.8,153.56,159.95,160.22,165.3,172.14,184.39,189.98,190.45,196.67,241.85,243.54}

and,

c_2 = {40.982,163.887,183.992,189.053,233.996}

s.t. $C = c_1 \cup c_2$ respectively and $P = 2$. Let u_j represent the j^{th} price in the set C, where $j = 1,\ldots\ldots C$.

The objective of price optimization modelling is to select a price range for each product type, which is referred to as status range.

Status range formulation

Status range formulation identifies the min and max range for a product type for each outlet. In other words, the price of product depends on average prices of all the similar products, not the individual price of each product. If S represents the sum of all products types with $P = 2$, then, for C

$s_1 = 3191.077$ and $s_2 = 811.910$. We can evaluate $S = avg\,(s_1, s_2) = 2001.494$.

Thus, the relative price of first product from c_1= $\frac{32.89}{(2001.494/2)} = 0.032$ and average price optimization for soft drink (Low, Fat, and Regular) for Outlet 1 are 2.47 and 2.85, respectively. Table IX depicts the complete price optimization for all the product categories across all outlets of ABC stores.Thus, mathematically the price optimization ranges from

$$(C * min_j\,\{u_j\} \text{ to } (C * max_j\,\{u_j\}) \qquad (11)$$

In order to maximize profit, the approach adopted for BREAD model is to fix the range as mentioned in (11). These status range are named as Inactive, Cold, Warm, and Active respectively. The assessment of the status is done in two parts as specified in Part 1 and Part 2.

Part 1: Assess the overall sales and outlet sales

This part gives analysis of the overall sales with respect to store sales, which customs the base of the observation criteria.

Part 2: Assessment of sales with respect to product types [Observation Criteria]

BREAD model in this phase assesses the status of different product type.

Stage 1: Identify the status range of sales with respect to overall sales based on criteria mentioned in Table III.

TABLE III.
STATUS RANGE OF OVERALL SALES

Status	Range 1 ($C * min_j\,\{u_j\}$)	Range 2 ($C * max_j\,\{u_j\}$)
Inactive	5000	10000
Cold	10001	25000
Warm	25001	50000
Active	50001	

Source:Author's calculation

Stage 2: Identify the status range of sales with respect to store sales based on table III. It is also used to categorize the status range of sales with respect to stores are given in table IV.

TABLE IV.
STATUS RANGE OF STORE SALES

STATUS	Range 1 ($C * min_j\,\{u_j\}$)	Range 2 ($C * max_j\,\{u_j\}$)
Inactive	0	5000
Cold	5001	10000
Warm	10001	15000
Active	15001	

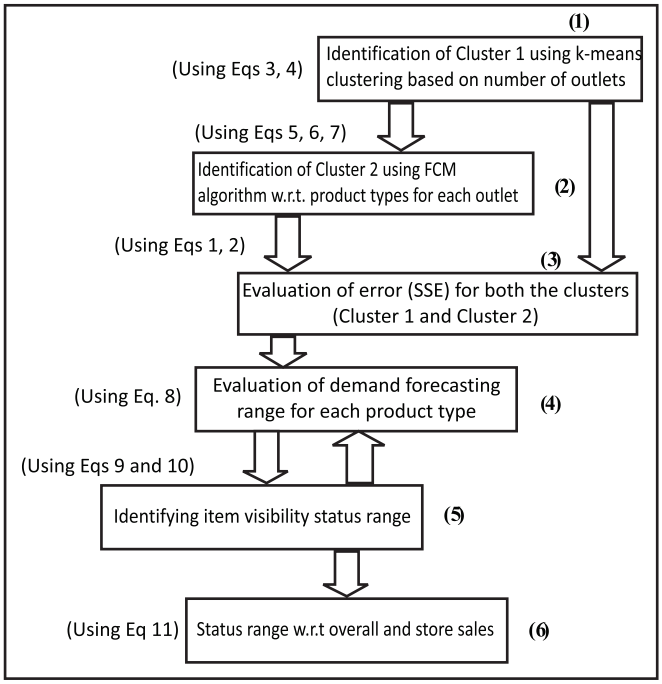
Source:Author's calculation

Step 5: Integration Phase

The last step takes individual cluster c from n clusters and executes product visibility (using demand forecasting) and status range of all product types as depicted in fig. 3.

Also phases (4) and (5) are implemented continuously for all clusters c $(1\ldots.c)$ to find demand forecasting (thereby status range of product type both-store wise and overall) and also product visibility. This is shown in (12) and (13), respectively.

Fig. 3. BREAD model Integration phase



Source:Author's calculation

III. BREAD STRATEGIES

BREAD strategies for commendations are alienated into 12 rules grounded on results of (8), (9), (10), (11), (12) and (13). Note that in Table V,↑ sign denotes that the sales is higher and ↓ denotes lower sales. As an example, Active ↑ and Active ↑ in columns 1 and 2 indicate that status range are Active at both individual and overall level. Similarly, Active ↑ Warm ↓ and Active ↑ in column 1 and 2 specifies that the status range at individual stores are a combination of Active and Warm though the overall sales is Active. If the sales are not as expected, we need to endorse actions, and, these actions are termed BREAD strategies. BREAD strategies comprise product status and product visibility strategies.

IV. IMPLEMENTATION OF BREAD MODEL

BREAD model was implemented for ABC stores, which has 10 different outlets with 16 product types (a total of 8523 records). The outcomes are mentioned in tables VI to IX.

TABLE V.
PRODUCT STATUS STRATEGIES

Rules	Status Range of Store Sales	Status Range of Overall Sales	Product Visibility Strategy	Price Optimization Strategy
R1#	Active ↑	Active ↑	No Change	No evaluation required
R2#	Active ↑ Warm ↓	Active ↑	Increment PV for stores with Warm sales by 1	Evaluate maximum range
R3#	Active ↑ Cold ↓	Active ↑	Increment PV for stores with Cold sales by 2	Evaluate maximum range
R4#	Warm ↑	Warm ↑	No Change	Evaluate minimum range
R5#	Active ↑ Warm ↓	Warm ↓	Increment PV for stores with Warm sales by 1 level	Evaluate minimum range
R6#	Warm ↓ Cold ↓	Warm ↑	Increment PV for stores with Cold sales by 2 levels	Evaluate minimum range
R7#	Cold ↓	Cold ↑	Increment PV for stores with Cold sales by 2 levels	Evaluate minimum and maximum range
R8#	Active ↑ Cold ↓	Cold ↓	Increment PV for stores with Cold sales by 2 levels	Evaluate minimum and maximum range
R9#	Inactive ↓ Cold ↑	Cold ↑	Increment PV for stores with Inactive sales by 2 levels	Evaluate minimum and maximum range
R10#	Inactive ↓	Inactive ↓	Discontinue the product	Not Required
R11#	Inactive ↓ Warm ↑	Inactive ↓	Discontinue the product	Not Required
R12#	Inactive ↓ Cold ↑	Inactive ↓	Discontinue the product	Not Required

Source:Author's calculation

$$\Sigma_{c\in n}\Sigma^{c}_{i=1}df_i \qquad (12)$$

$$\Sigma_{c\in n}\Sigma^{c}_{i=1}PV_i \qquad (13)$$

Table VI	Total sales at individual stores and overall sales for each product type
Table VII	Range status for individual stores and overall for each product type
Table VIII	Product visibility status and score (with respect to demand forecasting) for each product type
Table IX	Price optimization (as an average) for each product type

TABLE VI.
SALES OF PRODUCT TYPES

Product Type	Outlet 1	Outlet 2	Outlet 3	Outlet 4	Outlet 5	Outlet 6	Outlet 7	Outlet 8	Outlet 9	Outlet 10	Overall Sales
Baking Goods	8263.9	11397.7	12180.3	13095	6316.3	12497.7	12962.8	11020.8	11338.6	13976.4	113049.5
Breads	3696.5	5956.9	6258.7	5601.6	3636.4	3651.5	4258	4023.5	4414	5380.6	46877.7
Breakfast	824.9	1621.5	1763.7	2340.6	1370.9	2196.8	2566.3	2302.8	3057	2092.2	20136.7
Canned	8840.4	15305.4	13647. 6	11008.5	7923.5	12947.6	10063.1	11273	13129.7	14977.5	119116.3
Dairy	8132.5	11763.8	14759.5	15256.1	7444.2	16807.9	14665.5	16492	13669.8	12964.3	131955.6
Frozen Foods	10635	17241.7	13971.9	15593.6	12827.1	17925.8	15915.1	19037.3	12623.5	16315.8	152086.8
Fruits and Vegetables	16489.6	22346.1	25341.3	25470.3	15316.5	23035.7	26574.9	22093.2	24154.6	23898.3	224720.5
Hard Drinks	2892.6	4446	4590.2	3457.7	3817.9	4349.1	5032.3	3619.6	5174.4	3352.1	40731.9
Health and Hygiene	4844.4	8568.6	8653.5	9945.3	5537.5	9979.6	11733.3	10966.7	10555.7	11576.8	92361.4
Household	10720.2	19932.5	22099.7	21374.1	10940.4	21718.2	18808.8	21446.6	19383.2	22111.1	188534.8
Meat	4930.6	11278.7	10705.4	9143.1	4484.8	6849	9633.1	11102.1	9674.6	9807	87608.4
Others	2275.6	3842	3449	2975.9	903.8	4893.4	3625.2	3034	1862.2	3132	29993.1
Seafood	-	1470.2	1087.4	649.6	-	1085.8	782	397.5	709.4	811.4	6993.3
Snack Foods	15568.6	27831.8	27620.6	24263.4	12420.6	24784.5	23293	24810.6	28770.8	22387.1	231751
Soft Drinks	4266.5	8750.03	7018.8	9036	4808.1	10539.3	7973.6	7985.8	9227.3	9972.1	79577.53
Starchy Foods	2501	2973.4	3766.4	3807.6	2360.7	4984.5	3337	4602.7	4576	3896.8	36806.1

Source:Author's calculation

TABLE VII.
PRODUCT STATUS-INDIVIDUAL STORES AND OVERALL

Product Type	Outlet 1	Outlet 2	Outlet 3	Outlet 4	Outlet 5	Outlet 6	Outlet 7	Outlet 8	Outlet 9	Outlet 10	Overall Status
Baking Goods	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Breads	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold
Breakfast	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Cold
Canned	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Dairy	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Frozen Foods	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Fruits and Vegetables	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Hard Drinks	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold
Health and Hygiene	Active	Active	Active	Active	Warm	Active	Active	Active	Warm	Warm	Warm
Household	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Meat	Active	Active	Active	Active	Warm	Active	Active	Active	Warm	Warm	Warm
Others	Cold	Cold	Cold	Cold	Inactive	Cold	Cold	Cold	Inactive	Inactive	Cold
Seafood	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Snack Foods	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active	Active
Soft Drinks	Active	Active	Warm	Active	Warm	Active	Warm	Warm	Warm	Warm	Warm
Starchy Foods	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold	Cold

Source:Author's calculation

TABLE VIII.

PRODUCT VISIBILITY SCORE AND STATUS (WITH RESPECT TO DEMAND FORECASTING)

Product Type	Product Visibility Score	Product Visibility Status
Baking Goods	57.83	Moderately Visible
Breads	24.61	Slightly Visible
Breakfast	11.82	Invisible
Canned	58.64	Moderately Visible
Dairy	63.73	Visible
Frozen Foods	79.63	Visible
Fruits and Vegetables	107.25	Demanding
Hard Drinks	19.78	Invisible
Health and Hygiene	39.24	Slightly Visible
Household	73.28	Visible
Meat	34.64	Slightly Visible
Others	11.57	Invisible
Seafood	3.05	Invisible
Snack Foods	107.18	Demanding
Soft Drinks	37.86	Slightly Visible
Starchy Foods	16.20	Invisible

Source:Author's calculation

TABLE IX.

PRICE OPTIMIZATION OF PRODUCT TYPES (AT INDIVIDUAL OUTLETS)

Product Type	Outlet 1	Outlet 2	Outlet 3	Outlet 4	Outlet 5	Outlet 6	Outlet 7	Outlet 8	Outlet 9	Outlet 10
Baking Goods	2	2	2	2	2	2	2	2	2	2
Breads	4.05	5.64	5.8	4.855	4.925	3.935	4.08	4.28	4.79	4.475
Breakfast	7.055	9.835	9.62	11.07	8.45	10.99	11.86	11.81	15.83	9.72
Canned	2.34	2.14	1.855	1.585	1.96	1.755	1.505	1.69	1.885	1.805
Dairy	1.585	1.535	1.615	1.615	1.585	1.79	1.58	1.835	1.68	1.375
Frozen Foods	1.445	1.5	1.245	1.295	1.695	1.46	1.285	1.605	1.29	1.245
Fruits and Vegetables	0.99	0.915	0.96	0.905	1.02	0.875	0.93	0.9	0.96	0.825
Hard Drinks	2.61	2.89	2.79	2.38	7.03	2.69	2.875	2.53	3.435	2.14
Health and Hygiene	1.065	1.145	1.105	1.17	1.365	1.16	1.295	1.325	1.285	1.185
Household	0.54	0.62	0.63	0.6	0.655	0.615	0.54	0.63	0.615	0.585
Meat	2.305	3.35	2.99	2.7	2.435	2.27	2.65	3.3	2.94	2.655
Others	4.29	4.445	3.815	3.365	2.435	5.55	3.815	3.65	3.03	3.265
Seafood	17.025	27.975	24.525	15.77	13.345	20.475	16.76	15.27	17.935	16.265
Snack Foods	1	1.035	0.975	0.86	1.015	0.895	0.835	0.955	1.1	0.78
Soft Drinks	2.66	2.92	2.43	2.735	2.95	3.205	2.49	2.865	3.045	2.835
Starchy Foods	8.535	6.28	6.805	6.66	8.005	8.875	6.025	8.455	8.595	6.405

Source:Author's calculation

V. OUTCOMES AND RECOMMENDATIONS

This section gives recommendations based on the executed model. The set of outcomes for 16 different product types are mentioned in Table X.

Product wise recommendations

For product types *Baking Soda, Canned, Diary, Frozen Foods, Fruits and Vegetables, Household, Snack Foods*, the demand is *Active* in both stores and overall. Thus, no action is required.

For product type *Breads*, the demand is *Cold* at both stores and overall. From Table V, Rule VII (R7#) is recommended.

Product visibility Strategy: Increment *PV* for stores with *Cold* sales by 2 levels, indicating

(Slightly visible) → *(Visible)*

Price optimization Strategy: Evaluate minimum and maximum range, specifying that price optimization range should be between {3.935,5.8}.

For product type *Breakfast,* the demand is *Inactive* at stores level. However, demand is *Cold o*verall. From Table V, Rule 9 (R9#) is recommended.

Product Visibility Strategy

Increment *PV* for stores with *Inactive* by 2 levels, indicating

(Invisible) → *(Moderately Visible)*

TABLE X.
BREAD OUTCOMES

Product Type	Status	Rules
Baking Soda	Active ↑ → Active ↑	R1#
Breads	Cold ↑→ Cold ↑	R7#
Breakfast	Inactive ↓ → Cold ↑	R9#
Canned	Active ↑ → Active ↑	R1#
Dairy	Active ↑ → Active ↑	R1#
Frozen Foods	Active ↑ → Active ↑	R1#
Fruits and Vegetables	Active ↑ → Active ↑	R1#
Hard Drinks	Cold ↑ → Cold ↑	R7#
Health and Hygiene	Active ↑ Warm ↓ → Warm ↑	R6#
Household	Active ↑ → Active ↑	R1#
Meat	Active ↑ Warm ↓ → Warm ↓	R5#
Others	Inactive ↓ Cold ↑ → Cold ↑	R9#
Seafood	Inactive ↓ → Inactive ↓	R10#
Snack Foods	Active ↑ → Active ↑	R1#
Soft Drinks	Active ↑ Warm ↑ → Warm ↑	R5#
Starchy Foods	Cold ↑ → Cold ↑	R7#

Source:Author's calculation

Price Optimization Strategy

Evaluate minimum and maximum range, specifying that price optimization range should be between {7.055,15.83}.

For product type *Hard Drinks,* the demand is *Cold* at both stores and overall. From Table V, Rule 7 (R7#) is recommended.

Product Visibility Strategy

Increment *PV* for stores with *Cold* sales by 2 levels, indicating

(*Invisible*) → (*Moderately Visible*)

Price Optimization Strategy

Evaluate minimum and maximum range, specifying that price optimization range should be between 2.14 and 7.03.

For product type *Health and Hygiene,* the demand is either *Active or Warm* at store level and *Warm* overall. From Table V, Rule 6 (R6#) is recommended.

Product Visibility Strategy

Increment *PV* for stores with *Warm* sales by 2 levels, indicating

(Slightly visible) → *(Visible)*

Price Optimization Strategy

Evaluate minimum range, specifying that price optimization should be {1.065}for the stores with *Warm* sales. This strategy may help ABC stores to increase the sale.

For product type *Meat*, the demand is either Active or Warm at store level and Warm overall. From Table V, Rule 6 (R6#) is recommended.

Product Visibility Strategy

Increment *PV* for stores with *Warm* sales by 2 levels, indicating

(Slightly visible) → *(Visible)*

Price Optimization Strategy

Evaluate minimum range, specifying that price optimization should be {2.27} for the stores with *Warm* sales, to increase the sales in Warm stores.

For product type *Others,* the demand is either *Inactive or Cold* at store level and *Cold o*verall. From Table V, Rule 9 (R9#) is recommended.

Product Visibility Strategy

Increment *PV* for stores with *Warm* sales by 2 levels, indicating

($Invisible$) → ($Moderately\ Visible$)

Price Optimization Strategy

Evaluate minimum and maximum range, specifying that price optimization range should be between {2.435,5.55}.

For product type *Seafood*, the demand is either *Inactive* at both store level and overall. From Table V, Rule 10 (R10#) is recommended.

The recommendation is to discontinue the product.

For product type *Soft Drinks,* the demand is either *Active* or *Warm* at store level and *Warm* overall. From Table V, Rule 6 (R6#) is recommended.

Product Visibility Strategy

Increment *PV* for stores with *Warm sales* by 2 levels, indicating

(*Slightly visible*) → (*Visible*)

Price Optimization Strategy

Evaluate minimum range, specifying that price optimization should be {0.78} for the stores with Warm sales. This strategy may help ABC stores to increase sale.

For Product type *Starchy foods,* the demand is *Cold* at both stores and overall. From Table V, Rule 7 (R7#) is recommended.

Product Visibility Strategy

Increment PV for stores with *Cold* sales by 2 levels, indicating

(Invisible) → *(Moderately Visible)*

Price Optimization Strategy

Evaluate minimum and maximum range, specifying that price optimization range should be between {6.025,8.875}.

VI. CONCLUSION

The model integrates two characteristics of retail analytics, demand forecasting (using product visibility) and price optimization and makes recommendations accordingly. The other two facets of retail analytics, namely, attract new customers and predicting customer trends are not assimilated in the study, and will be incorporated in the impending model expansion.

REFERENCES

[1] Statista, "Forecast for global retail sales growth from 2008 to 2018," 2016 [Online] Available: https://www.statista.com/Statist/232347/forecast-of-global-retail-sales-growth/

[2] S. Lavalle, M. S. Hopkins, E. Lesser, R. Shockley, and N. Kruschwitz, "Analytics: The new path to value, 2010," *MIT Sloan Manage. Rev.*

[3] E. H. Brynjolfsson, "Strength in numbers: How does data-driven decision making affect firm performance?," 2011. [Online] Available: ebusiness.mit.edu.

[4] C. Luckie, "Big Data," *Facts and Statist That Will Shock You*, 2012. Fathom.

[5] DeZyre, "5 Big Data and Hadoop Use Cases in Retail Analytics," 2015. [Online] Available: https://www.dezyre.com/article/5-big-data-and-hadoop-use-cases-in-retail-analytics/91

[6] IBM, "Analytics: The real-world use of big data in retail," 2012. Oxford: IBM Inst. for Bus. Value.

[7] EKNRes., The Future of Retail Analytics, 2013. New Jersey: SAS.

[8] Euromonitor, "Retailing in United Arab Emirates," 2016. [Online] Available: http://www.euromonitor.com/retailing-in-the-united-arab-emirates/report

[9] C. Wen, "The four types of retail analytics every bus. needs," 2015. [Online] Available: https://www.livetiles.nyc/four-types-retail-analytics-every-business-needs/

[10] M. Levya, D. Grewal, P. K. Kopalle, and J. D. Hess, "Emerging trends in retail pricing practice: Implications for Res.," *J. of Retailing*, vol. 80, no. 3, pp. 13-21, 2004. doi: https://doi.org/10.1016/j.jretai.2004.08.003

[11] P. S. Fader, B. G. S. Hardie, and K. L. Lee, "RFM and CLV: Using iso-value curves for customer base Anal.," *J. of Marketing Res.*, pp. 415-430, 2005

[12] R. Fildes, and P. Goodwin, "Producing more efficient demand forecasts, 2006. [Online] Available: https://www.Res.gate.net/profile/Robert_Fildes/publication/228881292_Producing_more_efficient_demand_forecasts/links/0fcfd512f47d691674000000.pdf?origin=publication_list

[13] I. Kolyshkina, and S. Simoff, "Customer analytics projects: Addressing existing problems with a process that leads to success," in *Proc. of the 6th Australasian Conf. on Data Mining and Analytics*, Vol. 70, *Australian Comput. Soc.,* pp. 13-19, 2007. ACM Digital Library.

[14] M. Adams, "Predictive Analytics for the Retail Industry (MS SQL Server Tech. Article)," *Microsoft*, 2008.

[15] W. P. Tyco, "Don't let inventory distortion steal your

profits, 2012. [Online] Available: http://www.tycoretailsolutions.com/Pages/solutionarea-II.aspx
[16] J. B. MacQueen, "Some methods for classification and anal. of multivariate observations," *Proc. of 5th Berkeley Symp. on Math. Statist and Probability*, pp. 281-297, 1967. Berkeley: University of California Press.
[17] J. C. Bezdek, R. Ehrlich, and W. Full, "The fuzzy c-means clustering algorithm," *Comput. & GeoSciences*, pp. 191-203, 1984.

About the Author

Riktesh Srivastava is Associate Professor, Information Systems with Skyline University College, Sharjah. He holds Ph.D. in Information Systems and Executive Qualification in Management from the prestigious Indian Institute of Management, Ahmedabad (IIMA). He also holds degrees in Electronics Engineering, M.Tech. (IT) and M.B.A.(Marketing).
Dr Srivastava had also accomplished certification programs in:
• Marketing from Wharton School, University of Pennsylvania
• Customer Analytics from Wharton School, University of Pennsylvania
• Electronic Commerce from NTU, Singapore
He has a total experience of 19 years, and has written 3 books (OS, C++ Programming and RDBMS), and published more than 75 papers in International Journals and Conferences. His research areas include Queuing Theory, Electronic Commerce, and Business Analytics.