

Implying Association Rule Mining and Market Basket Analysis for Knowing Consumer Behavior and Buying Pattern in Lockdown - A Data Mining Approach

Anurag Sinha

Department of information technology, Amity University Jharkhand

ABSTRACT

Buyer practices have changed as individuals are figuring out how to live with the new truth of COVID-19. Take-out and conveyance orders have expanded, and our customer has added new items to their menu because of new client inclinations. With every one of the continuous changes, the customer had numerous unanswered inquiries, for example, Smartbridge has broad involvement with café innovation development Café TECHNOLOGY CAPABILITIES :Are the most famous items as yet unchanged after COVID? :Which are the most sold item blends now? :What is the acknowledgment of new things? :What are clients purchasing alongside new things? :How have liquor deals changed? The customer previously had reports that followed item deals and operational measurements, notwithstanding, there was a need to get a more profound knowledge into item examination. The customer expected to recognize what items and introductions were being sold all the more frequently, measure the acknowledgment of new items, and figure out what items clients buy together to improve advertising efforts, advancements, and deals. The E-business industry is filling immensely in the Indian market. The modest 4G web bundles in India clearly gives a push to these ventures. Thus, as Covid19 first hit in Quite a while, individuals got terrified to go out from their homes in light of the fact that, in their mind, it's a dread of Covid. They even wonder whether or not to go out to purchase fundamental (FMCG) products. Frenzy purchasing additionally has seen and to stay away from this dread of COVID-19, individuals are offering inclinations to the E-Commerce destinations to purchase fundamental products and a few clients are new which joined to purchase fundamental merchandise during this Pandemic Lockdown period. Numerous clients are moving their purchasing conduct from disconnected retail locations to online stores. This paper examines the customer buying pattern during lockdown.

Keywords - Market basket analysis, association rule mining, buying pattern, data mining.

I. Introduction

Shopper conduct is a customer movement in choosing to buy, use, just as burn-through the bought labor and products remembering for the client factors which can give an ascent to their choices regardless of whether to buy and utilize items. Each client have various necessities and tendencies too as have various practices in satisfying those things. Notwithstanding, in case of various practices to satisfy their requirements, they actually share a few similitudes, one of them is craving to expand their fulfillments in devouring a vital item or administration. Of that utilization action, that can be surmised concerning the conduct, example, or propensity that the clients do in satisfying their necessities and wants. That conduct can be distinguished via the logging did by the go-between supplier of purchaser needs (grocery store).

The logging is brought off considering the imperative of documentation and history information recognizable proof over the conveying out of exchange action. In these new years, exchange information have been regularly utilized as examination and investigation objects for scientists. In this investigation, additionally, exchange information are to be re-prepared/re-investigated to produce more important data. For example, data of a thing whose deals is the most noteworthy. Additionally, data can be used in respect with the stock expansion of that thing. Besides, from exchange information there can be used concerning the connection of each bought thing inside the client's bushel. By that data, we can utilize it for viable item show/arrangement to draw in clients' interest. The ordinarily utilized application to break down exchange information clients' shopping bushel is market bushel examination. Market bushel examination is one of the modes from information mining method predominantly utilized to break down things/merchandise in at least one shopping bushels that a client has in one specific second Exchange information is a bunch of recording information bring about associations with deals buy exercises at a specific organization. In these new years, exchange information has been predominantly utilized as exploration objects in methods for finding new data. One of the potential endeavors is to plan an application that can be utilized to dissect the current exchange information. That application has the nature of market bin investigation. In expansion, the application is intended to be work area based whose segments are capable to measure just as re-log the current exchange information. The pre-owned strategy in planning this application is via following the current strides on information mining method. The preliminary outcome showed that the turn of events and the execution of market crate investigation application through affiliation rule technique utilizing apriori calculation could function admirably. With the methods for certainty worth of 46.69% and backing worth of 1.78%, also, the measure of the produced rule was 30 standards.

The additional time and cash devouring cycle in retail industry is extracting more valuable data from retail location data set, to acquire cutthroat advantage. India is a market zone which contains gigantic number Kirana and frivolous shops and huge number of retail locations what's more, grocery stores. In this examination an endeavor has been made through strategically pitching investigation to distinguish the relationship among items and to recommend essential item advancement methodologies. Likewise utilizing this examination, the client inclinations and purchasing behavior have been concentrated to propose item situating to improve the deals. The result of the investigation can be utilized for valuing.

A. Market Basket Analysis

Market Basket Analysis is an investigation mode performed on client conduct while shopping at a grocery store through the methods distinguishing affiliation and associations among different things put by the clients in their shopping bins [2]. In explicit, Market Basket Analysis focuses on all the while distinguishing the most oftentimes bought things by clients. Here, thing is portrayed as a few sorts of items in grocery store. Utilizing market bin examination mode, an information on what are the things oftenly bought by the clients all the while and having a chance to be advanced can be gotten. Concerning the target of market container examination mode to choose which items that clients buy at the same time, whereby the name of this mode is taken from the conduct of the clients in setting shopping items into their shopping containers or shopping list.

Over recognizing shopping bin example of a client can altogether help an organization in utilizing that data in regard of business system needs, one of them is setting the most oftentimes bought items all the while into one explicit territory.

B. Association Rule

Association rule is identified with the assertion of "what' with what". This matter can be in a type of explanation on exchange movement completed by the clients at a store. From that assertion, there has a solid connection to the investigation of client exchange information data set to decide the propensity for a bought item with what item, hence, affiliation rule is as often as possible alluded as market bin investigation [1], [2]. The meaning of an acquainted principle can be figured within the sight of two boundaries, to be specific backing and certainty. Backing (supporting worth) is the level of blends of item things in the data set. While certainty (assurance esteem) is a worth to decide the strength of between thing connections in affiliation rules.

$$S = \frac{\sum(Ta+Tc)}{\sum(T)} \quad (1)$$

$$\text{Support (A)} = \frac{\text{The Number of Transaction that Contains A}}{\text{Total Transaction}} \quad (2)$$

where:

S = Support

$\sum(Ta + Tc)$ = the number of transaction that contains *antecedent* and *consequent*.

$\sum(T)$ = the number of transaction.

$$C = \frac{\sum(Ta+Tc)}{\sum(Ta)} \quad (3)$$

$$\text{Confidence} = P(A | B) = \frac{\text{The Number of Transaction that Contains A and B}}{\text{Total Transaction that Contains A}} \quad (4)$$

where:

C = Confidence

$\sum(Ta + Tc)$ = the number of transaction that contains *antecedent* and *consequent*.

$\sum(Ta)$ = the number of transaction that contains *antecedent*.

C. Customer Behaviour

Definition of customer behavior is a dynamic interaction between cognition, affection, behavior, and its environment whereby someone performs exchange activities in their regular basis [3]. In view of this statement, there are three significant matters to grasp, namely:

1. Customer behavior bears dynamic characteristic, thus, hard to predict.
2. Involving interaction, like cognition, affection, behavior, and the occurrences around customers,
3. Involving exchange, like the exchange of item and money from merchant to customer Four factors that could give a rise to customer purchase in shopping.

some of them were:

1. Cultural Factor
2. Social Factor
3. Personal Factor
4. Psychological Factor

There were three variables that must be regarded in understanding customer behavior, namely stimulus variable, response variable, and intervening variable

Data Mining

The researcher argued that Data Mining had garnered the interest of information industry and public during in that recent years, this was brought about by massive data availability and the necessities to render information data to usable knowledge [6], [7]. Data Mining is prevalently depicted as retrieval pattern on a large set of crude data so that some unveiled knowledge in the data can be found. Data Mining is considered a major step in the process of knowledge retrieval in the database. Data Mining inherits many aspects and techniques from established fields of science first. Data Mining also has long roots from fields such as artificial intelligence (Artificial Intelligent), machine learning, statistics, database and also information retrieval.

Apriori Algorithm

Apriori is a powerful calculation utilized for mining the itemsets and use of affiliation rules on the conditional data sets. The affiliation rules controlled by Apriori feature the general pattern in retail location data set. The Key issue in the retail area is to discover helpful covered up designs for business application. Apriori are designed to work on data set containing exchanges (the assortment of things by faithful customers). When incessant thing sets are shown up at the affiliation rules are shaped with certainty bigger than or equivalent to a client characterized least certainty. Items are the elements that we are distinguishing connection between. A gathering of items is called a thing set.

$$P=\{p1, p2...pn\}$$

Exchanges are cases of gatherings of items coinciding together. For a retail location, an exchange is, essentially, a, exchange. For an online book shop, an exchange may be the group of articles read in a solitary visit to the site. For every exchange, at that point, we have a thing set.

$$tn=\{pi,pj,...,pk\}$$

Rules are articulations of the structure

$$\{p_1, p_2, \dots\} \Rightarrow \{p_k\}$$

for example in the event that you have the items in the dataset (on the left hand side (LHS) of the standard for example $\{p_1, p_2, \dots\}$), at that point almost certainly, a client will be sought after in the items on the privilege hand side (RHS for example $\{p_k\}$). In our model over, our standard would be:

$$\{\text{Coffee powder, milk}\} \Rightarrow \{\text{sugar}\}$$

The yield of a strategically pitching examination is essentially a bunch of rules, that assist us with making business choices. (For example Showcasing, Product Assortment).

Support : Percentage or absolute number of exchanges that contain the entirety of the things in a thing set (e.g., espresso powder, milk and sugar). The rules with high help, will be appropriate to an enormous number of exchanges. E.g., Retail general store is probably going to include fundamental items, for example, espresso powder, milk & sugar that are mainstream across a whole clients. If there should be an occurrence of Electronics things, like printer toner, might not have items with a high help, on the grounds that every client as it were purchases toner for their own utilization and it requires some investment to supplant.

Confidence: Probability that an exchange that contains the things on the LHS of the standard (espresso powder, milk) likewise contains the thing on the RHS (sugar). Higher certainty esteem infers more noteworthy that the thing on the correct hand side will be bought.

Lift: Probability of the entirety of the things in a standard buying together (otherwise called the help) separated by the result of the probabilities of the things on the LHS and RHS happening as though there was no connection between them. For instance, if espresso powder, milk and sugar happened together in 3.5% of all exchanges, espresso powder and milk in 20% of exchanges and sugar in 6% of exchanges, at that point the lift would be: $0.035 / (0.2 * 0.06) = 2.917$. A lift of multiple proposes that the presence of espresso powder and milk expands the likelihood that a sugar will additionally happen in the exchange. All in all, lift tells about the strength of relationship between the items on the LHS of the standard; bigger lift esteem infers the more noteworthy connection between two items.

2. Background

5 Critical Ways the Pandemic is Changing Consumer Preferences

Change is hard, ordinarily. A great many people won't change since a long time ago held mentalities, propensities, or inclinations — except if they must choose between limited options. Also, that is the thing that occurred in the previous few months. The Coronavirus pandemic prompted far reaching social separating and sanctuary set up standards, upsetting stocks, deals, and supply chains —

and subsequently, shopper buy designs. Brands and retailers mixed to enhance stock and reevaluate conveyance. In the interim, shoppers everywhere on the world had auto-left on perhaps the main mass changes in current history. For the most recent decade at Course5 Intelligence, we have been observing and empowering worldwide customers, and intently following shopper experiences utilizing our Audience Graph arrangement [A worldwide large scale and miniature fragments checking stage at scale]. We found that customers are developing as time passes of the pandemic, moving from a demeanor of abdication in the underlying weeks to a resurging office lately. Then, something changed unalterably. The shopper experienced and acknowledged essential necessities, new items, an adjusted way of life, and new mentalities to life. Brands need to comprehend these progressions to have the option to oblige and interface with the late – and post-pandemic buyer. Here are a portion of the basic ways the pandemic has influenced shopper inclinations up to this point –

1. Affinity vs. Availability –

Across groceries, home cleaning products, and other essentials of daily life, consumers found that the brands they loved, preferred, or were accustomed to, were often out of stock and inevitably started using brands that were on the shelf. In many cases, the new brands have stuck.

Over ~35% of consumers said they would continue to buy new products they started using during the crisis. Consumers who shopped online for the first time or shopped more extensively online during the pandemic mostly said this would continue to do so for at least some of the purchases that were previously offline. One brand's (or shopping outlet/platform's) loss is another brand's gain. Both brands that lost and gained customers during the pandemic will need to understand and address consumers very carefully in the coming months to either hold on to newly acquired customers or woo back old ones.

2. Localism –

With restricted global movement and a sharp urge for self-preservation in the rapids of the pandemic, consumers globally started falling back on their local communities and championing local produce. The trend could mean consumers opting for products made locally and buying from community stores. The takeaway for brands here is to highlight their local connection and community engagement in meaningful and resonant ways. In the coming months, brands will need to follow consumer sentiment carefully as they make production and marketing decisions. Brands would need to look at their entire supply chains and marketing value chains to reach out to more local and micro-communities rather than focussing on national or a broader ecosystems. The better brands than can address exact consumer needs, the better their chances of winning consumers' hearts. As Zig Zagler said, "People don't buy for logical reasons. They buy for emotional reasons." More or less.

3. Safety Over Comfort –

Safety has been one of the biggest concerns on consumers' minds in the past few months, and continues to remain so. Critical trends in the safety needs based on our 'Audience Graph' –

(a) shopping at local stores,

- (b) choosing to go to a superstore optimized for social distancing to complete all purchases at one go,
- (c) making purchases online
- (d) opting for contactless delivery – are some of the critical micro trends that we predict will continue in the foreseeable future.

Consumer Preferences in Pandemic

As a brand or shopping outlet, this means you need to optimize your business model for safety and communicate the measures you have taken to your target customers. In our recent Compass Virtual (Digital) Conference, Alex Robertson shared: when David's Bridal reopens:

- (a) they will be sanitizing outlets regularly
- (b) meeting clients by appointment only, and
- (c) organizing home trials in addition to shifting parts of the purchase cycle online.

In short, doing everything they can to assuage customers' fears and ensure a safe shopping experience.

4. Prioritizing Health –

Factors like:

- (a) being physically confined (restricted movement and socializing),
- (b) spending more time online (being sedentary),
- (c) worries about possible infection (need to build immunity) and
- (d) growing anxiety about economic depression (need to stay positive) has made health a top priority for consumers.

From taking up some form of indoor or socially distanced fitness routine to home-cooking healthy meals to overall slowing down and reconsidering life priorities, consumers have distinctly become more health-conscious. More than ever, they carefully examine their product and lifestyle choices in the context of their health goals.

Socially distanced fitness routine It is increasingly crucial for brands to support healthy lifestyles for their customers and employees, stay carefully tuned to the apprehensions of consumers, and find ways to assure safety. Having a 'health strategy' will be a strategic differentiator for brands in the foreseeable future.

5. Patience and Value-Consciousness –

Consumers amid a lockdown and economic uncertainty are turning both more patient and more demanding at the same time. They are patient with delivery delays and logistical inconveniences, but they are more demanding in terms of quality and the value they are getting for their money. A growing number of consumers say “price matters” and may matter even more in the coming months if the economy worsens. Many consumers expect brands to help them by offering discounts and promotions. Brands will need to explore new ways to enable this. Sarda Farms, a local Milk supplier in India, changed its packaging to cut costs, which, in turn, lowered the price of milk. Similarly, technology companies offering their solutions free or at highly discounted rates to students, universities, and business organizations right now which would help these users adopt their solutions and price would not be a hindrance later.

Consumers say price matters-One of the trends accelerated by the current pandemic is also a rise in mindful consumption. Consumers are more attentive to what they’re buying, and strive to limit food waste, shop more cost consciously and buy more sustainable options. Brands will need to weave one or more of these considerations prominently into their offerings. These are some of the trends we can see and predict for now. But just as the pandemic situation remains in flux, consumer preferences will continue to change and evolve in the next 12-18 months. These are crucial habit-forming months for consumers, and brands will do well to continuously monitor consumer sensibilities and adapt their offerings to provide maximum comfort and utility in challenging times.

UNDERSTANDING CONSUMER BEHAVIOR: MARKET BASKET ANALYSIS

In the examination space, we call these retail shopping events market bins. They comprise of blends of items or potentially benefits that a customer may buy together. In the model over, the guarantee, the holster and the charger go with the telephone into the bin. In an eatery, some espresso may supplement the cake. Furthermore, in an auto assistance station, a tire pivot might be important for a similar container as an oil change. Market bushel examination considers mixes of purchased together things to give the correct items to the perfect client at the perfect time. The key is recognizing the buy designs, deciding those patterns that are of worth and introducing those items in an effectively available style. These affinities can be utilized to improve efficiencies and primary concern productivity. Extra proposals, strategically pitching, couponing and advancements give further freedoms to expand income.

An early illustration of this is the legend of lager and diapers, broadly revealed about twenty years prior. In 1998, Forbes magazine summed up as follows: "A corporate store put all its checkout-counter information into a goliath advanced distribution center and set the circle drives turning. Out popped a most unforeseen relationship: deals of diapers and brew. Clearly, youthful dads would make a late-night race to the store to get Pampers and get some Bud Light while they were there. Benefiting from the disclosure, the store put the divergent things together. Deals zoomed." In this way, we witness one more use of market bin examination: arrangement of items inside a retail, or so far as that is concerned, advanced stag

Also, the hidden proposition in the investigation is that in the event that we can distinguish bunches of items that are purchased together, we can use that data by offering different things that are related with what was simply bought. This examination is frequently utilized by internet business locales to offer buy ideas to purchasers. For instance, when an individual purchases toothpaste, the retailer may propose different items like a toothbrush, mouthwash, flossing supplies, and so forth. This is because of the predominance with which different customers purchased these things in a similar exchange as the toothpaste.

Sounds consistent, isn't that right? However it's not generally direct. For huge firms, the sheer number of blends can be overpowering — effectively ascending into the billions, making a logical bad dream. Notwithstanding, there are instruments and approaches that make market bushel more reasonable and helpful. A few specialists use affiliation examination to look through the different item exchange mixes. Integral to this methodology are two basic inquiries that should be brought up in analyzing market bushels. These are:

Which level of the time are items X and Y found in a similar market crate?

Given that a shopper buys item X, what is the likelihood that he will likewise buy item Y?

How about we attempt to explain these issues, and afterward give the classification that is ordinarily utilized to allude to every one of these ideas. Take a gander at the graph underneath, which addresses eight exchanges arising out of a book shop's information stockroom. They could be a physical retailer or an online business one. It records the containers of eight clients, with a limit of five things (books) each. Obviously, a similar methodology can be utilized to address every one of the mixes that are of interest. The primary passage, the recurrence of the event, is assigned as the "support". The last passage portrays what specialists allude to as the "certainty".

Directors utilize an assortment of measurements to check the strength of any found connections. For instance, expect that, in the wake of looking at our book retailer's exchanges, we notice that 6% of the multitude of bushels contain Technology as a bought classification. Advertisers are generally keen on introducing this class of book to their crowd. How would it be a good idea for them to respond? Whom will they target? Upon additional information examination, they see that 18% of the individuals who have purchased History likewise bought Technology. Amazing! That is a 300% improvement (18%/6%). This 300% estimation is alluded to as "lift."

The investigation likewise shows that among those that had Fiction in their bushels, just 4% bought Technology. This addresses a 66% lift (4%/6%). Plainly, the retailer would consider utilizing History as a trigger to rouse Technology buys, instead of Fiction. The following is a visual rundown for every one of these people.

Transactions at a bookseller

Customer ID	First Book	Second Book	Third Book	Forth Book	Fifth Book
103	Mystery	Action	Drama	History	Technology
277	Health	Drama	History		
432	Cookbook	Biography	Drama	Technology	
498	Religion				
619	Biography	Drama	Fiction	Mystery	
901	Mystery	Politics			
1045	Romance	Travel			
1222	Memoir	History			

Bought	Bought	LIFT
Technology	Technology and History	
6%	18%	300%

Bought	Bought	LIFT
Technology	Technology and Fiction	
6%	4%	67%

“So, what?” you might ask. You’re showing me two products. It doesn’t appear to be that complicated. Analysts are evaluating multiple-product baskets. Referring to the transaction record above, one who purchases Health and Drama categories has a 12.5% probability of adding History to his basket. Only customer 277 fulfills this rule (one basket out of eight is 12.5%). The fact is that market baskets may well differ by age, gender, occupation, etc. This doesn’t present any problem — all you need to do is incorporate it into the analysis. I have often found that this extra level of analysis provides a more insightful view of the customer journey. If we want to take this one step further, we can add yet another dimension: the time element. Let’s return to our illustration. We briefly discussed the relationships of the various products in the baskets. Recall that we concluded that, for example, Mystery was linked with Drama. A reasonable issue to examine is which one arrived in the market basket first. Was it Mystery or was it Drama? The purchase sequence can influence the tactics employed by the marketer. If many or all customers buy Drama before Mystery, that may provide the retailer with guidance about selling the subsequent product. Various cross-selling strategies can be implemented. To address the above sequence issue, we can incorporate a time factor into the data analysis. An ensuing conclusion may be the following: Customers who have bought a Mystery novel are 4 times as likely to buy an audiobook reader within 3 months. Sequence rules, however, are not always that clear. Sequence analysis provides a means to identify such rules, no matter how obscure they may be in your data. Market-basket type analysis can be extended to a host of different settings. Think of a physician confronted with a patient with multiple health issues. These health issues usually precede a more serious condition. Thus, the basket defined by the initial health problems may often be combined with the more serious event, enabling the physician to proactively treat the patient. Focusing in on which of these events occur together may assist in discerning patterns and correlations between individuals and a host of diverse actions.

So, the beer and diapers association from earlier tells us something about that consumer. A young adult purchasing the latest Xbox console, combined with a new subscription to Sports Illustrated, informs us a bit about his tendencies. A senior citizen having just paid for a health magazine, coupled with an overseas vacation, helps describe him, as well. Want to know what your preferences are? Just check what's in your basket.

3. Literature Review

Data mining has taken an important part of marketing literature for the last several decades. Market basket analysis is one of the oldest areas in the field of data mining and is the best example for mining association rules. Various algorithms for Association Rule Mining (ARM) and Clustering have been developed by researchers to help users achieve their objectives. Rakesh Agrawal and Usama Fayyad are one of the pioneers in data mining. They account for a number of developed algorithms and procedures. According to Shapiro, rule generating procedures can be divided into procedures that find quantitative rules and procedures that find qualitative rules. As proposed by Gregory Piatetsky-Shapiro, William Frawley, 1991 tested the effectiveness of the algorithm by applying it to data obtained from a large retailing company. They used attributes for mining qualitative rules for categorical data using association rules. Association rules found application in many research areas such as: market basket analysis, recommendation systems to Advertising and Strategy formulation. In marketing literature market basket analysis has been classified into two models: explanatory and exploratory. The main idea behind exploratory models is the discovering of purchase patterns from POS (point-of-sale) data. Exploratory approaches do not include information on consumer demographics or marketing mix variables summarise a vast amount of data into a fewer meaningful rules or measures. Such methods are quite useful for discovering unknown relationships between the items in the data. Moreover, these methods are computationally simple and can be used for undirected data mining. However, exploratory approaches are not appropriate for forecasting and finding the cause-roots of complex problems. They are just used to uncover distinguished cross-category interdependencies based on some frequency patterns for items or product categories purchased together. A typical application of these exploratory approaches is identifying product category relationships by simple association measures. Pairwise associations are used to compare entities in pairs and judge which entity is preferred or has greater amount of some quantitative property. Julander, 1992 compared the percentage of shoppers buying a certain product and the percentage of all total sales generated by this product. By making such comparisons, one can easily find out the leading products and what is their share of sales. Examining which the leading products are for consumers is extremely important since a large number of shoppers come into contact with these specific product types every day. As the departments with leading products generate much in-store traffic, it is crucial to use this information for placing other specific products nearby. Another significant stream of research in the field of exploratory analysis is the process of generating association rules. Data mining plays an important role in the marketing field to analyze the various areas such as Consumer Buying behavior, sales forecasting, etc. To achieve their objectives researchers have developed many algorithms for developing Association Rule, Mining and Clustering. Market Basket Analysis helps the retailers to position their product in the appropriate place so as to maximize the consumer satisfaction and for improving the revenue of the store.

In the research work by Loraine Charlet Annie M.C. and Ashok Kumar Dit was observed the cost effective sales promotional activities can be designed by identifying the specific needs of customers. It was also suggested that the consumers are to be segmented and association rules are separately generated to satisfy their specific needs in a cost effective manner using some special sales promotions.¹ In a study by Kavitha Venkatachari and Issac Davanbu Chandrasekaran it was observed that the product assortment in a grocery store are interrelated and these relations are exploited applying FP growth and Apriori algorithm.² Gajalakshmi.V and MuraliDhar M.S in their study found that every algorithm has its own advantages and limitations and has different approaches to association rule, discovery, pruning, ranking, prediction and evaluation.³ In a research work by Karthikeyan.T and Ravikumar.N it was found that the Efficiency of an Algorithm is measured based on its performance and scalability with other existing algorithm using a survey data.⁴ The study revealed by K. Antony ArokiaDurai Raj and Panchapakesan Padma is based on mining the relationship from data on performance made in One day International(ODI) by Indian cricket team.⁵

4. Methodology

AI is helping the retail business from various perspectives. From gauging the business execution to recognizing the forthcoming purchasers, there are a ton of uses of AI in the retail business. Market container investigation is one of the vital uses of AI in retail. By examining the previous purchasing conduct of clients, one can discover which are the items that are purchased together by the clients. For instance, bread and butter are sold together, infant diapers and child rub oil are sold together, and so forth That implies one can break down the relationship among items. On the off chance that the retailers the board can discover this affiliation, while putting the items in the shop, these related items can be assembled. Or on the other hand, when seeing that a client is purchasing an item, the sales rep can offer the related item to the client. This interaction of dissecting the affiliation is known as the Association Rule Learning and examining the items purchased together by the clients is known as the Market Basket Analysis. In this article, we will talk about the affiliation rule learning strategy with a functional execution of market container investigation in python. We will utilize the Apriori calculation as an affiliation rule strategy for market bin investigation.

What is Association Rule Learning?

The affiliation rule learning is a standard based AI approach that produces the connection between factors in a dataset. It has significant applications in the retail business including E-Commerce retail organizations. Utilizing this technique, the items sold in an affiliation can be investigated and can be offered to clients to purchase together. For instance, it tends to be found that assuming the clients have purchased onion and potato together, doubtlessly they have purchased tomato too. It very well may be given a standard as {onion, potato} - > tomato. These guidelines are called affiliation rules. The Association Learning techniques find these kinds of rules from the dataset containing the rundown of exchanges. The affiliation rule learning has three mainstream calculations – Apriori, Eclat, and FP-Growth. In this article, we will examine the Apriori technique for affiliation learning.

Apriori Algorithm in Market Basket Analysis

Apriori is a famous calculation utilized in market container examination. This calculation is utilized with social data sets for incessant itemset mining and affiliation rule learning. It utilizes a granular perspective where successive things are expanded each thing in turn and gatherings of up-and-comers are tried against the accessible dataset. This interaction proceeds until no further augmentations are found. It utilizes the idea of Support, Confidence and Lift where,

$$\text{Confidence (Item } I_1 \rightarrow \text{Item } I_2) = \frac{\text{No. of transactions with } I_1 \text{ and } I_2}{\text{No. of transactions with } I_1}$$

$$\text{Confidence (Item } I_1 \rightarrow \text{Item } I_2) = \frac{\text{No. of transactions with } I_1 \text{ and } I_2}{\text{No. of transactions with } I_1}$$

$$\text{Lift (Item } I_1 \rightarrow \text{Item } I_2) = \frac{\text{Confidence (Item } I_1 \rightarrow I_2)}{\text{Supprt (Item } I_2)}$$

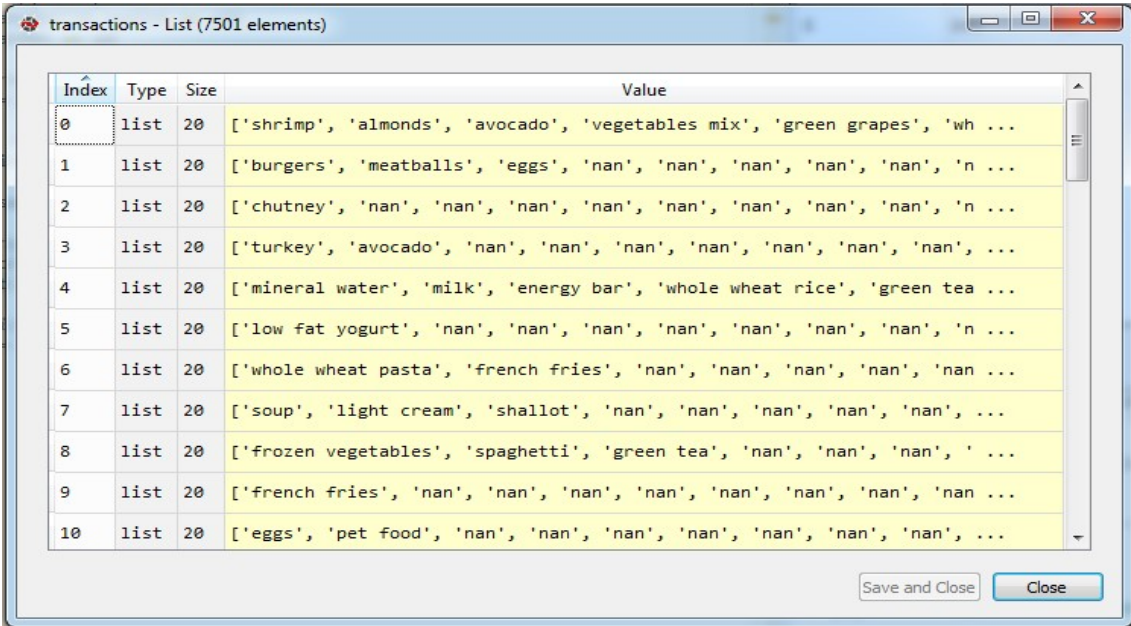
The Dataset

In this execution, we have utilized the Market Basket Optimization dataset that is openly accessible on Kaggle. This dataset contains the rundown of exchanges of a retail organization over the time of multi week. It contains an aggregate of 7501 exchange records where each record comprises of the rundown of things sold in one exchange. Utilizing this record of exchanges and things in every exchange, we will discover the affiliation rules between things.

Market Basket Analysis utilizing the Apriori technique

We need to import the necessary libraries. Python gives the apyori as an API which should be imported to run the apriori calculation.

Customer ID	First Book	Second Book	Third Book	Forth Book	Fifth Book
-------------	------------	-------------	------------	------------	------------



Index	Type	Size	Value
0	list	20	['shrimp', 'almonds', 'avocado', 'vegetables mix', 'green grapes', 'wh ...
1	list	20	['burgers', 'meatballs', 'eggs', 'nan', 'nan', 'nan', 'nan', 'nan', 'n ...
2	list	20	['chutney', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'n ...
3	list	20	['turkey', 'avocado', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', ' ...
4	list	20	['mineral water', 'milk', 'energy bar', 'whole wheat rice', 'green tea ...
5	list	20	['low fat yogurt', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'n ...
6	list	20	['whole wheat pasta', 'french fries', 'nan', 'nan', 'nan', 'nan', 'nan ...
7	list	20	['soup', 'light cream', 'shallot', 'nan', 'nan', 'nan', 'nan', 'nan', ' ...
8	list	20	['frozen vegetables', 'spaghetti', 'green tea', 'nan', 'nan', 'nan', ' ...
9	list	20	['french fries', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan ...
10	list	20	['eggs', 'pet food', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', 'nan', ' ...

Fig 1 - dataset for model

Stages of Algorithm

Stage 1 : In the first stage, we are having with us binary datasets of 4 windows with specifications as explained in previous section. Apriori Algorithm : In this part we just run the apriori algorithm on the binary datasets of all the windows and found frequent itemsets and further association rules from them.

Stage 2 : This stage can be divided in two sub-parts in which two algorithms are run alternatively. Part 1 - ARM-Update : This algorithm creates Score Table and the structure is shown in Fig.1(a) and then updating score table as the data from consecutive windows come.

ARM-Update(Window_i, ConfidenceToScoreTable, ScoreTable)

```
{ For ( i = start-of-Windowi ; i < end-of-Windowi ; i ++)  
    N = AssignScore (ith assoction rule, ConfidenceToScoreTable) ;  
    CreateEntryScoreTable(N,ith association rule); }  
    } where AssignScore (ith association rule, ConfidenceToScoreTable) :
```

It is a function which is taking Input some Association rule and Confidence to ScoreTable and this algorithm is used with algorithm with Part-2 Algorithm and supplies Information processed to Part-2 algorithm which further processes the information. CreateEntryScoreTable (N,ith association rule) : It is a function which create new entry in the score table if some ith association rule is not in the ScoreTable or if present, then just add score N to existing rule. Part 2 - ARM-Predictor : This part is run after we have run ARM-Update Algorithm, this algorithm find the outliers on the basis of some threshold value.

ARM-Outlier

(ScoreTable) { for (i=0;i\$

To complete a MBA you'll initially require an informational collection of exchanges. Every exchange addresses a gathering of things or items that have been purchased together and frequently alluded to as an "itemset". For instance, one itemset may be: {pencil, paper, staples, rubber} in which case these things have been purchased in a solitary exchange. In a MBA, the exchanges are broke down to distinguish rules of affiliation. For instance, one principle could be: {pencil, paper} => {rubber}. This implies that assuming a client has an exchange that contains a pencil and paper, they are probably going to be keen on additionally purchasing an elastic. Prior to following up on a standard, a retailer has to know whether there is adequate proof to propose that it will bring about a useful result. We subsequently measure the strength of a standard by figuring the accompanying three measurements (note different measurements are accessible, yet these are the three most generally utilized):

Backing: the level of exchanges that contain the entirety of the things in an itemset (e.g., pencil, paper and elastic). The higher the help the more much of the time the itemset happens. Rules with a high help are liked since they are probably going to be material to an enormous number of future exchanges.

Certainty: the likelihood that an exchange that contains the things on the left hand side of the standard (in our model, pencil and paper) additionally contains the thing on the correct hand side (an elastic). The higher the certainty, the more prominent the probability that the thing on the correct hand side will be bought or, as such, the more noteworthy the return rate you can expect for a given principle.

Lift: the likelihood of the entirety of the things in a standard happening together (also called the help) isolated by the result of the probabilities of the things on the left and right hand side happening as though there was no relationship between them. For instance, if pencil, paper and elastic happened together in 2.5% of all exchanges, pencil and paper in 10% of exchanges and elastic in 8% of exchanges, at that point the lift would be: $0.025/(0.1*0.08) = 3.125$. A lift of more than 1 proposes that the presence of pencil and paper expands the likelihood that an elastic will likewise happen in the exchange. In general, lift sums up the strength of relationship between the items on the left and right hand side of the standard; the bigger the lift the more noteworthy the connection between the two items.

To play out a Market Basket Analysis and distinguish expected principles, an information mining calculation called the 'Apriori calculation' is normally utilized, which works in two stages:

Methodically recognize itemsets that happen much of the time in the informational index with a help more prominent than a pre-indicated limit. Compute the certainty of all potential principles given the continuous itemsets and keep just those with a certainty more noteworthy than a pre-indicated edge. The edges at which to set the help and certainty are client indicated and are probably going to shift between exchange informational collections.

R has default esteems, however we prescribe that you explore different avenues regarding these to perceive what they mean for the quantity of rules returned (more on this beneath). At long last, albeit the Apriori calculation doesn't utilize lift to build up rules, you'll find in the accompanying that we use lift while investigating the standards that the calculation returns.

Performing Market Basket Analysis in R To exhibit how to complete a MBA we've decided to utilize R and, specifically, the arules bundle. For those that are intrigued we've incorporated the R code that we utilized toward the finish of this blog.

Here, we follow a similar model utilized in the arulesViz Vignette and utilize an informational collection of staple deals that contains 9,835 individual exchanges with 169 things. The primary thing we do is have a gander at the things in the exchanges and, specifically, plot the general recurrence of the 25 most continuous things in Figure 1. This is identical to the help of these things where each itemset contains just the single thing. This bar plot represents some goods that are habitually purchased at this store, and it is eminent that the help of even the most successive things is moderately low (for instance, the most continuous thing happens in just around 2.5% of exchanges). We utilize these bits of knowledge to educate the base limit when running the Apriori calculation; for instance, we realize that all together for the calculation to return a sensible number of rules we'll have to set the help edge at well underneath 0.025.

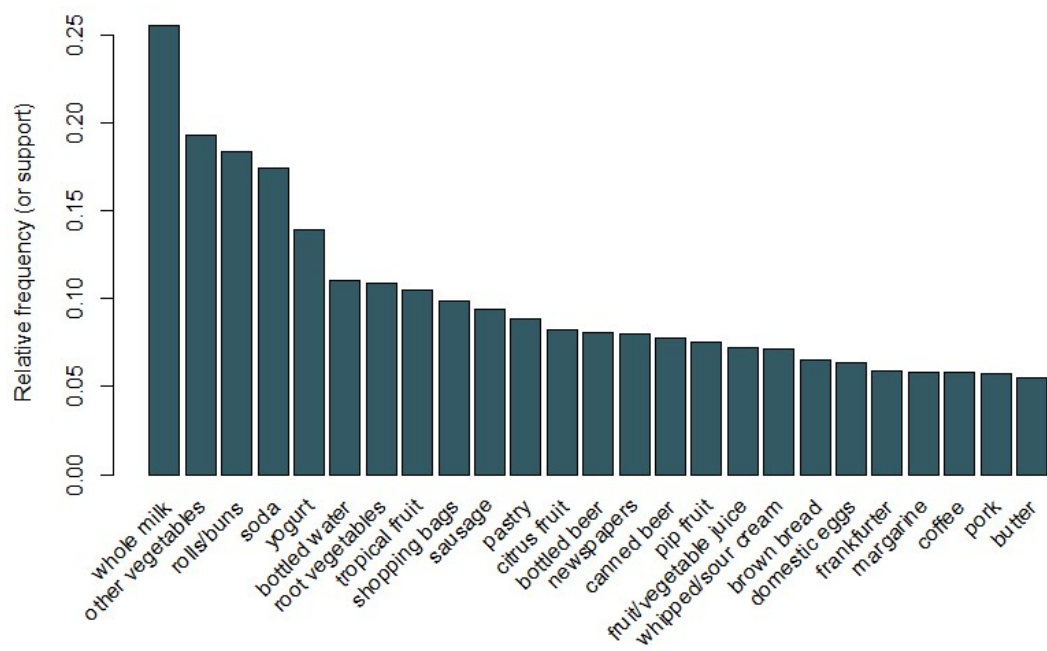


Fig 2 – a plot bar to show 25 frequent item bought in lockdown

these rules seem to make intuitive sense. For example, the first rule might represent the sort of items purchased for a BBQ, the second for a movie night and the third for baking. Rather than using the thresholds to reduce the rules down to a smaller set, it is usual for a larger set of rules to be returned so that there is a greater chance of generating relevant rules. Alternatively, we can use visualisation techniques to inspect the set of rules returned and identify those that are likely to be useful. Using the `arulesViz` package, we plot the rules by confidence, support and lift in Figure 2. This plot illustrates the relationship between the different metrics. It has been shown that the optimal rules are those that lie on what's known as the "support-confidence boundary". Essentially, these are the rules that lie on the right hand border of the plot where either support, confidence or both are maximised. The plot function in the `arulesViz` package has a useful interactive function that allows you to select individual rules (by clicking on the associated data point), which means the rules on the border can be easily identified.

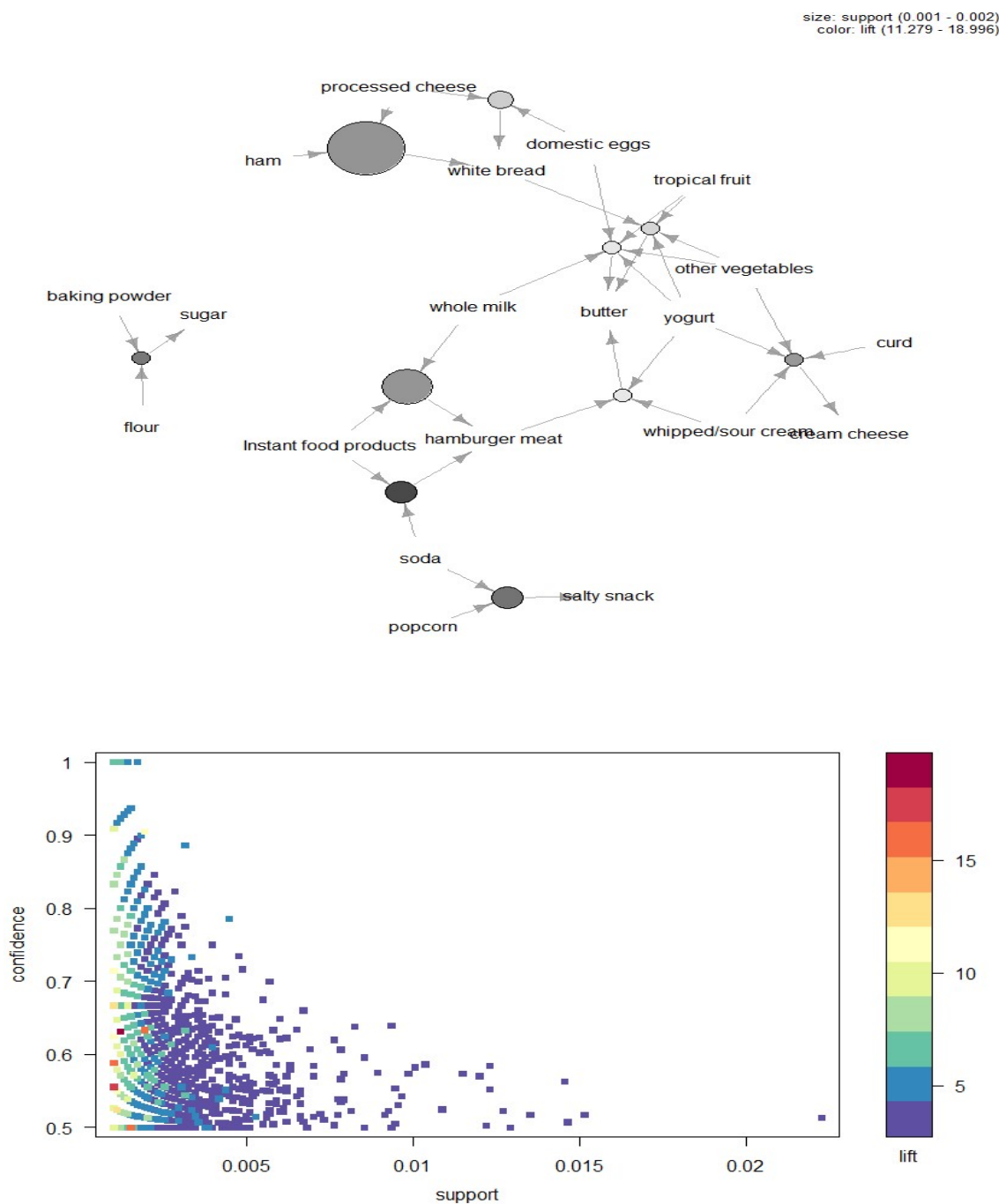


Fig 3 - A scatter plot of the confidence, support and lift metrics and graph-based visualisation of the top ten rules in terms of lift.

A Four Part Solution

After discussing the requirements with the client, Smartbridge designed four self-service reports that provide the data and information needed to answer their questions.

1. Affinity Report -The purpose of this report is to identify the relationships between two products and how often a menu item is sold in conjunction with another. With this report, stakeholders are able to find relationships that are not obvious and make decisions according to the results. market basket analysis pandemic

2. Control Group Report-This report tracks operational KPI's like sales, transactions, and cost by comparing transactions that contain one or several menu items of interest versus the transactions that don't include those menu items. The purpose of this report is to identify how a combination of menu items contribute to sales, transactions, and costs.

3. Combination Report-This report displays the most common combinations of products to determine what products customers purchase together.

4. Transactions Detail ReportThis report contains the details of each transaction such as products sold, quantity, price, sales amount, cashier, discounts, etc. This report is used to answer ad-hoc questions which require more detail.

Success Through Insightful Reporting

The client was able to identify changes and new patterns in products purchased together, track their sales, and identify some regions where they are performing better. The client also ran an analysis in the pre and post COVID period and found some products had unexpected sales growth. The client also detected the most common product combinations based on the sales channel and used all of this information for marketing and operations decisions.

ARM-Update Algorithm Input : Window, Look-Up Table Output : Score Table where Window: It contains association rules for some particular time period Look-Up Table: It contains Confidence to corresponding Score Values Score Table : Association rules along rows and their attributes in columns with their scores ARM-Predictor Algorithm Input : Score Table Output : Outliers Where Score Table : Association rules along rows and their attributes in columns with their scores Outliers: Set of Association Rules which are above score-threshold, Set of Association Rules which are below scorethreshold.

This application was carried out in private store of UIN Malang Business Center, over contributing the gathered exchange information of which the information utilized were exchange information in BC UIN Malang on October first, 2014 with 1553 gathered information. In increases of the put away exchange receipts of 890753 up to 891319 exchange receipts. The further advance was contributing least help and least worth of certainty. In representation, client needed to embed exchange information with receipts from 890753 up to 890853. From that receipt there were 56 exchange information with 20 things could be recovered. Those 56 exchange information were acquired from the exchange receipts whose thing sums were minimumly 2 out of 1 single exchange

The preliminary framework was completed in methods for framework preliminary of market bin examination application via affiliation rule tallying mode utilizing apriori calculation. In this investigation, a conversation of the tests that had been completed on the framework and assessment of the outcomes reacted by the framework. From this test, how designs contrasts of the created information would be known. The test was performed on 56 information by entering the underlying worth

Minimum support= 3

- Minimum confidence= 1

the information that would be utilized for the test, which were the receipts between 890753 up to 890853 with the quantity of exchange information was 56 information arranged in regard of "Receipts" The estimation of the help worth of 1 thing can be gotten through the accompanying recipe:

Support (A) = The Number of Transaction which contains A Total Transaction (5)

While the help an incentive for 2 things can be acquired with the accompanying equation: *Support(A, B) = Transaction Number Which Contains A and B □□□Total Transaction* Confidence esteem computation was produced of this after recipe: *Confidence = P(A|B) = The Number of Transaction Which Contains A and B Total Transaction which contains* ♦

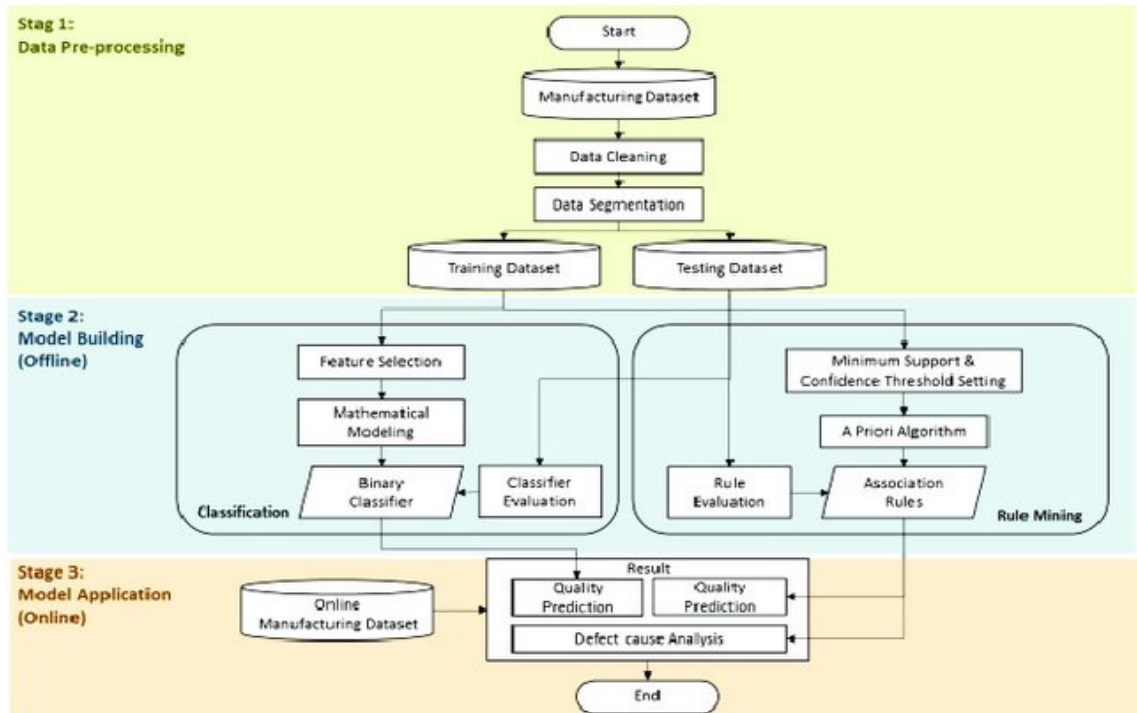
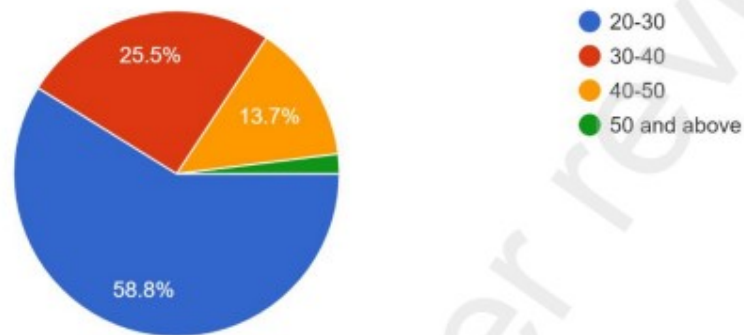


Fig 4 – system design flowchart

Points for analyzing results The results below are shown in order as follows: 1. In the Score Table as shown in Fig 1(a), the attributes are set aside row wise at the top and named them a,b,c and so on for the simplicity of the transaction. a,b,c and so on are the items which are kept in market basket. As shown in the after the top row of total number of items, there are association rules with their assigned score. 2. Upper Association Rules, meaning rules which are above threshold are being printed. 3. Lower Association Rules, meaning the association rules which are below threshold.

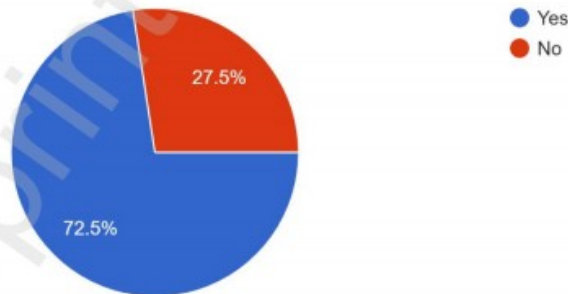
5. Survey

What is your age group?
153 responses



In this questionnaire, there are different age group respondents, as shown in the chart, 58.8 are between the 20-30 age group, which are the maximum number of respondents who participated in the survey and next to the maximum number of respondents are 30-40 of age group and 13.7% are between the 40-50 of age group and the least amount of respondents are 50 and above of age group.

Have you purchased essential (FMCG) goods online before COVID-19 Pandemic Lockdown?
153 responses



This question shows that how many customers used to buy before pandemic lockdown and how many customers have not bought anything through online platforms. So as in the figure, it shows that 72.5% of respondents have bought (FMCG) goods online before lockdown and 27.5% respondents have not even considered buying (FMCG) goods online.

Why do you purchase online (FMCG) goods online?

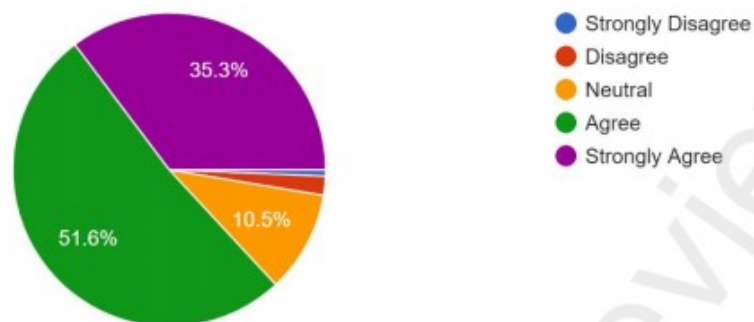
153 responses




Now this question asks respondents that, before they buy online (FMCG) goods, what is the thing which triggered them to use online platform rather than buying (FMCG) goods through offline retailers. So, in this 66.7% respondents say that they buy online (FMCG) goods online For Convenience 22.7% says that they buy online groceries for a better discount, 5.9% respondents say that they don't have any specific reason behind on buying goods online and only 5.2% respondents say that there are some other reasons which convince them to buy goods (FMCG) online.

Has COVID-19 pandemic lockdown increased your frequency of buying online?

153 responses



This chart shows that during the pandemic lockdown, the tendency of buying online (FMCG) goods has increased or not. So, the maximum number of respondents which is 51.6% says that they agree that during this lockdown they prefer to buy goods (FMCG) online, 35.3 respondents strongly agree on this

[illegible]

The screenshot shows the Microsoft Access 2003 interface. The 'Database Properties' dialog box is open, with the 'Tables' tab selected. The 'Table Properties' section is visible, showing the 'Table Name' as 'tbl_1' and the 'Table Type' as 'Table'. The 'Table Properties' section has two tabs: 'Table Properties' and 'Table Design'. The 'Table Design' tab is active, showing a table with one column named 'tbl_1' and one row with the value 'tbl_1'.

The screenshot shows the Orange3 software interface. The top menu bar includes File, Edit, View, Project, Build, Debug, Database, Mining Modeler, Tools, Window, and Help. Below the menu is a toolbar with various icons. The main window has a tabbed interface with several tabs: 'File', 'Mining Modeler', 'Prediction', 'Case1.dmn', 'Case2.dmn', 'Mining Structure', 'Mining Models', 'Mining Model Review', 'Mining Accuracy Chart', and 'Mining Model Prediction'. The 'Mining Modeler' tab is active. On the left side, there is a sidebar with a tree view showing the 'Expression' subcategory expanded. The main workspace displays a workflow diagram with three widgets: 'File' (orange), 'Mining Modeler' (blue), and 'Prediction' (green), connected by arrows indicating the data flow.

The best forecasts for singular clients, in view of past buys can be known with the assistance of numerous expectations inquiry. Various Predictions Query is made as demonstrated in screen capture 7 and expectations for singular clients is acquired as appeared in screen capture 8, where party ID is the contact quantities of the clients who make the buys and articulation gives the items that can be proposed to these clients dependent on past buys. This can likewise be useful for unwaveringness clients also, for making certain ideas to them.

Fig 9 Best predictions for individual customers

[illegible]

Outlier Detection

Confidence fluctuation of top 30 association rules with highest confidence in 3 months

Legend: 3 month (blue), April (orange), May (grey), June (yellow)

Rule Index	3 month	April	May	June
1	0.7000	0.7000	0.7000	0.7000
2	0.6500	0.6500	0.6000	0.6500
3	0.6200	0.6200	0.6000	0.6200
4	0.6200	0.5800	0.6200	0.6200
5	0.6000	0.5800	0.5800	0.6000
6	0.5800	0.5800	0.5800	0.5800
7	0.5800	0.5500	0.5800	0.5800
8	0.5800	0.5500	0.5800	0.5800
9	0.5800	0.5200	0.5800	0.5800
10	0.5500	0.5500	0.5500	0.5500
11	0.5500	0.5200	0.5500	0.5500
12	0.5200	0.5200	0.5200	0.5500
13	0.5200	0.5200	0.5200	0.5500
14	0.5200	0.5200	0.5200	0.5500
15	0.5200	0.5200	0.5200	0.5500
16	0.5200	0.5200	0.5200	0.5200
17	0.5200	0.5200	0.5200	0.5200
18	0.5000	0.4800	0.5000	0.5000
19	0.5000	0.4800	0.5000	0.5000
20	0.4800	0.4800	0.4800	0.4800
21	0.4800	0.4800	0.4800	0.4800
22	0.4800	0.4800	0.4800	0.4800
23	0.4800	0.4800	0.4800	0.4800
24	0.4800	0.4800	0.4800	0.4800
25	0.4800	0.4800	0.4800	0.4800
26	0.4800	0.4800	0.4800	0.4800
27	0.4800	0.4800	0.4800	0.4800
28	0.4800	0.4800	0.4800	0.4800
29	0.4800	0.4800	0.4800	0.4800
30	0.4800	0.4800	0.4800	0.4800

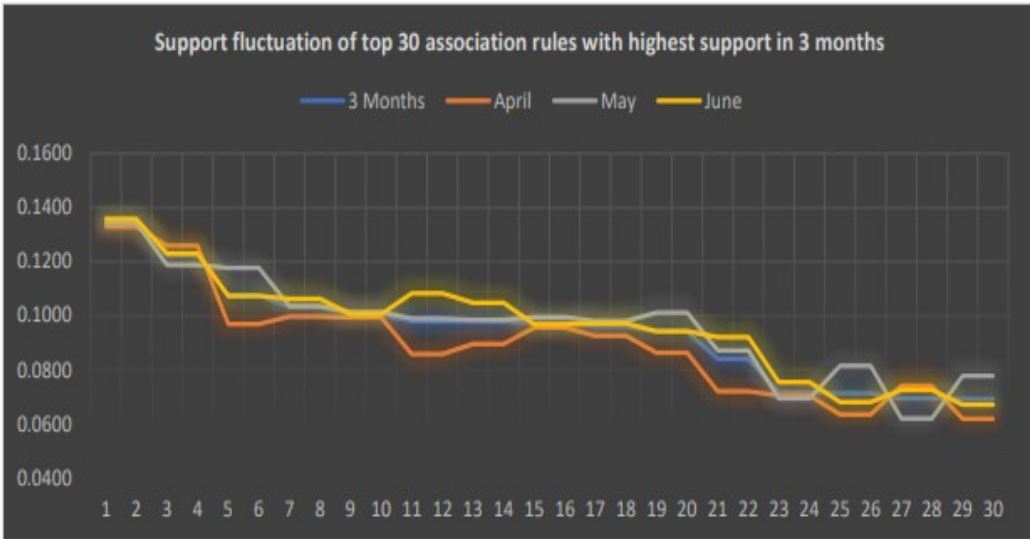
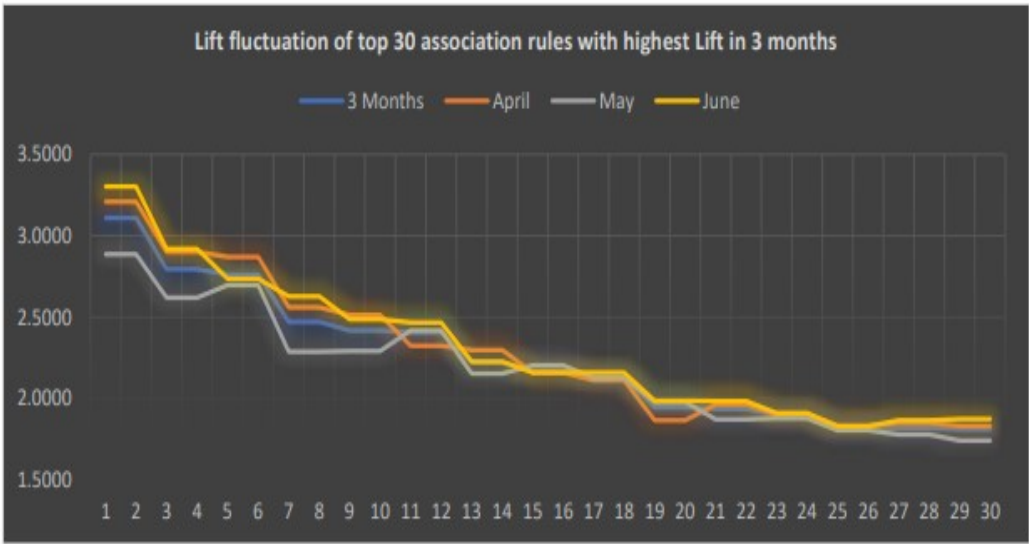


Table 1 : association rules that affect product quality.

LHS	Support	Confidence	Lift
{V65=very-high,V81=low}	0.0283	1	2.005
{V29=normal-high,V65=very-high}	0.0247	1	2.005
{V65=very-high,V320=normal-high}	0.0238	1	2.005
{V386=high,V547=low}	0.0224	1	2.005
{V354=low,V576=high}	0.0210	1	2.005
{V67=very-low,V512=high}	0.0206	1	2.005
{V67=very-low,V344=normal-high}	0.0206	1	2.005
{V274=very-high,V384=normal-low}	0.0206	1	2.005
{V246=normal-low,V274=very-high}	0.0206	1	2.005
{V274=very-high,V518=normal-low}	0.0206	1	2.005
{V274=very-high,V383=normal-low}	0.0206	1	2.005
{V245=normal-low,V274=very-high}	0.0206	1	2.005
{V274=very-	0.0206	1	2.005

References

1. Loraine Charlet Annie M.C. and Ashok Kumar D, "MARKET BASKET ANALYSIS FOR A SUPERMARKET BASED ON FREQUENT ITEM SET MINING", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, September 2012.
2. KavithaVenkatachari and IssacDavanbuChandrasekaran, "MARKET BASKET ANALYSIS USING FP GROWTH AND APRIORI ALGORITHM: A CASE STUDY OF MUMBAI RETAIL STORE", BVIMSR's Journal of Management Research Vol. 8 Issue- 1: April 2016.
3. R. Karpagam and Dr. S. Suganya, "Applications Of Data Mining And Algorithms In Education – A Survey", International Journal of Innovations in Scientific and Engineering Research (IJISER), Vol.3,no.4, pp.38-46, 2016.
- 4 M. Kaur and S. Kang, "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining," Procedia Comput. Sci., vol. 85, no. Cms, pp. 78–85, 2016.

5. A. Bertoni and T. Larsson, "ScienceDirect Data Mining in Product Service Systems Design: Literature Review and Research Questions," *Procedia CIRP*, vol. 64, pp. 306–311, 2017.
7. Wu X, Kumar V, Quilan JR., Ghosh J, Yang Q, Motoda H. *Top 10 Algorithms in Data Mining*. Springer-Verlay London Limited 2007:14:1-37.
8. Ngai EWT, Xiu Li, Chau DCK. *Application of Data Mining Techniques in Customer Relationship Management: A Literature Review and Classification*. Elsevier-Expert Systems with Applications 2009:36:2592-2602.
9. Ramageri Bharati M. *Data Mining Techniques and Applications*. Indian Journal of Computer Science and Engineering:1(4):301-305.
10. Hawkins D. *Identification of Outliers*. Chapman and Hall 1980.
11. Zhao Quiankun, Bhowmick Sourav. *Association Rule Mining: A Survey*. Technical Report CAIS Nanyan Technological University, Singapore 2003:1-20.
12. Trnka Andrej. *Market Basket Analysis with Data Mining Methods*. International Conference on Networking and Information Technology 2010:446-450.
13. Hilage Tejaswini A, Kulkarni RV. *Application of data mining techniques to a selected business organization with special reference to buying behavior*. International Journal of Database Management Systems 2011:3(4):169-181.
14. Gupta Savi, Mamtara Roopal. *A Survey on Association Rule Mining in Market Basket Analysis*. International Journal of Information and Computational Technology 2014:4(4):409-414.
15. He Zengyou, Xu Xiaofei, Huang Joshuaz, Deng Shengchun. *FP-Outlier: Frequent Pattern Based Outlier Detection*. ComSIS 2005:2(1):103- 118.
16. Khan Mohiuddin A, Pradhan Sateesh K, Khaleel MA. *Outlier Detection for Business Intelligence using Data Mining Techniques*. International Journal of Computer Applications 2014:106(2):28-31.
17. Kargupta H, Han J, Yu Philips S, Motwani R, Kuman Vipin. *Book: Next Generation of Data Mining*. CRC Press, Taylor and Francis Group, LLC 2008.
18. Verma Dipti, Nashine Rakesh. *Data Mining: Next Generation Challenges and Future Directions*. International Journal of Modeling and Optimization 2012:2(5):603-608.
19. Aneesh Reddy, (April, 14, 2020) Covid-19 impact: Consumers move towards digital <https://www.thehindubusinessline.com/opinion/covid-19-impact-consumers-move-more-towardsdigital/article31337127.ece>
20. Bansal, S., (April 02, 2020) Opinion|will consumer behaviour see shift post covid-19 <https://www.theweek.in/news/biz-tech/2020/04/23/covid-19-impact-permanent-shift-in-consumer-preferences-to-home-personal-hygiene-products.html>
21. Burgen, S.; Jones, S. Poor and Vulnerable Hardest Hit by Pandemic in Spain | World News | The Guardian. Available online: <https://www.theguardian.com/world/2020/apr/01/poor-and-vulnerable-hardest-hit-by-pandemic-in-spain> (accessed on 25 April 2020).

22. Todorov, T. *The Conquest of America: The Question of the Other*; Harper & Row: New York, NY, USA, 1984; ISBN 9780060151805.
23. Kupperberg, P. *The Influenza Pandemic of 1918–1919*; Chelsea House: New York, NY, USA, 2008; ISBN 978-0-7910-9640-6.
24. Kang, D.; Choi, H.; Kim, J.H.; Choi, J. Spatial epidemic dynamics of the COVID-19 outbreak in China. *Int. J. Infect. Dis.* 2020, 94, 96–102. [CrossRef]
25. Rocklöv, J.; Sjödin, H. High population densities catalyse the spread of COVID-19. *J. Travel Med.* 2020, 27. [CrossRef]
26. Praharaj, S.; Vaidya, H. The Urban Dimension of COVID-19 in India: COVID Outbreak and Lessons for Future Cities. Available online: https://www.researchgate.net/publication/341616744_The_urban_dimension_of_COVID-19_in_India_COVID_Outbreak_and_Lessons_for_Future_Cities?channel=doi&linkId=5ecb837492851c11a8880043&showFulltext=true (accessed on 21 May 2020).
27. Ito, H.; Hanaoka, S.; Kawasaki, T. The cruise industry and the COVID-19 outbreak. *Transp. Res. Interdiscip. Perspect.* 2020, 5, 100136. [CrossRef]
28. McCloskey, B.; Zumla, A.; Ippolito, G.; Blumberg, L.; Arbon, P.; Cicero, A.; Endericks, T.; Lim, P.L.; Borodina, M. Mass gathering events and reducing further global spread of COVID-19: A political and public health dilemma. *Lancet* 2020, 395, 1096–1099. [CrossRef]
29. Diele-Viegas, L.; Pereira, E. COVID-19 and Amazonia: Sustainable alternatives for an economic collapse. *Preprints* 2020, 1, 1–20. [CrossRef]
30. Dobos, S. On the need for interdisciplinarity and historiography in research on the rural space. *Ecoforum J.* 2016, 5, 259–265
31. [Figure 5, 6, 7, 9] Mohit Yadav. Retrieved from <https://www.geeksforgeeks.org/apriorialgorithm/45>
32. [Figure 8] Susan Li (September 24th, 2017). Retrieved from <https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysisassociation-rules-fa4b986a40ce>
33. [Table 1] University of Michigan, John W. Creswell (January 10th, 2017). Retrieved from *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* [Book]