Chapter 13
Market Basket Analysis

Abstract Market basket analysis scrutinizes the products customers tend to buy together, and uses the information to decide which products should be cross-sold or promoted together. The term arises from the shopping carts supermarket shoppers fill up during a shopping trip. The rise of the Internet has provided an entirely new venue for compiling and analyzing such data. This chapter discusses the key concepts of "confidence," "support," and "lift" as applied to market basket analysis, and how these concepts can be translated into actionable metrics and extended.

13.1 Introduction

Marketing researchers have been interested in studying product affinity for a long time. We have learned from the introductory economics or marketing course that coffee and sugar are complements while coffee and tea are substitutes. The price reduction of a product not only increases its own demand but also increases demand of its complementary product. That is, if two products are complements for each other, their demands tend to be positively associated. On the other hand, if two products are substitutes for each other, their demands tend to be negatively correlated since the price reduction of a product would decrease the demand of its substitute.

Marketing practitioners are interested in product affinities because they provide very useful information for designing various marketing strategies. It may not be surprising to a supermarket manager to see that coffee is purchased with coffee cream or sugar. In fact, an experienced manager may know lots of product pairs purchased together by consumers. However, considering that the typical supermarkets carry tens of thousands items, it is also likely that there are thousands of associated product pairs the manager may not have recognized. Maybe the best-known example in the data mining industry

is that beers and diapers tend to be purchased together in the supermarket.[1] Whatever the reasons are, the beer–diaper association is not obvious to the manager. Market basket analysis is designed to find these types of product associations with minimal human interaction.

Typically the input to a market basket analysis is point-of-sale (POS) transaction data at the customer level. Market basket analysis extracts many interesting product associations from transaction data. Hence its output consists of a series of product association rules: for example, if customers buy product A they also tend to buy product B. Market basket analysis alleviates managerial effort and automates the process for finding which products are purchased together. Let the data speak for itself.

Market basket analysis was originally applied to supermarket transaction data. Actually it takes its name from the fact that consumers in a supermarket place all of their purchased items into the shopping cart or the market basket. Nowadays the application of market basket analysis is not limited to the supermarket. It can be applied to any industry selling multiple products such as banks, catalogers, direct marketers and so on, and to new sales channels, especially the Internet.

13.2 Benefits for Marketers

The output of a market basket analysis is a series of association rules. These rules are used to improve the efficiency of marketing strategies and tactics. We learn from the analysis which products/services are purchased at the same time or in a particular sequence. Hence the rules can be very useful and actionable for firms dealing with multiple products/services. Examples are retailers, financial institutions (e.g., credit cards company), catalog marketers, direct marketers, Internet merchants, and so on (Berry and Linoff 1997). Market basket analysis is especially popular among retailers because of their large number of SKUs. In a recent Aberdeen Group survey, 38% of the retailers polled said they used market basket analysis and felt it had a positive effect on their business (Nishi 2005).

Market basket provides valuable information for firms to develop various marketing strategies and tactics. First, association rules from a market basket analysis can be used for a supermarket to manage its shelf space. It may stock the associated items close together such that consumers would not forget to purchase both items. On the other hand, it may stock the associated items far apart such that consumers would spend more time browsing aisle by

[1] Thomas Blischok first discovered this interesting statistical pattern. As vice president of industry consulting for NCR, he did a study for Osco Drug in 1992 when he discovered dozens of correlations, including one connecting beer and diapers in transactions between 5 p.m. and 7 p.m. Blischok recounted the tale in a speech, and the story became the legend in data mining industry (Forbes 1998).

aisle (Chain Store Age 1998).[2] Other types of merchants such as retailers, catalogers and Internet may realize similar benefits.

Second, market basket analysis can be used for designing various promotional strategies. It will provide ideas on product bundling. In addition, it can be used to design a cross-coupon program where consumers purchasing an item A get the (discount) coupon for an item B.[3] Or it will help managers to select appropriate items to be loss leaders.

Third, market basket analysis with temporal components can be very useful to various marketers for selecting cross-selling items. For example, market basket analysis might indicate that customers who have purchased whole life insurance tend to purchase property insurance within 6 months. It suggests a cross-selling possibility – the insurance salesperson should contact his/her current customers with whole life insurance (within 6 months) and try to cross-sell the property insurance.

13.3 Deriving Market Basket Association Rules

Since the seminal paper by Agrawal et al. (1993), the problem of deriving association rules have been widely studied within the field of knowledge discovery (Agrawal and Srikant 1994; Mannila et al. 1994; Silverstein et al. 1998; Zhang 2000), and is often called the market basket problem. In this section, we study how a market basket analysis works and derives various association rules that are "interesting."

13.3.1 Setup of a Market Basket Problem

The input for a market basket analysis is customer-level transactions data, although it is not necessary that each customer be explicitly identified. For example, grocery stores record each customer's transaction data ("market basket") with their scanning device even though they do not know the customer's name, address, etc. For each transaction, the store knows the date, the casher number, items purchased, prices of each item, coupons redeemed, and so on. Table 13.1 shows the hypothetical transaction data from a grocery store. There are five transactions and each transaction consists of a set of

[2] As discussed later, market basket analysis is an exploratory data mining tool. Once an association between two products is identified, we should test two different shelf strategies (stocking two products adjacent or far apart) with control groups.

[3] Dhar and Raju (1998) have developed a model to study the effects of cross-ruff coupons on consumer choice behavior and derived conditions under which cross-ruff coupons can lead to higher sales and profits than other types of package coupons. However, they did not employ market basket analysis for their empirical application.

Table 13.1 Transaction data from a grocery store

Transactions	Items purchased (market basket)
1	Milk, orange juice, ice cream, beer, soap
2	Milk, ice cream, beer
3	Milk, orange juice, detergent
4	Milk, ice cream, pizza
5	Milk, orange juice, soap

items. The main focus of market basket analysis is the set of items purchased for each transaction. From this transaction data, market basket analysis provides a series of association rules where we infer which items are purchased together.

Each association rule consists of an antecedent and a consequent. For example, consider the association rule, "if a consumer purchases item A, s/he also tends to purchase item B." Here item A is the *antecedent* while item B is the *consequent*. Note that both antecedent and consequent can contain multiple items.

13.3.2 Deriving "Interesting" Association Rules

Let us intuitively derive a few association patterns from Table 13.1. At first glance, we can see that milk and orange juice are purchased together in three out of the five transactions. This observation may tell us that there is a cross-selling possibility between milk and orange juice. Anything else? Ice cream and beer are purchased together in two out of the five transactions. Again from this pattern, we may suggest an association rule like: "if a customer purchases ice cream, then s/he also purchases beer," or more compactly, "if ice cream then beer." Similarly, we can formulate an association rule between orange juice and soap.

We can generate many association rules from Table 13.1 but we are only interested in selecting "interesting" rules. That is, how managerially relevant are the rules we have generated? It is difficult to come up with a single metric quantifying the "interestingness" or "goodness" of an association rule (Bayardo and Agrawal 1999). Hence, researchers have proposed several different metrics. There are three most popular criteria evaluating the quality or the strength of an association rule: support, confidence and lift.

Support is the percentage of transactions containing a particular combination of items relative to the total number of transactions in the database. We can think of the support for an individual item A, which would just be the probability a transaction contains item A, or "P(A)". However, when we are interested in associations, we are concerned with multiple items, so the support for the combination A and B would be P(AB). For example, consider the association rule "if milk then beer" from Table 13.1. Support

measures how often milk and beer are purchased together, as a percentage of the total number of transactions. They are purchased together two out of five transactions. Hence, support for the association rule is 40%.

Support for multiple items can be interpreted as a joint probability. It measures the probability that a randomly selected basket contains item A and item B together. Hence it is symmetric and does not hint at cause-and-effect. We know that the joint probability of A and B, $P(\text{AB})$, is no different than the joint probability of B and A, $P(\text{BA})$. For example, support for the association rule "if milk then beer" would be the same as the support for the association rule "if beer then milk."

Support has one critical disadvantage in evaluating the quality of an association rule. The example in Table 13.1 shows that the association rule "if beer then milk" has support of 40%. However, is the association rule "if beer then milk" an interesting rule? The answer is yes if this means that 40% of customers buy beer and milk together and no one buys milk without buying beer. However, Table 13.1 shows that all the transactions contain milk. All customers buy milk and only 40% of those buy beer. Hence, the association rule "if beer then milk" is not interesting even if its support is 40%. Milk is so popular in grocery shopping (by itself it has very high support) that the support for milk plus any other item can be large.

Confidence measures how much the consequent (item) is dependent on the antecedent (item). In other words, confidence is the conditional probability of the consequent given the antecedent, $P(\text{B}|\text{A})$. For example, the confidence for the association rule "if ice cream then beer" is 66% since three transactions contain ice cream (the antecedent) and two among the three transactions also contain beer (the consequent). In other words, given that the baskets containing ice cream is selected, there is 66% chance that the same basket also contains beer. Different from support, confidence is asymmetric. For example, the confidence of "if beer then ice cream" is 100% while the confidence of "if ice cream then beer" is 66%.

The law of conditional probability states that $P(\text{B}|\text{A}) = P(\text{AB})/P(\text{A})$. That is, confidence is equal to the support of the association rule divided by the probability or the support of the antecedent. For example, the support of an association rule "if ice cream then beer" is 40% (two out of five transactions) while the support or the probability of ice cream is 60% (three out of five). Hence, its confidence is 66% (40%/60%).

Confidence surely is a good criterion for selecting interesting rules but is not a perfect criterion. Consider a rule "if ice cream then orange juice." Its confidence or $P(\text{B}|\text{A})$ is 33% so you may think it is an interesting rule. However, there is 60% chance (e.g., $P(\text{B}) = 60\%$) that a randomly chosen transaction contains orange juice. Hence, ice cream is not a powerful antecedent for identifying an orange juice purchase – it has lower than a random chance of identifying an orange juice purchase. Thus there is no cross-selling opportunity.

Lift (also called improvement or impact) is a measure to overcome the problems with support and confidence. Consider an association rule "if A then B." The lift for the rule is defined as $P(\text{B}|\text{A})/P(\text{B})$ or $P(\text{AB})/[P(\text{A})P(\text{B})]$. As shown in the formula, lift is symmetric in that the lift for "if A then B" is the same as the lift for "if B then A."

$P(\text{B})$ is the probability that a randomly chosen transaction contains item B. In other words, it is an unconditional (or baseline) probability of purchasing item B regardless of other items purchased. Practitioners often use the term, "expected confidence" for $P(\text{B})$ instead of unconditional probability.

Hence, lift is said to measure the difference – measured in ratio – between the confidence of a rule and the expected confidence. For example, the lift of an association rule "if ice cream then beer" is 1.67 because the expected confidence is 40% and the confidence is 67%. This means that consumers who purchase ice cream are 1.67 times more likely to purchase beer than randomly chosen customers. That is, larger lift means more interesting rules.

A lift of 1 has a special meaning. We know that $P(\text{AB}) = P(\text{A})P(\text{B})$ if A and B are independent. Therefore, lift equals one if the event A is independent of the event B. Lift greater than 1 indicates that the item A and the item B tend to occur together more often would be predicted by random chance. Similarly, lift smaller than 1 indicates that the item A and item B are purchased together less likely than would be predicted by random chance.

Lift has little practical value when the support for the antecedent item is very low. For example, suppose that $P(\text{mushroom pizza \& ice cream}) = 0.01$, $P(\text{mushroom pizza}) = 0.01$ and $P(\text{ice cream}) = 0.25$. The association rule "if mushroom pizza then ice cream looks like a good rule based on its lift of 4. However, only a small number of customers purchase mushroom pizza. A co-marketing program designed to encourage mushroom pizza buyers to purchase ice cream may not have a high impact. This problem can be partially resolved by taxonomies described in Sect. 13.4.1.

Summarizing, we have introduced three popular criteria for evaluating association rules in market basket analysis, defined as follows:

$$\text{Confidence} = P(\text{B}|\text{A}) \tag{13.1a}$$

$$\text{Support} = P(\text{BA}) \tag{13.1b}$$

$$\text{Lift} = P(\text{B}|\text{A})/P(\text{B}) \tag{13.1c}$$

Each criterion has its advantages and disadvantages but in general we would like association rules that have high confidence, high support, and high lift. Association rules with high support are potentially interesting rules. Similarly, rules with high confidence would be interesting rules. Or you may look for association rules with very high or very low lift.[4] Practitioners generally

[4] The very low lift implies that the two products "repel" each other. Substitutes (e.g., Coke and Pepsi) tend not to be in the same basket. Knowing that two products "repel" each other can often suggest actionable recommendations. For example, Coke should not be promoted together with Pepsi.

employ all three together in generating a set of interesting association rules. They might set a threshold for each rule and let the market basket software choose rules to meet the condition (see Chapter 21 for further discussion). For example, practitioners might ask the software to find all associations so that support, confidence, and lift are all greater than some minimum threshold specification (e.g., see Yan et al. 2005).

13.3.3 Zhang (2000) Measures of Association and Dissociation

Other than three metrics discussed above, researchers have proposed a number of measures including chi-square value (Morishita 1998), entropy gain (Morimoto et al. 1998; Morishita 1998), gini (Morimoto et al. 1998) and laplace (Webb 1995). More recently, Zhang (2000) proposed a new metric that was theoretically shown to be better than traditional measures such as the confidence and/or the χ^2 test. He also applied his new measure (along with traditional measures) to a POS transaction data and a donation data, and showed that his measure could identify association patterns not discovered by traditional measures. Considering the importance of finding a good measure of association rules, we describe his measure with comparing others.

Zhang's point of departure is to recognize the difference between association and *disassociation*. If the probability of co-occurrence $P(A|B)$ for patterns A and B is larger than probability of no co-occurrence $P(A|\overline{B})$, then the relationship of A with B is association (attractive). Otherwise, the relationship is disassociation (repulsive). Association is described by $P_A(B \Rightarrow A) = 1 - P(A|\bar{B})/P(A|B)$ if $P(A|\bar{B}) < P(A|B)$. Disassociation is described by $P_D(B \Rightarrow A) = P(A|\bar{B})/P(A|B) - 1$ if $P(A|\bar{B}) \geq P(A|B)$. Combining the two formulas, we obtain

$$\begin{aligned} P(B \Rightarrow A) &= \frac{P(A|B) - P(A|\bar{B})}{Max[P(A|B), P(A|\bar{B})]} \\ &= \frac{P(AB) - P(A)P(B)}{Max[P(AB)(1 - P(B)), P(B)(1 - P(A))]} \end{aligned} \tag{13.2}$$

where $B \Rightarrow A$ (e.g., B implies A) describes the association of A with B. For example, let us calculate P(beer $\Rightarrow$ ice cream) in Table 13.1. Since P(ice cream|beer) = 1 is larger than P(ice cream|not beer) = 1/3, so the relationship of ice cream with beer is association. And P(beer $\Rightarrow$ ice cream) is equal to 2/3.

The association metric in Equation 13.2 is asymmetric. That is, $P(B \Rightarrow A)$ can be different from $P(A \Rightarrow B)$. Zhang's metric has several other good properties. For example, consider three extreme cases: perfect association, perfect disassociation, and random or independent association. A good measure of

association should yield a definitive result for each case. In other words, the result a measure of association should yield a constant number, independent of $P(A)$ and/or $P(B)$, for perfect association, perfect disassociation, or independent association.

The following table calculates the value for support, confidence, lift, and Zhang's measure for each of the three cases:

	Support	**Confidence**	**Lift**	**Zhang**
Perfect Association	$P(A)(=)P(B)$	1	$1/P(B)$	1
Perfect Dissociation	0	0	0	-1
Independence	$P(A)P(B)$	$P(B)$	1	0

The table shows that only Zhang's measure provides a unique number for all three cases. This means that Zhang's measure has a similar interpretation as a correlation coefficient: values close to 1 signify almost perfect positive association, values close to -1 signify almost perfect negative association, and values close to zero mean very little relationship. None of the other measures has this nice numerical interpretation.

13.4 Issues in Market Basket Analysis

13.4.1 Using Taxonomies to Overcome the Dimensionality Problem

The typical supermarket in the US carries about 30,000 items or SKUs (stock keeping units). It means that we need to evaluate about 4.5×10^8 potential association rules such as the ones "if A then B." Furthermore, as discussed later, you may be interested in association rules involved in more than two items. The "curse of dimensionality" comes into play unless we control the number of items to a manageable size.

Much research has focused on algorithms for computing all relevant associations (Agrawal et al. 1993; Agrawal and Srikant 1994; Hu et al. 2000; Yan et al. 2005). However, another way to overcome dimensionality problem is to aggregate items into some manageable number of categories. For example, Tropicana orange juices with various sizes can be grouped into the Tropicana orange juice category. Or different types of Tropicana juices such as orange and grape juice may be aggregated. More generalized categories such as the juice category (after aggregating all juices with different brands, sizes and types) can also be used as the input for market basket analysis.

There is another benefit from item aggregation. Unit sales of many SKUs in the original market basket data are so small. Hence, their supports are

extremely low. As shown in the previous section, the presence of low-support items may make it difficult to find good association rules. For example, suppose that there is only one transaction including Brand A orange juice in the entire basket data. And the transaction containing Brand A orange juice also includes yogurt. The confidence of the association rule, "if Brand A orange juice then yogurt," for this basket data is 100%. But it is not an interesting association rule. We can easily avoid this problem through item aggregation.

Obviously, as we employ higher levels of aggregation, the computational burden for the market basket analysis is diminished. However, item aggregation often leads to the loss of transaction details useful for developing actionable marketing strategies. Suppose that the result of the market basket analysis suggests the cross-promotion opportunity between beer and orange juice. It is unusual for a supermarket to promote all items in the orange juice category together. Instead, they promote a particular brand of orange juice. The market basket data aggregated across brands does not allow the manager to select a brand for target promotion.

What is the right level of item aggregation for a market basket analysis? Practitioners often suggest aggregating items such that each of the resulting aggregates have roughly the same level of appearance or support in the market basket data (for example, see www.megaputer.com/html/mba.html). As a result, items with smaller unit sales will be grouped together so that we can avoid the problem of bad association rules due to the low support items. However, one should not apply this suggestion too strictly. The needs of the end user are more important in deciding the level of aggregation. For example, the marketing manager in a discount store will be more interested in selling a television than a DVD. That is, it may be more reasonable to aggregate cheap items than expensive items.

13.4.2 Association Rules for More than Two Items

So far, we have investigated association rules with two items – one antecedent and one consequent. However, managers might be interested in association rules involving more than two items. The idea behind market basket analysis with two items can be easily extended to the analysis of more than two items. For example, consider an association rule, "if A and B then C." The support of this association rule is $P(\text{ABC})$ and its confidence is $P(\text{C}|\text{AB})$. And $P(\text{C}|\text{AB})/P(\text{C})$ is its lift. Similar analysis can be performed for the sets of four items, five and so on.

As discussed above, the curse of dimensionality comes into play as the number of items considered simultaneously increases. The number of calculations to perform the market basket analysis increases exponentially with the number of items to be considered together. For example, going back to the supermarket with 30,000 items, we need to evaluate ${}_{30000}C_3 (\approx 4.5 \times 10^{12})$

potential association rules such as the ones "if A and B then C." And about 3.4×10^{16} calculations are required for the sets of four items.

Researchers have suggested various pruning methods to overcome the dimensionality problem associated with the market basket analysis of multiple items (Agrawal et al. 1993). One easy pruning method is to generate association rules that satisfy a given support constraint. In addition, is the pruning is performed iteratively so that the number of calculations can be minimized. For example, given the support constraint of 1%, any items less than this minimum support are first eliminated and only the remaining items are used for the analysis of two items. For association rules for three items, any pairs of items less than the support constraint are eliminated and only the remaining pairs of items are used as antecedents. Similar iterative pruning is applied for generating association rules involved with more than three items. Yan et al. (2005) assert that even simple pruning rules that use thresholds can result in too many calculations, and propose a genetic algorithm for producing interesting associations.

13.4.3 Adding Virtual Items to Enrich the Quality of the Market Basket Analysis

Market basket analysis has originally been developed to study association patterns among items sold in supermarket. However, it becomes a much more useful data mining tool when items considered are not restricted to real products. Virtual items are not real products sold in retail stores, but they are treated as items in the market basket analysis. For example, marketing managers may be interested in knowing which items are sold well with male customers. The market basket analysis can provide this information simply by adding one more virtual item (sex identifier: "male" or "female") to each transaction basket.

Practically, the number of virtual items can be unlimited. They may include customer demographic information such as income, household size, education and so on. Sometimes customer's purchase behavioral information – for example, the type of payment (e.g., cash or credit cards), the day of the week that the purchase is made, etc. – is used as virtual items. Or marketing variables such as the indicator for temporary price reductions and special display are often used as virtual items.

Creating relevant virtual items definitely enriches the quality of the market basket analysis. However, it does run into the curse of dimensionality problem described earlier. Therefore, before you decide to add the virtual items to the market basket data, you should have some idea or hypothesis on how the results of analysis associated with virtual items help marketing managers to solve their decision making problems. For example, supermarkets typically select a set of items every week and discount their prices significantly – called

loss leaders – to increase the number of shoppers visiting their stores. Supermarket managers know that they lose money by selling loss leader items. But most of supermarkets employ this strategy since they expect that consumers would shop a lot of other (non-discounted) products. We can detect various issues associated with the loss leader strategy by adding the indicator of the loss leader item as virtual item.

13.4.4 Adding Temporal Component to the Market Basket Analysis

Market basket analysis was originally designed to analyze which products are purchased together at a given shopping trip. However, it can be applied to broader marketing problems if we incorporate a temporal component. This makes it more applicable to identifying cross-selling possibilities. For example, a segment of bank customers might open a savings account *after* they open checking accounts. Or customers who have purchased personal computers may tend to purchase printers within the next 3 months.

Researchers have attempted to accommodate a time-series component into market basket analysis to broaden its application domain (Agrawal and Srikant 1995; Chen et al. 1998; Ramaswamy et al. 1998). They have shown that temporal components can be incorporated into the existing association rule algorithm with minor modification. However, there is one big difference in terms of the data required. In particular, we need panel data whereby particular customers are identified and observed over time. Previously, each transaction was treated independently and there was no need to track *whose* transaction it is. To conduct a temporal analysis, a data-gathering system must track customer identification in order to relate transactions occurring at different times. For example, the traditional scanning device in the supermarket may provide transaction data with anonymous customer identity that is not appropriate for the market basket analysis with temporal component. To incorporate the temporal component, a customer identification device such as a store loyalty card is required where a cash register first scans the customer's store card and scans items purchased.

A temporal association rule can be considered a traditional association rule with some temporal relationships between items in the antecedent and the consequent. Theoretically, we need to consider all possible pairwise combinations among all transactions made by a given customer. As a result, we have all possible "before item(s)" (in the antecedent) and "after item(s)" (in the consequent) pairs. Because of this combinatorial nature of the problem, again the curse of dimensionality comes into play. For example, we need to consider $450 (= {}_{100}C_2)$ paired combinations for a customer with 100 transactions.

Table 13.2 Association rules for two items (in tabular form)

Antecedent	Consequent	Support	Confidence	Lift
Orange juice	Soap	0.40	0.67	1.67
Orange juice	Detergent	0.20	0.33	1.67
Ice cream	Beer	0.40	0.67	1.67
Ice cream	Pizza	0.20	0.33	1.67
Beer	Ice cream	0.40	1.00	1.67
Soap	Orange juice	0.40	1.00	1.67
Detergent	Orange juice	0.20	1.00	1.67
Pizza	Ice cream	0.20	0.50	1.25
Beer	Soap	0.20	0.50	1.25

An easy way to reduce the number of paired combinations is to restrict the temporal space of interest. For example, we may focus on temporal association for the "next shopping trips" where we now consider 99 pairwise comparisons for a customer with 100 transactions. Or we may restrict our attention to the transactions "within 2 months" from the transaction date of the antecedent.

13.5 Conclusion

There are several commercially available data mining software packages for performing market basket analysis. Examples are Integral Solutions' Clementine (marketed by SPSS), Silicon Graphics' MineSet, etc. that provide market basket analysis as a differentiating feature from other data mining products. Marketing managers without much statistical expertise can perform market basket analysis by clicking icons and interpreting the output without much difficulty.

Most market basket analysis software presents its output or association rules either in tabular form or in plain English. Most software allows users to specify selection criteria and sort the resulting association rules by support, lift, confidence, antecedent or consequent. Table 13.2 shows the output example in compact tabular form. We have here applied market basket analysis to the transaction data given in Table 13.1, selected the association rules with lifts greater than one, and sorted them by lift. Also note that we have limited the association rules to two items.

Some software presents results in plain English. For example, the association rules in Table 13.2 might be presented as the following:

- When a customer buys ***Orange Juice*** then the customer also buys ***Soap*** in 67% of cases. This pattern is present in 40% of transactions.[5]

[5] This paragraph means that the confidence of the association 'if orange juice then beer' is 67% and its support is 40%.

- When a customer buys ***Orange Juice*** then the customer also buys ***Detergent*** in 33% of cases. This pattern is present in 20% of transactions.
- When a customer buys ***Ice Cream*** then the customer also buys ***Beer*** in 67% of cases. This pattern is present in 40% of transactions.

Market basket analysis is an attractive data mining tool for several reasons. First, relative to other data mining tools, it is computationally simple. Second, its outputs are easy to understand because they are expressed in the form of association rules. Third, it is actionable in that it is easy for marketing managers to turn the association rules into marketing strategies and tactics.

Market basket analysis is particularly well suited to the problems without well-defined marketing objectives. You simply have a large set of data (e.g., POS transaction data from a supermarket) and you do not have specific hypothesis to test because you do not have much experience analyzing them. That is, it is a good undirected data mining technique. Market basket analysis can also be used for directed data mining tasks (Zhang 2000). But we suggest other statistically sound techniques when you have clear hypothesis to test.