ISSN 2347 - 3983
Volume 8. No. 5, May 2020
International Journal of Emerging Trends in Engineering Research
Available Online at http://www.warse.org/IJETER/static/pdf/file/ijeter108852020.pdf
https://doi.org/10.30534/ijeter/2020/108852020

Mining Negative Associations from Frequent and Regular Patterns through Application of Maximal Property

Kamma Chaitanya[1], GuntaVenkata Lakshmi Narasimha Reddy[2], Dr. JKR Sastry[3], NVSPawan Kumar[4]
[1]Koneru Lakshmaiah Education Foundation University, Vaddeswaram, chaitanyakrkl@gmail.com
[2]Koneru Lakshmaiah Education Foundation University, Vaddeswaram, narasimhareaddy.gunta@gmail.com [3]Koneru
[3]Lakshmaiah Education Foundation University, Vaddeswaram, drsastry@kluniversity.in
[4]Koneru Lakshmaiah Education Foundation University, Vaddeswaram, nvspawankumar@kluniversity.in

ABSTRACT

Knowledge hidden in databases normally is mined to find frequently occurring Item-Sets and then move on to find the positive associations that exists among those patterns. It is challenging to fix frequency threshold for finding such patterns. The spread of the item sets could be huge even though the Frequency of occurrence of the Item-sets is high, Regular occurrence of the Item sets is also very important while achieving high threshold on the frequency. Considering Both frequency and regularity may involve in generation of more item sets than expected.

Most of the focus is on finding the positive associations that exists among the frequently and regularly occurring patterns. Sometime negative associations that exeunt among the frequently and regularly occurring patterns have high significance in the medical field.

Reducing Item-sets to a minimum number will be helpful in finding the most accurate and important associations from Frequent and regular frequent Item-set. Maximal property can be applied such that minimum patterns that occur frequently and regularly that represents whole lot of patterns can be found. These pruned patterns can be used for finding effective negative associations.

This paper presents an algorithm that finds negative associations that exists among the frequent and regular patterns through application of Maximal property. The Algorithm is implemented on a sample e-commerce data and by using IBM supplied 100K data and the accuracy of finding the negative associations that exists among the Frequent and regular Item-sets is nearing 99%

Keywords : Frequent Patterns, Regular patterns, Positive associations, Negative associations, Maximal property

1. INTRODUCTION

Transactional data can be stored either in centralised database or distributed databases based on the way business is conducted centrally or at distributed locations. Data sometimes moves in streams and the data in the streams generally takes more time to be placed into a database before it can be accessed. The nature of the data patterns existing in the data base keeps changes as modifications to the databases are undertaken especially when new data records are added. Incrementing the existing data with new data alter the nature of the patterns existing in the database.

A set of data items keeps occurring again and again and appears in different transactions. Frequency of occurrence of a set of items considering different transactions that happen over a period of time differ from item set to item set. The set of data items together contained in a set is called a Pattern. There will be many such patterns in databases actually many in number. Not all the patterns are interesting

Frequent patterns (FP) are one of the fundamental knowledge discoveries recognised through transactional databases. Frequent itemset is the set of items occur more than a user given minimum support threshold. Confidence and correlation are two more parameters included to increase the reliability of the frequent itemsets.

A frequent item set (also called large item set is an item-set that meets the user-specified minimum support. Accordingly, we define infrequent item set (or small item set) as an item set that does not meet the user-specified minimum support.

However the frequent item sets relate to occurrence frequency (number of occurrences) only but not how they are spread in the distribution. Some of the patterns are sporadic and very rarely occur. Such kind of patterns is called outliers. They influence mining process to move away from real scenario. Many types of outliers exists in reality which could be categorised as global, contextual, collective and many more. The outliers must be eliminated before one arrives at frequent patterns.

Frequent Patterns play important role in knowledge discovery databases. The patterns that can be mined from a database can be classified as Basic patterns, Multi-level and

Multi-dimensional patterns and extended patterns. The Basic patterns include frequent patterns, Pattern of association rules, Closed Patterns, Maximum patterns and generator patterns. Multi-level patterns are uniform, varied patterns or patterns that depend on the support of item-sets. Multi-dimensional patterns are high dimensional patterns. Some of the patterns could be classified as continuous patterns which can be formed through the process of discretisation or through the application of statistical methods.Extend patterns include approximate patterns, uncertain patterns, compressed patterns, rare or negative patterns and colossal patterns, high dimensional-colossal patterns.

Most attention is made on finding the frequent items which focus on the frequency of occurrence of an item set. Not much focus is used on finding the regular item sets. Regular item sets are those that occur many times within a specific period of time. The periodicity could be absolute or just relative which can be measured as the distance between the transactions that contain an item set. Regularity of the item set is one of the most important aspects that have bearing on the business decisions.

There could be association between a set of patterns. The association between the patterns could be either positive or negative. In positively associated patterns, same set of business rules equally apply which means one pattern support the other. Most of the time, business managers look for positively associated patterns

The association between the frequently occurring patterns will be more interesting and important as there is an issue of interestingness associated with such patterns. In addition, the interestingness will be quite high if the issue of regularity is also considered. Positive association between regular and frequent will help taking many important decisions.

The negation of an item set A is indicated by ¬A. The support of ¬A, sup (¬A) = 1 – sup (A). In particular, for an item set i 1 ¬i 2 i3 , its support is sup (i 1 ¬i 2 i 3) = sup (i 1 i 3) –sup (i 1 i 2 i 3).We call a rule of the form A⇒ B a positive rule, and rules of the other forms (A ⇒ ¬B, ¬A ⇒ B and ¬A ⇒ ¬B) negative rules. For convenience, we often use only the form A⇒ ¬B to represent and describe negative association rules in this paper.

Like positive rules, a negative rule A ⇒ ¬B also has a measure of its strength, confidence, defined as the ratio supp(A ∪ ¬B) / sup (A). By extending the definition in [Agrawal et al. 1993b and Chen et al. 1996], negative association rule discovery seeks rules of the form A ⇒ ¬B with their support and confidence greater than, or equal to, user-specified minimum support and minimum confidence thresholds respectively, where A and B are disjoint item sets, that is, $A \cap B = \emptyset$;

—supp(A) ≥ ms, supp(B) ≥ ms and supp(A ∪ B) <ms;
—supp(A ⇒ ¬B) = supp(A ∪ ¬B);
—conf(A ⇒ ¬B) = supp(A ∪ ¬B)/supp(A) ≥ mc.

The rule A⇒ ¬B is referred to as an interesting negative rule.

Association rules are one of the precious outcomes of the pattern mining. Several algorithms have been developed to mine positive association rules. The term positive indicates the togetherness of the set of items wherever they appear in the transactional databases. These association rules are developed from positive patterns or item set. If there exists two item sets A and B such that whenever an item set A is purchased B is also purchased then we define a positive association rule A =>B.

The strength of an association rule can be measured in terms of its support "s", and confidence "c". The support "s" is the percentage of transactions in a database that contain all the elements i.e., A Ų B. The confidence "c" is the percentage of transactions in a database containing A that also contain B. The formal definitions of these measures are,

A lot of new problems may occur when we simultaneously study positive and negative association rules (PNARs), i.e., the forms A⇒B, A⇒¬B, ¬A⇒B and ¬A⇒¬B. These problems include how to discover infrequent item sets, how to generate PNARs correctly, how to solve the problem caused by a single minimum support and so on. Infrequent item sets become very important because there are many valued negative association rules (NARs) in them.

Many mining methods exist in literature for mining different types of patterns. The Mining methods that are in existence as on date can be classified into basic mining methods that include candidate generation methods that have been many in number (Apriori, partitioning sampling etc.), pattern Growth methods (FP-growth, HMine, FP max, Closttt+ etc.), Vertical format methods (Eclat, CHARM etc.)

The Mining methods can also be classified based on the interestingness which include interestingness (subjective Vs Objective), Constraint based mining, mining correlation rules, and exception rules. The mining methods can be also classified considering the way the database is organised that includes distributed data base, incremental database and streamed database.

Data mining methods greatly differ based on the type of database that must be mined. Data mining methods can be classified based on the type of data that must be mined. Various types of extended data that can be mined include sequential, time series, Structural (Tree, lattice, graph), spatial (Colocation), temporal (evolutionary and periodic0, Image Video and Multimedia and network patterns)

Mining methods can also be classified that include pattern based classification, clustering, semantic annotations, collaborative filtering and Privacy preserving. There are some approaches that have been presented in the literature for mining negative patterns. No method as such has been available that aims at mining regular frequent negative patterns which is the focus of this research. As said earlier, there are many types of databases exists and the mining methods largely vary from type of database to type of database. In this chapter mining regular, frequent negative patterns through use of vertical format which is the basic mining method has been explored and presented.

Not much attention is made on finding the negative patterns. Negative patterns are quite important even more than the positive patterns due to the kind of impact that it creates when such patterns exits. This is true in medical field, share market field, financial filed and whether forecasting sector. In medical filed two drugs having different chemicals may contradict each other. A rule that applies to a temperature zone may not be applicable to a cool zone. Thus the issue of finding the patterns that are regular, frequent and those that heave negative associations are most important which the focus of this paper. The use of maximal property to prune the frequent and Regular Item set is the most focussed aspect of this paper.

2. PROBLEM DEFINITION

Thus problem is to find from frequent and regular pattern, the patterns that have negativewhilemaximising the Patterns using Maximal property.

3.RELATED WORK

[Ming-Syan Chen 1996] [1] have presented a comprehensive survey on the availability of different mining methods and the purpose for which the mining methods are used. A purposed based classification of the mining methods and the purpose for which the mining methods are used has been presented.

A basic method to mine transactional databases leads to mining too many patterns that will reflect into too many association rules that were not quite interesting to the end user. [Ashok Savasere 1998] [2] Presented a mining method that combines positive associations with domain knowledge so that very few negative associations can be found which can be easily evaluated and presented.

[BalajiPadmanabhan 1998] [3]Have presented that pattern mining generally leads to too many patterns and do not take into account the domain knowledge that the decision makers have. Decision makers have prior knowledge about the data in terms of precepts and beliefs. They have also presented a method that mines the unexpected patterns through use of beliefs and perceptions and have experimented the method WEB log files and proved that efficient mining can be done using the user's perception.

Many mining methods existing in the literature have used the concept of candidate get generation approaches which actually uses Apriori like method. This approach has been proved to be time consuming and costly especially when many long patterns are involved. [JiaweHan 2000] [4] Have proposed a novel method which uses frequent pattern tree (FP-tree) which is an extension of Prefix-structure. The FP-tree structure is used for storing crucial information about the frequent patterns and the information is used for pattern mining. The method used the concept of FP-growth for mining complete set of frequent patterns. They have proposed by used three main techniques that include database compression, a pattern fragment growth method and divide and conquer method for decomposing mining tasks into small number of tasks which can carry the mining considering the constraints attached to the small tasks. The three methods used by them reduce the search space dramatically. Pattern mining can be undertaken either through horizontal or vertical mining approaches. The Vertical mining approaches have been found to be quite effective when compared to horizontal approaches. Fast frequency counting through intersecting operations on transactions IDs and pruning of the patterns that are irrelevant are the main advantages of vertical format methods. However these methods suffer from lack of memory when the when the entries to be made into vertical format table is too heavy.

[Mohammed J. Zaki, 2003] [5] have presented a novel way of presenting the vertical data called Di-Set which considers the differences between the transactions of a candidate patterns the very patterns itself. They have shown how Di-sets can drastically cut down the memory requirement for storing the vertical table entries,

In transactional data bases many patterns exist that can be used to generate both positive and negative association rules. A method has been proposed by [XINDONG WU 2004] [6] that can be used for generating both negative and positive association rules. The negative associations between the patterns can be evaluated through checking the expressions like $A \Rightarrow \neg B$, $\neg A \Rightarrow B$ and $\neg A \Rightarrow \neg B$. The rules can be mined from a large database through constraining the patterns using the interesting patterns.

Some of the association rules that can be generated could be exceptional in the sense that the rules are less interested or have too high a confidence. [Daly et al., 2004] [7] Have presented a method for mining exceptional rules that can be evaluated. They have considered the relationships between the exceptional rules and negative association rules. Exception rules are generated based on the knowledge gained through negative association rules. They have also defined a new measure that can be used to evaluate the interestingness of the exceptional rules. The exceptional rules that meet the exceptional measures are the candidate's exceptional rules that can be used for evaluation of the patterns and decision making.

Most of the methods proposed in the literature use the interestingness measures to prune the most wanted patterns for decision making. However precisely defining the interestingness measure is quite complicated and sometimes must be found using the trial and error method. There is no as such exact method that can be used for determining the interesting measures. A method that uses the inputs provided by the user in terms of number of rules that they user requires and the kind of constraints / interestingness that must be satisfied has been presented by [DR Thiruvady2004] [8]. AN algorithm called GRD that discovers M-most interesting rules has been has been presented.

Correlations are statistical measures that find how good a set of data records are related another set of records. A method that finds negative association rules which is based on the kind of existence correlation between two item sets has been presented by [Maria-Luiza, Antonie 2004] [9]. Negative rules between the item sets can be extracted if the correlation between the Item sets is negative and that the confidence of the items sets is quite high. The extracted negative association rules can have either consequents or antecedents (- Y¿ and $\neg Y$¿) even when the computed support value from the item sets is less than the threshold value of the support. The algorithm presented by them generates all the negative and positive. The correlation between the item sets could be positive when the interesting measures hold good the interestingness measures that include support and confidence. If the support is less and confidence is high then the negative correlation exists between the item sets leading to generation of negative association rules that negates both the antecedents and consequents.The method proposed by them] generates all negative and positive associations rules out of the patterns which have strong association between them. If no association rules can be generated, the threshold Value of correlation will have to be lowered, thereby reducing the strength of correlations between the items set.

A survey has been presented by [Chris Cornelis 2006] [10] citing several algorithms that mine both negative and positive association rules and have described several situations wherein the algorithms presented in the literature could not satisfy certain situations. They have classified and catalogued several mining algorithms based on some parameters and could figure out the drawbacks of each of the algorithms. They have also presented a modified mining algorithm based on Apriori approach that can find both negative associations with interesting attached through confidence framework. They have used upward closure property that confirms to the support based interestingness of negative associations under validity definitions.

Usually the interestingness parameter “Support” is defined for enter dataset. The data records cloud be recognised a hiercahy of records having occurrence of set of records at each level having a specific support value. Several support values can be defined at each level. A model is proposed by [Xiangjun Dong 2006] [11] called MLMS (Multi Level minimum support) that considers defining minimum support value at each of the level of the records. MLMS is used to discover both frequent and infrequent itemsets. They have considered both correlation and confidence interesting measures and proposed yet another interesting measure to mine the both frequent and infrequent itemsets. An algorithm called PNAR-MLMS has been proposed that can be used to generate both positively and negatively associated patterns from frequent and infrequent itemsets that are generated through MNMS model.

[Xiangjun Dong 2007-1] [12] have also developed PNAR based Classifiers using which the association rules can be classified into some known categories. The classifiers then can be used to find whether a pattern leads to negative or positive association. Discovering K-Most intersecting rule requires the minimum support value that should be used as threshold value. It is rather difficult to define minimum thresh hold value as the user has no real idea about the support value. Rather the users can define interestingness and the number of rules that the users expects from the mining system.

Another method called GRD which does not require minimum support value has also be presented in the literature. It mealy requires the user to define the measure of interestingness and the number of rules in which the user is interested in. [Xiangjun Dong 2007-2] [13] have extended the GRD method which can be used form for mining positive and negative rules.

Both positive and negative association rules can be mined through transactions. Negative association rules explain how one pattern negates some other patterns. Many applications exists that needs the mining of the negative association rules especially the negative association rules will help in carrying the market-basket analysis.

The negative association rules can be used to develop classifiers as well using which classification models can also be built. Mining negative association rules requires exploration of large data space. Many of the algorithms proposed in the literature are not being used due the reason that large data spaces have to be explored.

[Xiangjun Dong 2007-3] [14] Have extended the support confidence frame work through addition of correlation coefficient threshold that keeps sliding as the data accessing keeps moving. Essentially they have used correlation coefficients that can be calculated considering different patterns. The patterns that are positively correlated and the patterns that are negatively correlated can be found form antecedents and consequents..

Most of the work is focussed on frequent items till 2009 and no focus is made till such time on regular items which are those that occur in regular intervals. Regular items are the most important than the frequent items, the occurrence of which has no time limitation. **[Tanbeer 2008] [15]**are the first group of authors who have focussed on regular itemsets. They have proposed a “Regular pattern tree” which is a tree structure to discover regular patterns. The algorithm scans the database twice. In the first sct regularity and support values of the item sets is determined and in the scond scan an regular Pattern tree is constructed. The process adopted by them is similar to cyclic and periodic patterns.

In many transactional databases, data is hidden in sequence as a structure. Bio Technology based sequences are hidden in medical related databases. Mining sequential patterns reveal many interesting facts that when evaluated yield important decisions. Generally the sequential patterns are mined using defined minimum support threshold defined by the users. Use of the minimum support threshold assumes that all frequent sequences have the same frequency which is actually not the case in real world. If the frequencies of the pattern sequences vary greatly even though they meet the minimum threshold value, then rare item problem arises**.**

Mining negative associations is as important as mining positive associations among the frequent patterns. [Idheba Mohamad Ali 2012] [16] Have presented new models which can be used for mining interesting negative and positive associations among transactional data records. They have considered the merging of two algorithms that include mining interesting negative and positive association rules (PNAR) and mining interesting multiple level supports algorithm (IMLMS). The algorithm proposed by them helps mining positive and negative association rules from interesting frequent and in-frequent item sets using multiple support values.

[NVSPavan 2012] [17] [NVSPavan 2012] [18] have presented a method of finding both positive and negative association considering the regularity of the item set using vertical Table mining method. [NVSPavan 2019] [20] [NVSPavan 2019] [21] [NVSPavan 2019] [22] [NVSPavan 2019] [23] have presented methods that can be used for mining negative associations from frequent and Regular patterns that exists within fixed databases, Incremental database, distributed data bases and Data streams finding using vertical Table mining method.

Many other mining techniques have been presented considering different purpose for undertaking Patten mining and non-have considered the issue of Maximal property [24] [25] [26][27][28][29][30][31][32][33] and also the issues of regularity and primarily focussing on the issue of negative associations.

4. COMPARATIVE ANALYSIS – FINDING ASSOCIATIONS

Comparison of exiting algorithms has been done considering various as aspects that must be considered for generating negative associations considering regular and frequent item sets in order to assess the adequacy of those algorithms. **Table 1** shows the comparison, from the table it can be seen that none of the existing algorithms are dealing with the most important aspects of the negative associations that include regularity, positive/negative associations, frequency and interestingness measures.

5. INVESTIGATIONS AND FINDINGS

5.1 Algorithm

1. Read the support value that dictates the threshold value of the frequency of the patterns and also the regularity as defined by the user. Read the data in flat file / DBMS Table into an Array as shown in Table 2
2. Convert the data in Table 2 into vertical format as shown in Table 3
3. Prune the Initial Irregular Items and non-frequent items
4. Apply the maximal property and add all those item sets that are connected with the considered item sets either directly or indirectly considering complete closure of the patterns
5. Repeat the following process

Consider the current Item

Select next item and prune it if is not regular or frequent and go to next item (self-Loop).

If the next item is regular and frequent, get the intersection of the transactions of the current item and next item

If the intersection is null, then enter the current and next items into a negative set array

If the intersection is not null, then get the common elements and see if the count of elements is > regularity threshold decided by the user

If the common elements satisfy the regularity and frequency constraint, then add the common elements into vertical table as a new row as they are regular and frequent at the end of the vertical table.

If the common elements do not satisfy the regularity or frequency constraint, then ignore them

If all the elements in the vertical table are exhausted, then convert the next item which is next to current item as current item and then LOOP

If all the elements in the vertical table are not exhausted, then move to the next item and LOOP When no item is left in the vertical table, then the process terminates

6. Find the Patterns from the negative pattern list

5.2 Experimentation with sample data and results

The proposed algorithm has been applied on sample data extracted from IBM supplied 100K data. Following is the sequence of steps executed

Step-1

Transaction IDS are added to the Extracted data. Sample first 20 records of IBM supplied data are shown in Table 2. Here frequency is the count of transactions in which the item set appears

Table 2: Sample Transaction data

TrId	Item Set
1	I1 I2 I3 I4 I5 I9 I10 I14
2	I4 I5 I6 I10 I15
3	I2 I3 I7 I13 I14 I15
4	I5 I8 I10 I11 I15
5	I1 I3 I5 I6 I9
6	I4 I5 I6 I15
7	I2 I3 I7 I11 I12 I13
8	I5 I8 I11 I12 I14 I15
9	I1 I3 I5 I8 I9
10	I4 I5 I6 I10 I15
11	I2 I3 I7 I8 I13 I14 I15
12	I5 I8 I11 I15
13	I1 I3 I5 I9 I11
14	I4 I5 I6 I14 I15
15	I2 I3 I6 I7 I12 I13
16	I5 I8 I11 I12 I14 I15
17	I1 I3 I5 I6 I9 I10
18	I4 I5 I6 I12 I14 I15
19	I2 I3 I4 I7 I13
20	I5 I8 I11 I12 I15
21	I1 I3 I5 I9 I14

Step-2

Convert the Transaction filled IBM data to Vertical format. The vertical format data is shown Table 3. In the vertical data format, for each of the Item, in the data repository, the transactions that contain the Items are found and mapped. The frequency of an item is the count of transactions in which the item appears.

Table 3: Transaction Data in vertical format

Item Code	TrId	Frequency
I1	1 5 9 13 17 21	6
I2	1 3 7 11 15 19	6
I3	1 3 5 7 9 11 13 15 17 19 21	11
I4	1 2 6 10 14 18 19	7
I5	1 2 4 5 6 8 9 10 12 13 14 16 17 18 20 21	16
I6	2 5 6 10 14 15 17 18	8
I7	3 7 11 15 19	5
I8	4 8 9 11 12 16 20	7
I9	1 5 9 13 17 21	6
I10	1 2 4 10 17	5
I11	4 7 8 12 13 16 20	7
I12	7 8 15 16 18 20	6
I13	3 7 11 15 19	5
I14	1 3 8 11 14 16 18 21	8
I15	2 3 4 6 8 10 12 14 16 18 20	11

Step-3

Find the first regular item by pruning all the previous items whose regularity is > User given Maximum Regularity threshold (γ_{min_reg}) and also the item that satisfies the minimum support value. Here regularity implies the relative occurrence of the Item, computed as the distance between two successive transactions. Consider $(\gamma_{min\_reg}) = 4$ and the minimum frequency be 3 for sample data. The First regular and frequent Item is called Previous-Item. The list of items that will be left over in the vertical format table is shown in Table 4

Step-4

Consider each item starting from Previous-item and repeat the following procedure.

1. Consider the next item and let that be current-item
2. Find if the current-item is regular and frequent. If the current-item is not regular or frequent prune it.
3. If the current- item is regular and frequent, find Intersection of the transactions of the current item with the previous-item.
4. If the intersection is null, then add the Item set into negative item-set list.
5. If the intersection is not null, find the regularity considering the common elements.
6. If the regularity is $<(\gamma_{min\_reg}),$ then add previous item and the current item set along with its related transaction as an additional record to the vertical database since they are regularly and frequently associated.
7. If the next item is not the lost entry in the vertical table, Make Current Item as the next Item and loop.
8. If the next item is the last item in the Vertical table then Previous Item = Previous Item +1 and then Loop.

Step-5

Find the Maximum Item sets from negatively associated patterns

After this step is completed, the pruned items are shown in Table 5 and the negatively associated items are shows in Table 6 and positively associated items are shown in Table 7 and the Maximum Item sets are in Table 8

Table 7: Negatively Associated Item set

Item Set-1	Item-set 2
1	7
1	13
1	15
1	7,13
4	8
4	11

Item Set-1	Item-set 2
4	8, 11
5	7
5	7,13
6	8
6	8
6	11
6	8, 11
7	9
7	1,5
7	1,9
7	5,9
7	1,5,9
8	4,6
9	13
9	15
9	7,13
11	4,6
13	1,5
13	1,9
13	1,5,9
15	1,9
1,5	7,13
1,9	7,13
4,6	8,11
5,9	7,13
1,5,9	7,13

Table 8: Maximum negatively associated Item sets

Item Set-1	Item-set 2
7	1,5,9
13	1,5,9
1,5	7,13
1,9	7,13
4,6	8,11
5,9	7,13
1,5,9	7,13

5.3 Experimentation with IBM data and results

The above mentioned algorithm has been applied on to the IBM supplied data. A sample of 20,000 records has been selected and the regularity of the item set is fixed at 350. List of base data supplied by IBM, the vertical format of the same and the negatively associated Item sets are shown in Tables 9, Table 10 and table 11. Table 12 Shows the Maximal Negative patterns / Item sets

Table 11: Negatively Associated Item set

Itemst1	Itemset2
12	569
12	975
38	205
38	862
38	885
38	975
38	998
39	631
54	975
71	569
72	285
72	614
72	919
132	593
132	620
132	798
151	183
151	631
177	975
236	523

6.DATA ANALYSIS AND INTERPRETATIONS

The IBM supplied data has been analysed for different sizes of the samples drawn in terms of 40,000, 60,000 and 80,000 records. The data is analysed with different regularity percentages. The number of negative frequent regular patterns, for different maximum regularity and minimum frequency is shown in Table12

Table 12: Analysis of Negative Frequent Regular Itemsets with TD 100K Transaction data

Total Transactions	%Max Regularity	%Support Count	Number of Negative Frequent Regular Itemset
40000	3.00	2.000	6
	3.00	1.750	42
	3.00	1.625	154
	3.00	1.500	461
40000	2.50	1.750	41
	2.50	1.625	154
	2.50	1.500	352
	2.50	1.125	981
40000	2.00	1.750	35
	2.00	1.625	118
	2.00	1.500	181
40000	1.50	1.750	2
	1.50	1.625	3
	1.50	1.500	3

Total Transactions	%Max Regularity	%Support Count	Number of Negative Frequent Regular Itemset
60000	1.65	1.650	13
	1.65	1,250	41
	1.65	1.000	150
	1.65	1.000	352
60000	1.35	1.650	35
	1.35	1.350	118
	1.35	1.000	181
80000	1.00	0.875	35
	1.00	0.815	118
	1.00	0.750	181

Figure 1, refers to the number of negative frequent regular and negative regular items in relation to % regularity. The numbers of frequent regular items sets are narrowing down compared to regular frequent sets as the % of regularity and support count increases.

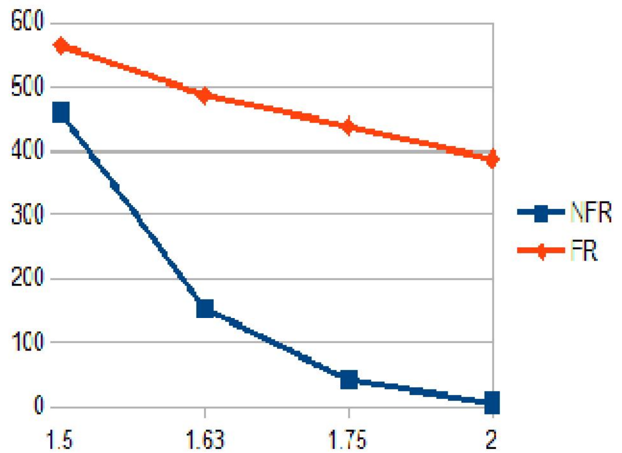

Figure 1: Negative Regular and Negative Frequent Regular with 40000 transactions at different Regularity and Frequency values.

Figure 2, Figure 3, Figure 4 shows the relationship between the negative frequent regular item sets and frequency plotted on the x-axis for different regularity %. It could be noted from the figures that as the frequency increases the number of item sets decreases.

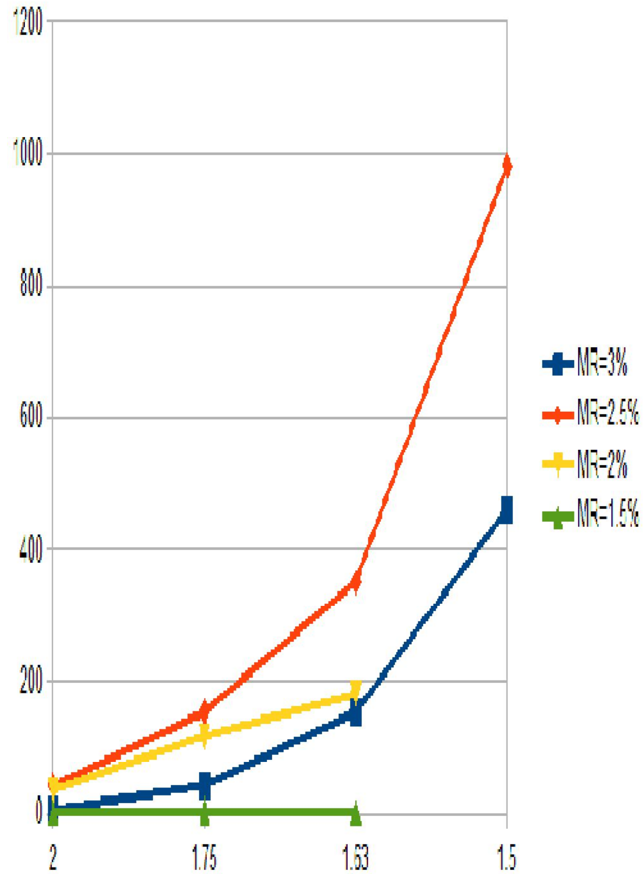

Figure 2: Negative Frequent Regular Itemset with 60000Transactions at different Regularity and Frequency values

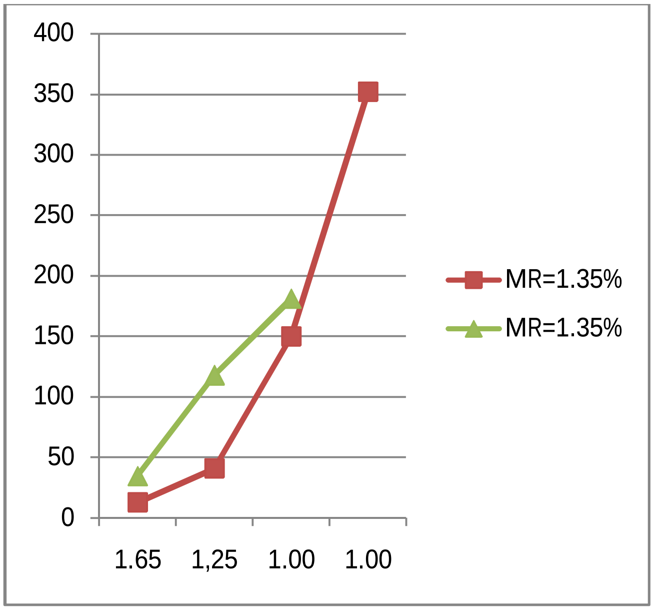

Figure 3: Negative Frequent Regular Itemset with 60000Transactions at different Regularity and Frequency values

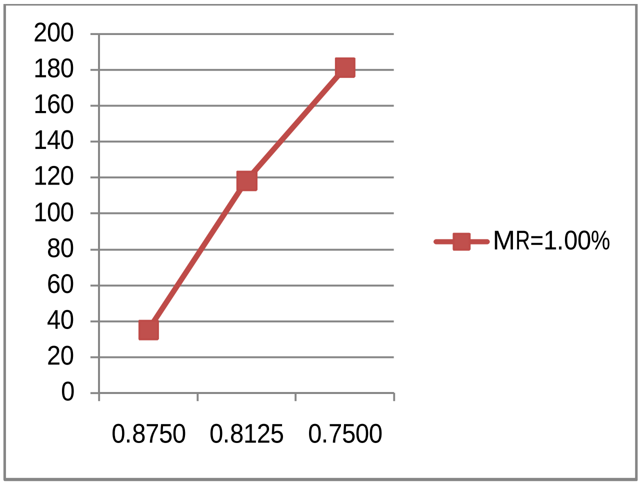

Table 4: Negative Frequent Regular Itemset with 80000 Transactions at different Regularity and Frequency values

6. CONCLUSION

Many type of databases exists which differ in the kind of data stored. Transactional data could be stored either as flat files or relational data. Patterns are the knowledge hidden in any database. A set of data items that frequently occurs in different records in a database is called as patterns. Knowledge of patterns is very important as one can take excellent decisions by analysing and visualising the patterns

A set of patterns can be either frequent or regular. Frequency of patterns can be computed considering the entire database. Regularity is a kind of distance between to like patterns. The frequency of the patterns within Regularity is also equally important.

Most concentration as on date is finding the frequent patterns which are positively associated. Regularity is also most important in addition to the frequency. Many negatively associated patterns exist in nature without which the ill effects of patterns on the business cannot be assessed and corrective actions taken.

Lot of importance for the negatively associated patterns exists in the Medical and atmospheric studies fields. Regularly negative, regularly frequent and negative, regularly frequent maximal and negative patterns are to be mined and analysed so that analysis can be made and proper decisions are taken based on those patterns. It is to be noted that the frequent patterns may not be regular and Vice versa.

Maximal property will help considering all the patterns which are connected either directly or indirectly so that all relavant patterns can be found. The more the patterns, the most important hidden negative associations can be found.

REFERENCES

1. Ming-Syan Chen, Jiawei Han, Philip S. Yu, Datamining an overview from a database perspective, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL 8, NO 6, DECEMBER 1996
2. Ashok SAVASERE, A., OMIECINSKI, E., AND NAVATHE, S. 1998. Mining for strong negative associations in a large database of customer transactions. In Proceedings of the Fourteenth International Conference on Data Engineering. IEEE Computer Society, Orlando, Florida, 494–502.
3. BalagiPADMANABHAN, B. AND TUZHILIN, A. 1998. A belief-driven method for discovering unexpected patterns. In Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98). AAAI, Newport Beach, California, USA, 94–100.
4. Jaiwe Han, Yin, Y. Yin, “Mining Frequent Patterns without candidate generation”, In Proc. ACM SIGMOD international Conference on management of Data, PP. 1-12 (2000). https://doi.org/10.1145/335191.335372
5. Mohammed J. Zaki, Fast Vertical Mining Using Diffsets, *SIGKDD*2003, ACM Copyright 2003
6. XINDONG Wu, C. Zhang, and S. Zhang, "Efficient Mining of both Positive and Negative Association Rules," ACM Transactions on Information Systems, 22, 2004, pp. 381-405. https://doi.org/10.1145/1010614.1010616
7. O. Daly and D. Taniar, ”Exception Rules Mining Based On Negative Association Rules”, Lecture Notes in Computer Science, Vol. 3046, 2004, pp 543–552
8. D.R. Thiruvady and G.I. Webb, ”Mining Negative Association Rules Using GRD”, Proc. Pacific-Asia Conf. on Advances in Knowledge Discovery and Data Mining, 2004, pp 161–165. https://doi.org/10.1007/978-3-540-24775-3_20
9. Maria LuizaAntonie, O. Zaiane, "Mining Positive and Negative Association Rules An Approach for Confined Rules," Proceedings of 8th European Conference on Principles and Practice of Knowledge Discovery in Databases(PKDD04), LNCS 3202, Springer-Verlag Berlin Heidelberg, Pisa, Italy, 2004, pp.27-38.
10. Chris Cornelis, Ghent, Peng Yan, Xing Zhang, Guoqing Chen, Mining Positive and Negative Association Rules from Large Databases, IEEE 2006 https://doi.org/10.1109/ICCIS.2006.252251
11. Xiangjun Dong, Sun, F., Han, X., Hou, R.: Study of Positive and Negative Association Rules Based on Multi-confidence and Chi-Squared Test. In: Li, X., Zaïane, O.R., Li, Z. (eds.) ADMA 2006. LNCS (LNAI), vol. 4093, pp. 100–109. Springer, Heidelberg (2006)
12. Xiangjun Dong, Z Niu, X Shi, X Zhang, D. Zhu, "Mining Both Positive and Negative Association Rules from Frequent and infrequent Itemsets," Proceedings of the Third International Conference on Advanced Data Mining and Applications (ADMA 2007), Harbin, 519 China, August 6-8, 2007. Lecture Notes in Computer Science 4632, Springer 2007, pp. 122-133. https://doi.org/10.1007/978-3-540-73871-8_13
13. Xiangjun Dong, Zheng, Z., Niu, Z., Jia, Q.: Mining Infrequent Item sets based on Multiple Level Minimum Supports. In: ICICIC07. Proceedings of the Second International Conference on Innovative Computing,

Information and Control, Kumamoto, Japan, September 2007
14. Xiangjun Dong, ZhendongNiu, Donghua Zhu, Zhiyun Zheng[4], and QiutingJia, Mining Interesting Infrequent and Frequent Itemsets Based on MLMS Model, C. Tang et al. (Eds.): ADMA 2008, LNAI 5139, pp. 444–451, 2008
15. TANBEER Syed Khairuzzaman, Chowdary Farhan Ahmed, Byeong-Soo Jeong, Young-Koo Lee, Mining regular patterns in transcational databases, IEICETranscations, INF, & SYST., Vol. E91-D, Issu. 11, 2008
16. Weimin Ouyang, Qinhua Huang, Mining Positive and Negative Sequential Patterns with Multiple Minimum Supports in Large Transaction Databases, 2010 Second WRI Global Congress on Intelligent Systems
17. Idheba Mohamad Ali O. Swesi, Azuraliza Abu Bakar, AnisSuhailis Abdul Kadir, Mining Positive and Negative Association Rules from Interesting Frequent and Infrequent Itemsets, 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2012)
18. N.V.S. Pavan Kumar, K Rajasekhara Rao, Mining Positive and Negative Regular Item-Sets using Vertical Databases, DOI 10.5013/IJSSST.a.17.32.33 33.1 ISSN: 1473-804x online, 1473-8031
19. N.V.S.Pavan Kumar, L. Jaga Jeevan Rao, G. Vijay Kumar, **A Study on Positive and Negative Association rule mining,** International Journal of Engineering Research & Technology (IJERT), Vol. 1 Issue 6, August – 2012
20. NVSPavan Kumar, Dr.JKRSastry, Dr. K Raja Sekhara Rao, Mining Negative Associations between Regular and Frequent Patterns hidden in Static Databases, International Journal of Emerging Trends in Engineering Research, 2019, Volume 7, Issue 7, pp. 54-67 https://doi.org/10.30534/ijeter/2019/04772019
21. NVSPavan Kumar, Dr.JKRSastry, Dr. K Raja Sekhara Rao, Mining Distributed Databases for Negative Associations from Regular and Frequent Patterns, International Journal of Advanced trends in computer science and engineering, 2019, Volume 8, Issue 4, pp. 1449-1463 https://doi.org/10.30534/ijatcse/2019/64842019
22. NVSPavan Kumar, Dr.JKRSastry, Dr. K Raja Sekhara Rao, Mining Negative Frequent regular Itemsets from Data Streams, International Journal of Emerging Trends in Engineering Research, 2019, Volume 7, Issue 8, pp. 85-98 https://doi.org/10.30534/ijeter/2019/02782019
23. NVSPavan Kumar, Dr.JKRSastry, Dr. K Raja Sekhara Rao, On mining Incremental Databases for Regular and Frequent Patterns, International Journal of Emerging Trends in Engineering Research, 2019, Volume 7, Issue 9 pp. 291-305 https://doi.org/10.30534/ijeter/2019/12792019
24. JKRSastry, M TrinathBasu, Securing SAAS service under cloud computing-based multi-tenancy systems, Indonesian Journal of Electrical Engineering and Computer Science, Volume 13, Issue 1, Page 65-71, 2019
25. JKRSastry, M TrinathBasu, Securing Multi-tenancy systems through multi DB instances and multiple databases on different physical servers, International Journal of Electrical and Computer Engineering (IJECE), Volume 9, Issue 2, Pages 1385-1392, 2019 https://doi.org/10.11591/ijece.v9i2.pp1385-1392
26. M. TrinathBasu, Dr.JKRSastry, A full security included Cloud Computing architecture, International Journal of Engineering & Technology, Volume 7, Issue 2.7, Page 807-812, 2018
27. JKRSastry, M TrinathBasu, Securing Multi-tenancy systems through user spaces defined within the database level, Jour of Adv Research in Dynamical & Control Systems, Volume 10, issue 7, Page 405-412, 2018
28. J. K. R. Sastry, K. Sai Abhigna, R. Samuel and D. B. K. Kamesh, Architectural models for fault tolerance within clouds at the infrastructure level, ARPN Journal of Engineering and Applied Sciences, VOL. 12, NO. 11, 2017, Pages 3463-3469,
29. DBK Kamesh, JKRSastry, Ch. Devi Anusha, P. Padmini, G. Siva Anjaneyulu, Building Fault Tolerance within Clouds at Network Level, International Journal of Electrical and Computer Engineering (IJECE), Vol. 6, No. 4, pp. 1560~1569, 2016 https://doi.org/10.11591/ijece.v6i4.10676
30. S. L. SUSHMITHA, Dr. D. B. K. J.K. R. SASTRY, V. V. N. SRI RAVALI, Y.SAI KRISHNA REDDY, building fault tolerance within clouds for providing uninterrupted software as service, Journal of Theoretical and Applied Information Technology, Vol.88. No.1, Pages 65-76, 2016
31. M. TrinathBasu, JKRSastry, Improving the Open Stack Authentication system through federation with JASON Tokens, International Journal of Advanced Trends n Computer Science and Engineering, 3596-3614,2019. https://doi.org/10.30534/ijatcse/2019/143862019
32. JKRSastry, M TrinathBasu, Multi-Factor Authentication through Integration with IMS System, International Journal of Emerging Trends in Engineering Research, Volume 8, Issue 1, 2020, PP. 87-113 https://doi.org/10.30534/ijeter/2020/14812020
33. JKRSastry, M TrinathBasu, Strengthening Authentication within OpenStack Cloud Computing System through Federation with ADDS System International Journal of Emerging Trends in Engineering Research, Volume 8, Issue 1, 2020, PP. 213-238 https://doi.org/10.30534/ijeter/2020/29812020
34. JKRSastry, M TrinathBasu, Enhancing Data Security under Multi-Tenancy within Open Stac, International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, Issue 1, 2020, PP. 533-544 https://doi.org/10.30534/ijatcse/2020/73912020

Table 1: Comparative Analysis of Pattern finding Algorithms

Algorithm Serial Number	Main Author	Interestingness measures					Occurrence Behaviour					Type of Associations		Extension to Mining technique	Use of domain Knowledge
		Support	Confidence	Correlation	Multi support	Multi Correlation	Regularity	Irregularity/ Rare	Frequent	Maximal	Unexpected	Positive Associations	Negative Associations		
1	Ashok Savasere	√										√			√
2	BalajiPadmanabhan	√									√	√			√
3	Jiawe Han 2000	√							√			√		FP Tree	
4	J. Zaki	√							√			√		DI-SET	
5	XINDONG WU	√							√			√	√		
6	Daly	√							√				√	Exception rule Mining	
7	DR Thiruvady	√											√		√
8	Maria-Luiza, Antonie			√					√			√	√		
9	Xiangjun Dong			√	√							√	√		
10	**Tanbeer**						√								
11	**Weimin Ouyang**				√									Sequential Mining	
12	Idheba Mohamad Ali				√		√					√	√		
13	PavanNVS	√					√	√	√	√	√	√	√	Veridical Tab	

Table 4: Pruning the Items in the vertical table till the first regular and frequent Item is traced

Item Code	TrId	Regularity (Periods)	Maximum Regularity of the Item	Frequency of the Item
I1	1 5 9 13 17 21	1 4 4 4 4	4	6
I2	1 3 7 11 15 19	1 4 4 4 4 2	4	6
I3	1 3 5 7 9 11 13 15 17 19 21	1 2 2 2 2 2 2 2 2 2 2	2	11
I4	1 2 6 10 14 18 19	1 1 4 4 4 4 1 2	4	7
I5	1 2 4 5 6 8 9 10 12 13 14 16 17 18	1 1 2 1 1 2 1 1 2 1 1 2 1 1 2 1	2	16
I6	2 5 6 10 14 15 17 18	2 3 1 4 4 1 2 1 3	4	8
I8	4 8 9 11 12 16 20	4 4 1 2 1 4 4 1	4	7
I9	1 5 9 13 17 21	1 4 4 4 4 4	4	6
I11	4 7 8 12 13 16 20	4 3 1 4 1 3 4 1	4	7
I15	2 3 4 6 8 10 12 14 16 18 20	2 1 1 2 2 2 2 2 2 2 1	2	11

Table 5: List of Pruned item list

Item Code	TrId	Regularity (Periods)	Maximum Regularity of the Item	Frequency of the item set
I7	3 7 11 15 19	3 4 4 4 4 2	4	5
I10	1 2 4 10 17	1 1 2 6 7 4	7	5
I12	7 8 15 16 18 20	7 1 7 1 2 2 1	7	6
I13	3 7 11 15 19	3 4 4 4 4 2	4	5
I14	1 3 8 11 14 16 18 21	1 2 5 3 3 2 2 3	5	8

Table 6: Positively Associated Item set

Itemset	Trids	Periods	Max Regularity	Support Count
1	1 5 9 13 17 21	4 4 4 4 4	4	6
2	1 3 7 11 15 19	2 4 4 4 4	4	6
3	1 3 5 7 9 11 13 15 17 19 21	2 2 2 2 2 2 2 2 2 2	2	11
4	1 2 6 10 14 18 19	1 4 4 4 4 1	4	7
5	1 2 4 5 6 8 9 10 12 13 14 16 17 18 20 21	1 2 1 1 2 1 1 2 1 1 2 1 1 2 1	2	16
6	2 5 6 10 14 15 17 18	3 1 4 4 1 2 1	4	8
7	3 7 11 15 19	4 4 4 4	4	5
8	4 8 9 11 12 16 20	4 1 2 1 4 4	4	7
9	1 5 9 13 17 21	4 4 4 4 4	4	6
11	4 7 8 12 13 16 20	3 1 4 1 3 4	4	7
13	3 7 11 15 19	4 4 4 4	4	5
15	2 3 4 6 8 10 12 14 16 18 20	1 1 2 2 2 2 2 2 2 2	2	11
1,3	1 5 9 13 17 21	4 4 4 4 4	4	6
1,5	1 5 9 13 17 21	4 4 4 4 4	4	6
1,9	1 5 9 13 17 21	4 4 4 4 4	4	6
2,3	1 3 7 11 15 19	2 4 4 4 4	4	6
2,7	3 7 11 15 19	4 4 4 4	4	5
2,13	3 7 11 15 19	4 4 4 4	4	5
3,5	1 5 9 13 17 21	4 4 4 4 4	4	6
3,7	3 7 11 15 19	4 4 4 4	4	5
3,9	1 5 9 13 17 21	4 4 4 4 4	4	6
3,13	3 7 11 15 19	4 4 4 4	4	5
4,5	1 2 6 10 14 18	1 4 4 4 4	4	6
4,6	2 6 10 14 18	4 4 4 4	4	5
4,15	2 6 10 14 18	4 4 4 4	4	5
5,6	2 5 6 10 14 17 18	3 1 4 4 3 1	4	7
5,8	4 8 9 12 16 20	4 1 3 4 4	4	6
5,9	1 5 9 13 17 21	4 4 4 4 4	4	6
5,11	4 8 12 13 16 20	4 4 1 3 4	4	6
5,15	2 4 6 8 10 12 14 16 18 20	2 2 2 2 2 2 2 2 2	2	10
6,15	2 6 10 14 18	4 4 4 4	4	5
7,13	3 7 11 15 19	4 4 4 4	4	5
8,11	4 8 12 16 20	4 4 4 4	4	5
8,15	4 8 12 16 20	4 4 4 4	4	5
1,3,5	1 5 9 13 17 21	4 4 4 4 4	4	6
1,3,9	1 5 9 13 17 21	4 4 4 4 4	4	6
2,3,7	3 7 11 15 19	4 4 4 4	4	5
2,3,13	3 7 11 15 19	4 4 4 4	4	5
11,15	4 8 12 16 20	4 4 4 4	4	5
1,3,5	1 5 9 13 17 21	4 4 4 4 4	4	6
1,5,9	1 5 9 13 17 21	4 4 4 4 4	4	6
3,5,9	1 5 9 13 17 21	4 4 4 4 4	4	6
4,5,6	2 6 10 14 18	4 4 4 4	4	5
4,5,15	2 6 10 14 18	4 4 4 4	4	5
4,6,15	2 6 10 14 18	4 4 4 4	4	5
5,6,15	2 6 10 14 18	4 4 4 4	4	5
2,3,7	3 7 11 15 19	4 4 4 4	4	5
2,7,13	3 7 11 15 19	4 4 4 4	4	5

Itemset	Trids	Periods	Max Regularity	Support Count
5,8,11	4 8 12 16 20	4 4 4 4	4	5
5,8,15	4 8 12 16 20	4 4 4 4	4	5
1,3,9	1 5 9 13 17 21	4 4 4 4 4	4	6
1,5,9	1 5 9 13 17 21	4 4 4 4 4	4	6
3,5,9	1 5 9 13 17 21	4 4 4 4 4	4	6
5,8,11	4 8 12 16 20	4 4 4 4	4	5
5,11,15	4 8 12 16 20	4 4 4 4	4	5
8,11,15	4 8 12 16 20	4 4 4 4	4	5
2,3,13	3 7 11 15 19	4 4 4 4	4	5
2,7,13	3 7 11 15 19	4 4 4 4	4	5
3,7,13	3 7 11 15 19	4 4 4 4	4	5
4,5,15	2 6 10 14 18	4 4 4 4	4	5
4,6,15	2 6 10 14 18	4 4 4 4	4	5
5,6,15	2 6 10 14 18	4 4 4 4	4	5
5,8,15	4 8 12 16 20	4 4 4 4	4	5
5,11,15	4 8 12 16 20	4 4 4 4	4	5
8,11,15	4 8 12 16 20	4 4 4 4	4	5
4,5,6,15	2 6 10 14 18	4 4 4 4	4	5
2,3,7,13	3 7 11 15 19	4 4 4 4	4	5
5,8,11,15	4 8 12 16 20	4 4 4 4	4	5
2,3,7,13	3 7 11 15 19	4 4 4 4	4	5
4,5,6,15	2 6 10 14 18	4 4 4 4	4	5
5,8,11,15	4 8 12 16 20	4 4 4 4	4	5
1,3,5,9	1 5 9 13 17 21	4 4 4 4 4	4	6

Table 9: List of Based data supplied by IBM

TrId	Itemset
1	25 52 164 240 274 328 368 448 538 561 630 687 730 775 825 834
2	39 120 124 205 401 581 704 814 825 834
3	35 249 674 712 733 759 854 950
4	39 422 449 704 825 857 895 937 954 964
5	15 229 262 283 294 352 381 708 738 766 853 883 966 978
6	26 104 143 320 569 620 798
7	7 185 214 350 529 658 682 782 809 849 883 947 970 979
8	227 390
9	71 192 208 272 279 280 300 333 496 529 530 597 618 674 675 720 855 914 932
10	183 193 217 256 276 277 374 474 483 496 512 529 626 653 706 878 939
11	161 175 177 424 490 571 597 623 766 795 853 910 960
12	125 130 327 698 699 839
13	392 461 569 801 862
14	27 78 104 177 733 775 781 845 900 921 938
15	101 147 229 350 411 461 572 579 657 675 778 803 842 903
16	71 208 217 266 279 290 458 478 523 614 766 853 888 944 969
17	43 70 176 204 227 334 369 480 513 703 708 835 874 895
18	25 52 278 730
19	151 432 504 830 890
20	71 73 118 274 310 327 388 419 449 469 484 706 722 795 810 844 846 918
21	130 274 432 528 967
22	188 307 326 381 403 523 526 722 774 788 789 834 950 975
23	89 116 198 201 333 395 653 720 846
24	70 171 227 289 462 538 541 623 674 701 805 946 964
25	143 192 317 471 487 631 638 640 678 735 780 865 888 935
26	17 242 471 758 763 837 956
27	52 145 161 283 375 385 676 721 731 790 792 885
28	182 229 276 529
29	43 522 565 617 859
30	12 296 350 354 401 548 684 740 774 775 782 841 937

Table 10 : Vertical format

Item Number	Transaction ID
1	55 136 152 187 227 236 414 557 595 659 745 775 887...
2	96 162 179 313 341 578 915 1189 1269 1278 1404 147...
3	737 1057 1108 1179 1530 1823 2024 2300 2415 2494 2...
4	159 197 266 379 445 491 791 982 1082 1174 1223 126...
5	35 59 143 165 292 360 388 393 471 635 693 702 871 ...
6	56 161 219 255 322 323 349 361 723 734 736 750 802...
7	7 178 252 397 754 769 868 970 1094 1162 1251 1275 ...
8	91 121 130 215 257 324 326 390 438 439 448 455 464...
10	44 147 224 288 354 366 381 396 498 507 529 595 623...
11	248 261 482 672 779 838 1063 1360 1459 1568 1781 1...
12	30 83 86 101 176 185 235 254 259 291 327 350 395 4...
13	5497 5586 7010 8372 10015 13935 16400 17122
14	834 882 890 1215 1439 1968 2556 2661 2809 3211 378...
15	5 421 876 911 1099 1278 1516 1544 2004 2040 2754 2...
16	305 1247 1355 2095 2597 2746 2885 3198 4506 5016 5...
17	26 77 78 82 102 186 188 416 483 509 620 660 682 73...
18	86 144 347 380 472 717 931 1092 1123 1154 1249 136...
19	495 1415 1986 2082 3825 3853 4378 4881 5102 5195 5...
20	433 512 6779 8099 8852 15866 15966 16429
21	46 67 71 97 133 159 187 270 277 293 330 512 632 64...
22	77 597 859 1160 1399 2740 2835 2896 2956 3232 3293...
23	377 400 1016 1232 1248 1659 1891 2405 2857 4472 47...
24	73 553 831 1995 2258 2919 3309 3336 4733 5348 6650...
25	1 18 172 308 382 400 637 658 675 781 828 849 979 1...
26	6 47 221 594 825 879 1303 1500 1644 1660 1718 1748...
27	14 112 113 133 215 260 287 292 399 437 474 497 502...
28	34 165 253 264 322 353 376 388 393 525 530 564 718...
29	968 3154 3986 4680 5563 5584 5945 6788 7282 9120 9...
31	39 53 114 194 221 237 241 322 325 327 403 426 471 ...