



Extração de Conhecimento por meio de Regras de Associação

Alex Werneck Ferreira

JUIZ DE FORA
DEZEMBRO, 2016

Extração de Conhecimento por meio de Regras de Associação

ALEX WERNECK FERREIRA

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Bacharelado em Ciência da Computação

Orientador: Victor Ströele de Andrade Menezes

JUIZ DE FORA
DEZEMBRO, 2016

EXTRAÇÃO DE CONHECIMENTO POR MEIO DE REGRAS DE ASSOCIAÇÃO

Alex Werneck Ferreira

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Victor Ströele de Andrade Menezes
Doutor em Engenharia de Sistemas e Computação (UFRJ)

Luciana Conceição Dias Campos
Doutora em Engenharia Elétrica (PUC-Rio)

Heder Soares Bernardino
Doutor em Modelagem Computacional (LNCC/MCTI)

JUIZ DE FORA
15 DE DEZEMBRO, 2016

"Todos os nossos sonhos podem-se realizar, se tivermos a coragem de persegui-los."

— WALT DISNEY

Resumo

Atualmente, muitas informações são geradas pelas mais variadas fontes de dados. Empresas, independentemente do seu porte, têm percebido o quão importante os dados são. A partir dessa premissa, o que se tem observado é que, nos dias atuais, tem havido um grande investimento das empresas em relação ao armazenamento e ao tratamento dos dados para posteriormente transformá-los em informações relevantes, pois essas informações geram um certo conhecimento que, mais tarde, pode ser usado por elas como um diferencial competitivo, auxiliando na tomada de decisões ou até mesmo em alguma definição estratégica. No presente trabalho, serão aplicadas técnicas de mineração de dados, em uma base de dados de um mercado de médio porte, utilizando Regras de Associação, onde será aplicado o algoritmo Apriori para a obtenção de conhecimento.

Palavras-chave: Apriori, Regras de Associação, Mineração de Dados.

Abstract

Currently, many information is generated for a variety of data sources. In the business environment, companies regardless of their size, have realized how important the data is. From this premise, which has been observed is that, today, there has been a great business investment in relation to the storage and processing of data to later turn them into relevant information, as this information can generate knowledge, and that can later be used by these companies as a competitive advantage, helping in decision making or even in some strategic definition. In this paper, we apply data mining techniques in a database of a midsize supermarket using Association Rules, which will apply the Apriori algorithm for obtaining knowledge.

Keywords: Apriori, Association Rules, Data Mining.

Agradecimentos

Quero agradecer, em primeiro lugar, a Deus, pela força e coragem durante toda esta longa caminhada.

Aos meus pais, por me apoiarem nos momentos mais difíceis e a sempre me encorajarem a seguir em frente.

Aos amigos, por todo carinho e compreensão.

A minha namorada, pelo companheirismo e amor.

Ao professor Victor, pela orientação, amizade e principalmente, pela paciência, sem a qual este trabalho não se realizaria.

Aos professores do Departamento de Ciência da Computação, pelos seus ensinamentos e aos funcionários do curso, que durante esses anos, contribuíram de algum modo para o nosso enriquecimento pessoal e profissional.

*”Sonhos determinam o que você quer. Ação
determina o que você conquista”.*

Aldo Novak.

Sumário

Lista de Figuras	7
Lista de Tabelas	8
Lista de Abreviações	9
1 Introdução	10
1.1 Apresentação do Tema	10
1.2 Problema	10
1.3 Justificativa	11
1.4 Objetivos	11
1.5 Metodologia	12
2 Fundamentação Teórica	14
2.1 Revisão Bibliográfica	14
2.2 Data Mining	15
2.2.1 Regras de Associação	16
2.2.2 Conceitos	18
2.2.3 Algoritmo Apriori	18
2.2.4 Funcionamento	19
2.2.5 Ferramenta WEKA	20
3 Análise dos Dados	25
3.1 Considerações	25
3.1.1 Base de Dados e Extração	25
3.1.2 Pré-Processamento	27
3.1.3 Mineração de Dados	33
3.2 Conhecimentos Adquiridos	34
3.2.1 Resultados	37
3.2.2 Trimestre 1	37
3.2.3 Trimestre 2	39
3.2.4 Trimestre 3	41
3.2.5 Trimestre 4	44
3.2.6 Trimestre 5	46
3.2.7 Trimestre 6	48
3.2.8 Trimestre 7	49
3.2.9 Trimestre 8	51
3.2.10 Trimestre 9	53
3.2.11 Trimestre 10	55
3.3 Considerações Finais	57
4 Conclusão	60
4.1 Trabalhos Futuros	60

Lista de Figuras

1.1	Etapas para busca de conhecimento	12
2.1	Processo para extração de conhecimento	15
2.2	Análise do Cesto de Compras	17
2.3	Interface WEKA	21
2.4	Funcionalidade WEKA	21
2.5	Arquivo ARFF	22
2.6	Cabeçalho do arquivo ARFF	23
2.7	Declaração dos dados ou instâncias no arquivo ARFF	24
3.1	Pequena amostra dos dados disponibilizados	26
3.2	Remoção de atributos no processo de limpeza	28
3.3	Itens do mesmo tipo com grupos diferentes.	30
3.4	Itens do mesmo tipo numa mesma categoria.	30
3.5	Itens iguais com nomes diferentes.	31
3.6	Itens com nomes iguais.	31
3.7	Itens de um mesmo tipo com marcas diferentes.	32
3.8	Itens de acordo com tipo.	33
3.9	Parâmetros Algoritmo Apriori.	35
3.10	Resultados do Primeiro Trimestre.	38
3.11	Amostras dos Resultados do Segundo Trimestre	40
3.12	Amostras dos Resultados do Terceiro Trimestre	42
3.13	Amostras dos Resultados do Quarto Trimestre	44
3.14	Amostras dos Resultados do Quinto Trimestre	46
3.15	Amostras dos Resultados do Sexto Trimestre	48
3.16	Amostras dos Resultados do Sétimo Trimestre	50
3.17	Amostras dos Resultados do Oitavo Trimestre	52
3.18	Amostras dos Resultados do Nono Trimestre	53
3.19	Amostras dos Resultados do Décimo Trimestre	55
3.20	Total de Compras por Trimestre	57

Lista de Tabelas

3.1	Atributos de um item comprado	27
3.2	Produtos com maior frequência de compras	34
3.3	Parâmetros de entrada para o processo de mineração.	36

Lista de Abreviações

DCC Departamento de Ciência da Computação

UFJF Universidade Federal de Juiz de Fora

KDD Knowledge Discovery in Database (Descoberta de Conhecimento em Base de Dados)

PDV Ponto de Venda

1 Introdução

1.1 Apresentação do Tema

Ao longo dos anos no mundo empresarial, percebe-se o aumento cada vez maior da busca de informações em sistemas informatizados, conforme Campos Filho (1994). Essa procura ocorre, pois muitas informações relevantes podem ser descobertas nas bases de dados. Como consequência, percebe-se que a aplicação de métodos ou técnicas de extração de conhecimento em base de dados vem sendo amplamente utilizados na obtenção dessas informações. Dessa forma, dados históricos passam a ter um valor fundamental no processo, pois é a partir deles que é possível extrair essas informações, afim de gerar conhecimento.

Diversas técnicas são abordadas no contexto geral, porém, antes de sua implantação, o domínio do problema deve ser compreendido. Dessa forma, será possível escolher a melhor delas para sua adequada aplicação e obtenção de resultados satisfatórios. Mañas (2005) reforça que o gerenciamento do conhecimento obtido pela empresa ou instituição será como total afirmação, o seu mais valioso bem. Empresas que conseguem fazer bom uso das tecnologias do mercado, aliado ao conhecimento adquirido a partir da aplicação de técnicas de extração de conhecimento, de fato estão um passo a frente no quesito competitividade.

1.2 Problema

A partir do que foi abordado em 1.1, é pretendido saber, se o uso de técnicas de extração de conhecimento será positivo no contexto de um mercado varejista de médio porte. A aplicação das técnicas terá como objetivo a extração das informações relevantes da base de dados do mercado, para que assim, seja possível definir algumas estratégias que possam ser usadas em uma posterior tomada de decisão. Para tanto, serão utilizados algoritmos baseados em regras de associação, que é uma das tarefas mais utilizadas em mineração de dados, conforme de Vasconcelos (2004).

Os algoritmos têm como premissa, encontrar elementos que implicam na presença de outros elementos, encontrando relacionamentos e padrões frequentes na base de dados analisada. O problema é de grande complexidade, pois pode-se não encontrar relações entre os itens, já que não se sabe a frequência de compras que são realizadas em um determinado período, por exemplo. Este fator pode afetar diretamente os resultados, fazendo com que se consiga pouco ou nenhum conhecimento. Para isso, deverá ser feita uma análise de como poderá ser obtido melhor relacionamento entre os produtos da base de dados disponibilizada, de modo que seja possível obter informações que irão dar suporte a uma posterior tomada de decisão.

1.3 Justificativa

De acordo com Quesado (2014), Inteligência Competitiva é uma maneira proativa de se obter e articular informações que gerem valor para a organização através de análises de tendências e cenários, sendo que essas informações podem ser referentes ao comportamento da concorrência, clientes e do mercado em geral.

Como salientado por Han et al. (2011), o uso de mineração de dados, em especial o que aqui será abordado, tem sido bastante utilizado e tem gerado resultados satisfatórios, os quais apoiaram em tomadas de decisão e contribuíram para gerar um potencial fator competitivo para as organizações. A partir dessa premissa, justifica-se como aplicação do trabalho, a obtenção de resultados que irão auxiliar o mercado a definir certas estratégias, com o intuito de proporcionar uma maior captação de renda, que será impulsionado pelo aumento do número de vendas do mercado em questão.

1.4 Objetivos

O objetivo principal desse trabalho é obter informações úteis através do uso de Mineração de Dados, explorando o uso de algoritmos, em especial o APRIORI, com a finalidade de se obter Regras de Associações que serão traduzidas em conhecimento a respeito dos produtos consumidos, tornando possível a identificação de padrões que poderão auxiliar os gestores em futuras tomadas de decisões, fazendo disso um eminente diferencial competitivo para

o mercado.

Os objetivos específicos do trabalho são:

- Descobrir padrões de consumo nas compras realizadas, avaliando o comportamento de consumo dos clientes em diferentes épocas do ano ao longo de 10 trimestres, de forma que possa ser feita uma analogia do poder de compra dos mesmos, para um determinado intervalo de tempo especificado;
- Avaliar o conhecimento obtido para cada época do ano de modo que possibilite a elaboração de estratégias temporais e a análise do relacionamento entre os produtos existentes no mercado, a fim de caracterizar uma venda estratégica.

1.5 Metodologia

De acordo com Agrawal et al. (1993), o processo de busca de conhecimento é composto por várias importantes etapas, conforme Figura 1.1.

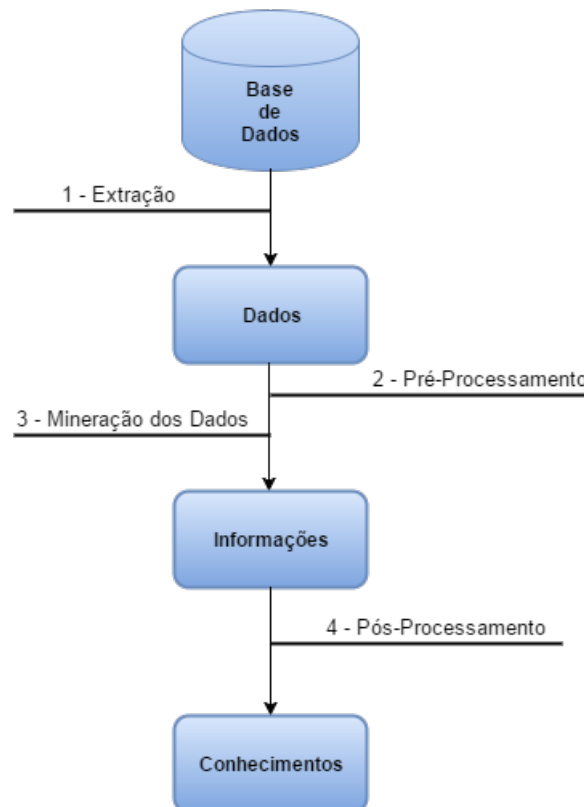


Figura 1.1: Etapas para busca de conhecimento

Fonte: Baseado em Agrawal et al. (1993).

Conforme Agrawal et al. (1993), as etapas de Extração e Pré-Processamento demonstradas na Figura 1.1, são fundamentais para o processo de mineração de dados. Portanto, foi necessário executar estas tarefas, dentre as quais foram utilizadas, limpeza, transformação e redução dos dados, com o intuito de se obter uma melhor definição dos dados, tornando possível a utilização dos mesmos, com o objetivo de facilitar a conclusão das outras etapas importantes dentro do processo de *Descoberta de Conhecimento em bases de dados* (KDD), para posterior aplicação de técnicas de extração de conhecimento na base adquirida.

A partir da fase de Extração e Pré-Processamento, é feita a parte de mineração de dados. Nesta etapa, são apresentados os padrões obtidos no conjunto de dados, onde as informações geradas irão refletir mais tarde em conhecimentos sobre os dados que poderão ser úteis no auxílio de definições a respeito de estratégias para vendas.

Portanto, será feito um levantamento das relações existentes entre os produtos do mercado observando padrões existentes na base de dados, através da aplicação de algoritmos baseados em regras de associações. Dessa forma, será possível fazer algumas análises, visando obtenção de associações diretas entre os produtos listados na base de dados, de forma a obter conhecimento necessário que, mais tarde, poderá ser utilizado numa eventual tomada de decisão. A partir desse estudo, será possível fazer um levantamento mais específico, o qual ajudará a relacionar quais produtos são mais consumidos em determinadas fases do ano.

Para tal atividade, todo o trabalho de manipulação de dados e extração de conhecimento será feito utilizando o algoritmo APRIORI, que é um algoritmo baseado em regras de associação, amplamente utilizado para descobertas de conhecimento em base de dados.

2 Fundamentação Teórica

2.1 Revisão Bibliográfica

Conforme enunciado anteriormente e apresentado por Han et al. (2011), a utilização de algoritmos baseados em regras de associação, tem ajudado empresas a obter um conhecimento importante sobre informações que antes, eram dadas como inutilizáveis. A percepção do quão importante são as informações históricas, fazem com que as empresas passem a investir em sistemas de apoio a decisão, pois a utilização deles fazem com que as empresas tenham um diferencial competitivo.

O caso mais conhecido da aplicação da mineração de dados, em especial, do uso de algoritmos baseados em regras de associação, é exemplificado na literatura através da relação fralda \rightarrow cerveja, como denotado por Gahane (2013). Foi desenvolvido um estudo em uma base de dados, onde foi possível observar o comportamento de compra dos clientes que inferia na seguinte constatação, às sexta-feiras, no horário da tarde, quem comprava fralda, comprava cerveja. A partir dessa percepção, foi possível elaborar algumas estratégias com o intuito de alavancar a venda desses produtos para o mercado em questão.

Vários trabalhos em diversas áreas apresentam a aplicação de algoritmos baseados em regras de associação. Conforme Kaur and Kang (2016), a mineração de dados pode ser aplicada em áreas como: bio-informática, comércios, indústrias e etc. O foco da aplicação é, prover conhecimento acerca das informações contidas na bases de dados, de forma que seja possível analisar o comportamento relacional dos itens contidos no domínio, possibilitando uma posterior tomada de decisão.

Na aplicação desenvolvida no contexto de um e-commerce apresentado por Yan (2015), foi possível analisar o comportamento dos usuários através de dados históricos, com o intuito de se tomar medidas que auxiliariam em definições de estratégias de vendas futuras.

O estudo realizado tem como base utilização dos princípios adotados nesta seção,

utilizando de uma abordagem um pouco mais específica acerca dos produtos comprados, onde será analisada a frequência dos produtos e suas relações como outros produtos em determinadas épocas do ano, que para este trabalho será compreendido dentro de um período de 10 trimestres.

2.2 Data Mining

Conforme apresentado por Han et al. (2011), pessoas fazem a associação de mineração de dados como se tivessem o mesmo significado de KDD (Knowledge-Discovery in Databases). Isto é um pensamento equivocado da maioria delas, pois KDD é um processo iterativo que é constituído por várias etapas, dentre elas a mineração de dados, onde o objetivo geral é extrair conhecimentos da base de dados, observando padrões que não são facilmente detectados, para posterior avaliação dos mesmos.



Figura 2.1: Processo para extração de conhecimento

A Figura 2.1, representa as diversas etapas para a descoberta de conhecimento. Como pode ser observado, o processo descrito na Figura 2.1 pode ser dividido nas seguintes etapas: Extração dos Dados, Pré-Processamento, Mineração de Dados, Pós-Processamento e Modelo de Regras.

Cada uma das etapas possuem subetapas, que irão compor o modelo ou o conjunto de regras que espera-se obter. A ligação exemplificada entre a etapa da obtenção de um Modelo/Regras e o Pré-Processamento, demonstram as limitações que podem ser encontradas no processo e que impede a obtenção de bons resultados. As primeiras etapas, principalmente a etapa de Pré-Processamento, serve para dar sustentação a uma das principais etapas do processo, que é a etapa de mineração de dados, pois é nela que serão feitas algumas sub-tarefas que irão garantir a qualidade dos dados que serão

minerados.

Segundo Agrawal et al. (1993), Mineração de Dados é o processo de análise dos dados provenientes de uma fonte, analisando suas diversas perspectivas, com o intuito de transformá-los em informações relevantes, sempre observando padrões, os quais podem ser usados para a geração de conhecimento.

Data Mining é usado principalmente por grandes companhias, as quais utilizam ferramentas para extração de conhecimento de seus dados históricos, visando obter um diferencial competitivo. Para isso Data Mining oferece diversas ferramentas para a descoberta de padrões e associações, que podem vir a ter um significado muito importante para elas.

2.2.1 Regras de Associação

A análise de dados baseada em regras de associação, tem como objetivo descobrir relacionamentos relevantes no conjunto de dados contidos em uma base de dados, tornando possível assim, descobrir relações entre dados que antes passavam despercebidas, justamente pelo fato de muitas das relações apresentadas não serem tão óbvias, para posterior avaliação das mesmas.

De acordo com Han et al. (2011) a descoberta de padrões utilizando algoritmos baseados em regras de associação podem ser divididas em duas partes:

1. Encontrar todos os *itemsets* frequentes, que são conjuntos de itens que são vendidos juntamente e com mais frequência. Para cada um desses *itemsets* encontrados, irá ser verificado na base de dados, quais irão satisfazer frequência mínima estipulada, que é determinada suporte.

2. Descobrir relações entre os itens do conjunto, baseando-se nas frequências do *itemsets*. Por definição essas regras precisam satisfazer o suporte e a confiança mínima que serão apresentados adiante.

Um dos problemas da utilização de algoritmos baseados em regras de associação em um grande conjunto de dados é que pode ser gerado um amplo conjunto de *itemsets* que satisfazem o suporte mínimo, tornando difícil a tarefa de buscar informações ou associações relevantes para pesquisa.

Ao pesquisar trabalhos relacionados, foi verificada a vasta aplicabilidade a respeito das regras de associações. Muitos dos trabalhos encontrados, revelam a utilidade da aplicação de regras de associações, com o intuito de se descobrir relações que, como dito anteriormente, podem passar despercebidas. Este fato acontece principalmente quando está a se fazer uma análise de uma base de dados que possui um volume de dados considerável, conforme em Liu (2010).

Um dos casos mais conhecidos na literatura com relação às regras de associações, certamente é o caso da Análise do Cesto de Compras (*Market Basket Analysis*). Conforme descrito por Han et al. (2011), trata-se de um caso onde pretende-se analisar o hábito de compras dos clientes, analisando a cesta de compras, com o intuito de conseguir informações sobre quais itens são comprados em conjunto, podendo auxiliar na criação de estratégias para vendas futuras.

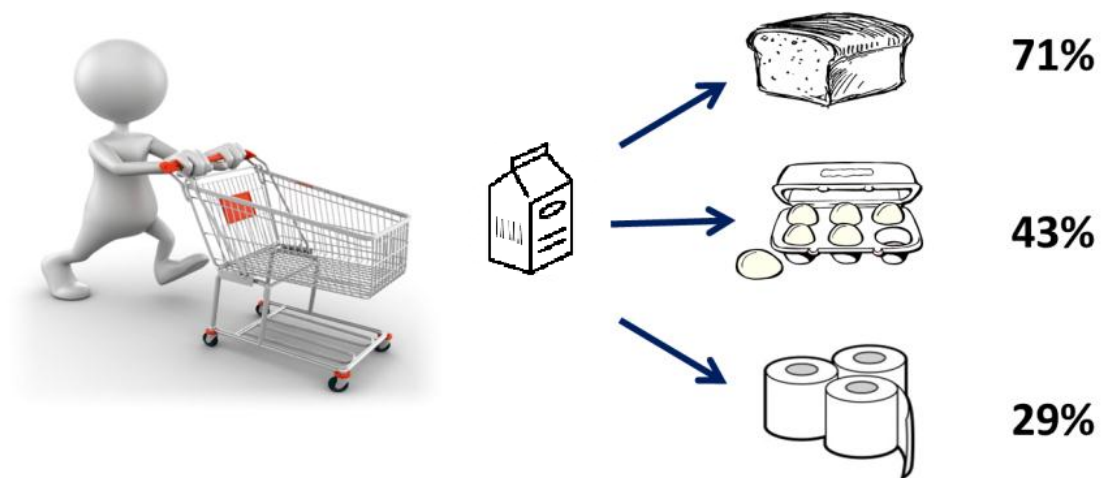


Figura 2.2: Análise do Cesto de Compras

Na Figura 2.2, pode-se ter uma noção do que é esperado como interpretação dos dados na abordagem da análise do cesto de compras. No presente trabalho, busca-se a interpretação dos dados através de uma abordagem diferente do que é comumente utilizada na maioria dos problemas relacionados a análise do cesto de compras. Conforme abordado no capítulo 1.4, a abordagem que será utilizada, visa identificar padrões distribuídos ao

longo de períodos determinados, de forma que seja possível encontrar padrões, para que posteriormente fosse viável de analisar o comportamento de compra relacionado àquela mesma fase do ano.

2.2.2 Conceitos

Os algoritmos baseados em regras de associações, em especial o que será abordado, APRIORI, possui alguns conceitos como suporte mínimo e confiança mínima, que podem ser explicados pelas seguintes definições:

- Suporte Mínimo (minsup): Tem como definição a frequência das transações (Vendas constituídas por itens) que satisfazem a união dos itens de um determinado conjunto (Frequência do item X e Frequência do item Y), com relação ao total de transações contidas na base de dados.

$$\text{Suporte} = \frac{\text{Frequência de X e Y}}{\text{Total de Atributos}} \quad (2.1)$$

- Confiança Mínima: É a relação entre a quantidade total de um determinado conjunto e a quantidade total de dados analisados de um determinado item pertencente ao conjunto na análise do total de transações contidas na base de dados.

$$\text{Confiança} = \frac{\text{Frequência de X e Y}}{\text{Frequência de X}} \quad (2.2)$$

2.2.3 Algoritmo Apriori

Baseado em regras de associação e que foi proposto por Agrawal et al. (1993). É um algoritmo utilizado para descoberta de conhecimento em grandes conjunto de dados, conforme Liu (2010).

De acordo com o que foi apresentado por Agrawal et al. (1993), o primeiro passo consiste em determinar a quantidade de ocorrência de cada item, isso se faz pela determinação dos itens relativamente mais frequentes do conjunto. O próximo passo, denominado passo k, consiste em duas etapas. Na primeira, todos os grupos mais frequentes de

itens são encontrados a partir dos itens denominados no passo k-1, onde formarão grupos de itens de C_k , isso acontece com a utilização da função *apriorigen*. Após essa etapa, toda base de dados é percorrida, com o intuito de verificar o suporte mínimo dos candidatos do grupo C_k . Para uma análise rápida, deve-se escolher os grupos C_k de forma que se possa obter os melhores resultados em cada passagem do algoritmo.

Algorithm 1 APRIORI

 C_k : itemset candidato de tamanho k

 L_k : itemset frequente de tamanho k

 T : total de itens da base de dados

início

```

   $L_1 = \{\text{itens frequentes de tamanho } 1\}$ 
  para ( $k = 1; L_k \neq \text{vazio}; k_{++}$ ) faça
     $C_{k+1} = \text{apriorigen}(L_k)$  ▷ Gerar itemsets Candidatos
    para cada  $t \in T$  faça
       $C_t = \text{subset}(C_k, t)$  ▷ Identificar todos os candidatos que pertencem a t
      para cada  $c \in C_t$  faça
         $\sigma(c) = \sigma(c) + 1$  ▷ Incrementar o Contador de Suporte
      fim para
       $L_{k+1} = \text{candidatos em } C_{k+1} \text{ com minsup}$  ▷ Extrair os k+1 - itemsets frequentes
    fim para
  fim para
fim
retorna  $\cup_k L_k$ 

```

2.2.4 Funcionamento

O algoritmo APRIORI é de fácil aplicação e divide-se em duas etapas:

1. Primeiro tem-se a geração dos conjuntos candidatos através da verificação do Suporte Mínimo estabelecido. Nessa etapa, se faz a separação de todas as possíveis combinações de itens, verificando se o conjunto gerado satisfaz o suporte mínimo pré estabelecido.

Inicialmente, um conjunto de apenas dois itens é analisado, sendo que esses elementos pertencentes a este conjunto já foram previamente analisados e que satisfizeram o suporte mínimo estabelecido. Posteriormente, é verificado a validade do suporte mínimo para este grupo contendo apenas dois itens, se é um conjunto válido, é gerado um novo conjunto agora contendo três elementos fazendo a mesma validação. Esta validação é feita

até que não existam mais combinações para serem realizadas com os itens presentes na base de dados.

Os conjuntos que não satisfizeram o suporte mínimo estabelecido são descartados a cada passagem do algoritmo. Dessa maneira, evita-se o acúmulo de dados, possibilitando uma maior eficiência na descoberta de informações.

2. Após o passo exemplificado, é feita então a criação das Regras de Associações. Dado o conhecimento obtido dos conjuntos que satisfizeram o suporte mínimo, verifica-se então se os mesmos, que são apresentados por regras, satisfazem agora a confiança mínima. As regras que satisfizerem a confiança mínima estabelecida, serão o resultado da execução do algoritmo.

A partir da obtenção dos conjuntos que se adequaram às métricas de suporte mínimo e confiança mínima estabelecidas, deve-se analisar quais regras geradas a partir desses conjuntos, irão possibilitar uma interpretação em que possa ser útil identificar particularidades a respeito da relação entre os itens encontrados na base de dados.

2.2.5 Ferramenta WEKA

Segundo Garner et al. (1995), a ferramenta foi criada para trazer uma variedade de técnicas de aprendizado de máquina, os quais são construídos por algoritmos de forma que fosse possível aplicar em grandes conjuntos de dados, de uma forma fácil, e que gerasse informações necessárias para dar suporte a uma análise exploratória.

Desenvolvido na linguagem JAVA, e com uma interface de fácil entendimento, a ferramenta consegue manipular arquivos de dados e produzir resultados de uma forma visualmente simples.

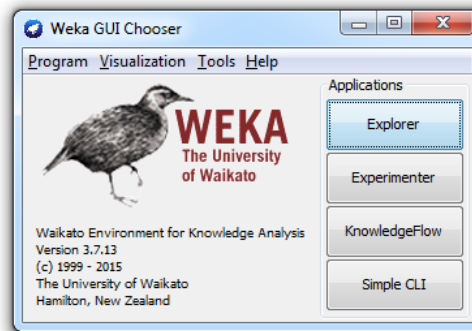


Figura 2.3: Interface WEKA

Para o presente trabalho, irá ser abordada a funcionalidade “Explorer” conforme Figura 2.3. Essa funcionalidade fornece um ambiente onde será possível explorar os dados fornecidos. Particularmente, será utilizada a opção “Associate” conforme Figura 2.4, que é uma implementação do Algoritmo APRIORI citado anteriormente, onde através do arquivo de dados carregado, é pretendido obter um certo aprendizado sobre as regras de associações presentes no conjunto de dados fornecido.

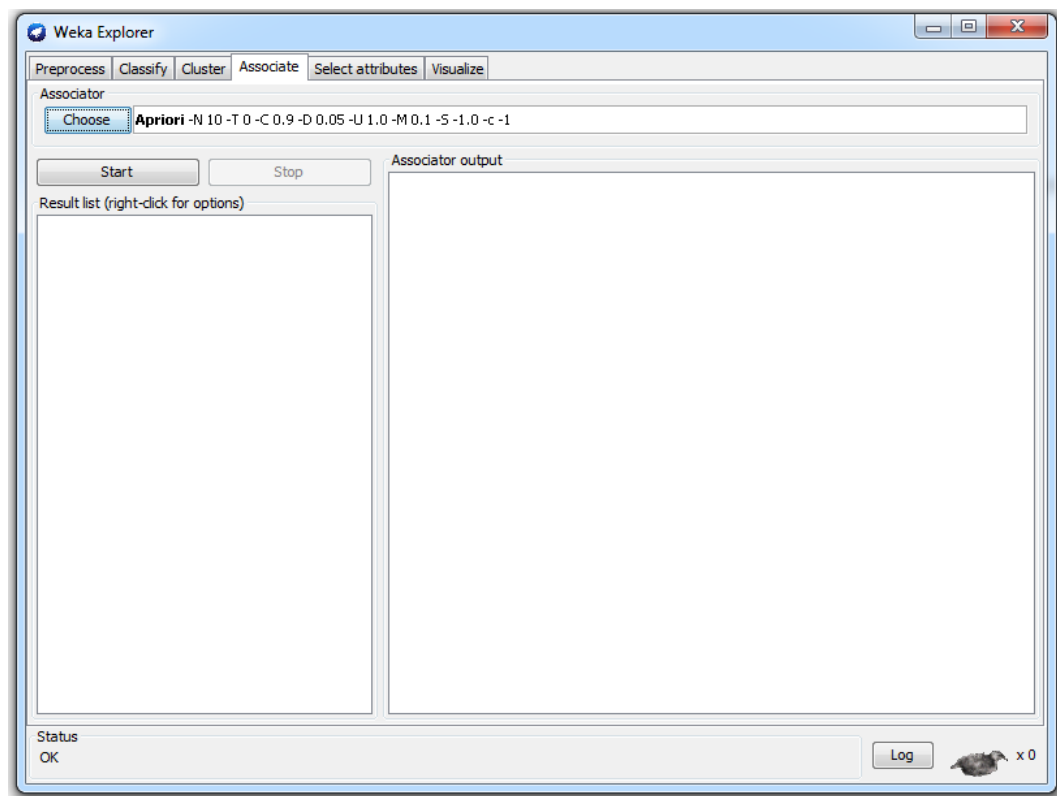


Figura 2.4: Funcionalidade WEKA

Antes de se iniciar o processo de associação, foi preciso elaborar os arquivos de entrada para consumo do software. A aplicação pode importar diversos tipos de arquivos de dados, como por exemplo: CSV, C4.5, arquivos Binários e ARFF. Este último, denominado como Attribute-Relation File Format (ARFF) ou, na tradução literal, Formato de Arquivo Atributo-Relação, representa uma das formas mais utilizadas para manipulação de dados dentro da aplicação, e foi o tipo de arquivo de entrada de dados que utilizamos neste trabalho. No arquivo encontra-se a definição de cada atributo e o que cada atributo contém. A Figura 2.5 apresenta uma exemplificação do conteúdo do arquivo de entrada da aplicação.

```
% 1. Title: Iris Plants Database
%
% 2. Sources:
%   (a) Creator: R.A. Fisher
%   (b) Donor: Michael Marshall (MARSHALL%PLU@io.arc.nasa.gov)
%   (c) Date: July, 1988
%
@RELATION iris

@ATTRIBUTE sepallength REAL
@ATTRIBUTE sepalwidth REAL
@ATTRIBUTE petallength REAL
@ATTRIBUTE petalwidth REAL
@ATTRIBUTE class {Iris-setosa,Iris-versicolor,Iris-virginica}

@DATA
5.1,3.5,1.4,0.2,Iris-setosa
4.9,3.0,1.4,0.2,Iris-setosa
4.7,3.2,1.3,0.2,Iris-setosa
4.6,3.1,1.5,0.2,Iris-setosa
5.0,3.6,1.4,0.2,Iris-setosa
5.4,3.9,1.7,0.4,Iris-setosa
```

Figura 2.5: Arquivo ARFF

O arquivo da Figura 2.5, é uma exemplificação do formato do arquivo que se deve criar para todas as relações que irão ser feitas através da abordagem explicitada anteriormente. O arquivo possui duas partes bem distintas. A primeira parte é composta pelo cabeçalho de informações, nessa parte é necessário explicitar alguns dados como: o nome da relação (encabeçado por *@relation* - Nome da Relação) , a lista de atributos

(denominado por *@attribute* - Nome do atributo) e conseqüentemente seus tipos, como demonstrado na Figura 2.6 a partir dos arquivos gerados para este trabalho.

```
@relation TRIMESTRE1

@attribute ABSORVENTE {1,0}
@attribute ACETONA {1,0}
@attribute ACHOCOLATADO {1,0}
@attribute ACUCAR {1,0}
@attribute ADOCANTE {1,0}
@attribute AGUA_OXIGENADA {1,0}
@attribute AGUA {1,0}
@attribute ALGODAO {1,0}
@attribute ALICATE {1,0}
@attribute AMENDOIN {1,0}
@attribute AMONIA {1,0}
@attribute AMPOLA {1,0}
@attribute ANILINA {1,0}
@attribute APRESUNTADO {1,0}
@attribute ARROZ {1,0}
@attribute AZEITONA {1,0}
```

Figura 2.6: Cabeçalho do arquivo ARFF

A segunda parte do arquivo contém a declaração dos dados em si. Essa parte é precedida pela declaração *@data*, que é uma marcação para o programa saber onde começa a declaração dos dados. Cada linha é representada pela presença ou não dos atributos anteriormente denotados no cabeçalho conforme foi exemplificado nas figuras 2.5 e 2.6.

No arquivo elaborado para o trabalho, foi indicado a presença de um atributo booleano 1 ou 0, o valor 1 indicará a presença de um determinado produto em uma determinada transação (Venda), já o valor 0, indicará a não presença de um determinado produto em uma transação. Para este trabalho não teve-se o intuito de demonstrar quais produtos não participam de uma determinada venda, e sim quais estariam presentes, os quais seriam representados pelo valor 1.

A não verificação de um determinado item de uma compra é representado pelo caracter '?'. Dessa forma o programa consegue entender que um atributo estará presente quando ele encontrar o valor 1, e saberá também que deve ser ignorado um atributo que for representado pelo caracter '?', conforme exemplificado na Figura 2.7, que é proveniente

de um dos arquivos criados para a elaboração das regras de associação presentes entre os itens do mercado, e que estão relacionados ao primeiro semestre.

```

@data
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?
?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?

```

Figura 2.7: Declaração dos dados ou instâncias no arquivo ARFF

Para este trabalho, foram criados 10 arquivos ARFF , um para cada trimestre, com o intuito de se obter regras de associações que possam auxiliar na análise dos produtos e seus relacionamentos em determinadas fases do ano, e que servirá de suporte para uma posterior tomada de decisão ou até mesmo para uma eventual elaboração estratégica.

3 Análise dos Dados

3.1 Considerações

Os dados utilizados para essa pesquisa são de um mercado de médio porte. O mercado utiliza de sistemas tipos PDV's, que são integrados a um banco de dados transacional, onde são armazenadas todas as informações sobre compras que foram realizadas no estabelecimento. O mercado tem um moderado volume de compras e é esperado que a base de dados possa fornecer bastante conhecimento.

Ao todo, foram fornecidos mais de 300 mil registros de compras, distribuídos ao longo de um intervalo de dois anos e meio, e que foram utilizados nesse trabalho para serem analisados ao longo de trimestres.

Conforme mencionado no capítulo 2, seção 2.2.1, este problema é conhecido como análise da cesta de compras (Market Basket Analysis). Este problema consiste em verificar as transações efetuadas (compras) e seus atributos constituintes (cestas), as quais são constituídas por itens, que são os produtos do mercado em questão. O problema tem como objetivo, encontrar itens que são frequentes em uma mesma cesta, de forma que se possa observar que, a presença de um determinado produto, pode vir a implicar na presença de outro produto para uma mesma cesta.

3.1.1 Base de Dados e Extração

Por ser um banco de dados transacional e possuir uma grande quantidade de dados heterogêneos, é esperado que seja possível aplicar as técnicas descritas no processo KDD, principalmente para a etapa de pré-processamento, visando obter uma maior qualidade nos dados para posterior mineração.

O conjunto de dados do mercado que foi fornecido veio através de um arquivo em formato de planilha. Com o arquivo em mãos, foi possível identificar e avaliar os atributos presentes, assim como possíveis "erros" causados pela não padronização no que

se diz respeito à inserção dos dados na base. Em uma primeira análise, foi verificado uma quantidade elevada de dados inconsistentes, assim como atributos predispostos em colunas dentro do arquivo, que armazenavam informações nulas a respeito de um determinado produto.

Na Figura 3.1, pode ser visto uma pequena amostra dos dados disponibilizados para este estudo.

Número da Venda	Data venda	nome estacao	Unitario	Quantidade	Valor	Hora	Descricao reduzida	Produto cancelado	Forma pag
6200	01/01/2013	CAIXA5	5,5	1	5,5	10:47:28	AGUA MIN DEMINAS 20	NAO	CONVENIOS
6217	01/01/2013	CAIXA5	1	1	1	11:18:09	AGUA MINERAL S/GAS	NAO	Dinheiro
6219	01/01/2013	CAIXA5	3,299999952	1	3,3	11:22:37	GATORADE SABOR UVA	NAO	CONVENIOS
6213	01/01/2013	CAIXA5	1,5	1	1,5	11:09:37	REFRIG PETBIZ 2L LI	NAO	Dinheiro
057109	01/01/2013	CAIXA1	4,5	1	4,5	08:11:36	GUARANA ANTARCTIC	NAO	
6187	01/01/2013	CAIXA5	6,900000095	1	6,9	10:08:18	COCA COLA 3LT	NAO	Dinheiro
057120	01/01/2013	CAIXA1	3,5	1	3,5	08:50:44	COCA COLA 1,25LT	NAO	
6194	01/01/2013	CAIXA5	5,5	1	5,5	10:36:11	COCA COLA 2LT	NAO	CONVENIOS
6197	01/01/2013	CAIXA5	4	1	4	10:40:01	COCA COLA 1,5LT	NAO	Dinheiro
6197	01/01/2013	CAIXA5	2,799999952	1	2,8	10:40:01	REFRIGERANTE SCHIN	NAO	Dinheiro
6201	01/01/2013	CAIXA5	2,799999952	1	2,8	10:49:21	REFRI SCHIN COLA 2	NAO	Dinheiro
6202	01/01/2013	CAIXA5	5,5	1	5,5	10:51:45	COCA COLA 2LT	NAO	CONVENIOS
6211	01/01/2013	CAIXA5	5	1	5	11:04:59	REFRIG KUAT 2LT	NAO	Dinheiro
6211	01/01/2013	CAIXA5	5	1	5	11:04:59	FANTA UVA 2LT	NAO	Dinheiro
6211	01/01/2013	CAIXA5	6,900000095	1	6,9	11:04:59	COCA COLA 3LT	NAO	Dinheiro
6214	01/01/2013	CAIXA5	4,5	1	4,5	11:10:38	GUARANA ANTARCTIC	NAO	Dinheiro
6217	01/01/2013	CAIXA5	4	1	4	11:18:09	GUAR ANTARCTICA 2L	NAO	Dinheiro
6220	01/01/2013	CAIXA5	5,5	1	5,5	11:23:44	COCA COLA 2LT	NAO	Dinheiro
057125	01/01/2013	CAIXA1	4,949999809	1	4,95	08:58:49	SUCO BELA IS MARA	NAO	
021359	01/01/2013	CAIXA2	1,5	1	1,5	11:39:39	SUCO B ISCHIA UVA	NAO	
6201	01/01/2013	CAIXA5	4,75	1	4,75	10:49:21	SUCO ADES UVA 1L	NAO	Dinheiro

Figura 3.1: Pequena amostra dos dados disponibilizados

Cada linha da planilha, dentre os mais de 300 mil registros disponibilizados, equivalem a um item que foi comprado no mercado. Para cada linha, foi possível identificar que as compras realizadas possuíam um código agrupador, que os caracterizavam como produtos que foram levados em uma mesma transação.

Ao todo, 15 atributos caracterizam os itens que foram comprados em uma determinada data. A descrição dos atributos encontrados na planilha estão em detalhes na tabela 3.1.

Tabela 3.1: Atributos de um item comprado

Atributo	Descrição
Número da Venda	Número da venda, numeral.
Data da Venda	Data em que o produto foi vendido.
Nome Estação	Caixa que efetuou a venda.
Unitário	Preço unitário do produto.
Quantidade	Quantidade comprada.
Valor	Valor da compra em relação ao preço e quantidade.
Hora	Hora da compra.
Descrição Resumida	Descrição resumida do produto.
Produto Cancelado	Informação se o produto foi cancelado.
Forma Pagamento	Forma de pagamento da compra.
Nome Grupo	Grupo o produto pertence. Ex: Açougue.
Nome da Marca	Nome da Marca.
Custo	Custo do Produto.
Código do Grupo	Código que identifica o grupo.
Código da Marca	Código que identifica a marca.

3.1.2 Pré-Processamento

Para a seleção dos dados, buscou-se, dentro dos 15 atributos distintos, aqueles que mais poderiam agregar valor para a pesquisa em questão. Como o objetivo do trabalho é analisar padrões de compra efetuadas em diferentes épocas do ano, verificou-se que os atributos, Número da Venda, Data da Venda, Hora da Venda, Descrição Reduzida e Nome do Grupo, poderiam caracterizar de forma absoluta as compras que foram e são realizadas no mercado, observando a que período de tempo elas estariam associadas. Portanto, a caracterização da base de dados ficou conforme amostra apresentada na Figura 3.2.

Número da Venc	Data venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
6201	01/01/2013	10:49:21	CERV ITAIPAVA 350ML	CERVEJA
6204	01/01/2013	10:53:22	CR LEITE NESTLE .	CREME DE LEITE
6203	01/01/2013	10:52:29	OVOS LIDER BCO DZ	OVOS
6203	01/01/2013	10:52:29	SAL MARFIM 1KG	EMPACOTADOS
6203	01/01/2013	10:52:29	CR LEITE PIRACANJUB	CREME DE LEITE
6202	01/01/2013	10:51:45	PAULISTA MOR 540G	LATICINIOS
6201	01/01/2013	10:49:21	PAPEL HIG PERSONAL	HIGIENE
6202	01/01/2013	10:51:45	COCA COLA 2LT	REFRIGERANTE
6204	01/01/2013	10:53:22	PAPEL ALUMINIO 4M	UTILIDADES
6201	01/01/2013	10:49:21	SAL GROSSO SOSAL 1K	SAL
6201	01/01/2013	10:49:21	PAO DE FORMA NETTO	PADARIA
6201	01/01/2013	10:49:21	SAL MARFIM 1KG	EMPACOTADOS
6201	01/01/2013	10:49:21	PIPOCA MICRO C BOM	EMPACOTADOS
6201	01/01/2013	10:49:21	REFRI SCHIN COLA 2	REFRIGERANTE
6208	01/01/2013	10:56:41	ACHOC NESCAU 800G	ACHOCOLATADO
6202	01/01/2013	10:51:45	MUSSARELA	AÇOUQUE
6206	01/01/2013	10:55:33	PRESUNTO	AÇOUQUE
057126	01/01/2013	09:04:00	ALPISTE PEREIRA 500	EMPACOTADOS
6207	01/01/2013	10:56:20	ACHOC PIRAKIDS 200M	ACHOCOLATADO
6207	01/01/2013	10:56:20	AGUA DE COCO SOCOCO	SUCOS
6206	01/01/2013	10:55:33	MAC INSTANT DOMANI	MAC INSTANTANEO
6206	01/01/2013	10:55:33	MIOJO NISSIN LAMEN	MAC INSTANTANEO

Figura 3.2: Remoção de atributos no processo de limpeza

Além disso, a base de dados foi submetida a uma prévia limpeza, conforme processo descrito em Zhang et al. (2003). Para este processo, não foi eliminado apenas os dados incoerentes, também foi efetuado a transformação de alguns deles. Isso se deu pois havia uma grande variação de marcas que representavam um mesmo tipo de produto por exemplo, e isso poderia impactar no objetivo do trabalho. Como solução, buscou-se simplificar a nomeação dos produtos assim como fazer uma caracterização dos mesmos, com o intuito de representá-los unicamente. Por exemplo, produtos como Nescau e Toddy, dois produtos amplamente reconhecidos no âmbito comercial, e que representam um mesmo tipo de produto em questão, foram transformados em um único tipo de produto, caracterizado por achocolatado. Dessa forma, foi entendido que seria a melhor maneira de se obter boas relações para os itens comprados, possibilitando assim, um melhor entendimento de todo o processo de busca de conhecimento. Portanto, as principais modificações feitas, foram em cima de dois atributos apresentados na Figura 3.2, Descrição Resumida e Nome do Grupo.

O atributo Nome do Grupo passou por uma melhor categorização. Dessa forma, foi possível alocar os itens representados pelo atributo Descrição Resumida às categorias que melhor os representaria. Isso possibilitou agrupar itens que representam um mesmo tipo de produto, indiferentemente de sua marca em categorias definidas, ajudando assim

na tarefa de pré-processamento.

Pesou-se em atribuir os produtos em categorias das quais é de costume encontrar em qualquer mercado. Essa subdivisão feita nos mercados, tem como finalidade orientar os clientes a buscarem facilmente um determinado tipo de produto. Com base nisso, foi feita a classificação de cada produto contido no arquivo, classificando-os em oito grupos distintos. Dessa forma, percebeu-se que os grupos criados trariam uma facilidade maior no processo de limpeza de dados, pois agora seria possível agrupá-los através dos seguintes grupos genéricos:

- AÇOUGUE
- BAZAR
- FARMACIA
- FRIOS
- HORTI FRUT
- LIMPEZA
- MERCEARIA
- PADARIA

Assim, espera-se que seja possível entender a relação entre os produtos das oito categorias criadas. De acordo com Angelo et al. (2003), a disposição de alguns produtos podem vir a auxiliar na compra de outros produtos que não estavam programados e, portanto, essa categorização poderá ser útil em ajudar a definir algumas estratégias importantes ao longo do processo.

O processo de categorização descrito, pode ser observada nas figuras 3.3 e 3.4.

Número da Venda	Data da Venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
49052	14/05/2015	13:52:51	COND HEAD&SHOULDERS	PERFUMARIA
028042	13/07/2013	13:51:17	COND KANECHOMN KIDS	PERFUMARIA
31727	01/03/2014	14:50:40	COND MONANG AMARELO	FARMACIA
072611	03/06/2014	10:53:10	COND MONANG AMARELO	FARMACIA
028042	13/07/2013	13:51:17	COND MONANGE ANTIFR	PERFUMARIA
16582	11/08/2013	08:56:16	COND MONANGE ANTIFR	PERFUMARIA
24694	24/11/2013	10:41:55	COND MONANGE ANTIFR	PERFUMARIA
41652	18/10/2014	11:45:08	COND MONANGE BRILH	HIGIENE
47695	28/03/2015	16:43:53	COND MONANGE BRILH	HIGIENE
036980	11/05/2014	13:16:57	COND MONANGE EXTRAP	HIGIENE

Figura 3.3: Itens do mesmo tipo com grupos diferentes.

Na Figura 3.3, é possível observar que, produtos notoriamente descritos em "Descrição Resumida" e, que claramente pertencem a um mesmo tipo, são representados por diferentes grupos, ou melhor, categorias diferentes. Já a Figura 3.4, demonstra o processo de caracterização descrito anteriormente, onde produtos que representam um mesmo tipo, passam a pertencer a uma mesma categoria.

Número da Venda	Data da Venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
49052	14/05/2015	13:52:51	COND HEAD&SHOULDERS	FARMACIA
028042	13/07/2013	13:51:17	COND KANECHOMN KIDS	FARMACIA
31727	01/03/2014	14:50:40	COND MONANG AMARELO	FARMACIA
072611	03/06/2014	10:53:10	COND MONANG AMARELO	FARMACIA
028042	13/07/2013	13:51:17	COND MONANGE ANTIFR	FARMACIA
16582	11/08/2013	08:56:16	COND MONANGE ANTIFR	FARMACIA
24694	24/11/2013	10:41:55	COND MONANGE ANTIFR	FARMACIA
41652	18/10/2014	11:45:08	COND MONANGE BRILH	FARMACIA
47695	28/03/2015	16:43:53	COND MONANGE BRILH	FARMACIA
036980	11/05/2014	13:16:57	COND MONANGE EXTRAP	FARMACIA

Figura 3.4: Itens do mesmo tipo numa mesma categoria.

Para o atributo "Descrição Resumida", foi feito primeiramente uma limpeza para a verificação de inconsistências na base. Dessa forma, foram encontradas amostras que possuíam nomes diferentes, mas que se referiam a um mesmo produto conforme exemplificado na Figura 3.5.

Número da Venda	Data da Venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
40099	04/09/2014	18:09:07	ACH NESCAL ACTIGEN	MERCEARIA
076159	05/09/2014	09:11:35	ACH NESCAL ACTIGEN	MERCEARIA
41162	04/10/2014	14:41:28	ACH NESCAL ACTIGEN	MERCEARIA
42378	08/11/2014	12:15:44	ACH NESCAU 2.0 SACH	MERCEARIA
42378	08/11/2014	12:15:44	ACH NESCAU 2.0 SACH	MERCEARIA
21219	12/10/2013	10:24:35	ACHOC EMBAREKIDS	MERCEARIA
21148	11/10/2013	16:11:41	ACHOC EMBAREKIDS	MERCEARIA
031691	15/10/2013	17:48:16	ACHOC EMBAREKIDS	MERCEARIA
21509	16/10/2013	18:51:04	ACHOC EMBAREKIDS	MERCEARIA
21509	16/10/2013	18:51:04	ACHOC EMBAREKIDS	MERCEARIA

Figura 3.5: Itens iguais com nomes diferentes.

Na Figura 3.5, é possível verificar que o mesmo produto foi representado de formas diferentes, através dos produtos com os seguintes nomes “ACH NESCAL ACTIGEN” e “ACH NESCAL 2.0 SACH”. Com isso, buscou-se padronizar todos os produtos da base, com intuito de aumentar a qualidade do estudo em questão, visto que passaria a representar unicamente o produto. Portanto, foi feita uma varredura na base de dados, afim de se obter tipos de produtos genéricos, com a finalidade de descrever facilmente a qual tipo aquele determinado produto representaria. Para contextualizar o que foi dito, pode ser observado como pequenos exemplos, o resultado do processo aplicado na Figura 3.6, com relação ao que foi apresentado na Figura 3.5.

Número da Venda	Data da Venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
40099	04/09/2014	18:09:07	ACHOCOLATADO	MERCEARIA
076159	05/09/2014	09:11:35	ACHOCOLATADO	MERCEARIA
41162	04/10/2014	14:41:28	ACHOCOLATADO	MERCEARIA
42378	08/11/2014	12:15:44	ACHOCOLATADO	MERCEARIA
42378	08/11/2014	12:15:44	ACHOCOLATADO	MERCEARIA
21219	12/10/2013	10:24:35	ACHOCOLATADO	MERCEARIA
21148	11/10/2013	16:11:41	ACHOCOLATADO	MERCEARIA
031691	15/10/2013	17:48:16	ACHOCOLATADO	MERCEARIA
21509	16/10/2013	18:51:04	ACHOCOLATADO	MERCEARIA
21509	16/10/2013	18:51:04	ACHOCOLATADO	MERCEARIA

Figura 3.6: Itens com nomes iguais.

Também foi observado que, além dos produtos terem nomes diferentes, os quais representavam um mesmo tipo de produto, alguns produtos vinham descritos juntamente com sua marca. Como não é de interesse do estudo identificar qual marca é a mais vendida, e sim qual tipo de produto é mais consumido em diferentes épocas do ano, foi necessário fazer um processo para a tipificação dos produtos contidos na base. Essa tipificação, teve como o intuito, assim como foi apresentado em etapas anteriores, a representação dos

produtos em um tipo único. Essa tipificação, não levou em conta a marca do produto, e sim o que ele representa. Este processo pode ser observado nas figuras 3.7 e 3.8.

Número da Venda	Data da Venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
43475	05/12/2014	09:11:51	ARROZ CAMERA 5KG	MERCEARIA
079747	18/12/2014	13:52:53	ARROZ CAMERA 5KG	MERCEARIA
44273	23/12/2014	10:36:06	ARROZ CAMERA 5KG	MERCEARIA
44858	02/01/2015	10:13:10	ARROZ CAMERA 5KG	MERCEARIA
44876	02/01/2015	11:58:42	ARROZ CAMERA 5KG	MERCEARIA
067214	27/06/2013	12:56:34	ARROZ CAMIL 1K	MERCEARIA
026874	06/06/2013	13:22:52	ARROZ CAMIL 1K	MERCEARIA
066614	15/06/2013	11:46:38	ARROZ CAMIL 1K	MERCEARIA
066938	20/06/2013	18:31:01	ARROZ CAMIL 1K	MERCEARIA
027564	29/06/2013	11:54:53	ARROZ CAMIL 1K	MERCEARIA
067354	01/07/2013	13:02:50	ARROZ CAMIL 1K	MERCEARIA
027870	08/07/2013	19:39:24	ARROZ CAMIL 1K	MERCEARIA
14874	17/07/2013	18:36:20	ARROZ CAMIL 1K	MERCEARIA
068760	07/02/2014	15:07:41	ARROZ CARIJO 5KG	MERCEARIA
30732	13/02/2014	08:27:03	ARROZ CARIJO 5KG	MERCEARIA
30733	13/02/2014	08:29:25	ARROZ CARIJO 5KG	MERCEARIA
30615	09/02/2014	10:00:59	ARROZ CARIJO 5KG	MERCEARIA
068953	11/02/2014	14:46:08	ARROZ CARIJO 5KG	MERCEARIA
069002	12/02/2014	10:11:40	ARROZ CARIJO 5KG	MERCEARIA

Figura 3.7: Itens de um mesmo tipo com marcas diferentes.

Na Figura 3.7, pode-se observar que o produto "ARROZ" é descrito na base juntamente com sua marca. A limpeza efetuada, se deu a partir da transformação dessas descrições, em descrições de um tipo único, conforme descrito anteriormente, e que pode ser observado na Figura 3.8.

Número da Venda	Data da Venda	Hora da Venda	Descricao Reduzida	Nome do Grupo
43475	05/12/2014	09:11:51	ARROZ	MERCEARIA
079747	18/12/2014	13:52:53	ARROZ	MERCEARIA
44273	23/12/2014	10:36:06	ARROZ	MERCEARIA
44858	02/01/2015	10:13:10	ARROZ	MERCEARIA
44876	02/01/2015	11:58:42	ARROZ	MERCEARIA
067214	27/06/2013	12:56:34	ARROZ	MERCEARIA
026874	06/06/2013	13:22:52	ARROZ	MERCEARIA
066614	15/06/2013	11:46:38	ARROZ	MERCEARIA
066938	20/06/2013	18:31:01	ARROZ	MERCEARIA
027564	29/06/2013	11:54:53	ARROZ	MERCEARIA
067354	01/07/2013	13:02:50	ARROZ	MERCEARIA
027870	08/07/2013	19:39:24	ARROZ	MERCEARIA
14874	17/07/2013	18:36:20	ARROZ	MERCEARIA
068760	07/02/2014	15:07:41	ARROZ	MERCEARIA
30732	13/02/2014	08:27:03	ARROZ	MERCEARIA
30733	13/02/2014	08:29:25	ARROZ	MERCEARIA
30615	09/02/2014	10:00:59	ARROZ	MERCEARIA
068953	11/02/2014	14:46:08	ARROZ	MERCEARIA
069002	12/02/2014	10:11:40	ARROZ	MERCEARIA
30821	14/02/2014	15:38:38	ARROZ	MERCEARIA

Figura 3.8: Itens de acordo com tipo.

Ao final de todas as etapas demonstradas anteriormente, foi obtido como resultado as seguintes informações a respeito dos produtos e, conseqüentemente, de suas categorias:

- Foram comprados 305.391 itens dispostos em um período de 10 trimestres
- Ao todo existem 157 tipos diferentes de produtos que foram comercializados em oito categorias diferentes

Observa-se, vide Tabela 3.2, a quantidade dos produtos mais comprados durante o período de estudo. É esperado que eles se façam presente na grande maioria das regras de associação que irão ser geradas na etapa de mineração de dados.

3.1.3 Mineração de Dados

Nesta etapa, foi utilizada a ferramenta WEKA (Waikato Environment for Knowledge Analysis), que é uma das ferramentas que podem ser utilizadas para a extração de conhe-

Tabela 3.2: Produtos com maior frequência de compras

Produto	Quantidade
Biscoito	23944
Legumes	21908
Produtos Limpeza	17973
Refrigerante	14250
Pães	11141
Sucos	10050
Frutas	9415
Queijos	7835
Limpeza Roupas	7613
Macarrão	7333

cimento. Através dessa ferramenta, busca-se explorar a utilização do algoritmo APRIORI, o qual foi mencionado anteriormente. Dessa forma, espera-se encontrar padrões nas compras realizadas, onde de fato seria humanamente impossível de se detectar, pois conforme Tanna and Ghodasara (2014), as técnicas de mineração de dados foram criadas para serem utilizadas e servirem de auxílio na detecção de padrões em grandes conjuntos de dados.

Nesta fase é importante salientar todo o processo envolvido anteriormente, principalmente a etapa de pré-processamento dos dados. Para que se possa conseguir bons resultados, deve-se garantir a qualidade do que foi feito anteriormente, pois dessa forma, será garantido que o processo de mineração de dados não apresentará dados inconsistentes, ou, como é de objetivo desse trabalho, a apresentação errônea de itens e suas devidas associações para posterior interpretação.

Com o escopo definido, e a base de dados pronta para a utilização após o processo de limpeza, será aplicada a técnica de mineração de dados, utilizando a ferramenta WEKA.

3.2 Conhecimentos Adquiridos

Nesta seção, será apresentado o conhecimento adquirido através de análises feitas a partir das informações encontradas pela ferramenta WEKA, com base na relação dos dados de entrada referente aos 10 trimestres especificados ao longo do trabalho. Para isso, será dividido e apresentado cada trimestre separadamente, apresentando uma análise pontual para determinado arquivo de entrada.

A princípio, foram feitos vários testes com os arquivos de entradas gerados para

o estudo, com intuito de se fazer uma verificação de quais parâmetros associados ao algoritmo APRIORI iriam gerar bons resultados, ou melhor, boas regras. Essas regras, posteriormente seriam analisadas para uma eventual tomada de decisão.

Foram feitos vários testes utilizando valores distintos para os parâmetros de confiança e suporte mínimo. Esses parâmetros foram descritos anteriormente e o local para defini-los na ferramenta Weka pode ser visto na Figura 3.9. Conforme dito anteriormente, eles são responsáveis por dizer o percentual mínimo de correlação que cada produto tem que ter em relação aos outros, o que foi denominado como suporte, e também a confiança necessária da correlação que foi obtida, que foi denominado como confiança mínima.

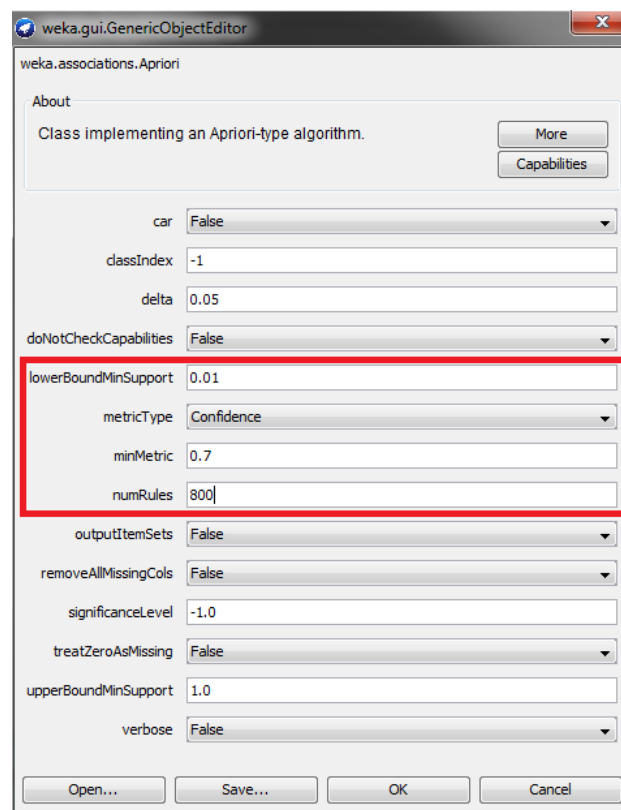


Figura 3.9: Parâmetros Algoritmo Apriori.

Na Tabela 3.3, é dada uma breve descrição dos parâmetros observados na Figura 3.9.

Cada parâmetro é responsável por uma configuração que se deseja aplicar para a obtenção de regras de associações. Embora haja métricas, como lift, conviction e leverage, que também podem definir o grau de interesse de uma regra baseada nas relações

dos produtos contidos na base de dados, para este trabalho será utilizado como métrica denominado na ferramenta por "metricType", o parâmetro de confiança (confidence).

Tabela 3.3: Parâmetros de entrada para o processo de mineração.

Parâmetros	Descrição
car	Se ativado, irá minerar as classes das regras de associações.
delta	Reduz o suporte estipulado até que o limite inferior seja atingido.
doNotCheckCapabilities	Se ativado, a capacidade de associação não serão verificadas.
lowerBoundMinSupport	Suporte mínimo a ser definido.
metricType	Métrica que validará as regras. Confiança será utilizada.
minMetric	Valor mínimo para a métrica adotada.
numRules	Número de regras a serem geradas.
outputItemSets	Se ativado, exibe os conjuntos de itemsets encontrados.
removeAllMissingCols	Remove coluna sem informações.
significanceLevel	Nível de significância.
treatZeroAsMissing	Se ativado, trata 0 como valor insignificante.
upperBoundMinSupport	Limite superior do suporte.
Verbose	Não foi encontrada definição, manter padrão.

Na execução de testes para a padronização dos parâmetros a serem utilizados neste estudo, foi definido o uso do suporte mínimo (`lowerBoundMinSupport`), com o valor de 1%, e a confiança mínima (`minMetric`), com o valor de 70%. Também foi padronizado a quantidade de regras que deveriam ser geradas, para isso foi estipulado um valor de 800 regras. O restante dos parâmetros apresentados pela Figura 3.9, foram deixados com os valores padrão da ferramenta.

O valor de suporte mínimo adotado é baixo, mas pode ser explicado devido à base de dados possuir um número pequeno de transações efetivas, além de uma grande variedade de produtos distintos. Portanto, o valor de 1% empregado ao suporte, irá referir-se a seleção de produtos que em conjunto com outros produtos, aparecem juntos em pelo menos 1% de todas as transações (Vendas) efetuadas para o determinado conjunto de dados estipulado no período. Então, por exemplo, caso a base de dados contenha informações

a respeito de 100 transações de compras efetuadas e nela contenha a aparição de relacionamento entre produtos que formam um conjunto, como exemplo "ARROZ" → "FEIJAO" em apenas uma das 100 transações, será obtido um suporte mínimo de 1%. Dentre as 100 transações, onde cada transação contém um conjunto de itens, deverá ser verificado o número de transações em que aparecem o produto "ARROZ". Supondo, que nas 100 transações, o produto "ARROZ" apareça 2 vezes, a confiança será obtida pelo número de vezes em que aparece a relação "ARROZ" → "FEIJAO", dividida pelo total de aparição do produto "ARROZ", conforme explicado na seção 2.2.2. Sendo assim, será obtido como confiança para a relação "ARROZ" → "FEIJAO", um valor de 50% das amostras. Este valor indicará que 50% dos compradores que compraram "ARROZ", também compraram "FEIJAO".

3.2.1 Resultados

Conforme explicado anteriormente, todos os resultados aqui demonstrados seguiram um padrão de configuração para a ferramenta utilizada. A partir dessa premissa, foram obtidos os resultados das compras realizadas no mercado por trimestre, num intervalo de tempo de dois anos e meio. Cada compra é representada por uma linha no arquivo de entrada da ferramenta, a qual contém informações de cada produto que está presente ou não na compra, seguindo o que foi explicitado na apresentação da seção 2.2.5.

Serão apenas demonstradas alguns exemplos das regras geradas pela ferramenta, pois não é o intuito do trabalho representar todas visto que foram apresentados os padrões utilizados para a geração das regras para cada trimestre de estudo, dessa forma, todas as regras geradas para cada arquivo serão analisadas e de forma que seja possível identificar padrões de consumo dos clientes para o determinado trimestre de estudo.

3.2.2 Trimestre 1

Para o primeiro trimestre tem-se um total de 9454 compras realizadas. Esse primeiro trimestre é correspondente ao período de Janeiro/2013 até final de Março/2013.

A Figura 3.10, apresenta as regras de associação produzidas para os itens comprados nesse período.

Best rules found:

```

1. CREME_DE_LEITE=1 273 ==> LEITE=1 273 <conf:(1)> lift:(8.62) lev:(0.03) [241] conv:(241.32)
2. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 99 ==> LEITE=1 99 <conf:(1)> lift:(8.62) lev:(0.01) [87] conv:(87.51)
3. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 90 ==> LEITE=1 90 <conf:(1)> lift:(8.62) lev:(0.01) [79] conv:(79.56)
4. BISCOITO=1 CREME_DE_LEITE=1 89 ==> LEITE=1 89 <conf:(1)> lift:(8.62) lev:(0.01) [78] conv:(78.67)
5. CREME_DE_LEITE=1 QUEIJOS=1 85 ==> LEITE=1 85 <conf:(1)> lift:(8.62) lev:(0.01) [75] conv:(75.14)
6. ACUCAR=1 CAFE=1 FEIJAO=1 FUBA=1 83 ==> ARROZ=1 83 <conf:(1)> lift:(25.21) lev:(0.01) [79] conv:(79.71)
7. CREME_DE_LEITE=1 REFRIGERANTE=1 82 ==> LEITE=1 82 <conf:(1)> lift:(8.62) lev:(0.01) [72] conv:(72.49)
8. CREME_DE_LEITE=1 SUCO=1 81 ==> LEITE=1 81 <conf:(1)> lift:(8.62) lev:(0.01) [71] conv:(71.6)
9. ACUCAR=1 BISCOITO=1 FEIJAO=1 FUBA=1 77 ==> ARROZ=1 77 <conf:(1)> lift:(25.21) lev:(0.01) [73] conv:(73.95)
10. ACUCAR=1 CAFE=1 FEIJAO=1 FUBA=1 PROD_LIMPEZA=1 76 ==> ARROZ=1 76 <conf:(1)> lift:(25.21) lev:(0.01) [72] conv:(72.99)
11. CREME_DE_LEITE=1 PROD_LIMPEZA=1 75 ==> LEITE=1 75 <conf:(1)> lift:(8.62) lev:(0.01) [66] conv:(66.3)
12. EXTRATO_DE_TOMATE=1 MACARRAO=1 PAPEL_HIGIENICO=1 75 ==> PROD_LIMPEZA=1 75 <conf:(1)> lift:(7.74) lev:(0.01) [65] conv:(65.31)
13. ACUCAR=1 EXTRATO_DE_TOMATE=1 PAPEL_HIGIENICO=1 73 ==> PROD_LIMPEZA=1 73 <conf:(1)> lift:(7.74) lev:(0.01) [63] conv:(63.57)
14. CAFE=1 FEIJAO=1 FUBA=1 MACARRAO=1 73 ==> ARROZ=1 73 <conf:(1)> lift:(25.21) lev:(0.01) [70] conv:(70.1)
15. ACUCAR=1 BISCOITO=1 CAFE=1 FEIJAO=1 FUBA=1 72 ==> ARROZ=1 72 <conf:(1)> lift:(25.21) lev:(0.01) [69] conv:(69.14)
16. ACUCAR=1 CAFE=1 FEIJAO=1 FUBA=1 MACARRAO=1 72 ==> ARROZ=1 72 <conf:(1)> lift:(25.21) lev:(0.01) [69] conv:(69.14)
17. ACUCAR=1 BISCOITO=1 FEIJAO=1 FUBA=1 PROD_LIMPEZA=1 70 ==> ARROZ=1 70 <conf:(1)> lift:(25.21) lev:(0.01) [67] conv:(67.22)
18. CAFE=1 FEIJAO=1 FUBA=1 MACARRAO=1 PROD_LIMPEZA=1 70 ==> ARROZ=1 70 <conf:(1)> lift:(25.21) lev:(0.01) [67] conv:(67.22)
19. CREME_DE_LEITE=1 MOLHO=1 69 ==> LEITE=1 69 <conf:(1)> lift:(8.62) lev:(0.01) [60] conv:(60.99)
20. ACUCAR=1 CAFE=1 FEIJAO=1 FUBA=1 MACARRAO=1 PROD_LIMPEZA=1 69 ==> ARROZ=1 69 <conf:(1)> lift:(25.21) lev:(0.01) [66] conv:(66.26)
21. ARROZ=1 EXTRATO_DE_TOMATE=1 PAPEL_HIGIENICO=1 68 ==> PROD_LIMPEZA=1 68 <conf:(1)> lift:(7.74) lev:(0.01) [59] conv:(59.22)
22. CAFE=1 EXTRATO_DE_TOMATE=1 MACARRAO=1 PAPEL_HIGIENICO=1 68 ==> PROD_LIMPEZA=1 68 <conf:(1)> lift:(7.74) lev:(0.01) [59] conv:(59.22)
23. BISCOITO=1 CAFE=1 FEIJAO=1 FUBA=1 PROD_LIMPEZA=1 68 ==> ARROZ=1 68 <conf:(1)> lift:(25.21) lev:(0.01) [65] conv:(65.3)
24. ACUCAR=1 EXTRATO_DE_TOMATE=1 MACARRAO=1 PAPEL_HIGIENICO=1 67 ==> PROD_LIMPEZA=1 67 <conf:(1)> lift:(7.74) lev:(0.01) [58] conv:(58.35)
25. ACUCAR=1 BISCOITO=1 CAFE=1 FEIJAO=1 FUBA=1 PROD_LIMPEZA=1 67 ==> ARROZ=1 67 <conf:(1)> lift:(25.21) lev:(0.01) [64] conv:(64.34)
26. ARROZ=1 CAFE=1 FEIJAO=1 FUBA=1 84 ==> ACUCAR=1 83 <conf:(0.99)> lift:(19.54) lev:(0.01) [78] conv:(39.88)
27. CAFE=1 FEIJAO=1 FUBA=1 PROD_LIMPEZA=1 78 ==> ARROZ=1 77 <conf:(0.99)> lift:(24.89) lev:(0.01) [73] conv:(37.45)
28. ARROZ=1 CAFE=1 FEIJAO=1 FUBA=1 PROD_LIMPEZA=1 77 ==> ACUCAR=1 76 <conf:(0.99)> lift:(19.52) lev:(0.01) [72] conv:(36.55)
29. ACUCAR=1 CAFE=1 FUBA=1 PAPEL_HIGIENICO=1 75 ==> PROD_LIMPEZA=1 74 <conf:(0.99)> lift:(7.64) lev:(0.01) [64] conv:(32.66)
30. ARROZ=1 BISCOITO=1 FUBA=1 PROD_LIMPEZA=1 74 ==> ACUCAR=1 73 <conf:(0.99)> lift:(19.51) lev:(0.01) [69] conv:(35.13)

```

Figura 3.10: Resultados do Primeiro Trimestre.

Mediante a apresentação resumida de tais regras, é possível observar que, várias regras que contém PROD_LIMPEZA como consequente. Regras como:

ACUCAR FEIJAO, PAPEL_HIGIENICO → PROD_LIMPEZA

ARROZ FEIJAO, PAPEL_HIGIENICO → PROD_LIMPEZA

ACUCAR, MACARRAO, PAPEL_HIGIENICO → PROD_LIMPEZA=1

Tais regras, podem indicar fortemente que, quando os clientes compram o produto PAPEL_HIGIENICO, constuma-se ser realizado também a compra de PROD_LIMPEZA.

Outras Regras:

CAFE, ESPECIARIAS, FUBA → ARROZ

ACUCAR, BISCOITO, CAFE, FUBA → ARROZ

ACUCAR, BISCOITO, FEIJAO, FUBA → ARROZ

CAFE, FEIJAO, FUBA, PROD_LIMPEZA → ARROZ

ARROZ, CAFE, FUBA → ACUCAR

ARROZ, CAFE, FEIJAO → ACUCAR

ARROZ, CAFE, ESPECIARIAS → ACUCAR

ARROZ, CAFE, FEIJAO MACARRAO → ACUCAR

Percebe-se através das regras que, ARROZ e CAFE implicam diretamente na compra de ACUCAR, e que há indícios de que a compra de FUBA está diretamente relacionada com a compra do produto ARROZ.

A partir da observação de todas as regras geradas, foi percebido que os itens, ACUCAR, MACARRAO, PROD_LIMPEZA, CAFE, PAPEL_HIGIENICO, ARROZ e FEIJAO, eram os produtos que mais apareciam e, muitas das vezes, apareciam correlacionados entre si, mostrando a forte ligação que eles possuem na realização das compras do mercado para aquele trimestre. Portanto, pode-se presumir, a partir do que foi analisado, que produtos considerados óbvios, se tratando do consumo de produtos de necessidades básicas, irão aparecer com uma maior frequência ao longo dos trimestres que serão observados.

Fugindo da análise dos produtos que são de conhecimento comum, foi observado o relacionamento entre CREME_DE_LEITE e LEITE. A relação desses produtos traz um importante conhecimento adquirido para este trimestre, pois será possível analisar se o mesmo irá ocorrer, ao fazer a comparação dos trimestres que correspondem à mesma fase de um ano.

CREME_DE_LEITE → LEITE

CREME_DE_LEITE, LEGUMES_EM_LATA → LEITE

CREME_DE_LEITE, LEITE_CONDENSADO → LEITE

CREME_DE_LEITE, REFRIGERANTE → LEITE

CREME_DE_LEITE, SUCO → LEITE

Observou-se também, a relação encontrada entre CREME_DE_LEITE e produtos como: LEGUMES, LEITE_CONDENSADO, REFRIGERANTE e SUCO. Essas relações irão servir como parâmetro, afim de se observar a manutenção dos mesmos para algum trimestre posterior.

3.2.3 Trimestre 2

Para o trimestre 2, foi obtido um total de 8863 compras realizadas, para análise compreendida no intervalo ABRIL/2013 até JUNHO/2013.

Best rules found:

```

1. CREME_DE_LEITE=1 253 ==> LEITE=1 253 <conf:(1)> lift:(7.24) lev:(0.02) [218] conv:(218.03)
2. ACUCAR=1 LIMP_ROUPAS=1 MACARRAO=1 OLEO_VEGETAL=1 105 ==> PROD_LIMPEZA=1 105 <conf:(1)> lift:(8.18) lev:(0.01) [92] conv:(92.17)
3. ACUCAR=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 104 ==> PROD_LIMPEZA=1 104 <conf:(1)> lift:(8.18) lev:(0.01) [91] conv:(91.29)
4. ACUCAR=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PROD_BANHO=1 103 ==> PROD_LIMPEZA=1 103 <conf:(1)> lift:(8.18) lev:(0.01) [90] conv:(90.41)
5. CAFE=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 102 ==> PROD_LIMPEZA=1 102 <conf:(1)> lift:(8.18) lev:(0.01) [89] conv:(89.54)
6. FEIJAO=1 LIMP_ROUPAS=1 MACARRAO=1 OLEO_VEGETAL=1 101 ==> PROD_LIMPEZA=1 101 <conf:(1)> lift:(8.18) lev:(0.01) [88] conv:(88.66)
7. ACUCAR=1 ARROZ=1 LIMP_ROUPAS=1 PROD_BANHO=1 100 ==> PROD_LIMPEZA=1 100 <conf:(1)> lift:(8.18) lev:(0.01) [87] conv:(87.78)
8. ACUCAR=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PAPEL_HIGIENICO=1 100 ==> PROD_LIMPEZA=1 100 <conf:(1)> lift:(8.18) lev:(0.01) [87] conv:(87.78)
9. ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 PROD_BANHO=1 100 ==> PROD_LIMPEZA=1 100 <conf:(1)> lift:(8.18) lev:(0.01) [87] conv:(87.78)
10. CAFE=1 LIMP_ROUPAS=1 MACARRAO=1 OLEO_VEGETAL=1 100 ==> PROD_LIMPEZA=1 100 <conf:(1)> lift:(8.18) lev:(0.01) [87] conv:(87.78)
11. CAFE=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PROD_BANHO=1 100 ==> PROD_LIMPEZA=1 100 <conf:(1)> lift:(8.18) lev:(0.01) [87] conv:(87.78)
12. ACUCAR=1 LIMP_ROUPAS=1 MACARRAO=1 PROD_BANHO=1 99 ==> PROD_LIMPEZA=1 99 <conf:(1)> lift:(8.18) lev:(0.01) [86] conv:(86.9)
13. LIMP_ROUPAS=1 MACARRAO=1 OLEO_VEGETAL=1 PROD_BANHO=1 98 ==> PROD_LIMPEZA=1 98 <conf:(1)> lift:(8.18) lev:(0.01) [86] conv:(86.03)
14. LIMP_ROUPAS=1 MACARRAO=1 PAPEL_HIGIENICO=1 PROD_BANHO=1 98 ==> PROD_LIMPEZA=1 98 <conf:(1)> lift:(8.18) lev:(0.01) [86] conv:(86.03)
15. ACUCAR=1 ESPECIARIAS=1 LIMP_ROUPAS=1 PROD_BANHO=1 97 ==> PROD_LIMPEZA=1 97 <conf:(1)> lift:(8.18) lev:(0.01) [85] conv:(85.15)
16. ARROZ=1 LIMP_ROUPAS=1 MACARRAO=1 PROD_BANHO=1 97 ==> PROD_LIMPEZA=1 97 <conf:(1)> lift:(8.18) lev:(0.01) [85] conv:(85.15)
17. ACUCAR=1 BISCOITO=1 LIMP_ROUPAS=1 PROD_BANHO=1 96 ==> PROD_LIMPEZA=1 96 <conf:(1)> lift:(8.18) lev:(0.01) [84] conv:(84.27)
18. LIMP_ROUPAS=1 OLEO_VEGETAL=1 PAPEL_HIGIENICO=1 PROD_BANHO=1 96 ==> PROD_LIMPEZA=1 96 <conf:(1)> lift:(8.18) lev:(0.01) [84] conv:(84.27)
19. ACUCAR=1 ESPECIARIAS=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 95 ==> PROD_LIMPEZA=1 95 <conf:(1)> lift:(8.18) lev:(0.01) [83] conv:(83.39)
20. ACUCAR=1 ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 95 ==> PROD_LIMPEZA=1 95 <conf:(1)> lift:(8.18) lev:(0.01) [83] conv:(83.39)
21. ACUCAR=1 ARROZ=1 LIMP_ROUPAS=1 MACARRAO=1 OLEO_VEGETAL=1 95 ==> PROD_LIMPEZA=1 95 <conf:(1)> lift:(8.18) lev:(0.01) [83] conv:(83.39)
22. ACUCAR=1 CAFE=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 95 ==> PROD_LIMPEZA=1 95 <conf:(1)> lift:(8.18) lev:(0.01) [83] conv:(83.39)
23. ACUCAR=1 CAFE=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PROD_BANHO=1 95 ==> PROD_LIMPEZA=1 95 <conf:(1)> lift:(8.18) lev:(0.01) [83] conv:(83.39)
24. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 94 ==> LEITE=1 94 <conf:(1)> lift:(7.24) lev:(0.01) [81] conv:(81.01)
25. CAFE=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PAPEL_HIGIENICO=1 94 ==> PROD_LIMPEZA=1 94 <conf:(1)> lift:(8.18) lev:(0.01) [82] conv:(82.51)
26. ACUCAR=1 CAFE=1 LIMP_ROUPAS=1 MACARRAO=1 OLEO_VEGETAL=1 94 ==> PROD_LIMPEZA=1 94 <conf:(1)> lift:(8.18) lev:(0.01) [82] conv:(82.51)
27. ARROZ=1 BISCOITO=1 LIMP_ROUPAS=1 PROD_BANHO=1 93 ==> PROD_LIMPEZA=1 93 <conf:(1)> lift:(8.18) lev:(0.01) [81] conv:(81.64)
28. ACUCAR=1 ARROZ=1 CAFE=1 LIMP_ROUPAS=1 PROD_BANHO=1 93 ==> PROD_LIMPEZA=1 93 <conf:(1)> lift:(8.18) lev:(0.01) [81] conv:(81.64)
29. ACUCAR=1 ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 PROD_BANHO=1 92 ==> PROD_LIMPEZA=1 92 <conf:(1)> lift:(8.18) lev:(0.01) [80] conv:(80.76)
30. ACUCAR=1 CAFE=1 LIMP_ROUPAS=1 MACARRAO=1 PROD_BANHO=1 92 ==> PROD_LIMPEZA=1 92 <conf:(1)> lift:(8.18) lev:(0.01) [80] conv:(80.76)

```

Figura 3.11: Amostras dos Resultados do Segundo Trimestre

Para este trimestre, assim como no anterior, existe uma grande quantidade de produtos relacionados à PROD_LIMPEZA. Foi observado que quando um dos produtos, ARROZ, CAFE, FEIJAO, LIMP_ROUPAS está presente no cesto, na maioria dos casos, obtém-se a compra de PROD_LIMPEZA. Tal fato pode ser observado pelas regras:

LEGUMES, LIMP_ROUPAS, OLEO_VEGETAL → PROD_LIMPEZA

ACUCAR, FEIJAO, MACARRAO, PROD_BANHO → PROD_LIMPEZA

CAFE, MACARRAO, PAPEL_HIGIENICO, PROD_BANHO → PROD_LIMPEZA

ARROZ, OLEO_VEGETAL, PAPEL_HIGIENICO, PROD_BANHO → PROD_LIMPEZA

Também foi percebido que produtos comprados em combinação, como ARROZ, CAFE e FEIJAO, seguindo o trimestre anterior, levam à compra de ACUCAR, como pode ser observado nas regras:

ARROZ, CAFE, ESPECIARIAS, FEIJAO → ACUCAR

ARROZ, CAFE, FEIJAO, PAPEL_HIGIENICO → ACUCAR

Algumas relações diferentes do conhecimento óbvio, que constata a compra de produtos considerados de necessidades básicas, e que foram observadas no primeiro trimestre, também foram encontradas nesse segundo semestre, como são os casos de:

CREME.DE.LEITE \rightarrow LEITE

CREME.DE.LEITE, LEGUMES_EM.LATA \rightarrow LEITE

A partir da relação dos produtos encontrados anteriormente, juntamente com a análise feita para este segundo trimestre, foi possível observar que, quem comprou um dos produtos ARROZ, CAFE, FEIJAO, LIMP_ROUPAS também comprou PROD.LIMPEZA. Essa regra foi mantida, porém, foi observado que para o segundo trimestre houve a adição de mais dois produtos, no caso, PROD.BANHO e PAPEL.HIGIENICO, que implicavam na compra de PROD.LIMPEZA, bastava a ocorrência de um deles para se obter o resultado mencionado. Essa relação pode ser logicamente denotada por: $ARROZ \vee CAFE \vee FEIJAO \vee LIMP_ROUPAS \vee PROD_BANHO \vee PAPEL_HIGIENICO \rightarrow PROD.LIMPEZA$

A combinação de dois ou mais produtos entre, ARROZ, CAFE, OLEO VEGETAL e PROD.LIMPEZA, gerou como conseqüente o produto ACUCAR.

A manutenção de algumas regras presentes no primeiro trimestre, como: "CREME.DE.LEITE, LEGUMES_EM.LATA \rightarrow LEITE" e "CREME.DE.LEITE \rightarrow LEITE", que fogem da necessidade de consumo básico apresentadas até aqui, adiciona um conhecimento muito importante, pois foi possível notar um comportamento de compra similar ao do trimestre passado com relação a compra desses produtos. Essa comparação é válida pois auxilia na busca de informações ao se fazer uma análise para um trimestre posterior, e que deverá estar contido na mesma faixa de tempo dos trimestres aqui já avaliados

3.2.4 Trimestre 3

Para o trimestre 3, foi obtido um total de 9112 compras realizadas para análise do efetivo de compra do período compreendido em JULHO/2013 até SETEMBRO/2013.

Best rules found:

```

1. LEITE_CONDENSADO=1 320 ==> LEITE=1 320 <conf:(1)> lift:(5.94) lev:(0.03) [266] conv:(266.09)
2. LEGUMES_EM_LATA=1 308 ==> LEGUMES=1 308 <conf:(1)> lift:(5.37) lev:(0.03) [250] conv:(250.67)
3. CREME_DE_LEITE=1 284 ==> CREME=1 284 <conf:(1)> lift:(22.12) lev:(0.03) [271] conv:(271.16)
4. CREME_DE_LEITE=1 284 ==> LEITE=1 284 <conf:(1)> lift:(5.94) lev:(0.03) [236] conv:(236.16)
5. CREME_DE_LEITE=1 LEITE=1 284 ==> CREME=1 284 <conf:(1)> lift:(22.12) lev:(0.03) [271] conv:(271.16)
6. CREME_DE_LEITE=1 CREME=1 284 ==> LEITE=1 284 <conf:(1)> lift:(5.94) lev:(0.03) [236] conv:(236.16)
7. CREME_DE_LEITE=1 284 ==> CREME=1 LEITE=1 284 <conf:(1)> lift:(29.21) lev:(0.03) [274] conv:(274.28)
8. COMIDA_CONGELADA=1 280 ==> GEL=1 280 <conf:(1)> lift:(23.79) lev:(0.03) [268] conv:(268.23)
9. CREME_DE_LEITE=1 LEGUMES=1 133 ==> CREME=1 133 <conf:(1)> lift:(22.12) lev:(0.01) [126] conv:(126.99)
10. CREME_DE_LEITE=1 LEGUMES=1 133 ==> LEITE=1 133 <conf:(1)> lift:(5.94) lev:(0.01) [110] conv:(110.59)
11. LEGUMES_EM_LATA=1 LEITE=1 133 ==> LEGUMES=1 133 <conf:(1)> lift:(5.37) lev:(0.01) [108] conv:(108.24)
12. CREME_DE_LEITE=1 LEGUMES=1 LEITE=1 133 ==> CREME=1 133 <conf:(1)> lift:(22.12) lev:(0.01) [126] conv:(126.99)
13. CREME_DE_LEITE=1 CREME=1 LEGUMES=1 133 ==> LEITE=1 133 <conf:(1)> lift:(5.94) lev:(0.01) [110] conv:(110.59)
14. CREME_DE_LEITE=1 LEGUMES=1 133 ==> CREME=1 LEITE=1 133 <conf:(1)> lift:(29.21) lev:(0.01) [128] conv:(128.45)
15. BISCOITO=1 LEITE_CONDENSADO=1 116 ==> LEITE=1 116 <conf:(1)> lift:(5.94) lev:(0.01) [96] conv:(96.46)
16. CREME=1 LEITE_CONDENSADO=1 116 ==> LEITE=1 116 <conf:(1)> lift:(5.94) lev:(0.01) [96] conv:(96.46)
17. CREME=1 LEGUMES_EM_LATA=1 112 ==> LEGUMES=1 112 <conf:(1)> lift:(5.37) lev:(0.01) [91] conv:(91.15)
18. LEGUMES=1 LEITE_CONDENSADO=1 111 ==> LEITE=1 111 <conf:(1)> lift:(5.94) lev:(0.01) [92] conv:(92.3)
19. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 105 ==> CREME=1 105 <conf:(1)> lift:(22.12) lev:(0.01) [100] conv:(100.25)
20. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 105 ==> LEITE=1 105 <conf:(1)> lift:(5.94) lev:(0.01) [87] conv:(87.31)
21. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 LEITE=1 105 ==> CREME=1 105 <conf:(1)> lift:(22.12) lev:(0.01) [100] conv:(100.25)
22. CREME_DE_LEITE=1 CREME=1 LEITE_CONDENSADO=1 105 ==> LEITE=1 105 <conf:(1)> lift:(5.94) lev:(0.01) [87] conv:(87.31)
23. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 105 ==> CREME=1 LEITE=1 105 <conf:(1)> lift:(29.21) lev:(0.01) [101] conv:(101.4)
24. CREME=1 LEGUMES_EM_LATA=1 LEITE=1 105 ==> LEGUMES=1 105 <conf:(1)> lift:(5.37) lev:(0.01) [85] conv:(85.46)
25. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 98 ==> CREME=1 98 <conf:(1)> lift:(22.12) lev:(0.01) [93] conv:(93.57)
26. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 98 ==> LEGUMES=1 98 <conf:(1)> lift:(5.37) lev:(0.01) [79] conv:(79.76)
27. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 98 ==> LEITE=1 98 <conf:(1)> lift:(5.94) lev:(0.01) [81] conv:(81.49)
28. LEGUMES_EM_LATA=1 QUEIJOS=1 98 ==> LEGUMES=1 98 <conf:(1)> lift:(5.37) lev:(0.01) [79] conv:(79.76)
29. CAFE=1 MACARRAO=1 PAPEL_HIGIENICO=1 98 ==> PROD_LIMPEZA=1 98 <conf:(1)> lift:(9.23) lev:(0.01) [87] conv:(87.38)
30. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 LEGUMES=1 98 ==> CREME=1 98 <conf:(1)> lift:(22.12) lev:(0.01) [93] conv:(93.57)

```

Figura 3.12: Amostras dos Resultados do Terceiro Trimestre

Para este trimestre, pode-se ver a manutenção de algumas regras que foram encontradas nos trimestres anteriores, e que não são de conhecimento comum, como pode-se observar como exemplo as regras:

CREME_DE_LEITE → LEITE

CREME_DE_LEITE, LEGUMES_EM_LATA → LEITE

Percebe-se cada vez mais a constatação de que produtos como CAFE, ARROZ, FEIJAO LIMP_ROUPAS, PROD_BANHO e PAPEL_HIGIENICO, geram como conseqüente a possível compra de PROD_LIMPEZA, como pode ser observado na regra 29 indicado na Figura 3.12, que diz:

ARROZ, FEIJAO, PAPEL_HIGIENICO → PROD_LIMPEZA

CAFE, FEIJAO, PAPEL_HIGIENICO → PROD_LIMPEZA

ARROZ, FEIJAO, LIMP_ROUPAS → PROD_LIMPEZA

Novas associações entre produtos foram descobertas nesse trimestre. Por exemplo a associação entre produtos como:

COMIDA_CONGELADA → GEL

COMIDA_CONGELADA, REFRIGERANTE → GEL

BISCOITO, COMIDA_CONGELADA → GEL

CREME.DE.LEITE → CREME

CREME, LEGUMES_EM.LATA → LEGUMES

CREME.DE.LEITE, CREME_QUEIJOS 93 → LEITE

Com isso, constata-se também o conhecimento adquirido a respeito desses novos produtos, onde é possível afirmar que, para este trimestre, a compra do produto COMIDA_CONGELADA implica na compra do produto GEL. Da mesma maneira, pode se entender que a compra do produto CREME, para este trimestre, implica na compra de produtos como, CREME.DE.LEITE, LEITE, LEGUMES e LEGUMES_EM.LATA.

Foi observado em certas regras, que a presença do produto ESPECIARIAS, caracterizado por temperos em geral, implica em muitas das vezes na compra de LEGUMES, essa constatação pode ser observada através das regras:

ESPECIARIAS, FEIJAO_MACARRAO → LEGUMES

CAFE, ESPECIARIAS, LIMP_ROUPAS → LEGUMES

CAFE, ESPECIARIAS, OLEO_VEGETAL → LEGUMES

CAFE, ESPECIARIAS, OLEO_VEGETAL → LEGUMES

BISCOITO, ESPECIARIAS, LIMP_ROUPAS → LEGUMES

ESPECIARIAS, MACARRAO, OLEO_VEGETAL → LEGUMES

ACUCAR, ESPECIARIAS, PROD_LIMPEZA → LEGUMES

BISCOITO, ESPECIARIAS, OLEO_VEGETAL → LEGUMES

ESPECIARIAS, OLEO_VEGETAL, PROD_LIMPEZA → LEGUMES

ESPECIARIAS, LIMP_ROUPAS, PAPEL_HIGIENICO → LEGUMES

3.2.5 Trimestre 4

Para o trimestre 4, foi obtido um total de 9125 compras realizadas para análise das compras realizadas no período correspondente a OUTUBRO/2013 até DEZEMBRO/2013.

Best rules found:

```

1. COTONETE=1 1870 ==> BISCOITO=1 1870 <conf:(1)> lift:(4.88) lev:(0.16) [1486] conv:(1486.78)
2. BISCOITO=1 1870 ==> COTONETE=1 1870 <conf:(1)> lift:(4.88) lev:(0.16) [1486] conv:(1486.78)
3. DESODORANTE=1 492 ==> CAFE=1 492 <conf:(1)> lift:(18.55) lev:(0.05) [465] conv:(465.47)
4. CAFE=1 492 ==> DESODORANTE=1 492 <conf:(1)> lift:(18.55) lev:(0.05) [465] conv:(465.47)
5. FARINHA=1 490 ==> CHOCOLATE=1 490 <conf:(1)> lift:(18.62) lev:(0.05) [463] conv:(463.69)
6. CHOCOLATE=1 490 ==> FARINHA=1 490 <conf:(1)> lift:(18.62) lev:(0.05) [463] conv:(463.69)
7. COTONETE=1 REFRIGERANTE=1 371 ==> BISCOITO=1 371 <conf:(1)> lift:(4.88) lev:(0.03) [294] conv:(294.97)
8. BISCOITO=1 REFRIGERANTE=1 371 ==> COTONETE=1 371 <conf:(1)> lift:(4.88) lev:(0.03) [294] conv:(294.97)
9. COTONETE=1 SUCO=1 334 ==> BISCOITO=1 334 <conf:(1)> lift:(4.88) lev:(0.03) [265] conv:(265.55)
10. BISCOITO=1 SUCO=1 334 ==> COTONETE=1 334 <conf:(1)> lift:(4.88) lev:(0.03) [265] conv:(265.55)
11. COTONETE=1 PAES=1 315 ==> BISCOITO=1 315 <conf:(1)> lift:(4.88) lev:(0.03) [250] conv:(250.45)
12. BISCOITO=1 PAES=1 315 ==> COTONETE=1 315 <conf:(1)> lift:(4.88) lev:(0.03) [250] conv:(250.45)
13. COTONETE=1 LEGUMES=1 295 ==> BISCOITO=1 295 <conf:(1)> lift:(4.88) lev:(0.03) [234] conv:(234.55)
14. BISCOITO=1 LEGUMES=1 295 ==> COTONETE=1 295 <conf:(1)> lift:(4.88) lev:(0.03) [234] conv:(234.55)
15. COTONETE=1 LEITE=1 289 ==> BISCOITO=1 289 <conf:(1)> lift:(4.88) lev:(0.03) [229] conv:(229.77)
16. BISCOITO=1 LEITE=1 289 ==> COTONETE=1 289 <conf:(1)> lift:(4.88) lev:(0.03) [229] conv:(229.77)
17. FERMENTO=1 281 ==> COMIDA_CONGELADA=1 281 <conf:(1)> lift:(32.47) lev:(0.03) [272] conv:(272.35)
18. COMIDA_CONGELADA=1 281 ==> FERMENTO=1 281 <conf:(1)> lift:(32.47) lev:(0.03) [272] conv:(272.35)
19. COTONETE=1 PROD_LIMPEZA=1 273 ==> BISCOITO=1 273 <conf:(1)> lift:(4.88) lev:(0.02) [217] conv:(217.05)
20. BISCOITO=1 PROD_LIMPEZA=1 273 ==> COTONETE=1 273 <conf:(1)> lift:(4.88) lev:(0.02) [217] conv:(217.05)
21. ACHOCOLATADO=1 COTONETE=1 214 ==> BISCOITO=1 214 <conf:(1)> lift:(4.88) lev:(0.02) [170] conv:(170.14)
22. ACHOCOLATADO=1 BISCOITO=1 214 ==> COTONETE=1 214 <conf:(1)> lift:(4.88) lev:(0.02) [170] conv:(170.14)
23. FRUTAS=1 206 ==> COMIDA_PET=1 206 <conf:(1)> lift:(44.3) lev:(0.02) [201] conv:(201.35)
24. COMIDA_PET=1 206 ==> FRUTAS=1 206 <conf:(1)> lift:(44.3) lev:(0.02) [201] conv:(201.35)
25. COTONETE=1 IOGURTE=1 206 ==> BISCOITO=1 206 <conf:(1)> lift:(4.88) lev:(0.02) [163] conv:(163.78)
26. BISCOITO=1 IOGURTE=1 206 ==> COTONETE=1 206 <conf:(1)> lift:(4.88) lev:(0.02) [163] conv:(163.78)
27. COTONETE=1 QUEIJOS=1 205 ==> BISCOITO=1 205 <conf:(1)> lift:(4.88) lev:(0.02) [162] conv:(162.99)
28. BISCOITO=1 QUEIJOS=1 205 ==> COTONETE=1 205 <conf:(1)> lift:(4.88) lev:(0.02) [162] conv:(162.99)
29. COTONETE=1 LIMP_ROUPAS=1 202 ==> BISCOITO=1 202 <conf:(1)> lift:(4.88) lev:(0.02) [160] conv:(160.6)
30. BISCOITO=1 LIMP_ROUPAS=1 202 ==> COTONETE=1 202 <conf:(1)> lift:(4.88) lev:(0.02) [160] conv:(160.6)

```

Figura 3.13: Amostras dos Resultados do Quarto Trimestre

Conforme a Figura 3.13, é possível notar que houve uma alteração no comportamento de compra para este trimestre. Novas relações entre produtos apareceram e algumas se mantiveram conforme análise a seguir.

COTONETE → BISCOITO

BISCOITO → COTONETE

DESODORANTE → CAFE

CAFE → DESODORANTE

FARINHA → CHOCOLATE

CHOCOLATE → FARINHA

COTONETE, REFRIGERANTE → BISCOITO

BISCOITO, REFRIGERANTE → COTONETE

COTONETE, SUCO 4 → BISCOITO

BISCOITO, SUCO → COTONETE

COTONETE, PAES → BISCOITO

BISCOITO, PAES → COTONETE

COTONETE, LEGUMES → BISCOITO

BISCOITO, LEGUMES → COTONETE

COTONETE, LEITE → BISCOITO

BISCOITO, LEITE → COTONETE

FERMENTO, → COMIDA_CONGELADA

COMIDA_CONGELADA → FERMENTO

FRUTAS → COMIDA_PET

COMIDA_PET → FRUTAS

ESPUMA_DE_BARBEAR → CEREAL

CEREAL → ESPUMA_DE_BARBEAR

Algumas das regras apresentadas fornecem informações diretas sobre o comportamento de compras de determinado produto. Elas mostram a relação direta entre dois produtos e, o que pode-se esperar é que a aparição de um determinado produto implicará diretamente na compra de outro, como demonstrado anteriormente nas regras. O surgimento dessas regras, demonstram até então a relação entre produtos bem distintos categoricamente, e que podem ser explicados devido ao baixo índice de suporte estabelecido. Como exemplo, tem-se a seguinte regra: DESODORANTE, MACARRAO, PROD_LIMPEZA → CAFE.

Regras definidas anteriormente para a determinação da compra de produtos PROD_LIMPEZA, foram novamente encontradas neste trimestre, o que caracteriza a consolidação do conhecimento adquirido sobre os grupos de produtos que ao serem comprados implicam na compra de PROD_LIMPEZA, como pode ser observado:

CAFE, LIMP_ROUPAS, PAPEL_HIGIENICO, PROD_BANHO → PROD_LIMPEZA

ARROZ, LIMP_ROUPAS → PROD_LIMPEZA

Regras consolidadas presentes nos três trimestres anteriores não foram encontradas para este trimestre, como é o caso da regra de associação: CREME_DE_LEITE → LEITE.

A implicação de compras de LEGUMES encontrada no trimestre anterior não manteve o padrão especificado anteriormente, assim como regras que envolviam os produtos CREME_DE_LEITE, LEITE e LEGUMES_EM_LATA.

A partir dos resultados obtidos para este primeiro ano, foi observado que a maioria das regras encontradas estavam ligadas a produtos de limpeza (PROD.LIMPEZA). Conforme citado anteriormente, a aparição de produtos como ARROZ, CAFE, FEIJAO, LIMP.ROUPAS, PROD.BANHO e PAPEL.HIGIENICO, por exemplo, constata que há um padrão nas compras realizadas pelos clientes para o mercado em questão. Essas compras, são consideradas neste trabalho, como produtos de uso básico da sociedade, e que validam o estudo até então. É de interesse analisar os outros trimestres restantes, afim de se obter relações entre produtos que não são de conhecimento óbvio, afim de se adquirir conhecimentos necessários que irão ajudar em posteriores tomadas de decisão.

3.2.6 Trimestre 5

Para o trimestre 5, foi obtido um total de 8027 compras realizadas para análise compreendidas no período de JANEIRO/2014 até MARÇO/2014.

Best rules found:

```

1. COMIDA_PET=1 569 ==> BEBIDA_ALCOOLICA=1 569 <conf:(1)> lift:(14.11) lev:(0.07) [528] conv:(528.67)
2. BEBIDA_ALCOOLICA=1 569 ==> COMIDA_PET=1 569 <conf:(1)> lift:(14.11) lev:(0.07) [528] conv:(528.67)
3. CAFE=1 392 ==> ACUCAR=1 392 <conf:(1)> lift:(20.48) lev:(0.05) [372] conv:(372.86)
4. ACUCAR=1 392 ==> CAFE=1 392 <conf:(1)> lift:(20.48) lev:(0.05) [372] conv:(372.86)
5. BOMBOM=1 374 ==> ACHOCOLATADO=1 374 <conf:(1)> lift:(21.46) lev:(0.04) [356] conv:(356.57)
6. ACHOCOLATADO=1 374 ==> BOMBOM=1 374 <conf:(1)> lift:(21.46) lev:(0.04) [356] conv:(356.57)
7. CHA=1 316 ==> ARROZ=1 316 <conf:(1)> lift:(25.4) lev:(0.04) [303] conv:(303.56)
8. ARROZ=1 316 ==> CHA=1 316 <conf:(1)> lift:(25.4) lev:(0.04) [303] conv:(303.56)
9. CHOCOLATE=1 265 ==> BALA=1 265 <conf:(1)> lift:(30.29) lev:(0.03) [256] conv:(256.25)
10. BALA=1 265 ==> CHOCOLATE=1 265 <conf:(1)> lift:(30.29) lev:(0.03) [256] conv:(256.25)
11. COMIDA_CONGELADA=1 261 ==> BATATA_PALHA=1 261 <conf:(1)> lift:(30.75) lev:(0.03) [252] conv:(252.51)
12. BATATA_PALHA=1 261 ==> COMIDA_CONGELADA=1 261 <conf:(1)> lift:(30.75) lev:(0.03) [252] conv:(252.51)
13. CREME_DE_LEITE=1 214 ==> LEITE=1 214 <conf:(1)> lift:(10.48) lev:(0.02) [193] conv:(193.58)
14. CANUDO=1 178 ==> AGUA=1 178 <conf:(1)> lift:(45.1) lev:(0.02) [174] conv:(174.05)
15. AGUA=1 178 ==> CANUDO=1 178 <conf:(1)> lift:(45.1) lev:(0.02) [174] conv:(174.05)
16. COMIDA_PET=1 REFRIGERANTE=1 159 ==> BEBIDA_ALCOOLICA=1 159 <conf:(1)> lift:(14.11) lev:(0.02) [147] conv:(147.73)
17. BEBIDA_ALCOOLICA=1 REFRIGERANTE=1 159 ==> COMIDA_PET=1 159 <conf:(1)> lift:(14.11) lev:(0.02) [147] conv:(147.73)
18. CAFE=1 PROD_LIMPEZA=1 150 ==> ACUCAR=1 150 <conf:(1)> lift:(20.48) lev:(0.02) [142] conv:(142.67)
19. ACUCAR=1 PROD_LIMPEZA=1 150 ==> CAFE=1 150 <conf:(1)> lift:(20.48) lev:(0.02) [142] conv:(142.67)
20. CORTADOR_DE_UNHA=1 146 ==> BEBIDAS_SEM_ALCOOL=1 146 <conf:(1)> lift:(54.98) lev:(0.02) [143] conv:(143.34)
21. BEBIDAS_SEM_ALCOOL=1 146 ==> CORTADOR_DE_UNHA=1 146 <conf:(1)> lift:(54.98) lev:(0.02) [143] conv:(143.34)
22. CAFE=1 OLEO_VEGETAL=1 146 ==> ACUCAR=1 146 <conf:(1)> lift:(20.48) lev:(0.02) [138] conv:(138.87)
23. ACUCAR=1 OLEO_VEGETAL=1 146 ==> CAFE=1 146 <conf:(1)> lift:(20.48) lev:(0.02) [138] conv:(138.87)
24. CHA=1 OLEO_VEGETAL=1 146 ==> ARROZ=1 146 <conf:(1)> lift:(25.4) lev:(0.02) [140] conv:(140.25)
25. ARROZ=1 OLEO_VEGETAL=1 146 ==> CHA=1 146 <conf:(1)> lift:(25.4) lev:(0.02) [140] conv:(140.25)
26. ARROZ=1 CAFE=1 141 ==> ACUCAR=1 141 <conf:(1)> lift:(20.48) lev:(0.02) [134] conv:(134.11)
27. ACUCAR=1 ARROZ=1 141 ==> CAFE=1 141 <conf:(1)> lift:(20.48) lev:(0.02) [134] conv:(134.11)
28. ACUCAR=1 CHA=1 141 ==> ARROZ=1 141 <conf:(1)> lift:(25.4) lev:(0.02) [135] conv:(135.45)
29. ARROZ=1 CHA=1 141 ==> ACUCAR=1 141 <conf:(1)> lift:(25.4) lev:(0.02) [135] conv:(135.45)
30. CAFE=1 CHA=1 141 ==> ACUCAR=1 141 <conf:(1)> lift:(20.48) lev:(0.02) [134] conv:(134.11)

```

Figura 3.14: Amostras dos Resultados do Quinto Trimestre

Para este trimestre foi obtido novos relacionamentos entre produtos do mercado, porém, por estarmos adotando um valor de suporte especificado anteriormente em 1%, a análise para este trimestre pode não ser muito boa, visto que podem aparecer muitas

regras que irão representar a compra efetuada em si e não um padrão frequentemente encontrado nas compras, como pode ser observado na imagem 3.14.

Com isso, foi observado novamente que as regras que possuíam como consequente PROD_LIMPEZA, tinham como associação de produtos, os produtos anteriormente citados, como: ARROZ, CAFE, LIMP_ROUPAS, ACUCAR entre alguns outros. Isso mostra um certo padrão de compra, que de fato será muito útil para uma posterior análise, conforme pode-se observar nas regras: ACUCAR, CAFE, CHA LIMP_ROUPAS → PROD_LIMPEZA e também ACUCAR, CAFE, LIMP_ROUPAS PROD_BANHO → PROD_LIMPEZA.

Foi observado também, a frequência entre a associação entre os produtos, e a partir dessa análise foi possível identificar que, quando ACUCAR e ARROZ estavam relacionados em um cesto de compra, consequentemente, na maioria dos casos implicavam na compra de CAFE ou de CHA, como pode-se observar nas regras: ACUCAR, ARROZ, LEGUMES → CAFE ACUCAR, ARROZ, CAFE, LEGUMES → CHA ACUCAR, ARROZ, LEGUMES → CAFE CHA.

A implicação específica da compra de ACUCAR foi analisada, e pode-se observar que a compra de CAFE, sugere que se compre ACUCAR, como pode-se notar nas regras: CAFE, MACARRAO, OLEO_VEGETAL → ACUCAR
CAFE, FARINHA → ACUCAR
CAFE, ESPECIARIAS, LEGUMES → ACUCAR
CAFE, LIMP_ROUPAS, MACARRAO → ACUCAR
CAFE, UTEN_COZINHA → ACUCAR.

Para este trimestre também foi possível identificar associação entre produtos como BEBIDA_ALCOOLICA e COMIDA_PET, como pode ser observado nas regras:

COMIDA_PET, REFRIGERANTE → BEBIDA_ALCOOLICA
COMIDA_PET → BEBIDA_ALCOOLICA
COMIDA_PET, LEGUMES → BEBIDA_ALCOOLICA
COMIDA_PET, SUCO → BEBIDA_ALCOOLICA.

Regras estabelecidas anteriormente para o mesmo período (Trimestre 1), e até mesmo produtos que foram encontrados nesse período para o ano anterior (Trimestre 1), aqui se mantiveram, como por exemplo, os produtos SUCO, CREME_DE_LEITE e LEITE. Isso gera um importante conhecimento a respeito de produtos que são consumidos para o

período estabelecido. Dessa forma, pode-se pensar em sugerir algumas estratégias que irão ter como base os produtos que foram aqui encontrados, e obviamente os que mantiveram um padrão para o mesmo período analisado.

3.2.7 Trimestre 6

Para o trimestre 6, foi obtido um total de 5686 compras realizadas para análise compreendidas no período de ABRIL/2014 até JUNHO/2014.

Best rules found:

1. CREME_DE_LEITE=1 175 ==> LEITE=1 175 <conf:(1)> lift:(9.77) lev:(0.03) [157] conv:(157.09)
2. FEIJAO=1 OLEO_VEGETAL=1 PROD_BANHO=1 PROD_LIMPEZA=1 60 ==> ARROZ=1 60 <conf:(1)> lift:(25.05) lev:(0.01) [57] conv:(57.6)
3. ARROZ=1 ESPECIARIAS=1 LIMP_ROUPAS=1 57 ==> OLEO_VEGETAL=1 57 <conf:(1)> lift:(16.43) lev:(0.01) [53] conv:(53.53)
4. ARROZ=1 ESPECIARIAS=1 LIMP_ROUPAS=1 PROD_LIMPEZA=1 56 ==> OLEO_VEGETAL=1 56 <conf:(1)> lift:(16.43) lev:(0.01) [52] conv:(52.59)
5. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 55 ==> LEITE=1 55 <conf:(1)> lift:(9.77) lev:(0.01) [49] conv:(49.37)
6. FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PROD_BANHO=1 PROD_LIMPEZA=1 55 ==> ARROZ=1 55 <conf:(1)> lift:(25.05) lev:(0.01) [52] conv:(52.8)
7. ARROZ=1 ESPECIARIAS=1 FEIJAO=1 LIMP_ROUPAS=1 54 ==> OLEO_VEGETAL=1 54 <conf:(1)> lift:(16.43) lev:(0.01) [50] conv:(50.71)
8. ARROZ=1 CAFE=1 MACARRAO=1 OLEO_VEGETAL=1 PROD_LIMPEZA=1 53 ==> FEIJAO=1 53 <conf:(1)> lift:(24.61) lev:(0.01) [50] conv:(50.85)
9. ESPECIARIAS=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PROD_LIMPEZA=1 53 ==> ARROZ=1 53 <conf:(1)> lift:(25.05) lev:(0.01) [50] conv:(50.88)
10. ARROZ=1 ESPECIARIAS=1 FEIJAO=1 LIMP_ROUPAS=1 PROD_LIMPEZA=1 53 ==> OLEO_VEGETAL=1 53 <conf:(1)> lift:(16.43) lev:(0.01) [49] conv:(49.77)
11. BISCOITO=1 CREME_DE_LEITE=1 51 ==> LEITE=1 51 <conf:(1)> lift:(9.77) lev:(0.01) [45] conv:(45.78)
12. ACUCAR=1 FEIJAO=1 PROD_BANHO=1 PROD_LIMPEZA=1 51 ==> ARROZ=1 51 <conf:(1)> lift:(25.05) lev:(0.01) [48] conv:(48.96)
13. ARROZ=1 ESPECIARIAS=1 FEIJAO=1 MACARRAO=1 51 ==> OLEO_VEGETAL=1 51 <conf:(1)> lift:(16.43) lev:(0.01) [47] conv:(47.9)
14. ARROZ=1 ESPECIARIAS=1 MACARRAO=1 PROD_LIMPEZA=1 51 ==> OLEO_VEGETAL=1 51 <conf:(1)> lift:(16.43) lev:(0.01) [47] conv:(47.9)
15. ARROZ=1 ESPECIARIAS=1 PROD_BANHO=1 PROD_LIMPEZA=1 51 ==> OLEO_VEGETAL=1 51 <conf:(1)> lift:(16.43) lev:(0.01) [47] conv:(47.9)
16. ARROZ=1 ESPECIARIAS=1 PAPEL_HIGIENICO=1 50 ==> OLEO_VEGETAL=1 50 <conf:(1)> lift:(16.43) lev:(0.01) [46] conv:(46.96)
17. ARROZ=1 EXTRATO_DE_TOMATE=1 PROD_LIMPEZA=1 50 ==> FEIJAO=1 50 <conf:(1)> lift:(24.61) lev:(0.01) [47] conv:(47.97)
18. ACUCAR=1 ARROZ=1 MACARRAO=1 PAPEL_HIGIENICO=1 50 ==> OLEO_VEGETAL=1 50 <conf:(1)> lift:(16.43) lev:(0.01) [46] conv:(46.96)
19. ARROZ=1 ESPECIARIAS=1 LIMP_ROUPAS=1 PROD_BANHO=1 50 ==> OLEO_VEGETAL=1 50 <conf:(1)> lift:(16.43) lev:(0.01) [46] conv:(46.96)
20. FEIJAO=1 MACARRAO=1 OLEO_VEGETAL=1 PROD_BANHO=1 50 ==> ARROZ=1 50 <conf:(1)> lift:(25.05) lev:(0.01) [48] conv:(48)
21. ACUCAR=1 FEIJAO=1 OLEO_VEGETAL=1 PROD_BANHO=1 PROD_LIMPEZA=1 50 ==> ARROZ=1 50 <conf:(1)> lift:(25.05) lev:(0.01) [48] conv:(48)
22. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 49 ==> LEITE=1 49 <conf:(1)> lift:(9.77) lev:(0.01) [43] conv:(43.98)
23. CREME_DE_LEITE=1 QUEIJOS=1 49 ==> LEITE=1 49 <conf:(1)> lift:(9.77) lev:(0.01) [43] conv:(43.98)
24. ACUCAR=1 ESPECIARIAS=1 FEIJAO=1 PROD_LIMPEZA=1 49 ==> ARROZ=1 49 <conf:(1)> lift:(25.05) lev:(0.01) [47] conv:(47.04)
25. ARROZ=1 ESPECIARIAS=1 FEIJAO=1 PROD_BANHO=1 49 ==> OLEO_VEGETAL=1 49 <conf:(1)> lift:(16.43) lev:(0.01) [46] conv:(46.02)
26. ESPECIARIAS=1 FEIJAO=1 MACARRAO=1 OLEO_VEGETAL=1 PROD_LIMPEZA=1 49 ==> ARROZ=1 49 <conf:(1)> lift:(25.05) lev:(0.01) [47] conv:(47.04)
27. ARROZ=1 ESPECIARIAS=1 FEIJAO=1 MACARRAO=1 PROD_LIMPEZA=1 49 ==> OLEO_VEGETAL=1 49 <conf:(1)> lift:(16.43) lev:(0.01) [46] conv:(46.02)
28. ARROZ=1 ESPECIARIAS=1 LIMP_ROUPAS=1 PROD_BANHO=1 PROD_LIMPEZA=1 49 ==> OLEO_VEGETAL=1 49 <conf:(1)> lift:(16.43) lev:(0.01) [46] conv:(46.02)
29. FEIJAO=1 LEGUMES=1 OLEO_VEGETAL=1 PROD_BANHO=1 PROD_LIMPEZA=1 49 ==> ARROZ=1 49 <conf:(1)> lift:(25.05) lev:(0.01) [47] conv:(47.04)
30. FEIJAO=1 MACARRAO=1 OLEO_VEGETAL=1 PROD_BANHO=1 PROD_LIMPEZA=1 49 ==> ARROZ=1 49 <conf:(1)> lift:(25.05) lev:(0.01) [47] conv:(47.04)

Figura 3.15: Amostras dos Resultados do Sexto Trimestre

Mesmo este trimestre tendo uma significativa queda com relação ao quantitativo de transações de compras, foi possível notar que alguns comportamentos já analisados anteriormente. Para o mesmo trimestre no ano anterior (Trimestre 2), foi possível perceber que o comportamento de compras, analisado anteriormente continuou praticamente o mesmo. Regras anteriormente definidas, com uma maior atenção para o mesmo trimestre relacionado ao ano anterior (trimestre 2), irão ajudar a consolidar um padrão de compras para o período analisado, pois, conforme dito anteriormente, a maioria das regras já foram analisadas e encontram-se explicitadas novamente neste sexto trimestre, como pode ser possível observar na regra: CREME_DE_LEITE \rightarrow LEITE e também em

ACUCAR FEIJAO LIMP_ROUPAS OLEO_VEGETAL → PROD_LIMPEZA.

Houve um grande número de implicações de associações tendo como consequente o produto OLEO_VEGETAL. Com isso, foi percebido que a os mesmos produtos que originariam a possível compra de PROD_LIMPEZA, estariam relacionados para gerar a implicação de compra do produto OLEO_VEGETAL, como pode ser observado nas regras:

ACUCAR, ARROZ, FEIJAO MACARRAO → OLEO_VEGETAL

ARROZ, LEGUMES, LIMP_ROUPAS → OLEO_VEGETAL

ARROZ, CAFE, MACARRAO → OLEO_VEGETAL

CAFE, FEIJAO, LIMP_ROUPAS → OLEO_VEGETAL.

Também foi possível reparar no grande número de regras onde estavam presentes produtos como, ACUCAR, ARROZ e também regras que continham a união dos produtos ARROZ e FEIJAO. A aparição da união entre ARROZ e FEIJAO em uma mesma regra, foi subentendido que já é de grande conhecimento comum, já a aparição deles separadamente em uma regra, fornece uma relação que poderá ser analisada com mais atenção posteriormente.

A aparição do produto ESPECIARIAS, se manteve com um número elevado nas regras geradas, como pode ser comparado ao mesmo período do ano anterior (Trimestre 3), portanto, é levado a entender, que esse produto pode ser mais consumido neste determinado período do ano.

3.2.8 Trimestre 7

Para o trimestre 7, tem-se um total de 5643 compras realizadas, para análise compreendidas no período de JULHO/2014 até SETEMBRO/2014.

Neste sétimo trimestre, apareceram regras bem distintas do que foi apresentado até agora, conforme pode ser observado na Figura 3.16. Para este trimestre, foi observado regras onde existia a relação direta entre produtos que antes não foi possível observar e que demonstram ser interessantes relacionamentos a serem analisados, como a relação entre CREME (Creme Hidratante) e CAFE, FRUTAS e CHOCOLATE, GORDURA_VEGETAL e PAES, e também GELATINA e COMIDA_CONGELADA conforme é apresentado através das seguintes regras:

Best rules found:

```

1. CREME=1 376 ==> CAFE=1 376 <conf:(1)> lift:(15.01) lev:(0.06) [350] conv:(350.95)
2. CAFE=1 376 ==> CREME=1 376 <conf:(1)> lift:(15.01) lev:(0.06) [350] conv:(350.95)
3. FRUTAS=1 305 ==> CHOCOLATE=1 305 <conf:(1)> lift:(18.5) lev:(0.05) [288] conv:(288.51)
4. CHOCOLATE=1 305 ==> FRUTAS=1 305 <conf:(1)> lift:(18.5) lev:(0.05) [288] conv:(288.51)
5. GORDURA_VEGETAL=1 168 ==> PAES=1 168 <conf:(1)> lift:(9.56) lev:(0.03) [150] conv:(150.43)
6. GELATINA=1 164 ==> COMIDA_CONGELADA=1 164 <conf:(1)> lift:(34.41) lev:(0.03) [159] conv:(159.23)
7. COMIDA_CONGELADA=1 164 ==> GELATINA=1 164 <conf:(1)> lift:(34.41) lev:(0.03) [159] conv:(159.23)
8. CREME=1 OLEO_VEGETAL=1 128 ==> CAFE=1 128 <conf:(1)> lift:(15.01) lev:(0.02) [119] conv:(119.47)
9. CAFE=1 OLEO_VEGETAL=1 128 ==> CREME=1 128 <conf:(1)> lift:(15.01) lev:(0.02) [119] conv:(119.47)
10. ACUCAR=1 CREME=1 124 ==> CAFE=1 124 <conf:(1)> lift:(15.01) lev:(0.02) [115] conv:(115.74)
11. ACUCAR=1 CAFE=1 124 ==> CREME=1 124 <conf:(1)> lift:(15.01) lev:(0.02) [115] conv:(115.74)
12. CREME_DE_LEITE=1 120 ==> BOLOS=1 120 <conf:(1)> lift:(47.02) lev:(0.02) [117] conv:(117.45)
13. BOLOS=1 120 ==> CREME_DE_LEITE=1 120 <conf:(1)> lift:(47.02) lev:(0.02) [117] conv:(117.45)
14. ESPELHO=1 112 ==> CARNE_SUINA=1 112 <conf:(1)> lift:(50.38) lev:(0.02) [109] conv:(109.78)
15. CARNE_SUINA=1 112 ==> ESPELHO=1 112 <conf:(1)> lift:(50.38) lev:(0.02) [109] conv:(109.78)
16. GELEIA=1 111 ==> COMIDA_PET=1 111 <conf:(1)> lift:(50.84) lev:(0.02) [108] conv:(108.82)
17. COMIDA_PET=1 111 ==> GELEIA=1 111 <conf:(1)> lift:(50.84) lev:(0.02) [108] conv:(108.82)
18. BISCOITO=1 CREME=1 106 ==> CAFE=1 106 <conf:(1)> lift:(15.01) lev:(0.02) [98] conv:(98.94)
19. BISCOITO=1 CAFE=1 106 ==> CREME=1 106 <conf:(1)> lift:(15.01) lev:(0.02) [98] conv:(98.94)
20. FAROFA=1 105 ==> CEREAL=1 105 <conf:(1)> lift:(53.74) lev:(0.02) [103] conv:(103.05)
21. CEREAL=1 105 ==> FAROFA=1 105 <conf:(1)> lift:(53.74) lev:(0.02) [103] conv:(103.05)
22. CREME=1 PALMITO=1 104 ==> CAFE=1 104 <conf:(1)> lift:(15.01) lev:(0.02) [97] conv:(97.07)
23. CAFE=1 PALMITO=1 104 ==> CREME=1 104 <conf:(1)> lift:(15.01) lev:(0.02) [97] conv:(97.07)
24. BISCOITO=1 FRUTAS=1 98 ==> CHOCOLATE=1 98 <conf:(1)> lift:(18.5) lev:(0.02) [92] conv:(92.7)
25. BISCOITO=1 CHOCOLATE=1 98 ==> FRUTAS=1 98 <conf:(1)> lift:(18.5) lev:(0.02) [92] conv:(92.7)
26. CREME=1 PROD_LIMPEZA=1 97 ==> CAFE=1 97 <conf:(1)> lift:(15.01) lev:(0.02) [90] conv:(90.54)
27. CAFE=1 PROD_LIMPEZA=1 97 ==> CREME=1 97 <conf:(1)> lift:(15.01) lev:(0.02) [90] conv:(90.54)
28. CREME=1 QUEIJOS=1 90 ==> CAFE=1 90 <conf:(1)> lift:(15.01) lev:(0.01) [84] conv:(84)
29. CAFE=1 QUEIJOS=1 90 ==> CREME=1 90 <conf:(1)> lift:(15.01) lev:(0.01) [84] conv:(84)
30. CREME=1 PALITO_PARA_UNHAS=1 89 ==> CAFE=1 89 <conf:(1)> lift:(15.01) lev:(0.01) [83] conv:(83.07)

```

Figura 3.16: Amostras dos Resultados do Sétimo Trimestre

CREME → CAFE

FRUTAS → CHOCOLATE

GORDURA_VEGETAL → PAES

GELATINA → COMIDA_CONGELADA

CREME_DE_LEITE → BOLOS

ESPELHO → CARNE_SUINA

GELEIA → COMIDA_PET

CEREAL → FAROFA

FARINHA → CASADINHA

É evidente que a ocorrência de um desses produtos, para este trimestre, implicará no aparecimento de seu subsequente, conforme apresentado. A avaliação mais pertinente a ser feita neste caso, é avaliar o aparecimento de alguns padrões de comportamento já observado nos trimestres anteriores e a relação entre os produtos para este trimestre.

A relação anteriormente encontrada entre ARROZ e ACUCAR se mantém para este trimestre como pode ser observado nas regras:

ARROZ, CAFE → ACUCAR

ARROZ, CREME \rightarrow ACUCAR

ARROZ, CAFE, CREME \rightarrow ACUCAR

ACUCAR, PALMITO \rightarrow ARROZ

A relação de produtos comprados e PROD.LIMPEZA para este trimestre também se mantém, como: ARROZ, LEITE.CONDENSADO \rightarrow PROD.LIMPEZA.

A maioria das regras deste trimestre, tem como consequentes o produto CAFE ou o produto CREME, portanto regras que possuem qualquer um desses produtos apresentados, vão gerar na maioria das vezes como consequente, o produto que está relacionado pelas regras diretas apresentadas, como: BISCOITO, CREME, OLEO-VEGETAL \rightarrow CAFE e também na regra representada por: ARROZ, CAFE, LEITE.CONDENSADO \rightarrow CREME.

Para este trimestre foi possível observar alguns padrões que vem sido encontrados ao longo do estudo. Porém, o que mais chamou a atenção foi o fato do produto CREME, para este período específico, manter um padrão com relação ao período anterior (Trimestre 3), além da sua significância no conjunto de dados, semelhante ao que foi observado no período anterior (Trimestre 3). Com isso, é possível consolidar este conhecimento para que possa ser utilizado posteriormente em futuras estratégias de vendas.

3.2.9 Trimestre 8

Para o trimestre 8, tem-se um total de 5565 compras realizadas, para análise compreendidas no período de OUTUBRO/2014 até DEZEMBRO/2014.

Para o oitavo trimestre, foi possível perceber o padrão demonstrado pelos trimestres anteriores. O produto PROD.LIMPEZA aparece na maioria das regras geradas efetivando cada vez mais o conhecimento adquirido em seus relacionamentos. Relacionamentos que foram frequentes nos primeiros três trimestres, apareceram novamente no oitavo, como é o caso da relação entre CREME.DELEITE e LEITE como se pode observar na Figura 3.17.

Uma das implicações de compra para CAFE se manteve neste trimestre, conforme apresentado anteriormente na análise do trimestre 5. Os produtos ACUCAR e ARROZ, são frequentemente presentes em uma cesta de compra no trimestre 8, com isso é obtido como consequente o produto CAFE, conforme regras ACUCAR, ARROZ, PROD.LIMPEZA \rightarrow CAFE

Best rules found:

```

1. CREME_DE_LEITE=1 171 ==> LEITE=1 171 <conf:(1)> lift:(11.29) lev:(0.03) [155] conv:(155.85)
2. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 64 ==> LEITE=1 64 <conf:(1)> lift:(11.29) lev:(0.01) [58] conv:(58.33)
3. BISCOITO=1 CREME_DE_LEITE=1 60 ==> LEITE=1 60 <conf:(1)> lift:(11.29) lev:(0.01) [54] conv:(54.68)
4. BATATA_PALHA=1 CREME_DE_LEITE=1 51 ==> LEITE=1 51 <conf:(1)> lift:(11.29) lev:(0.01) [46] conv:(46.48)
5. CREME_DE_LEITE=1 LEGUMES=1 51 ==> LEITE=1 51 <conf:(1)> lift:(11.29) lev:(0.01) [46] conv:(46.48)
6. CREME_DE_LEITE=1 PROD_LIMPEZA=1 51 ==> LEITE=1 51 <conf:(1)> lift:(11.29) lev:(0.01) [46] conv:(46.48)
7. CREME_DE_LEITE=1 REFRIGERANTE=1 51 ==> LEITE=1 51 <conf:(1)> lift:(11.29) lev:(0.01) [46] conv:(46.48)
8. CREME_DE_LEITE=1 SUCO=1 51 ==> LEITE=1 51 <conf:(1)> lift:(11.29) lev:(0.01) [46] conv:(46.48)
9. ARROZ=1 FUBA=1 LIMP_ROUPAS=1 47 ==> PROD_BANHO=1 47 <conf:(1)> lift:(17.84) lev:(0.01) [44] conv:(44.36)
10. CAFE=1 FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 47 ==> PROD_BANHO=1 47 <conf:(1)> lift:(17.84) lev:(0.01) [44] conv:(44.36)
11. FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 PROD_LIMPEZA=1 47 ==> PROD_BANHO=1 47 <conf:(1)> lift:(17.84) lev:(0.01) [44] conv:(44.36)
12. FUBA=1 LIMP_ROUPAS=1 MACARRAO=1 46 ==> PROD_BANHO=1 46 <conf:(1)> lift:(17.84) lev:(0.01) [43] conv:(43.42)
13. ARROZ=1 CAFE=1 FUBA=1 LIMP_ROUPAS=1 46 ==> PROD_BANHO=1 46 <conf:(1)> lift:(17.84) lev:(0.01) [43] conv:(43.42)
14. CREME_DE_LEITE=1 ESPECIARIAS=1 45 ==> LEITE=1 45 <conf:(1)> lift:(11.29) lev:(0.01) [41] conv:(41.01)
15. CREME_DE_LEITE=1 QUEIJOS=1 45 ==> LEITE=1 45 <conf:(1)> lift:(11.29) lev:(0.01) [41] conv:(41.01)
16. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 44 ==> LEITE=1 44 <conf:(1)> lift:(11.29) lev:(0.01) [40] conv:(40.1)
17. ARROZ=1 BISCOITO=1 FUBA=1 PROD_BANHO=1 44 ==> CAFE=1 44 <conf:(1)> lift:(17.78) lev:(0.01) [41] conv:(41.53)
18. ARROZ=1 FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 44 ==> PROD_BANHO=1 44 <conf:(1)> lift:(17.84) lev:(0.01) [41] conv:(41.53)
19. CAFE=1 FUBA=1 LIMP_ROUPAS=1 MACARRAO=1 44 ==> PROD_BANHO=1 44 <conf:(1)> lift:(17.84) lev:(0.01) [41] conv:(41.53)
20. FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 MACARRAO=1 44 ==> PROD_BANHO=1 44 <conf:(1)> lift:(17.84) lev:(0.01) [41] conv:(41.53)
21. FUBA=1 LIMP_ROUPAS=1 MACARRAO=1 PROD_LIMPEZA=1 44 ==> PROD_BANHO=1 44 <conf:(1)> lift:(17.84) lev:(0.01) [41] conv:(41.53)
22. CAFE=1 FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 PROD_LIMPEZA=1 44 ==> PROD_BANHO=1 44 <conf:(1)> lift:(17.84) lev:(0.01) [41] conv:(41.53)
23. ARROZ=1 CAFE=1 FUBA=1 OLEO_VEGETAL=1 43 ==> PROD_BANHO=1 43 <conf:(1)> lift:(17.84) lev:(0.01) [40] conv:(40.59)
24. ARROZ=1 FUBA=1 LIMP_ROUPAS=1 PROD_LIMPEZA=1 43 ==> PROD_BANHO=1 43 <conf:(1)> lift:(17.84) lev:(0.01) [40] conv:(40.59)
25. BISCOITO=1 LIMP_ROUPAS=1 PROD_BANHO=1 PROD_ODONTO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(7.67) lev:(0.01) [37] conv:(37.39)
26. ESPECIARIAS=1 FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 43 ==> PROD_BANHO=1 43 <conf:(1)> lift:(17.84) lev:(0.01) [40] conv:(40.59)
27. FEIJAO=1 FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 43 ==> PROD_BANHO=1 43 <conf:(1)> lift:(17.84) lev:(0.01) [40] conv:(40.59)
28. FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 43 ==> PROD_BANHO=1 43 <conf:(1)> lift:(17.84) lev:(0.01) [40] conv:(40.59)
29. ARROZ=1 CAFE=1 FUBA=1 LEGUMES=1 LIMP_ROUPAS=1 43 ==> PROD_BANHO=1 43 <conf:(1)> lift:(17.84) lev:(0.01) [40] conv:(40.59)
30. CREME_DE_LEITE=1 LIMP_ROUPAS=1 42 ==> LEITE=1 42 <conf:(1)> lift:(11.29) lev:(0.01) [38] conv:(38.28)

```

Figura 3.17: Amostras dos Resultados do Oitavo Trimestre

e ACUCAR, ARROZ, LEGUMES → CAFE.

Produtos como ARROZ e CAFE foram frequentemente encontrados em uma mesma cesta de compra para este período, conforme observado nas regras:

ARROZ, CAFE, LEGUMES → ACUCAR

ARROZ, CAFE, PROD_LIMPEZA → BISCOITO

ARROZ, CAFE, PROD_LIMPEZA → FEIJAO

Também foi percebido a manutenção de compras de LEGUMES com relação ao trimestre correspondente ao mesmo período de estudo (Trimestre 4), assim como sua correlação com o produto PROD_LIMPEZA, observado nas regras:

CAFE, LEGUMES, PROD_LIMPEZA → LIMP_ROUPAS

MACARRAO, PROD_BANHO, PROD_LIMPEZA → LEGUMES

ESPECIARIAS, LEGUMES, PROD_LIMPEZA → PROD_BANHO

Este trimestre mostrou o comportamento que se esperava em relação as compras comuns realizadas, pois está de acordo com o que foi analisado para os trimestres anteriores, ajudando a consolidar o conhecimento adquirido até aqui. Para produtos fora desse padrão, foi possível observar que o produto PROD_BANHO, foi mais frequentemente

encontrado, para ambos os trimestres que estão relacionados ao mesmo período de tempo analisado, que neste caso compreende ao período correspondente ao trimestre 4.

3.2.10 Trimestre 9

Para o trimestre 9 tem-se um total de 4784 compras realizadas, para análise compreendidas no período de JANEIRO/2015 até MARÇO/2015.

Best rules found:

```

1. CREME_DE_LEITE=1 168 ==> LEITE=1 168 <conf:(1)> lift:(11.75) lev:(0.03) [153] conv:(153.71)
2. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 73 ==> LEITE=1 73 <conf:(1)> lift:(11.75) lev:(0.01) [66] conv:(66.79)
3. BISCOITO=1 CREME_DE_LEITE=1 56 ==> LEITE=1 56 <conf:(1)> lift:(11.75) lev:(0.01) [51] conv:(51.24)
4. CREME_DE_LEITE=1 LEGUMES=1 54 ==> LEITE=1 54 <conf:(1)> lift:(11.75) lev:(0.01) [49] conv:(49.41)
5. CREME_DE_LEITE=1 SUCO=1 52 ==> LEITE=1 52 <conf:(1)> lift:(11.75) lev:(0.01) [47] conv:(47.58)
6. CREME_DE_LEITE=1 LEGUMES_EM_LATA=1 46 ==> LEITE=1 46 <conf:(1)> lift:(11.75) lev:(0.01) [42] conv:(42.09)
7. FEIJAO=1 LIMP_ROUPAS=1 SUCO=1 46 ==> PROD_LIMPEZA=1 46 <conf:(1)> lift:(6.99) lev:(0.01) [39] conv:(39.42)
8. ARROZ=1 CAFE=1 LIMP_ROUPAS=1 PROD_BANHO=1 45 ==> PROD_LIMPEZA=1 45 <conf:(1)> lift:(6.99) lev:(0.01) [38] conv:(38.57)
9. ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 PROD_BANHO=1 45 ==> PROD_LIMPEZA=1 45 <conf:(1)> lift:(6.99) lev:(0.01) [38] conv:(38.57)
10. LEGUMES=1 LIMP_ROUPAS=1 PROD_BANHO=1 SUCO=1 45 ==> PROD_LIMPEZA=1 45 <conf:(1)> lift:(6.99) lev:(0.01) [38] conv:(38.57)
11. CREME_DE_LEITE=1 PROD_LIMPEZA=1 44 ==> LEITE=1 44 <conf:(1)> lift:(11.75) lev:(0.01) [40] conv:(40.26)
12. CREME_DE_LEITE=1 REFRIGERANTE=1 44 ==> LEITE=1 44 <conf:(1)> lift:(11.75) lev:(0.01) [40] conv:(40.26)
13. ARROZ=1 BISCOITO=1 CAFE=1 LIMP_ROUPAS=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(6.99) lev:(0.01) [37] conv:(37.71)
14. ARROZ=1 CAFE=1 LIMP_ROUPAS=1 PAPEL_HIGIENICO=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(6.99) lev:(0.01) [37] conv:(37.71)
15. ARROZ=1 FEIJAO=1 MACARRAO=1 PROD_BANHO=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(6.99) lev:(0.01) [37] conv:(37.71)
16. ARROZ=1 LEGUMES=1 MACARRAO=1 PROD_BANHO=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(6.99) lev:(0.01) [37] conv:(37.71)
17. ARROZ=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PAPEL_HIGIENICO=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(6.99) lev:(0.01) [37] conv:(37.71)
18. FEIJAO=1 LEGUMES=1 LIMP_ROUPAS=1 SUCO=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(6.99) lev:(0.01) [37] conv:(37.71)
19. CREME_DE_LEITE=1 QUEIJOS=1 43 ==> LEITE=1 43 <conf:(1)> lift:(11.75) lev:(0.01) [39] conv:(39.34)
20. ARROZ=1 BISCOITO=1 LEGUMES=1 LIMP_ROUPAS=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(6.99) lev:(0.01) [36] conv:(36.85)
21. ARROZ=1 CAFE=1 FEIJAO=1 PROD_BANHO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(6.99) lev:(0.01) [36] conv:(36.85)
22. ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 PAPEL_HIGIENICO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(6.99) lev:(0.01) [36] conv:(36.85)
23. ARROZ=1 LEGUMES=1 LIMP_ROUPAS=1 PROD_BANHO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(6.99) lev:(0.01) [36] conv:(36.85)
24. ARROZ=1 MACARRAO=1 PAPEL_HIGIENICO=1 PROD_BANHO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(6.99) lev:(0.01) [36] conv:(36.85)
25. CAFE=1 ESPECIARIAS=1 LIMP_ROUPAS=1 PROD_BANHO=1 43 ==> MACARRAO=1 43 <conf:(1)> lift:(12.93) lev:(0.01) [39] conv:(39.67)
26. FEIJAO=1 LIMP_ROUPAS=1 MACARRAO=1 SUCO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(6.99) lev:(0.01) [36] conv:(36.85)
27. ARROZ=1 EXTRATO_DE_TOMATE=1 LIMP_ROUPAS=1 42 ==> MACARRAO=1 42 <conf:(1)> lift:(12.93) lev:(0.01) [38] conv:(38.75)
28. ARROZ=1 CAFE=1 ESPECIARIAS=1 LIMP_ROUPAS=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(6.99) lev:(0.01) [35] conv:(35.99)
29. ARROZ=1 ESPECIARIAS=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(6.99) lev:(0.01) [35] conv:(35.99)
30. ARROZ=1 LIMP_ROUPAS=1 PAPEL_HIGIENICO=1 PROD_BANHO=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(6.99) lev:(0.01) [35] conv:(35.99)

```

Figura 3.18: Amostras dos Resultados do Nono Trimestre

No nono trimestre, o número de compras foi reduzido praticamente pela metade em relação ao primeiro (9454 compras realizadas) e também ao quinto trimestre (8027 compras realizadas), porém, este trimestre mantém o comportamento padrão analisado até o momento. Para este trimestre, mais uma vez foi obtido um grande número de produtos que relacionados implicavam na compra de PROD_LIMPEZA. A relação com esse produto, já foi analisado em trimestres anteriores e só consolida a descoberta de informações importantíssimas geradas através desse estudo.

Como pode ser observado na Figura 3.18, algumas regras mostram as relações que já foram observadas em outros períodos e que aqui se mantiveram, como é o caso da regra: CREME.DE.LEITE → LEITE.

Foi possível perceber também, a frequência com que produtos como ACUCAR, FEIJAO e LIMP_ROUPAS, foram apresentados para este trimestre, conforme pode ser observado nas regras:

ACUCAR, FEIJAO, LIMP_ROUPAS, PROD_LIMPEZA → CAFE

ACUCAR, FEIJAO, LIMP_ROUPAS, PROD_LIMPEZA → LEGUMES

ACUCAR, FEIJAO, LEGUMES, PROD_LIMPEZA → LIMP_ROUPAS

ACUCAR, FEIJAO, LIMP_ROUPAS, PROD_LIMPEZA → MACARRAO

ACUCAR, FEIJAO, LIMP_ROUPAS, → CAFE

ACUCAR, FEIJAO, LIMP_ROUPAS, → LEGUMES

Também foi observado a relação entre os produtos ARROZ e BISCOITO, exemplificado nas regras:

ARROZ, BISCOITO, LEGUMES, PROD_LIMPEZA → MACARRAO

ARROZ, BISCOITO, MACARRAO, → PROD_LIMPEZA

ARROZ, BISCOITO, PROD_LIMPEZA → MACARRAO

ARROZ, BISCOITO, OLEO_VEGETAL → PROD_LIMPEZA

Uma relação aparente no trimestre 3, aparece bem recorrente no trimestre 9, que é a relação entre ESPECIARIAS E LEGUMES, exemplificada pelas regras:

ESPECIARIAS, LEGUMES, PROD_BANHO → LIMP_ROUPAS

ESPECIARIAS, LEGUMES, PAPEL_HIGIENICO → LIMP_ROUPAS

ESPECIARIAS, LEGUMES, PAPEL_HIGIENICO → PROD_LIMPEZA

ESPECIARIAS, LEGUMES, PROD_BANHO → PROD_LIMPEZA

O relacionamento entre ACUCAR e ARROZ anteriormente analisado, também se faz presente neste trimestre:

ACUCAR, ARROZ, FEIJAO, PROD_LIMPEZA → CAFE

ACUCAR, ARROZ, CAFE, FEIJAO → PROD_LIMPEZA

ACUCAR, ARROZ, CAFE, LEGUMES → LIMP_ROUPAS

ACUCAR, ARROZ, CAFE, MACARRAO → LEGUMES

ACUCAR, ARROZ, CAFE, LEGUMES → MACARRAO

A partir desses dados, podemos confirmar um comportamento ideal para a fundamentação da conclusão teórica, onde irá procurar filtrar essas informações com intuito de

apresentar os padrões obtidos nas compras, levando em conta o período que se encontra. Essas informações poderão vir a auxiliar as vendas do mercado em questão, tanto baseando em informações dos produtos de uso básico, quanto a produtos que possuem uma forte pertinência para o determinado período de tempo avaliado. Isso pode ser observado através da aparição da relação entre produtos como CREME.DE.LEITE e LEITE, que fogem do padrão de consumo comum, e que foram comumente encontrados nesse mesmo intervalo de tempo.

3.2.11 Trimestre 10

Para o trimestre 10 tem-se um total de 3882 compras realizadas, para análise compreendidas no período de ABRIL/2015 até JUNHO/2015.

Best rules found:

```

1. CREME_DE_LEITE=1 105 ==> LEITE=1 105 <conf:(1)> lift:(11.52) lev:(0.02) [95] conv:(95.88)
2. CAFE=1 LIMP_ROUPAS=1 PROD_BANHO=1 51 ==> PROD_LIMPEZA=1 51 <conf:(1)> lift:(7.37) lev:(0.01) [44] conv:(44.08)
3. ACUCAR=1 ESPECIARIAS=1 LIMP_ROUPAS=1 50 ==> PROD_LIMPEZA=1 50 <conf:(1)> lift:(7.37) lev:(0.01) [43] conv:(43.21)
4. FEIJAO=1 LIMP_ROUPAS=1 PROD_BANHO=1 50 ==> PROD_LIMPEZA=1 50 <conf:(1)> lift:(7.37) lev:(0.01) [43] conv:(43.21)
5. ACUCAR=1 BISCOITO=1 LIMP_ROUPAS=1 49 ==> PROD_LIMPEZA=1 49 <conf:(1)> lift:(7.37) lev:(0.01) [42] conv:(42.35)
6. LIMP_ROUPAS=1 PAPEL_HIGIENICO=1 PROD_BANHO=1 49 ==> PROD_LIMPEZA=1 49 <conf:(1)> lift:(7.37) lev:(0.01) [42] conv:(42.35)
7. CAFE=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 48 ==> PROD_LIMPEZA=1 48 <conf:(1)> lift:(7.37) lev:(0.01) [41] conv:(41.48)
8. ACUCAR=1 CAFE=1 LIMP_ROUPAS=1 PROD_BANHO=1 47 ==> PROD_LIMPEZA=1 47 <conf:(1)> lift:(7.37) lev:(0.01) [40] conv:(40.62)
9. ACUCAR=1 ESPECIARIAS=1 LIMP_ROUPAS=1 PROD_BANHO=1 47 ==> PROD_LIMPEZA=1 47 <conf:(1)> lift:(7.37) lev:(0.01) [40] conv:(40.62)
10. ACUCAR=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 47 ==> PROD_LIMPEZA=1 47 <conf:(1)> lift:(7.37) lev:(0.01) [40] conv:(40.62)
11. ACUCAR=1 FEIJAO=1 LIMP_ROUPAS=1 PROD_BANHO=1 47 ==> PROD_LIMPEZA=1 47 <conf:(1)> lift:(7.37) lev:(0.01) [40] conv:(40.62)
12. BISCOITO=1 LIMP_ROUPAS=1 PROD_BANHO=1 46 ==> PROD_LIMPEZA=1 46 <conf:(1)> lift:(7.37) lev:(0.01) [39] conv:(39.76)
13. CAFE=1 ESPECIARIAS=1 LIMP_ROUPAS=1 46 ==> PROD_LIMPEZA=1 46 <conf:(1)> lift:(7.37) lev:(0.01) [39] conv:(39.76)
14. ACUCAR=1 BISCOITO=1 PROD_BANHO=1 45 ==> PROD_LIMPEZA=1 45 <conf:(1)> lift:(7.37) lev:(0.01) [38] conv:(38.89)
15. ACUCAR=1 LIMP_ROUPAS=1 MACARRAO=1 45 ==> PROD_LIMPEZA=1 45 <conf:(1)> lift:(7.37) lev:(0.01) [38] conv:(38.89)
16. ACUCAR=1 CAFE=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 45 ==> PROD_LIMPEZA=1 45 <conf:(1)> lift:(7.37) lev:(0.01) [38] conv:(38.89)
17. ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(7.37) lev:(0.01) [38] conv:(38.03)
18. CAFE=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 44 ==> PROD_LIMPEZA=1 44 <conf:(1)> lift:(7.37) lev:(0.01) [38] conv:(38.03)
19. ACUCAR=1 BISCOITO=1 LIMP_ROUPAS=1 PROD_BANHO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(7.37) lev:(0.01) [37] conv:(37.16)
20. ACUCAR=1 CAFE=1 ESPECIARIAS=1 LIMP_ROUPAS=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(7.37) lev:(0.01) [37] conv:(37.16)
21. CAFE=1 FEIJAO=1 LIMP_ROUPAS=1 PROD_BANHO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(7.37) lev:(0.01) [37] conv:(37.16)
22. FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 PROD_BANHO=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(7.37) lev:(0.01) [37] conv:(37.16)
23. ACUCAR=1 ARROZ=1 FEIJAO=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 43 ==> PROD_LIMPEZA=1 43 <conf:(1)> lift:(7.37) lev:(0.01) [37] conv:(37.16)
24. BISCOITO=1 CREME_DE_LEITE=1 42 ==> LEITE=1 42 <conf:(1)> lift:(11.52) lev:(0.01) [38] conv:(38.35)
25. CREME_DE_LEITE=1 LEITE_CONDENSADO=1 42 ==> LEITE=1 42 <conf:(1)> lift:(11.52) lev:(0.01) [38] conv:(38.35)
26. LEGUMES=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(7.37) lev:(0.01) [36] conv:(36.3)
27. ARROZ=1 LIMP_ROUPAS=1 PROD_BANHO=1 PROD_LIMPEZA=1 42 ==> ACUCAR=1 42 <conf:(1)> lift:(17.33) lev:(0.01) [39] conv:(39.58)
28. ACUCAR=1 FEIJAO=1 LIMP_ROUPAS=1 MACARRAO=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(7.37) lev:(0.01) [36] conv:(36.3)
29. ACUCAR=1 LIMP_ROUPAS=1 MACARRAO=1 PROD_BANHO=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(7.37) lev:(0.01) [36] conv:(36.3)
30. ARROZ=1 CAFE=1 LIMP_ROUPAS=1 OLEO_VEGETAL=1 42 ==> PROD_LIMPEZA=1 42 <conf:(1)> lift:(7.37) lev:(0.01) [36] conv:(36.3)

```

Figura 3.19: Amostras dos Resultados do Décimo Trimestre

Neste último trimestre a ser analisado, percebe-se que o número de vendas diminuiu drasticamente com relação a trimestres anteriores. Foi o período com o menor quantitativo de compras ao ser comparado com todos os outros anteriormente analisados. Mesmo com essa queda, foi possível observar o comportamento das compras realizadas e observar que, mais uma vez, tem-se que o maior número de implicações de compras para

este período, referem-se a compra do produto PROD_LIMPEZA, como pode se observar na imagem 3.19.

Os produtos que implicam na compra de PROD_LIMPEZA ainda mantêm um padrão observado anteriormente. Foi observado que tal produto relaciona-se frequentemente com o grupo de produtos formado por produtos como, ARROZ, CAFE, FEIJAO, LIMP_ROUPAS, PROD_BANHO e PAPEL_HIGIENICO. Alguns trimestres anteriormente analisados, apresentaram a adição de novos produtos a esse grupo, como o caso de produtos como MACARRAO, OLEO_VEGETAL e ESPECIARIAS. Para este trimestre não foi diferente, como pode ser observado através das regras:

CAFE, LIMP_ROUPAS, MACARRAO → PROD_LIMPEZA

ESPECIARIAS, LIMP_ROUPAS, MACARRAO, PROD_BANHO → PROD_LIMPEZA

LIMP_ROUPAS, OLEO_VEGETAL, PROD_BANHO → PROD_LIMPEZA

As associações entre os produtos, ARROZ e CAFE, ARROZ e FEIJAO e também ARROZ e ACUCAR, foram frequentemente encontradas neste trimestre, o que leva a consolidação do conhecimento para estes produtos, como podemos ver nas regras:

ARROZ, CAFE, FEIJAO, PROD_BANHO → PROD_LIMPEZA

ACUCAR, ARROZ, LIMP_ROUPAS → FEIJAO

ARROZ, FEIJAO, PROD_BANHO, PROD_LIMPEZA → OLEO_VEGETAL

As relações entre produtos como, ESPECIARIAS e CAFE, ESPECIARIAS e FEIJAO, denotam um padrão observado nos trimestres analisados, representado pelas regras:

CAFE, ESPECIARIAS, MACARRAO, OLEO_VEGETAL → FEIJAO

CAFE, ESPECIARIAS, FEIJAO, MACARRAO → OLEO_VEGETAL

ACUCAR, CAFE, ESPECIARIAS, FEIJAO → PROD_BANHO

ACUCAR, ESPECIARIAS, FEIJAO, PROD_BANHO → OLEO_VEGETAL

A relação entre CREME_DE_LEITE e LEITE também demonstra um padrão observado anteriormente em outros trimestres, o que nos leva a consolidação da implicação de que quem compra CREME_DE_LEITE possivelmente comprará LEITE, e que pode ser observado pela a regra: CREME_DE_LEITE → LEITE

Também foi analisado um comportamento frequente ao longo do estudo, ao se observar a associação entre produtos como FEIJAO e MACARRAO, conforme regras:

FEIJAO MACARRAO PROD_BANHO PROD_LIMPEZA → ACUCAR

ACUCAR FEIJAO MACARRAO PROD_BANHO → PROD_LIMPEZA

ARROZ FEIJAO MACARRAO PROD_LIMPEZA → OLEO_VEGETAL

Apesar desse trimestre possuir um número de compras menor do que outros trimestres, foi possível identificar padrões no comportamento das vendas realizadas no mercado, principalmente ao se encontrar produtos como ESPECIARIAS, que são amplamente encontrados justamente no mesmo período de estudo, correspondente ao trimestre 2 e também ao trimestre 3. Dessa forma, será possível assimilar todo o conhecimento adquirido dentro desses dez trimestres analisados, com intuito de distribuir o conhecimento adquirido ao longo deles, sendo possível demonstrar quais produtos frequentemente são comprados em cada época do ano, tornando possível assim a fomentação de estratégias futuras para o mercado.

3.3 Considerações Finais

Ao fazer as análises dos trimestres, foi possível perceber que a quantidade de compras realizadas ao longo dos trimestres diminuíram bruscamente. Não se sabe os fatores que levaram a essa diminuição contínua, mas é interessante ressaltar, que foi possível analisar um comportamento de compras frequentes, mesmo diante a diminuição de vendas demonstrada pelo total de compras realizadas para cada trimestre, como pode ser observado na Figura 3.20.

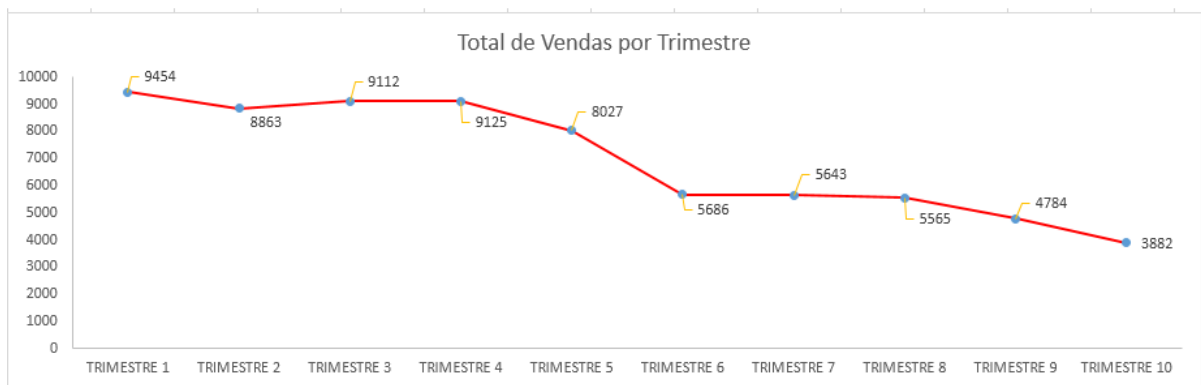


Figura 3.20: Total de Compras por Trimestre

Diante desta situação, foi possível analisar as compras e extrair algum conhecimento do comportamento de compras dos clientes nesses 10 trimestres. Para muitos dos trimestres analisados, conseguiu-se identificar uma relação forte entre itens de consumo em geral, como por exemplo: ARROZ, FEIJAO, CAFE, MACARRAO, PROD_BANHO entre outros, com o produto PROD_LIMPEZA. Isso sugere como conhecimento adquirido que, a compra de produtos como ARROZ, CAFE, FEIJAO, LIMP_ROUPAS, PROD_BANHO, PAPEL_HIGIENICO, MACARRAO, OLEO_VEGETAL implicaria na compra de produtos de limpeza em geral.

As relações entre ARROZ e FEIJAO e, ARROZ e ACUCAR, também apareceram quase que constantemente em todos os trimestres analisados, porém, é de sabedoria comum que estes produtos, assim como os aqui já mencionados, estão dentro de um padrão de compra nacional, visto que são uns dos alimentos mais consumidos no país e que fazem parte de um conjunto de alimentos de consumo básico.

A relação do produto ESPECIARIAS, que constituem temperos em geral, com produtos como CAFE e FEIJAO, apresentou duas relações bastante interessantes. A relação entre ESPECIARIAS e CAFE, é um aprendizado que pode-se absorver como conhecimento adquirido, visto que ela não apresenta uma certa obviedade quanto a relação entre ESPECIARIAS e FEIJAO.

Um acontecimento importante, foi a constatação de que alguns produtos, diferentes dos que já foram mencionados como uso básico, seguiam um padrão de consumo específico. Portanto, ao se fazer uma análise mais específica entre trimestres pertencentes ao mesmo período do ano foi possível identificar e tomar conhecimento que, para uma determinada época do ano, adquiriu-se os seguintes conhecimentos:

Trimestre - De Janeiro até Março: Para este trimestre, foi possível identificar relações antes não observadas em outros trimestres. Chamou a atenção, a frequência com que alguns produtos como: CREME_DE_LEITE, LEGUMES, LEITE_CONDENSADO, REFRIGERANTE e SUCO apareceram. Para estes produtos, foi notado que havia uma presença maior nas regras geradas pela ferramenta WEKA, e com isso, pode-se entender que esses produtos são mais consumidos nessa determinada faixa do ano. Portanto é de grande importância, utilizar estratégias baseando-se nesse conhecimento adquirido, pois

de fato, foi possível captar essa informação através da análise dos seguintes trimestres: Trimestre 1, Trimestre 5 e Trimestre 9.

Trimestre - De Abril até Junho: Neste período estabelecido, foi observado um padrão referente ao consumo do produto OLEO_VEGETAL, e que não foi encontrado ao se comparar com trimestres que não fazem parte desse período do ano. Portanto, foi gerado um conhecimento específico para este produto mencionado, que podem auxiliar na definição de estratégias de vendas futuras.

Trimestre - De Julho até Setembro: Para os trimestres analisados (Trimestre 3 e Trimestre 7), e que fazem parte do período do ano mencionado, foi percebido um consumo específico de produtos como CREME e CAFE. Os dois produtos em conjunto puderam ser encontrados somente neste período do ano. Uma possível explicação para tal, é que este período compreende a estação do inverno, portanto, se crê que o consumo de produtos como CREME, que são produtos relacionados a hidratação da pele, esteja relacionado com o frio que é acometido por esta determinada fase do ano. A mesma análise foi feita para o produto CAFE. Em muitas das regras encontradas para este período, quase sempre esses dois produtos apareciam juntos, o que demonstra a validade do que foi analisado.

Trimestre - De Outubro até Dezembro: Para o período em questão, pode-se observar que o produto PROD_BANHO, teve um grande número de aparição nas regras geradas. Isso de fato foi percebido para os dois trimestres analisados e que correspondem a faixa do ano especificada, que são o trimestre 4 e o trimestre 8. Ambos mostraram regras já mencionadas antes a respeito do consumo de produtos de uso comum, porém, a maioria delas continham o produto mencionado fazendo com que fosse possível estabelecer essa análise e que podem auxiliar em posteriores tomadas de decisão.

A conclusão que se pode tirar a partir das análises é que o mercado vende muitos produtos que são considerados de uso comum. Também foi observado que o mercado possui um padrão de vendas bem definido. Isso foi possível observar pois praticamente os mesmos produtos eram consumidos em um mesmo período do ano.

4 Conclusão

Como conclusão de caso de estudo, percebe-se que foi possível analisar os dados cedidos pelo mercado através da utilização de técnicas de extração de conhecimento, em especial a mineração de dados, a qual possibilitou a percepção de relacionamentos entre os produtos consumidos do mercado em estudo, tendo como intuito encontrar padrões de consumo de seus clientes. Esta abordagem é bastante conhecida na literatura como Análise do cesto de compras (Market Basket Analysis).

Através das análises das relações aqui apresentadas, foi possível demonstrar que na maioria das vezes a compra de um determinado conjunto de produtos, induziam na compra de um produto de limpeza e até mesmo foi possível demonstrar quais os produtos melhor se relacionavam para cada trimestre analisado.

Foi possível também fazer uma análise pontual para cada trimestre de um ano, o que possibilitou entender o padrão de consumo dos clientes dentro de um ano. Para isso, foi apresentado uma série de produtos e seus relacionamentos, de forma que fosse possível ter, por exemplo, a noção do comportamento de compra entre os trimestres de um mesmo período e também o comportamento de compra dos produtos num âmbito geral. Assim, foi identificado que as relações iriam ser interpretadas de forma clara e coesa, com a intenção de gerar um certo conhecimento para a eventual tomada de decisão.

Os resultados atingidos foram satisfatórios, e permite através do conhecimento gerado, gerar subsídios necessários para o mercado, de forma que seja possível definir algumas estratégias, tendo como objetivo sua utilização para posteriores tomadas de decisões, podendo fazer disso um eminente diferencial competitivo.

4.1 Trabalhos Futuros

Espera-se que através dos conhecimentos obtidos pela aplicação do algoritmo APRIORI na base de dados do mercado, possa ser possível definir algumas estratégias que poderão abordar vendas casadas de produtos que se relacionam conforme demonstrados neste

trabalho. Por exemplo, conforme conhecimentos obtido a respeito do padrão de compra dos clientes para os trimestres que correspondem ao período de Julho até Setembro, poderá ser definido algo para que se possa alavancar as vendas dos produtos como CAFE e CREME, visto que esses produtos foram mais encontrados nessa época do ano.

Referências Bibliográficas

- Agrawal, R., Imieliński, T., and Swami, A. (1993). Mining association rules between sets of items in large databases. In *ACM SIGMOD Record*, volume 22.
- Angelo, C. F. d., João Paulo de Lara, S., and Fávero, L. P. L. (2003). As compras não planejadas em supermercados: a importância do tempo e da organização da loja na determinação dos gastos. *Revista de Administração Contemporânea*.
- Campos Filho, M. P. d. (1994). Os sistemas de informação e as modernas tendências da tecnologia e dos negócios. *Revista de Administração de Empresas*, 34(6):33–45.
- de Vasconcelos, Lívia Maria Rocha e de Carvalho, C. L. (2004). Aplicação de regras de associação para mineração de dados na web. *Instituto de Informática da Universidade Federal de Goiás*.
- Gahane, S. T. (2013). The conceptual overview: Challenges and opportunities with big data.
- Garner, S. R. et al. (1995). Weka: The waikato environment for knowledge analysis. In *Proceedings of the New Zealand computer science research students conference*.
- Han, J., Kamber, M., and Pei, J. (2011). *Data mining: concepts and techniques: concepts and techniques*. Elsevier.
- Kaur, M. and Kang, S. (2016). Market basket analysis: Identify the changing trends of market data using association rule mining. *Procedia Computer Science*, 85:78–85.
- Liu, Y. (2010). Study on application of apriori algorithm in data mining. In *Computer Modeling and Simulation, 2010. ICCMS'10. Second International Conference on*, volume 3, pages 111–114. IEEE.
- Mañas, A. V. (2005). *Administração de sistemas de informação*, volume 1. Érica.

- Quesado, F. J. (2014). O desafio da inteligência competitiva. *Revista Inteligência Competitiva*, 4(1):75–78.
- Tanna, P. and Ghodasara, Y. (2014). Using apriori with weka for frequent pattern mining. *arXiv preprint arXiv:1406.7371*.
- Yan, L. (2015). The application of data mining technology in e-commerce. In *2015 International Conference on Education, Management, Information and Medicine*. Atlantis Press.
- Zhang, S., Zhang, C., and Yang, Q. (2003). Data preparation for data mining. *Applied Artificial Intelligence*, 17(5-6):375–381.