



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE DESENVOLVIMENTO SUSTENTÁVEL DO SEMIÁRIDO
UNIDADE ACADÊMICA DE ENGENHARIA DE PRODUÇÃO
CURSO DE ENGENHARIA DE PRODUÇÃO**

RAVENNA LINS RODRIGUES

**UMA ABORDAGEM DE *DATA MINING* E *BUSINESS INTELLIGENCE* PARA
OTIMIZAÇÃO DO PROCESSO PRODUTIVO E APOIO À TOMADA DE
DECISÃO NA AGROINDÚSTRIA**

**SUMÉ - PB
2023**

RAVENNA LINS RODRIGUES

**UMA ABORDAGEM DE *DATA MINING* E *BUSINESS INTELLIGENCE* PARA
OTIMIZAÇÃO DO PROCESSO PRODUTIVO E APOIO À TOMADA DE
DECISÃO NA AGROINDÚSTRIA**

Monografia apresentada ao Curso Superior de Engenharia de Produção do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharela em Engenharia de Produção.

Orientador(a): Professora Dr^a. Cecir Barbosa de Almeida Farias.



R696a Rodrigues, Ravenna Lins.

Uma abordagem de data mining e business intelligence para otimização do processo produtivo e apoio à tomada de decisão na agroindústria. / Ravenna Lins Rodrigues. - 2023.

63 f.

Orientadora: Professora Dra. Cecir Barbosa de Almeida Farias.

Monografia - Universidade Federal de Campina Grande; Centro de Desenvolvimento Sustentável do Semiárido; Curso de Engenharia de Produção.

1. Data mining. 2. Business intelligence. 3. Agroindústria - processo produtivo. 4. Tomada de decisão. 5. Mineração de dados. 6. Modelagem de dados. 7. Agricultura familiar. 8. Agrupamento com k-means. 9. Dashboard. 10. Diagrama de use case. I. Farias, Cecir Barbosa de Almeida. II. Título.

CDU: 658:004(043.1)

Elaboração da Ficha Catalográfica:

Johnny Rodrigues Barbosa
Bibliotecário-Documentalista
CRB-15/626

RAVENNA LINS RODRIGUES

**UMA ABORDAGEM DE *DATA MINING* E *BUSINESS INTELLIGENCE* PARA
OTIMIZAÇÃO DO PROCESSO PRODUTIVO E APOIO À TOMADA DE
DECISÃO NA AGROINDÚSTRIA**

Monografia apresentada ao Curso Superior de Engenharia de Produção do Centro de Desenvolvimento Sustentável do Semiárido da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharela em Engenharia de Produção.

BANCA EXAMINADORA:

Professora Dra. Cecir Barbosa de Almeida Farias
Orientadora – UAEP/CDSA/UFCG

Professor Me. Josean da Silva Lima Júnior
Examinador I – UAEP/CDSA/UFCG

Professor Dr. Tiago Gonçalves Pereira Araújo
Examinador II – UATEC/CDSA/UFCG

Trabalho aprovado em: 06 de fevereiro de 2023.

SUMÉ - PB

“A homeostase é fundamental para resiliência”.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me conceder saúde, força, sabedoria e discernimento, mantendo-me firme diante dos obstáculos e me acolhendo na sua divina bondade. Aos meus pais, Maria do Socorro Lins Rodrigues e Jairo Rodrigues da Silva que desde o princípio investiram na minha educação, além de se fazerem presentes me incentivando e confiando, dedico minha eterna gratidão. Ao meu irmão Rodrigo Lins Rodrigues que me apoia, inspira e motiva para o alcance dos meus objetivos. Ao meu esposo Vinicius Costa Amador por todo apoio, auxílio, amor, companheirismo e compreensão. À minha filha Sophia Lins Amador, minha maior fortaleza e motivação, nasceu durante essa trajetória e ressignificou a minha vida. À minha cunhada Simone Lins que sempre me incentivou e com palavras de carinho acalmou os dias de maior turbulência. Às minhas primas Suelen Lins e Izabel Barbosa pelo apoio e companheirismo e aos demais familiares que se fizeram presentes independente da distância. À minha orientadora Prof^a. Dr^a. Cecir Barbosa de Almeida Farias, por todas as oportunidades, por confiar e acreditar no meu potencial, pela orientação, dedicação, e, principalmente, pela amizade. Aos professores que atuam como porta-voz do conhecimento, ponte que nos leva não só por rotas já conhecidas, mas por caminhos nunca antes percorridos. À Universidade Federal de Campina Grande que foi essencial para o meu processo de desenvolvimento pessoal e profissional através da formação acadêmica. Aos órgãos de financiamento à pesquisa e desenvolvimento que contribuíram no incentivo e apoio do meu processo de aprendizagem. À todos os colaboradores da instituição, que se dedicam diariamente às suas funções, proporcionando um ambiente confortável e seguro. À gestão da organização, onde a presente pesquisa foi realizada, por confiar e ser solícita quanto às informações fornecidas. À cidade de Sumé onde fui imensamente feliz, dedico meu amor e consideração. Agradeço, também, aos meus amigos de turma, em especial, a João Vitor da Silva, Gabriel Mendes, Yasmin Siqueira, Lucas Matheus, Danniely Silva, Laís Vaz, Eliel Gomes e George Lacerda. À Cida, Iudo, Stefanny Nogueira, Gilson Simões, Aldeâne Braz e em especial à Francisca Alves por me acolherem e estarem presentes no meu dia-a-dia contribuindo de forma significativa. À Celma Mendes (*in memoriam*) que apesar de não estar presente fisicamente sei que está me guiando e protegendo. À Antônio Marques e Rafael Maia que estão sempre vibrando com as minhas conquistas. E por fim, às pessoas que conheci ao longo dessa trajetória e que de alguma forma contribuíram para essa realização.

RESUMO

A evolução e o uso massivo das tecnologias elucidam o crescimento exponencial de dados. Esse volume de dados tem se traduzido para as organizações como desafios e oportunidades, na medida em que, extrair informações relevantes para suportar a tomada de decisão exige mudanças em processos e atuações. No entanto, tais mudanças de comportamento propiciam vantagens competitivas e sustentáveis às organizações. Dessa forma, a aplicação de técnicas de mineração de dados aliada ao *business intelligence* possibilita uma ampla visão sobre a organização, a partir da descoberta de novos modelos de atuação, monitoramento e melhoria dos processos já existentes. Nesse contexto, o presente estudo teve como objetivo a exploração de dados, por meio de aprendizagem não supervisionada, para identificação de informações implícitas capazes de subsidiar o processo decisório. Para isto foi desenvolvido um mapeamento sistemático de literatura, tendo em vista, a apropriação de conhecimento literário capaz de auxiliar nas fases da metodologia *CRISP-DM*. Como ferramenta de operação foram utilizados o *Orange Data Mining*, por apresentar-se como expansível e de baixo custo, além de suportar uma ampla variedade de técnicas de análise de dados, o *Google Sheets* e *Power BI* da *Microsoft* que contribuíram de forma efetiva para informatização e automatização do processo produtivo. Concluiu-se que os resultados alcançados atingiram os objetivos propostos no estudo, apresentando informações e possibilidades relevantes para otimizações dentro do processo produtivo. O modelo gerado foi capaz de transformar o volume de dados em informações importantes e o *dashboard* desenvolvido possibilitou a fácil interpretação dos resultados, influenciando positivamente nas decisões da gestão e gerando conhecimento agregado e de valor para a organização.

Palavras-chave: Mineração de Dados, Business Intelligence, Power Bi, Dashboard, Tomada de Decisão.

RODRIGUES, Ravenna Lins. **A data mining and business intelligence approach to optimize the production process and support decision-making in agroindustry.** 2023. 63f. (Bachelor Thesis). Industrial Engineering Course de Engenharia de Produção, Centro de Desenvolvimento Sustentável do Semiárido, Universidade Federal de Campina Grande, Sumé – Paraíba – Brazil, 2023.

ABSTRACT

The evolution and massive use of technologies elucidate the exponential growth of data. This volume of data has been translated into challenges and opportunities for organizations, as extracting relevant information to support decision-making requires changes in processes and actions. However, such behavior changes provide organizations with competitive and sustainable advantages. In this way, the application of data mining techniques combined with business intelligence enables a broad view of the organization, based on the discovery of new models of action, monitoring and improvement of existing processes. In this context, the present study aimed to explore data, through unsupervised learning, to identify implicit information capable of supporting the decision-making process. A systematic literature mapping was developed, considering the appropriation of literary knowledge capable of helping in the phases of the CRISP-DM methodology. As an operation tool, Orange Data Mining was used, as it is expandable and low cost, in addition to supporting a wide variety of data analysis techniques, Google Sheets and Microsoft Power BI, which contributed effectively to computerization and automation of the production process. It was concluded that the achieved results achieved the objectives proposed in the study, presenting relevant information and possibilities for optimizations within the production process. The generated model was capable of transforming the volume of data into important information and the developed dashboard made it possible to easily interpret the results, positively influencing management decisions and generating added and valuable knowledge for the organization.

Keywords: Data Mining, Business Intelligence, Dashboard, Power Bi, Decision Making.

LISTA DE FIGURAS

	Pg.
Figura 1 - Etapas do método <i>CRISP-DM</i>	30
Figura 2 - Diagrama de <i>Use Case - Dashboard</i> AGUBEL.....	43
Figura 3 - <i>Layout</i> da tela inicial do <i>dashboard</i>	44
Figura 4 - Ilustração do painel referente à análise do leite in natura.....	45
Figura 5 - Representação do quantitativo total de leite.....	45
Figura 6 - Visual adotado para apresentar a distribuição do quantitativo de leite por época.....	46
Figura 7 - Ilustração do filtro de data.....	46
Figura 8 - Representação do quantitativo de litros de leite por município.....	47
Figura 9 - <i>Layout</i> do visual adotado para representação da média dos parâmetros físico-químicos.....	48
Figura 10 - Ilustração dos tipos de não conformidades.....	48
Figura 11 - <i>Layout</i> da análise do leite pasteurizado.....	49
Figura 12 - Representação do painel referente à análise de <i>clusterização</i> do leite	50

LISTA DE QUADROS

	Pg.
Quadro 1 - As modalidades do PAA (2014-2019)	21
Quadro 2 - Identificação das necessidades e desejos dos usuários por meio da técnica de <i>brainstorming</i>	42

LISTA DE TABELAS

	Pg.
Tabela 1 - Distribuição dos estabelecimentos agropecuários existentes no Nordeste.....	23
Tabela 2 - Requisitos mínimos de qualidade físico-química do leite de cabra pasteurizado segundo a IN 37 do MAPA.....	25
Tabela 3 - Agrupamento com <i>dataset</i> pertencente a Sumé.....	37
Tabela 4 - Agrupamento com <i>dataset</i> referente aos municípios associados	38
Tabela 5 - Agrupamento com <i>dataset</i> pertencente ao leite pasteurizado.....	40

LISTA DE SIGLAS E ABREVIATURAS

AGUBEL - Associação Gestora da Usina de Beneficiamento de Lácteos

BI - *Business Intelligence*

CONAB - Companhia Nacional de Abastecimento

CONTAG - Confederação Nacional dos Trabalhadores da Agricultura

CUT Rural - Departamento Nacional dos Trabalhadores Rurais da Central Única dos Trabalhadores Rurais

CRISP-DM - *Cross Industry Standard Process for Data Mining*

DM - *Data Mining*

MAPA - Ministério da Agricultura, Pecuária e Abastecimento

MD - Mineração de Dados

MG - Gordura Mínima

MP - Medida Provisória

MST - Movimento dos Trabalhadores Rurais Sem-Terra

PAA - Programa de Aquisição de Alimentos

PNAE - Programa Nacional de Alimentação Escolar

PRONAF - Programa Nacional de Fortalecimento da Agricultura Familiar

SAN - Segurança Alimentar e Nutricional

SI - Sistemas de Informação

TI - Tecnologia da Informação

SUMÁRIO

1	INTRODUÇÃO.....	14
1.1	OBJETIVO GERAL.....	16
1.1.1	Objetivos específicos.....	16
1.2	JUSTIFICATIVA.....	17
2	REFERENCIAL TEÓRICO.....	18
2.1	AGRICULTURA FAMILIAR E MERCADOS INSTITUCIONAIS.....	18
2.2	PROGRAMA DE AQUISIÇÃO DE ALIMENTOS.....	20
2.3	CAPRINOCULTURA LEITEIRA E O ASSOCIATIVISMO.....	22
2.4	CARACTERÍSTICAS FÍSICO-QUÍMICAS DO LEITE DE CABRA.....	24
2.4.1	Fatores que Influenciam a Qualidade do Leite.....	25
2.5	O EMPREGO DA TECNOLOGIA NO MERCADO AGROPECUÁRIO.....	26
2.5.1	Gestão da Informação.....	26
2.5.2	Mineração de Dados.....	27
2.5.3	<i>Business Intelligence</i>.....	28
2.5.4	Tomada de Decisão.....	29
3.	PROCEDIMENTOS METODOLÓGICOS.....	29
3.1	ENTENDIMENTO DO NEGÓCIO (<i>BUSINESS UNDERSTANDING</i>).....	30
3.2	ENTENDIMENTO DOS DADOS (<i>DATA UNDERSTANDING</i>).....	30
3.3	PREPARAÇÃO DOS DADOS (<i>DATA PREPARATION</i>).....	31
3.4	MODELAGEM (<i>MODELING</i>).....	31
3.4.1	Exploração de dados.....	31
3.4.2	Agrupamento com o algoritmo <i>K-means</i>.....	32

3.5	AVALIAÇÃO (<i>EVALUATION</i>) E IMPLEMENTAÇÃO (<i>DEPLOYMENT</i>).....	32
4.	RESULTADO E DISCUSSÃO.....	33
4.1	CARACTERIZAÇÃO DO NEGÓCIO.....	33
4.1.1	Descrição dos Dados.....	34
4.2	ANÁLISE EXPLORATÓRIA E MODELAGEM DOS DADOS.....	34
4.2.1	Análise Exploratória de Dados.....	34
4.2.2	Análise de Agrupamento com <i>K-means</i>.....	36
4.3	<i>BUSINESS INTELLIGENCE</i> NA CONSTRUÇÃO DE UM <i>DASHBOARD</i>	42
4.3.1	<i>Brainstorming</i> das Necessidades e Desejos dos Usuários....	42
4.3.2	Diagrama de <i>Use Case</i>.....	43
5	DESENVOLVIMENTO DO <i>DASHBOARD</i>.....	44
6	CONSIDERAÇÕES FINAIS.....	50

REFERÊNCIAS

APÊNDICE

ANEXO

1 INTRODUÇÃO

A pecuária de pequenos ruminantes é uma atividade incontestável quando se tomam por referência as zonas áridas e semiáridas do planeta, seja no âmbito sociocultural, pela fixação do homem ao campo, ou no econômico, como fonte de renda. Esta representa, para o semiárido brasileiro, uma das mais importantes atividades do agronegócio e tem se constituído em um dos principais fatores para segurança alimentar e geração de emprego das famílias rurais (LIMA 2009).

Os caprinos são animais de dupla aptidão, porém, são mais eficientes como animais leiteiros do que como animais de corte, apresentando resultados rápidos em comparação às outras atividades pecuárias (SIMPLÍCIO *et al.*, 2000). No Brasil a caprinocultura leiteira é bastante presente em pequenas propriedades já que a criação de cabras necessita de pouco espaço e se identifica como sendo de fácil manejo (FELISBERTO, 2016). Desse modo, a caprinocultura leiteira apresenta-se como uma alternativa para regiões áridas e semiáridas, tornando-se assim uma estratégia para melhoria da renda de seus produtores.

Conforme o censo brasileiro de 2017, a caprinocultura leiteira tem se tornado importante no âmbito social e econômico, principalmente no que se refere à região Nordeste, que possui neste ano possuía 92,8% do total de 8.260.607 caprinos presente em território nacional (IBGE, 2017).

O rebanho paraibano é o quinto maior do Brasil (546.036 cabeças), e o primeiro em produção leiteira com cerca de 5.627.000 litros/ano (22% da produção nacional), tendo o município de Monteiro como o maior produtor (684 mil litros/ano) e detentor do maior rebanho, seguido de Sumé (491 mil litros/ano) (IBGE, 2017).

Contudo, destaca-se que a caprinocultura na região do Cariri Paraibano, em sua maioria, está relacionada aos pequenos produtores e agricultura familiar, sem grandes investimentos tecnológicos, uma vez que, inicialmente era produzido leite apenas para o sustento da família.

De acordo com Souza e colaboradores (2011), a região do Cariri Paraibano apresentava baixos rendimentos em relação à produção de leite, não ultrapassando o total de um litro cabra/dia. No anseio, ou pela necessidade, de melhorar a renda familiar, o produtor rural buscou explorar as oportunidades ao seu entorno, visando a lógica comercial que a região oferece para a caprinocultura leiteira.

Nesse contexto, a abertura para comercialização do leite caprino produzido por agricultores familiares para o Programa de Aquisição de Alimentos (PAA) foi determinante para impulsionar essa atividade (BATISTA *et al.*, 2012). Dessa forma, a compra garantida de

quase totalidade do leite caprino produzido no Nordeste pelo referido programa fez com que vários sistemas pecuários mistos de base familiar passassem a explorar a caprinocultura leiteira como atividade comercial na Paraíba, promovendo, assim, significativas mudanças nos sistemas de produção do Estado (SAMPAIO *et al.*, 2009; MENESES, 2015).

Entretanto, apesar do avanço da caprinocultura leiteira Paraibana, com o incentivo de programas governamentais (GOVERNO DA PARAÍBA, 2019) e mesmo possuindo um ambiente flexível pela interação dos interesses coletivos, nem sempre a produtividade corresponde às expectativas do homem do campo. Tal fato está correlacionado à entressa que dificultam o crescimento da atividade, como: a ampla visão acerca do processo produtivo, fato que pode ser determinante na produtividade e resultados financeiros inexpressivos.

De acordo com Farias e colaboradores (2019), a falta de informações pertinentes que possibilitem a caracterização do sistema produtivo limita o potencial de crescimento e autocorreção das ações, sendo altamente prejudicial (FARIAS *et al.*, 2019). Em contrapartida, o diagnóstico, de forma rápida e precisa, das práticas realizadas nas propriedades permite determinar os principais obstáculos para o crescimento da atividade (SILVA *et al.*, 2015; CARVALHO *et al.*, 2020). Desse modo, torna-se imprescindível a busca por mecanismos que otimizem o sistema a fim de alcançar melhores resultados e conseqüentemente uma maior rentabilidade para o produtor.

1.1 OBJETIVO GERAL

Otimizar o processo produtivo e subsidiar a tomada de decisão em uma agroindústria, por meio do uso de ferramentas de *business intelligence* e da aplicação de técnicas de mineração de dados.

1.1.1 Objetivos Específicos

- Explorar, identificar e analisar informações implícitas em banco de dados;
- Construir modelos de aprendizado não supervisionado por meio do agrupamento de amostras similares;
- Gerar indicadores que permitam avaliar a performance do processo;
- Garantir facilidade operacional na combinação com outras informações assim como torna-las de simples entendimento e compreensão;
- Avaliar, implementar e validar os modelos construídos.

1.2 JUSTIFICATIVA DO ESTUDO

As organizações têm passado por um crescimento exponencial de geração de dados a uma velocidade e variedade cada vez maiores (ATZMUELLER, 2016; JANVRIN e WATSON, 2017). Tal fato, tem sido posto como um dos principais desafios enfrentados pelas organizações, na medida em que extrair informações significativas para apoiar no processo de tomada de decisão (BUGNION, MANIVANNAN e NICOLAS, 2017), tem exigido significantes modificações nos modelos, processos e visão de negócios das mesmas (LOPES e SOARES, 2016).

Nesse sentido, compreendendo o contexto de mercado como incerto e em constantes modificações, evidencia-se que não adianta somente armazenar dados, se não forem reunidos esforços para selecioná-los e utilizá-los de forma organizada e estruturada, como recurso para gerar vantagem e competitividade no mercado.

Dessa forma, tendo em vista que essa mudança de atitude pode ser crítica para o desenvolvimento e permanência das organizações no mercado (MARQUES *et al.*, 2017), gradativamente as empresas estão internalizando em sua cultura o uso de dados para tomada de decisão (PEREIRA, 2016).

Face a esta realidade, surge a *Business Intelligence* (BI), que é um conjunto de ferramentas que possibilita criar, explorar e visualizar dados de forma interativa e dinâmica, além de oferecer aos gestores e analistas de negócios uma análise adequada na condição de servir como base ao refinamento e aprimoramento das atuais decisões (TURBAN, 2009).

A partir dessas perspectivas surge então este trabalho de conclusão de curso, com o intuito de aplicar técnicas de mineração de dados, no processamento e geração de informações, e utilizar ferramentas de BI na construção de painéis interativos. Esses painéis terão como finalidade a exposição de informações, acerca do planejamento e controle da qualidade, de forma concisa e objetiva tendo em vista subsidiar tomadas de decisões ágeis e seguras, garantindo assim, a excelência e a melhoria contínua dos processos organizacionais.

2. REFERENCIAL TEÓRICO

2.1 AGRICULTURA FAMILIAR E MERCADOS INSTITUCIONAIS

Para Abramovay (1998), a agricultura familiar é mais que um segmento econômico e social, sendo classificada como um valor. Esta representa um importante pilar para a economia e o abastecimento do Brasil, sendo responsável, segundo o Censo Agropecuário 2017, por 77% dos estabelecimentos agropecuários do país, 23% do valor da produção agropecuária total e 67% do total de trabalhadores nos estabelecimentos agropecuários (IBGE, 2019).

No entanto, o tratamento dado aos agricultores familiares no Brasil permaneceu sem avanços significativos durante muito tempo, até então preponderava uma forte associação entre essa forma social de produção e a precariedade estrutural. A ênfase, portanto, recaiu na instituição de políticas compensatórias voltadas ao que então se denominava “pequena produção”, isto é, à trabalhadores rurais, meeiros, lavradores e outros atores sociais marcados pelo estigma da pobreza e pela irrelevância enquanto objeto da intervenção dos poderes públicos.

Na década de 1990 ocorreram sucessivos anos de seca que afetaram de forma expressiva a agricultura dos estados meridionais do país. A crise experimentada pelos pequenos produtores desaguou em manifestações multitudinárias lideradas pela Confederação Nacional dos Trabalhadores da Agricultura (CONTAG), Movimento dos Trabalhadores Rurais Sem-Terra (MST), Departamento Nacional dos Trabalhadores Rurais da Central Única dos Trabalhadores Rurais (CUT Rural) e pelos setores progressistas da igreja. As jornadas nacionais de luta emergem dentro do que se veio a chamar “Gritos da Terra Brasil”. E as reivindicações foram direcionadas, sobretudo, para a problemática que tratava de terra, crédito rural e justiça social.

Do ponto de vista acadêmico, os anos noventa convergiram para a aparição de estudos (ABRAMOVAY, 1992; VEIGA, 1991, LAMARCHE 1993) dedicados a demonstrar que a agricultura familiar consistiu na forma social dominante para o desenvolvimento agrário dos países capitalistas avançados. Surgiram pesquisas (FAO, 1994; 1996; BERGAMASCO, 1993) evidenciando que os governos haviam subestimado a importância desse tipo de exploração para o abastecimento alimentar do país.

Desse modo, conforme destacou Leite (2004), a publicação destes estudos supôs uma mudança radical na concepção do papel da agricultura familiar na economia brasileira, rompendo-se a associação pejorativa que a vinculava com a produção de subsistência e com a implantação de políticas de caráter social e meramente compensatórias.

O resultado da pressão exercida desde o âmbito político, mas também acadêmico, foi a criação de programas sociais e políticas públicas, específicos e abrangentes, voltados para a esfera da agricultura familiar, sendo eles: o Programa Nacional de Fortalecimento da Agricultura Familiar (PRONAF), originado em 1995 com o qual o Brasil passa a contar, pela primeira vez em sua história, com uma política específica de crédito (custeio e investimento); o Programa de Aquisição de Alimentos (PAA), instituído em 2003 por meio do qual o governo federal assegura a compra antecipada dos produtos da agricultura familiar, ou como se denomina, dos chamados “beneficiários produtores” que são destinados aos “beneficiários consumidores” (asilos, creches, hospitais) e o Programa Nacional de Alimentação Escolar (PNAE), criado em 2009 o qual oferece alimentação escolar e ações de educação alimentar e nutricional a estudantes de todas as etapas da educação básica pública (CAMARGO; BACCARIN; SILVA, 2016).

Diante do exposto, Amartya Sen (2000) enfatiza em sua obra, que os mercados representam uma oportunidade singular de realizar a inclusão social, com o qual os os setores mais frágeis da sociedade devem ter o direito de exercer sua cidadania através da possibilidade de comercializar o fruto do seu trabalho e/ou de sua produção.

Fornazier e Belik (2019), destacam que “o mercado institucional tem sido descrito como uma possibilidade de os governos atuarem no incentivo de determinadas atividades; principalmente, promovendo a inserção de atores sociais locais”. Para Vogt (2009), esses mercados abrangem as três esferas governamentais (municipal, estadual e federal) em todas as suas realizações de compra de alimentos, configurando-se como oportunidade para o segmento em questão.

Institucionalizados por meio de programas federais, tais mercados geram efeitos positivos aos agricultores familiares. Assim, programas como o PAA e o PNAE induzem o poder público a desenvolver os próprios mecanismos de compras públicas, colaborando com a valorização da produção local/regional e ecológica/orgânica, promovendo ações de justiça social, equidade e qualificação dos produtos da agricultura familiar (GRISA; SCHNEIDER, 2014).

O acesso aos mercados institucionais suscita a aproximação do consumo com a produção familiar. Das iniciativas vindas das políticas públicas, a que possivelmente possibilitou maior melhoria na qualidade de vida de diversos produtores rurais foi o PAA, que se baseou nos seguintes objetivos:

i) incentivar a produção de alimentos pelos agricultores familiares mais pobres, contribuindo assim para sua segurança alimentar; ii) gerar renda entre os agricultores familiares mais pobres com a venda do excedente de sua produção ao governo federal; iii) incentivar a criação ou o desenvolvimento de canais de comercialização da produção familiar nas comunidades onde estes eram frágeis ou inexistentes; iv) ampliar os estoques de alimentos para a distribuição pelos programas alimentares, procurando garantir o acesso aos alimentos em quantidade, qualidade e regularidade necessárias às populações em situação de insegurança alimentar e nutricional) promover a inclusão social no campo por meio do fortalecimento da agricultura familiar (PERACI; BITTENCOURT, 2010).

2.2 PROGRAMA DE AQUISIÇÃO DE ALIMENTOS

Em 2 de julho de 2003, o Governo Federal instituiu por meio do art. 19 da Lei nº 10.696, o Programa de Aquisição de Alimentos (PAA) visando executar ações relacionadas à política agrícola e de segurança alimentar e nutricional (SAN), com vistas a fomentar a agricultura familiar (ALMEIDA, PERIN. e SAMBUICHI, 2019). O programa, operado principalmente pela Companhia Nacional de Abastecimento (CONAB) executou recursos ao longo de dezoito anos de existência, até que a Medida Provisória (MP) no 1.061, de 9 de agosto de 2021, o substituiu pelo Programa Alimenta Brasil, o qual foi criado nos mesmos moldes do anterior e manteve a maior parte das suas modalidades e finalidades (BRASIL, 2021).

O PAA teve como público alvo agricultores familiares, assentados da reforma agrária, silvicultores, aquicultores, pescadores artesanais, indígenas e integrantes de comunidades de quilombolas rurais e demais povos e comunidades (BRASIL, 2015).

Uma das principais inovações trazidas pelo PAA foi a realização de compras públicas com dispensa de licitação, cujas normas estão previstas na Lei nº 8.666/1993.2 A lei de criação do programa definiu esse mecanismo na compra de alimentos, desde que os preços praticados não sejam superiores aos dos mercados regionais (BRASIL, 2003). Esse instrumento facilita a inserção de agricultores familiares no programa, visto que reduz os trâmites burocráticos inerentes à lei de licitações (SAMBUICHI *et al.*, 2014).

As modalidades de operação do programa, foram definidas na época, como: Compra direta; Compra com doação simultânea; PAA-Leite; Formação de estoque; Compra institucional e Aquisição de sementes; operar em seis modalidades distintas. O Quadro 1 apresenta uma descrição das modalidades do PAA previstas nos Decretos nos 7.775/2012 e 8.293/2014, incluindo os executores, os limites de compras, o tipo de acesso e as fontes de recursos (BRASIL, 2014).

Quadro 1 - As modalidades do PAA (2014-2019)

Modalidade	Descrição	Executor	Tipo de acesso do agricultor	Fonte de recurso	Limite de compra – por beneficiário/ano (R\$)
Compra Direta	Compra de produtos definidos pelo GGPA; objetivo de sustentação de preços	Conab	Individual ou em grupo (formal ou informal)	Ministério da Cidadania ou Mapa	8.000,00
Compra com Doação Simultânea	Compra de alimentos para doação imediata às entidades	Estados e municípios e/ou Conab	Individual ou em grupo (formal ou informal)	Ministério da Cidadania	6.500,00 ¹ ou 8.000,00 ²
PAA-Leite	Compra de leite para doação às pessoas em situação de Insan, operada nos estados do Nordeste e municípios do semiárido mineiro	Governos estaduais	Individual ou em grupo (formal ou informal)	Ministério da Cidadania	9.500,00
Formação de Estoque	Apoio financeiro para a formação de estoques de alimentos para posterior comercialização e devolução de recursos ao poder público	Conab	Cooperativa e associação	Mapa	8.000,00
Compra Institucional	Compra de produtos da agricultura familiar via chamada pública; atendimento de demandas do órgão comprador (consumo de alimentos, de sementes e de outros materiais)	Ente interessado	Cooperativa e associação	Ente interessado	20.000,00
Aquisição de Sementes	Aquisição de sementes, mudas e materiais propagativos para alimentação humana ou animal; doação a beneficiários consumidores ou fornecedores.	Conab	Cooperativa e associação	Ministério da Cidadania	16.000,00

Fontes: Brasil (2014)

A organização em cooperativas ou associações foi incentivada direta ou indiretamente pelo PAA, ao exigir que os fornecedores estivessem organizados institucionalmente para o acesso em algumas modalidades. Ademais, o limite de venda para as cooperativas é superior àquele praticado nas vendas individuais, o que também estimula a participação dos agricultores familiares em organizações formais.

A relação entre o acesso ao programa e a participação em organizações parece se retroalimentar. É importante destacar que o programa atua tanto fortalecendo as cooperativas e as associações existentes, havendo resultados mais efetivos do PAA em municípios com organizações estruturadas (GRISA *et al.*, 2010; SILVA e FERREIRA, 2016), quanto estimulando a formação de novas organizações dos agricultores (SILVA e SILVA, 2011; SANTOS *et al.*, 2019; OLIVEIRA E FERRANTE, 2013).

2.3 CAPRINOCULTURA LEITEIRA E O ASSOCIATIVISMO

A pecuária representa a criação de animais, sejam eles bovinos, suínos, equinos, caprinos, entre outros, com o intuito de fornecer alimentos para a população e matérias-primas para as indústrias. A pecuária costuma ser classificada em dois tipos: a pecuária de corte, destinada ao abastecimento de carne aos mercados interno e externo e a pecuária de leite, destinada à produção de leite e derivados (RESENDE, 2012). Dentro da pecuária brasileira, uma das práticas que cresceu ao longo dos anos e vem tendo grande desenvolvimento é a Caprinocultura.

Segundo Fonseca e colaboradores (2012), a criação de caprinos no Brasil teve início no período da colonização e o rebanho nacional era constituído, basicamente, por animais sem raça definida. Entretanto, foi só a partir da década de 70, que surgiram as primeiras associações de produtores de leite de cabra e também ocorreram as primeiras importações de animais de raças leiteiras. Com isso, a atividade se desenvolveu no Brasil (FONSECA; BRUSCHI, 2009) e na década de 80 a criação caprina se direcionou também para produção de leite em sua forma líquida e posteriormente em pó.

Nesse sentido, a formação de associações, surgiu como uma forma de fortalecer essa atividade e garantir a permanência do agricultor familiar no meio em que vive, consolidando planos estratégicos para o desenvolvimento dessa categoria.

O associativismo funciona como um espaço de articulação e reflexão sobre diversas ações, buscando o desenvolvimento social, econômico e cultural dos associados (LISBOA e ALCANTARA, 2015). Nas associações comunitárias, os associados compartilham problemas, produzem e comercializam, através da cooperação mútua, que é a base da sua sustentação (FAGOTTI, 2017).

Ao se associar em grupos, os agricultores podem visualizar novos caminhos para o mercado, sendo um meio facilitador para obtenção de créditos agrícolas, permitindo assim transformar a realidade local a partir do esforço coletivo, tornando-se um caminho eficaz para participar do mercado em melhores condições de concorrência.

Desse modo, as usinas de beneficiamento de leite ganharam espaço e a confiança dos produtores através do fornecimento do leite como matéria- prima, e da população que é beneficiada. Estas usinas são fontes de escoamento certo da produção leiteira e geralmente são vinculadas a associações comunitárias.

No nordeste brasileiro a caprinocultura leiteira vem recebendo importante incremento no número de criatórios e na qualidade dos rebanhos (ARAÚJO, 2008). Na tabela 1 pode-se

observar a quantidade de estabelecimentos agropecuários existentes no Nordeste, com suas respectivas quantidades de propriedades produtoras de leite caprino e quantidade produzida por estado.

Tabela 1 - Distribuição dos estabelecimentos agropecuários existentes no Nordeste

Estados	Quantidade de propriedades com caprinos	Quantidade de propriedade produtoras de leite de cabra	Quantidade de leite de cabra produzido por estados (mil litros)
Maranhão	13.356	126	116
Piauí	65.065	1.326	708
Ceará	40.498	980	937
Rio Grande do Norte	10.853	397	1.494
Paraíba	31.068	2.677	5.627
Pernambuco	58.212	2.027	3.417
Alagoas	4.038	397	464
Sergipe	1.913	180	266
Bahia	71.390	4.991	4.665

Fonte: IBGE – Censo Agropecuário (2017)

Pode-se observar que todos os estados da região apresentam uma considerável quantidade de estabelecimentos voltados para caprinocultura leiteira, com destaque para o estado da Paraíba, Bahia, Pernambuco e Rio Grande do Norte como maiores produtores.

Dando ênfase ao estado da Paraíba como maior produtor do país, onde a caprinocultura sempre representou uma atividade potencial, especialmente para a região do Cariri, apresenta pontos de estrangulamento em todo o seu arranjo produtivo, principalmente em relação à ausência de modelos de sistemas de produção (SOUZA, *et al.*, 2016). Assim, pode-se dizer que a região apresenta fatores propícios para criação de cabras, a lacuna está justamente na falta de tecnologia e no controle da produção, pouco realizado.

2.4 CARACTERÍSTICAS FÍSICO-QUÍMICAS DO LEITE DE CABRA

O leite de cabra é o produto oriundo da ordenha completa, ininterrupta, em condições de higiene, de animais da espécie caprina sadios, bem alimentados e descansados (BRASIL, 2000). A qualidade do leite pode ser definida em termos de sua integridade, ou seja, sem adição de substâncias e/ou remoção de componentes, de sua composição química, características físicas, livre de deterioração microbiológica e presença de patógenos (DÜRR, 2004).

Dessa forma, correta caracterização físico-química do leite de cabra é necessária para assegurar que o produto apresente os padrões mínimos de composição, que não foi adulterado, e que não contém contaminantes (RICHARDS *et al.*, 2001), sendo assim, boa parte do controle de qualidade é baseada nestas características e em alterações destas (BELOTI e TAMANINI, 2015).

Do ponto de vista físico-químico, o leite é uma mistura homogênea de grande número de substâncias (lactose, glicérides, proteínas, sais, vitaminas, enzimas), das quais algumas estão em emulsão (gordura, vitaminas lipossolúveis), algumas em suspensão (caseínas ligadas a sais minerais) e outras em dissolução (lactose, vitaminas hidrossolúveis, proteínas do soro, sais e outros) (ORDÓNEZ, 2005).

Entre as análises físico-químicas do leite pode-se citar: acidez, densidade, gordura, sólidos totais, sólidos não gordurosos, crioscopia, lactose, proteína e caseína. Sendo a análise do ponto de congelamento (índice crioscópico) e da acidez do as mais utilizadas para a averiguação de adição de água bem como da conservação, respectivamente.

Nesse sentido, de acordo com regulamento técnico de identidade e qualidade de leite de cabra aprovado pela Instrução Normativa nº 37 do Ministério da Agricultura, Pecuária e Abastecimento (MAPA) fixa as condições de produção, a identidade e os requisitos mínimos de qualidade do leite de cabra destinado ao consumo humano (BRASIL, 2000). A tabela 2 apresenta os requisitos e os valores padrão das características físico químicas do leite de cabra de acordo com o exigido na Instrução Normativa nº 37 para leite caprino.

Tabela 2 - Requisitos mínimos de qualidade físico-química do leite de cabra pasteurizado segundo a IN 37 do MAPA

Requisitos	Leite Integral	Leite Semi-Desnatado	Leite Desnatado
Gordura (%)	Teor original	0,6 a 2,9	máx 0,5
Acidez (% ácido láctico)	0,13 a 0,18 para todas as variações		
Sólidos Não Gordurosos (%)	mínimo 8,20 para todas as variedades		
Densidade (g/mL 15°C)	1,028 a 1,034 para todas as variedades		
Índice Crioscópico (°H)	-0,550 a -0,585 para todas as variedades		
Proteína Total (Nx6,28) (%)	mínimo 2,8 para todas as variedades		
Lactose (%)	mínimo 4,3 para todas as variedades		
Cinzas (%)	mínimo 0,70 para todas as variedades		

Fonte: Brasil (2000)

É válido salientar que o leite cru em relação a esses requisitos presentes no quadro anterior difere do pasteurizado só em relação ao teor de gordura, que neste caso será de igual ou superior a 3%.

2.4.1 Fatores que Influenciam a Qualidade do Leite

O leite de boa qualidade é definido como sendo altamente nutritivo, com sabor agradável, ausência de microrganismos patogênicos e contaminantes como antibióticos, pesticidas, adição de água e sujidades, mínima contagem de células somáticas, sanidade do rebanho, limpeza dos equipamentos e utensílios, a higienização do local da obtenção do leite, e também, a qualidade da água que é utilizada na propriedade, podendo transmitir doenças tanto para humanos quanto para animais, devem ser observadas a fim de evitar estes contratemplos (VENTURINI *et al.*, 2007).

Segundo Dürr (2005), a produção de leite com qualidade reduz a ocorrência de não conformidades, bem como, a presença de doenças, resultando em menores custos e uma maior produção. Para isso, é necessário estar atento a alterações na composição físico-química do leite que podem ser ocasionadas devido a fatores como origem do animal, raça, alimentação, idade, higiene da ordenha e dos utensílios, manejo, clima, estágio de lactação, genética dos rebanhos,

armazenagem e transporte do leite. Por esta razão, existem padrões que são estabelecidos para detecção de erros que interferem na qualidade do produto final (AMORIM, 2017).

Destaca-se que uma das principais causas que desempenham influência negativa sobre a qualidade e produção do leite é a infecção da glândula mamária, conhecida como mastite (COSTA *et al.*, 2017). Além da mastite, existem outras doenças que acometem o rebanho leiteiro; porém com menos incidência, como a brucelose, tuberculose, salmonelose (também conhecida como paratifo dos bezerros), raiva, diarreia viral bovina (BVD), cisticercose bovina, manqueira entre outros (DUNCAN e PRASSE, 1982).

De acordo com Agnese e colaboradores (2002), a qualidade do leite está ligada à disseminação de doenças tanto para o homem quanto para os animais. Assim, faz-se necessário a avaliação das características físico-químicas do produto, tendo em vista verificar o estado de conservação até uma possível ocorrência de fraudes.

2.5 O EMPREGO DA TECNOLOGIA NO MERCADO AGROPECUÁRIO

Atualmente, organizações, públicas ou privadas, estão cada vez mais dependentes das tecnologias da informação (TI) para alcançar os seus objetivos estratégicos e operacionais. Em face a esta realidade, tem-se notado uma evolução do papel da TI nas organizações, em sincronia com a evolução das próprias organizações (BROWN *et. al.*, 2012).

Observa-se que o emprego das TI em apoio aos processos organizacionais gera uma quantidade substancial de informação, que deve ser utilizada a fim de aumentar o conhecimento organizacional. Assim, a correta gestão de informações e conhecimentos que potencializam a importância do capital intelectual, é fundamental no desenvolvimento organizacional.

2.5.1 Gestão da informação

Quando se fala em informação nas organizações é natural que também se mencione os Sistemas de Informação (SI) que permitem coletar, armazenar, recuperar e disseminar informações para fins específicos (TURBAN *et al.*, 2007). Estes apoiam funções operacionais, gerenciais e de tomada de decisão (RAINER JR. e CEGIELSKI, 2011).

Em termos básicos, um sistema de informação em uma organização é compreendido como todos os registros e documentos gerados nas operações desenvolvidas por ela (CASSARRO, 2010), podendo ser manual ou informatizado. Assim, como fator de ordenação de um sistema, pode-se afirmar que a informação constitui-se como um elemento fundamental e indispensável para a organização.

Spinola (2013), destaca que o controle da informação ainda é um recurso inacessível em muitas organizações que não desenvolveram habilidades para armazenar e processar os dados gerados, tendo como consequência inúmeras ocorrências de falhas e/ou falências motivadas pelo mau uso das informações, ineficaz gestão e administração dos recursos informacionais.

Nesse sentido, a aplicação de técnicas de mineração e exploração de dados gera conexões e possibilita a construção de modelos capazes de apoiar o processo decisório por meio do processamento e geração de informações.

2.5.2 Mineração de dados

Segundo Nhacuonge (2015), Mineração de dados (MD) - do inglês *Data Mining* - (DM), originou-se como uma resposta face ao desenvolvimento de tecnologias da informação. Esta consiste na identificação de padrões consistentes, tendências e relacionamentos sistemáticos entre variáveis através da exploração de grandes quantidades de dados armazenados em repositórios. Desse modo, a descoberta de informações significativas ocorre por meio do uso de reconhecimento de padrões, estatística e técnicas matemáticas (MALHOTRA E BIRKS, 2007).

De acordo com Goldschmidt, Passos e Bezerra (2015), esse processo é responsável por gerar um modelo de conhecimento que é visualizado, analisado e interpretado no Pós-Processamento. Gopalkrishnan e colaboradores (2012), enfatizam que a correta análise dos dados proporciona vantagens sustentáveis em relação à concorrência, impulsionando as corporações a darem valor aos seus dados.

Para Cardoso e Machado (2008), a mineração de dados é capaz de revelar, automaticamente, o conhecimento que está implícito em grandes quantidades de informações armazenadas, comumente desestruturadas, nos bancos de dados de uma organização. Essa técnica pode fazer, entre outras, uma análise antecipada dos eventos, possibilitando prever tendências e comportamentos futuros, permitindo aos gestores a tomada de decisões baseada em fatos e não em suposições (TWO CROWNS CORPORATIONS, 2005).

Neste contexto, destacam-se os sistemas de informação, que utilizam os dados existentes nas organizações para disponibilizar informações relevantes para a tomada de decisão (SANTOS e RAMOS, 2017). Logo, o desenvolvimento de tecnologias capazes de realizar análises eficientes desses dados torna-se necessária, tendo em vista que os mesmos possuem alto valor agregado.

2.5.3 Business Intelligence

O termo *Business Intelligence* (BI) foi proposto por Howard Dresner, do Grupo Gartner, no ano de 1996, quando foi definido como o uso de tecnologias e aplicações que consistem em *data warehouse* ou *data mart*, relatórios de consulta, análise de dados, mineração de dados e *backup* de dados (LU, 2014). Atualmente o Grupo Gartner conceitua BI como sendo um termo abrangente que inclui as aplicações, as infraestruturas, ferramentas e as melhores práticas que permitem o acesso e análise de informação para melhorar e otimizar decisões e o desempenho (GARTNER, 2019).

Os sistemas de BI têm como propósito, com base em dados, apresentar informação aos decisores, que estes usam e transformam em conhecimento (SHERMAN, 2015) para efeitos de apoio à tomada de decisão. O BI pode, assim, ser descrito como:

[...] um conjunto de técnicas e ferramentas para a aquisição e transformação de dados brutos em informações significativas e úteis para fins de análise de negócios. [...]” Assim, o BI pode ser utilizado para apoiar uma vasta gama de decisões de negócio, como decisões operacionais e estratégicas (BENTLEY, 2017).

Empresas, sejam elas de pequeno, médio ou grande porte, em algum momento geram registros, dados sobre funcionários, momentos econômicos, registros de produção mensal, porém esta crescente quantidade e complexidade dos dados não trazem valor por si só, é necessário o uso ou desenvolvimento de ferramentas que possibilitem a análise e obtenção de informações de forma ágil (FERRARI e RUSSO, 2016).

A implementação de um serviço de BI não traz só vantagens apenas em questões de armazenamento e pesquisa rápida de grandes conjuntos de dados, mas também na procura inteligente de informações que permite a flexibilização de consultas e análises de informações consoante a necessidade do utilizador (CASTRO, 2016).

Posto isto, é possível maximizar a utilidade de informação recolhida e o suporte aos processos de tomada de decisão nos negócios, através da adoção de um serviço de BI obtendo uma visão integrada do negócio que disponibiliza informação relevante para o decisor de forma rápida (MALLADI, 2013).

2.5.4 Tomada de Decisão Orientada por Dados

Tradicionalmente, organizações conduziram a maior parte de suas decisões por meio de regras simples, baseadas na experiência ou prática (CALLINGHAM, 2004). Contudo, o uso de dados tem mudando completamente o processo de tomada de decisão nas empresas (CHAN; PEYNE, 2017), e diversos estudos têm apresentado a importância da decisão baseada em dados para o bom desempenho das organizações (BRYNJOLFSSON; HITT; KIM, 2011).

Gerenciar processos sem ter indicadores que descrevam sua viabilidade de forma clara e precisa geralmente tem consequências indesejáveis no desempenho não apenas do processo em questão, mas também na escala de toda a organização (FORTULAN; GONÇALVES FILHO, 2005).

Além disso, os autores supracitados afirmam que para fazer frente ao aumento da concorrência, regulamentação e globalização, o acesso às informações confiáveis de forma intuitiva e rápida é necessário em todos os níveis da empresa, seja para decisão operacional, tática ou estratégica (FORTULAN; GONÇALVES FILHO, 2005).

Ademais, destaca-se que o uso de dados na tomada de decisão leva as empresas a uma nova realidade, tornando possíveis práticas que até mesmo os funcionários mais experientes não poderiam adotar sozinhos. Assim, as decisões obtidas têm maior potencial de imparcialidade e objetividade (LONG, 2018), bem como, conferem uma melhor performance e eficiência. Logo, o paradigma de tomada de decisão orientado a dados torna-se predominante (XU *et al.*, 2014).

Entendendo-se que boas escolhas dependem de informações precisas e atuais, recursos de coleta e processamento de dados ganharam avanços significativos em campos como análise de *Big Data*, *Business Intelligence* e Mineração de dados.

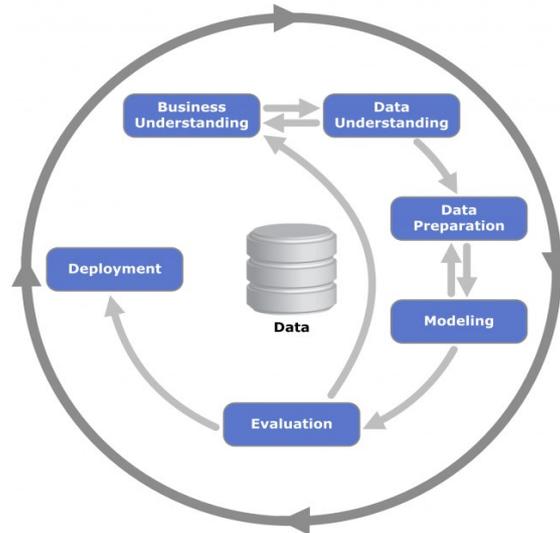
3 PROCEDIMENTOS METODOLÓGICOS

O método adotado para realização do estudo foi o Crisp-DM (*Cross Industry Standard Process for Data Mining*), o qual se apresenta como um dos mais utilizados na área de Mineração de Dados devido fornecer uma abordagem estruturada para o planejamento de projetos.

Segundo Chapman (2000), a metodologia é composta por seis fases, organizadas de maneira cíclica, cujo fluxo é unidirecional, possibilitando ir e voltar entre as suas fases e tarefas. São estas: Entendimento do Negócio (*Business Understanding*), Entendimento dos Dados (*Data Understanding*), Preparação dos Dados (*Data Preparation*), Modelagem (*Modeling*),

Avaliação (*Evaluation*) e Aplicação (*Deployment*). A Figura 1, demonstra o ciclo de vida original do modelo CRISP-DM.

Figura 1 - Etapas do método *CRISP-DM*



Fonte: IBM Knowledge Center (2012)

3.1 ENTENDIMENTO DO NEGÓCIO (*BUSINESS UNDERSTANDING*)

Na primeira fase, ocorreu uma imersão na organização para melhor e maior entendimento do negócio. Inicialmente foi necessário entender todo o processo de manufatura com foco no controle de qualidade, visto que nessa etapa ocorrem as análises dos aspectos visuais e físico-químicos, as quais detectam se o leite está apto para consumo. Identificou-se, também, questões relacionadas à confiabilidade, segurança, disponibilidade, manutenção, governança, ferramentas e tecnologias envolvidas. Ademais, ocorreu a definição da base de dados a ser trabalhada.

3.2 ENTENDIMENTO DOS DADOS (*DATA UNDERSTANDING*)

A segunda fase compreendeu a coleta inicial dos dados, prosseguindo de atividades para familiarização. Os dados foram coletados a partir de planilhas físicas (em papel), cedidas pela organização. Estas englobaram cerca de 3 mil folhas referentes aos anos de 2019 e 2020, as quais foram posteriormente digitadas em planilhas eletrônicas do *Google Sheets*.

3.3 PREPARAÇÃO DOS DADOS (*DATA PREPARATION*)

Com o entendimento dos dados, foi possível organizá-los e realizar o processo de tratamento. Esse processo buscou reduzir ou eliminar anomalias e/ou discrepâncias que poderiam comprometer a análise como um todo. Dessa forma, dados considerados anômalos (valores nulos, inconsistentes, duplicados) foram removidos ou tratados. Além disso, respostas em branco foram removidas, ou seja, linhas que apresentavam células em branco e/ou informações incompletas.

É importante destacar que os dados foram organizados formando três *datasets* diferentes no *Google Sheets*, isto é, o primeiro com os dados referentes ao município de Sumé, o segundo com os dados pertencentes aos demais municípios (Coxixola, Livramento, Parari, São José dos Cordeiros, São João do Cariri, Serra Branca, Camalaú, Boa Vista) e o terceiro referente aos dados das amostras de leite pasteurizado, este engloba as amostras dos 9 municípios associados.

Salienta-se que o *dataset* do município de Sumé não foi unificado à dos municípios associados devido possuir dados referentes a cada agricultor associado, enquanto o *dataset* dos municípios possui valores médios, isto é, referente ao leite fornecido por todos os agricultores associados a cada município.

3.4 MODELAGEM (*MODELING*)

Após o processo de tratamento dos dados, deu-se início a fase de modelagem. Essa etapa foi subdividida em duas, sendo a primeira (3.4.1), referente à análise exploratória dos dados e a segunda (3.4.2) a aplicação da técnica de agrupamento para construção de modelos.

3.4.1 Exploração de dados

A análise exploratória dos dados buscou direcionar posteriores análises, e portanto, foram abordados questionamentos referentes à natureza dos dados, tendo em vista obter uma visão ampla e direcionada acerca da base de dados. Neste sentido, a caracterização e análise descritiva dos dados foi elucidada pelas seguintes questões:

1. Existe algum mês do ano em que ocorre maior ou menor volume de leite entregue à organização?

2. Existe algum município que entrega maior ou menor volume de leite à organização?
3. Algum dos anos, 2019 ou 2020, apresentou maior ou menor volume de leite entregue à organização? Se sim, tal fato pode ter se dado em decorrência de que?
4. Qual não conformidade se mostrou mais presente nas amostras de leite?

3.4.2 Agrupamento com o algoritmo *K-means*

Após a exploração dos dados e um maior entendimento acerca do processo produtivo da organização, foi aplicada a técnica de agrupamento aos *datasets*, tendo em vista identificar informações relevantes a partir da geração de grupos formados por registros similares, ou seja, que compartilham de propriedades comuns.

O algoritmo adotado para a análise de agrupamento foi o *K-means*, este avaliou e agrupou os dados de acordo com suas características. E a ferramenta utilizada foi o Orange Data Mining, este é um software de código aberto para aprendizado de máquina e mineração de dados. Por meio de sua interface gráfica denominada *Orange Canvas* é possível, criar todo fluxo de trabalho de um projeto de *Data Mining*, sem necessidade de código, de forma simples, visto que este apresenta um *design* visual e inovador. Além do suporte às tarefas que vão desde o pré-processamento de dados até modelagem e avaliação.

3.5 AVALIAÇÃO (*EVALUATION*) E IMPLEMENTAÇÃO (*DEPLOYMENT*)

A avaliação, de forma detalhada, das estatísticas e dos modelos de mineração construídos, ocorreu através da construção de um *dashboard* para o setor de qualidade da organização. Dessa forma, para o desenvolvimento do *dashboard* inicialmente foi realizado um *brainstorming*, tendo em vista verificar as necessidades e desejos dos usuários. Em seguida foi desenvolvido o diagrama de *Use Case* do sistema proposto, visando a definição e melhor compreensão do funcionamento do sistema. Ademais, foi construída uma planilha personalizada através do *Google Sheets*, a qual teve como objetivo funcionar como banco de dados, armazenando e alimentando o *dashboard*, a representação desta pode ser visualizada em apêndice 3. Por fim, o *dashboard* foi apresentado à gestão e posteriormente implementado aos computadores da organização.

4 RESULTADO E DISCUSSÃO

4.1 CARACTERIZAÇÃO DO NEGÓCIO

O presente estudo foi realizado na Associação Gestora da Usina de Beneficiamento de Lácteos – AGUBEL. Está encontra-se localizada na Fazenda Agreste, zona rural de Sumé –PB e atualmente presta serviço para o Programa de Aquisição de Alimentos (PAA). Tal programa é resultante de um convênio entre o Governo Federal e Estadual, e tem como objetivo incentivar a produção e o consumo de leite, sobretudo, o fortalecimento da cadeia produtiva por meio da geração de renda do agricultor, bem como, o abastecimento familiar a partir da distribuição gratuita do leite para as famílias em situação de vulnerabilidade social.

Atualmente a associação tem beneficiado leite de cerca de 250 agricultores familiares. Seu quadro social composto por 23 associados, além da parceria com 09 (nove) associações de produtores dos municípios de: Coxixola, Livramento, Parari, São José dos Cordeiros, São João do Cariri, Serra Branca, Camalaú, Boa Vista e Sumé.

Em cada município, integrante da associação, existe um local de coleta do leite. Dessa forma, diariamente o responsável pelo ponto de coleta recebe o leite dos produtores associados, analisa a temperatura, a acidez e a densidade. Se o leite estiver dentro dos parâmetros legais, ocorre a condução para o tanque de resfriamento, caso não esteja de acordo com os parâmetros definidos é comunicado ao produtor para que, se for do interesse deste, ele retorne ao local e recolha. Não havendo interesse é realizado o descarte.

Esses pontos de coleta não detêm os mecanismos necessários para realização de todas as análises físico-químicas necessárias, isto é, que verifiquem a qualidade e integridade do leite. Sendo assim, duas vezes na semana ocorre o recolhimento do leite que é transportado para a usina AGUBEL onde ocorre a análise físico-química, a qual compreende a determinação da temperatura, acidez, proteína, densidade, gordura mínima, sólidos não gordurosos e da crioscopia. Em seguida o leite segue para as próximas etapas do processo, conforme detalhado no Anexo 2.

A organização busca utilizar o máximo possível de estratégias que melhorem a gestão, estabelecendo metas e objetivos a serem alcançados. Esta conta apenas com um sistema de informação para emissão de nota fiscal, e destacou a importância e necessidade de um sistema informatizado, para o setor de qualidade, tendo em vista agregar valor ao gerenciamento e funcionamento da organização.

4.1.1 Descrição dos Dados

Os dados contidos nas planilhas físicas, referentes aos anos de 2019 e 2020, foram digitados em planilhas eletrônicas, do *Google Sheets*, integrando 39.715 linhas, com informações acerca dos parâmetros físico-químicos, bem como, data de entrega, quantidade, presença de não conformidades e ações corretivas.

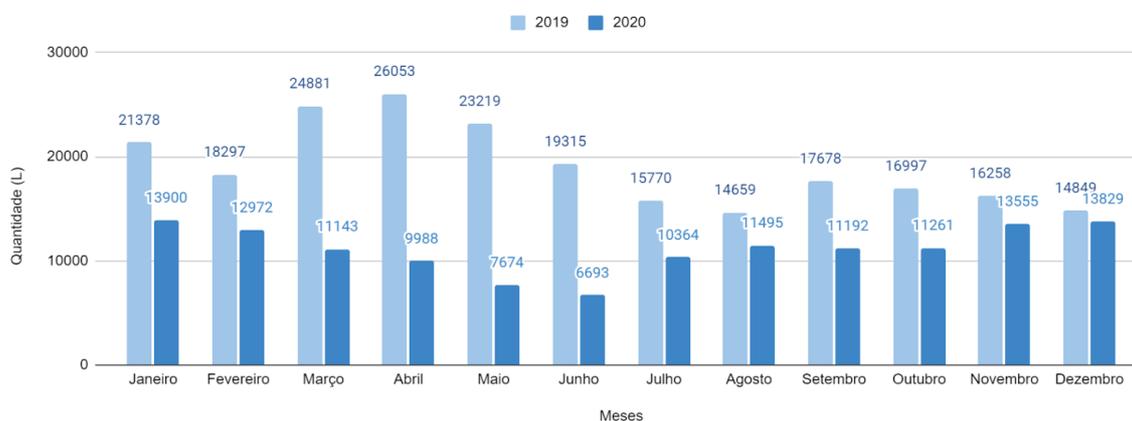
4.2 ANÁLISE EXPLORATÓRIA E MODELAGEM DOS DADOS

Esse subtópico apresenta a análise exploratória de dados realizada a partir das variáveis previamente selecionadas, discutida no subtópico 4.2.1, assim como os resultados da aplicação de técnicas de Mineração de Dados e construção de modelos utilizando algoritmos não supervisionados, apresentado no subtópico 4.2.2.

4.2.1 Análise exploratória de dados

A comparação do quantitativo de leite fornecidos mensalmente nos anos de 2019 e 2020 mostrou a atipicidade do ano 2020 com destaque para o contraste entre meses de março a junho, conforme apresenta a Gráfico 1. Tal cenário relaciona-se com o enfrentamento da crise sanitária da Covid 19, visto que as medidas de contingenciamento adotadas pelo país iniciaram por volta da segunda quinzena de março de 2020 e conforme demonstrado no gráfico a partir desse mesmo mês ocorreu uma baixa no fornecimento de leite.

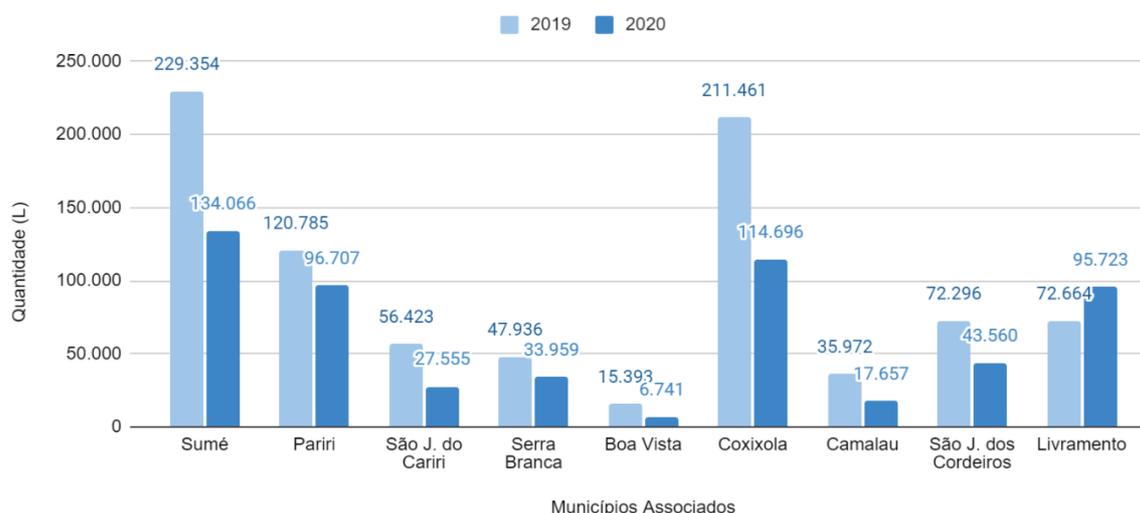
Gráfico 1 - Comparativo da distribuição quantitativa de leite fornecido mensal nos anos de 2019 (azul claro) e 2020 (azul escuro).



Fonte: Autoria própria (2023)

Adicionalmente, foi analisado o comparativo da quantidade de leite fornecida mensalmente entre os anos de 2019 e 2020, por cada município integrante da associação (Gráfico 2). Destaca-se, dentre os nove municípios, uma média inferior de fornecimento de leite no ano de 2020 em comparação com o ano anterior, com exceção do município de Livramento.

Gráfico 2 - Comparativo da distribuição quantitativa (litros) de leite fornecido pelos municípios associados entre os anos de 2019 (azul claro) e 2020 (azul escuro).

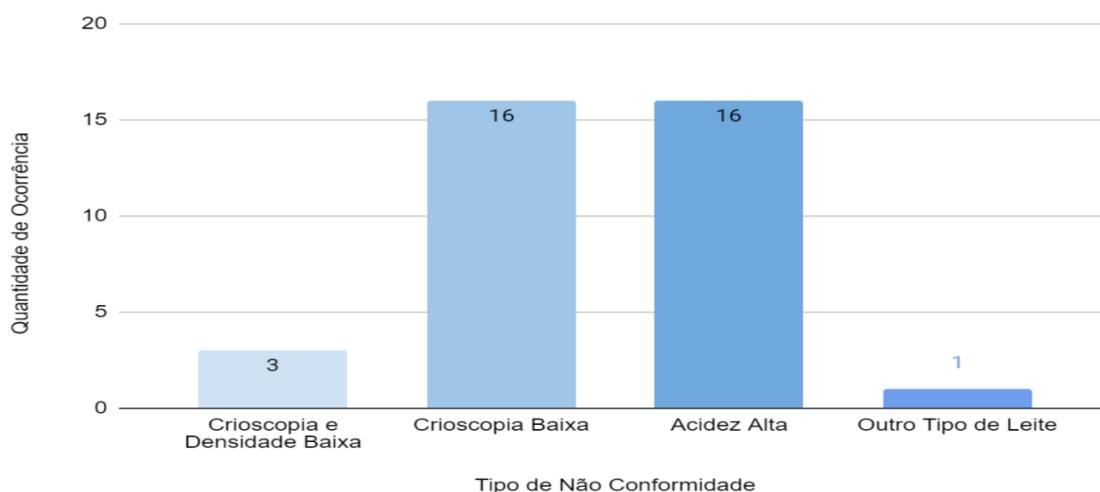


Fonte: Autoria Própria (2023)

Foram triadas, também, as não conformidades presentes no leite fornecido pelo município de Sumé, a triagem de não conformidade para os demais municípios não foi realizada, visto que a organização não possui os referidos dados.

Dessa forma, os resultados, exibidos no Gráfico 3, apresentam que 44,4% das ocorrências foram por crioscopia baixa, 44,4% por acidez alta, 8,3% por crioscopia baixa e densidade baixa e apenas 2,8% por presença de adição de leite de vaca. Tais eventos comprometeram 998 litros de leite, sendo 40 litros referente a crioscopia baixa e densidade baixa, 218 litros correspondente a crioscopia baixa, 190 litros com presença de leite de vaca e 550 litros referentes a acidez alta. Dessa forma, nota-se que houve maior desperdício de leite em decorrência de altas taxas de acidez. É interessante salientar que o ano que mais apresentou não conformidades de leite foi o de 2019 com um percentual de 71,5%, em comparação com o ano de 2020 que retratou 28,5% das ocorrências registradas.

Gráfico 3 -Triagem da ocorrência de não conformidade no leite fornecido por Sumé à associação nos anos de 2019 e 2020.



Fonte: Autoria Própria (2023)

Apesar da baixa ocorrência de não conformidades, isto é, 36 eventos durante o período de dois anos, a quantidade de litros de leite comprometida (998 litros) é compreendida como alta.

4.2.2 Análise de Agrupamento com *K-means*

Na análise de agrupamento, referente ao município de Sumé, foram formados dois grupos, representados na Tabela 3. No entanto, os grupos foram formados apenas com três parâmetros (densidade, acidez e proteína) tendo em vista atender uma medida de qualidade efetiva para a estrutura, conforme apresentado no Anexo 1. O *dataset* compreendeu 10.784 amostras, sendo 8.247 pertencentes ao grupo 1 e 2.537 referentes ao grupo 2. Ressalta-se que o agrupamento do município de Sumé foi realizado de forma individual, visto que a base de dados fornecida pela organização compreendia informações detalhadas de cada agricultor associado, enquanto o *dataset* dos demais municípios possuía quantitativos gerais e valores médios para os parâmetros.

Tabela 3 - Agrupamento com *dataset* pertencente a Sumé

Sumé		C1	C2	Geral
Densidade	Média	1028,66	1028,41	1028,60
	Desvio Padrão	1,38	1,51	1,42
Acidez	Média	16,15	14,78	15,83
	Desvio Padrão	0,37	0,56	0,72
Proteína	Média	3,08	3,11	3,09
	Desvio Padrão	0,11	0,41	0,22

Fonte: Autoria Própria (2023)

Nota-se a partir das médias que não há diferença expressiva entre os grupos do parâmetro densidade, apresentando-se dentro dos valores estabelecidos pela legislação. Contudo, o desvio do grupo 2 foi mais expressivo em comparação com o grupo 1 e o desvio geral dos dados (Tabela 3). Tal resultado indica uma maior heterogeneidade dos dados.

O *cluster* 1 detentor de 76% da população do *dataset*, apresentou um índice de acidez superior ao *cluster* 2 e a média geral. Tal resultado, pode estar relacionado, principalmente, com a quantidade de ácido lático presente no leite devido à presença de bactérias que degradam a lactose até este componente (ALVES, 2018). É válido salientar que 44,4% das não conformidades do leite e 55,1% da quantidade de litros perdidos ocorreram em função de altas taxas de acidez presentes no leite do município de Sumé. Neste sentido, se faz necessário um acompanhamento mais próximo e cauteloso acerca do processo como um todo, tendo em vista minimizar e/ou erradicar as chances de situações de não conformidade do leite.

Quanto ao parâmetro proteína foi identificado um valor médio aproximado entre os grupos. O teor médio proteico das amostras variou entre 3,08 e 3,11%, portanto, acima do preconizado na legislação brasileira (acima de 2,8%) e superior aos resultados verificados por Silva e colaboradores (2011) e Pinheiro e colaboradores (2014), que foram de 2,90 e 2,85%, respectivamente.

Ademais, é evidenciado, a partir do diagnóstico promovido pelos resultados, que o grupo 2 apresenta valores mais heterogêneos e dispersos em todos os parâmetros indicando menor qualidade em comparação ao grupo 1. Entretanto, ambos os grupos apresentaram valores médios dentro dos valores exigidos pela legislação IN 37 (BRASIL, 2000).

A Tabela 4 apresenta o agrupamento realizado com o *dataset* pertencente aos demais municípios associados, o qual foi formado com cinco parâmetros, sendo estes, temperatura, densidade, acidez, gordura mínima (MG) e proteína.

Tabela 4 - Agrupamento com *dataset* referente aos municípios associados

Municípios		C1	C2	Geral
Temperatura	Média	4,41	29,26	6,07
	Desvio Padrão	0,72	2,77	6,28
Densidade	Média	1028,54	1029,53	1028,61
	Desvio Padrão	0,59	6,11	1,67
Acidez	Média	16,01	16,07	16,01
	Desvio Padrão	0,66	0,36	0,64
MG	Média	3,81	3,99	3,82
	Desvio Padrão	0,36	0,59	0,38
Proteína	Média	3,09	3,16	3,10
	Desvio Padrão	0,11	0,16	0,12

Fonte: Autoria Própria (2023)

O parâmetro temperatura apresentou valor médio bastante superior no grupo 2 em comparação ao grupo 1 e a média geral das amostras. A média geral apontou uma alarmante dispersão, evidenciando a alta heterogeneidade entre os dados. Tal fato está relacionado à gestão, armazenamento e logística no transporte do leite entre os municípios e a usina situada em Sumé. Já as amostras de leite que chegam aos pontos de recebimento nos dias de coleta, próximo ao horário de saída do veículo transportador, só são refrigeradas durante o percurso.

O grupo das bactérias mesófilas inclui a maioria dos contaminantes do leite e a quase totalidade das bactérias patogênicas. Estas se multiplicam em temperaturas de 20° a 40°C, sendo a temperatura ideal em torno de 30°C. Além de atingirem altas contagens, quando o leite é mantido em temperatura ambiente (BRASIL, 2017). Desta forma, salienta-se a necessidade de uma atenção especial ao grupo 2, tendo em vista a alta probabilidade de contaminação do leite, através do desenvolvimento de bactérias.

De acordo com a instrução normativa nº 37, o leite caprino deve apresentar densidade entre 1,028 a 1,034 g/cm³ (BRASIL, 2000). Cruz (1998), em pesquisa com leite de cabra, relata valores médios de densidade variando entre 1030,8 a 1032,2 g/cm³, enquanto nesta análise foram encontrados valores médios que variaram entre 1028,5, para o grupo 1, a 1029,5 g/cm³ para o grupo 2.

Quanto ao parâmetro acidez os grupos apresentaram valores médios aproximados, sendo 16,01 °D para o C1 (grupo 1) e de 16,07 °D para o C2 (grupo 2), enquadrando-se dentro dos valores exigidos pela legislação. Entretanto, evidenciou-se uma dispersão significativamente superior no grupo 1 em comparação ao desvio do C2 e do geral, bem como, uma maior heterogeneidade dos dados.

Os grupos não apresentaram diferença expressiva com relação ao valor médio referente ao parâmetro gordura. É válido salientar que diversos autores encontraram resultados semelhantes aos valores médios do presente estudo (3,82%). O estudo de Pereira (2016) com cabras das raças Saanen e Parda Alpina apresentou valor médio de 3,5% de gordura, enquanto Fernandes e colaboradores (2008), trabalharam com cabras mestiças Moxotó no estado da Paraíba e verificaram teor de gordura médio de 3,89%. No entanto, a variação para este componente é ampla, sendo reportada na literatura com valores de 2,6 a 5,5% (QUEIROGA *et al.*, 1998; FONSECA *et al.*, 2006; PEREIRA *et al.*, 2005; QUEIROGA *et al.*, 2007; ANDRADE *et al.*, 2008; ALMEIDA *et al.*, 2013).

O parâmetro proteína expressou valores médios aproximados, isto é 3,09 para o grupo 1 e 3,16 para o grupo 2, bem como baixa dispersão. O autor Cruz (1998), verificou em seus estudos um intervalo médio de 2,62 a 3,88 g/100g de proteína em amostras de leite caprino comercializado no estado da Paraíba. Valores similares foram encontrados no estudo de Benedet (1996), em relação ao leite produzido no Cariri paraibano, com variação de 2,6 a 3,8 g/100mL e valor médio de proteína de 3,29%. Em um trabalho mais recente desenvolvido por Queiroga (2016), foi verificado em suas análises, valor médio inferior (2,7 g/100mL) ao obtido no presente na presente análise (3,1 g/100mL) conforme demonstrado na Tabela 4.

De forma geral, quanto à análise de agrupamento realizada com os municípios associados, com exceção de Sumé, pôde-se observar que o grupo 2 apresentou uma maior heterogeneidade dos dados, sendo a temperatura um valor criticamente disforme no grupo 2. Salienta-se que o grupo 2 pertencente a essa análise corresponde a apenas 7% dos dados do referido *dataset*.

A terceira análise realizada com o *dataset* pertencente ao leite pasteurizado, representada na Tabela 5, conta com valor médio e desvio padrão para os quatro grupos

formados. É válido destacar que o processo de pasteurização engloba todo o leite fornecido à organização, isto é, o leite fornecido por todos os municípios associados.

Tabela 5 - Agrupamento com *dataset* pertencente ao leite pasteurizado

Leite Pasteurizado	C1	C2	C3	C4	Geral	
Temperatura	Média	3,75	4,00	3,77	3,00	6,07
	Desvio Padrão	0,43	0,00	0,42	0,00	6,28
Densidade	Média	1029	1028	1028	1028	1028,20
	Desvio Padrão	0,00	0,00	0,00	0,00	0,40
Acidez	Média	16,03	16,01	14,87	16,01	15,94
	Desvio Padrão	0,23	0,12	0,43	0,12	0,34
MG	Média	3,90	3,81	3,75	3,83	3,82
	Desvio Padrão	0,26	0,20	0,18	0,19	0,22
Proteína	Média	3,10	3,05	3,04	3,05	3,06
	Desvio Padrão	0,05	0,04	0,04	0,03	0,05

Fonte: Autoria Própria

A formação dos quatro grupos contou com os parâmetros temperatura, densidade, acidez, gordura mínima (MG) e proteína. Em relação à temperatura, os valores dos grupos 2 e 4 apresentaram desvios nulos, enquanto os grupos 1 e 3 apresentaram valores aproximados. Ademais, todos os grupos apresentaram valores dentro do estabelecido pela legislação.

Torna-se importante salientar que a temperatura para o leite refrigerado após processo de pasteurização, de acordo com a legislação, deve ser igual ou inferior a 4°C (BRASIL, 2000). Neste sentido, nota-se que o grupo 2, correspondente a 58% das amostras do *dataset*, possui uma temperatura média de 4° C, ou seja, o valor máximo aceitável pela legislação. Dessa forma, evidencia-se a relevância em dedicar uma maior atenção quanto a esse parâmetro, especialmente em relação ao grupo 2.

O parâmetro densidade apresentou valor médio próximo para todos os grupos, bem como, valores de dispersão nulos, sendo entendidos como grupos homogêneos. É importante destacar que o peso específico do leite é determinado pela densidade, onde o resultado está fortemente ligado à concentração de elementos solúveis e porcentagem de gordura (SANTOS, 2020). Dessa forma, observou-se que as amostras atendem às normas legislativas.

Quanto à acidez das amostras, estas apresentaram valores médios aproximados para os grupos 1, 2 e 4. Já o grupo 3 expressou um valor médio inferior, entretanto foi evidenciado um desvio padrão superior em comparação aos demais grupos, denotando em uma maior heterogeneidade dos dados.

Os parâmetros gordura mínima (MG) e proteína apresentaram valores médios e desvios aproximados em todos os grupos, apresentando conformidade das amostras. Albuquerque e colaboradores (2009), trabalhando com cabras Saanen e mestiças obteve média de proteína (2,97%) inferior ao encontrado neste estudo (3,06%).

Assim como a gordura, a proteína do leite pode ser influenciada pela raça, estágio de lactação, alimentação, clima, número de partos, época do ano, e estado de saúde do úbere (MENDES *et al.*, 2009).

De forma geral, foi possível observar, que a análise exploratória dos dados evidenciou informações importantes, como: quantitativo de litros de leite fornecidos por período de tempo e por municípios que mais forneceram leite, tipos de não conformidades presentes e quantidade de leite desperdiçado em função da inadequação dos parâmetros.

Por meio da análise de agrupamento, verificou-se os grupos que apresentaram maior ou menor conformidade nas amostras, contribuindo assim, para identificação dos parâmetros carentes de maior atenção. Além disso, comparando as Tabelas 6, 7 e 8 observou-se que o processo de pasteurização atribui maior conformidade ao produto final.

Diante disso, evidenciou-se a relevância da exploração e processamento de dados para identificação, visualização e análise de informações relevantes. Sendo assim, a partir da compreensão por meio dos resultados obtidos até essa etapa da pesquisa, deu-se início ao processo de desenvolvimento do sistema informatizado (*dashboard*) por meio de um *brainstorming* tendo em vista definir as funcionalidades do sistema.

4.3 BUSINESS INTELLIGENCE NA CONSTRUÇÃO DE UM DASHBOARD

4.3.1 Brainstorming das Necessidades e Desejos dos Usuários

Os principais pontos abordados e definidos no *brainstorming*, para serem implementados como funcionalidades do *dashboard*, estão detalhados no Quadro 2. Um total de 5 sugestões foram selecionadas em consenso.

Quadro 2 - Identificação das necessidades e desejos dos usuários por meio da técnica de *brainstorming*

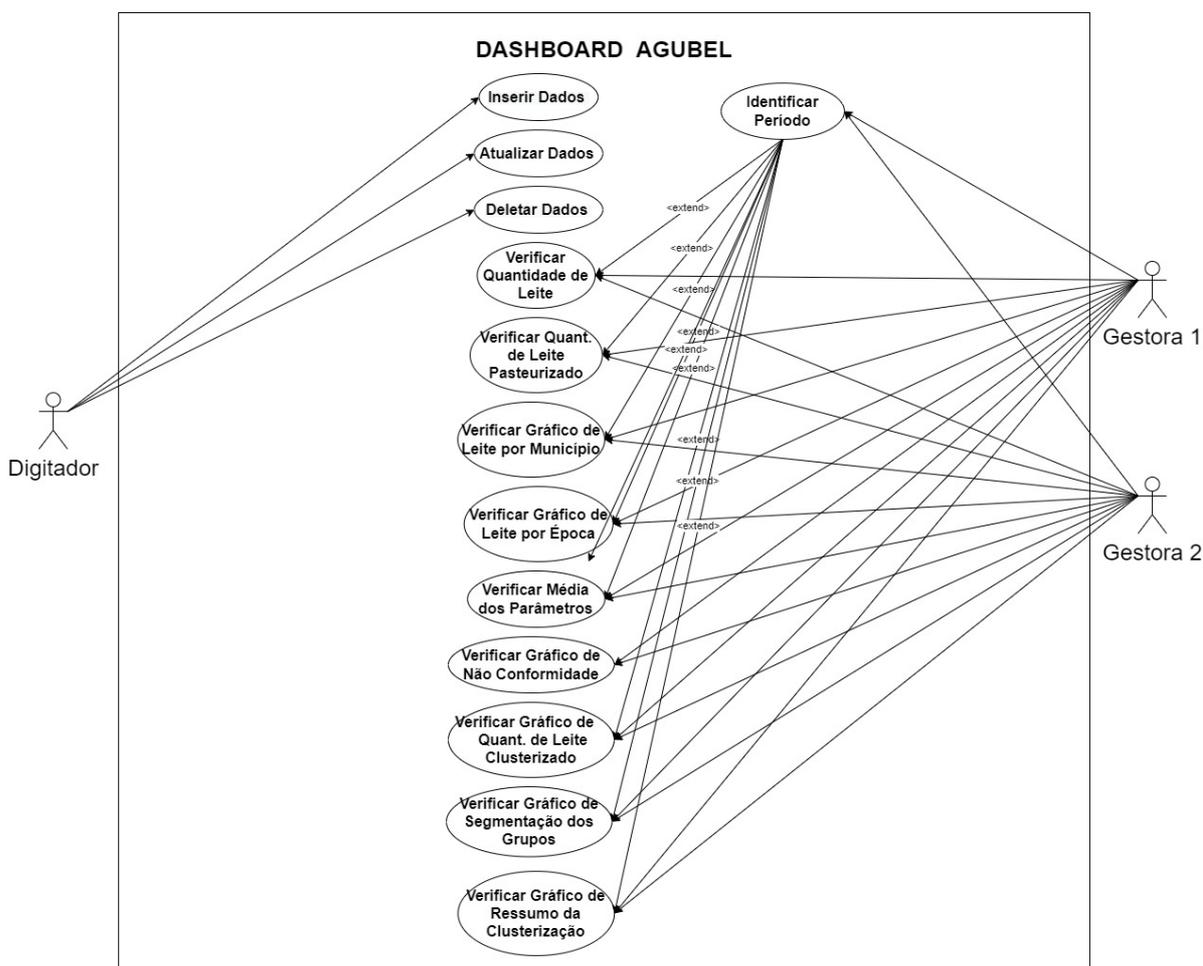
Sugestão da Gestão	Sugestão Pessoal	Discussão
-	Representação da quantidade de leite fornecida (distribuição por municípios; por época)	Sugestão selecionada em consenso, vista como função básica, necessária e de alta relevância.
-	Representação da quantidade e tipo de não conformidade (distribuição por município e época);	Sugestão não selecionada. A gestão citou a baixa frequência desse evento, porém em virtude da alta relevância da informação, não descartou a possibilidade de implementação.
-	Representação do valor médio e desvios padrão dos parâmetros físico-químicos do leite in natura e do leite pasteurizado (distribuição por município e época).	Sugestão selecionada e complementada. A gestão mencionou a importância da incorporação dos parâmetros microbiológicos, nessa representação.
Representação da quantidade total de leite pasteurizado (distribuição por época).	-	Sugestão selecionada em consenso. Definida como relevante no que diz respeito à visão pontual e histórica da informação.
Representação de <i>clusterização</i> (identificação e agrupamento dos produtores, municípios e amostras de leite)	-	Sugestão selecionada em consenso. Definida como de alta relevância, no que diz respeito à identificação dos grupos de indivíduos/amostras com características semelhantes, bem como o mapeamento do perfil.
Representações de gráficos do tipo: barra, coluna, linha ou pizza.	-	Sugestão selecionada em consenso. No caso de representações por meio de gráficos, a gestão relatou ter maior facilidade de interpretação em gráficos do tipo sugerido.

Fonte: Autoria Própria (2023)

4.3.2 Diagrama de Use Case

O diagrama de caso de uso tem como finalidade descrever todas as funcionalidades do sistema do ponto de vista dos usuários. Dessa forma, a Figura 2 apresenta o diagrama de caso de uso do sistema proposto.

Figura 2 - Diagrama de Use Case - Dashboard AGUBEL



Fonte: Autoria Própria (2023)

O sistema proposto conta com as funções de: inserir, atualizar e deletar dados referente ao leite *in natura*, leite pasteurizado e *clusterização*. Ressalta-se que o sistema será destinado ao setor de qualidade da organização. Dessa forma, observa-se que a gestora 1 possui permissões e funcionalidades diferentes da gestora 2 da empresa, conforme apresentado abaixo:

- **Gestora 1:** Inserir, atualizar e deletar dados, além de participar na análise das informações;
- **Gestora 2:** Atualizar e analisar as informações.

5. DESENVOLVIMENTO DO *DASHBOARD* AGUBEL

O *dashboard* constituiu-se de quatro abas, com o intuito de segmentar as informações, contribuindo para um visual claro e objetivo. Nesse sentido, a Figura 3 apresenta o *layout* da tela inicial a qual conta com três botões que direcionam o usuário para os painéis referentes às análises do leite in natura, pasteurizado e clusterizado.

Figura 3 - *Layout* da tela inicial do *dashboard*



Fonte: Autoria Própria (2023)

O painel referente à análise do leite *in natura*, ilustrado na Figura 4, teve como objetivo fornecer informações precisas acerca da quantidade total de leite fornecida à organização, bem como, a distribuição desse quantitativo por município e ao longo do tempo. Ademais, estão presentes informações referentes aos valores médios dos parâmetros físico-químicos, assim como, a respeito dos tipos de não conformidade ocorridas e seus respectivos quantitativos de aparição.

Figura 4 - Ilustração do painel referente à análise do leite *in natura*



Fonte: Autoria Própria (2023)

Para apresentar o valor total da quantidade de leite fornecida à organização em determinado período, optou-se por o uso do visual “Cartão”, visto que este é ideal para informar valores únicos de forma rápida e direta. No painel este encontra-se do lado superior esquerdo. A Figura 5 ilustra a adoção desse visual.

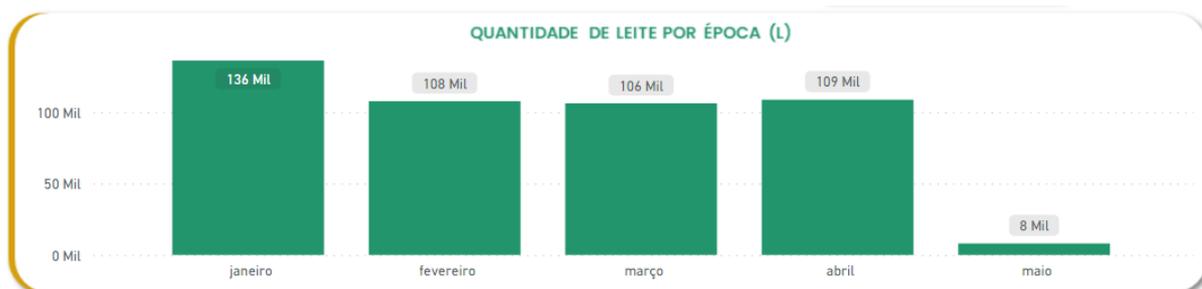
Figura 5 - Representação do quantitativo total de leite



Fonte: Autoria Própria (2022)

Tendo em vista proporcionar a identificação da quantidade de litros de leite fornecidos ao longo do tempo, foi utilizado o “Gráfico de Colunas Empilhadas”. A ferramenta *Power BI* conta com uma funcionalidade que possibilita visualizar tais dados de acordo com as hierarquias de ano, trimestre, mês e dia. Por meio da Figura 6 é possível visualizar o gráfico adotado.

Figura 6 - Visual adotado para apresentar a distribuição do quantitativo de leite por época



Fonte: Autoria Própria (2023)

Adicionalmente, ao lado superior direito encontra-se o visual “Segmentação de Dados”, conforme demonstrado na Figura 7. Através dessa funcionalidade o usuário poderá filtrar os dados de acordo com o período histórico de interesse.

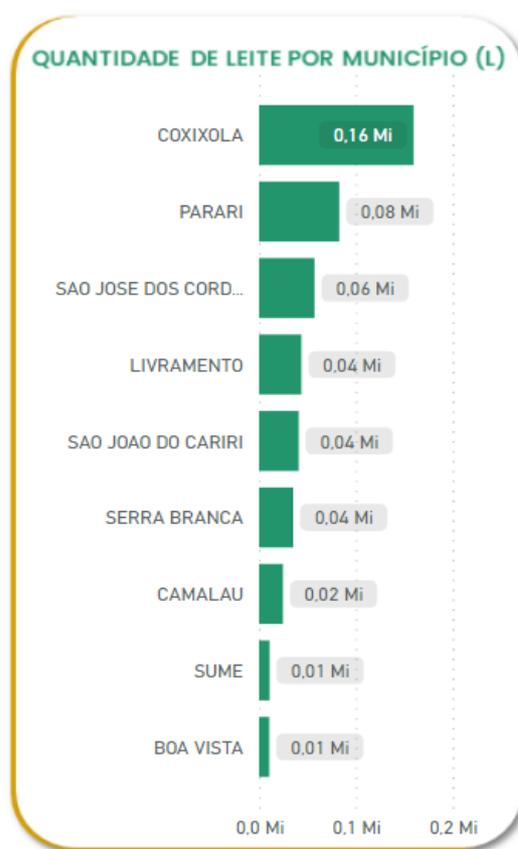
Figura 7 - Ilustração do filtro de data



Fonte: Autoria Própria (2023)

Considerando a importância de fornecer informações da quantidade de leite fornecida por cada município, foi adotado o “Gráfico de Barras Empilhadas”. Este apresenta as informações de forma decrescente, isto é, do município que forneceu uma maior quantidade de litros de leite para o que forneceu uma menor quantidade em determinado período de tempo. Na Figura 8 é possível verificar as informações presentes nesse visual.

Figura 8 - Representação do quantitativo de litros de leite por município



Fonte: Autoria Própria (2023)

Quanto à identificação dos valores médios dos parâmetros físico-químicos, estes foram representados por meio do visual “Cartão de Linhas Múltiplas” conforme ilustrado na Figura 9.

Figura 9 - *Layout* do visual adotado para representação da média dos parâmetros físico-químicos



Fonte: Aatoria Própria (2023)

Por fim, foi utilizado o visual “*Treemap*” destacado na Figura 10, tendo em vista informar ao usuário os tipos de não conformidades ocorridas, assim como as respectivas quantidades de ocorrência. É válido ressaltar que tais informações só são registradas a respeito do leite in natura.

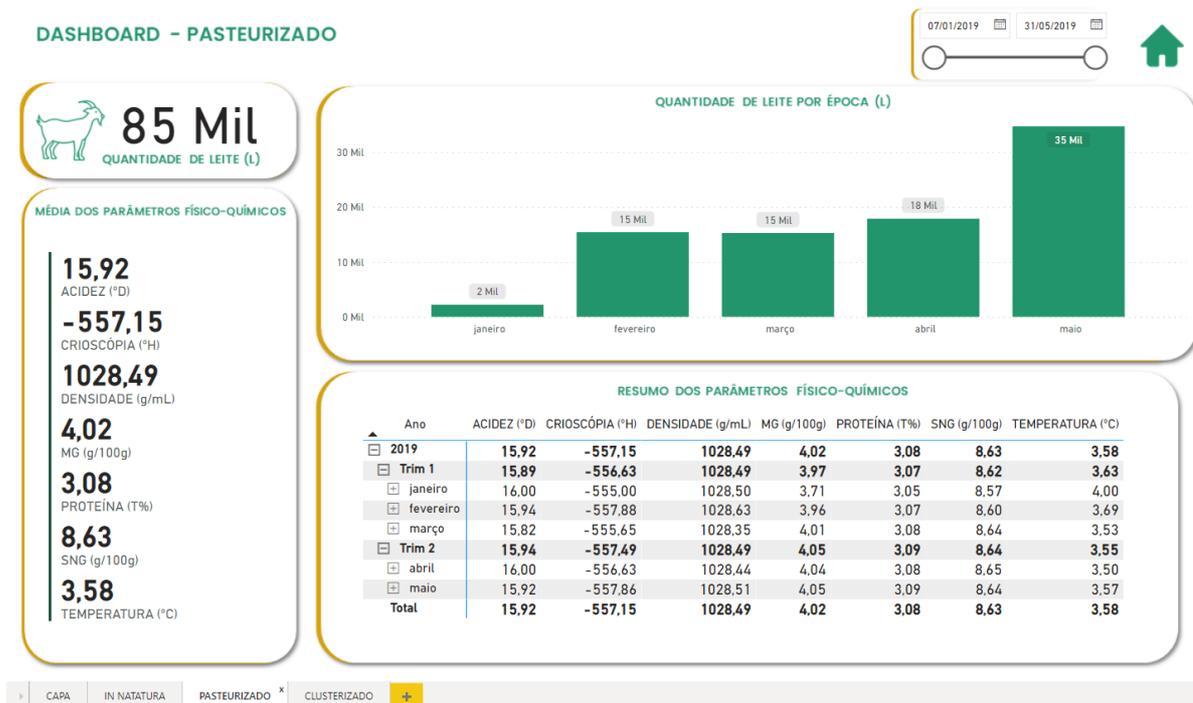
Figura 10 - Ilustração dos tipos de não conformidades



Fonte: Aatoria Própria (2023)

Para construção da terceira tela do *dashboard*, referente à análise do leite pasteurizado, optou-se por utilizar os mesmos visuais para representação dos mesmos tipos de informações. Na Figura 11 é possível visualizar a disposição dessas informações.

Figura 11 - Layout da análise do leite pasteurizado



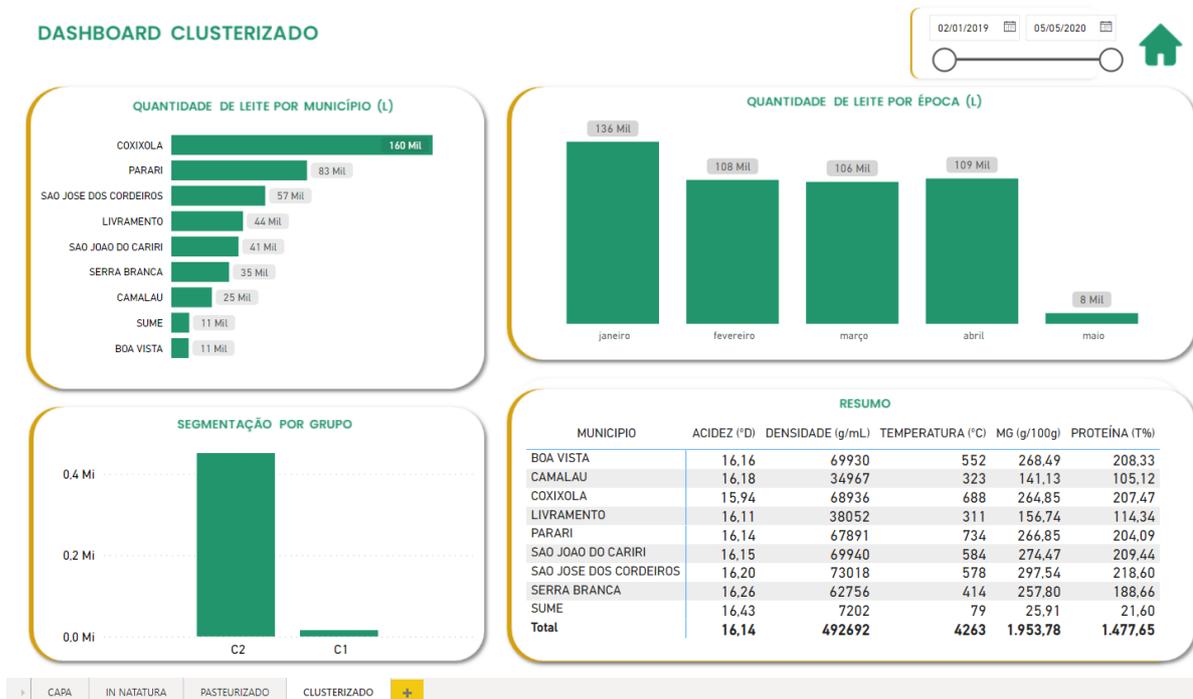
Fonte: Autoria Própria (2023)

O quarto painel do *dashboard*, representado a partir da Figura 12, contou com as informações referentes à *clusterização* das amostras de leite. Para construção deste foi utilizado o visual “Gráfico de Barras Empilhadas” para representação do quantitativo de litros fornecidos por município.

O “Gráfico de Colunas Empilhadas” foi adotado para apresentação da quantidade de litros de leite fornecidos por época. Esse mesmo visual, também, foi utilizado para apresentar os resultados da *clusterização*, conforme é possível visualizar no visual denominado “Segmentação por Grupos”.

Por fim, o último visual utilizado para esse painel foi a “**Matriz**” a qual apresenta uma ampla visão acerca dos parâmetros físico-químicos e seus respectivos municípios.

Figura 12 - Representação do painel referente à análise de *clusterização* do leite



Fonte: Autoria Própria (2023)

Por fim, após a apresentação do sistema à organização e implantação. A gestão relatou que a ferramenta será de grande valia para o monitoramento da qualidade do leite, bem como do processo. Atuando também no suporte à tomada de decisão, visto que possibilitará uma ampla visão acerca do processo produtivo.

6. CONSIDERAÇÕES FINAIS

A presente pesquisa buscou, inicialmente, aplicar técnicas de mineração de dados, tendo em vista a descoberta de informações úteis, implícitas nos dados que a princípio eram apenas armazenados. Com a validação dessa primeira fase por meio dos resultados obtidos com a análise exploratória e de agrupamento dos dados, deu-se início a construção de painéis de controle dinâmicos e interativos que possibilitaram a otimização do processo produtivo da organização.

Através dos indicadores disponibilizados no *dashboard* foi possível acompanhar o quantitativo de litros de leite fornecidos por período de tempo e por município, bem como, observar o valor médio para cada parâmetro físico-químico, as não conformidades presentes e os grupos compostos por amostras que possuem características semelhantes.

Dessa forma, considera-se que o método adotado (*CRISP-DM*) apresentou eficiência no que diz respeito aos resultados obtidos, configurando-se como um eficaz balizador metodológico para processos de preparação, modelagem e análise de dados. Enquanto a ferramenta *Orange Canvas* destacou-se devido a rapidez na emissão de resultados e na construção de modelos. Contudo, é importante salientar que as ferramentas não substituem a necessidade de um conhecimento prévio e profundo do domínio de exploração, por parte dos pesquisadores.

Quanto à construção do *dashboard*, a integração do *Google Sheets* com a ferramenta *Power BI*, demonstrou ser uma boa alternativa, demonstrando transparência das informações, capacidade de interpretação, facilidade no acesso, avaliação e junção de informações, mobilidade, independência, colaboração e organização. Ou seja, informações que anteriormente estavam espalhadas por diversos documentos e arquivos distintos, dificultando o encontro de informações relativas ao semestre de vendas analisado.

Para trabalhos futuros seguem-se novos desafios de expansão e integração dos painéis de controle em outros setores da organização. Ademais, espera-se que as informações fornecidas a partir desse estudo possam sustentar discussões e questionamentos quanto à utilização de *business intelligence* apoiado por técnicas de mineração de dados para extração de conhecimentos capazes de subsidiar a tomada de decisão.

REFERÊNCIAS

ABRAMOVAY, R. Agricultura familiar e serviço público: novos desafios para a extensão rural. **Cadernos de Ciência e Tecnologia**, Brasília, v. 15, n. 1, p. 132-152, 1998.

AGNESE, A. P.; NASCIMENTO, A. M. D. do; VEIGA, F. H. A.; PEREIRA, B. M.; OLIVEIRA, V. M. de. Avaliação físico química do leite cru comercializado informalmente no Município de Seropédica – RJ. **Revista Higiene Alimentar**, v.16, n. 94. p. 58-61, 2002.

ALMEIDA, J. F., Aquino, M. H. C, Magalhães, H. *et al.* Principais alterações no leite por agentes causadores de mastite no rebanho caprino dos estados de Minas Gerais e Rio de Janeiro. **Arquivos do Instituto Biológico**. 80(1), 2013, p. 13-18.

ALMEIDA, A. F.; PERIN, G.; SAMBUICHI, R, HALVES, Laís Santos. COMPOSIÇÃO FÍSICO-QUÍMICA E CONTAGEM DE CÉLULAS SOMÁTICAS EM LEITE DE CABRAS NO MUNICIPIO DE GURJÃO/PB. 2018.

AMORIM, A. L. B. do C. **Avaliação da presença de substâncias químicas em leite cru e beneficiado produzidos e comercializados no Distrito Federal e Entorno**. 2017. Dissertação de mestrado em saúde animal. Faculdade de Agronomia e Medicina Veterinária – Universidade de Brasília, 2017.

ANDRADE, P. V. D., Souza, M. R., Penna, C. F. A. M. E Ferreira, J. M. 2008. Características microbiológicas e físico-químicas do leite de cabra submetido à pasteurização lenta pósenvase e ao congelamento. **Ciênc. Rural**.

ARAÚJO, Vinicius José Apropriano de. **Qualidade do leite de cabra in natura processado em mini-usinas do Médio Sertão e Cariri Paraibano - estudo comparativo**. Trabalho de Conclusão de Curso (Monografia), Curso de Bacharelado em Medicina Veterinária, Centro de Saúde e Tecnologia Rural, Universidade Federal de Campina Grande - Patos - Paraíba - Brasil, 2008. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/24385>

ATZMUELLER, M. (2016), Enterprise Big Data Engineering, Analytics, and Management, Advances in Business Information Systems and Analytics, IGI Global.

BATISTA, N. L., SOUZA, B. B., LOPES, J. J. Efeito do ambiente sobre as respostas fisiológicas de caprinos Saanen e seus mestiços com a raça Boer no semiárido Paraibano. **Revista Agropecuária Científica no Semiárido**. v.8, n.3, p.83-89. 2012.

BELOTI, V.; TAMANINI, R. Propriedades Organolépticas e físico-química do leite. In: BELOTI, V. (Org). Leite: obtenção, inspeção e qualidade. Londrina: **Editora planta**, 2015, p. 51-108.

BENEDET, H.D.; CARVALHO, M.W. Caracterização do Leite de Cabra no Estado de Santa Catarina, Brasil. **Ciênc. Tecnol. Aliment**. Campinas, 1996, v.16, n.2, p.116- 119.

BENTLEY, D. Business Intelligence and Analytics. (D. Bentley, Ed.) New York, Estado Unidos da América: **Library Press**, 2017.

BERGAMASCO, Sônia M. P. P. Família e trabalho rural no Brasil e no Estado de São Paulo. **Informações Econômicas**, São Paulo, vol. 23: 1-58: 7-16, 1993.

BUGNION, Pascal; MANIVANNAN, Arun; NICOLAS, Patrick R. (2017). **Scala: Guide for Data Science Professionals**. Birmingham: Packt Publishing, 2017.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento. Secretaria Nacional de Agricultura. Instrução Normativa nº 37 de 31 de outubro de 2000. Regulamento técnico de produção, identidade e qualidade do leite de cabra. Diário Oficial da República Federativa do Brasil. Brasília, DF, 8 nov. 2000. Seção 1, p. 23.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento.. Resolução nº 1 do Grupo Gestor do Programa de Aquisição de Alimentos, de 31 de julho de 2003. Define a sistemática de aquisição da produção da agricultura familiar para os produtores enquadrados no Programa Nacional de Fortalecimento da Agricultura Familiar – Pronaf e estabelece critérios para a aquisição e doação de gêneros alimentícios. Diário Oficial da União, Brasília, 31 jul. 2003.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento.. Decreto nº 7.775, de 4 de julho de 2012. Regulamenta o art. 19 da Lei nº 10.696, de 2 de julho de 2003, que institui o Programa de Aquisição de Alimentos, e o Capítulo III da Lei nº 12.512, de 14 de outubro de 2011, e dá outras providências. Diário Oficial da União, Brasília, p. 3, 5 jul. 2012.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento.. Resolução no 62, de 24 de outubro de 2013. Dispõe acerca da destinação dos alimentos adquiridos com recurso do PAA (revogada pela Resolução no 72, de 9 de outubro de 2015). Diário Oficial da União, 25 out. 2013.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento.. Resolução nº 69 do Grupo Gestor do Programa de Aquisição de Alimentos, de 18 de setembro de 2014. Altera os arts. 2º e 3º da Resolução nº 62, de 24 de outubro de 2013, do Grupo Gestor do Programa de Aquisição de Alimentos. Diário Oficial da União, Brasília, 22 set. 2014.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento.. Ministério da Agricultura, Pecuária e Abastecimento. **Regulamento da Inspeção Industrial e Sanitária de Produtos de Origem Animal (RIISPOA)**. Da classificação dos derivados lácteos, Seção II, Subseção IV. Brasília, 2017.

BRASIL, Ministério da Agricultura, Pecuária e Abastecimento. Medida Provisória no 1.061, de 9 de agosto de 2021. Institui o Programa Auxílio Brasil e o Programa Alimenta Brasil, e dá outras providências. Diário Oficial da União, Brasília, p. 6, 10 ago. 2021.

BROWN, C. V., DEHAYES, D. W., HOFFER, J. A., MARTIN, E. W., & PERKINS, W. C. **Managing Information Technology** (7.ª Edição ed.). Upper Sadlle River: Pretince Hall, 2012.

BRYNJOLFSSON, E.; HITT, L. M.; KIM, H. H. Strength in numbers: How does datadriven decisionmaking affect firm performance?. Available at SSRN 1819486, v. 1, n. 1, p. 1-33, 2011.

CALLINGHAM, M. **Market intelligence: how and why organizations use market research.** London: Kogan Page Publishers, 2004.

CAMARGO, R. A. L. De; BACCARIN, J. G.; SILVA, D. B. P. Da. O papel do Programa de Aquisição de Alimentos (PAA) e do Programa Nacional de Alimentação Escolar (PNAE) no fortalecimento

CARDOSO, Olinda N. P., MACHADO, Rosa T. M. Gestão do conhecimento usando data mining: estudo de caso na Universidade Federal de Lavras. **Revista de Administração Pública.** Rio de Janeiro 42(3) : 495-528, Maio/Jun. 2008 da agricultura familiar e promoção da segurança alimentar. TEMAP, 2013. v. 8, n. 2.

CARVALHO, J. S., SILVA, T. R., SANTOS, P. V. M., Almeida, F. F., Jesus, T. K. F., & Rizzo, H (2020). Characterization of goat and sheep production in the state of Sergipe, Northeast of Brazil. *Acta Veterinaria Brasilica*, 14(2), 121-131. <https://doi.org/10.21708/avb.2020.14.2.9247>.

CASARRO, A.C. *Sistemas de informações para tomada de decisões.* São Paulo: Pioneira Thomson Learning, 2010.

CASTRO, J. M. L. T. de. (2016). Tendências de Business Intelligence. Retrieved from <<https://run.unl.pt/bitstream/10362/19434/1/TGI0065.pdf>>.

COSTA, H. N., Molina, L. R., Lage, C. F. A., Malacco, V. M. R., Facury Filho, E. J. & Carvalho, A. Ú. 2017. Estimativa das perdas de produção leiteira em vacas mestiças Holandês x Zebu com mastite subclínica baseada em duas metodologias de análise. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia*, 69, 579-586.

CHAN, A.; PEYNE, B. *Data-Driven decision making in Marketing: A theoretical approach.* Halmstad University, School of Business, Engineering and Science. 2017.

CHAPMAN, Pete et al. The CRISP-DM process model. The CRIP-DM Consortium, v. 310, p. 91, 1999.

CRUZ GRB, Costa RG, Queiroga RCRE. Características físicas do leite de cabra produzido no Estado da Paraíba. In: Anais da 35ª Reunião Anual da Sociedade Brasileira de Zootecnia. Botucatu: Sociedade Brasileira de Zootecnia, 1998.

DUNCAN, J. & Prasse, K. 1982. *Patologia clínica veterinária.* UFSM, Santa Maria, 1982.

DÜRR, J. W. Programa Nacional de Melhoria da Qualidade do Leite: uma oportunidade única. In: CARVALHO, M. P.; SANTOS, M. V. dos S. O compromisso com a qualidade do leite no Brasil. Passo Fundo: Upf, 2004. p. 38- 55.

DÜRR, J. W. *Como produzir leite de alta qualidade.* Brasília: SENAR, 2005. 28 p.

FAGOTTI, L. N. O 'fazer-se' associativo: associativismo e agricultura familiar no interior paulista. Repositório Institucional UNESP, 2017. Disponível em: <http://hdl.handle.net/11449/150693>. Acesso em: 02 de abril de 2020.

FAO/INCRA. Diretrizes de política agrária e desenvolvimento sustentável. Versão Resumida do Relatório Final do Projeto UFT/BRA/036, Novembro. Brasília: Incra, 1994.

FELISBERTO, N. D. O., OLIVEIRA, L., CORDEIRO, A. Sistemas de produção de caprinos 50 leiteiros. In: Embrapa Caprinos e Ovinos-Artigo em anais de congresso (ALICE). In: WORKSHOP SOBRE PRODUÇÃO DE CAPRINOS NA REGIÃO DA MATA ATLÂNTICA, 13., 2016, Coronel Pacheco. Anais... Sobral: Embrapa Caprinos e Ovinos; Coronel Pacheco: Embrapa Gado de Leite, 2016. p. 11-35.

FERRARI, A., & Russo, M. (2016). Introducing Microsoft Power BI. (R. Caperton, O. P. Russel, Diane, & O. P. Russel, Bob, Eds.). Redmond: Microsoft Press, 2016.

FERNANDES, M. F., Queiroga, R. C. R. E., Medeiros, A. N., Costa, R. G., Bonfim, M. A., & Braga, A. A. (2008). Características físico-químicas e perfil lipídico do leite de cabras mestiças Moxotó alimentadas com dietas suplementadas com óleo de semente de algodão ou de girassol. **Revista Brasileira de Zootecnia**, 37(4), 2008. 703-710.

FONSECA, C. E. M.; SILVA, T. L.; OLIVEIRA, C. A. Caprinocultura. Niterói: Programa Rio Rural, 2012.

FONSECA, C. R.; PORTO, E.; DIAS, C. T. S.; SUSIN, I. Qualidade do leite de cabra in natura e do produto pasteurizado armazenados por diferentes períodos. **Ciência e Tecnologia Alimentar**, v. 26, n. 4, p. 944-949, 2006.

FONSECA, J. F. da; BRUSCHI, J. H. **A caprinocultura leiteira no Brasil: uma visão histórica**. Produção de caprinos na região da Mata Atlântica. Juiz de Fora: Embrapa Gado de Leite; Sobral: Embrapa Caprinos e Ovinos, 2009, p. 15-24.

FORNAZIER, A.; BELIK, W. Articulação entre políticas de compras governamentais da agricultura familiar e políticas territoriais. **Planejamento e Políticas Públicas**. n. 52, jan/jun. 2019.

FORTULAN, M. R.; GONÇALVES FILHO, E. V. Uma proposta de aplicação de Business Intelligence no chão-de-fábrica. **Revista Gestão e Produção**, v.12, n.1, pp. 55-66, jan.-abr., 2005.

GOLDSCHMIDT, R.; PASSOS, E.; BEZERRA, E. **Data mining: conceitos, técnicas, algoritmos, orientações e aplicações**. Rio de Janeiro: Elsevier, 2015.

GOPALKRISHNAN, V., STEIER, D., Lewis, H., GUSZCZA, J. Big data, big business: Bridging the gap. Preceedings of the International Workshop on Big Data, Streams and Heterogeneous Source Mining: Algorithms, Systems, Programming Models and Applications, p. 7-11, 2012.

GARTNER. (2019). Analytics and Business Intelligence (ABI). Retrieved October 12, 2019, from <<https://www.gartner.com/en/information-technology/glossary/businessintelligence-bi>>.

GOVERNO DA PARAÍBA (2019). Sedh apresenta novas diretrizes para execução do Programa Leite da Paraíba. <https://paraiba.pb.gov.br/noticias/sedh-apresentanovas-diretrizes-para-execucao-do-programa-leite-da-paraiba>.

GRISA, C. As redes e as instituições do Programa de Aquisição de Alimentos (PAA). **Revista Brasileira de Gestão e Desenvolvimento Regional**, Taubaté, v. 6, n. 2, p. 97-129, maio-ago. 2010.

GRISA, C.; SCHNEIDER, S. Três gerações de políticas públicas para a agricultura familiar e formas de integração entre sociedade e Estado no Brasil. **Rev. RESR**, v. 52, n. 1, p. 125-146, 2014.

IBM Knowledge Center. CRISP-DM Help Overview. In.: IBM SPSS Modeler CRISP-DM Guide. 2012 Disponível em: <https://www.ibm.com/support/knowledgecenter/SS3RA7_sub/modeler_crispdm_ddit_a/clementine/crisp_help/crisp_overview.html> Acessado em: 06/12/2022.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Censo Agropecuário. 2017. Disponível em: <https://sidra.ibge.gov.br/home/pms/brasil>. Acesso em 03 de janeiro de 2023.

JANVRIN, D. & WATSON, M. Big Data: A new twist to accounting. **Journal of Accounting Education**. 38, 2017. 3-8.

LAMARCHE, Hugues. (Coord.). **A agricultura familiar**. Comparação internacional - uma realidade multiforme. Tradução: A. Tijiwa. Coleção Repertórios Campinas: Ed. UNICAMP, 1993.

LEITE, S. Autoconsumo y sustentabilidad en la agricultura familiar: una aproximación a la experiencia brasileña. In: W. Belik (COORD.), **Políticas de seguridad alimentaria y nutrición en América Latina**, São Paulo: Hucitec, 2004.

LIMA G.F.C. Reservas estratégicas de forragem de boa qualidade para bovinos leiteiros, p.11-35. In: Brito A.S., Nobre F.V. & Fonseca J.R.R. (Eds), **Bovinocultura Leiteira: informações técnicas e de gestão**. SEBRAE/RN, Natal, 2009. 320p.

LISBOA, A. S, ALCANTARA, F. V. Espaço Rural e Associativismo no Fortalecimento na Agricultura Familiar. VII SINGA, 2015.

LONG, Q. Data-driven decision making for supply chain networks with agent-based computational experiment. **Knowledge-Based Systems**, v. 141, n. 1, p. 55-66, 2018.

LOPES, J. E. F.; SOARES, D. J. Empresas Orientadas a Dados e Análises: as pessoas estão preparadas? In: ENCONTRO DE GESTÃO E NEGÓCIOS - EGEN, 2016, Uberlândia - MG. Anais...Uberlândia, MG. 2016

LU, M. . Discovering Microsoft Self-service BI solution : Power BI. Haaga-Helia - University of Applied Sciences. 2014.

MALHOTRA, N. K.; BIRKS, D. F. Marketing Research: an applied approach. 3ed. Harlow: Prentice Hall, 2007.

MALLADI, S. (2013). Adoption of Business Intelligence & Analytics in Organizations – An Empirical Study of Antecedents. AMCIS 2013 Proceedings, 2016, 1–11. Retrieved from <<http://aisel.aisnet.org/amcis2013/BusinessIntelligence/GeneralPresentations/3>>.

MARQUES, M. et al. Decentralized decision support for intelligent manufacturing in Industry 4.0. **Journal of Ambient Intelligence and Smart Environments**, v. 9, n.3, p.299-313, 2017.
MENEZES, V. F. “Miunça” e caprinocultura: entrelaçamento de lógicas sociais da pecuária caprina e o PAA/Leite no Cariri Paraibano. *Raízes*. v.35, n.2, p.66-82. 2015.

MENDES, C. G., Silva, J. B. A., & Abrantes, M. R. (2009). Caracterização organoléptica, físico-química, e microbiológica do leite de cabra: uma revisão. **Acta Veterinaria Brasilica**, 3(1), 5-12.

NHACUONGE, J. A. O campo da Ciência da Informação: contribuições, desafios e perspectivas da mineração de dados para o conhecimento pós-moderno. 2015. 194 f. Tese (Doutorado em Ciência da Informação)– Programa de Pós-Graduação em Ciência da Informação, Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília – SP. 2015.

OLIVEIRA, P. M. de; FERRANTE, V. L. S. B. O Programa de Aquisição de Alimentos no PDS Sepé Tiarajú: bloqueios e perspectivas. **Retratos de Assentamentos**, v. 16, n. 1, p. 77-110, 2013.

ORDÓÑEZ, J. A. **Tecnologia de alimentos: alimentos de origem animal**, v. 2. Porto Alegre: Artmed, 2005. 279p.

PERACI, A. S.; BITTENCOURT, G. A. Agricultura familiar e os programas de garantia de preços no Brasil: o Programa de Aquisição de Alimentos (PAA). In: SILVA, J. G. da; DEL GROSSI, M. E.; FRANÇA, C. G. de. **Fome Zero: a experiência brasileira**. Brasília, DF: Ministério do Desenvolvimento Agrário, 2010. p. 191-222.

PEREIRA, Daniel. Dados e Análise Preditiva na tomada de decisão (Big Data, IOT e Predictive Analytics), 16 de março de 2016.

PEREIRA, C. S. **Qualidade do leite de cabra in natura pela detecção de microrganismos, susceptibilidade antimicrobiana, parâmetros físico-químicos, contagem de células somáticas, contagem total bacteriana e resíduo antimicrobiano**. Universidade Federal Fluminense, Rio de Janeiro, 102f. Tese (Doutorado em Higiene Veterinária e 52 Processamento Tecnológico de Produtos de Origem Animal), 2016.

PEREIRA, R. A. G.; QUEIROGA, R. C. R. E.; VIANNA, R. P. T.; OLIVEIRA, M. E. G. Qualidade química e física do leite de cabra distribuído no Programa Social “Pacto Novo Cariri”

no Estado da Paraíba. **Revista do Instituto Adolfo Lutz**, São Paulo, v. 64, n. 2, p. 205-211, 2005.

PINHEIRO, J. G., Aroucha, E. M. M., Abrantes, M. R., Figueredo, J. P., Góis, V. A., & Silva, J. B. A. (2014). Características físico-químicas do leite caprino na época seca e chuvosa na microrregião de Mossoró-RN. **Acta Veterinaria Brasilica**, 8(3), 2014. 192-200.

QUEIROGA, R. C. R. E., Trigueiro, I. N. S., Ferreira, M. C. C. 1998. Caracterização do leite de cabras mestiças do Brejo Paraibano, durante o período de lactação. **Hig. Aliment.** 12:77-80.

QUEIROGA, R. C. R. E., Costa, R. G., Biscotini, T. M. B., Medeiros, A. N., Madruga, M. S. e Shuler, A. R. P. Influência do manejo do rebanho, das condições higiênicas da ordenha e da fase de lactação na composição química do leite de cabras Saanen. **Rev. Bras. Zootec.** 2007.

RAINER JR, R.K.; CEGIELSKI, C.G. **Introdução a Sistemas de Informação**. Rio de Janeiro: Campus, 2012.

RESENDE, D. **Agricultura e Pecuária no Norte e Nordeste**. São Paulo: Rideel, 2012.

SANTOS, A. O. et al. O Programa aquisição de alimentos (PAA) na região Metropolitana de Santarém (Pará): o caso das mulheres agricultoras da COOMAPLAS. **Brazilian Journal of Development**, v. 5, n. 7, p. 11090-11106, 2019.

RICHARDS, N.S.P.S.; PINTO, A.T.; SILVA, M.E.; CARDOSO, V.C. Avaliação físicoquímica da qualidade do leite de cabra pasteurizado comercializado na Grande Porto Alegre, RS. **Revista do Instituto de Laticínios Cândido Tostes**, Juiz de Fora/MG, v.56. n.321, 2001.

SAMBUICHI, R. H. R. et al. Compras públicas sustentáveis e agricultura familiar: a experiência do Programa de Aquisição de Alimentos (PAA) e do Programa Nacional de Alimentação Escolar (PNAE). In: SAMBUICHI, R. H. R. et al. (Orgs.). **Políticas agroambientais e sustentabilidade: desafios, oportunidades e lições aprendidas**. Brasília: Ipea, 2014. p. 75-104.

SANTOS, M. Y., & RAMOS, I. **Business Intelligence da Informação ao Conhecimento**. Lisboa, Lisboa, Portugal: FCA - Editora de Informática, Lda, 2017.

SAMPAIO, B., SAMPAIO, Y., LIMA, R. C., AIRES, A., SAMPAIO, G. A economia da caprinocultura em Pernambuco: problemas e perspectivas. **Revista de Economia**. v. 35, n. 2, p. 137-159. 2009.

SEN, Amartya Kumar. **Desenvolvimento como liberdade**. Tradução: Laura Teixeira Motta. São Paulo: Companhia das Letras, 2000.

SILVA, M. G. DA; SILVA, S. P. Para além do acesso: uma análise da relação entre mercados institucionais e empreendimentos de economia solidária. **Mercado de Trabalho**, n. 49, 2011.

SILVA, G. M., Nardelli, M. J., Silva, A. C., Souza, M. S., Mesquita, E. P., & Carvalho, M. G. X. (2015). Avaliação da qualidade do leite de cabra in natura produzido no estado de Sergipe. **Ciência Veterinária nos Trópicos**, 18(2), 2015. 284-287.

SILVA, J. N., Araújo, A. C., Santos, E. P., Holanda Neto, J. P., & Alves, T. T. L. (2011). Parâmetros e determinantes da qualidade físico-química do leite caprino. **Revista Verde**, 6(3), 2011. 32-38.

SILVA, T. C.; FERREIRA, P. R. O Programa de Aquisição de Alimentos sob a ótica dos atores sociais envolvidos. **Desenvolvimento em questão**, v. 14, n. 33, p. 301-329, 2016.

SIMPLÍCIO, A. A., *et al.* Manejo de caprinos para produção de leite em regiões tropicais. **Revista Ciência Animal**, p.13-27, 2000.

SOUZA, B. B. ; SILVA, E. M. N. ; SILVA, G. A. A CAPRINOCULTURA DE CORTE NO BRASIL: Raças especializadas e adaptadas às condições tropicais. 2011. Disponível >http://www.cstroid.sti.ufcg.edu.br/bioclimateologia/artigos_tecnicos/caprinocultura_corte_Brasil_racas_especializadas_adaptadas.pdf< Acesso em 15 de janeiro de 2019

SOUZA, W. H.; RAMOS, J. P. F.; FILHO, E. C. P.; MENEZES, L. M. Potencialidades da exploração de Caprinos Leiteiros na Paraíba. Sobral: Anais do 13º Workshop sobre Produção de Caprinos na Região da Mata Atlântica – EMBRAPA, 2016.

TURBAN, E.; RANIER JR., R. K.; POTTER, R.E. **Introdução a sistemas de Informação uma abordagem gerencial**. Tradução Daniel Vieira. Rio de Janeiro: Elsevier, 2007.

TURBAN, E. Business intelligence: um enfoque gerencial para a inteligência do negócio. São Paulo: Bookman, 2009.

TWO CROWNS CORPORATIONS. Introduction to Data Mining and Knowledge Discovery. 3 ed. Potomac: Two Crowns, 2005.

VEIGA, José Eli da. **O desenvolvimento agrícola: uma visão histórica**. São Paulo: Edusp/Hucitec, 1991. 219 p.

VENTURINI, K. S. SARCINELLI, M. F. SILVA, L. C. Obtenção do leite. Universidade Federal do Espírito Santo – UFES, 2007. Disponível em:<http://www.agais.com/telomc/b01207_obtencao_leite.pdf> Acesso em: 22/12/2022.

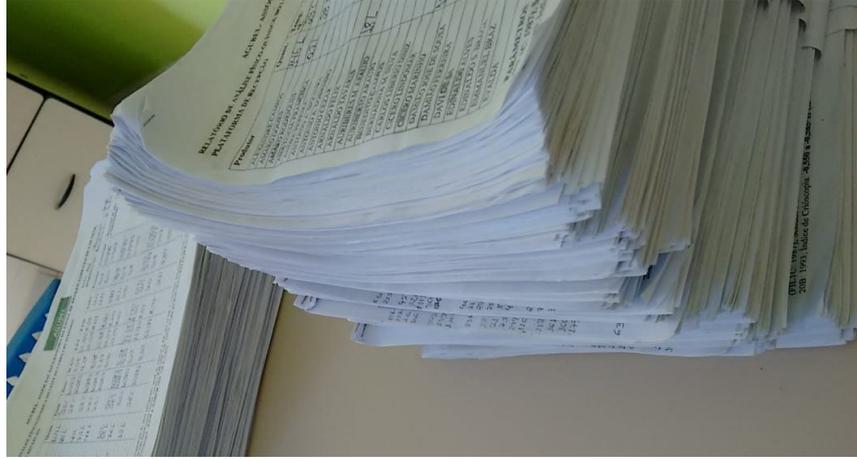
VOGT, S. P. C. Mercados Institucionais Locais como instrumento de fortalecimento da agricultura familiar: uma análise do Programa de Aquisição de Alimentos na Região Ceilândia – RS. In: CONGRESSO DA SOCIEDADE BRASILEIRA DE ECONOMIA, ADMINISTRAÇÃO E SOCIOLOGIA RURAL, 47, Porto Alegre, 2009. Anais... Brasília: Sober, 2009.

SPINOLA, L. H. O. Gestão da informação: Conceitos, aplicabilidade, desafios e perspectivas da área – a ótica do bibliotecário. Brasília, 2013. 70f. Trabalho de Conclusão de Curso (monografia). Universidade de Brasília, Faculdade de Ciência da Informação, Curso de Biblioteconomia. 2013. Disponível em: http://bdm.unb.br/bitstream/10483/7102/1/2013_LeandroHenriquedeOliveiraSpinola.pdf Acesso em 11/12/2022.

XU, Z. B. *et al.* The frontiers in big-data-driven management and decision making. *Manage. World China*, v. 11, n.1, p. 158-163, 2014.

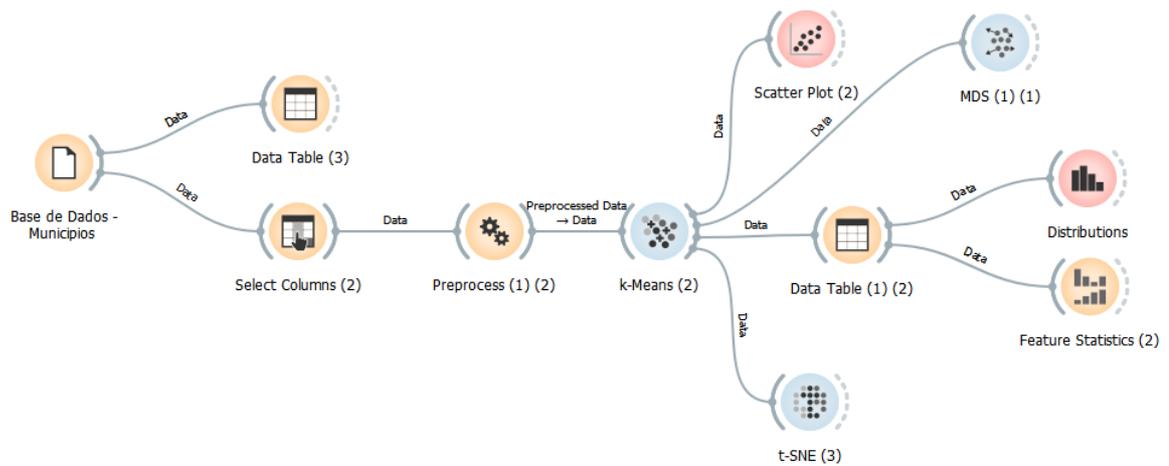
APÊNDICES

Apêndice A - Planilhas físicas das análises físico-químicas



Fonte: Autoria própria (2023)

Apêndice B - Análises gráficas a partir dos grupos formados



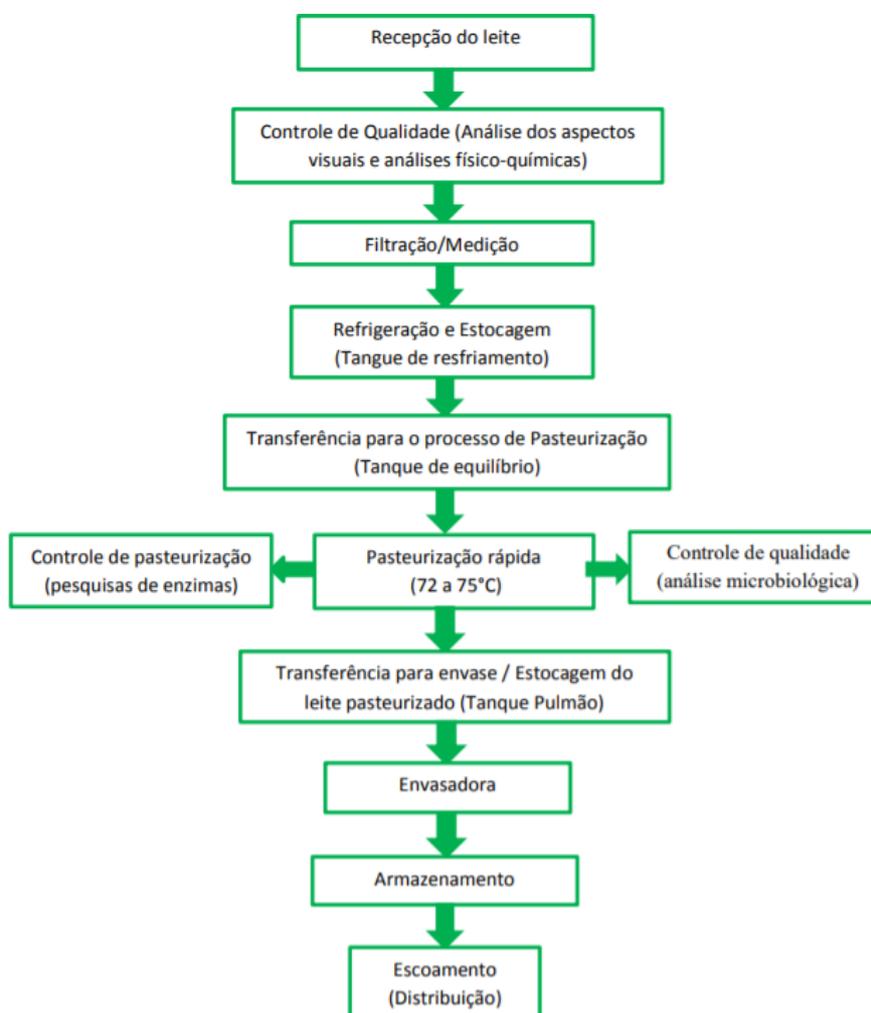
Fonte: Autor da pesquisa (2023)

ANEXOS

Anexo A - Índice de Validação - Silhueta

SM	Interpretação
0.71 - 1.00	Grupos produzidos pelo algoritmo possuem uma estrutura muito robusta
0.51 - 0.70	Grupos possuem uma estrutura razoável
0.26 - 0.50	Os grupos encontrados possuem uma estrutura fraca e pode ser artificial
< 0.25	Nenhuma estrutura foi descoberta

Anexo B - Fluxograma do processo produtivo do leite de cabra



Fonte: AGUBEL, 2020.