

Ciência de dados aplicada à pecuária de corte

71

Data science applied to beef cattle production

Thaís Basso Amaral¹, Guilherme Jordão de Magalhães Rosa², Marcos Cezar Visoli³, Fabio Martins Guerra Nunes Dias⁴

¹ Pesquisadora Embrapa Agricultura Digital, Campinas (SP), Brasil, thais.amaral@embrapa.br

² Professor, Department of Animal and Dairy Sciences, University of Wisconsin, Madison (WI), EUA, grosa@wisc.edu

³ Pesquisador Embrapa Agricultura Digital, Campinas (SP), Brasil, marcos.visoli@embrapa.br

⁴ Friboi - JBS S.A., São Paulo (SP), Brasil, fabio.dias@friboi.com.br

RESUMO

Com o crescimento da utilização de dispositivos móveis, do uso de sensores, do aumento da conectividade no campo, dados estão sendo coletados a uma taxa exponencialmente crescente nos sistemas de produção pecuários. Diariamente são gerados dados climáticos, provenientes de satélites, índices de vegetação, dados econômicos, isso sem contar os dados coletados nas propriedades, como ganho de peso, consumo de alimentos, qualidade de carcaça etc. De forma isolada esses dados podem auxiliar o produtor na tomada de decisão, porém quando combinados podem gerar *insights* valiosos para toda a cadeia. Ao projetar tendências futuras, pode-se otimizar a alocação de recursos, tornando toda a cadeia produtiva mais sustentável. Neste capítulo será abordado como o big data e o analytics podem contribuir para a geração de valor na cadeia da carne e superar os desafios que a pecuária de corte vem enfrentando.

Palavras-chave: análise de dados; big data; inteligência artificial; produção pecuária.

ABSTRACT

With the increasing use of mobile devices and sensors, and growing connectivity in the field, data is being collected at an exponentially increasing rate in livestock production systems. Climate data are generated daily from satellites, vegetation indices, and economic data, as well as data collected from properties, including weight gain, food consumption, carcass quality, etc. In isolation, such data can assist the producer during decision-making, but in aggregate, they provide valuable insights for the entire production chain. When projecting future trends, it is possible to optimize resource allocation, making the entire production chain more sustainable and efficient. This chapter discusses how big data and analytics can contribute to generating value in the meat production chain and help solve the challenges that beef cattle production faces.

Keywords: analytics; artificial intelligence; beef production; big data.

<https://doi.org/10.4322/978-65-86819-38-0.1000063>

 Este é um capítulo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution-NonCommercial-NoDerivatives, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que sem fins comerciais, sem alterações e que o trabalho original seja corretamente citado.



1 INTRODUÇÃO

A sustentabilidade da pecuária de corte depende da otimização contínua de um sistema extremamente complexo envolvendo técnicas de manejo animal, melhoramento genético, instalações e equipamentos, insumos, transporte animal e processamento e comercialização dos produtos de origem animal. Tal otimização envolve não somente a melhoria da produtividade (i.e., produção por unidade de área) mas também a eficiência da produção em termos da utilização de insumos, incluindo energia e água, de maneira a garantir retorno econômico satisfatório ao produtor com mínimo impacto ambiental. Além disso, a produção animal moderna deve corresponder também às novas exigências de consumidores em termos de qualidade de produto, bem-estar animal e preservação da biodiversidade.

Dada essa complexidade do sistema produtivo e dos vários fatores que devem ser considerados para a sustentabilidade da pecuária, o produtor rural e a indústria pecuária em geral têm que tomar decisões continuamente e em tempo real, num contexto dinâmico de disponibilidade e de mudanças de preço de insumos e produtos, de novas exigências de mercado, interno e externo, da incidência de doenças, mudanças de clima etc. Tais decisões só podem ser otimizadas se tomadas de maneira objetiva e baseadas em dados.

A cadeia da carne bovina é longa e complexa, caracterizando-se pela enorme quantidade de fornecedores de animais, de ambientes de produção, raças e tecnologias de criação muito diferentes. Com o aumento da capacidade de armazenamento e processamento de dados, o uso de sensores no campo, o acesso à internet, a pecuária vem se tornando cada vez mais orientada a dados, sendo fundamental para apoiar a tomada de decisões a nível da fazenda (Morota et al., 2018).

Os grandes volumes ou coleções de dados podem ser uma fonte valiosa de informações para efetivamente gerar impacto e apoiar as demandas que vêm sendo enfrentadas pelo setor (Kamilaris et al., 2017; Liakos et al., 2018; Morota et al., 2018). As organizações terão de desenvolver abordagens inteligentes para traduzir os dados em percepções acionáveis e monetizáveis (Empresa Brasileira de Pesquisa Agropecuária, 2022). Plataformas digitais que geram dados estão crescendo constantemente. Big data, modelos preditivos, ciência de dados e data analytics estão ajudando a revolucionar o agro, revelando que é possível otimizar tudo com maior capacidade de armazenamento, processamento e análise, produzir novos conhecimentos para o negócio e para o produtor, assim como traçar tendências e prever o futuro. O

desafio é fazer o big data chegar aos pequenos e aos grandes produtores, pois isso exigirá crescentemente profissionais com habilidades digitais, de reflexão (questionadores) e atentos ao ambiente de negócios (Empresa Brasileira de Pesquisa Agropecuária, 2022).

Neste capítulo serão abordados alguns conceitos relacionados à temática, além de demonstrar alguns exemplos de como a ciência de dados pode contribuir para os desafios atuais da cadeia da carne bovina.

2 TIPOS E FONTES DE DADOS

A ciência de dados consiste na tomada de decisão baseada em dados. Isso se dá a partir do tratamento de uma imensidão e diversidade de dados brutos, a fim de transformá-los em informações de valor, para realizar análises escaláveis, velozes e precisas, e, dessa forma, gerar insights para decisões mais assertivas, consequentemente, otimização e expansão dos negócios.

Nesse contexto, torna-se imprescindível a mensuração e obtenção de dados de várias fontes, não somente de informação de “dentro da porteira”, como inventário de rebanho, taxas produtivas e reprodutivas, compra e venda de insumos e produção, mas também dados macroeconômicos, dados climáticos, dados de sensores orbitais, de transporte de animais entre fazendas e de fazendas para a indústria de processamento, e tantos outros (Figura 1). Tais conjuntos de dados são gerados por agentes e instituições independentes, como associações de criadores, agências do governo e agências não governamentais, outras entidades públicas e privadas etc. Um primeiro passo nesse contexto de tomadas de decisão baseadas em dados requer a junção de tais repositórios de dados independentes para a geração de data lakes, ilustrados na Figura 2, os quais se referem a banco de dados multidimensionais com informação espaço-temporal. Tal junção ou pareamento de repositórios de dados pode se dar tendo entidade jurídica como variável comum (e.g. nome da fazenda ou CNPJ), ou data (como no caso de valores econômicos), ou informação espacial (como município ou coordenada geográfica), ou combinação de tais variáveis, como no caso de tempo, temperatura e índice pluviométrico (em determinado local e data) ou imagens de satélite.

A diversidade de dados em relação ao seu tipo, origem, tamanho, frequência de atualização e demais características inerentes dos processos que os originam leva ao desafio de como organizá-los e adequá-los para potencializar análises e obtenção de informações úteis para tomada de decisão no momento necessário.



Figura 1. Representação de diferentes fontes de dados e como elas se relacionam de forma a compor um grande conjunto de dados ou big data.

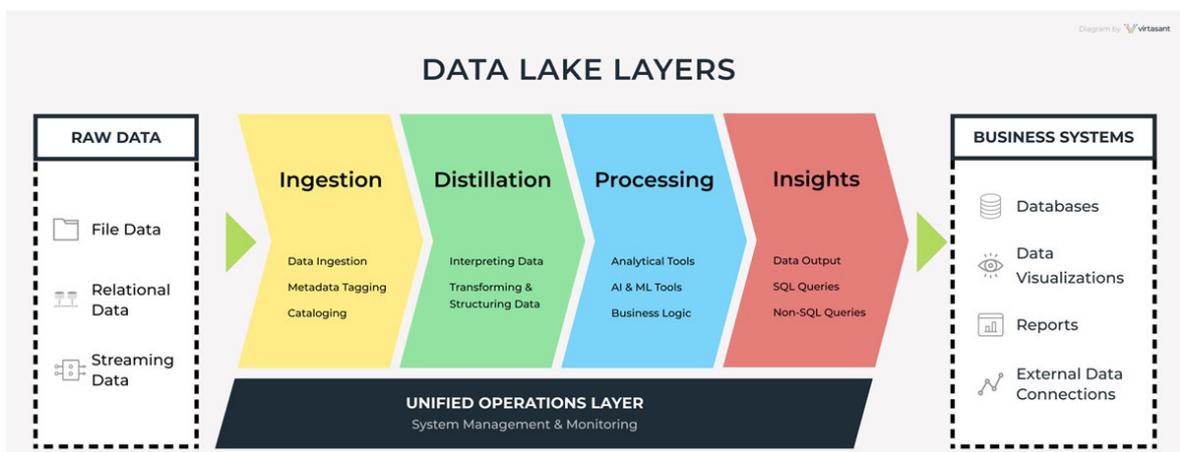


Figura 2. Esquemática da estruturação de dados em data lakes.
Fonte: Hajdarbegovic (2020).

Nesse contexto, a área de gestão de dados vem ganhando destaque, seja em ambientes empresariais quanto nas comunidades e instituições de pesquisa, onde o termo utilizado é gestão de dados de pesquisa.

Segundo o Guia Dama-Dmbokv2 (DAMA International, 2017), a gestão de dados é a função na organização que cuida do planejamento, controle e entrega de ativos de dados e de informação. Esta função inclui as disciplinas do desenvolvimento, execução e supervisão de planos, políticas, programas, projetos, processos, práticas e procedimentos que controlam, protegem, distribuem e aperfeiçoam o valor dos ativos de da-

dos e informações. Pode-se entender também gestão de dados como uma área ou conjunto de disciplinas responsável por prezar pelos dados da organização, fazendo com que sejam únicos, íntegros, confiáveis e disponíveis para as pessoas que realmente precisam ter acesso.

Baseado no Dama-Dmbok v2, foram definidas 11 áreas de atuação para a sua efetiva implementação, representadas na Figura 3 e descritas posteriormente.

- Governança de dados, que envolve planejamento, supervisão e controle sobre o gerenciamento de dados e o uso de dados e recursos relacionados a dados.



Figura 3. Framework de gestão de dados. Figura adaptada/traduzida de DAMA-DMBOK®V2 (DAMA International, 2017).

- Arquitetura de dados, que define a estrutura geral de dados e recursos relacionados a dados como uma parte da arquitetura da empresa.
- Modelagem e design de dados, que envolvem as tarefas de análise, projeto, construção, testes e manutenção.
- Armazenamento e operações de dados, com as ações de implantação de armazenamento de ativos de dados físicos estruturados e gestão.
- Segurança de dados, que envolve a privacidade, confidencialidade, acesso apropriado, questões legais ligadas a legislações como a LGPD.
- Integração de dados e interoperabilidade, que consiste nos processos de aquisição, extração, transformação, movimento, entrega, replicação, federação, virtualização e suporte operacional.
- Documentos e conteúdo, com as funções relacionadas aos processos de armazenar, proteger, indexar e habilitar o acesso aos dados encontrados em fontes não estruturadas e de disponibilização desses dados para integração e interoperabilidade com dados estruturados.
- Dados mestre e referência, que envolvem o gerenciamento dos dados compartilhados para reduzir a redundância e garantir a qualidade de dados através da definição padronizada e do uso de valores de dados.
- Data Warehousing e Business Intelligence: item amplo, que envolve o gerenciamento de processamento de dados analíticos e permitindo acesso a dados de suporte à decisão para relatórios e análises.

- Metadados, envolvendo coleta, categorização, integração, controle e entrega de metadados.
- Qualidade dos dados, com a definição, monitoramento e manutenção da integridade dos dados e melhoria da qualidade dos dados.

A implantação dessas áreas e funções pode demandar planejamento e investimento a médio e longo prazo em instituições. Abordagens mais simplificadas podem ser aplicadas de acordo com o tamanho e finalidade da instituição ou do trabalho a ser realizado, mas é importante destacar que a gestão de dados é fundamental na atualidade.

3 TIPOS DE ANÁLISES

Após a junção dos dados e editoração dos mesmos para eliminação de redundâncias e controle de qualidade para redução de possíveis erros e valores discrepantes, tais data lakes podem então ser utilizados para a tomada de decisões. Para tanto, eles precisam ser processados utilizando-se métodos estatísticos e computacionais adequados para a análise de dados de grandes dimensões, i.e., big data, os quais requerem recursos computacionais apropriados para manejo e administração dos dados, bem como técnicas eficientes de mineração de dados. A análise de dados e geração de informação necessária para a tomada eficiente de decisões pode ser dividida em três grupos, como ilustrado na Figura 4. Numa primeira etapa, os dados são sumarizados utilizando-se técnicas de *estatística descritiva*, como o cálculo de medidas de tendência central (como média e mediana), medidas de variabilidade (como desvio padrão, diferença entre maiores e menores valores etc.), detecção de valores discrepantes e construção de gráficos e tabelas. Ferramentas modernas para análise descritiva eficiente de dados de grande dimensão referem-se a planilhas dinâmicas como Power BI, Tableau etc.

Como esse é um campo em franco desenvolvimento, grande quantidade de tecnologias é lançada todos os anos, porém dificilmente existirá uma ferramenta que supra todas as funcionalidades necessárias e se integre com as soluções já em uso. Por isso, o mais comum é combinar soluções *open source* com aquelas comerciais. Dentre os softwares *open source* mais utilizados estão o RStudio e o Python. O RStudio sendo mais utilizado para análises estatísticas e o Python mais utilizado por desenvolvedores de *machine learning* (aprendizado de máquina) e inteligência artificial. Para interpretação e visualização dos dados, dentre as ferramentas comerciais destacam-se o Microsoft Power BI e o Tableau. Os destaques dessas ferramentas são a capacidade de gerar representações gráficas

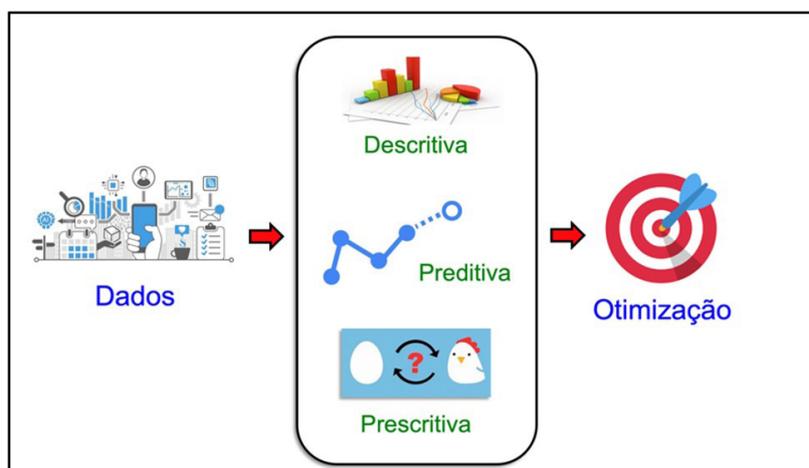


Figura 4. Dos dados à tomada de decisões: análise descritiva, preditiva e prescritiva para a tomada de decisões e otimização do sistema produtivo.

de dados e a interface que facilita a tomada de decisões sobre a forma mais adequada de extrair e interpretar dados.

Estatísticas descritivas são importantíssimas para tomadas de decisão cotidianas, como monitoramento de inventários e decisões de compra e venda, controle produtivo de rebanhos e decisões de descarte de animais, seleção genética, decisões de manejo como vacinação, inseminação, alimentação etc. Além disso, estatísticas descritivas, por mais simples que sejam, são parte fundamental de uma análise estatística mais profunda dos dados, pois auxiliam a escolha de modelos estatísticos apropriados, os quais dependem de pressuposições em relação aos dados, como distribuições probabilísticas, dependências e colinearidades, e tantas outras.

Além do componente descritivo, a análise de dados envolve modelos de predição e modelos que objetivam um melhor entendimento dos fatores envolvidos no sistema. Os modelos de predição, ou a chamada **análise preditiva**, objetivam a previsão de resultados ou valores futuros de variáveis de interesse. Um exemplo do nosso cotidiano são os modelos utilizados na previsão do tempo. Tais modelos utilizam dados de temperatura diária, umidade, pressão e ventos para a predição de chuvas, sol e outras tantas variações climáticas. Modelos preditivos são utilizados também em econometria, por exemplo para decisão e compra e venda futura de insumos e produção. Outros exemplos que podem ser citados são a predição do ganho de peso de bovinos dado o potencial genético e o manejo alimentar dos mesmos. Modelos de predição exploram a associação da variável de interesse com as covariáveis disponíveis no modelo. Técnicas eficientes de predição são modelos de aprendizado de máquina, como máquina de vetores de suporte (SVM, do

inglês *Support Vector Machine*) e rede neural artificial (ANN, do inglês *Artificial Neural Network*). Modelos de predição proporcionam informação importantíssima para a tomada de decisões. Tais decisões, entretanto, são mais no sentido de adaptação de sistemas, administração de recursos etc. do que de mudança do resultado da predição propriamente dito. Por exemplo, a partir de uma previsão de chuva forte para determinado dia, pode-se decidir adiar a distribuição de ração num confinamento bovino. Portanto, há uma adaptação da rotina do manejo alimentar do confinamento, dada a impossibilidade de mudança da previsão de chuva. O mesmo ocorre com decisões de compra e venda em mercados futuros etc. Em outras situações, se conhecemos relações de causa e efeito entre os componentes de um sistema, podemos fazer determinadas intervenções no mesmo para alterar potenciais resultados futuros. Exemplos desse tipo de decisão são: mudança na constituição da ração animal para aumento do ganho de peso ou deposição de gordura (tais decisões envolvem o conhecimento do efeito de determinados componentes da ração, como proteína, carboidrato, fibra e gordura no metabolismo animal), vacinação ou tratamento profilático para evitar doenças, fertilização de pastagens etc. Relações causais entre variáveis são comumente investigadas utilizando-se experimentos randomizados. Tais experimentos controlados são em geral conduzidos em fazendas experimentais, em universidades e instituições de pesquisa. Apesar de muito úteis, eles têm algumas desvantagens, como custo, escopo reduzido, e seus resultados nem sempre são diretamente aplicáveis no campo, em fazendas comerciais (Rosa; Valente, 2013). Nesse contexto, os data lakes proporcionam a possibilidade de se investigar relações de causa e efeito entre as variáveis presentes, utilizando-

do-se técnicas de análise de dados observacionais, muito comuns em estudos epidemiológicos, como técnicas de redes bayesianas (BN, do inglês *Bayesian Networks*) e modelos de equações estruturais (SEM, do inglês *Structural Equation Models*). Esse grupo de técnicas de análise de dados, aqui denominado **análise prescritiva**, propicia a tomada de decisão num contexto de intervenção no sistema, para alteração e resultados (Rosa; Valente, 2013; Bello et al., 2018). Por exemplo, a partir de análise prescritiva, pode-se determinar que tipo de equipamento e instalações, protocolos de manejo reprodutivos e de saúde animal, manutenção de pastos, etc. melhoram a eficiência de produção, retorno econômico, bem-estar animal e outros critérios de performance do sistema produtivo, de maneira que intervenções específicas podem ser feitas para a otimização dos mesmos. Alguns exemplos de aplicação de análise de dados em pecuária de corte são apresentados a seguir.

4 EXEMPLOS DE APLICAÇÃO E PERSPECTIVAS NA PECUÁRIA DE CORTE

Como exemplos de aplicação de análise de dados em pecuária, podemos dividi-los em dois grupos: análise de dados históricos e análise de dados em tempo real. O primeiro grupo refere-se à análise de dados obtidos ao longo do tempo e armazenados por determinado período. A partir da análise dos mesmos, podem-se investigar os fatores que afetam o sistema, de maneira que intervenções específicas podem ser efetuadas no futuro para melhorar o sistema. Além disso, utilizando-se dados históricos podem ser desenvolvidos modelos preditivos para aplicação futura. Num segundo grupo tem-se a análise de dados em tempo real, onde decisões devem ser tomadas de imediato ou em curto prazo. Exemplos desse tipo de aplicação referem-se em geral a dados de sensores, como termômetros, acelerômetros, GPS, imagens etc. A partir de tais dados, decisões relativas ao tratamento dos animais, à inseminação de vacas, irrigação etc. podem ser tomadas.

Como exemplo de dados históricos têm-se as informações de rotina coletadas diariamente nas fazendas, incluindo informações de saúde e desempenho animal, vacinação e tratamento veterinário, bem como informações detalhadas sobre instalações/equipamentos e práticas de manejo. Juntamente com os dados demográficos e climáticos publicamente disponíveis, grandes coleções de dados podem ser geradas, as quais constituem uma valiosa fonte de informação para ajudar a melhorar os sistemas de produção em termos de saúde e eficiência animal e sustentabilidade

da produção (Kamilaris et al., 2017; Liakos et al., 2018; Morota et al., 2018; Rosa, 2021). Como mencionado anteriormente, um primeiro desafio com dados históricos é a junção dos dados de diferentes fontes, os quais incluem diversos formatos e distribuições espaço-temporais.

Um exemplo aparentemente trivial mas muito importante na pecuária de corte foi apresentado por Aiken et al. (2019), os quais desenvolveram um procedimento automático baseado em aprendizado de máquina para combinar dados de fazendas oriundos de diferentes bancos de dados. O algoritmo desenvolvido foi aplicado na integração de dois grandes conjuntos de dados, um compreendendo 44.566 fazendas que utilizaram suplementação alimentar de uma empresa de nutrição animal (DSM) e outro com 32.776 fazendas que abateram os animais em uma empresa frigorífica (JBS). Abordagens determinísticas, estocásticas e de aprendizado de máquina foram utilizadas para classificar pares de fazendas como compatíveis ou não. As técnicas foram avaliadas em termos de qualidade de resultados (acurácia, precisão, sensibilidade e especificidade), bem como eficiência (tempo de execução computacional). Em geral, os métodos de ML superaram as outras estratégias, com as técnicas supervisionadas geralmente superando as não supervisionadas. Ao final do trabalho foi possível extrair de forma automática e muito precisa as 5.204 fazendas que apareciam nos dois bancos de dados, conectando assim o consumo de produtos de nutrição animal de fazendas com os dados de qualidade de carcaça antes isolados em duas bases de dados independentes. O conjunto de dados mesclado continha informações sobre mais de 4 milhões de bovinos de corte terminados, oriundos de 5.204 fazendas presentes em ambos os conjuntos de dados originais. Em um segundo estágio de integração de dados, o conjunto de dados foi ainda enriquecido com variáveis econômicas relacionadas à produção de carne bovina, fertilidade do solo e dados climáticos de cada local de fazenda (Aiken et al., 2019).

O conjunto de dados integrado de Aiken et al. (2019) foi usado por Aiken et al. (2020) para prever a produção futura bem como a qualidade de bovinos de corte usando três metodologias: regressão linear múltipla, floresta aleatória (RF, do inglês *Random Forest*) e ANN. A técnica de validação cruzada foi utilizada para avaliar a capacidade preditiva de cada modelo, bem como a importância de cada variável. Os resultados indicaram que, para a maioria das características, o melhor modelo preditivo foi o RF. Em termos de importância variável, os modelos puderam identificar os principais fatores determinantes de varia-

ções na produção e qualidade (preço de venda de bovinos e milho, bem como consultoria técnica, além de clima e solo) e quantificar sua importância relativa. Por fim, o tempo de execução computacional e o espaço em disco necessários variaram muito entre os algoritmos, com RF e ANN sendo consideravelmente mais exigentes do que os modelos de regressão, exigindo processamento paralelo.

Outro exemplo de integração e análise de big data no contexto da otimização das práticas de gestão foi apresentado por Passafaro et al. (2019). Neste estudo, dados obtidos durante o transporte final dos suínos até o abatedouro foram relacionados com a morte ou perda total de animais. A análise dos dados utilizando técnicas de regressão com modelos mistos aditivos generalizados (Wood et al., 2017) indicou importante efeito das condições climáticas (vento/temperatura/umidade) e motorista (proprietário/contratado).

Na cadeia de bovinos, menos integrada e padronizada que a dos suínos, podemos esperar efeitos significativos das variáveis relacionadas ao transporte como distância, tempo de viagem e condições de estradas, chuvas, temperatura, tipo do caminhão e motorista sobre bem-estar dos animais, qualidade das carcaças e couro. Gomes et al. (2018) compilaram informações relativas ao abate de gado no estado de Mato Grosso do Sul representando 1,74 milhão de animais abatidos em 12 diferentes unidades frigoríficas, distribuídas por todas as regiões do estado, e conseguiram demonstrar relação entre a origem geográfica dos animais; observaram que o nível de contusões foi de 0,48% na mesorregião leste e em todas as demais menor que 0,09%, o que indica grande dependência regional também para a ocorrência de contusões nas carcaças.

Imagens de satélite, dados cadastrais das fazendas e as informações de trânsito dos bovinos relacionando viagens entre fazendas nas etapas anteriores ao abate têm se mostrado fundamentais na identificação de riscos ambientais nas diversas fazendas por onde o animal passou durante as fases de cria, recria e terminação.

Amaral et al. (2021) utilizaram ferramentas de *Business Intelligence* (BI) para análise da qualidade de carcaça dos animais oriundos do programa Proape-Precoce de Mato Grosso do Sul. O software Tableau foi utilizado para elaborar um painel para visualização e análise dos dados de forma dinâmica e pode ser acessado publicamente por meio do link <https://public.tableau.com/app/profile/thais.basso.amaral/viz/PrecoceMS2/Propriedades>.

Este painel vem sendo atualizado e conta atualmente com as informações relacionadas aos abates

dos animais desde 2017, quando iniciou o programa até 2021. As análises podem ser feitas no painel por meio da aplicação de filtros, dependendo do nível de detalhamento que o usuário deseja. É possível filtrar um ou mais municípios, e todos os demais indicadores do painel serão ajustados para mostrar os dados referentes ao município selecionado. Também é possível selecionar somente produtores que possuem um determinado sistema de produção e da mesma forma visualizar os indicadores de qualidade, peso dos animais, número de produtores, entre outros indicadores referentes àquele sistema de produção. Inúmeras possibilidades de análises que podem ser feitas pelo usuário de forma rápida e dinâmica (Figura 5).

A Internet das Coisas, conhecida também como IoT, surgiu como termo para identificar os dispositivos de tecnologias embarcadas e a interação entre eles e a internet como constituída hoje. Conforme citado no projeto Swamp (Kamienski; Visoli, 2018), o uso das tecnologias de informação e comunicação, com esta grande diversidade e quantidade de sensores existentes, já está se disseminando no campo. Segundo pesquisa realizada por Bolfe et al. (2020), IoT está entre as tecnologias digitais emergentes preferidas para serem incorporadas ao portfólio de empresas e prestadores de serviço que atuam no agronegócio brasileiro. Na mesma pesquisa, bem-estar animal e modelos de estimativa de produtividade são assuntos que estão entre os mais atendidos por essas empresas no que tange ao uso de tecnologias digitais. Em Speranza et al. (2022) está descrita a proposta de uma arquitetura baseada em IoT para coleta, armazenamento e processamento de dados oriundos de um sistema de integração lavoura-pecuária-floresta (ILPF), com foco no monitoramento de produtividade e bem-estar animal de bovinos. A coleta automatizada e a devida organização dos dados em tempo real, como pesagens diárias, umidade e temperatura de solo, índices de conforto térmico, por exemplo, criam as condições necessárias para a aplicação de técnicas de previsão de ganho de peso animal e da produtividade, fornecendo ao produtor critérios mais apurados para tomada de decisão em seu negócio.

O controle de doenças e sua propagação tem apresentado bons resultados quando da inclusão de informações relativas ao transporte dos animais. Passafaro et al. (2020) investigaram dados de um sistema integrado de produção suína em que os animais são transportados entre granjas após a conclusão de cada fase de produção. Relatórios de 76.566 remessas entre as granjas foram obtidos e utilizados para caracterizar o fluxo de movimento de suínos foram analisados e indicaram, de maneira geral, que as unidades de de-

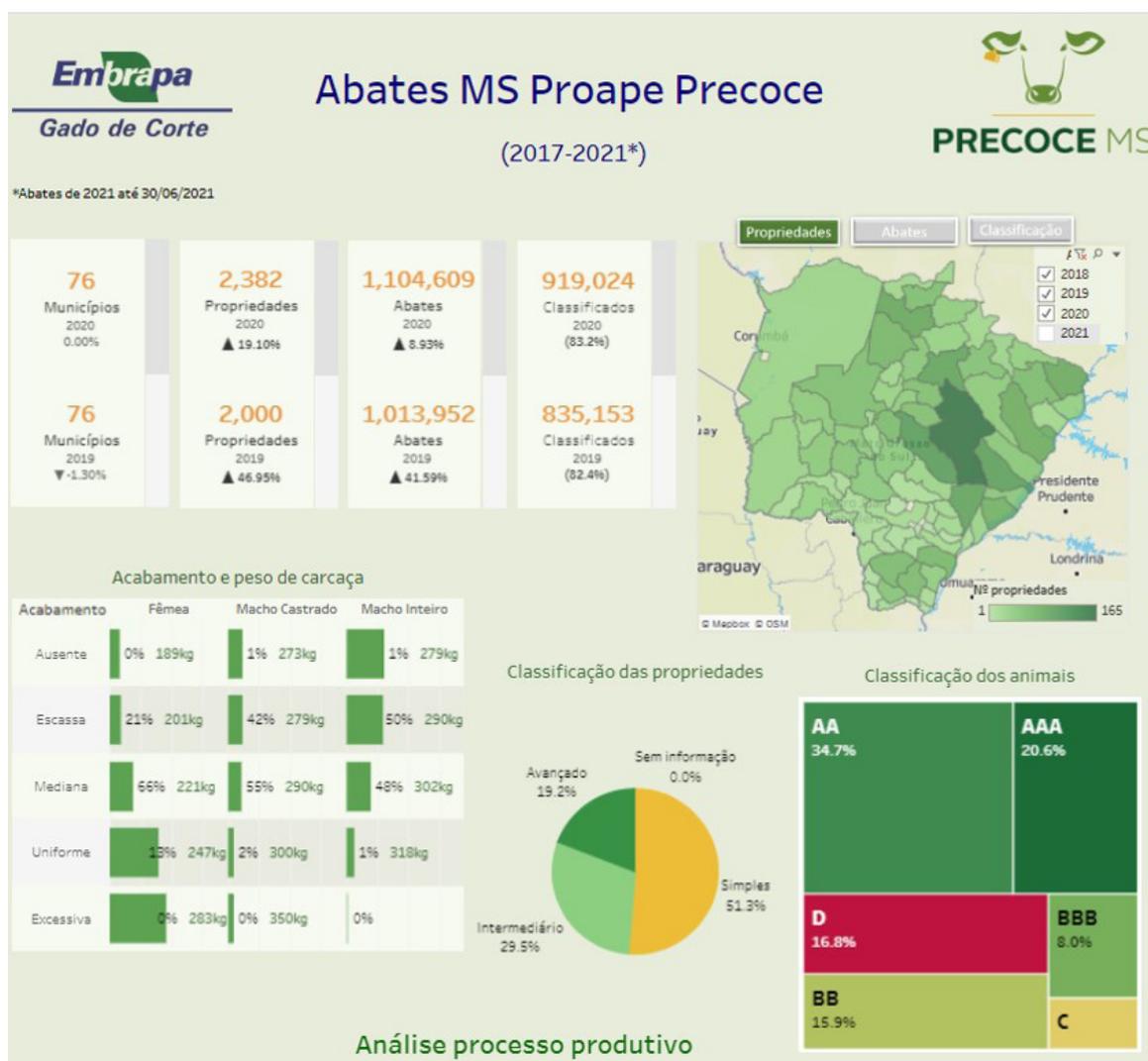


Figura 5. Painel BI dos animais abatidos no programa Proape-Precoce MS nos anos de 2017 a 2021.

envolvimento de leitões de reposição, maternidades e granjas de matrizes são potencialmente os principais fatores de introdução e disseminação de doenças no sistema e que somadas às características de cada doença devem certamente ser foco das políticas de biossegurança para todos os núcleos de produção.

Gomes et al. (2018) compilaram informações relativas ao abate e transporte de gado no estado de Mato Grosso do Sul, representando 1,74 milhão de animais abatidos em 12 diferentes unidades frigoríficas, distribuídas por todas as regiões do estado, e conseguiram demonstrar relação entre a origem geográfica dos animais e o descarte de carcaças pelo SIF devido à ocorrência de contaminação por cisticercose bovina, uma importante zoonose. Considerando a microrregião onde a fazenda de origem do gado estava localizada observou-se incidência média de 0,88% em todo o estado, sendo que nas duas piores microrregiões este índice chegou a 1,81% e na melhor delas apenas

0,23%. Oferecer dados de risco sanitário para cada lote embarcado nas fazendas pode melhorar a precisão das políticas de aquisição de matéria-prima da indústria e de fazendas de terminação.

Para finalizar, a análise de dados de sensores em geral coletados em tempo real tem sido cada vez mais comum com o advento de técnicas de pecuária de precisão, como monitoramento e tomada de decisão em nível individual de cada animal ou de grupos de animais (Berckmans, 2017; Benjamin; Yik, 2019; Rosa, 2021). Dentre as várias tecnologias de sensores, a utilização de análise de imagens e visão computacional tem se mostrada bastante promissora (Fernandes et al., 2020), como por exemplo com imagens de profundidade ou 3D, imagens digitais coloridas (ou RGB, do inglês *Red Green Blue*), imagens termais, ou imagens de infravermelho. Exemplos de aplicações em bovinos podem ser encontrados em Cominotte et al. (2020), com a utilização de imagens 3D para a predição de

peso corporal e ganho de peso de bovinos de corte, e em Ferreira et al. (2022) no reconhecimento de bovinos a partir de imagens 3D.

5 CONSIDERAÇÕES FINAIS

Um aumento na aplicação de big data analytics em ciências animais tem sido observado nos últimos anos. Ao projetar tendências futuras, pode-se otimizar a alocação de recursos, tornando toda a cadeia produtiva mais sustentável. Apesar dos desafios enfrentados na análise de grandes coleções de dados, existe um grande potencial para gerar valor para a cadeia da carne, e nesse sentido produtores, técnicos e a indústria precisam estar preparados para se beneficiarem da digitalização da pecuária e por meio desta agregarem valor ao negócio.

REFERÊNCIAS

- AIKEN, V. C. F.; DÓREA, J. R. R.; ACEDO, J. S.; SOUSA, F. G.; DIAS, F. G.; ROSA, G. J. M. Record linkage for farm-level data analytics: comparison of deterministic, stochastic and machine learning methods. **Computers and Electronics in Agriculture**, v. 163, p. 104857, 2019. DOI: <http://doi.org/10.1016/j.compag.2019.104857>.
- AIKEN, V. C. F.; FERNANDES, A. F. A.; PASSAFARO, T. L.; ACEDO, J. S.; DIAS, F. G.; DÓREA, J. R. R.; ROSA, G. J. D. M. Forecasting beef production and quality using large-scale integrated data from Brazil. **Journal of Animal Science**, v. 98, n. 4, p. 89, 2020. DOI: <http://doi.org/10.1093/jas/skaa089>.
- AMARAL, T. B.; GOMES, R. C.; ROSA, G. J. M. **Produção de novilho precoce no Estado de Mato Grosso do Sul**: análise exploratória de dados do Programa PROAPE-Precoce MS. Brasília: Embrapa, 2021. Comunicado Técnico, 162. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1138385/1/COT-162-Final-em-alta1.pdf>. Acesso em: 30 dez. 2022.
- BELLO, N. M.; FERREIRA, V. C.; GIANOLA, D.; ROSA, G. J. M. Conceptual framework for investigating causal effects from observational data in livestock. **Journal of Animal Science**, v. 96, n. 10, p. 4045-4062, 2018. DOI: <http://doi.org/10.1093/jas/sky277>.
- BENJAMIN, M.; YIK, S. Precision livestock farming in swine welfare: a review for swine practitioners. **Animals**, v. 9, n. 4, p. 133, 2019. DOI: <http://doi.org/10.3390/ani9040133>.
- BERCKMANS, D. General introduction to precision livestock farming. **Animal Frontiers**, v. 7, n. 1, p. 6-11, 2017. DOI: <http://doi.org/10.2527/af.2017.0102>.
- BOLFE, É. L.; JORGE, L. A. C.; SANCHES, I. D. A.; LUCHIARI JÚNIOR, A.; COSTA, C. C.; VICTORIA, D. C.; INAMASU, R. Y.; GREGO, C. R.; FERREIRA, V. R.; RAMIREZ, A. R. Precision and digital agriculture: adoption of technologies and perception of Brazilian farmers. **Agriculture**, v. 10, n. 12, p. 653, 2020. DOI: <http://doi.org/10.3390/agriculture10120653>.
- COMINOTTE, A.; FERNANDES, A. F. A.; DÓREA, J. R. R.; ROSA, G. J. M.; LADEIRA, M. M.; VAN CLEEF, E. H. C. B.; PEREIRA, G. L.; BALDASSINI, W. A.; MACHADO NETO, O. R. Automated computer vision system to predict body weight and average daily gain in beef cattle during growing and finishing phases. **Livestock Science**, v. 232, p. 103904, 2020. DOI: <http://doi.org/10.1016/j.livsci.2019.103904>.
- DAMAINTERNATIONAL. **DAMA-DMBOK: Data Management Body of Knowledge**. 2nd ed. Denville, NJ: Technics Publications, 2017.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. **Visão de futuro do Agro**. Brasília: EMBRAPA, 2022. Disponível em: <https://www.embrapa.br/en/visao-de-futuro>. Acesso em: 30 dez. 2022.
- FERNANDES, A. F. A.; DÓREA, J. R. R.; ROSA, G. J. M. Image analysis and computer vision applications in animal sciences: an overview. **Frontiers in Veterinary Science**, v. 7, p. 551269, 2020. DOI: <http://doi.org/10.3389/fvets.2020.551269>.
- FERREIRA, R. E. P.; BRESOLIN, T.; ROSA, G. J. M.; DÓREA, J. R. R. Using dorsal surface for individual identification of dairy calves through 3D deep learning algorithms. **Computers and Electronics in Agriculture**, v. 201, p. 107272, 2022. DOI: <http://doi.org/10.1016/j.compag.2022.107272>.
- GOMES, R. C.; NICACIO, A. C.; NOGUEIRA, É.; COSTA, F. P.; DIAS, F. R. T.; FEIJÓ, G. L. D.; MENEZES, G. R. O.; SILVA, J. C. B.; OLIVEIRA, L. O. F.; SILVA, L. O. C.; GOMES, M. N. B.; MEDEIROS, S. R.; ABREU, U. G. P. **Novilho precoce**: demandas e caminhos para sua produção e valorização. Campo Grande: Embrapa Gado de Corte, 2018. 48 p. Disponível em: <https://www.infoteca.cnptia.embrapa.br/infoteca/bitstream/doc/1102353/1/Novilhoprecocedemandasecaminhos.pdf>. Acesso em: 20 dez. 2022.
- HAJDARBEGOVIC, M. **Data lake architecture: a comprehensive guide**. Austin: Virtasant, 2020. Disponível em: <https://www.virtasant.com/blog/data-lake-architecture>. Acesso em: 30 dez. 2022.
- KAMIENSKI, C.; VISOLI, M. C. Swamp: uma plataforma para irrigação de precisão baseada na Internet das Coisas. **Fonte**, v. 15, n. 20, p. 76-84, 2018. Disponível em: <http://www.alice.cnptia.embrapa.br/alice/handle/doc/1105001>. Acesso em: 23 nov. 2022.
- KAMILARIS, A.; KARTAKOULLIS, A.; PRENAFETA-BOLDÚ, F. X. A review on the practice of big data analysis in agriculture. **Computers and Electronics in Agriculture**, v. 143, p. 23-37, 2017. DOI: <http://doi.org/10.1016/j.compag.2017.09.037>.
- LIAKOS, K. G.; BUSATO, P.; MOSHOU, D.; PEARSON, S.; BOCHTIS, D. Machine learning in agriculture: a review. **Sensors**, v. 18, n. 8, p. 2674, 2018. DOI: <http://doi.org/10.3390/s18082674>.

- MOROTA, G.; VENTURA, R. V.; SILVA, F. F.; KOYAMA, M.; FERNANDO, S. C. Machine learning and data mining advance predictive big data analysis in precision animal agriculture. **Journal of Animal Science**, v. 96, n. 4, p. 1540-1550, 2018. DOI: <http://doi.org/10.1093/jas/sky014>.
- PASSAFARO, T. L.; FERNANDES, A. F. A.; VALENTE, B. D.; WILLIAMS, N. H.; ROSA, G. J. M. Network analysis of swine movements in a multi-site pig production system in Iowa, USA. **Preventive Veterinary Medicine**, v. 174, p. 104856, 2020. DOI: <http://doi.org/10.1016/j.prevetmed.2019.104856>.
- PASSAFARO, T. L.; VAN DE STROET, D.; BELLO, N. M.; WILLIAMS, N. H.; ROSA, G. J. Generalized additive mixed model on the analysis of total transport losses of market-weight pigs. **Journal of Animal Science**, v. 97, n. 5, p. 2025-2034, 2019.
- ROSA, G. J. M. Grand challenge in precision livestock farming. **Frontiers in Animal Science**, v. 2, p. 650324, 2021. DOI: <http://doi.org/10.3389/fanim.2021.650324>.
- ROSA, G. J. M.; VALENTE, B. D. Inferring causal effects from observational data in livestock. **Journal of Animal Science**, v. 91, n. 2, p. 553-564, 2013. DOI: <http://doi.org/10.2527/jas.2012-5840>.
- SPERANZA, E. A.; HIGA, R. H.; VISOLI, M. C.; MARTINS, C.; CARROMEU, C. Plataforma de internet das coisas para monitoramento de produtividade e bem-estar animal em sistemas de integração lavoura-pecuária-floresta. In: CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO (ConBAP), 2022, Campinas. **Anais [...]**. Campinas: Associação Brasileira de Agricultura de Precisão, 2022. p. 378-385.
- WOOD, S. N.; LI, Z.; SHADDICK, G.; AUGUSTIN, N. H. Generalized additive models for Gigadata: modeling the U.K. black smoke network daily data. **Journal of the American Statistical Association**, v. 112, n. 519, p. 1199-1210, 2017. DOI: <http://doi.org/10.1080/01621459.2016.1195744>.