

VISUALIZAÇÃO E INTEGRAÇÃO DE DADOS ZOOTÉCNICOS VIA LINGUAGEM R E FERRAMENTAS DE BUSINESS INTELLIGENCE

Gustavo Henrique Tito da Luz Anezio da Silva^{1*}

Ricardo Vieira Ventura¹

¹Faculdade de Medicina Veterinária e Zootecnia - Universidade de São Paulo

*gustavo.tito.silva@usp.br

Objetivos

Este estudo teve por objetivo a geração de dados em larga escala, via processo *in-silico* de simulação computacional, e posterior utilização de ferramentas Business Intelligence (BI) como forma de integração de dados zootécnicos provenientes de fontes e formatos distintos. Tal modelo computacional almeja favorecer a interpretação biológica, feita pelo usuário final, por meio de um dashboard interativo.

Métodos e Procedimentos

Estatísticas descritivas sobre ganho de peso, resultantes de análises prévias conduzidas em um rebanho de bovinos de corte canadense, foram compartilhadas com os autores do presente estudo para base de um novo processo de simulação computacional, codificada em linguagem R, via RStudio (RStudio Team, 2020). Rotinas gráficas e interativas para fins de visualização e adequação dos dados simulados, foram desenvolvidas por meio do pacote ggplot2 (H. Wickham, 2016), além de ferramentas BI (PowerBI - Microsoft). Após a etapa preliminar de geração de dados, novas métricas e informações pormenorizadas (que mimetizassem o recebimento de dados a cada ida do animal ao cocho) foram calculadas. Como premissa fundamental, a característica ganho de peso diário foi modelada com base em uma curva Gaussiana, construindo-se então um *dataframe* contendo valores simulados sobre ganho de peso de cada animal para cada dia de observação. Posteriormente, o consumo diário de matéria seca foi

estipulado com base em valores da literatura, de forma que valores de ingestões diárias (em peso bruto da dieta de cada animal) pudessem ser atribuídos para cada indivíduo.

Resultados

Como resultados do processo de simulação, um total de 6005 arquivos foram gerados, contemplando um total de 30031 idas ao cocho, oriundos de 200 animais, o que se assemelha a uma real entrada de dados obtida por meio de um cocho automatizado. Adicionalmente, um modelo integrativo computacional, flexível e capaz de trabalhar com elevado volume de dados oriundos de diversos tipos de arquivos (CSV, TXT, entre outros), foi desenvolvido via ferramenta BI. Tal ferramenta auxilia o melhor entendimento de possíveis contratempos enfrentados em um confinamento de gado de corte. Dados adicionais, tais como origem do animal, preço da arroba do boi (R\$/@) e custo da dieta (R\$/ton), foram inseridos para fins do desenvolvimento de relatórios dinâmicos.

Conclusões

A integração de diferentes dados via ferramentas BI pode elucidar potenciais problemas em operações comerciais de gado de corte, favorecendo-se então futuras tomadas de decisão no gerenciamento do rebanho.

Referências Bibliográficas

H. WICKHAM. ggplot2: Elegant Graphics for Data analysis. Springer-Verlag New York, 2016.

VISUALIZATION AND INTEGRATION OF BIOLOGICAL DATA RELATED TO ANIMAL SCIENCE USING R LANGUAGE AND BUSINESS INTELLIGENCE TOOLS (BI)

Gustavo Henrique Tito da Luz Anezio da Silva^{1*}

Ricardo Vieira Ventura¹

¹Faculdade de Medicina Veterinária e Zootecnia - Universidade de São Paulo

*gustavo.tito.silva@usp.br

Objectives

This study aimed to generate data on a large scale, through an *in-silico* computer simulation process, as well as to adopt Business Intelligence (BI) tools as a potential mechanism of integrating feedlot data from different sources and formats. This computational model aims to enhance the biological interpretation at the user level through the development of interactive dashboard.

Materials and Methods

Descriptive statistics on weight gain, resulting from previous analysis conducted on a Canadian beef cattle herd, shared with the authors of the present study as a base information for the computer simulation process, coded in R language, via RStudio (RStudio Team, 2020). Graphical and interactive routines were developed using the ggplot2 package (H. Wickham, 2016), as well as BI tools (PowerBI - Microsoft), both created for calibrating and checking the simulated data. After the preliminary stage of basic data generation, new metrics and detailed information were calculated (to mimic the process of receiving data for every single time the animal approached the automated feeding system). As a fundamental assumption, the trait average daily gain was modeled based on a Gaussian curve, and a dataframe was built containing simulated values on the weight gain of each animal for each day/moment of collection. Subsequently, the daily consumption of dry matter was defined based on literature

values, in order that daily intakes values could be assigned to each individual.

Results

As a result of the simulation process, a total of 6,005 files were generated, comprising a total of 30,031 visits to the feeding system (meals), considering a total of 200 animals, which is similar to a real data entry obtained through an automated process. In addition, an integrative computational model, flexible and capable of working with a high volume of data from different types of files (CSV, TXT, among others), was developed via the BI tool. Such a tool helps to better understand possible problems faced in a beef cattle feedlot. Additional data, such as the animal origin, the production cost based on the live weight (R\$/arroba); and the final feed cost (R\$ / ton), were included in our visual tool with the goal of developing dynamic reports.

Conclusions

The integration of different data via BI tools can elucidate potential problems in commercial beef cattle operations, which will allow farmers on future decisions.

References

H. WICKHAM. ggplot2: Elegant Graphics for Data analysis. Springer-Verlag New York, 2016.