

Análise de Dados de Vendas de Feiras Livres: Um Estudo com a Polifeira da UFSM

Gabriel Mello Porcher¹, Gustavo Pinto da Silva¹, Daniel Lichtnow¹

¹Colégio Politécnico – Universidade Federal de Santa Maria (UFSM)
Av. Roraima, nº1000, Campus UFSM – 97.105-900 – Santa Maria – RS – Brasil
gabrielporcher95@gmail.com, gustavo.pinto@politecnico.ufsm.br,
dlichtnow@politecnico.ufsm.br

Abstract. *This work presents analysis on sales data from a free fair using Business Intelligence tools. The analyzes were carried out through the elaboration of dashboards and the use of data mining algorithms. Initial results indicate that it is possible to estimate the demand for products, something that can be applied in other free fairs.*

Resumo. *Este trabalho apresenta análises feitas sobre os dados das vendas de uma feira livre utilizando ferramentas de Business Intelligence. As análises foram feitas mediante a elaboração de dashboards e uso de algoritmos de mineração de dados. Os resultados iniciais indicam que é possível estimar a demanda por produtos, algo que pode ser aplicado em outras feiras livres.*

1. Introdução

As feiras livres são espaços de comércio varejista, que ocorrem normalmente ao ar livre e em espaços públicos. Nelas são comercializados especialmente alimentos de pequenos produtores rurais e/ou pequenas agroindústrias. Um exemplo de feira livre é a Polifeira da Universidade Federal de Santa Maria (UFSM), projeto de extensão que tem por objetivo a comercialização direta dos produtos da agricultura familiar de produtores de Santa Maria. Na Polifeira da UFSM é utilizado um sistema para registrar as vendas.

A partir disto, este trabalho tem por objetivo apresentar a análise dos dados das vendas da Polifeira da UFSM, feita com ferramentas de *Business Intelligence*. *Business Intelligence – BI* (Inteligência de Negócios) foi definido pelo Gartner em 1990 como um termo guarda-chuva que inclui aplicações, infraestrutura, ferramentas e práticas que visam propiciar o acesso à informação de forma a facilitar o processo decisório [Watson e Wixon 2007]. As análises foram realizadas em um primeiro momento com o *Power BI* e depois com ferramentas de mineração de dados. Estas análises podem auxiliar a coordenação da Polifeira e os próprios feirantes na previsão da demanda por produtos, além de fornecer subsídios para outras feiras livres, onde é praticamente inexistente o uso de ferramentas de análise de dados.

Inicialmente, são descritos alguns trabalhos onde foram feitos estudos com dados das feiras livres (Seção 2). Após é feito o relato sobre os dados coletados na Polifeira, sua análise, bem como limitações encontradas (Seção 3). Finalmente, são feitas considerações, fruto da análise realizada, sobre aspectos que devem ser levados em conta para análise de dados de feiras livres (Seção 4).

2. Análise de Dados em Feiras Livres

O número de trabalhos voltados para análise de dados de feiras livres que foi identificado é pequeno. Em [Costa et al 2016], por exemplo, o perfil dos consumidores é analisado a partir das respostas dadas a um questionário, sendo os consumidores agrupados, usando *K-means* [MacQueen 1967], a partir de atributos como idade, escolaridade, profissão, sexo, residência, renda, frequência na feira, dentre outros. Alguns trabalhos focam em determinar razões para escolha de produtos a partir do perfil dos clientes [Silva et al 2016]. Já em [Pereira et al 2016] a correlação entre preços de produtos orgânicos em feiras livres e nos supermercados é analisada. Não foram identificados trabalhos onde, a partir dos registros das vendas, fossem realizadas análises que possam fornecer subsídios aos feirantes ou responsáveis pela coordenação das feiras livres.

3. Análise de Dados da Polifeira da UFSM

Os dados analisados correspondem ao período anterior à pandemia (2018-2019), quando a Polifeira da UFSM ocorria duas vezes por semana no Campus da UFSM, sendo constituída por cerca de 20 bancas. Na Polifeira da UFSM o registro das vendas existe pelo fato de que ao aderir, o feirante deve concordar em registrar suas vendas usando um formulário. Assim, ao final de cada dia, cada feirante entrega um boletim de vendas onde consta a quantidade trazida de cada produto, a quantidade vendida e o valor cobrado. Estes boletins são digitados em um sistema que permite a emissão de relatórios básicos. Não é feito o registro individualizado das vendas para cada cliente. Cabe ressaltar que na Polifeira da UFSM os dados das vendas de um feirante não são compartilhados com os demais – apenas o volume global.

Para realizar a análise dos dados, inicialmente foi construída uma série de *dashboards* com a ferramenta *Power BI*. O *Power BI* é uma ferramenta de *Business Intelligence* da *Microsoft*, consistindo de um conjunto de serviços que permite acessar múltiplas fontes de dados bem como criar *dashboards* e realizar seu compartilhamento na Web¹. A construção destes *dashboards* permitiu avaliar o desempenho de vendas (totais e variação ao longo do tempo). A partir disto, para alguns feirantes foi feita a proposta de mudar os produtos ofertados.

No que se refere à variação do volume de vendas de produtos ao longo do tempo, como a Polifeira ocorre dentro do Campus da UFSM, é esperado que nos meses em que não estão ocorrendo aulas o volume de vendas caia e tem se observado alterações na demanda de alguns produtos, dado que o perfil do frequentador muda. Existem ainda produtos que são mais consumidos em determinadas épocas ou condições meteorológicas e para estes estimar o número de vendas reduz o risco da oferta ser muito maior que a demanda. Trazer para a feira produtos que não serão vendidos é um problema, dado que muitos produtos são perecíveis.

A Figura 1 ilustra a variação das vendas com o caso do agnolini (capeletti), que conforme esperado, possui uma demanda maior no período de inverno. Já no caso do caldo de cana (Figura 2), nos meses com temperatura mais elevada ocorre um aumento no consumo. Cabe ressaltar que, em alguns casos, a temperatura pode não ser a esperada

¹ Disponível em: <https://powerbi.microsoft.com/pt-br/> Acesso em 10/06/2021

para um determinado período (temperaturas mais elevadas em julho, por exemplo) ou as condições meteorológicas podem não favorecer o comparecimento na feira.

Neste sentido, foram recuperados dados do CPTEC – Centro de Previsão de Tempo e Estudos Climáticos do INPE que indicam de hora em hora as condições do tempo na estação da Base Aérea de Santa Maria (vizinha ao Campus da UFSM). Estes dados podem ser recuperados via requisição a um *Web Service*². Como padrão, optou-se por considerar os dados gerados às 14:00, uma vez que a Polifeira normalmente ocorria na parte da tarde. Os dados recuperados foram armazenados em uma base de dados.

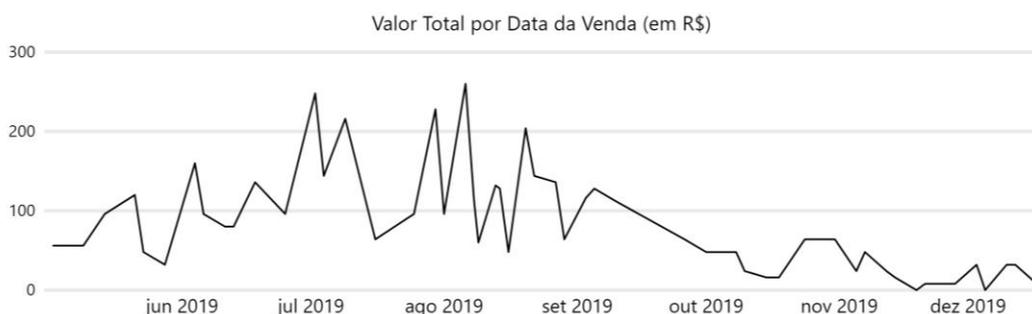


Figura 1. Variação das vendas de agnolini – elaborado com Power BI

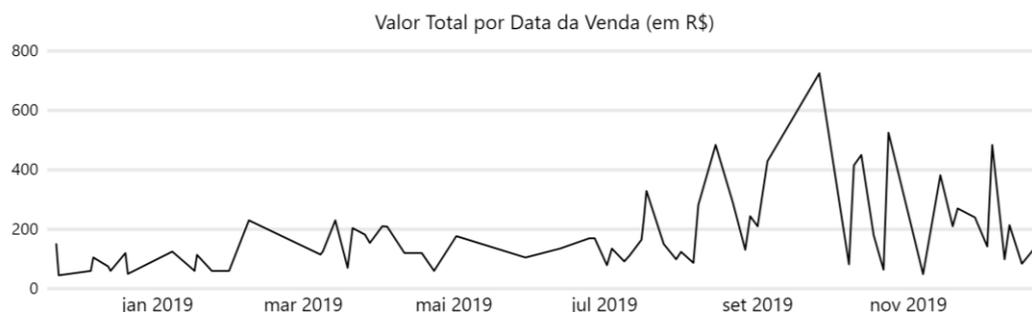


Figura 2. Variação das vendas de caldo de cana – elaborado com Power BI

A coleta dos dados sobre o clima e a variação de vendas ao longo do tempo motivou o uso de algoritmos de mineração de dados. O ainda pequeno volume de dados coletados (Tabela 1), limitou o processo de análise de dados. Além disto, qualquer análise feita a partir de 2020 é prejudicada em função da crise sanitária.

Tabela 1. Números da base de dados (entre 01/11/2018 e 31/12/2019)

Número de dias de feira	162
Número de Boletins	1.714
Número de registros existentes nos boletins (em média 10,52 produtos por boletim)	18.025

² Disponível em: <http://servicos.cptec.inpe.br/XML/> Acesso em: 13/09/2020

Uma análise foi feita usando o algoritmo *Apriori* [Agrawal et al., 1994] disponível no *WEKA* [Frank et al 2016]. O objetivo era estimar o volume de vendas em função de alguns atributos. Inicialmente os atributos considerados foram: produto, tipo de produto, temperatura, condição meteorológica, trimestre, mês, dia, valor das vendas (valores do total de vendas do dia foram discretizados, usando o *WEKA*, em 4 faixas de valores). Inicialmente, a granularidade considerada era o total de vendas por produto e depois por tipo de produto. Como mesmo com suporte de 1% e confiança de 50% não foram encontradas regras relevantes, foi feita análise do total de vendas por dia considerando todos os produtos. Nesta análise inicial, descartadas regras irrelevantes como *Neblina->Julho*, foi possível constatar/comprovar que nos dias em que a condição do tempo era desfavorável, conforme dados recuperados do CPTEC/INPE, as vendas ficaram dentro da menor faixa de valor (até R\$ 2.012,00). A regra relacionada a este fato tem um grau de confiança de 100%, um *Lift* de 3,92 e um suporte de 2%. A redução de vendas sob condições meteorológicas adversas é algo esperado. Mas esta análise pode ser aprofundada de forma a determinar a demanda que um determinado tipo de produto terá, levando em conta a previsão do tempo.

4. Considerações Finais

Este trabalho descreveu o uso de ferramentas de *Business Intelligence* na análise de dados de uma feira livre. Embora preliminares, os resultados das análises permitem vislumbrar a possibilidade de auxiliar os feirantes na definição da oferta de produtos.

Agradecimentos. Trabalho apoiado pelo programa de bolsas de ensino, de pesquisa e de extensão do Colégio Politécnico da UFSM.

Referências

- Agrawal, Rakesh et al. (1994) Fast algorithms for mining association rules. In: Proc. 20th int. conf. Very Large Data Bases, VLDB. p. 487-499.
- Chaudhuri, Surajit; Dayal, Umeshwar; Narasayya, Vivek. An overview of business intelligence technology. *Communications of the ACM*, v. 54, n. 8, p. 88-98, 2011.
- Costa, D. M. D., et al (2016). Análise do comportamento dos consumidores de feiras livres na microrregião de Formiga (MG). *Revista Agrogeoambiental*, 8(4).
- Frank, Eibe; et al (2016) The WEKA Workbench. Online Appendix for Data Mining: Practical Machine Learning Tools and Techniques, Morgan Kaufmann, 4th Ed.
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proc. of the fifth Berkeley symposium on mathematical statistics and probability v. 1, No. 14, pp. 281-297
- Pereira, B. D., et al (2016). Análise de correlação dos preços de produtos orgânicos no varejo supermercadista e feira livre. *Revista ESPACIOS* | v. 37 n 13.
- Silva, F. J. F. D., et al (2016). Compra do pescado na feira de Juruá: fatores que influenciam na tomada de decisão. *Revista Igapó Revista de Educação Ciência e Tecnologia do IFAM*, 10(1), 16-24.
- Watson, Hugh J.; Wixon, Barbara H. The current state of business intelligence. *Computer*, v. 40, n. 9, p. 96-99, 2007.