

Uso de Variáveis de Calendário na Previsão de Séries Temporais: Conceito e Aplicação

Felipe Freitas Viveiros Ribeiro, Gabriel Alvaro Batista

Introdução

A previsão de **séries temporais** é um importante auxiliar na tomada de decisões no mercado, na indústria ou nas políticas públicas. Analisar e modelar padrões temporais permite entender e prever acontecimentos futuros baseando-se em **dados históricos**.

Aliando dados passados de uma série à **variáveis preditoras**, é possível criar previsões acuradas e assertivas, utilizando a informação temporal de outras fontes para explicar os dados de interesse. Dentre as variáveis que podem ser utilizadas nas séries temporais, as **variáveis de calendário** destacam-se por suas características intrínsecas: são variáveis **fixas, determinísticas**, amplamente **conhecidas** e sempre disponíveis para uso, sem a necessidade de tratamentos e transformações complexas.

A influência de **feriados, fins de semana**, ou mesmo **períodos sazonais pré-definidos** não têm incerteza associada e são consistentes ao longo do tempo. Isso permite com que tenhamos um grau de controle maior sobre os parâmetros, além de trazer **interpretabilidade** para os resultados do modelo e os coeficientes estimados.

Metodologia

O objeto de estudo foi uma série referente ao **M1**, um **agregado monetário** de alta liquidez que inclui o papel-moeda em circulação e os depósitos à vista em bancos. A hipótese central é que esta série está correlacionada com as variáveis de calendário, refletindo o **padrão de consumo** da população, como maior demanda e circulação em datas próximas a dias de pagamento, vésperas de feriados e início do mês, e menor circulação em feriados, final de mês e domingos, por exemplo. Esse dado é público e divulgado pelo Banco Central do Brasil (BCB).

Para a modelagem, foi utilizado um **modelo linear dinâmico (DLM)** a fim de obtermos resultados **interpretáveis** e maior **explicabilidade** ao validar os coeficientes estimados.

Agregado Monetário M1

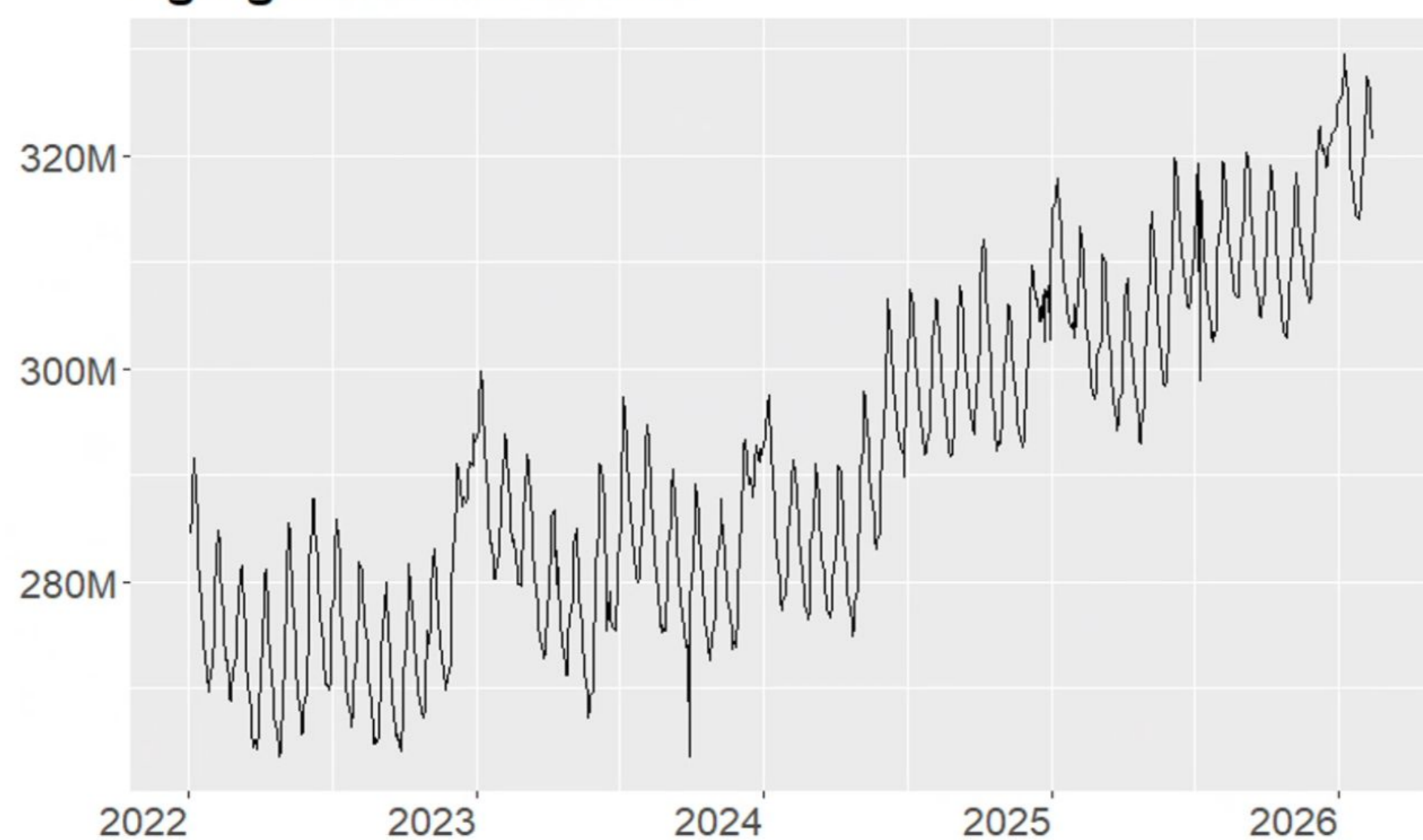


Figura 1: Série temporal representando o agregado monetário M1, em bilhões, de janeiro/2022 a janeiro/2026.

Como entrada do modelo, temos uma tabela com datas indo de **01/01/2022 até 31/01/2026**, os **valores diários** correspondentes ao agregado monetário (variável resposta), a **sazonalidade** da série decomposta em séries de Fourier (senos e cossenos mensais e anuais), **dummies** referentes a cada uma das variáveis de calendário elencadas, que incluem feriados, datas de véspera e pós feriado, emendas, demarcações de períodos, dentre outras variáveis relacionadas. Todas as variáveis consideradas no calendário social encontram-se na tabela abaixo.

Réveillon	Ano Novo	Pós Ano Novo	Pré Carnaval Quarta	Pré Carnaval Quinta
Pré Carnaval Sexta	Pré Carnaval Sábado	Pré Carnaval Domingo	Carnaval Segunda	Carnaval Terça
Carnaval Quarta de Cinzas	Véspera Feriado Nacional	Feriado Nacional	Pós Feriado Nacional	Emenda Feriado Nacional
Duas Semanas Antes Natal	Uma Semana Antes Natal	Ante Véspera Natal	Véspera Natal	Pós Natal
Natal	Dezembro 27	Dezembro 28	Dezembro 29	Dezembro 30

Resultados

O modelo capturou um aumento de aproximadamente **43%** no M1 nos dias de véspera e **39%** nas emendas de feriados nacionais. Já no dia posterior ao feriado houve um aumento mas em magnitude bem menor, em torno de **2%**, que pode ser visto na Figura 2.

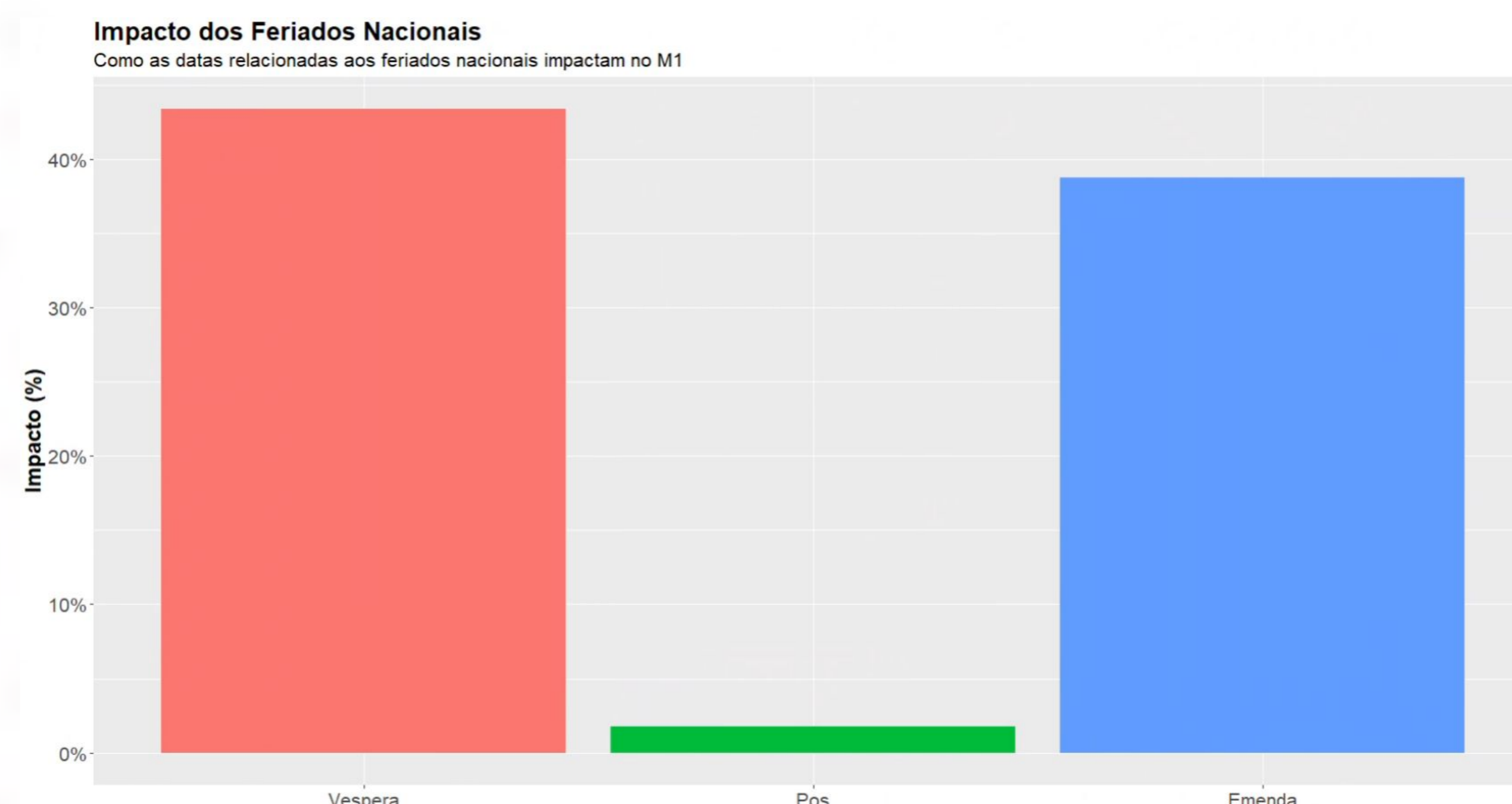


Figura 2: Impacto das variáveis relacionadas aos feriados sobre a variável resposta (M1).

Já para o Natal e períodos relacionados, o modelo encontrou um padrão típico em vários mercados de varejo no Brasil, como mostrado na Figura 3. Um **aumento gradativo** na medida em que o Natal se aproxima, seguido de **forte queda** nos dias anteriores, e retomada no dia posterior.

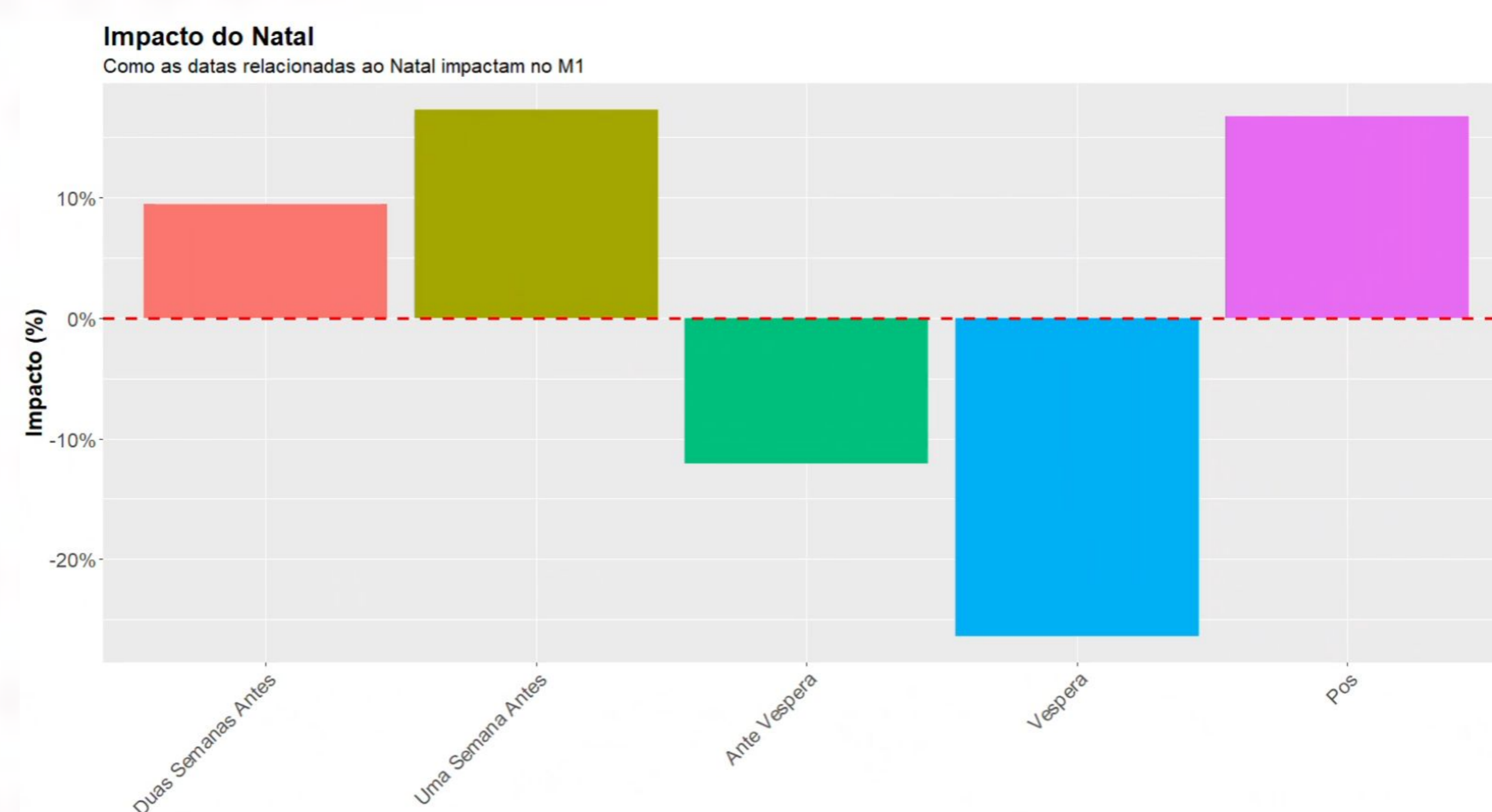


Figura 3: Impacto das variáveis relacionadas ao Natal sobre a variável resposta (M1).

Conclusões

O emprego do **calendário social**, isto é, feriados e datas comemorativas, através de variáveis *dummies* como regressoras, permite controlar a data de interesse, bem como medir o impacto dos dias anteriores e posteriores ao evento.

Este tipo de variável possui boa aplicabilidade como variável preditiva, uma vez que os eventos por ela descritos são bem conhecidos, o que reduz sobremaneira a incerteza contida na modelagem, vis-à-vis, outros tipos de variáveis preditivas, que possuem trajetória futura desconhecida.

Pelos coeficientes estimados e resultados obtidos, podemos afirmar que essas variáveis trazem um **impacto significativo** sobre a variável resposta, reforçando a importância e o potencial das variáveis de calendário na previsão de séries temporais, sobretudo séries fortemente relacionadas à **movimentações financeiras** e **padrões de consumo**.

Bibliografia

- West, M., and P. J. Harrison. 1986. "Monitoring and Adaptation in Bayesian Forecasting Models." *Journal of the American Statistical Association* 81(395): 741-750.
- E. G. Pinheiro and A. F. B. Menezes, "Pybats-detection: A python package for outlier and structural changes detection in time series analysis", 2016.