



UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES  
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO

IGOR ANTUN DA COSTA GAGO

**Os efeitos do agendamento midiático nas redes sociais: um estudo de caso no  
contexto das eleições de 2022 no Brasil**

São Paulo

2023

IGOR ANTUN DA COSTA GAGO

**Os efeitos do agendamento midiático nas redes sociais: um estudo de caso no  
contexto das eleições de 2022 no Brasil**

Monografia apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, como parte dos requisitos exigidos na disciplina ACH 2018 - Projeto Supervisionado ou de Graduação II, para obtenção do título de Bacharelado em Sistemas de Informação.

Modalidade: TCC longo (1 ano) - individual

Orientador: Prof. Dr. Marcio Moretto Ribeiro

São Paulo

2023

## Resumo

GAGO, Igor Antun da Costa. **Os efeitos do agendamento midiático nas redes sociais**: um estudo de caso no contexto das eleições de 2022 no Brasil. 2023. 22 f. Monografia (Bacharelado em Sistemas de Informação) – Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, São Paulo, 2022.

A hipótese do agendamento propõe que grandes veículos midiáticos, apesar de não influenciarem diretamente *o que* as pessoas pensam sobre determinadas questões, possuem o grande poder de determinar *quais* serão as questões de destaque na sociedade em um dado momento. No entanto, com a adoção em massa das mídias sociais, o público é exposto a uma variedade maior de fontes de notícia e possui influência direta na relevância que é dada a um tema. O presente estudo buscou entender, através da aplicação de técnicas automatizadas de raspagem de dados, a correlação entre a importância dada pelos jornais à uma notícia e a que lhe é dada nas redes sociais (Facebook e Twitter), no contexto das eleições presidenciais de 2022 no Brasil. Dentre os 10 cenários analisados, apenas 1 apontou para uma possível correlação entre estas. Isto é, dentro dos dados obtidos, verificou-se que na maior parte dos casos tal correlação é inexistente.

Palavras-chaves: Teoria do agendamento; Mídia; Redes sociais; Agenda pública; Eleições;

## Abstract

GAGO, Igor Antun da Costa. **The effects of media agenda-setting on social networks**: A case study in the context of 2022 Brazilian general elections. 2023. 22 p. Monograph (Bachelor of Information Systems) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2022.

The agenda-setting theory suggests that media institutions — despite being unable to directly influence *what* people think about a set of issues — have the great power of determining *which* issues are most important to society in a given moment. With the mass adoption of social networks, though, the public is exposed to a greater amount of news sources and have direct influence in the relevance given to a topic. The present study aimed to understand, through the use of automatic data scraping techniques, the correlation between the relevance given by newspapers to an article and the one given to it in the social networks (Facebook and Twitter), in the context of 2022 Brazilian general elections. Out of 10 analyzed scenarios, only 1 pointed towards a possible correlation between those. That is, given the obtained data, it was concluded that said correlation is inexistent in the vast majority of scenarios.

Keywords: Agenda-setting; Media; Social networks; Public agenda; Elections;

## Lista de figuras

- Figura 1 – Correlação entre saliência nos jornais online e interações no Facebook . 17
- Figura 2 – Correlação entre saliência nos jornais online e interações no Twitter . . 19

## Lista de tabelas

Tabela 1 – Correlação de Spearman - Facebook . . . . .	18
Tabela 2 – Correlação de Spearman - Twitter . . . . .	20

## Lista de abreviaturas e siglas

API	Interface de programação de aplicações, em inglês
-----	---

## Lista de símbolos

$\rho$  Coeficiente de correlação de Spearman



## Sumário

<b>1</b>	<b>Introdução</b>	9
1.1	<i>Justificativa</i>	9
1.2	<i>Proposição</i>	10
1.3	<i>Objetivos</i>	10
1.3.1	Objetivo geral	10
1.3.2	Objetivos específicos	10
1.4	<i>Método</i>	11
1.5	<i>Limitações</i>	11
<b>2</b>	<b>Método</b>	13
2.1	<i>Raspagem de dados dos jornais</i>	13
2.2	<i>Raspagem de dados das redes sociais</i>	14
2.2.1	Extração de dados do Twitter	14
2.2.2	Extração de dados do Facebook	14
2.3	<i>Tratamento e análise dos dados</i>	14
2.3.1	Tratamento dos dados	15
2.3.2	Análise dos dados	15
<b>3</b>	<b>Resultados</b>	16
3.1	<i>Facebook</i>	17
3.2	<i>Twitter</i>	19
<b>4</b>	<b>Conclusão</b>	21
	<b>REFERÊNCIAS</b>	22

## 1 Introdução

A hipótese do agendamento (*agenda-setting*) propõe que grandes veículos midiáticos, apesar de não influenciarem diretamente *o que* as pessoas pensam sobre determinadas questões, possuem o grande poder de determinar *quais* serão as questões de destaque na sociedade em um determinado momento, em detrimento de outras (MCCOMBS, 2005).

Historicamente o consumo de notícias se deu majoritariamente de forma passiva, através de meios como o jornal impresso, o rádio ou a televisão (AHLERS, 2006). Nesse contexto, as priorizações feitas pela equipe editorial de um jornal possuem efeitos diretos na disseminação de informação e percepção de relevância de uma pauta. Com a adoção em massa das mídias sociais, no entanto, o público geral é exposto a uma variedade crescente de fontes de notícias (STERRETT *et al.*, 2019). Sejam essas legítimas ou falsas, de jornais tradicionais ou de páginas pessoais, todas competem pelos mesmos espaços de destaque.

A decisão editorial, agora descentralizada, é diretamente influenciada pelo engajamento dado por usuários nas redes. A relevância que de outra forma seria atribuída pela decisão de poucas pessoas, agora é ditada pelas interações de inúmeras pessoas. Essa configuração pode gerar um descasamento entre o que é destacado pela mídia e o que é priorizado no debate público, como previamente observado nas eleições brasileiras de 2010 por Nunomura (2013).

Este trabalho busca entender, através de um estudo longitudinal e da aplicação de técnicas automatizadas de raspagem de dados, a correlação entre a importância dada pela agenda midiática em portais jornalísticos e a que é atribuída pela agenda pública nas redes sociais, no contexto das eleições presidenciais de 2022 no Brasil.

### 1.1 Justificativa

Ao ditar quais serão as questões de maior relevância para o debate público, um agente efetivamente limita as opções de temas que serão tratados na sociedade e em suas políticas públicas (MCCOMBS, 2005). Estudar o impacto do agendamento nas plataformas digitais sociais e midiáticas, portanto, é uma importante ferramenta para entendermos sua influência no contexto da mudança social.

Com o crescente aumento do uso de redes sociais e de *fake news* com intuito político, cresce também a preocupação acerca de tentativas de manipulação da opinião pública (ZHURAVSKAYA; PETROVA; ENIKOLOPOV, 2020). Há, então, a oportunidade de reavaliar estudos como o de Nunomura (2013) com um novo olhar, trazendo-os para o período eleitoral do Brasil de 2022.

## 1.2 Proposição

O presente estudo de caso possui a intenção de analisar os efeitos do agendamento midiático no contexto das eleições presidenciais no Brasil em 2022. Através da observação conjunta e sistemática do destaque dado a uma publicação nos cinco principais portais jornalísticos do país (*Estadão*, *Folha de S. Paulo*, *g1*, *UOL* e *Veja*) e o engajamento obtido por estas nas redes sociais Facebook e Twitter, pretende-se mensurar e analisar a eventual correlação entre suas agendas.

Ainda que uma correlação não implique relação de causalidade, seu resultado pode dar pistas de como as agendas públicas e midiáticas interagem entre si. Trabalhos futuros podem explorar a análise dos resultados aqui obtidos em conjunto com novos dados e observações, de modo a identificar outros fatores que potencialmente influenciem a relação entre essas.

## 1.3 Objetivos

### 1.3.1 Objetivo geral

Mensurar e analisar a correlação entre o destaque atribuído à uma matéria em portais jornalísticos (agenda midiática) e o engajamento obtido por ela nas redes sociais Facebook e Twitter (agenda pública).

### 1.3.2 Objetivos específicos

- Desenvolver uma aplicação para execução automatizada da raspagem de dados e notícias de *sites* de jornais. Disponibilizar publicamente seu código-fonte para que possa ser utilizada em outros contextos.

## 1.4 Método

O presente trabalho trata de uma pesquisa aplicada e exploratória, realizada através de estudo de caso empírico. Seus cálculos e conclusões foram alcançados através da observação direta e da análise quantitativa de indicadores de saliência de temas em jornais digitais e redes sociais durante o período eleitoral brasileiro de 2022, na intenção de estudar a correlação entre as agendas dessas plataformas.

A fim de comparar a saliência dada a um determinado tema nos portais jornalísticos e nas redes sociais, é necessário primeiro obter os dados brutos necessários e tratá-los de forma a possibilitar a análise. Cada meio aqui analisado necessitou de uma forma distinta de obtenção de dados.

Para o monitoramento dos portais jornalísticos, foi desenvolvido um *script* automatizado que aplica técnicas de raspagem de dados objetivando a guardar o histórico das matérias publicadas na página inicial dos jornais analisados.

Na rede social Twitter foi utilizada sua API de busca histórica, a fim de encontrar publicações com referência direta ao *link* das matérias identificadas pelo *script* anterior e armazenar métricas acerca de seu engajamento. As métricas do Facebook, por sua vez, foram obtidas com o uso da ferramenta CrowdTangle, que oferece acesso ao arquivo histórico de publicações do Facebook e também a seus dados atualizados.

Por fim, com todas os dados necessários obtidos dos portais jornalísticos, do Facebook e do Twitter, foi realizada a análise de correlação das métricas de saliência. Para isso, foi escrito um *script* na linguagem R.

## 1.5 Limitações

- A observação automatizada dos portais jornalísticos compreendeu o período de 16 de outubro à 13 de novembro de 2022. Foram observadas as duas semanas anteriores e posteriores ao segundo turno do período eleitoral.
- Para mensurar o engajamento obtido por uma matéria jornalística nas redes sociais, foi realizada uma busca em cada rede (Facebook e Twitter) com base no *link* da matéria do portal. Ou seja, o cálculo de engajamento se restringe às publicações que continham em seu conteúdo o endereço específico da matéria em questão.

- Os dados de engajamento foram obtidos e calculados após a janela de observação dos jornais, de forma pontual. Isto é, foram coletados os números totais e acumulados de interações durante todo o período, uma única vez, sem dimensão temporal.

## 2 Método

### 2.1 Raspagem de dados dos jornais

O *script* desenvolvido foi configurado de forma a ser executado em intervalos regulares de 1 hora (podendo ter o intervalo facilmente alterado). A cada execução, são acessados os cinco principais portais jornalísticos do país: *Estadão*, *Folha de S. Paulo*, *g1*, *UOL* e *Veja*. Uma vez acessada a página inicial do jornal, são coletados dados acerca das 5 principais publicações.

Ao fim de cada execução, são exportados relatórios em 4 formatos diferentes:

- *.json*: Relatório principal, contendo o resultado do processamento da raspagem de dados. Inclui, para cada publicação encontrada, seu *link*, título, descrição, chapéu (uma ou duas palavras usadas para definir o assunto da matéria, geralmente acima da manchete), espaço ocupado na tela em *pixels*, coordenadas na tela e posição hierárquica entre as demais matérias.
- *.html*: Página original na íntegra, acessível localmente com imagens, *links*, menus e botões em sua maioria funcionais. Especialmente importante como forma de *backup*, uma vez que caso alguma falha ocorra na raspagem de dados é possível revisitar o conteúdo original da página e analisá-lo novamente. Sem isso, uma falha no *script* de raspagem poderia ocasionar perda de dados importantes, dada a constante rotatividade de conteúdo dos jornais.
- *.pdf*: Página exportada em formato de documento, com imagens, texto selecionável, rolagem de conteúdo e *links* funcionais.
- *.webp*: Captura de tela da página no momento de acesso, limitada ao tamanho da janela de exibição configurada

Finalmente, uma vez gerados os relatórios, os mesmos são armazenados para posterior análise.

## 2.2 *Raspagem de dados das redes sociais*

Para obter as métricas de engajamento das matérias jornalísticas dentro das redes sociais, técnicas diferentes precisaram ser empregadas. Em cada caso foi necessário levar em consideração as limitações das APIs disponíveis e a diferença entre os dados disponibilizados.

### 2.2.1 Extração de dados do Twitter

Na rede social Twitter foi utilizada sua API pública, que permite a busca de publicações específicas e a quantidade de interações obtidas por estas. Seu método de busca foi utilizado para encontrar todas as publicações (*tweets*) contendo referência direta ao *link* das matérias identificadas pelo *script* supracitado. As métricas de engajamento, como “curtidas”, respostas e republicações (*retweets*) são, por fim, armazenados.

### 2.2.2 Extração de dados do Facebook

Os dados do Facebook, por sua vez, foram obtidos com o uso da API da plataforma CrowdTangle, que oferece acesso ao arquivo histórico de publicações do Facebook e também a seus dados atualizados. Através da mesma, também é possível buscar publicações que contenham um *link* específico, e, assim como com o Twitter, armazenar suas métricas de engajamento: reações (curtidas e outras), comentários e compartilhamentos.

## 2.3 *Tratamento e análise dos dados*

Ao final do período de obtenção de dados foram geradas duas bases separadas de dados: uma para a rede social Facebook, e outra para o Twitter. Cada uma dessas listas contendo: endereço da matéria, portal jornalístico, saliência mensurada no portal, saliência mensurada na rede social.

### 2.3.1 Tratamento dos dados

Em ambas as redes sociais, observou-se um número significativo de casos em que publicações contendo o *link* da matéria não foram encontradas, ou em que não houve nenhum engajamento. Para que uma análise mais clara pudesse ser feita, optou-se por remover das bases de dados as matérias cuja saliência mensurada nas redes sociais fosse nula.

### 2.3.2 Análise dos dados

Em razão das especificidades das diferentes redes sociais e jornais analisados, os dados obtidos apresentaram grandes variações em magnitude. Isto é, o valor absoluto da saliência calculada para uma dada matéria varia significativamente a depender do jornal ou rede social em questão.

Dado isso, optou-se por utilizar o método de Spearman para identificar eventuais correlações entre as aparições de uma notícia nos portais jornalísticos e o engajamento obtido pela mesma nas redes sociais. A correlação de Spearman utiliza a posição ordenada dos itens (*ranking*), ao invés de seus valores absolutos. Dessa forma foi possível focar na proporcionalidade das saliências, e não na relação exata entre seus valores numéricos.



### 3 Resultados

Nessa seção exploraremos os dados gerados pela pesquisa e entraremos nos detalhes dos resultados obtidos com suas análises. Em todos os casos, empregou-se a mesma fórmula de correlação e portanto os mesmos critérios de interpretação. Para o método de Spearman, isso significa:

- Analisar o coeficiente de correlação ( $\rho$ ). Um valor 0 aponta inexistência de correlação, enquanto -1 ou +1 equivalem à uma correlação perfeita negativa ou positiva.
- Considerar a probabilidade de significância (p-valor). Este varia entre 0 e 1 e indica a probabilidade da correlação observada se dar ao acaso. Um valor próximo a 1 sugere que não há correlação fora o acaso, e portanto valida a hipótese nula. Um valor próximo a 0 (nesse caso, abaixo de 0.05), no entanto, rejeita a hipótese nula e reforça a evidência obtida pelo coeficiente de correlação.

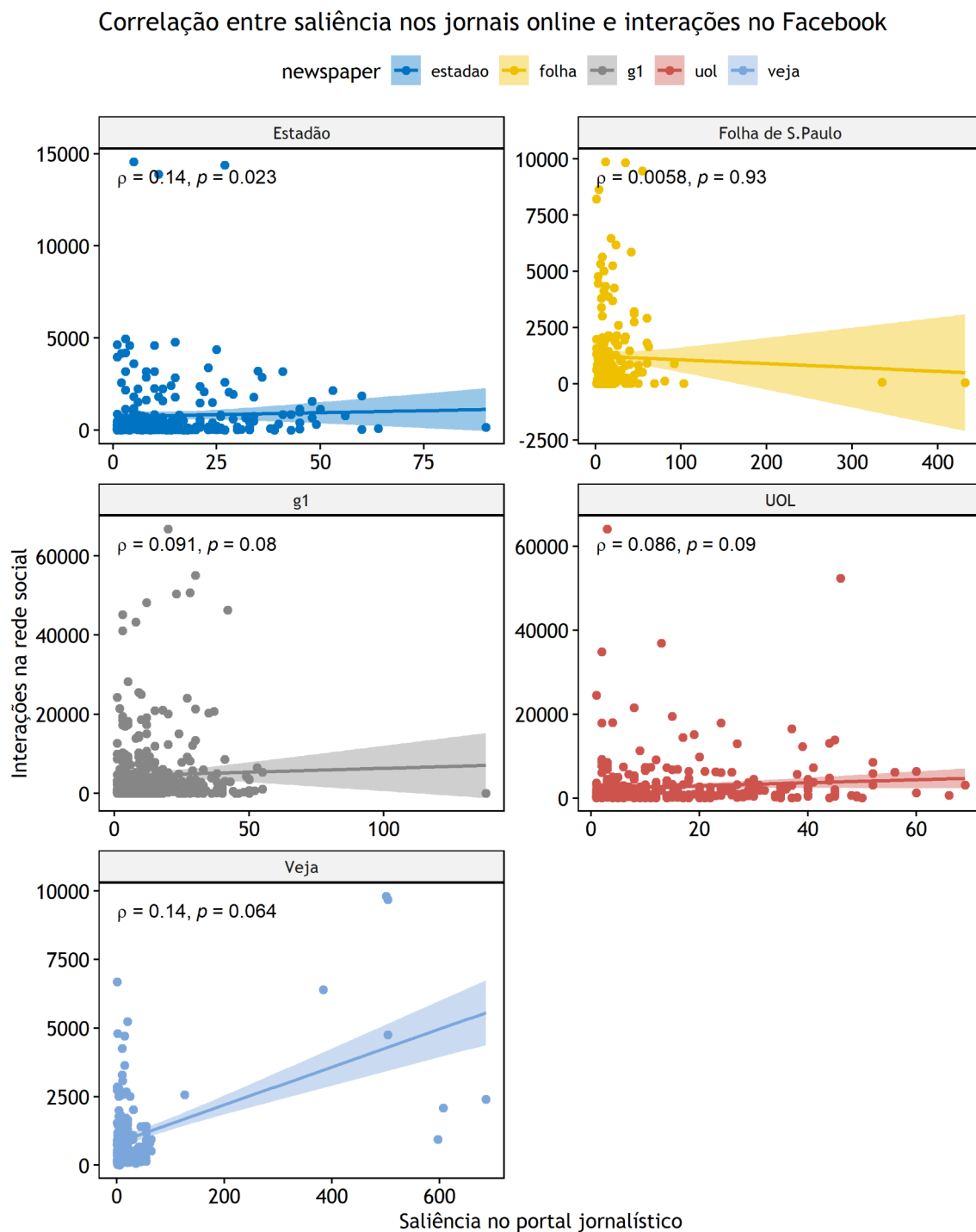
Coeficientes de correlação individuais foram calculados para cada jornal analisado, dentro de cada rede social. Dessa forma foi possível isolá-los e compreender também suas diferenças quanto ao uso das plataformas sociais e o engajamento que possuem em cada. Estes estão apresentados abaixo em forma de gráficos independentes, e também em tabelas equivalentes.

Nos gráficos, cada ponto representa uma matéria jornalística. O eixo horizontal representa a saliência conferida a ela no portal jornalístico, considerando seu tempo e hierarquia de exibição. O eixo vertical, por sua vez, indica o número total de interações acumuladas na rede social em questão por publicações que contenham o endereço da mesma matéria. Há também, em cada gráfico, a regressão linear advinda de sua correlação, junto a seu intervalo de confiança.

Como poderá ser observado a seguir, os dados não apresentam tendência clara e em muitos casos estão espalhados de forma aparentemente desordenada. Isso pode, em certa medida, ser atribuído a algumas limitações do estudo como a janela de tempo observada e dificuldades na busca de publicações. No entanto, é possível que haja relação também com a forma como os jornais utilizam as redes sociais e a presença que têm nas mesmas: frequência de publicação, número de seguidores, etc.

## 3.1 Facebook

Figura 1 – Correlação entre saliência nos jornais online e interações no Facebook



Fonte – Igor Antun, 2023

Tabela 1 – Correlação de Spearman - Facebook

Jornal	$\rho$	p-valor
Estadão	0.14	0.023
Folha de S.Paulo	0.0058	0.93
g1	0.091	0.08
UOL	0.086	0.09
Veja	0.14	0.064

Fonte – Igor Antun, 2023

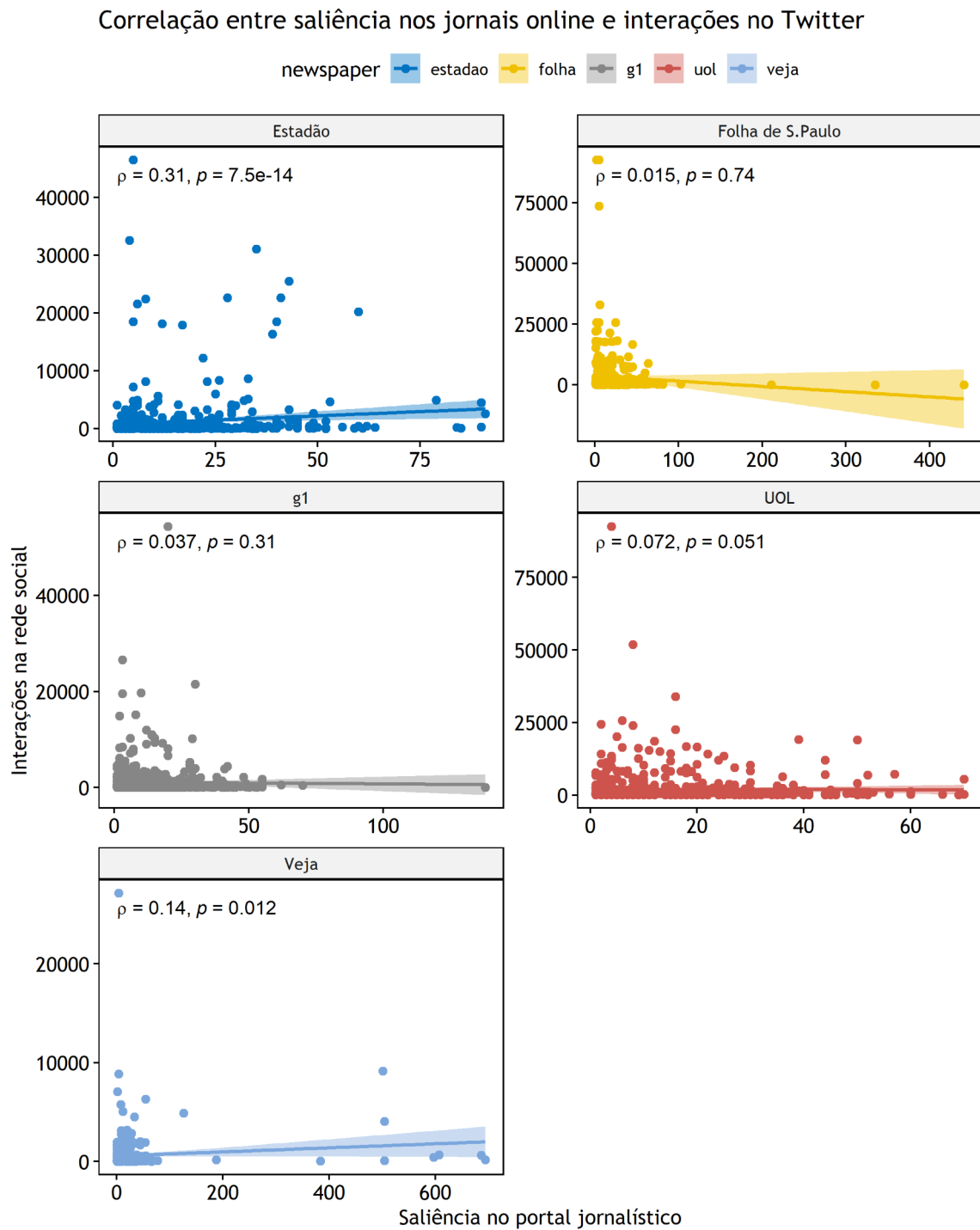
Aqui podemos observar que somente o Estadão e a Veja obtiveram coeficientes de correlação acima de 0.1, o que no entanto ainda evidencia a inexistência de uma correlação. Em todos os outros casos, os valores foram ainda menores.

É importante notar também que, para quase todos os jornais, o p-valor obtido também estava acima do limiar aceitável de 0.05 (ou 5%), novamente tendo o Estadão como exceção.

Considerando os coeficientes de correlação quase nulos e também os p-valores altos, os dados apontam para a ausência de qualquer correlação entre os eixos.

## 3.2 Twitter

Figura 2 – Correlação entre saliência nos jornais online e interações no Twitter



Fonte – Igor Antun, 2023

Tabela 2 – Correlação de Spearman - Twitter

Jornal	$\rho$	p-valor
Estadão	0.31	7.5e-14
Folha de S.Paulo	0.015	0.74
g1	0.037	0.31
UOL	0.072	0.051
Veja	0.14	0.012

Fonte – Igor Antun, 2023

No caso do Twitter, novamente encontramos um cenário onde o Estadão se mostra exceção frente aos outros jornais. Seu coeficiente de 0.31 do Estadão aponta para a existência de uma correlação fraca a moderada. Ademais, o p-valor próximo a zero também corrobora essa hipótese, afastando a possibilidade de isso se dar ao acaso.

Cabe mencionar também o caso da Veja, onde mais uma vez se conferiu um coeficiente de 0.14 e, nesse caso, com p-valor que possibilita a interpretação da existência de uma correlação fraca.

Os outros jornais, por sua vez, não apresentaram qualquer correlação.

## 4 Conclusão

Os resultados gerais apontam para a inexistência de correlação entre as agendas social e midiática no contexto das eleições de 2022 no Brasil. Ou seja, dentro do período analisado não comprovou-se relação significativa entre o destaque dado pelos jornais digitais às matérias observadas e o número de interações que estas obtiveram no Facebook ou no Twitter.

## Referências

- AHLERS, D. News consumption and the new electronic media. *Harvard International Journal of Press/Politics*, v. 11, n. 1, p. 29–52, 2006. Disponível em: <https://doi.org/10.1177/1081180X05284317>. Citado na página 9.
- MCCOMBS, M. A look at agenda-setting: past, present and future. *Journalism Studies*, Informa UK Limited, v. 6, n. 4, p. 543–557, 2005. Citado na página 9.
- NUNOMURA, E. A imprensa, o twitter e as eleições de 2010 no brasil. *Revista Brasileira de Ciências da Comunicação [online]*, v. 36, n. 1, p. 103–126, 2013. Disponível em: <https://doi.org/10.1590/S1809-58442013000100006>. Citado 2 vezes nas páginas 9 e 10.
- STERRETT, D.; MALATO, D.; BENZ, J.; KANTOR, L.; TOMPSON, T.; ROSENSTIEL, T.; SONDERMAN, J.; LOKER, K. Who shared it?: Deciding what news to trust on social media. *Digital Journalism*, Routledge, v. 7, n. 6, p. 783–801, 2019. Disponível em: <https://doi.org/10.1080/21670811.2019.1623702>. Citado na página 9.
- ZHURAVSKAYA, E.; PETROVA, M.; ENIKOLOPOV, R. Political effects of the internet and social media. *Annual Review of Economics*, v. 12, n. 1, p. 415–438, 2020. Disponível em: <https://doi.org/10.1146/annurev-economics-081919-050239>. Citado na página 10.