

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ECONOMIA

BRUNA DOS SANTOS ALVES

**ELEIÇÕES EM REDE: IMPACTO DOS GASTOS DIGITAIS NO DESEMPENHO DE
CANDIDATOS A DEPUTADO FEDERAL E A DEPUTADO ESTADUAL EM 2022.**

JUIZ DE FORA - MG

2026

BRUNA DOS SANTOS ALVES

ELEIÇÕES EM REDE: IMPACTO DOS GASTOS DIGITAIS NO DESEMPENHO DE CANDIDATOS A DEPUTADO FEDERAL E A DEPUTADO ESTADUAL EM 2022.

Monografia apresentada ao curso de Ciências Econômicas da Universidade Federal de Juiz de Fora, como requisito à obtenção do título de bacharel em Ciências Econômicas.

Orientador: Prof. Dr. Sidney Martins Caetano.

JUIZ DE FORA – MG

2026

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Alves, Bruna dos Santos.

Eleições em rede: impacto dos gastos digitais no desempenho de candidatos a deputado federal e a deputado estadual em 2022 / Bruna dos Santos Alves. -- 2026.

56 f.

Orientador: Sidney Martins Caetano

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Juiz de Fora, Faculdade de Economia, 2026.

1. Financiamento eleitoral. 2. Comunicação política digital. 3. Eleições brasileiras de 2022. 4. Desempenho eleitoral. 5. Modelos de escolha discreta. I. Caetano, Sidney Martins, orient. II. Título.



UNIVERSIDADE FEDERAL DE JUIZ DE FORA
REITORIA - FACECON - Depto. de Economia

FACULDADE DE ECONOMIA / UFJF
ATA DE APROVAÇÃO DE MONOGRAFIA II (MONO B)

NA DATA DE 16/06/2026, ÀS NOVE HORAS, A BANCA EXAMINADORA, COMPOSTA PELOS PROFESSORES E DOUTORANDA

- 1 - SIDNEY MARTINS CAETANO - ORIENTADOR;
- 2 - WILSON LUIZ ROTATORI CORRÊA; E
- 3 - ANA LUCRÉCIA DUTRA DE SOUZA.

REUNIU-SE NA FACULDADE DE ECONOMIA, SALA 02/4º ANDAR, PARA AVALIAR A MONOGRAFIA DA ACADÊMICA **BRUNA DOS SANTOS ALVES**, INTITULADA: **ELEIÇÕES EM REDE: IMPACTO DOS GASTOS DIGITAIS NO DESEMPENHO DE CANDIDATOS A DEPUTADO FEDERAL E A DEPUTADO ESTADUAL EM 2022**. APÓS PRIMEIRA AVALIAÇÃO, RESOLVEU A BANCA SUGERIR ALTERAÇÕES AO TEXTO APRESENTADO, CONFORME RELATÓRIO SINTETIZADO PELO ORIENTADOR. A BANCA, DELEGANDO AO ORIENTADOR A OBSERVÂNCIA DAS ALTERAÇÕES PROPOSTAS, RESOLVEU **APROVAR** A REFERIDA MONOGRAFIA.

ASSINATURA ELETRÔNICA DOS AVALIADORES



Documento assinado eletronicamente por **Sidney Martins Caetano, Professor(a)**, em 19/06/2026, às 15:18, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Wilson Luiz Rotatori Correa, Professor(a)**, em 19/06/2026, às 15:21, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Lucrécia Dutra de Souza, Usuário Externo**, em 19/06/2026, às 16:59, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **3034733** e o código CRC **F40E2076**.

AGRADECIMENTOS

Dedico esta monografia à minha mãe, pela luta diária ao longo de toda a nossa infância para que eu e meu irmão tivéssemos a oportunidade de estudar e alcançar oportunidades que não lhe foram dadas.

Agradeço à Cleide, ao meu tio Luiz, Laurita, José e amigos, pelo apoio, incentivo e companheirismo durante toda a minha graduação.

Em especial, agradeço ao meu irmão Nathan, que partiu neste ano e que, ao longo do desenvolvimento deste trabalho, percebi o quanto foi minha inspiração para os estudos.

Agradeço também ao meu orientador, professor Sidney, por toda a paciência, dedicação e apoio para que este trabalho pudesse ser realizado.

RESUMO

Este trabalho analisa a relação entre os gastos em comunicação digital e o desempenho eleitoral de candidatos nas eleições brasileiras de 2022. Inserido no contexto de crescente digitalização das campanhas, o estudo investiga como diferentes categorias de investimento em mídia digital influenciam a probabilidade de eleição.

A pesquisa utiliza dados públicos do Tribunal Superior Eleitoral (TSE), consolidados em nível individual, que reúnem informações de votação, perfil dos candidatos e registros detalhados de despesas de campanha. Os gastos digitais são classificados em impulsionamento de conteúdo, produção de programas de rádio, televisão ou vídeo, criação de páginas na internet e publicidade em jornais e revistas. Metodologicamente, empregam-se modelos de escolha discreta do tipo Probit, com especificações Logit como teste de robustez, em abordagens agregada e desagregada, com interpretação baseada em efeitos marginais.

Os resultados indicam forte concentração regional e partidária dos gastos digitais e mostram que o impacto do investimento varia conforme o nível da disputa. Para deputados estaduais, o gasto individual apresenta efeito positivo e estatisticamente significativo sobre a probabilidade de eleição. Para deputados federais, destacam-se os investimentos em produção de programas de rádio, televisão ou vídeo, enquanto algumas categorias de gasto apresentam efeitos negativos. Os resultados se mostram robustos entre as especificações, evidenciando que o financiamento digital influencia o desempenho eleitoral, ainda que de forma heterogênea entre cargos e tipos de despesa.

PALAVRAS-CHAVE: Financiamento eleitoral, comunicação digital, eleições de 2022, desempenho eleitoral.

ABSTRACT

This study analyzes the relationship between digital communication expenditures and the electoral performance of candidates in Brazil's 2022 elections. Situated within the context of the increasing digitalization of electoral campaigns, the research investigates how different categories of investment in digital media influence the probability of election.

The analysis uses public data from the Superior Electoral Court (TSE), consolidated at the individual level, combining information on vote totals, candidate profiles, and detailed records of campaign expenditures. Digital spending is classified into content boosting, production of radio, television, or video programs, website creation, and advertising in newspapers and magazines. Methodologically, the study employs discrete choice models of the Probit type, with Logit specifications used as robustness checks, in both aggregated and disaggregated approaches, with interpretation based on marginal effects.

The results indicate a strong regional and partisan concentration of digital expenditures and show that the impact of investment varies according to the level of the electoral race. For state deputy candidates, individual spending has a positive and statistically significant effect on the probability of election. For federal deputy candidates, investments in the production of radio, television, or video programs stand out, while some spending categories exhibit negative effects. The findings are robust across specifications, demonstrating that digital campaign financing influences electoral performance, albeit heterogeneously across offices and types of expenditure.

KEYWORDS: Campaign finance, digital political communication, 2022 elections, electoral performance.

LISTA DE GRÁFICOS

GRÁFICO 1 – PROPORÇÃO DE DOMICÍLIOS COM ACESSO A INTERNET – PNAD (2022)	14
GRÁFICO 2 - PERCENTUAIS DE GASTOS (DEPUTADOS ESTADUAIS).....	25
GRÁFICO 3 - PERCENTUAIS DE GASTOS (DEPUTADOS FEDERAIS)	26
GRÁFICO 4 - EFEITOS MARGINAIS ESTIMADOS A PARTIR DO MODELO PROBIT.....	36
GRÁFICO 5 - CURVA DE PROBABILIDADE DE VITÓRIA POR GASTO DIGITAL	38

LISTA DE TABELAS

TABELA 1 - ESTATÍSTICAS DESCRITIVAS DOS GASTOS DIGITAIS	22
TABELA 2 - GASTOS TOTAIS POR UF (ESTADUAIS E FEDERAIS).....	24
TABELA 3 - GASTOS POR CATEGORIA (DEPUTADOS ESTADUAIS)	28
TABELA 4 - GASTOS POR CATEGORIA (DEPUTADOS FEDERAIS).....	29
TABELA 5 - RANKING DE PARTIDOS (DEPUTADOS ESTADUAIS).....	31
TABELA 6 - RANKING DE PARTIDOS (DEPUTADOS FEDERAIS)	32
TABELA 7 - EFEITOS MARGINAIS DOS MODELOS PROBIT	35

LISTA DE ABREVIATURAS E SIGLAS

TSE – Tribunal Superior Eleitoral

IBGE – Instituto Brasileiro de Geografia e Estatística

PNAD – Pesquisa Nacional por Amostra de Domicílios Contínua

TIC – Tecnologias da Informação e Comunicação

UFJF – Universidade Federal de Juiz de Fora

SUMÁRIO

1.	INTRODUÇÃO	10
2.	REVISÃO DA LITERATURA	12
2.1.	Financiamento Eleitoral e Transformações na Comunicação Política no Brasil	12
2.2.	Marketing Político Digital e Evidências Empíricas sobre Uso de Recursos.....	13
2.3.	Modelos de Escolha Discreta e a Relação Gasto–Desempenho Eleitoral	15
3.	METODOLOGIA E BASE DE DADOS.....	17
3.1.	Metodologia e Modelo Econométrico	17
3.2.	Base de Dados.....	18
3.3.	Procedimentos de Tratamento dos Dados e Implementação Empírica.....	19
4.	MAPEAMENTO E DISTRIBUIÇÃO DOS GASTOS DIGITAIS.....	22
4.1.	Estatísticas Descritivas dos Gastos Digitais	22
4.2.	Distribuição dos Investimentos Digitais por Unidade da Federação (UF)	23
4.3.	Gastos por Categoria de Despesa.....	26
4.4.	Ranking dos Partidos	30
5.	RESULTADOS.....	34
5.1.	Análise dos Determinantes do Sucesso Eleitoral (Modelos Probit e Logit).....	34
5.2.	Modelos Agregados e Estratégias por Categoria	35
5.3.	Intensidade Financeira e Relação entre Gastos e Probabilidade Estimada.....	37
6.	CONCLUSÃO	39
	REFERÊNCIAS	41
	APÊNDICE A – Script em Python para Consolidação das Bases Eleitorais.....	43
	APÊNDICE B – Efeitos Marginais dos Modelos Logit para Probabilidade de Eleição	54

1. INTRODUÇÃO

A relevância desta pesquisa justifica-se pela transformação digital do eleitorado brasileiro. Segundo dados da PNAD Contínua (IBGE, 2022), a internet já alcançava 91,5% dos domicílios no Brasil durante o ano eleitoral. Esse cenário de elevada conectividade contribui para que o ambiente digital se consolide como um dos principais espaços de disputa política contemporânea, ampliando o alcance das campanhas e modificando as estratégias adotadas por candidatos e partidos. Nesse contexto, o financiamento eleitoral assume papel central na organização das campanhas, uma vez que o volume e a alocação dos recursos influenciam diretamente a capacidade de comunicação dos candidatos. Com a crescente importância das plataformas digitais, parte relevante desses recursos passou a ser direcionada para estratégias de impulsionamento de conteúdo, produção audiovisual e engajamento nas redes sociais. Assim, compreender como o investimento em comunicação digital se relaciona com o desempenho eleitoral torna-se fundamental para a análise das eleições recentes no Brasil.

Apesar da ampla literatura que investiga os efeitos do financiamento de campanhas sobre os resultados eleitorais, ainda há desafios importantes relacionados à mensuração desse impacto, especialmente diante de possíveis problemas de endogeneidade e da heterogeneidade entre candidatos. Além disso, as diferenças regionais no acesso à internet e no uso das mídias digitais indicam que os efeitos desses investimentos podem variar significativamente entre distintas realidades socioeconômicas. Diante disso, o presente estudo tem como objetivo analisar a relação entre o investimento em comunicação digital e o sucesso eleitoral nas eleições de 2022 no Brasil. Busca-se, especificamente, avaliar de que forma os gastos com mídia digital influenciam a probabilidade de eleição dos candidatos, considerando características individuais, partidárias e regionais.

Para atingir esse objetivo, serão utilizados modelos de escolha discreta, em particular os modelos Probit e Logit, adequados para a análise de variáveis dependentes binárias, como a condição de eleito ou não eleito. A estratégia empírica incorpora variáveis de controle e procedimentos de robustez, de modo a lidar com possíveis vieses e garantir maior consistência aos resultados obtidos. Por fim, esta pesquisa contribui para a literatura ao integrar a análise do financiamento eleitoral com o contexto contemporâneo das campanhas digitais, oferecendo evidências empíricas sobre o papel dos investimentos em comunicação política no ambiente online. Além disso, os resultados podem auxiliar na compreensão das dinâmicas recentes das eleições brasileiras, bem como subsidiar o debate sobre regulação e transparência no uso de recursos em campanhas eleitorais.

Embora a base consolidada reúna informações referentes a todos os cargos eletivos das eleições de 2022, a etapa econométrica concentra-se exclusivamente nos candidatos aos cargos de deputado federal e deputado estadual (incluindo deputados distritais). Essa escolha decorre do maior número de observações disponíveis e da maior comparabilidade entre as campanhas proporcionais, que compartilham características operacionais semelhantes. Além disso, esse recorte permite avaliar de forma mais precisa a relação entre os gastos em comunicação digital e o desempenho eleitoral, contribuindo para maior consistência estatística das estimativas e melhor interpretação dos resultados.

No que se refere à estrutura da monografia, o trabalho está organizado em seis capítulos, além dos apêndices. Após esta introdução, o Capítulo 2 apresenta a revisão da literatura, abordando o financiamento eleitoral, as transformações da comunicação política digital e os modelos de escolha discreta aplicados ao estudo do desempenho eleitoral. O Capítulo 3 descreve a metodologia adotada, a construção da base de dados e os procedimentos de tratamento e modelagem econométrica. O Capítulo 4 apresenta a análise descritiva dos gastos digitais, contemplando sua distribuição regional, por categoria de despesa e por partido político. O Capítulo 5 discute os resultados dos modelos econométricos, avaliando a relação entre os investimentos em comunicação digital e a probabilidade de eleição. Por último, o Capítulo 6 apresenta as conclusões do estudo, suas limitações e sugestões para pesquisas futuras. Complementarmente, os apêndices contêm os scripts utilizados na consolidação da base de dados e os resultados adicionais dos modelos estimados.

2. REVISÃO DA LITERATURA

Nesta seção, serão analisados estudos científicos que abordam o cenário eleitoral brasileiro, com ênfase nas eleições de 2022 e no papel crescente das mídias digitais nas campanhas políticas. O objetivo é estabelecer as bases conceituais e teóricas que fundamentam a análise da relação entre o investimento digital e os resultados eleitorais, sob uma perspectiva econômica e econométrica.

2.1. Financiamento Eleitoral e Transformações na Comunicação Política no Brasil

O financiamento eleitoral é um dos principais instrumentos estratégicos das campanhas eleitorais: quanto mais recursos, maior o alcance e diversificação das ações em comunicação, principalmente em relação às plataformas digitais. No Brasil, o financiamento público concentra-se sobretudo no Fundo Especial de Financiamento de Campanha (FEFC), utilizado em anos eleitorais, e no Fundo Partidário, repassado mensalmente para a manutenção dos partidos políticos, conforme a legislação vigente (BRASIL, Lei nº 9.504/97).

Ademais, o Brasil adota um sistema combinado de financiamento, que contempla também o financiamento privado, formado por doações de pessoas físicas, limitadas a 10% da renda bruta anual do doador. É importante evidenciar que a proibição das doações empresariais está vigente desde 2015, em decorrência da decisão do Supremo Tribunal Federal (STF) na Ação Direta de Inconstitucionalidade nº 4650, retificada pela Lei nº 13.165/2015.

Segundo Mancuso et al. (2023), o financiamento desempenha um papel central nas disputas eleitorais. Devido a esse cenário, a Justiça Eleitoral assume uma função fundamental como órgão regulador e fiscalizador do processo eleitoral, uma vez que a distribuição de recursos públicos pode tornar-se sujeita a irregularidades caso não seja adequadamente supervisionada.

Speck e Mancuso (2014) realizam uma análise empírica da relação entre financiamento e desempenho eleitoral, evidenciando que o financiamento é um importante determinante nas campanhas para a eleição dos candidatos. Entretanto, os autores demonstram que essa não é a única variável importante para alcançar uma vitória nas eleições, visto que características individuais, como o gênero, entre outras, podem causar efeitos variados em relação ao desempenho de candidatos, o que indica heterogeneidade no efeito dos gastos ligados às campanhas eleitorais.

De acordo com Samuels (2001), o financiamento de campanhas exerce impacto direto sobre o processo democrático, pois os valores investidos influenciam o nível de competitividade eleitoral e o sucesso nas urnas. O autor destaca que o volume de recursos amplia a capacidade

estratégica das campanhas e contribui para diferenciar os candidatos no processo eleitoral. Sob uma perspectiva contemporânea, esse argumento reforça a importância dos investimentos em comunicação, especialmente no ambiente digital, como estratégia para ampliar o alcance e a visibilidade de candidatos e partidos.

A partir de 2018, a autorização para impulsionar conteúdos nas redes sociais reconfigurou a forma como os políticos planejam seus gastos. Antes, investir em mídia digital era mais um complemento, mas agora tornou-se a principal estratégia para aumentar a visibilidade e o engajamento. Entender como esses recursos são administrados e distribuídos é, portanto, essencial para a contextualização deste estudo, visto que tanto o volume quanto a forma de repasse influenciam diretamente as estratégias adotadas nas campanhas, inclusive o nível de investimento em mídia digital.

Gomes (2018) evidencia as transformações promovidas pela internet e pelas redes sociais, destacando, inicialmente, como a tecnologia passou a integrar o cotidiano dos indivíduos e suas relações sociais. Segundo o autor, a facilidade de acesso à informação, frequentemente a apenas um clique de distância, ajuda a explicar a forte presença da internet no campo político, ainda que a política não seja um tema central na rotina da maioria das pessoas. Além disso, o autor ressalta o papel das mobilizações online promovidas por partidos e candidatos.

Além disso, Levitt (1994) apresenta uma estratégia inovadora ao introduzir o conceito de endogeneidade em relação a estudos anteriores sobre financiamento eleitoral, que associavam o aumento dos gastos de campanha a maiores chances de vitória, constantemente superestimando o efeito causal desses gastos. O autor destaca três pontos principais: primeiro, a qualidade do candidato, na medida em que candidatos mais qualificados tendem a arrecadar mais recursos, recebem maior apoio partidário e apresentam maior probabilidade de vitória; segundo, a competitividade do distrito, pois eleições mais disputadas geram maior arrecadação e incentivam níveis mais elevados de gasto; terceiro, o problema da simultaneidade, já que o volume de arrecadação pode ser influenciado pela expectativa de vitória, gerando causalidade reversa. Assim, este estudo considera a possibilidade de endogeneidade entre investimento digital e desempenho eleitoral, adotando especificações com variáveis de controle e testes de robustez por meio dos modelos Probit e Logit.

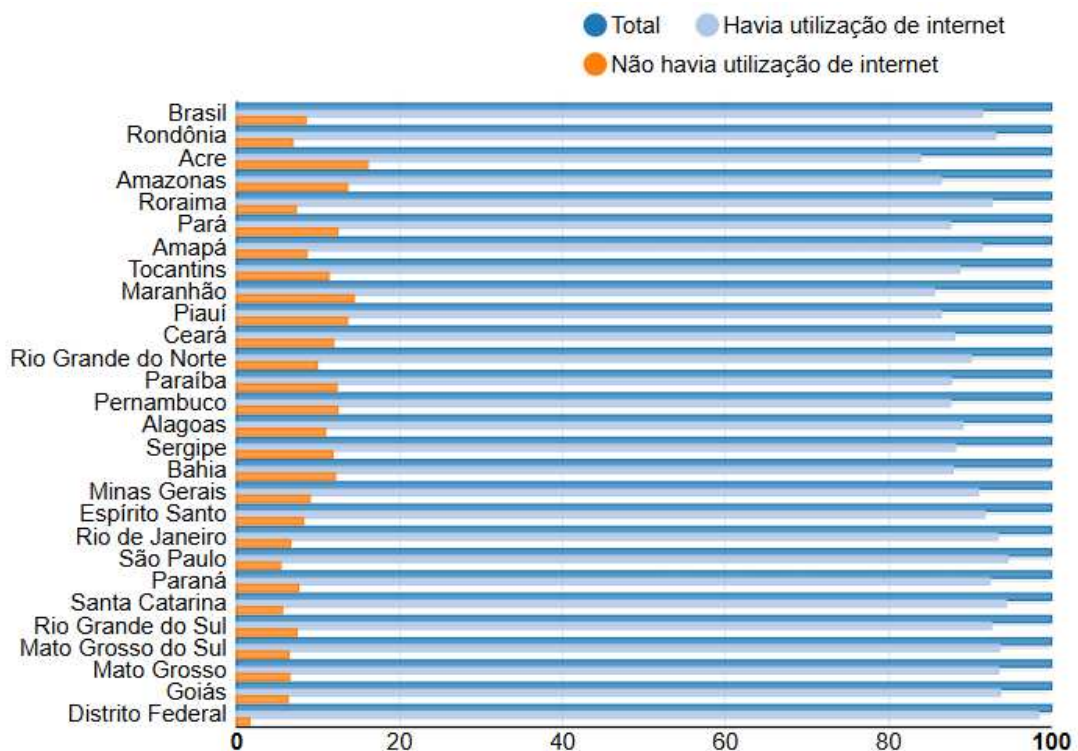
2.2. Marketing Político Digital e Evidências Empíricas sobre Uso de Recursos

O marketing político transformou-se em uma das principais ferramentas estratégicas das campanhas eleitorais, especialmente em um cenário marcado pela elevada conectividade digital contemporânea. A estratégia de investimento em comunicação digital nas eleições de 2022

encontra respaldo estatístico nos dados da PNAD Contínua (IBGE, 2022). De acordo com o Gráfico 1, a internet já alcançava 91,5% dos domicílios brasileiros no período.

No entanto, a análise regional revela diferenças significativas: enquanto o Distrito Federal apresentava uma taxa de conectividade de 98,4%, estados como o Acre registravam 83,9%. Esses dados evidenciam que, embora o ambiente digital tenha se consolidado como um dos principais meios para o debate político, a eficácia dos investimentos em impulsionamento e mídias digitais encontra limites geográficos e socioeconômicos que devem ser considerados pelos estrategistas eleitorais.

GRÁFICO 1 – PROPORÇÃO DE DOMICÍLIOS COM ACESSO À INTERNET – PNAD (2022)



Fonte: IBGE – Pesquisa Nacional por Amostra de Domicílios Contínua Anual – 4º trimestre

Chadwick (2013) refuta a ideia de que as mídias tradicionais foram substituídas pelas mídias digitais, propondo a existência de um sistema híbrido, no qual ambas coexistem e geram uma relação de interdependência, operando de forma complementar. O autor explora como jornalistas e políticos utilizam simultaneamente os dois sistemas de comunicação, evidenciando que não há uma hierarquia entre eles, por exemplo, a noção de que as mídias digitais teriam necessariamente mais poder do que a televisão. Segundo o autor, os candidatos mais bem-

sucedidos são aqueles que conseguem utilizar estrategicamente ambas as formas de mídia. Contudo, o sistema híbrido pode ampliar desigualdades, uma vez que candidatos com mais recursos tendem a expandir seu alcance comunicacional com maior eficiência.

Kreiss (2016) ressalta como as tecnologias transformaram a comunicação política, tornando-se uma ferramenta essencial para que os candidatos obtenham vantagens competitivas nas eleições e alcancem o maior número possível de eleitores. Para o autor, os investimentos expressivos tanto em campanhas quanto em partidos, no contexto tecnológico contemporâneo, não têm como objetivo apenas acompanhar a realidade vigente, mas também criar novos cenários digitais, destacando a importância da circulação de profissionais políticos entre diferentes partidos, a fim de levarem os conhecimentos adquiridos para outras campanhas. O termo “protótipos”, sinalizado pelo autor, refere-se a um novo tipo de política, no qual determinados modelos de campanha passam a servir de referência para outras, como ocorreu no caso da campanha de Barack Obama em 2012, que se tornou um modelo do uso de tecnologias em campanhas eleitorais.

Nesse cenário, os gastos com impulsionamento, produção audiovisual e presença digital deixam de ser custos secundários e tornam-se investimentos estratégicos fundamentais nas campanhas atuais.

2.3. Modelos de Escolha Discreta e a Relação Gasto–Desempenho Eleitoral

Greene (2012) e Wooldridge (2010) destacam que modelos de escolha discreta são apropriados em situações nas quais a variável dependente assume valores binários, como a condição de um candidato ser eleito ou não. Nesses casos, os autores apontam limitações relevantes no uso de modelos lineares tradicionais, como o método de Mínimos Quadrados Ordinários, uma vez que esse procedimento pode gerar previsões fora do intervalo entre 0 e 1, além de violar pressupostos fundamentais, especialmente o da homocedasticidade dos erros. Em razão disso, os modelos de escolha discreta, especialmente Probit e Logit, tornaram-se amplamente utilizados na literatura aplicada à ciência política e à economia eleitoral.

Segundo Greene (2012) e Wooldridge (2010), esses modelos permitem estimar a probabilidade de ocorrência de determinado evento, captando relações não lineares entre as variáveis explicativas e a probabilidade de êxito eleitoral. No campo do financiamento de campanhas, essa abordagem é particularmente relevante para investigar de que forma diferentes tipos de investimento influenciam o desempenho dos candidatos, considerando, simultaneamente, características individuais, partidárias e institucionais.

Gujarati e Porter (2009) ressaltam que, no que se refere às variáveis de gasto, é comum a utilização de transformações logarítmicas, sobretudo em razão da elevada assimetria e da presença de observações extremas, características recorrentes em dados financeiros de campanhas eleitorais. De acordo com os autores, a transformação logarítmica contribui para reduzir a influência desses valores, além de facilitar a interpretação econômica dos coeficientes estimados em termos de variações relativas.

Gujarati e Porter (2009) e Wooldridge (2010) também discutem o uso de variáveis indicadoras (dummies) para representar características qualitativas dos indivíduos. No contexto eleitoral, essas variáveis podem representar características como gênero, escolaridade, estado civil e raça/cor dos candidatos. Segundo os autores, esse procedimento possibilita a comparação entre grupos distintos dentro da amostra, desde que uma categoria seja definida como referência, evitando problemas de multicolinearidade perfeita.

Cameron e Miller (2015) destacam ainda que os estudos eleitorais frequentemente lidam com bases de dados que apresentam agrupamentos naturais, como candidatos pertencentes ao mesmo partido, estado ou coligação. Nesses contextos, os autores argumentam que a correlação dos erros dentro dos grupos pode comprometer a validade da inferência estatística caso não seja devidamente tratada. Dessa forma, a utilização de erros-padrão robustos e clusterizados constitui prática consolidada na literatura para lidar com esse tipo de dependência.

Por fim, Long e Freese (2014) enfatizam que, em razão da natureza não linear dos modelos Probit e Logit, os coeficientes estimados não possuem interpretação direta em termos de variação na probabilidade do evento de interesse. Assim, os autores recomendam a utilização de efeitos marginais como principal instrumento de interpretação dos resultados, uma vez que esses efeitos quantificam o impacto das variáveis explicativas sobre a probabilidade estimada, mantendo as demais constantes.

A estrutura inicial de implementação em Python inspirou-se nas aplicações apresentadas em Heiss e Brunner (2024), sendo posteriormente adaptada às especificidades da base de dados e dos objetivos desta pesquisa.

Dessa forma, os procedimentos metodológicos constituem a base analítica predominante nos estudos empíricos que buscam compreender o papel dos gastos de campanha no sucesso eleitoral, oferecendo um arcabouço consistente para a análise da relação entre investimento em comunicação política e desempenho nas urnas.

3. METODOLOGIA E BASE DE DADOS

3.1. Metodologia e Modelo Econométrico

Nesta seção, apresenta-se o modelo econométrico utilizado, sua especificação, interpretação e resultados esperados. Para tanto, foram utilizados dois modelos de escolha discreta: o Probit, estabelecido como especificação principal, e o Logit, empregado como teste de robustez. O uso combinado dessas duas formas funcionais é comum na literatura, pois a convergência dos resultados aumenta a validade das estimativas e evita que as conclusões sejam sensíveis à escolha do modelo.

A probabilidade de o candidato i ser eleito $P(Y_i = 1 | X_i)$ é formalmente modelada pela função de distribuição acumulada da normal padrão $\Phi(\cdot)$ (Modelo Probit):

$$P(Y_i = 1 | X_i) = \Phi(\beta_0 + \beta_1 \cdot Investimento_i + \beta_2 \cdot Controles_i),$$

onde Y_i é a variável binária de desempenho eleitoral (1 se eleito, 0 caso contrário), $Investimento_i$ representa os gastos em comunicação digital do candidato i , transformados pela função logarítmica $\ln(1 + x)$ para lidar com zeros e assimetrias, e $Controles_i$ é o vetor de variáveis sociodemográficas e políticas, como gênero, escolaridade, estado civil e raça, que atuam como fatores de controle. Os coeficientes β são os parâmetros a serem estimados.

O coeficiente β_1 , associado à variável de investimento, indica como a propensão de um candidato ser eleito se altera quando os gastos em comunicação digital aumentam uma unidade. Para interpretação direta em termos de probabilidade, serão calculados os efeitos marginais posteriormente. Espera-se que $\beta_1 > 0$, ou seja, maiores investimentos em comunicação digital aumentam a probabilidade de um candidato ser eleito, mantendo todas as demais variáveis constantes.

A estratégia empírica foi organizada em duas abordagens:

- 1) Modelo Agregado: consolida os gastos por categoria (Criação de Páginas, Impulsionamento de Páginas, Produção de programas de rádio, TV ou vídeo e Publicidade em Jornais/Revistas), permitindo identificar o impacto de cada tipo de investimento;
- 2) Modelo Desagregado: trata cada despesa individual como uma observação própria,

representada pela variável `Total_Pago_log` (versão logarítmica do valor pago). Essa abordagem permite avaliar o efeito marginal de cada transação isolada e lidar com um número maior de observações.

3.2. Base de Dados

A construção da base de dados utilizada na análise empírica envolveu a consolidação de diferentes conjuntos de dados públicos disponibilizados pelo Tribunal Superior Eleitoral (TSE). Todo o processamento foi conduzido em Jupyter Notebook, utilizando Python 3.11 e as bibliotecas Pandas, NumPy e Matplotlib. A partir desses procedimentos, foi gerado o arquivo `Relatorio_Completo_Consolidado_2022`, que reúne, em nível individual, informações de votação, características pessoais dos candidatos e registros detalhados dos gastos individuais de campanha. O processo completo de integração, tratamento e consolidação dos dados encontra-se reproduzido integralmente nos scripts disponibilizados no Apêndice A, assegurando transparência metodológica e permitindo a replicação dos resultados por outros pesquisadores.

O processo de integração das informações foi realizado de forma sequencial, respeitando a coerência lógica da unidade de análise.

Inicialmente, realizou-se a preparação e integração de quatro bases centrais:

- 1) Despesas pagas (`despesas_pagas_candidatos_2022_BRASIL`);
- 2) Despesas contratadas (`despesas_contratadas_candidatos_2022_BRASIL`);
- 3) Votação por município e zona eleitoral (`votacao_candidato_munzona_2022_BRASIL`); e
- 4) Cadastro de candidaturas (`consulta_cand_2022_BRASIL`).

A vinculação entre essas bases seguiu uma ordem predefinida, de modo a garantir consistência e permitir que cada etapa agregasse novas dimensões informacionais ao nível individual de análise. Primeiramente, integrou-se o arquivo de despesas pagas ao de despesas contratadas; em seguida, o resultado desse cruzamento foi associado ao desempenho eleitoral proveniente da base de votação; por fim, foram incorporadas as informações pessoais e características socioeconômicas dos candidatos. Essa ordem permitiu construir uma estrutura analítica coerente, relacionando gastos digitais, atributos individuais e resultados eleitorais.

Antes da etapa de unificação das bases, foi realizado o filtro das categorias diretamente relacionadas à comunicação digital (o filtro dos códigos foi realizado antes do cruzamento entre

despesas pagas e contratadas, garantindo que apenas registros pertinentes à comunicação digital fossem levados para as etapas subsequentes), de forma a isolar os investimentos relevantes para a estratégia empírica deste estudo.

Dentro da base de despesas pagas, foram selecionados os seguintes códigos: 20420000 (impulsioneamento de conteúdos), 20220000 (produção de programas para rádio, TV ou vídeo), 20260000 (criação e inclusão de páginas na internet) e 20130000 (publicidade em jornais e revistas), este último incluído por sua função complementar na estratégia de comunicação e exposição pública dos candidatos. Esses códigos compõem a variável de gasto digital, utilizada tanto na análise descritiva quanto nos modelos econométricos.

Posteriormente, procedeu-se à regionalização das despesas. Como a localização geográfica não era uniforme entre as bases originais, adotou-se um critério hierárquico para identificação da unidade da federação associada a cada registro: (1) SG_UF (UF da despesa); (2) SG_UE (unidade eleitoral da contratação); (3) SG_UF_FORNECEDOR; e, caso todas estivessem ausentes, atribuiu-se a categoria “BR”. A variável resultante, denominada SG_UF_FINAL, permitiu gerar agregações regionais, tais como o total investido por candidato, a média de gasto digital por candidato e a distribuição territorial dos votos, mantendo comparabilidade entre diferentes cargos.

A etapa seguinte consistiu na integração das informações de votação. Como a base de votações apresenta grande volume de dados, sua importação foi realizada por meio de leitura segmentada (chunksize), técnica que permite processar arquivos extensos sem sobrecarregar a memória do ambiente. Essa etapa resultou na variável “Total_Votos”, utilizada posteriormente tanto nas análises descritivas quanto na construção da variável dependente binária dos modelos de escolha discreta. Em sequência, foram incorporados os dados de resultado da candidatura (situação eleito/não eleito), extraídos da base de votação.

A base consolidada incluiu ainda as características individuais dos candidatos: gênero, escolaridade, estado civil, raça/cor e ocupação. Para assegurar consistência na construção de variáveis dummy utilizadas nos modelos, categorias esparsas como “NÃO INFORMADO” e “NÃO DIVULGÁVEL” foram tratadas e, quando inviabilizavam a criação adequada das variáveis, foram removidas especificamente na etapa de modelagem econométrica, preservando-as na estrutura original da base para garantir transparência e reprodutibilidade.

Por fim, para fins de organização analítica e compatibilidade com a etapa econométrica, o arquivo foi exportado em abas distintas, separando eleições federais e estaduais.

3.3. Procedimentos de Tratamento dos Dados e Implementação Empírica

A etapa econométrica foi implementada integralmente em ambiente Python, utilizando as bibliotecas *Pandas*, *NumPy* e *Statsmodels*. O procedimento iniciou-se com a importação da base consolidada de dados eleitorais e financeiros referentes às eleições de 2022, construída a partir de informações públicas do Tribunal Superior Eleitoral (TSE).

Inicialmente, os valores monetários das despesas de campanha foram padronizados, convertendo-se os registros textuais para formato numérico. Em seguida, a amostra foi restrita aos candidatos aos cargos de deputado estadual e deputado federal, com a unificação da denominação “Deputado Distrital” à categoria de deputado estadual. A variável dependente foi definida de forma binária, assumindo valor igual a 1 para candidatos eleitos e 0 caso contrário.

Com o objetivo de lidar com a elevada assimetria dos gastos de campanha e com a presença de valores nulos, aplicou-se a transformação logarítmica $\ln(1 + x)$ à variável de despesa paga (*Total_Pago*). Essa transformação permitiu reduzir a influência de valores extremos e garantir a inclusão de observações sem gastos registrados.

As variáveis sociodemográficas dos candidatos, gênero, escolaridade, estado civil e raça/cor, passaram por tratamento prévio, com a exclusão de categorias pouco informativas (“NÃO INFORMADO” e “NÃO DIVULGÁVEL”). Posteriormente, essas variáveis foram convertidas em indicadores binários (dummies), adotando-se uma categoria de referência para evitar problemas de multicolinearidade.

Conforme descrito na seção 3.1, a estratégia empírica foi estruturada em duas abordagens complementares. No modelo agregado, os gastos em comunicação digital foram consolidados por candidato e por tipo de despesa, por meio de uma tabela de agregação que soma os valores pagos em cada categoria. Essas variáveis de gasto agregado também foram transformadas em $\ln(1 + x)$ e incorporadas ao modelo juntamente com o perfil sociodemográfico dos candidatos.

No modelo desagregado, cada observação corresponde a uma despesa individual de campanha. Nessa especificação, a variável explicativa central é o valor logaritmizado da despesa paga (*Total_Pago_log*), permitindo avaliar o efeito marginal de cada transação isolada sobre a probabilidade de êxito eleitoral. Dado que múltiplas despesas pertencem a um mesmo candidato, os erros-padrão foram estimados de forma robusta e clusterizados por candidato (*SQ_CANDIDATO*), corrigindo a dependência intraindivíduo e assegurando inferência estatística consistente.

A estimação foi realizada por meio de modelos Probit, com interpretação dos resultados baseada nos efeitos marginais, os quais permitem quantificar o impacto das variáveis explicativas sobre a probabilidade de eleição. Adicionalmente, os coeficientes estimados, os erros-padrão e os efeitos marginais foram organizados e exportados para arquivos externos,

assegurando transparência, reprodutibilidade e facilitando a construção das tabelas de resultados apresentadas no capítulo seguinte.

4. MAPEAMENTO E DISTRIBUIÇÃO DOS GASTOS DIGITAIS

Este capítulo examina como os candidatos a deputado estadual (incluindo distritais) e deputado federal alocaram gastos em categorias digitais nas eleições de 2022. A análise reúne estatísticas descritivas, distribuição territorial dos gastos, composição das categorias de despesa e volume investido por partido. O objetivo é mapear a utilização desses recursos, destacando diferenças relevantes entre estados e partidos. Essas informações permitem comparar unidades da federação, regiões e siglas partidárias, identificando padrões de uso das ferramentas digitais e preparando o terreno para a análise econométrica subsequente.

4.1. Estatísticas Descritivas dos Gastos Digitais

A Tabela 1 apresenta as estatísticas descritivas dos gastos digitais. Os dados mostram uma assimetria marcante nos investimentos digitais dos dois cargos. As médias ficam acima das medianas, sinalizando que poucos candidatos concentram a maior parte dos recursos, enquanto a maioria opera com valores bem menores. Esse resultado reflete como as campanhas operam na prática: alguns candidatos conseguem reunir estrutura e recursos suficientes para investir valores muito acima da média, enquanto a maior parte trabalha com orçamentos bastante limitados. A amplitude dos valores é mais evidente entre os candidatos a deputado federal, cujas campanhas tendem a alcançar um eleitorado maior e, portanto, exigem estruturas de comunicação digital mais robustas.

Tabela 1 — Estatísticas descritivas dos Gastos Digitais

	Deputado Estadual	Deputado Federal
Somatório	R\$ 142.478.194,71	R\$ 280.347.278,09
Média	R\$ 13.861,09	R\$ 32.769,99
Mediana	R\$ 5.000,00	R\$ 8.000,00
Desvio Padrão	R\$ 27.424,45	R\$ 78.491,04
Máximo	R\$ 500.013,00	R\$ 1.307.500,00
Mínimo	R\$ 3,38	R\$ 2,85
Quantidade de Candidatos	6598	5041

Fonte: Elaboração Própria com base nos dados do TSE.

A comparação entre os cargos confirma que a candidatura para a Câmara Federal demanda

maior investimento médio, maior dispersão dos gastos e valores máximos muito superiores aos observados nas campanhas estaduais e distritais.

4.2. Distribuição dos Investimentos Digitais por Unidade da Federação (UF)

A distribuição dos valores entre as UFs revela diferenças expressivas no território nacional. A Tabela 2 apresenta o total de gastos digitais por UF para deputados estaduais e federais, evidenciando que o Sudeste concentra os maiores montantes investidos. Esse padrão reflete não apenas o tamanho dos eleitorados, mas também a maior capacidade financeira das estruturas partidárias locais. Como o objetivo desta seção é examinar o volume total movimentado, a análise permanece focada nos valores absolutos, independentemente do número de candidaturas em cada estado.

No caso dos candidatos a deputado estadual e distrital, a Tabela 2 mostra que São Paulo concentra o maior volume, ultrapassando R\$ 29 milhões. Em seguida aparecem Minas Gerais, Rio Grande do Sul e Rio de Janeiro. Estados como Acre, Alagoas e Roraima apresentam valores absolutos mais baixos, o que pode estar associado ao menor tamanho do eleitorado e ao escopo das disputas locais. A comparação entre essas UFs evidencia a concentração regional dos recursos, indicando que campanhas em estados maiores tendem a demandar investimentos digitais significativamente superiores.

Tabela 2 — Gastos Totais por UF (Estaduais e Federais)

SG_UF_FINAL	Total Pago – Deputados Estaduais (R\$)	Total Pago – Deputados Federais (R\$)
AC	705.871,78	3.603.478,77
AL	1.100.330,73	4.718.322,13
AM	2.652.678,76	5.665.257,50
AP	2.711.858,88	6.486.567,62
BA	4.531.960,95	8.166.024,30
CE	3.875.257,41	8.368.675,80
DF	5.487.868,89	6.979.952,30
ES	2.641.398,95	7.740.651,93
GO	3.857.968,19	13.235.002,45
MA	2.392.185,71	7.316.733,65
MG	11.500.917,05	24.082.051,84
MS	4.484.761,61	4.932.054,36
MT	5.748.842,10	11.111.535,91
PA	3.174.006,51	7.879.037,47
PB	2.640.380,40	6.781.365,40
PE	4.531.067,65	10.835.994,43
PI	2.026.800,12	2.671.665,48
PR	9.330.372,99	17.027.846,13
RJ	10.288.160,15	21.761.611,39
RN	1.760.844,92	5.173.893,71
RO	2.684.914,49	6.943.945,98
RR	1.219.280,13	2.224.931,62
RS	10.757.784,62	11.377.935,60
SC	9.428.020,56	10.822.779,90
SE	1.739.869,03	1.917.883,30
SP	29.262.988,14	57.636.143,40
TO	1.941.803,99	4.885.935,72

Fonte: Elaboração Própria com base nos dados do TSE.

Ao avançar para as candidaturas a deputado federal, essas disparidades tornam-se ainda mais evidentes. Observa-se, na Tabela 2, que São Paulo supera R\$ 57 milhões, seguido por Minas Gerais e Rio de Janeiro. Estados menores, como Sergipe, Roraima e Piauí, apresentam valores absolutos reduzidos, embora alguns registrem gastos relativamente altos quando considerados em relação ao tamanho de seus eleitorados, indicando que mesmo em unidades menores há esforços concentrados de comunicação digital em campanhas específicas.

Para facilitar a visualização comparativa, foram produzidos gráficos que mostram a participação percentual de cada UF no total de gastos digitais. O Gráfico 2 apresenta os

percentuais para deputados estaduais, enquanto o Gráfico 3 apresenta os percentuais para deputados federais. Esses gráficos correspondem diretamente aos valores da Tabela 2, mas permitem observar de forma clara a participação relativa de cada estado no montante total, reforçando a concentração em unidades mais populosas e a menor presença em estados do Norte e Nordeste.

Gráfico 2 — Percentual de Gastos (Deputados Estaduais)



Fonte: Elaboração Própria com base nos dados do TSE.

Gráfico 3 — Percentual de Gastos (Deputados Federais)



Fonte: Elaboração Própria com base nos dados do TSE.

Os gráficos reforçam a concentração de gastos digitais nos estados mais populosos. Nas regiões Norte e parte do Nordeste, a presença reduzida de recursos evidencia desigualdades estruturais persistentes entre as UFs, aspecto que pode impactar a competitividade das campanhas e a eficácia das estratégias digitais.

4.3. Gastos por Categoria de Despesa

Ao analisar as categorias de despesa, as Tabelas 3 e 4 apresentam os gastos por categoria

para deputados estaduais e federais, respectivamente. Observa-se que o impulsionamento de conteúdo concentra a maior parcela dos recursos, especialmente nos estados com maior volume total de investimento, como São Paulo, Minas Gerais e Rio de Janeiro. Esse padrão indica que, para candidatos em regiões de maior competição, estratégias de alcance rápido e segmentado nas redes digitais são prioridade, sugerindo a necessidade de visibilidade ampla e intensa durante a campanha.

Entretanto, esse comportamento não se repete de forma uniforme no país. Em diversos estados, especialmente aqueles com disputas de menor escala, a produção de programas de rádio, televisão ou vídeo aparece como o item mais elevado, superando inclusive o impulsionamento, conforme evidenciado nas Tabelas 3 e 4. Esse resultado sugere que candidatos em regiões menores podem concentrar seus esforços em formatos mais tradicionais ou audiovisuais, que podem gerar maior impacto local mesmo com orçamentos menores.

As categorias de menor peso, como criação e inclusão de páginas na internet e publicidade por jornais e revistas, mantêm valores reduzidos nas duas bases (estaduais e federais), funcionando como componentes complementares das estratégias digitais. Apesar de não representarem a maior parte dos gastos, essas categorias desempenham papel estratégico, garantindo presença institucional online, principalmente em campanhas com menor orçamento.

Tabela 3 — Gastos por Categoria (Deputados Estaduais)

SG_UF_FINAL	Criação e inclusão de páginas na internet (R\$)	Despesa com Impulsionamento de Conteúdos (R\$)	Produção de programas de rádio, televisão ou vídeo (R\$)	Publicidade por jornais e revistas (R\$)
AC	31.954,63	111.367,15	562.550,00	0
AL	29.325,47	568.564,26	501.241,00	1.200,00
AM	186.703,78	1.675.205,52	788.969,46	1.800,00
AP	150.068,56	403.798,32	2.157.992,00	0
BA	582.429,92	1.681.126,03	2.139.916,00	128.489,00
CE	390.023,84	2.648.228,58	825.880,54	11.124,45
DF	557.300,44	3.867.900,16	903.673,79	158.994,50
ES	196.344,16	1.596.616,31	705.433,48	143.005,00
GO	281.822,43	2.356.743,92	1.109.364,57	110.037,27
MA	142.800,20	1.054.552,24	1.177.943,08	16.890,19
MG	498.991,03	8.299.744,78	1.664.205,47	1.037.975,77
MS	377.097,84	2.077.552,45	1.639.291,32	390.820,00
MT	199.879,58	3.012.478,08	2.404.729,44	131.755,00
PA	165.085,73	1.324.445,78	1.673.175,00	11.300,00
PB	124.916,83	573.243,40	1.915.820,17	26.400,00
PE	462.433,24	2.620.472,80	1.416.656,22	31.505,39
PI	29.474,19	613.008,05	1.318.117,88	66.200,00
PR	790.562,41	5.436.739,97	1.438.897,28	1.664.173,33
RJ	600.352,17	7.550.177,50	1.034.693,14	1.102.937,34
RN	207.533,68	856.889,24	658.082,00	38.340,00
RO	157.463,01	1.364.300,55	1.136.198,93	26.952,00
RR	59.566,99	401.890,89	710.933,25	46.889,00
RS	817.996,37	6.073.909,66	1.467.603,10	2.398.275,49
SC	470.233,16	5.513.294,64	2.082.238,34	1.362.254,42
SE	35.783,74	326.358,18	1.376.977,11	750
SP	1.508.177,13	22.035.124,34	2.332.055,12	3.387.631,55
TO	135.556,77	568.186,22	1.223.061,00	15.000,00

Fonte: Elaboração Própria com base nos dados do TSE.

Tabela 4 — Gastos por Categoria (Deputados Federais)

SG_UF_FINAL	Criação e inclusão de páginas na internet (R\$)	Despesa com Impulsioneamento de Conteúdos (R\$)	Produção de programas de rádio, televisão ou vídeo (R\$)	Publicidade por jornais e revistas (R\$)
AC	141.736,88	585.748,45	2.875.493,44	500
AL	69.865,68	1.698.027,06	2.896.929,39	53.500,00
AM	475.586,55	2.845.992,28	2.342.478,67	1.200,00
AP	161.264,09	612.419,95	5.552.883,58	160.000,00
BA	1.288.066,99	3.404.201,06	3.387.156,25	86.600,00
CE	522.371,28	5.461.395,41	2.321.556,94	63.352,17
DF	390.052,28	3.965.392,63	2.477.261,39	147.246,00
ES	186.113,23	4.172.584,52	3.140.720,40	241.233,78
GO	1.382.991,50	5.351.639,82	6.255.624,63	244.746,50
MA	516.923,51	1.909.620,04	4.867.790,10	22.400,00
MG	1.162.529,44	15.498.497,07	5.011.990,95	2.409.034,38
MS	256.233,47	2.615.971,63	1.820.799,26	239.050,00
MT	260.557,99	5.033.935,18	5.634.355,74	182.687,00
PA	496.268,54	2.745.279,81	4.556.039,12	81.450,00
PB	139.578,77	1.448.726,78	5.148.459,85	44.600,00
PE	886.363,70	5.840.431,29	3.885.076,22	224.123,22
PI	98.312,01	1.107.648,77	1.465.204,70	500
PR	1.103.072,33	11.192.634,19	3.038.498,68	1.693.640,93
RJ	1.386.713,02	15.470.390,45	2.850.829,11	2.053.678,81
RN	396.583,04	1.352.351,77	3.245.346,40	179.612,50
RO	645.984,78	3.068.743,15	3.191.828,05	37.390,00
RR	102.759,03	679.236,29	1.408.116,30	34.820,00
RS	1.163.636,35	5.657.905,47	2.636.227,73	1.920.166,05
SC	580.351,61	6.336.271,02	2.799.974,97	1.106.182,30
SE	15.069,10	553.926,20	1.336.628,00	12.260,00
SP	3.903.314,90	42.347.701,55	7.486.885,79	3.898.241,16
TO	250.839,70	923.360,62	3.668.555,20	43.180,20

Elaboração Própria com base nos dados do TSE.

Ao observar os estados individualmente, nota-se que aqueles com maior número de candidaturas e competição intensa, como SP, MG e RJ, apresentam valores expressivos em praticamente todas as categorias, com destaque para o impulsioneamento, indicando uma alocação estratégica ampla dos recursos. Nos estados menores, os totais são mais concentrados e seguem um padrão mais homogêneo, embora algumas campanhas isoladas se destaquem por investimentos acima da média local, refletindo iniciativas individuais de candidatos ou partidos.

4.4. Ranking dos Partidos

O ranking partidário apresentado nesta seção descreve a distribuição dos gastos digitais realizados pelos candidatos de cada sigla, permitindo comparar como diferentes partidos se posicionam em relação aos investimentos em comunicação digital. Embora o foco esteja na agregação por partido, a perspectiva adotada é individual, isto é, considera-se o conjunto dos gastos efetuados pelos candidatos vinculados a cada legenda e não desembolsos realizados diretamente pelas direções partidárias.

A Tabela 5 sintetiza os valores totais investidos pelos candidatos a deputado estadual e distrital, enquanto a Tabela 6 apresenta o mesmo recorte para os candidatos a deputado federal.

Entre os candidatos estaduais e distritais, observa-se que aqueles filiados a UNIÃO, PL e PP concentram os maiores gastos digitais. Esse padrão sugere que tais legendas combinam maior disponibilidade de recursos, maior estrutura organizacional e maior competitividade eleitoral. Partidos de porte intermediário, como MDB, PSD e PSDB, também apresentam investimentos expressivos, embora em patamar inferior ao dos líderes. Na outra extremidade, legendas menores, como PSTU, UP, PCB e PMB, registram gastos muito reduzidos, refletindo limitações financeiras e menor alcance das suas campanhas.

Tabela 5 — Ranking de Partidos (Deputados Estaduais)

SG PARTIDO	Total Pago
UNIÃO	18.508.087,90
PL	13.327.853,53
PP	11.120.882,20
MDB	10.770.802,20
PSD	10.059.431,04
PSDB	9.845.091,54
PT	9.603.483,69
REPUBLICANOS	8.090.634,88
PODE	6.067.473,95
PSB	5.797.165,88
PDT	5.088.926,82
PSOL	4.875.224,94
PTB	2.880.179,86
CIDADANIA	2.878.755,59
SOLIDARIEDADE	2.809.125,56
NOVO	2.692.302,01
PSC	2.628.028,56
PATRIOTA	2.528.492,77
PV	2.142.218,34
PC do B	2.041.292,43
REDE	1.994.754,66
AVANTE	1.681.316,90
PROS	1.591.055,94
PMN	997.881,34
AGIR	957.507,87
PRTB	629.813,52
DC	539.770,53
PMB	295.190,82
PSTU	35.097,44
UP	352

Elaboração Própria com base nos dados do TSE.

Entre os deputados federais, observa-se padrão semelhante: os partidos com maior capital político e financeiro, como UNIÃO, PL e PP, concentram a maior parte dos recursos, enquanto partidos pequenos registram valores modestos, reforçando a desigualdade na presença digital das campanhas. Esse contraste evidencia que a capacidade de investimento não depende apenas do número de candidaturas, mas também da estrutura interna e do acesso a recursos do Fundo Eleitoral, refletindo diferenças significativas de competitividade entre os partidos. A seguir, apresenta-se o ranking equivalente para os candidatos a deputado federal

Tabela 6 — Ranking de Partidos (Deputados Federais)

SG PARTIDO	Total Pago
UNIÃO	39.604.717,70
PL	24.732.597,53
PP	22.445.030,51
MDB	21.332.319,48
PT	17.227.914,73
PSD	16.871.177,62
PSDB	15.309.996,00
REPUBLICANOS	15.048.176,85
PODE	14.721.246,38
PSB	13.810.310,93
PSOL	8.522.610,01
PTB	8.439.066,58
PDT	7.649.999,41
AVANTE	7.025.723,16
NOVO	6.795.377,52
PSC	6.066.011,05
CIDADANIA	6.063.587,07
PROS	5.817.687,11
PATRIOTA	5.567.935,33
SOLIDARIEDADE	4.620.270,45
REDE	4.020.940,67
PC do B	3.989.350,72
PV	3.094.871,06
PMN	902.615,11
AGIR	295.120,18
PRTB	226.745,41
DC	59.506,10
PSTU	44.800,57
PMB	23.222,60
PCB	10.244,99
UP	8.105,26

Elaboração Própria com base nos dados do TSE.

A análise comparativa das Tabelas 5 e 6 permite identificar tendências estratégicas relevantes. Observa-se que partidos de maior porte tendem a apresentar investimentos mais diversificados e consistentes entre seus candidatos, refletindo estruturas organizacionais mais robustas e maior capacidade de coordenação interna. Já os partidos menores costumam registrar gastos esporádicos e concentrados em poucas categorias, frequentemente direcionados a candidaturas específicas ou ações pontuais de visibilidade local.

Esses padrões reforçam que, mesmo considerando-se a soma dos gastos realizados pelos candidatos de cada legenda, a presença digital das campanhas permanece desigual entre partidos, sendo condicionada por fatores organizacionais, territoriais, pelo acesso ao Fundo Eleitoral e investimento privado. Por fim, essas estatísticas descritivas revelam uma clara concentração regional e partidária dos investimentos em comunicação digital, oferecendo o contexto necessário para a seção econométrica subsequente, que busca mensurar o efeito marginal desses gastos sobre a probabilidade de eleição.

A análise descritiva apresentada anteriormente permitiu identificar a concentração regional e partidária dos recursos digitais. A partir desse panorama, a próxima etapa consiste em investigar se tais investimentos se traduzem em maior probabilidade de sucesso eleitoral, controlando-se pelas características individuais dos candidatos.

5. RESULTADOS

5.1. Análise dos Determinantes do Sucesso Eleitoral (Modelos Probit e Logit)

Esta seção detalha os efeitos dos investimentos em comunicação digital sobre a probabilidade de eleição para os cargos de deputado estadual e federal em 2022. Foram aplicados modelos Probit e Logit de forma comparativa. Enquanto o Probit assume uma distribuição normal dos erros, o Logit assume uma distribuição logística, a aplicação de ambos serve como teste de robustez, garantindo que os sinais e a significância das variáveis não sejam sensíveis à especificação funcional escolhida. A análise é dividida entre o modelo Agregado, que observa a estratégia de alocação por categorias (Despesa com Impulsionamento de Conteúdos, Criação e inclusão de páginas na internet, Produção de programas de rádio, televisão ou vídeo, Publicidade por jornais e revistas), e o modelo Desagregado, que foca no impacto de cada lançamento financeiro individual.

Como ponto de partida, a Tabela 7 consolida os coeficientes e erros-padrão de todas as especificações, servindo de base para as discussões subsequentes.

Tabela 7 – Efeitos Marginais dos Modelos Probit para Probabilidade de Eleição

Variáveis	Estadual Agregado	Estadual Desagregado	Federal Agregado	Federal Desagregado
Criação e Inclusão de Páginas	-0.0039 (0.001)***	—	-0.0020 (0.001)	—
Impulsionamento de Conteúdos	0.0021 (0.001)*	—	-0.0026 (0.001)*	—
Total Pago (Gasto Individual)	—	0.0085 (0.002)***	—	0.0031 (0.002)
Produção de programas de rádio, TV ou vídeo	0.0009 (0.001)	—	0.0050 (0.001)***	—
Publicidade por jornais e revistas	0.0008 (0.001)	—	-0.0070 (0.002)***	—
Gênero (Masculino)	0.0861 (0.013)***	0.0907 (0.014)***	0.0548 (0.015)***	0.0335 (0.017)*
Grau de Instrução (Lê e Escreve)	0.0545 (0.103)	0.0780 (0.112)	-0.3286 (0.124)***	-0.2672 (0.143)*
Grau de Instrução (Ensino Médio Incompleto)	-0.1585 (0.075)**	-0.1491 (0.081)	-0.1436 (0.079)	-0.1673 (0.087)*
Raça (Indígena)	0.0099 (0.114)	0.0339 (0.122)	0.2773 (0.144)*	0.3931 (0.162)**
Observações (N)	6.570	10.240	5.023	8.526
Pseudo R^2	0,0132	0,0141	0,0135	0,0061

Nota: Erros-padrão entre parênteses. *** $p < 0,01$; ** $p < 0,05$; * $p < 0,10$.

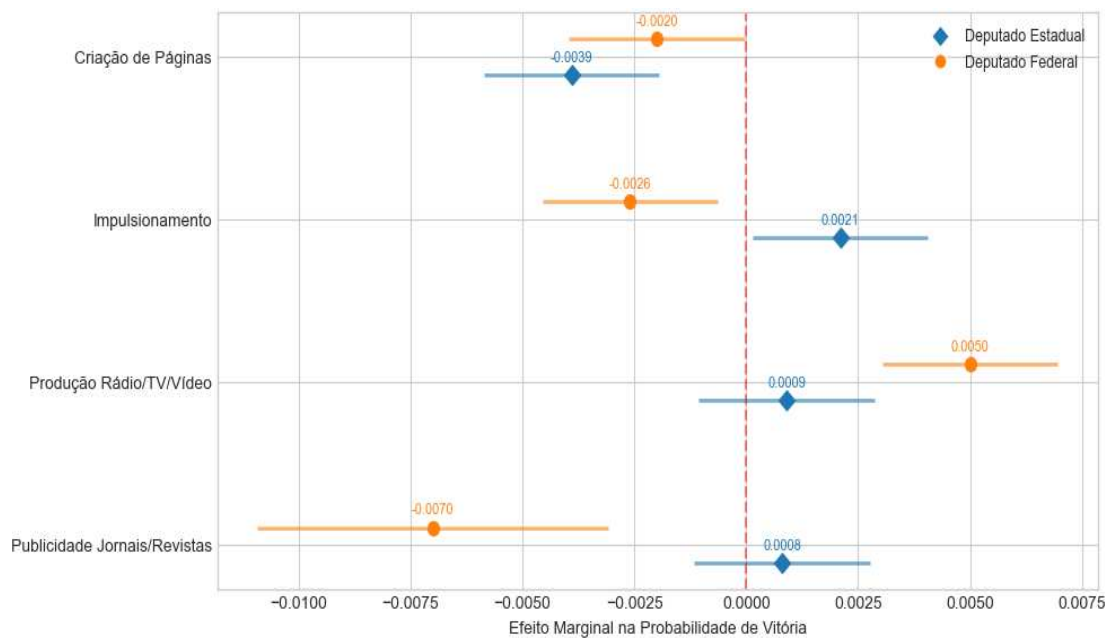
Os resultados apresentados na Tabela 7 referem-se aos efeitos marginais do modelo Probit. Ressalta-se que as estimativas obtidas via modelo Logit (omitidas por brevidade, mas disponíveis no Apêndice B) apresentaram convergência de sinal e significância em todas as variáveis testadas, confirmando a consistência dos achados independentemente da técnica de estimação utilizada. Observa-se ainda associação positiva para candidatos indígenas no nível federal, embora os resultados devam ser interpretados com cautela devido ao número reduzido de observações nessa categoria.

5.2. Modelos Agregados e Estratégias por Categoria

Os resultados do modelo agregado revelam divergências estruturais no retorno do investimento entre os cargos. No nível federal, a alocação de recursos em categorias associadas à produção de conteúdo audiovisual profissionalizado apresenta maior associação marginal estimada entre as categorias analisadas, sendo a Produção de Programas de Rádio, TV ou Vídeo a categoria com maior contribuição positiva ($dy/dx = +0,0050$; $P < 0,001$), indicando maior sensibilidade da probabilidade estimada de eleição a esse tipo de investimento. Em contrapartida, investimentos em Publicidade em Jornais e Revistas apresentaram a maior associação negativa para este cargo ($dy/dx = -0,0070$; $P < 0,001$).

Para os candidatos a Deputado Estadual, observa-se que o gasto em Criação de Páginas na Internet apresentou associação negativa ($dy/dx = -0,0039$; $P = 0,003$). Este achado sugere que, no âmbito estadual, a alocação de recursos exclusivamente na criação de páginas na internet, sem estratégias complementares de amplificação, está associada a menor probabilidade estimada de vitória. O Gráfico 4 sintetiza visualmente esses efeitos marginais, permitindo a comparação direta entre as categorias e os cargos.

GRÁFICO 4 - Efeitos Marginais Estimados a partir do Modelo Probit



Fonte: Elaboração própria com base em dados do TSE (2022).

Além das despesas, o perfil socioeconômico exerce influência robusta: ser do gênero masculino eleva a probabilidade de vitória em 8,6 p.p. para estadual e 5,5 p.p. para federal, evidenciando disparidades históricas de gênero na representação política.

Quanto ao nível educacional, os resultados revelam que a escolaridade exerce um papel diferente conforme o cargo. No nível federal, observa-se uma penalização estatisticamente significativa para candidatos com baixos níveis de instrução (Lê e Escreve: $dy/dx = -0,3286$; $P < 0,01$), sugerindo que a disputa federal se mostra mais seletiva em termos de escolaridade formal. Já no nível estadual, essa barreira é menos discriminante, com a maioria das faixas de instrução não apresentando coeficientes significativos, o que indica um cenário competitivo mais heterogêneo em termos de formação acadêmica. A categoria de referência para escolaridade é o ensino superior completo.

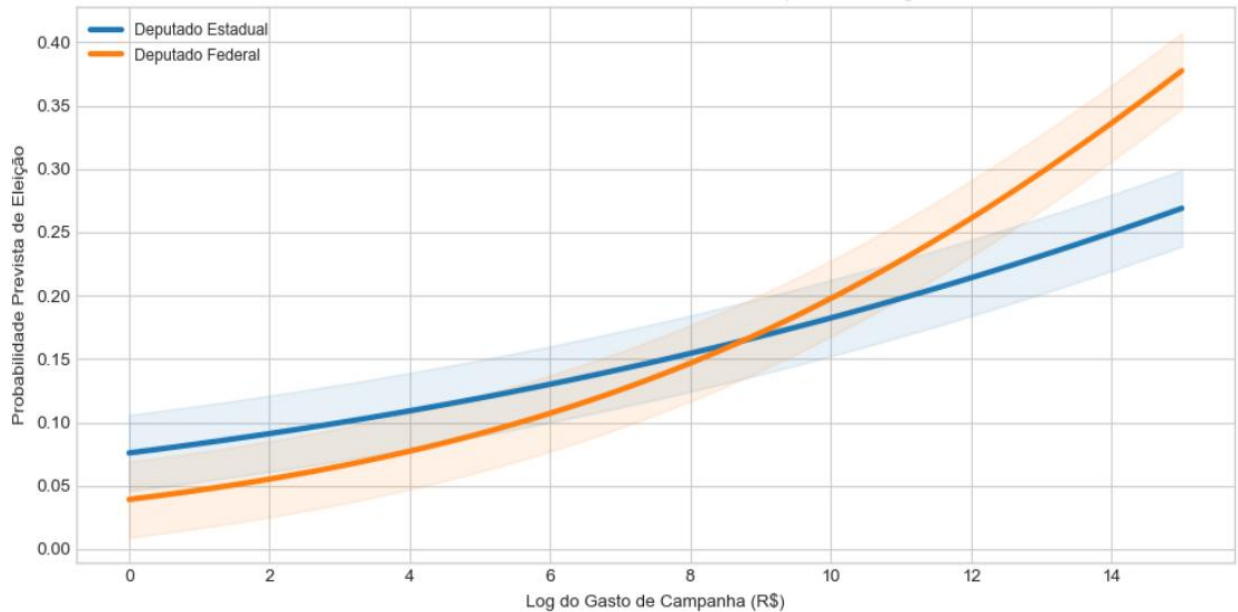
Esse resultado reforça que, embora o investimento digital seja um fator associado ao desempenho eleitoral, ele opera sobre uma estrutura de competição eleitoral que ainda preserva vantagens comparativas para candidatos do gênero masculino.

5.3. Intensidade Financeira e Relação entre Gastos e Probabilidade Estimada

Na análise dos modelos Desagregados, o foco desloca-se da "categoria do gasto" para a "intensidade de cada lançamento". Para Deputado Estadual, a variável de gasto individual (Total Pago) apresenta alta significância ($dy/dx = 0,0085$; $P < 0,01$), evidenciando que a magnitude financeira de cada transação está diretamente associada ao sucesso. No nível federal, contudo, esse impacto isolado perde significância, sugerindo que, na esfera federal, a estratégia de alocação (onde gastar) sobrepõe-se à intensidade bruta (quanto gastar por vez). Embora o modelo desagregado aumente o número de observações, a inferência permanece válida porque a dependência intraindivíduo é tratada por meio de erros-padrão clusterizados por candidato.

Por fim, a relação entre o volume total investido em campanhas digitais e a chance de êxito eleitoral não se apresenta de forma linear. O Gráfico 5 demonstra a probabilidade estimada de vitória seguindo uma curva em "S", característica dos modelos de escolha discreta. Observa-se a existência de um nível inicial de investimento a partir do qual a probabilidade estimada de eleição passa a crescer de forma mais acentuada, seguido por uma fase de retornos crescentes que, em níveis muito elevados de gasto, tende à estabilização, caracterizando retornos marginais decrescentes.

GRÁFICO 5 – Curva de Probabilidade de Vitória por Gasto Digital



Fonte: Elaboração própria com base em dados do TSE (2022).

Dessa forma, os dados sugerem que a eficácia do gasto digital é contingente ao cargo disputado. Enquanto candidatos a deputado estadual apresentam maior probabilidade inicial de eleição, observa-se que o efeito marginal do gasto digital é mais pronunciado para candidatos a deputado federal, especialmente em níveis mais elevados de investimento. Esses achados sugerem uma relação positiva e não linear entre o gasto digital e a probabilidade de vitória eleitoral.

A comparação sistemática entre os modelos Probit e Logit evidencia a estabilidade dos resultados. Em ambas as especificações, variáveis como os gastos com produção de conteúdo para rádio e televisão no nível federal e o total de recursos pagos no nível estadual mantiveram-se como preditores centrais do sucesso eleitoral. Essa convergência estatística fortalece a interpretação de que o impacto dos gastos de campanha, em especial os digitais, segue um padrão estrutural consistente na amostra analisada.

6. CONCLUSÃO

De maneira geral, os resultados indicam que os maiores investimentos em comunicação digital concentraram-se nos estados mais populosos e em partidos com maior capacidade financeira. Observa-se uma predominância de gastos com impulsionamento digital e produção de programas de rádio, televisão ou vídeo, o que evidencia a relevância do capital de campanha para o desempenho eleitoral. Por outro lado, candidatos de partidos menores e de estados com menor população apresentaram investimentos significativamente inferiores, e suas estratégias de comunicação se diversificam conforme o contexto regional e a posição política. Esses padrões sugerem que a alocação de recursos digitais é fortemente influenciada por fatores estruturais, como tamanho do partido e densidade eleitoral, destacando a importância do planejamento estratégico nas campanhas contemporâneas.

Os modelos Probit estimados indicam que as características socioeconômicas dos candidatos, principalmente gênero e escolaridade, exercem impacto muito mais relevante sobre a probabilidade de eleição do que os gastos em comunicação digital. Entre as categorias de gasto, apenas a categoria “Criação e Inclusão de Páginas na Internet” apresenta associação negativa com a probabilidade de êxito eleitoral, sugerindo estratégia potencialmente menos eficiente ou associada a candidaturas menos competitivas. O impulsionamento de conteúdos mostra efeito positivo, embora pequeno e estatisticamente frágil. No geral, a evidência empírica aponta que o simples aumento de recursos investidos em comunicação digital não significa maiores chances de sucesso eleitoral.

Este achado reforça a hipótese de que o investimento digital, quando isolado de uma estrutura de campanha profissionalizada ou focado em ferramentas de baixa interatividade (como sites estáticos), não se traduz em capital político efetivo.

Apesar do cuidado metodológico, a pesquisa apresenta algumas limitações. A primeira refere-se à ausência de dados de engajamento (curtidas, compartilhamentos e comentários) nas redes sociais. Além disso, no processo de tratamento e integração das bases, identificou-se uma discrepância entre os registros de despesas pagas e o cadastro de candidaturas. Algumas despesas presentes no arquivo *despesas_pagas* não possuem correspondência em *despesas_contratadas*, gerando registros sem candidatos vinculados.

Especificamente, foram excluídas da análise econométrica 39 linhas de despesa de candidatos a deputado estadual e distrital (totalizando R\$ 522.745,13) e 29 linhas de candidatos federais (R\$ 1.373.125,28), devido à ausência de informações socioeconômicas completas ou classificações como “NÃO DIVULGÁVEL”. Adicionalmente, observou-se a existência de

gastos significativos que não puderam ser atribuídos a indivíduos específicos, figurando na base como "NÃO INFORMADO" e "SEM PARTIDO". Esse fenômeno concentrou-se majoritariamente no Distrito Federal e em São Paulo, envolvendo montantes expressivos em categorias como Impulsionamento de Conteúdos (mais de R\$ 2 milhões somados entre DF e SP) e Criação de Páginas. Embora esses valores componham o volume total movimentado nas categorias digitais de 2022, a impossibilidade de vinculação ao *SQ_CANDIDATO* exigiu sua remoção da etapa de modelagem, sob o risco de comprometer a precisão das estimativas individuais.

Por fim, ressalta-se que os modelos estimados capturam associações estatísticas, e não causalidade estrita. Isso se deve a potenciais problemas de endogeneidade e variáveis omitidas; afinal, candidatos mais competitivos tendem a atrair mais recursos. Fatores não observados, como força partidária ou capital político prévio, podem influenciar simultaneamente o gasto e o desempenho eleitoral. Ainda assim, a abordagem adotada oferece evidências robustas e interessante sobre o papel dos recursos digitais nas estratégias políticas contemporâneas.

REFERÊNCIAS

- BRASIL. **Lei nº 9.504**, de 30 de setembro de 1997. Estabelece normas para as eleições. Diário Oficial da União, Brasília, DF, 1 out. 1997.
- BRASIL. **Lei nº 13.165**, de 29 de setembro de 2015. Altera as Leis nos 9.504/97, 9.096/95 e 4.737/65 para promover a reforma eleitoral. Diário Oficial da União, Brasília, DF, 30 set. 2015.
- CAMERON, A. Colin; MILLER, Douglas L. A practitioner's guide to cluster-robust inference. **Journal of Human Resources**, Madison, v. 50, n. 2, p. 317–372, 2015. Disponível em: https://cameron.econ.ucdavis.edu/research/Cameron_Miller_JHR_2015_February.pdf. Acesso em: 3 ago. 2025.
- CHADWICK, Andrew. **The Hybrid Media System: Politics and Power**. Oxford: Oxford University Press, 2013.
- GOMES, Wilson. **A democracia digital: história, problemas e temas**. Salvador: EDUFBA, 2018.
- GREENE, William H. **Econometric analysis**. 7. ed. Boston: Pearson, 2012.
- GUJARATI, Damodar N.; PORTER, Dawn C. **Basic econometrics**. 5. ed. New York: McGraw-Hill, 2009.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **PNAD Contínua: Distribuição percentual dos domicílios por existência de utilização da internet no domicílio - 2022**. Rio de Janeiro: IBGE, 2022. Disponível em: <https://www.ibge.gov.br/>. Acesso em: 3 ago. 2025.
- KREISS, Daniel. **Prototype politics: technology-intensive campaigning and the data of democracy**. New York: Oxford University Press, 2016.
- LEVITT, Steven D. Using repeat challengers to estimate the effect of campaign spending on election outcomes in the U.S. House. **Journal of Political Economy**, Chicago, v. 102, n. 4, p. 777–798, 1994.
- LONG, J. Scott; FREESE, Jeremy. **Regression models for categorical dependent variables using Stata**. 3. ed. College Station: Stata Press, 2014.
- MANCUSO, Wagner Pralon; OLIVEIRA, Vanessa Elias de; SPECK, Bruno Wilhelm; HOROCHOVSKI, Rodrigo Rossi. Money in electoral campaigns: the relationship between money and politics as a cause of the judicialization of electoral processes in Brazil. **Social Sciences**, Basel, v. 12, n. 12, p. 656, 2023. Disponível em: <https://doi.org/10.3390/socsci12120656>. Acesso em: 4 mar. 2026.
- SAMUELS, David. Money, elections, and democracy in Brazil. **Latin American Politics and Society**, v. 43, n. 2, p. 27–48, 2001.
- SPECK, Bruno Wilhelm; MANCUSO, Wagner Pralon. A Study on the Impact of Campaign Finance, Political Capital and Gender on Electoral Performance. **Brazilian Political Science**

Review, São Paulo, v. 8, n. 1, p. 34-57, 2014. Disponível em: <https://doi.org/10.1590/1981-38212014000100002>. Acesso em: 4 nov. 2025.

TRIBUNAL SUPERIOR ELEITORAL (TSE). **Candidatos – eleições 2022** (consulta_cand_2022_BRASIL). Dados Abertos TSE, Brasília, DF, 2022. Disponível em: <https://dadosabertos.tse.jus.br/dataset/candidatos-2022>>. Acesso em: 3 ago. 2025.

TRIBUNAL SUPERIOR ELEITORAL (TSE). **Votação por município e zona eleitoral – eleições 2022** (votacao_candidato_munzona_2022_BRASIL). Dados Abertos TSE, Brasília, DF, 2022. Disponível em: <https://dadosabertos.tse.jus.br/dataset/resultados-2022>>. Acesso em: 3 ago. 2025.

TRIBUNAL SUPERIOR ELEITORAL (TSE). **Despesas contratadas de candidatos – eleições 2022** (despesas_contratadas_candidatos_2022_BRASIL). Dados Abertos TSE, Brasília, DF, 2022. Disponível em: <https://dadosabertos.tse.jus.br/dataset/dadosabertos-tse-jus-br-dataset-prestacao-de-contas-eleitorais-2022>>. Acesso em: 3 ago. 2025.

TRIBUNAL SUPERIOR ELEITORAL (TSE). **Despesas pagas de candidatos – eleições 2022** (despesas_pagas_candidatos_2022_BRASIL). Dados Abertos TSE, Brasília, DF, 2022. Disponível em: <https://dadosabertos.tse.jus.br/dataset/dadosabertos-tse-jus-br-dataset-prestacao-de-contas-eleitorais-2022>>. Acesso em: 3 ago. 2025.

WOOLDRIDGE, Jeffrey M. **Econometric analysis of cross section and panel data**. 2. ed. Cambridge: MIT Press, 2010.

HEISS, Florian; BRUNNER, Daniel. **Using Python for Introductory Econometrics**. 2. ed. Düsseldorf, Germany, 2024. Disponível em: <http://www.UPfIE.net>. Acesso em: 3 jan. 2026.

APÊNDICE A – Script em Python para Consolidação das Bases Eleitorais

```
# =====  
# CONSOLIDAÇÃO DAS BASES ELEITORAIS - ELEIÇÕES 2022  
# Despesas, votação, resultado eleitoral e variáveis  
# demográficas dos candidatos  
#  
# Script desenvolvido em Python 3.11  
#  
# Bibliotecas utilizadas:  
# pandas, numpy e openpyxl  
#  
# Bases de dados:  
# - Despesas eleitorais (TSE)  
# - Votação nominal (TSE)  
# - Cadastro de candidatos (TSE)  
# =====  
  
# =====  
# 1. IMPORTAÇÃO DAS BIBLIOTECAS  
# =====  
  
import pandas as pd  
import numpy as np  
  
from openpyxl import Workbook  
from openpyxl.utils.dataframe import dataframe_to_rows  
  
# =====  
# 2. CONFIGURAÇÃO DOS ARQUIVOS  
# =====  
  
caminho_despesas_pagas = (  
    "dados/despesas_pagas_candidatos_2022_BRASIL.csv"  
)  
  
caminho_despesas_contratadas = (  
    "dados/despesas_contratadas_candidatos_2022_BRASIL.csv"  
)  
  
caminho_votacao = (  
    "dados/votacao_candidato_munzona_2022_BRASIL.csv"  
)  
  
caminho_candidatos = (  
    "dados/consulta_cand_2022_BRASIL.csv"  
)  
  
nome_arquivo_excel = (  
    "resultados/Relatorio_Completo_Consolidado_2022.xlsx"  
)
```

```

# Categorias de despesas analisadas
codigos_interesse = [
    20420000,
    20220000,
    20260000,
    20130000
]

# Eleições consideradas
NOMES_ELEICOES_INTERESSE = [
    "Eleição Geral Federal 2022",
    "Eleições Gerais Estaduais 2022"
]

# =====
# 3. FUNÇÃO AUXILIAR
# =====

def formatar_moeda(valor):

    if pd.isna(valor):
        return None

    return (
        f"R$ {valor:,.2f}"
        .replace(",", "X")
        .replace(".", ",")
        .replace("X", ".")
    )

# =====
# 4. IMPORTAÇÃO DAS BASES DE DADOS
# =====

def importar_base(caminho):

    df = pd.read_csv(
        caminho,
        sep=';',
        encoding='latin1',

        dtype={
            'VR_PAGTO_DESPESA': str,
            'VR_DESPESA_CONTRATADA': str,
            'DS_CARGO': object,
            'DS_ORIGEM_DESPESA': object,
            'SQ_CANDIDATO': object,
            'CD_CARGO': object,
            'DS_ELEICAO': object
        }
    )

```

```

# Padronização de valores ausentes
for col in [
    'SG_UF',
    'SG_UE',
    'SG_UF_FORNECEDOR',
    'SG_PARTIDO'
]:

    if col in df.columns:

        df[col] = (
            df[col]
            .replace({
                '': np.nan,
                'BR': np.nan,
                '#NULO#': np.nan
            })
            .astype('object')
        )

    return df

base_pagas = importar_base(caminho_despesas_pagas)

base_contratadas = importar_base(
    caminho_despesas_contratadas
)

# =====
# 5. TRATAMENTO DAS DESPESAS ELEITORAIS
# =====

base_pagas_filtrada = (

    base_pagas

    .query(
        "DS_ELEICAO.isin(@NOMES_ELEICOES_INTERESSE)"
    )

    .assign(

        Valor_Pago=lambda df:

            df['VR_PAGTO_DESPESA']
            .str.replace(',', '.', regex=False)
            .astype(float)
    )

    .query(
        "CD_ORIGEM_DESPESA.isin(@codigos_interesse)"
    )
)

```

```

        .filter(
            items=[
                'SQ_DESPESA',
                'DS_ELEICAO',
                'SG_UF',
                'Valor_Pago',
                'DS_ORIGEM_DESPESA'
            ]
        )
    )

base_contratadas_info = (
    base_contratadas
        .query(
            "DS_ELEICAO.isin(@MOMES_ELEICOES_INTERESSE)"
        )
        .query(
            "CD_ORIGEM_DESPESA.isin(@codigos_interesse)"
        )
        .groupby(
            'SQ_DESPESA',
            as_index=False
        )
        .agg({
            'SQ_CANDIDATO': 'first',
            'CD_CARGO': 'first',
            'DS_CARGO': 'first',
            'DS_ORIGEM_DESPESA': 'first',
            'SG_PARTIDO': 'first',
            'SG_UE': 'first',
            'SG_UF_FORNECEDOR': 'first'
        })
    )

# =====
# 6. CONSOLIDAÇÃO DAS BASES
# =====

base_pagas_filtrada['SQ_DESPESA'] = (
    base_pagas_filtrada['SQ_DESPESA']
        .astype(str)
    )

base_contratadas_info['SQ_DESPESA'] = (
    base_contratadas_info['SQ_DESPESA']
        .astype(str)
    )

base_mestra_final = base_pagas_filtrada.merge(

```

```

base_contratadas_info,

on='SQ_DESPESA',

how='left',

suffixes=('', '_contratada')
)

# Preenchimento de categorias ausentes
base_mestra_final['DS_ORIGEM_DESPESA'] = (

    base_mestra_final['DS_ORIGEM_DESPESA']

    .fillna(
        base_mestra_final[
            'DS_ORIGEM_DESPESA_contratada'
        ]
    )
)

base_mestra_final.drop(

    columns=['DS_ORIGEM_DESPESA_contratada'],

    inplace=True
)

# Padronização de valores faltantes
base_mestra_final = base_mestra_final.fillna({

    'DS_CARGO': 'NÃO INFORMADO',

    'CD_CARGO': 'NÃO INFORMADO',

    'SQ_CANDIDATO': 'NÃO INFORMADO',

    'SG_PARTIDO': 'SEM PARTIDO',

    'SG_UE': 'BR',

    'SG_UF_FORNECEDOR': 'BR'
})

# =====
# 7. DEFINIÇÃO DAS REGIÕES BRASILEIRAS
# =====

base_mestra_final['SG_UF_FINAL'] = (

    base_mestra_final['SG_UF']

    .fillna(base_mestra_final['SG_UE'])

    .fillna(
        base_mestra_final['SG_UF_FORNECEDOR']
    )
)

```

```

    )
    .fillna('BR')
)

base_mestra_final['Regiao'] = np.select(
    [
        base_mestra_final['SG_UF_FINAL'].isin([
            'RO', 'AC', 'AM', 'RR', 'PA', 'AP', 'TO'
        ]),
        base_mestra_final['SG_UF_FINAL'].isin([
            'MA', 'PI', 'CE', 'RN', 'PB',
            'PE', 'AL', 'SE', 'BA'
        ]),
        base_mestra_final['SG_UF_FINAL'].isin([
            'MS', 'MT', 'GO', 'DF'
        ]),
        base_mestra_final['SG_UF_FINAL'].isin([
            'MG', 'ES', 'RJ', 'SP'
        ]),
        base_mestra_final['SG_UF_FINAL'].isin([
            'PR', 'SC', 'RS'
        ]),
        base_mestra_final['SG_UF_FINAL'] == 'BR'
    ],
    [
        "Norte",
        "Nordeste",
        "Centro-Oeste",
        "Sudeste",
        "Sul",
        "Nacional"
    ],
    default="Outra / Não Identificada"
)

# =====
# 8. AGRUPAMENTO DOS GASTOS ELEITORAIS
# =====

relatorio_gastos = (
    base_mestra_final
    .groupby(

```

```

        [
            'DS_ELEICAO',
            'Regiao',
            'SQ_CANDIDATO',
            'CD_CARGO',
            'DS_CARGO',
            'DS_ORIGEM_DESPESA',
            'SG_PARTIDO',
            'SG_UF_FINAL'
        ],

        dropna=False
    )

    .agg(
        Total_Pago=('Valor_Pago', 'sum')
    )

    .reset_index()

    .assign(

        Total_Pago_F=lambda df:

            df['Total_Pago']
            .apply(formatar_moeda)
    )
)

# =====
# 9. IMPORTAÇÃO DA VOTAÇÃO NOMINAL
# =====

def importar_votacao_resumida(caminho):

    votos_por_candidato = {}

    for chunk in pd.read_csv(

        caminho,

        sep=';',

        encoding='latin1',

        dtype={
            'SQ_CANDIDATO': str,
            'QT_VOTOS_NOMINAIS': 'Int64'
        },

        usecols=[
            'SQ_CANDIDATO',
            'QT_VOTOS_NOMINAIS'
        ],

        chunksize=500_000

    ):

```

```

agrupado = (
    chunk
    .groupby('SQ_CANDIDATO')
    ['QT_VOTOS_NOMINAIS']
    .sum()
)

for candidato, votos in agrupado.items():

    votos_por_candidato[candidato] = (

        votos_por_candidato.get(
            candidato,
            0
        ) + votos
    )

return pd.DataFrame(

    list(votos_por_candidato.items()),

    columns=[
        'SQ_CANDIDATO',
        'Total_Votos'
    ]
)

base_votacao_total = importar_votacao_resumida(
    caminho_votacao
)

# =====
# 10. IMPORTAÇÃO DAS VARIÁVEIS DEMOGRÁFICAS
# =====

def importar_variaveis_candidatos(caminho):

    return (

        pd.read_csv(

            caminho,

            sep=';',

            encoding='latin1',

            usecols=[
                'SQ_CANDIDATO',
                'DS_GENERO',
                'DS_GRAU_INSTRUCAO',
                'DS_ESTADO_CIVIL',
                'DS_COR_RACA',
                'DS_OCUPACAO'
            ],

```

```

        dtype={
            'SQ_CANDIDATO': str
        }
    )

    .drop_duplicates(
        subset='SQ_CANDIDATO'
    )
)

base_candidatos_info = (

    importar_variaveis_candidatos(
        caminho_candidatos
    )
)

# =====
# 11. IMPORTAÇÃO DO RESULTADO ELEITORAL
# =====

def importar_resultado_eleicao(caminho):

    return (

        pd.read_csv(

            caminho,

            sep=';',

            encoding='latin1',

            usecols=[
                'SQ_CANDIDATO',
                'DS_SIT_TOT_TURNO'
            ],

            dtype={
                'SQ_CANDIDATO': str,
                'DS_SIT_TOT_TURNO': object
            }
        )

        .drop_duplicates(
            subset='SQ_CANDIDATO'
        )
    )

base_resultado = importar_resultado_eleicao(
    caminho_candidatos
)

```

```

# =====
# 12. CONSOLIDAÇÃO FINAL DOS DADOS
# =====

relatorio_completo_consolidado = (

    relatorio_gastos

    .merge(
        base_votacao_total,
        on='SQ_CANDIDATO',
        how='left'
    )

    .merge(
        base_resultado,
        on='SQ_CANDIDATO',
        how='left'
    )

    .merge(
        base_candidatos_info,
        on='SQ_CANDIDATO',
        how='left'
    )

    .assign(

        Total_Votos=lambda df:

        df['Total_Votos']
        .fillna(0)
        .astype(int),

        Resultado_Eleicao=lambda df:

        df['DS_SIT_TOT_TURNO']
        .fillna("NÃO INFORMADO"),

        Custo_por_Voto=lambda df:

        np.where(

            df['Total_Votos'] > 0,

            df['Total_Pago'] /
            df['Total_Votos'],

            0
        ),

        Custo_por_Voto_F=lambda df:

        df['Custo_por_Voto'].apply(

            lambda x:

            (

                f"R$ {x:,.2f}"
            )
        )
    )
)

```

```

        .replace(",","X")
        .replace(".",",")
        .replace("X",".")
    )
    if x > 0 else "N/A"
)
)
.drop(columns=['DS_SIT_TOT_TURNO'])
)

# =====
# 13. EXPORTAÇÃO DOS RESULTADOS
# =====

def exportar_para_excel(dataframes, filename):

    workbook = Workbook()

    if 'Sheet' in workbook.sheetnames:
        workbook.remove(workbook['Sheet'])

    for sheet_name, df in dataframes.items():

        ws = workbook.create_sheet(
            title=sheet_name
        )

        for row in dataframe_to_rows(
            df,
            header=True,
            index=False
        ):

            ws.append(row)

    workbook.save(filename)

abas_excel = {

    "Gastos_Estaduais_Com_Votos":

        relatorio_completo_consolidado.query(
            "DS_ELEICAO==@NOMES_ELEICOES_INTERESSE[1]"
        ),

    "Gastos_Federais_Com_Votos":

        relatorio_completo_consolidado.query(
            "DS_ELEICAO==@NOMES_ELEICOES_INTERESSE[0]"
        )
}

exportar_para_excel(
    abas_excel,
    nome_arquivo_excel
)

```

APÊNDICE B – Efeitos Marginais dos Modelos Logit para Probabilidade de Eleição

Variáveis	Estadual Agregado	Estadual Desagregado	Federal Agregado	Federal Desagregado
Criação e Inclusão de Páginas	-0.0039 (0.001)***	—	-0.0020 (0.001)	—
Impulsionamento de Conteúdos	0.0022 (0.001)*	—	-0.0027 (0.001)*	—
Total Pago (Gasto Individual)	—	0.0085 (0.002)***	—	0.0031 (0.002)
Produção de programas de rádio, TV ou vídeo	0.0009 (0.001)	—	0.0050 (0.001)***	—
Publicidade por jornais e revistas	0.0008 (0.001)	—	-0.0070 (0.002)***	—
Gênero (Masculino)	0.0871 (0.013)***	0.0919 (0.014)***	0.0548 (0.015)***	0.0335 (0.017)*
Raça (Indígena)	0.0097 (0.116)	0.0371 (0.125)	0.2822 (0.149)*	0.3931 (0.162)**
Observações (N)	6.570	10.240	5.023	8.526
Pseudo R ²	0,0132	0,0141	0,0135	0,0061

Nota: Erros-padrão entre parênteses. *** p < 0,01; ** p < 0,05; * p < 0,10.