

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO SOCIOECONÔMICO - CSE  
DEPARTAMENTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS - CNM  
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

Alessandro Silvério Jonsson

**Proposta de um Índice de Preços Online para Alimentos Básicos nas Mesorregiões de  
Santa Catarina**

Florianópolis

2024

Alessandro Silvério Jonsson

**Proposta de um Índice de Preços Online para Alimentos Básicos nas Mesorregiões de Santa Catarina**

Trabalho de Conclusão de Curso submetido ao curso de Graduação em Ciências Econômicas do Departamento de Economia e Relações Internacionais da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Bacharel em Ciências Econômicas.

Orientador(a): Prof. Arlei Luiz Fachinello, Dr.

Florianópolis

2024

Silvério Jonsson, Alessandro  
Proposta de um Índice de Preços Online  
para Alimentos Básicos nas Mesorregiões de  
Santa Catarina / Alessandro Silvério  
Jonsson ; orientadora, Arlei Luiz  
Fachinello, 2024.  
53 p.

Trabalho de Conclusão de Curso  
(graduação) - Universidade Federal de Santa  
Catarina, Centro Socioeconômico, Graduação  
em Ciências Econômicas, Florianópolis,  
2024.

Inclui referências.

1. Ciências Econômicas. 2. Inflação de  
Alimentos. 3. Web Scrapping. 4. Santa  
Catarina. I. Luiz Fachinello, Arlei. II.  
Universidade Federal de Santa Catarina.  
Graduação em Ciências Econômicas. III

Alessandro Silvério Jonsson

**Proposta de um Índice de Preços Online para Alimentos Básicos nas Mesorregiões de Santa Catarina**

Florianópolis, 12 de dezembro de 2024.

O presente Trabalho de Conclusão de Curso foi avaliado e aprovado pela banca examinadora composta pelos seguintes membros:

Prof.(a) Arlei Luiz Fachinello, Dr.

Universidade Federal de Santa Catarina

Prof.(a) Gueibi Peres Souza, Dr.

Universidade Federal de Santa Catarina

Prof.(a) Helberte João França Almeida, Dr.

Universidade Federal de Santa Catarina

Certifico que esta é a **versão original e final** do Trabalho de Conclusão de Curso que foi julgado adequado para obtenção do título de Bacharel em Economia por mim e pelos demais membros da banca examinadora.

---

Prof.(a) Arlei Luiz Fachinello, Dr.

Orientador

Florianópolis, 2024

## **AGRADECIMENTOS**

Agradeço à minha família, pela dedicação e suporte por tornar possível o meu acesso ao ensino superior, em uma prestigiada Universidade Federal. Agradeço também a Gabriela de Lima, por ter me motivado a continuar, apesar das dificuldades. Ao meu orientador, por ter me auxiliado a desenvolver esse projeto, e a Sociedade Brasileira, por tornar a Universidade Pública possível.

## RESUMO

Este trabalho teve como objetivo desenvolver um índice de preços para alimentos básicos nas mesorregiões de Santa Catarina, utilizando técnicas de web scraping para coleta de dados em supermercados online, entre janeiro e abril de 2024. O índice foi calculado com base na fórmula de Laspeyres e os resultados indicaram um aumento geral de 2,5% nos preços de alimentos no estado durante o período analisado, com variações regionais significativas. Embora o método automatizado tenha se mostrado eficiente e de baixo custo, a análise foi limitada pela ausência de dados para as regiões Oeste e Serrana, bem como pela exclusão de pequenos estabelecimentos sem presença online. A comparação com índices tradicionais, como o ICV da UDESC, evidenciou similaridades na dinâmica de preços, reforçando a representatividade parcial do índice desenvolvido. Conclui-se que o uso de dados online apresenta potencial como ferramenta de monitoramento de preços, especialmente com a crescente digitalização do comércio, mas melhorias na abrangência e representatividade são necessárias para ampliar sua aplicabilidade em estudos econômicos regionais.

**Palavras-chave:** Índice de Preços; Web Scraping; Inflação; Alimentos; Santa Catarina.

## **ABSTRACT**

This study aimed to develop a price index for basic food products in the mesoregions of Santa Catarina using web scraping techniques to collect data from online supermarkets between January and April 2024. The index was calculated based on the Laspeyres formula, and the results indicated a general increase of 2.5% in food prices in the state during the analyzed period, with significant regional variations. Although the automated method proved efficient and cost-effective, the analysis was limited by the absence of data for the Western and Serrana regions, as well as the exclusion of small establishments without an online presence. Comparison with traditional indices, such as UDESC's ICV, highlighted similarities in price dynamics, reinforcing the partial representativeness of the developed index. It is concluded that using online data shows potential as a price monitoring tool, especially with the increasing digitization of commerce, but improvements in scope and representativeness are necessary to enhance its applicability in regional economic studies.

**Keywords:** Price Index; Web Scraping; Inflation; Food; Santa Catarina.

## LISTA DE FIGURAS

Figura 1 – Índice Geral de Preços de Alimentos (Janeiro à Abril de 2024).....	35
Figura 2 – Variação Mensal do Índice Geral de Preços de Santa Catarina e Paraná.....	36
Figura 3 – Índice de Preços por Mesorregião.....	37
Figura 4 – Comparação da variação percentual mensal do ICV de alimentos e o índice elaborado.....	38
Figura 5 – Índice de Preços por Produto para o estado de Santa Catarina.....	39
Figura 6 – Índice por produto para região Norte.....	40
Figura 7 – Índice por Produto Para a Região do Vale do Itajaí.....	41
Figura 8 – Índice por produto para a região de Grande Florianópolis.....	41
Figura 9 – Índice por produto para a região do Sul Catarinense.....	42
Figura 10 – Índice por produto para a região do Sul Catarinense.....	43



## LISTA DE TABELAS

Tabela 1 – Produtos que compõem a cesta básica, quantidade consumida mensal e o peso do produto na cesta.....	32
Tabela 2 – Distribuição de estabelecimentos selecionados por região e leituras realizadas....	33
Tabela 3 – Variação da Média de Preços por Categoria - Região Norte.....	50
Tabela 4 – Variação da Média de Preços por Categoria - Região Vale do Itajaí.....	51
Tabela 5 – Variação da Média de Preços por Categoria - Grande Florianópolis.....	52
Tabela 6 – Variação da Média de Preços por Categoria - Sul.....	53

## LISTA DE ABREVIATURAS E SIGLAS

IBGE	Instituto Brasileiro de Geografia e Estatística
FGV	Fundação Getulio Vargas
IPCA	Índice de preços ao consumidor
INPC	Índice Nacional de Preços ao Consumidor
BPP	The Billion Prices Project
MIT	Massachusetts Institute of Technology
POF	Pesquisa de Orçamento Familiar
DIEESE	Departamento Intersindical de Estatística e Estudos Socioeconômicos
IGP-M	Índice Geral de Preços - Mercado
IPC	Índices de Preços ao Consumidor
COICOP6	Classification Of Individual Consumption by Purpose
INCC-M	Índice Nacional de Custo da Construção - Mercado
IPA-M	Índice de Preços ao Produtor Amplo - Mercado
PLC	Pesquisa de Locais de Compra
PEPS	Pesquisa de Especificação de Produtos e Serviços

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>16</b>
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>18</b>
2.1	Índices de Inflação no Brasil	19
2.1.1	IPCA (Índice Nacional de Preços ao Consumidor Amplo)	19
2.1.2	Índice Geral de Preços – Mercado	20
2.1.2.1	IPA-M	21
2.1.2.2	IPC-M	22
2.2	Índice de Custo de Vida de Florianópolis	23
2.3	Índice de Preços Regional do Paraná - IPR	24
2.5.	<i>Web Crawlers</i> e Ciências Sociais Aplicadas	25
2.5.1	The Billion Prices Project	26
2.5.2	Índice de Preços ao Consumidor Web (IPC-W)	27
2.5.3	Índice de Preços ao Consumidor Norueguês	28
<b>3</b>	<b>PROCEDIMENTOS METODOLÓGICOS ADOTADOS</b>	<b>31</b>
3.1	Cálculo do Índice de Preços	31
3.2	Cesta Utilizada	32
3.3	Coleta e tratamento de dados	33
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>35</b>
4.1	Índice Geral de Preços	35
4.2	Índice de Preços Por Produto	38
4.3	Limitações do Método	43
<b>5</b>	<b>CONCLUSÃO</b>	<b>45</b>
	<b>REFERÊNCIAS</b>	<b>47</b>
	<b>ANEXO I – Variação da Média de Preços por Categoria - Região Norte</b>	<b>50</b>
	<b>ANEXO II – Variação da Média de Preços por Categoria - Região Vale do Itajaí</b>	<b>51</b>
	<b>ANEXO III – Variação da Média de Preços por Categoria - Região Grande Florianópolis</b>	<b>52</b>
	<b>ANEXO IV – Variação da Média de Preços por Categoria - Região Sul</b>	<b>54</b>

## 1 INTRODUÇÃO

A estabilidade de preços é um dos principais objetivos da política monetária de um país, sendo essencial para a estabilidade da economia, uma vez que a inflação corrói o poder de compra da população, prejudicando o seu bem-estar. Diante disso, é de interesse para a sociedade, e especialmente para os formuladores de políticas públicas, que existam meios eficientes de mensurar a inflação para tornar possível o desenvolvimento de políticas de controle do nível de preços, de forma a garantir o bem-estar da população (Banco do Brasil, 2024).

No contexto da segurança alimentar, o aumento de preços de alimentos tende a prejudicar as famílias mais pobres, uma vez que os gastos com alimentos representam uma parcela considerável da renda dessas famílias, e, portanto, o aumento de preços leva à diminuição do seu poder de compra e à redução de seus rendimentos reais. Esse problema piora quanto menor for a renda das famílias, visto que quanto menor a renda, maior será a participação dos gastos em alimentação na renda total das famílias (Souza, 2012).

Com esse objetivo foram criados diversos índices de preços que são utilizados como base para a criação e manutenção de políticas públicas, entre elas o reajuste do salário mínimo e de programas previdenciários, a definição de metas de inflação e taxa de juros, e reajuste de contratos (Mendonça, 2007).

No contexto brasileiro, o levantamento de preços começou a ser realizado a partir da década de 1920, enquanto o primeiro índice de custo de vida remonta a década de 1940, visando atender a legislação do salário mínimo (Lei 2.162/40), o qual após algumas modificações, deu origem ao Índice Nacional de Preços ao Consumidor (INPC), que atualmente é gerido pelo Instituto Brasileiro de Geografia e Estatística (IBGE). A partir da implementação do INPC, a Fundação Getúlio Vargas (FGV) também passou a participar da formulação de indicadores de preços, utilizando como base os dados da Pesquisa de Orçamento Familiar (POF) realizada pelo IBGE. Em geral, os índices desenvolvidos por estas instituições seguem metodologias que variam de acordo com o objetivo de cada índice (Oliveira *et. al.*, 2010).

O índice de inflação oficial do Brasil é o Índice de Preços ao Consumidor Amplo (IPCA), o qual é gerido pelo IBGE, sendo este um índice fundamental para que as instituições possam diagnosticar a dinâmica do poder de compra da população. Além do IPCA, existem outros índices que são utilizados em diferentes contextos, como o Índice Geral de Preços do Mercado (IGP-M), e o INPC.

Por mais que estes índices sejam amplamente utilizados como referência para inflação, eles possuem algumas limitações quando se trata em comparar níveis de preço nas mesorregiões do estado de Santa Catarina. Isso acontece porque os índices de inflação são calculados levando-se em consideração a variação no nível de preços nas principais regiões metropolitanas de alguns estados brasileiros, desconsiderando as demais regiões. Um dos problemas dessa abordagem é que ela não possibilita diagnosticar com clareza a dinâmica da inflação nas mesorregiões catarinenses.

A partir da problemática exposta, surge a necessidade de desenvolver um índice que possa captar as variações no nível de preços de alimentos básicos de forma eficiente para diferentes mesorregiões do estado, uma vez que a dinâmica inflacionária pode variar dependendo dos padrões de consumo de cada região.

A construção e manutenção de um índice de preços envolve um esforço considerável, principalmente no que tange a coleta de dados. Nos índices tradicionais, a coleta de preços consiste na utilização de pesquisadores que vão até estabelecimentos pré-selecionados, e fazem entrevistas para coletar os dados necessários para se calcular os índices. Essa metodologia pode variar em alguns casos, em que são utilizados informantes que devem responder formulários periodicamente.

Nesse contexto, a introdução de técnicas de coleta de dados automatizada de preços praticados online, se mostram como uma alternativa eficiente para a resolução desses problemas (Cavallo, 2016).

O aumento significativo da acessibilidade e das condições de acesso à internet impulsionou a popularização da rede e, conseqüentemente, o crescimento do número de consumidores de lojas online. Esse cenário favoreceu a digitalização e indexação dos dados, tornando as técnicas de coleta de informações relevantes para o cálculo de índices de inflação, uma vez que essas técnicas em geral podem ser mais baratas e eficientes do que as formas tradicionais.

Portanto, tendo em vista a importância da quantificação dos preços de alimentos básicos para o bem-estar da sociedade, e as limitações dos índices tradicionais de inflação, o presente trabalho tem como objetivo desenvolver um índice de preços para itens básicos de alimentação, através da aplicação de técnicas de coleta automatizada de preços praticados por lojas virtuais de supermercados de Santa Catarina, uma vez que essas técnicas viabilizam a construção e manutenção do índice, devido ao seu baixo custo de manutenção (Cavallo, 2016).

A partir desse índice, espera-se estimar a inflação de alimentos agregada para o estado de Santa Catarina, assim como a inflação de alimentos para as mesorregiões do estado. Além disso, espera-se contribuir para a melhora da capacidade analítica da ciência econômica, investigando casos de uso de tecnologias de coleta de dados, e como essas técnicas podem ser aplicadas no contexto das ciências sociais, a fim de criar novos conhecimentos.

O presente estudo foi dividido em quatro capítulos. No primeiro capítulo, é apresentada a introdução ao tema, contextualizando a importância da estabilidade de preços, especialmente em relação à segurança alimentar, e a necessidade de novos índices de preços regionais mais eficientes para mensurar a inflação em Santa Catarina.

O segundo capítulo abrange o desenvolvimento teórico do estudo, com destaque para a revisão da literatura sobre índices de inflação no Brasil, a metodologia tradicional de coleta de dados de preços, e a relevância do comércio eletrônico como fonte de dados para pesquisas econômicas. Também são abordados os web crawlers e suas aplicações em pesquisas de preços online, com exemplos de projetos internacionais que utilizam essas tecnologias.

O terceiro capítulo descreve os procedimentos metodológicos adotados, detalhando o cálculo do índice de preços com base na metodologia de Laspeyres, a definição da cesta básica utilizada, a coleta de dados por meio de robôs desenvolvidos para acessar sites de supermercados, e o tratamento dos dados coletados.

O quarto capítulo apresenta os resultados obtidos, discutindo a evolução do índice de preços de alimentos em Santa Catarina e suas mesorregiões, bem como a inflação por categoria de produto. Além disso, são discutidas as limitações da pesquisa e sugestões para estudos futuros.

## **2 REVISÃO DA LITERATURA**

Os índices econômicos foram originados a partir da necessidade de quantificar aspectos importantes para o desenvolvimento e estabilidade de uma economia, e foram se materializando à medida que se tornou possível coletar informações para permitir o seu cálculo. Em um primeiro momento, a metodologia adotada nos índices de preços não era muito rigorosa, devido às grandes limitações para se coletar, calcular e processar os dados obtidos, de modo que os indicadores utilizados eram aqueles que eram possíveis de serem realizados, independente de serem “ideais”, do ponto de vista metodológico ou não (Souza, 2004).

O primeiro índice adotado foi o cálculo do Índice de Custo de Vida em 1920, com dados retroativos a 1912. Esse índice foi elaborado com base no padrão de consumo de uma única família de classe média alta, e os preços eram coletados diretamente pela própria família. Publicado pelo Ministério da Fazenda até 1939, o Índice de Custo de Vida foi precursor dos Índices de Preços ao Consumidor (IPC), que passaram a ser calculados de forma mais ampla pelo FGV a partir de 1946, com a inclusão de diferentes classes de renda e regiões metropolitanas. Esse desenvolvimento histórico marcou o início de uma trajetória contínua de aperfeiçoamento das metodologias de cálculo de índices de preços no Brasil, que evoluíram conforme a economia do país se tornou mais complexa.

Com o passar do tempo, as metodologias adotadas e a capacidade de coletar e processar dados se desenvolveram. Esse processo, naturalmente, levou a transformação de índices que foram sendo modificados, revisados ou substituídos ao longo do tempo, em um processo quase que “darwiniano” (Souza, 2004).

### **2.1 Índices de Inflação no Brasil**

#### **2.1.1. IPCA (Índice Nacional de Preços ao Consumidor Amplo)**

O IPCA (Índice de Preços ao Consumidor Amplo) tem como objetivo medir a variação dos preços de um conjunto de produtos e serviços comercializados no varejo, voltados para o consumo pessoal de famílias com rendimentos mensais entre 1 e 40 salários mínimos. Essa faixa de renda foi estabelecida para abranger 90% das famílias residentes em áreas urbanas, cobertas pelo Sistema Nacional de Índices de Preços ao Consumidor (SNIPC) (IBGE, s.d).

A metodologia do IPCA abrange a construção de uma estrutura de ponderadores, a definição dos produtos e locais de coleta, e o método de cálculo. A estrutura de pesos é organizada em grupos, subgrupos, itens e subitens, sendo esses últimos o nível mais detalhado e para os quais são calculados os pesos que refletem a importância de cada subitem na cesta de consumo das famílias, com base nos hábitos da população-alvo (IBGE, s.d).

A coleta de preços é feita a partir de um cadastro de informantes, obtido por dois procedimentos: o levantamento de locais de compra por meio da Pesquisa de Locais de Compra (PLC), que define onde coletar a maioria dos subitens; e um segundo procedimento específico para itens que exigem tratamento diferenciado, os chamados subitens extra-PLC, como aluguel de moradia, serviços públicos, entre outros. O cadastro de informantes inclui estabelecimentos comerciais, prestadores de serviços, concessionárias e profissionais autônomos. Para definir os produtos a serem pesquisados, considera-se a representatividade dos produtos consumidos pela população-alvo. Uma Pesquisa de Especificação de Produtos e Serviços (PEPS) é realizada para determinar o nível de especificação dos itens incluídos na coleta de preços (IBGE, s.d).

Para se calcular o índice de cada subitem, primeiramente calcula-se a média aritmética dos preços coletados em diferentes estabelecimentos para o produto no mês corrente, comparando com o mês anterior. Em seguida, aplica-se a média geométrica simples para agregação dos preços dos produtos de cada subitem. Para os níveis superiores, utiliza-se a fórmula de Laspeyres, ponderando os relativos de preços com base nas despesas obtidas pela POF. O índice nacional é uma média aritmética ponderada dos índices regionais de dez regiões metropolitanas, além de Goiânia, Campo Grande, Rio Branco, São Luís, Aracaju e Brasília. O ponderador utilizado é o rendimento total urbano com base na POF. A série histórica do índice tem base em dezembro de 1993, com valor de referência 100 (IBGE, s.d).

### **2.1.2. Índice Geral de Preços – Mercado**

O Índice Geral de Preços – Mercado (IGP-M) é um índice econômico calculado pelo Instituto Brasileiro de Economia (IBRE) da FGV. Esse índice teve início no ano de 1989, através de um contrato firmado entre a Confederação Nacional das Instituições Financeiras (CNF) e a FGV. Este índice tem como principal objetivo medir a variação dos preços em



diferentes setores da economia brasileira, desempenhando um papel fundamental na economia brasileira, uma vez que é amplamente utilizado como referência para correções de preços e ajustes contratuais. (FGV, 2019).

O IGP-M é baseado na estrutura metodológica do Índice Geral de Preços – Disponibilidade Interna (IGP-DI) e é composto por três subíndices: o Índice de Preços ao Produtor Amplo (IPA-M), o Índice de Preços ao Consumidor (IPC-M) e o Índice Nacional de Custo da Construção (INCC-M). Esses subíndices possuem pesos diferentes na composição final do IGP-M. O IPA-M, que reflete a variação dos preços de produtos agropecuários e industriais no nível do produtor, tem um peso de 60%. O IPC-M, que mede as variações de preços ao consumidor final, abrangendo bens e serviços destinados ao consumo das famílias, representa 30% do índice. O INCC-M, que avalia as variações nos custos da construção civil, com foco em insumos e mão-de-obra, tem peso de 10%. Esta combinação de índices aumenta a robustez do IGP-M, melhorando a sua representatividade na determinação da inflação (FGV, 2019).

O cálculo do IGP-M é feito a partir de uma média ponderada dos três subíndices mencionados. A fórmula de cálculo do índice é exibida na Equação 1:

$$IGP_m = 0.6 * IPAm + 0.3 * IPCm + 0.1 * INCCm \quad (1)$$

Onde:

*IGP<sub>m</sub>* é o índice final, e os outros componentes correspondem aos valores de cada subíndice no período de referência

*IPAm* é valor do o Índice de Preços ao Produtor Amplo

*IPC<sub>m</sub>* é Índice de Preços ao Consumidor

*INCC<sub>m</sub>* é Índice Nacional de Custo da Construção

### **2.1.2.1. IPA-M**

O IPA-M tem como principal objetivo quantificar as variações de preços das mercadorias no mercado de produção, tanto no setor agrícola quanto industrial.

No setor agrícola, o índice abrange produtos de três classes de atividades de acordo com sistema do CNAE, sendo essas, as lavouras temporárias, lavouras permanentes e pecuária. Destas classes, são selecionados os produtos com maior valor de produção, de acordo com as pesquisas Produção Agrícola Municipal (PAM) e Produção da Pecuária Municipal (PPM), do IBGE, sendo considerados 25 produtos para realizar o cálculo. No setor industrial, a seleção se baseia na Pesquisa Industrial Anual (PIA), sendo considerados 343 produtos, que representam atividades econômicas (FGV, 2016).

A coleta dos preços é feita sistematicamente em regiões-chave do país. Para produtos industriais, a coleta é mensal diretamente nas empresas produtoras, enquanto para produtos agropecuários é realizada diariamente por meio de diversas fontes, como cooperativas e centrais de abastecimento (FGV, 2016).

O IPA-M utiliza um sistema de ponderação que atribui diferentes pesos a cada categoria de produto com base em sua importância econômica, como o valor adicionado bruto de cada setor. O índice é calculado utilizando a fórmula de Laspeyres encadeada de base móvel, que permite comparar os preços atuais com os de períodos anteriores (FGV, 2016).

### **2.1.2.2. IPC-M**

A cesta de produtos do IPC-M é composta por 338 subitens, que são selecionados com base na POF realizada pelo IBGE. Esses itens representam os hábitos de consumo das famílias em sete capitais brasileiras: Belo Horizonte, Brasília, Porto Alegre, Recife, Rio de Janeiro, Salvador e São Paulo (FGV, 2016).

Este índice utiliza um sistema de pesos para atribuir a importância relativa de cada item dentro do índice. Esses pesos são derivados da POF, que reflete o gasto das famílias em diferentes categorias de consumo, como alimentação, habitação, vestuário, saúde, entre outras. A coleta de preços ocorre de forma sistemática entre os dias 21 do mês anterior ao dia

20 do mês de referência, abrangendo tanto bens quanto serviços. Os preços são levantados em estabelecimentos comerciais e de prestação de serviços nas capitais mencionadas (FGV, 2016).

O IPC-M é calculado utilizando a fórmula de Laspeyres, com base em preços coletados mensalmente. Ele é divulgado em três apurações ao longo do mês, conhecidas como decêndios. Cada decêndio reflete uma parte do mês, sendo o terceiro e último apurado o índice definitivo (FGV, 2016).

### **2.1.2.3. INCC-M**

O INCC-M tem como objetivo quantificar as variações nos preços de materiais, mão-de-obra e serviços relacionados à construção civil, com abrangência em sete regiões metropolitanas: Belo Horizonte, Brasília, Porto Alegre, Recife, Salvador, Rio de Janeiro e São Paulo. Este índice é utilizado como referência para ajustes de contratos no setor de construção civil. A pesquisa é realizada mensalmente entre os dias 21 do mês anterior e 20 do mês de referência, com coleta em diversos estabelecimentos relacionados ao setor de construção (FGV, 2016).

Para o cálculo desse índice, inicialmente, calcula-se o relativo de preços para cada insumo em cada região, comparando o preço atual com o preço do mês anterior. Os relativos de preços de cada insumo são agregados por média geométrica para chegar ao relativo médio de cada categoria (materiais, equipamentos, serviços e mão-de-obra). A partir dos relativos médios de cada categoria, calcula-se o índice geral de cada cidade usando uma média aritmética ponderada, baseada na estrutura de pesos de cada capital. Por fim, o INCC-M nacional é obtido pela média ponderada dos índices das sete capitais, de acordo com a importância relativa de cada uma no mercado nacional de construção habitacional (FGV, 2016).

## **2.2. Índice de Custo de Vida de Florianópolis**

O Índice de Custo de Vida (ICV) de Florianópolis foi criado em 1968 pelo Centro de Ciências da Administração e Socioeconômicas (ESAG) da Universidade Estadual de Santa Catarina. Este índice tem como objetivo medir a variação de preços incidentes sobre os orçamentos das famílias com renda entre 1 e 40 salários mínimos, especificamente na região de Florianópolis (UDESC, s.d).

Este índice foi criado a partir da necessidade de medir as variações de preços na cidade de Florianópolis, uma vez que o IPCA não abrange a capital catarinense na coleta de dados, permitindo assim uma análise mais precisa da inflação regional. Além disso, este índice acaba sendo mais preciso, pois se baseia nos padrões de consumo da população de Florianópolis. A metodologia deste índice é igual à metodologia do cálculo do IPCA e portanto os pesos relativos de cada produto utilizado no índice são baseados na POF realizada pelo IBGE (UDESC, s.d).

## **2.3. Índice de Preços Regional do Paraná - IPR**

O Índice de Preços Regional do Paraná – Alimentos e Bebidas (IPR – Alimentos e Bebidas), desenvolvido pelo Instituto Paranaense de Desenvolvimento Econômico e Social (IPARDES), tem como objetivo mensurar a variação mensal dos preços de uma cesta composta por 35 itens essenciais no orçamento familiar. Este índice abrange o estado do Paraná e municípios-pólo, como Curitiba, Londrina, Maringá, Cascavel, Ponta Grossa e Foz do Iguaçu, permitindo analisar o comportamento dos preços em diferentes regiões (IPARDES, 2022).

A metodologia do IPR envolve a seleção de localidades com relevância regional, considerando aspectos econômicos, demográficos e de influência urbana, bem como a escolha de produtos baseados na Pesquisa de Orçamentos Familiares (POF) de 2018. A amostragem dos estabelecimentos comerciais é definida por critérios de porte e localização, incluindo hipermercados e supermercados, com o objetivo de assegurar a consistência dos dados (IPARDES, 2022).

A coleta de preços ocorre por meio de notas fiscais eletrônicas (NFC-e), refletindo o valor pago pelos consumidores. O cálculo do índice utiliza o método de Divisia, baseado na média geométrica ponderada dos preços relativos entre períodos, o que garante comparabilidade ao longo do tempo e permite que o índice seja representativo da variação de preços de alimentos no estado (IPARDES, 2022).

## 2.4 O Comércio Eletrônico

O comércio eletrônico consiste na compra e venda de bens e serviços por meio de plataformas digitais, chamadas popularmente de *e-commerces*, utilizando a internet como meio de transação. Este modelo se popularizou com o advento da internet e tem evoluído desde a década de 1990, período em que essa modalidade de comércio surgiu (Laudon & Traver, 2021).

O crescimento do comércio eletrônico tem sido impulsionado por diversos fatores, como a expansão da conectividade à internet, a melhoria das tecnologias de pagamento e a mudança nos comportamentos dos consumidores. Segundo dados do IEDI (2022), o comércio global faturou em 2022 cerca de US\$31 trilhões, com participação de 18% do comércio eletrônico, ou US\$5,7 trilhões (Statista, 2023), o que demonstra a importância crescente desse setor na economia mundial. No Brasil, a expansão também é significativa, com um crescimento anual médio de 19,4% entre 2011 e 2020, atingindo o faturamento de R\$87 bilhões em 2020, indicando a adoção crescente dos consumidores brasileiros para essa modalidade de comércio (Ortega et al., 2021).

O advento da internet e consequente desenvolvimento de novas tecnologias e ferramentas relacionadas representou, principalmente nos últimos anos, uma mudança fundamental na vida das pessoas e, como efeito, da economia. Nesse aspecto, a expansão do comércio eletrônico representa um agente fundamental dessas mudanças, e, evidentemente, um tema imprescindível no desenvolvimento da ciência econômica. Tendo em vista que o comércio eletrônico gera um grande volume de informações que são digitalizadas e indexadas, esse mercado possibilita novas pesquisas baseadas em análise de dados e desenvolvimento de índices sociais, reforçando a necessidade de métodos eficazes para sua coleta e tratamento.

## 2.5 *Web Crawlers* e Ciências Sociais Aplicadas

Os *Web crawlers*, ou "rastreadores da web", são *softwares* que têm por objetivo fazer a coleta e indexação de informações disponibilizadas na internet. O funcionamento dos *web crawlers* é baseado em algoritmos que determinam quais páginas devem ser analisadas, e quais dados devem ser extraídos dessas páginas. As técnicas de coleta de dados, chamadas de *web scraping*, são amplamente utilizadas por diversas aplicações, como por exemplo, os motores de busca como o *Google*, que indexam páginas presentes na internet através de *web crawlers* (Baeza-Yates e Cambazoglu, 2011).

Essas técnicas se desenvolveram em paralelo ao avanço tecnológico ocorrido nas últimas décadas, que foi responsável por reduzir consideravelmente os custos de armazenamento e transmissão de dados. Como consequência da revolução digital, a quantidade de dados armazenados aumentou consideravelmente, passando de 0,8 Zettabyte (ZB) em 2009 para 33 ZB em 2018, com a previsão de atingir 175 ZB até 2025.

Como mostrado por Śpiewanowski e Talavera (2022), no contexto das ciências sociais, isso significa uma grande disponibilidade, e um crescimento considerável no volume de dados que podem ser utilizados para a realização de pesquisas empíricas através de técnicas de *web scraping*. A extração automatizada permite a criação de conjuntos de dados personalizados para atender às necessidades específicas de cada pesquisa, possibilitando a utilização de dados de menor granularidade em tempo real, algo antes disponível apenas de forma agregada e com atraso. Existem diversos estudos que fizeram análises utilizando dados extraídos na *web*, desde pesquisas envolvendo preços praticados online, vagas de emprego e dados do mercado imobiliário, que permitiram refinar as respostas para questões econômicas (Śpiewanowski e Talavera, 2022).

Nesta seção, são apresentados estudos e índices desenvolvidos a partir de dados obtidos por meio de *web crawlers*. A análise desses projetos é de grande relevância para compreender o impacto dessa tecnologia no desenvolvimento da ciência econômica.

### 2.5.1. The Billion Prices Project

Cavallo e Rigobon (2016) abordam o uso de dados de preços online, extraídos a partir de técnicas de *web scraping*, para o desenvolvimento de índices de preços para pesquisas econômicas. Os autores apresentam o projeto "*Billion Prices Project*" (BPP), que foi uma iniciativa *MIT Sloan* e da *Harvard Business School*, em que dados de preços online foram utilizados para calcular índices de inflação diários em vários países. Cavallo e Rigobon (2016) apontam que a utilização dos preços online se mostram vantajosas, uma vez que apresentam um custo menor de execução e maior eficiência, em relação aos índices métodos tradicionais utilizados por agências governamentais, onde os dados são coletados por um grande número de pesquisadores, em um grande número de estabelecimentos, e então processados para a construção de índices.

A motivação para a criação do primeiro índice de preços online se deu devido a manipulação de dados na Argentina durante o período de 2007 à 2015, em que os níveis oficiais de inflação não refletiam as variações reais nos preços. Enquanto os índices oficiais reportaram 8% de inflação, a coleta de preços online sugeriu que a inflação real era cerca de 20%.

A criação do primeiro índice de preços online foi motivada pela manipulação de dados inflacionários na Argentina entre 2007 e 2015. Durante esse período, os índices oficiais de inflação divulgados pelo governo indicavam uma taxa média de 8% ao ano. No entanto, dados coletados diariamente em sites de grandes varejistas sugeriam que a inflação real ultrapassava 20%, corroborando estimativas de governos provinciais, economistas locais e pesquisas de expectativa inflacionária das famílias (Cavallo e Rigobon, 2016).

A metodologia proposta pelos autores consiste em selecionar os varejistas e as categorias de produtos a serem monitorados. São selecionados principalmente grandes varejistas multicanal que vendem tanto online quanto offline, como o Walmart, em vez de varejistas exclusivamente online, como a Amazon. Isso ocorre porque os varejistas multicanal representam a maioria das transações de varejo na maioria dos países. Além disso, esses lojistas tendem a possuir preços similares tanto nas suas lojas físicas quanto nas lojas online, conforme demonstrado pela pesquisa de Cavallo (2016). Em sua pesquisa, o autor coletou preços de lojas online e físicas para mais de 24.000 produtos em 56 grandes varejistas em 10

países. Essa comparação utilizou aplicativos de smartphone e coleta de dados por meio de crowdsourcing, revelando um que cerca de 70% dos preços praticados eram idênticos nas lojas físicas e online. A maior semelhança foi encontrada em varejistas de eletrônicos e vestuário, enquanto as farmácias demonstraram mais variabilidade nos preços. O autor também constatou que embora o momento das mudanças de preços difira entre os ambientes online e offline, sua frequência e tamanho são comparáveis (Cavallo, 2016).

A escolha das categorias de produtos que serão analisados é feita utilizando como critério a sua presença nos índices de preços ao consumidor tradicionais, evitando categorias que são mais vendidas no comércio eletrônico (Cavallo e Rigobon, 2016).

Os dados são coletados diretamente dos sites dos varejistas, utilizando técnicas de *web scraping*, e então são tratados. Nesta etapa é feita a limpeza, padronização e categorização dos dados, com base nas categorias de índices de preços ao consumidor. Em seguida, são gerados indicadores simples para avaliar as características e desempenho dos preços ao longo do tempo. Cada varejista é tratado como uma unidade de amostragem independente, sendo monitorado por um ano antes de ser incluído nos índices (Cavallo e Rigobon, 2016).

### **2.5.2. Índice de Preços ao Consumidor Web (IPC-W)**

O Índice de Preços ao Consumidor Web (IPC-W) é um projeto desenvolvido por pesquisadores do FGV, baseado no BPP de Cavallo (2016). O objetivo deste projeto foi estimar a inflação para o Brasil, utilizando dados online, visando reduzir os custos associados à coleta de dados, oferecendo informações mais detalhadas e com maior frequência do que os índices tradicionais.

O ponto de partida da metodologia de cálculo do IPC-W é a coleta de dados feita diariamente em *websites* previamente selecionados utilizando técnicas de *web scraping*. Após a coleta, os produtos são classificados em subitens com base nas categorias da Pesquisa de Orçamentos Familiares (POF), permitindo que os dados online estejam em conformidade com aqueles utilizados nos cálculos de índices tradicionais. Após a coleta dos dados, é realizado o tratamento de dados ausentes, onde exclui-se da análise produtos que apresentam mais de 30% de dados faltantes, *i.e.* possuem preço nulo ou não foram encontrados, levando em



consideração o início do período base até o final do período referência. Além disso, também é feito o controle de outliers, através de um filtro baseado em distribuições normais log-transformadas, onde são excluídas leituras de produtos, cujas variações de preço excedem três desvios-padrão, evitando que alterações anômalas distorçam o índice (EMAp/IBRE, 2017).

Em seguida, o índice de cada subitem é calculado usando uma média geométrica dos preços relativos. Por fim, a inflação geral do IPC-W é calculada, agregando os subíndices gerados utilizando pesos derivados da POF nacional, similar ao método empregado no cálculo do IPC tradicional, garantindo consistência com outros indicadores de inflação (EMAp/IBRE, 2017).

### **2.5.3. Índice de Preços ao Consumidor Norueguês**

Com o objetivo de aumentar a representatividade dos preços online nos índices de preços, no ano de 2014, a Divisão de Estatísticas de Preços da Noruega iniciou um projeto para incorporar a análise de preços online no Índice de Preços ao Consumidor (CPI), e no *Harmonised Consumer Prices Index* (HICP). O projeto foi financiado pelo Gabinete de Estatísticas da União Europeia (Eurostat), que consiste na organização estatística da Comissão Europeia (Nygaard, 2015).

Esse projeto foi iniciado levando em consideração a tendência crescente do consumo de produtos e serviços online por parte da população norueguesa, onde até o ano de 2015, cerca de um terço da população fazia pelo menos uma compra online por mês. Desta forma, a incorporação de preços online nos índices oficiais de inflação, permitiram refletir a inflação com maior precisão, tendo em vista que as compras online podem ter níveis de preços e dinâmicas diferentes das lojas físicas (Nygaard, 2015).

O projeto focou em grupos de consumidores que representavam uma grande fatia das compras online na Noruega, focando em produtos de cuidados pessoais e eletrônicos. Foram coletados dados de quatro das maiores lojas online registradas na Noruega. Para reduzir a complexidade do cálculo de pesos, a coleta se concentrou nos produtos mais vendidos nas categorias selecionadas (Nygaard, 2015).

Com o objetivo de realizar testes da incorporação dos preços online à CPI, Nygaard (2015) desenvolveu três séries de índices diários. A primeira série consistiu em índices diários encadeados, que foram baseados na média geométrica dos preços relativos sem ponderação, utilizando a amostra casada de produtos mais vendidos em cada grupo do *Classification Of Individual Consumption by Purpose* (COICOP6). A segunda série consistiu em índices diários encadeados com estratificação que foram calculados a partir da razão entre a média geométrica dos preços de dois períodos consecutivos, levando em consideração o tamanho/volume dos produtos, como por exemplo, diferentes tamanhos de frascos de shampoo. A terceira série foi a de Índices diários encadeados sem estratificação, que seguiram o método anterior, mas sem levar em conta as diferenças de tamanho ou qualidade.

Ao comparar as três séries, foi observado que o índice baseado em médias simples de preços sem considerar as diferenças de qualidade, ficou abaixo das outras duas séries, mostrando uma redução de 5%, enquanto as outras séries indicam um aumento de preço entre 3% e 5% no mesmo período. Ambos os índices de preço médio são afetados por mudanças na composição e por variações de preço (Nygaard, 2015).

### 3 PROCEDIMENTOS METODOLÓGICOS ADOTADOS

Como já mencionado, este trabalho se propõe a elaborar um índice de inflação de alimentos para o estado de Santa Catarina e suas mesorregiões, através da utilização de técnicas de *web scraping*. Nesta seção, discute-se o método de cálculo utilizado, a metodologia de seleção da cesta, coleta e tratamento de dados.

#### 3.1 Cálculo do Índice de Preços

Por se tratar de um índice de inflação, é comum a utilização da técnica de Laspeyres para o cálculo do índice. Esse método se torna mais adequado, porque leva as quantidades do período-base como critério de ponderação, o que permite quantificar a variação monetária necessária para dispor das mesmas quantidades do período-base no período atual. O índice de Laspeyres é definido na Equação 2 (Pereira, 2017).

$$I = \frac{\sum P_t \times Q_0}{\sum P_0 \times Q_0} \quad (2)$$

Onde:

I é o índice de preços agregado

$P_t$  é o preço do produto no período t.

$Q_0$  é a quantidade no período base.

$P_0$  é o preço no período base.

Para este trabalho, foi considerado como período base janeiro de 2024, que ponderou o cálculo para todos os períodos subsequentes, conforme a fórmula de Laspeyres. As quantidades do período base foram definidas através da metodologia da Cesta Básica de Alimentos proposta pelo Departamento Intersindical de Estatística e Estudos Socioeconômicos (DIEESE) (2009), que estabelece uma cesta básica necessária pelo trabalhador a partir do Decreto Lei nº 399 (Brasil, 1938).

### 3.2. Cesta Utilizada

No cálculo realizado pelo DIEESE, os produtos de alimentação foram selecionados com base no Decreto Lei nº 399 (Brasil, 1938), que estabelece que o salário mínimo deve ser capaz de suprir as necessidades básicas de sobrevivência do trabalhador. Neste decreto, é estabelecida uma lista de alimentos que são considerados essenciais para que um trabalhador tenha uma dieta balanceada de proteínas, calorias e nutrientes.

Para cada produto é atribuída uma quantidade consumida, que indica a quantidade média consumida por trabalhador para cada item.

Nessa metodologia, a cesta de alimentos possui composições diferentes, dependendo da região analisada. Para esta pesquisa, utilizou-se a cesta da Região 3, formada pelos estados de Paraná, Santa Catarina, Rio Grande do Sul, Mato Grosso e Mato Grosso do Sul.

A composição dessa cesta, e as suas respectivas quantidades mensais são exibidas na Tabela 1.

Tabela 1 – Produtos que compõem a cesta básica, quantidade consumida mensal e o peso do produto na cesta.

<b>Produto</b>	<b>Quantidade Consumida Mensal</b>	<b>Peso</b>
Carne (Frango)	6,6 kg	23,28%
Leite (UHT)	7,5 Kg	26,45%
Feijão	4,5 kg	15,87%
Arroz	3,0 kg	10,58%
Farinha	1,5 kg	5,29%
Café Em Pó	600 g	2,12%
Açúcar	3,0 kg	10,58%
Óleo de Soja	900 g	3,17%
Manteiga	750 g	2,65%

Fonte: Adaptado de DIEESE (2009)

### 3.3. Coleta e tratamento de dados

Para atingir o objetivo proposto neste trabalho, foram desenvolvidos robôs para realizar a coleta de preços de produtos em 4 redes de supermercados no estado de Santa Catarina, para as regiões Norte, Vale do Itajaí, Grande Florianópolis e Sul Catarinense, totalizando 14 estabelecimentos analisados.

Foram selecionadas apenas as redes de supermercado que atendem tanto em lojas físicas quanto em lojas online, uma vez que, conforme demonstrado por Cavallo (2016), os preços em lojas exclusivamente online, como *marketplaces*, tendem a ser mais voláteis do que os preços praticados por lojas que atuam nos dois formatos, o que tende a melhorar a representatividade do índice. Na Tabela 2, é exibida a distribuição de estabelecimentos selecionados por região e a quantidade de leituras consideradas na análise, após o tratamento de dados.

Tabela 2 – Distribuição de estabelecimentos selecionados por região e leituras realizadas

<b>Rede</b>	<b>Norte</b>	<b>Vale do Itajaí</b>	<b>Sul</b>	<b>Grande Florianópolis</b>
Rede A	-	11,9 mil	-	13,9 mil
Rede B	13 mil	13,8 mil	13,7 mil	14,36 mil
Rede C	-	3,6 mil	3,8 mil	3,9 mil
Rede D	17,7 mil	17,3 mil	16,6 mil	17,1 mil

Fonte: Adaptado de DIEESE (2009)

Os robôs foram programados para pesquisar os termos específicos referente a cada categoria de produtos da cesta, e armazenar todos os resultados obtidos em um banco de dados.

Para coleta de preços de diferentes regiões, o robô fez as pesquisas de preço alterando a localização de busca para cidades de cada mesorregião analisada. Após o desenvolvimento dos robôs, foram coletados periodicamente de forma automatizada, os preços dos produtos que compõem a cesta básica definida.

Entre janeiro e abril de 2024, foram coletados 306 mil registros de preços. Após a coleta, os dados passaram por um tratamento para excluir produtos fora das categorias analisadas. Esse processo foi necessário devido à metodologia de coleta, que se baseia na pesquisa de termos-chave nos sites monitorados, resultando, eventualmente, em registros irrelevantes para as categorias em estudo. Além disso, foram removidos produtos com preços que ultrapassam um desvio padrão acima da média. Esse ajuste assegura que os itens comparados estejam em uma faixa de preços semelhante e compartilhem características compatíveis, conferindo maior representatividade ao índice.

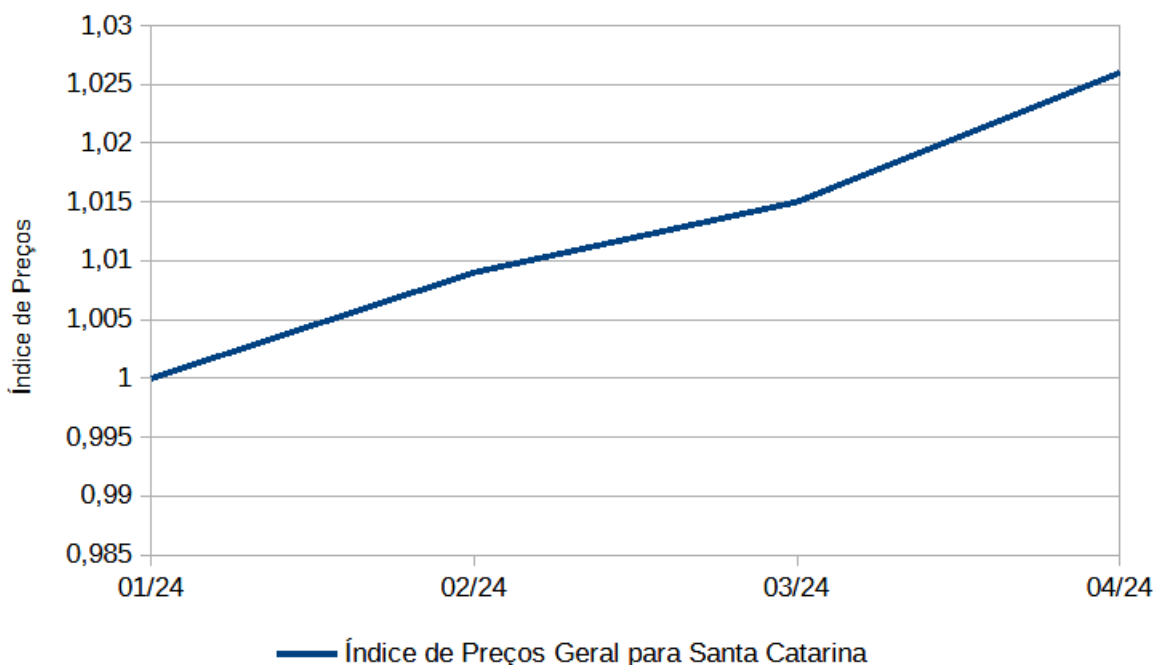
## 4 RESULTADOS E DISCUSSÃO

Neste capítulo, serão apresentados os resultados obtidos a partir da coleta de dados dos preços de alimentos básicos em Santa Catarina, realizada entre janeiro e abril de 2024. O objetivo desta seção é demonstrar a evolução dos índices de preços ao longo do período analisado. Esse capítulo será dividido em três etapas, onde inicialmente será feita uma análise do índice geral de preços, e das mesorregiões analisadas, que se seguirá de uma análise de preços por categoria de produto analisada, e por fim será discutido as limitações do método.

### 4.1. Índice Geral de Preços

Neste tópico, apresenta-se o índice geral de preços de alimentos para o estado de Santa Catarina, calculado a partir dos dados coletados via *crawlers* entre janeiro e abril de 2024. A análise considera as mesorregiões Norte, Vale do Itajaí, Sul e Grande Florianópolis. Os dados são visualizados no gráfico da Figura 1, ilustrando as variações de preços ao longo do período.

Figura 1 – Índice Geral de Preços de Alimentos (Janeiro à Abril de 2024).

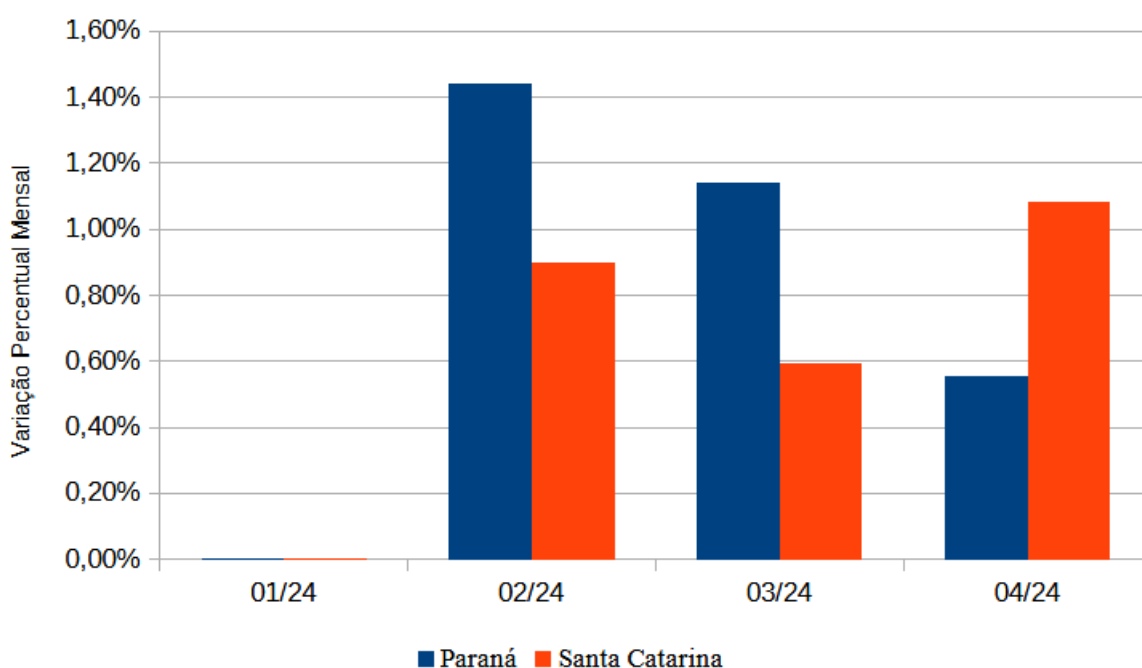


Fonte: Elaborado pelo autor

A análise dos dados revelou que o índice geral de preços apresentou crescimento contínuo ao longo do período analisado, culminando em um aumento total de 2,5%.

Na Figura 2, é exibido um gráfico comparando o índice calculado neste trabalho com o IPR - Alimentos desenvolvido pelo IPARDES para o estado do Paraná. Essa comparação tem como objetivo validar se o índice desenvolvido possui um comportamento similar ao índice paranaense.

Figura 2 – Variação Mensal do Índice Geral de Preços de Santa Catarina e Paraná.



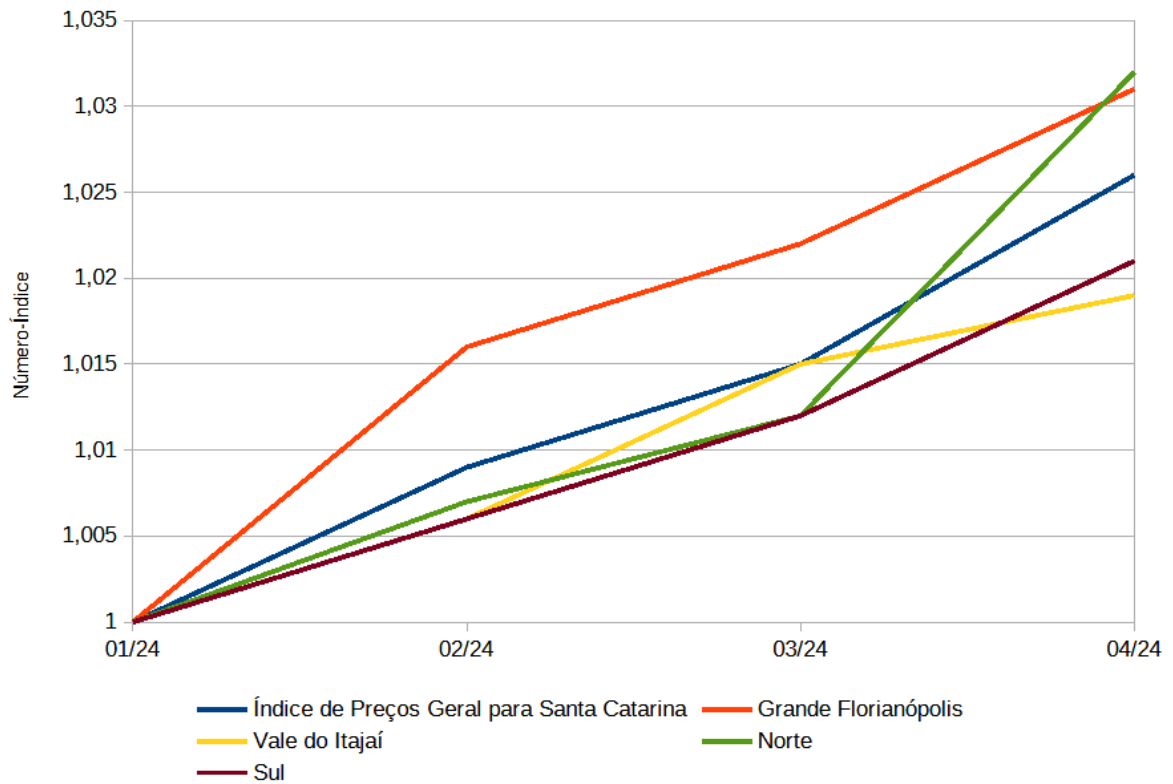
Fonte: Adaptado de IPARDES (2024)

A partir da análise, observou-se que o índice do estado do Paraná apresentou quedas contínuas ao longo do período, enquanto o índice de Santa Catarina apresentou quedas nos meses de fevereiro e março, e reverteu essa tendência no mês de abril.

Ao comparar o índice entre diferentes mesorregiões, observou-se uma tendência semelhante de aumento entre todas as regiões analisadas, com um aumento mais acentuado nas regiões Norte e Grande Florianópolis. O gráfico contendo os índices de preço por mesorregião são exibidos na Figura 3.



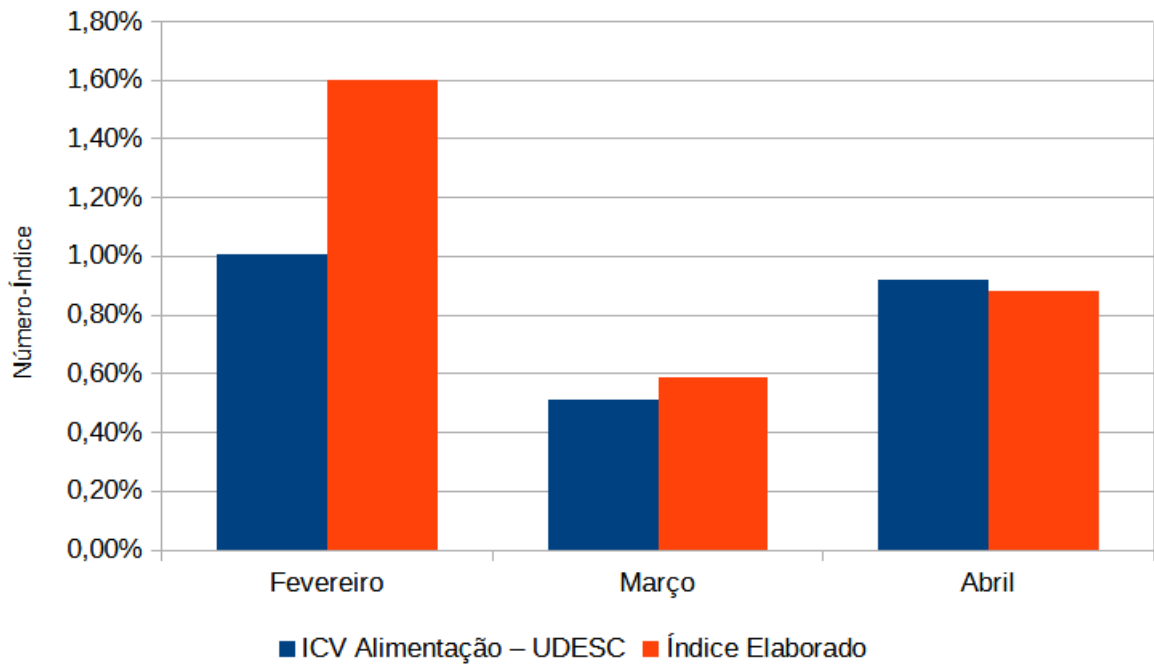
Figura 3 – Índice de Preços por Mesorregião



Fonte: Elaborado pelo autor

Para avaliar a representatividade do índice de preços com dados online, é essencial compará-lo com índices já estabelecidos. Espera-se que ambos apresentem comportamentos semelhantes. Assim, foi realizada a comparação entre o índice desenvolvido neste estudo e o subíndice de preços de alimentação e bebidas calculado pela UDESC. O gráfico comparando a variação percentual mensal dos dois índices é exibido na Figura 4.

Figura 4 – Comparação da variação percentual mensal do ICV de alimentos e o índice elaborado.



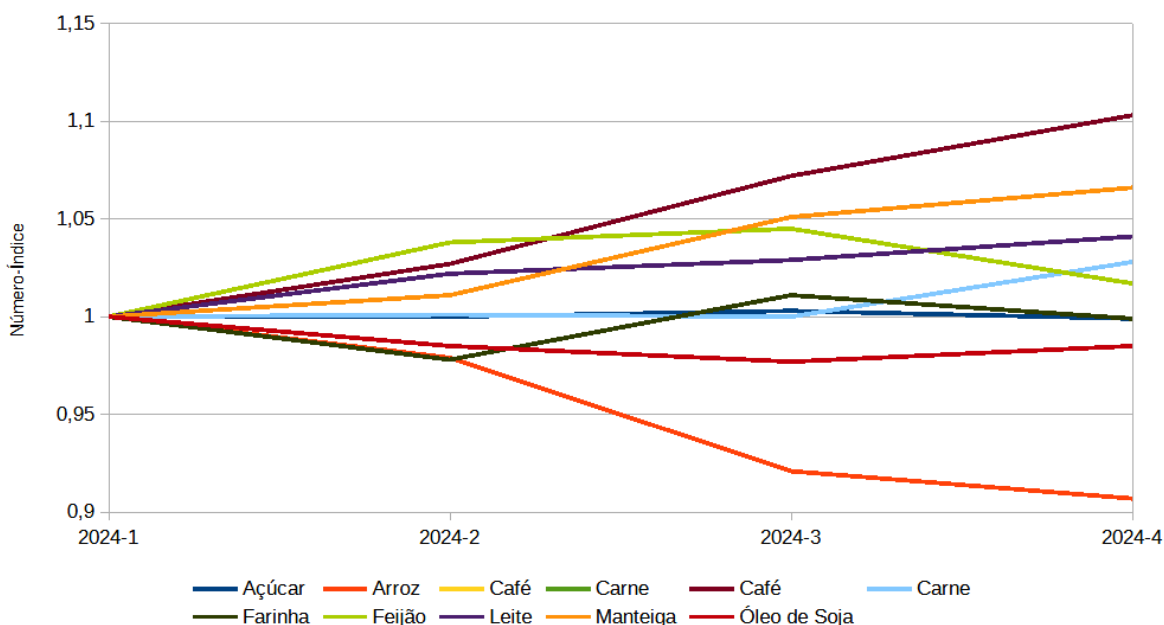
Fonte: Elaborado pelo autor.

A comparação entre o índice calculado neste estudo e o ICV da UDESC para a região de Florianópolis revela comportamentos semelhantes entre ambos. Esse resultado sugere que o índice desenvolvido neste trabalho pode ser representativo para a região.

## 4.2 Índice de Preços Por Produto

Para aprofundar a compreensão da dinâmica inflacionária no estado de Santa Catarina, analisou-se a variação nos preços para cada categoria da cesta selecionada. A partir desta análise, é possível identificar quais categorias de produtos tiveram maior impacto na variação do índice durante o período observado. O gráfico contendo o índice de preços por categoria é disponibilizado na Figura 5.

Figura 5 – Índice de Preços por Produto para o estado de Santa Catarina

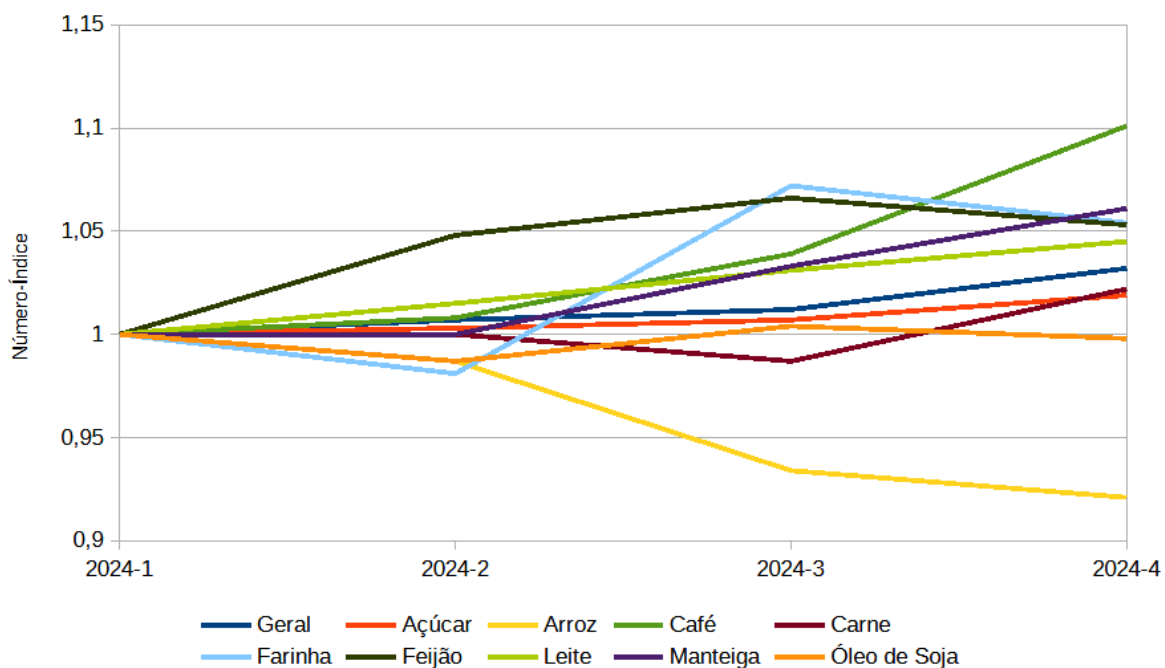


Fonte: Elaborado pelo autor

Ao examinar os produtos de forma individual, constatou-se que o índice de preços geral, bem como os índices regionais, foram influenciados por aumentos nos índices de preços de todos os produtos, com exceção do arroz e óleo de soja.

Em seguida, são disponibilizados os índices para cada categoria para a mesorregião Norte. A partir dessa análise, é possível examinar como os preços das categorias variaram ao longo do período analisado, e como influenciaram o índice na região. O gráfico contendo as variações de preços por categoria é disponibilizado na Figura 6.

Figura 6 – Índice por produto para região Norte

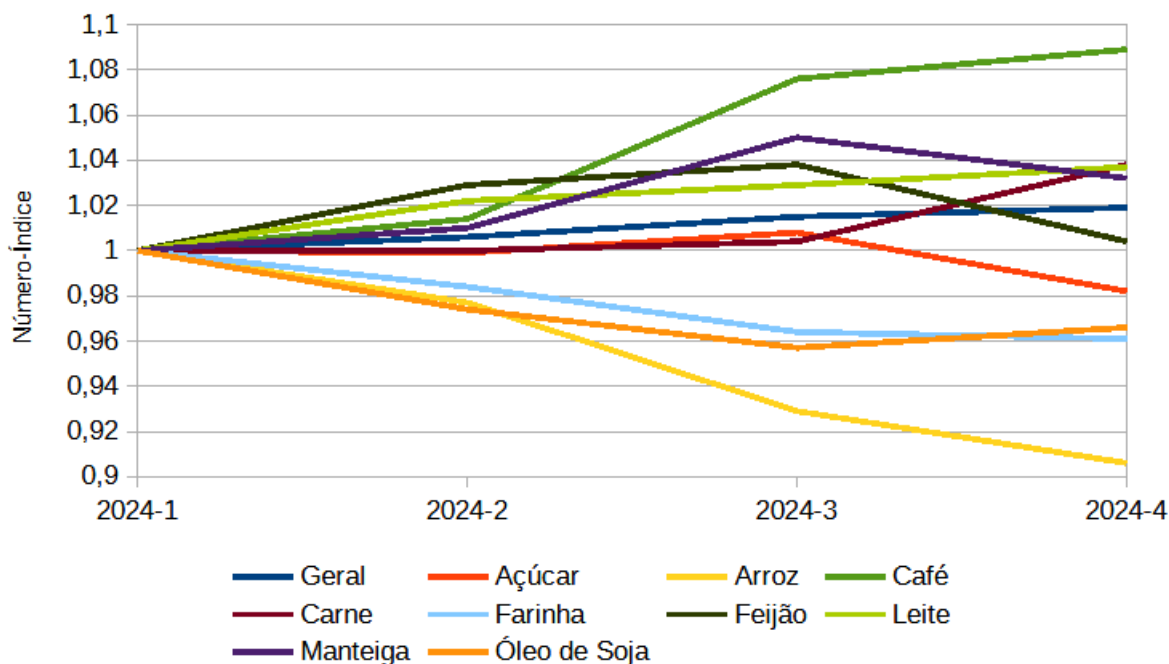


Fonte: Elaborado pelo autor

Ao analisar os dados obtidos, verificou-se que todas as categorias, com exceção do Arroz, apresentaram uma tendência de alta durante o período analisado, o que explica o aumento do índice de preços para a mesorregião.

Em seguida, são exibidos os dados obtidos de produtos por categoria para o Vale do Itajaí. O comportamento observado nesta região foi de uma queda mais acentuada do Arroz, Óleo de Soja, Farinha e Açúcar, enquanto as outras categorias apresentaram uma tendência de alta. O gráfico contendo os índices de preço por categoria para o Vale do Itajaí são exibidos na Figura 7.

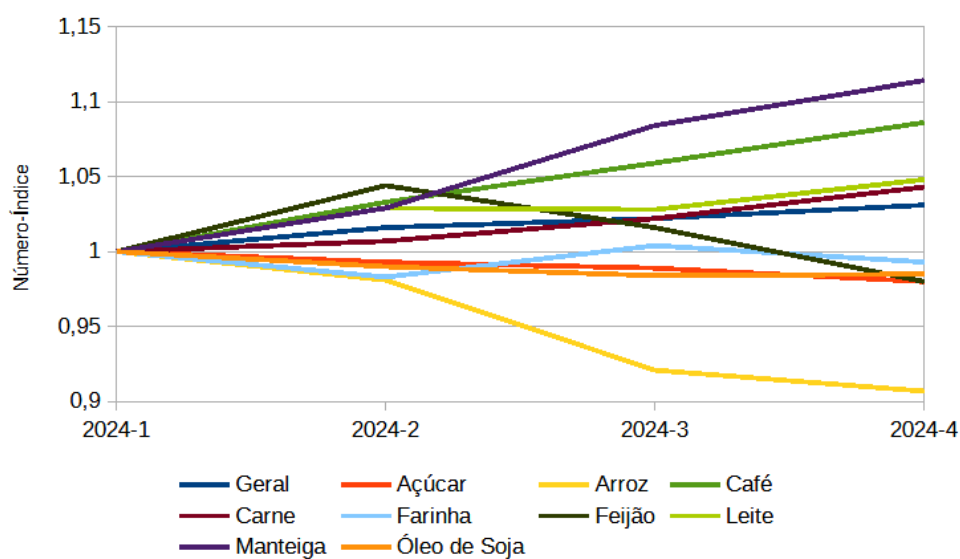
Figura 7 – Índice por Produto Para a Região do Vale do Itajaí



Fonte: Elaborado pelo autor

A região da Grande Florianópolis, por sua vez, apresentou uma tendência de alta em praticamente todas as categorias observadas, com exceção do Arroz. Isso pode indicar uma maior resistência para a redução de preços nesta região.

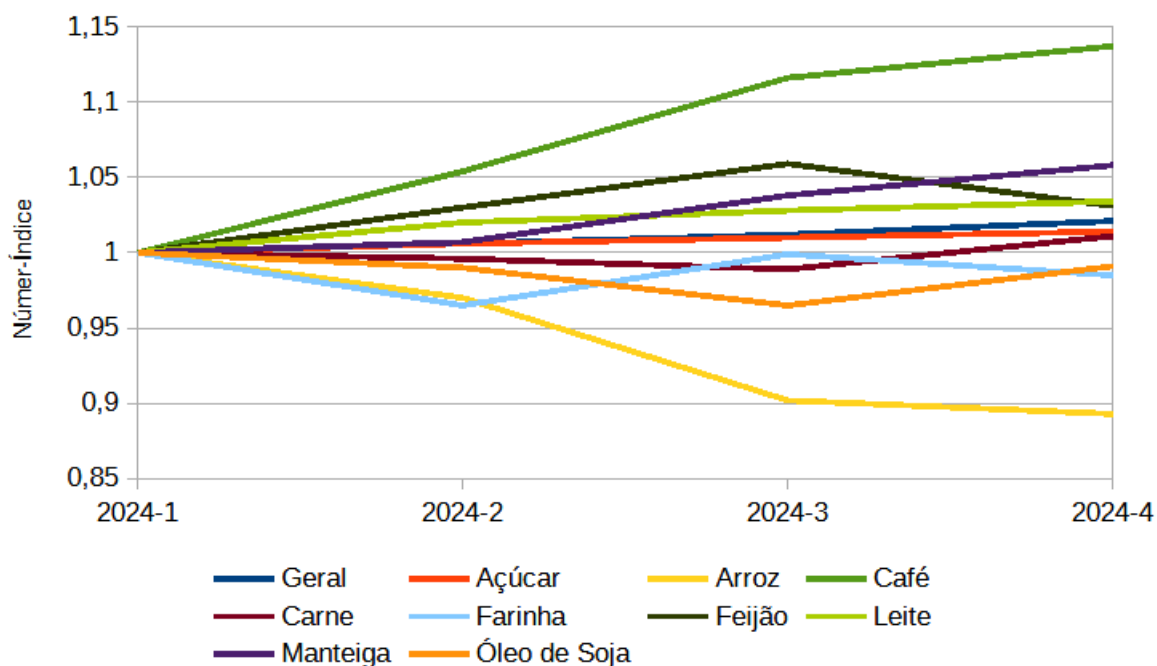
Figura 8 – Índice por produto para a região de Grande Florianópolis



Fonte: Elaborado pelo autor

Para a região Sul, o comportamento observado foi similar às outras regiões, apresentando queda significativa no preço do Arroz. O gráfico contendo o índice de preços por produto para a região Sul é exibido na Figura 8.

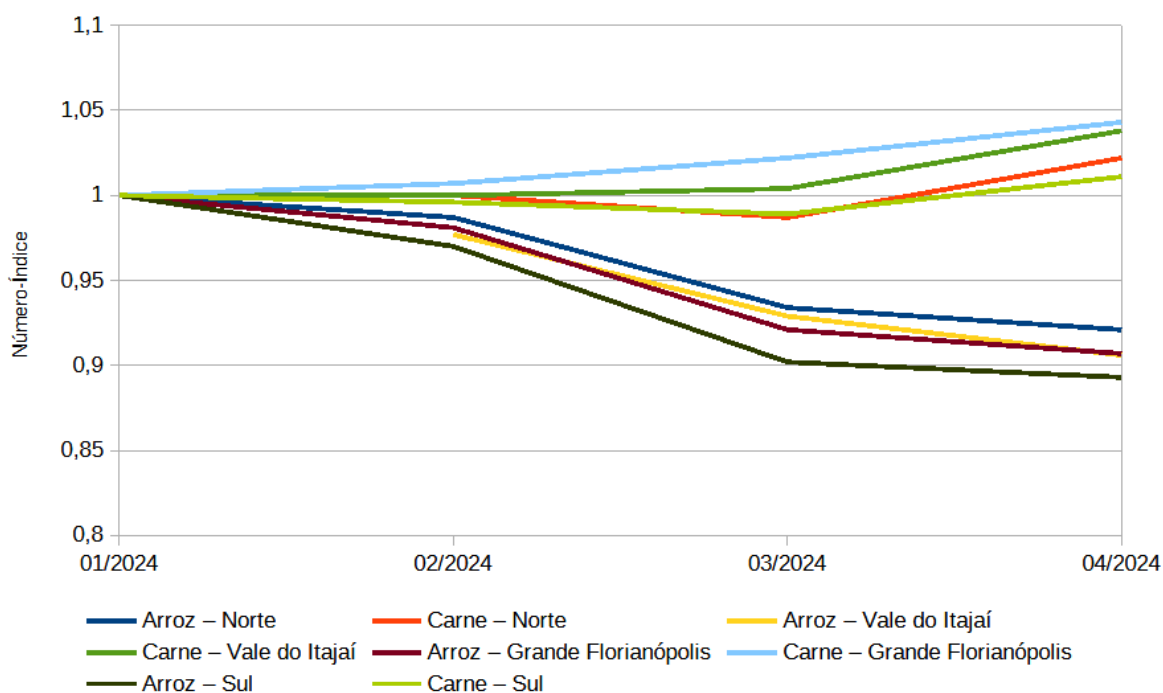
Figura 9 – Índice por produto para a região do Sul Catarinense



Fonte: Elaborado pelo autor

Para avaliar a consistência das variações no índice de preços, foi realizada uma comparação entre os itens Carne e Arroz. Observou-se um padrão similar em todas as regiões, embora com intensidades diferentes: enquanto o preço do arroz registrou queda em todas as áreas, o preço da carne apresentou aumento. Esse comportamento está ilustrado no gráfico da Figura 10.

Figura 10 – Índice por produto para a região do Sul Catarinense



Fonte: Elaborado pelo autor.

### 4.3. Limitações do Método

Embora a coleta automatizada de dados via *web scraping* ofereça vantagens consideráveis, algumas limitações dessa metodologia podem comprometer a robustez do índice de preços construído. Primeiramente, a amostra é restrita, pois a coleta abrange apenas supermercados que divulgam preços online. Essa limitação impossibilitou o cálculo da inflação para as regiões Oeste e Serrana devido à ausência de fontes de dados, o que reduz a abrangência da estimativa da inflação de alimentos para todo o estado. Conseqüentemente, a análise pode apresentar maior volatilidade, já que variações de preços em um número limitado de estabelecimentos podem não refletir com precisão o comportamento do mercado em geral. Expandir a amostra, incluindo mais supermercados e pontos de venda, contribuiria para a precisão da análise.

Além disso, a impossibilidade de coletar dados de estabelecimentos sem presença *online*, como açougues, padarias, mercearias e feiras, limita a representatividade do índice, especialmente em regiões menos urbanizadas, onde esses estabelecimentos compõem uma

parcela significativa do consumo de alimentos. A adesão crescente de pequenos negócios a plataformas digitais, como aplicativos de entrega (por exemplo, iFood), pode, no entanto, ajudar a mitigar essa limitação no futuro, ampliando o espectro de fontes de dados com a inclusão de pequenos mercados e padarias.

Outro ponto a ser observado é que o método de coleta, por ser generalista, busca produtos por palavras-chave da categoria, sem uma seleção específica. Uma das limitações desse método é que os resultados ficam sujeitos a presença de itens que não fazem parte das categorias que serão consideradas no índice, o que exige um tratamento mais rigoroso da amostra para evitar distorções. Para isso, foram aplicados filtros, como a remoção de produtos fora das categorias analisadas e a exclusão de preços fora de um desvio padrão da média. Embora essa abordagem torne o processo de coleta mais eficiente, permitindo o registro simultâneo de múltiplos produtos e agilizando o processo, ela também torna o tratamento da amostra mais complexo.



## 5. CONCLUSÃO

Tendo em vista a importância da mensuração da inflação de alimentos no contexto da segurança alimentar, principalmente para as famílias mais pobres, além do alto custo e complexidade de construção de um índice de preços seguindo a metodologia tradicional de coleta de dados, o presente trabalho teve como objetivo desenvolver um índice de preços para alimentos básicos nas mesorregiões do estado de Santa Catarina, utilizando dados coletados por meio de técnicas de *web scraping*. Este índice utilizou dados online coletados de supermercados com presença física e digital, em quatro mesorregiões do estado de Santa Catarina, permitindo uma análise de inflação segmentada por mesorregião.

Os resultados mostraram que, entre janeiro e abril de 2024, a inflação geral de alimentos em Santa Catarina apresentou uma queda nos primeiros meses, seguida de uma leve recuperação em abril. A análise regional demonstrou que a inflação de alimentos teve um comportamento similar entre as mesorregiões analisadas, com destaque para a região do Vale do Itajaí, que teve a maior redução no índice de preços.

Não foi possível calcular a inflação de alimentos para a região Oeste e Serrana, devido à falta de fontes de dados nestas regiões. Isso mostra que por mais que o método de coleta de dados automatizada através de *web crawlers* seja vantajoso em termos de custo, ele ainda depende de fontes de dados que estejam disponíveis online. Isso pode ser um desafio para se calcular um índice mais representativo, uma vez que uma parte considerável dos estabelecimentos ainda não vende online, e portanto não é possível coletar os seus dados. Para solucionar esse problema, é indicado para novos trabalhos, que sejam utilizados mais fontes de dados, como por exemplo o IFood. Neste aplicativo, é possível encontrar uma grande variedade de estabelecimentos, como mercearias, padarias, etc, o que torna a análise mais representativa.

Em conclusão, apesar das limitações, o índice desenvolvido se mostra como uma ferramenta promissora para monitorar a inflação de alimentos em Santa Catarina. Com o avanço da adesão ao comércio online, espera-se que essa metodologia possa ser expandida e aprimorada, permitindo uma análise cada vez mais precisa e abrangente da dinâmica de preços no estado. Além disso, o uso de tecnologias de coleta automatizada, como o *web scraping*, abre novas possibilidades para pesquisas econômicas no Brasil, contribuindo para a

construção de indicadores regionais mais representativos e úteis para a formulação de políticas públicas.

## REFERÊNCIAS

**BANCO CENTRAL DO BRASIL.** Relatório de Inflação – junho de 2024. Brasília, 2024. Disponível em: <https://www.bcb.gov.br>. Acesso em: 15 dez. 2024.

CAVALLO, Alberto; RIGOBON, Roberto. The Billion Prices Project: Using Online Prices for Measurement and Research. **Journal Of Economic Perspectives**. [S. L.], p. 78-151. maio 2016. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/jep.30.2.151>. Acesso em: 23 out. 2023.

COSTA, Paulo Thiago Gomes Camêllo da; ALMEIDA, Jose Frederico Fonseca Assis de; FERNANDES, Juliana Monteiro; ORTEGA, Luciane Meneguim. E-commerce no Brasil: revisão sistemática de literatura de 2011 a 2021 / e-commerce in brazil. **Brazilian Journal Of Business**, [S.L.], v. 3, n. 4, p. 2969-2982, 20 ago. 2021. South Florida Publishing LLC. <http://dx.doi.org/10.34140/bjbv3n4-014>.

CAMBAZOGLU, Berkant Barla; BAEZA-YATES, Ricardo. Scalability Challenges in Web Search Engines. **The Information Retrieval Series**, [S.L.], p. 27-50, 2011. Springer Berlin Heidelberg. [http://dx.doi.org/10.1007/978-3-642-20946-8\\_2](http://dx.doi.org/10.1007/978-3-642-20946-8_2).

**EMAp/IBRE.** Metodologia de Cálculo do IPC-W. Rio de Janeiro: EMap/IBRE, 21 nov. 2017. Disponível em: <<https://repositorio.fgv.br/items/593d6528-2328-4c2c-af9a-c084dfa9fb81>>. Acesso em: 08 out. 2024

Fundação Getulio Vargas (FGV). **Metodologia.** Índice Geral de Preços - Mercado. 2016. Disponível em: <https://portalibre.fgv.br/sites/default/files/2020-03/metodologia-igp-m-jul-2019.pdf>. Acesso em 01 de outubro de 2024.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Índice Nacional de Preços ao Consumidor Amplo - IPCA.** Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-indice-nacional-de-precos-ao-consumidor-amplo.html?=&t=o-que-e>>. Acesso em: 06 out. 2024.

INSTITUTO DE ESTUDOS PARA O DESENVOLVIMENTO INDUSTRIAL (IEDI). **Carta IEDI n.º 1279**: A evolução recente do comércio mundial e a posição do Brasil. São Paulo: IEDI, 2023. Disponível em: <[https://www.iedi.org.br/cartas/carta\\_iedi\\_n\\_1279.html](https://www.iedi.org.br/cartas/carta_iedi_n_1279.html)>. Acesso em: 06 out. 2024.

INSTITUTO PARANAENSE DE DESENVOLVIMENTO ECONÔMICO E SOCIAL – IPARDES. **Índice de Preços Regional do Paraná – Alimentos e Bebidas: Nota Metodológica**. Curitiba: IPARDES, 2022

MENDONÇA, Helder Ferreira de. Metas para inflação e taxa de juros no Brasil: uma análise do efeito dos preços livres e administrados. **Revista de Economia Política**, [S.L.], v. 27, n. 3, p. 431-451, set. 2007. FapUNIFESP (SciELO).

NYGAARD, Ragnhild. **The use of online prices in the Norwegian Consumer Price Index**. 2015. Disponível em: <[https://www.stat.go.jp/english/info/meetings/og2015/pdf/t1s2p5\\_pap.pdf](https://www.stat.go.jp/english/info/meetings/og2015/pdf/t1s2p5_pap.pdf)>. Acesso em: OLIVEIRA, Giovani; PEREIRA, André; ZILLI, Julcemar. O índice de preços de Passo Fundo (RS): análise e relevância. **Perspectiva Econômica**, [S.L.], v. 6, n. 2, p. 59-79, 1 dez. 2010. UNISINOS - Universidade do Vale do Rio Dos Sinos.

STATISTA. Retail e-commerce sales worldwide from 2014 to 2027. 2023. Disponível em: <https://www.statista.com/statistics/379046/worldwide-retail-e-commerce-sales>. Acesso em 02 de dezembro de 2024.

SOUZA, Janes Ângelo de. Para entender os índices de preços: uma visão moderna. **Revista de Economia da Mackenzie**, v. 2, n. 2. São Paulo, 2004. Disponível em: <http://editorarevistas.mackenzie.br/index.php/rem/article/view/769/460>. Acesso em: 01 nov. 2023.

SOUZA, Sabrina de Cássia Mariano de. **Insegurança alimentar, acesso e inflação de alimentos no Brasil**. 2012. 231 f. Tese (Doutorado) - Curso de Ciências Econômicas, Universidade Federal de Uberlândia, Uberlândia, 2012. Disponível em: <https://repositorio.ufu.br/handle/123456789/13448>. Acesso em: 23 out. 2023.

ŚPIEWANOWSKI, P.; TALAVERA, O.; VI, L. **Applications of Web Scraping in Economics and Finance**. *Oxford Research Encyclopedia of Economics and Finance*, 18 jul. 2022. Disponível em:  
<https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-652>. Acesso em: 14 out. 2024.

Universidade Estadual de Santa Catarina (UDESC). **Metodologia do Índice de Custo de Vida**. 2019. Disponível em: <<https://www.udesc.br/esag/custodevida/metodologia>>. Acesso em: 03 de outubro de 2024.

## ANEXO I

Tabela 3: Variação da Média de Preços por Categoria - Região Norte

	<b>Janeiro</b>	<b>Fevereiro</b>	<b>Março</b>	<b>Abril</b>
<b>Açúcar</b>		0,25%	0,46%	1,19%
<b>Arroz</b>		-1,26%	-5,40%	-1,35%
<b>Café</b>		0,77%	3,14%	5,90%
<b>Carne</b>		0,02%	-1,36%	3,57%
<b>Farinha</b>		-1,92%	9,28%	-1,71%
<b>Feijão</b>		4,75%	1,76%	-1,20%
<b>Leite</b>		1,54%	1,54%	1,33%
<b>Manteiga</b>		-0,02%	3,33%	2,74%
<b>Óleo de soja</b>		-1,34%	1,72%	-0,54%

Fonte: Elaborado pelo autor.

## ANEXO II

Tabela 4: Variação da Média de Preços por Categoria - Região Vale do Itajaí

	<b>Janeiro</b>	<b>Fevereiro</b>	<b>Março</b>	<b>Abril</b>
<b>Açúcar</b>		-0,08%	0,88%	-2,55%
<b>Arroz</b>		-2,30%	-4,95%	-2,44%
<b>Café</b>		1,44%	6,04%	1,20%
<b>Carne</b>		-0,05%	0,42%	3,42%
<b>Farinha</b>		-1,62%	-2,06%	-0,30%
<b>Feijão</b>		2,89%	0,90%	-3,30%
<b>Leite</b>		2,19%	0,65%	0,81%
<b>Manteiga</b>		1,02%	3,98%	-1,75%
<b>Óleo de soja</b>		-2,56%	-1,82%	1,02%

Fonte: Elaborado pelo autor.

## ANEXO III

Tabela 5: Variação da Média de Preços por Categoria - Grande Florianópolis

	<b>Janeiro</b>	<b>Fevereiro</b>	<b>Março</b>	<b>Abril</b>
<b>Açúcar</b>		-0,72%	-0,43%	-0,83%
<b>Arroz</b>		-1,91%	-6,10%	-1,50%
<b>Café</b>		3,34%	2,46%	2,54%
<b>Carne</b>		0,75%	1,46%	2,00%
<b>Farinha</b>		-1,67%	2,08%	-1,09%
<b>Feijão</b>		4,44%	-2,75%	-3,50%
<b>Leite</b>		2,87%	-0,08%	1,98%
<b>Manteiga</b>		2,89%	5,31%	2,82%
<b>Óleo de soja</b>		-1,00%	-0,56%	0,05%

Fonte: Elaborado pelo autor.



## ANEXO IV

Tabela 6: Variação da Média de Preços por Categoria - Sul

	<b>Janeiro</b>	<b>Fevereiro</b>	<b>Março</b>	<b>Abril</b>
<b>Açúcar</b>		0,65%	0,38%	0,40%
<b>Arroz</b>		-2,99%	-7,03%	-0,99%
<b>Café</b>		5,37%	5,93%	1,87%
<b>Carne</b>		-0,38%	-0,71%	2,26%
<b>Farinha</b>		-3,51%	3,57%	-1,49%
<b>Feijão</b>		3,00%	2,83%	-2,65%
<b>Leite</b>		2,02%	0,73%	0,58%
<b>Manteiga</b>		0,70%	3,08%	1,90%
<b>Óleo de soja</b>		-1,04%	-2,51%	2,68%

Fonte: Elaborado pelo autor.