



# Análise da Volatilidade de Preços e Inflação de Alimentos

Jhonnata Carvalho Santos ([jhonnatacarvalho00@gmail.com](mailto:jhonnatacarvalho00@gmail.com))

João Vitor Alcântara da Silva ([joao.vitoralcantara2001@gmail.com](mailto:joao.vitoralcantara2001@gmail.com))

Fonte dos Dados: SIDRA

Aplicação OLAP: <link para a aplicação OLAP>

Github: [analise-ipca-ufrpe](#)

# ETAPA 1 - PLANEJAMENTO

# 1. Contextualização

A volatilidade dos preços dos alimentos e a inflação são temas centrais na economia, afetando diretamente a vida de todos, usando dados do IPCA e ferramentas de análises de dados como o Power BI, é possível ajudar os consumidores, oferecendo indicadores para auxiliar na análise atual do mercado alimentício. Trazendo informações chave de forma prática e confiável dos preços dos últimos dois anos.

## 2. Escopo/objetivo do Data Mart

Entender o cenário econômico dos produtos nas principais regiões metropolitanas do Brasil:

- Belém (PA)
- Fortaleza (CE)
- Recife (PE)
- Salvador (BA)
- Belo Horizonte (MG)
- Grande Vitória (ES)
- Rio de Janeiro (RJ)
- São Paulo (SP)
- Curitiba (PR)
- Porto Alegre (RS)

Isso pode ajudar usuários a analisar a variação dos preços de produtos alimentícios, identificar padrões sazonais e entender o impacto da inflação.

Os dados analisados foram os valores dos últimos dois anos do Índice Nacional de Preços ao Consumidor Amplo (IPCA), um dos principais índices de inflação produzidos pelo IBGE.

O IPCA é calculado mensalmente, então o período utilizado para as análises consta de agosto de 2023 até agosto de 2025.

### 3. Arquitetura Tecnológica



Os dados são extraídos de arquivo CSV fornecidos pelo IBGE, tratados e transformados no Pentaho Data Integration(ETL), armazenados no PostgreSQL como Data Mart e, por fim, visualizados no PowerBI para a análise de indicadores de inflação.

## 4. Processo

**Escolha e montagem da base de dados:** antes de tudo, selecionamos uma fonte de dados pública e confiável. Os dados do IPCA disponibilizados pelo IBGE através da plataforma SIDRA. A base de dados foi baixada em formato CSV, contendo a série histórica de variação de preços, os produtos alimentícios e as regiões metropolitanas onde é realizada a pesquisa.

**Estudar os dados:** fizemos uma análise exploratória para entender a estrutura dos dados, como os tipos de cada coluna, a presença de valores nulos etc.

**Levantamento das entidades, atributos, da hierarquia e das métricas:** com base na análise anterior, definimos os componentes do nosso modelo dimensional, como: entidades (dimensões) e seus atributos, a tabela fato com suas métricas e as hierarquias.

**Criação do Modelo Dimensional:** consolidação final do modelo.

## 5. Abordagem

Para o desenvolvimento do Data Mart, adotamos a abordagem Top-Down, pois ela parte das necessidades do negócio e garante uma visão mais estruturada desde o início do projeto.

O modelo Star Schema foi escolhido por sua simplicidade e eficiência em consultas analíticas, facilitando o cruzamento de informações entre as dimensões de produto, tempo e local.

## 6. Usuários

Entre os possíveis usuários, temos os tomadores de decisão e analistas que precisam de insights sobre o comportamento dos preços no varejo alimentício, como:

- Diretores ou gerentes gerais, no nível estratégico.
- Gerentes de marketing, gerentes comerciais etc, no nível tático.
- Analistas de compras, no nível operacional.
- Economistas
- Pesquisadores



# ETAPA 2 - LEVANTAMENTO DAS NECESSIDADES

## 7. Consultas de Apoio à Decisão

- 1) Qual categoria de alimentos (frutas, carnes, pescados) sofreu a maior variação de preço nos últimos 3 meses?
- 2) A inflação dos produtos é maior na região metropolitana de São Paulo ou do Rio de Janeiro?
- 3) Dentre os 10 itens que mais subiram em determinado mês, quais deles têm o maior “Peso” no índice geral?
- 4) Quais subgrupos da alimentação no domicílio apresentam maior estabilidade de preços ao longo dos meses?
- 5) Produtos ligados a datas comemorativas mostram picos de preço claros nos meses anteriores ao evento? Com quanta antecedência os preços começam a subir?
- 6) Qual região metropolitana apresentou a maior deflação em hortaliças no último trimestre?

## 8. Indicadores do negócio

**Variação Mensal (%):** Inflação oficial do mês. Mede o quanto os preços de um conjunto produtos subiram ou desceram em relação ao mês anterior.

**Variação Acumulada Anual (%):** É a inflação acumulada do início do ano até o mês da análise, por exemplo, se a "Variação acumulada no ano" para o grupo "Alimentação e Bebidas" em setembro de 2025 for de +9%, significa que, desde janeiro de 2025 até setembro de 2025, o custo médio da alimentação subiu 7%.

**Variação Acumulada dos últimos 12 meses (%):** Também chamada de Inflação Interanual, ela compara os preços de hoje com os preços do mesmo mês do ano anterior.

**Peso Mensal (%):** Indica a importância ou impacto de cada item no cálculo geral do IPCA. A soma de todos os pesos é 100%. O peso é definido com base em quanto as famílias brasileiras (pesquisadas na POF - Pesquisa de Orçamentos Familiares) gastam, em média, com aquele item.

# ETAPA 3 - MODELAGEM

# 9. Modelo Relacional

## Entidade: Grupo

- **Descrição:** Armazena as categorias principais dos produtos.
- **Atributos:**
  - grupo\_codigo (Inteiro, PK): Código identificador único do grupo.
  - grupo\_nome (Texto): Nome do grupo.

[Link para dicionário de dados](#)

# 9. Modelo Relacional

## Entidade: Subgrupo

- **Descrição:** Agrupa produtos dentro de um mesmo grupo.
- **Atributos:**
  - subgrupo\_codigo (Inteiro, PK): Código identificador único do subgrupo.
  - subgrupo\_nome (Texto): Nome do subgrupo.
  - grupo\_codigo (Inteiro, FK): Referência à tabela Grupo.

# 9. Modelo Relacional

## Entidade: Produto

- **Descrição:** Armazena as informações de cadastro dos produtos.
- **Atributos:**
  - produto\_codigo (Inteiro, PK): Código identificador único do produto.
  - produto\_nome (Texto): Nome do produto.
  - subgrupo\_codigo (Inteiro, FK): Referência à tabela Subgrupo.

# 9. Modelo Relacional

## Entidade: Localidade

- **Descrição:** Contém as informações de identificação das regiões geográficas.
- **Atributos:**
  - `localidade_codigo` (Inteiro, PK): Código identificador único da região.
  - `nome_regiao` (Texto): Nome da região metropolitana.



# 9. Modelo Relacional

## Entidade: Tempo

- **Descrição:** Representa os períodos de referência das medições do IPCA.
- **Atributos:**
  - tempo\_codigo (Inteiro, PK): Código no formato AAAAMM (ex.: 202501).
  - ano (Inteiro): Ano da observação.
  - mes (Inteiro): Mês numérico.
  - nome\_mes (Texto): Nome do mês por extenso.
  - trimestre (Inteiro): Número do trimestre do ano.
  - semestre (Inteiro): Número do semestre do ano.

# 9. Modelo Relacional

## Entidade: **Metrica\_Mensal\_IPCA**

- **Descrição:** Armazena os valores das métricas calculadas mensalmente. É a entidade transacional que conecta as demais.
- **Atributos:**
  - `metrica_id` (Inteiro, PK): Identificador sequencial único da métrica.
  - `tempo_codigo` (Inteiro, FK): Referência à tabela Tempo.
  - `produto_codigo` (Inteiro, FK): Referência à tabela Produto.
  - `localidade_codigo` (Inteiro, FK): Referência à tabela Localidade.
  - `variacao_mensal_pct` (Decimal): Variação percentual mensal.
  - `variacao_acum_anual_pct` (Decimal): Variação percentual acumulada no ano.
  - `variacao_acum_12_meses_pct` (Decimal): Variação percentual acumulada nos últimos 12 meses.
  - `peso_mensal_pct` (Decimal): Peso percentual do item no cálculo do IPCA.

# 10. Modelo Dimensional

## A. Área de Negócio

Finanças / Planejamento.

## B. Processo

Estatísticas da inflação por tipo de produto alimentício.

## C. Granularidade

Dimensão Produto: produto.

Dimensão Local: nome da capital.

Dimensão Tempo: mês.

Tabela Fato: o conjunto de métricas para um produto específico, em um mês específico, em uma região metropolitana específica.

# 10. Modelo Dimensional

## D. Atributos e Hierarquia das Dimensões

Dimensão Produto: Cod\_dim\_produto, Grupo -> Subgrupo -> Produto

Dimensão Local: Cod\_dim\_local, Cod\_capital, Nome\_capital. Temos uma dimensão plana, ou seja, os atributos estão no mesmo nível de granularidade

Dimensão Tempo: Cod\_dim\_tempo, Nome\_mes, Mes\_ano, Ano -> Semestre -> Trimestre -> Mês

# 10. Modelo Dimensional

## E. Métricas da tabela Fato

Variação mensal (%): não-aditiva

Variação acumulada anual (%): não-aditiva

Variação acumulada dos últimos 12 meses (%): não-aditiva

Peso mensal (%): não-aditiva

# 10. Modelo Dimensional

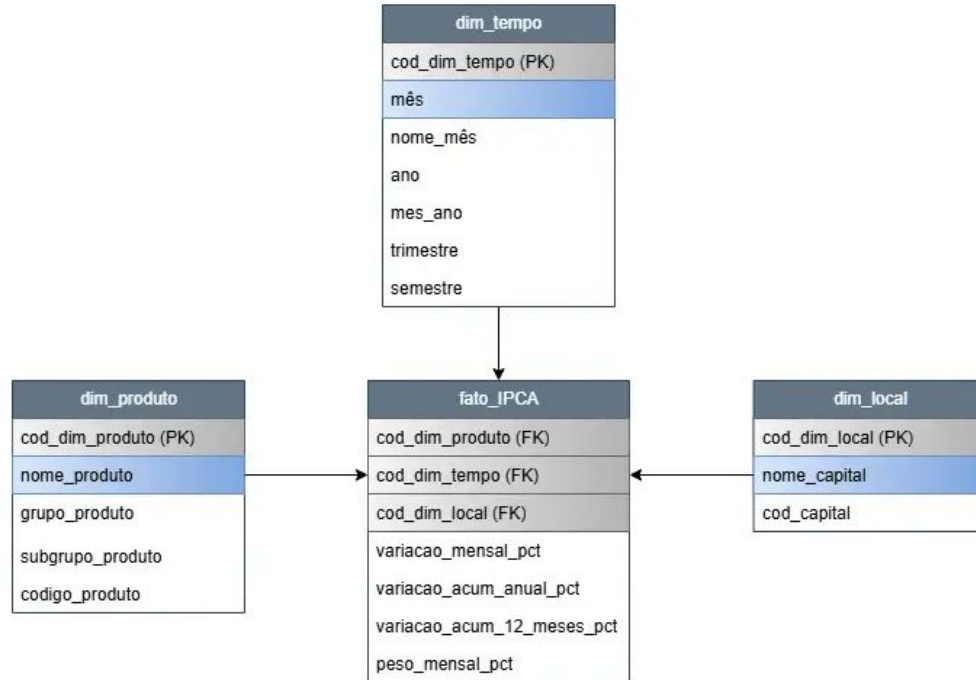
## E. Métricas da tabela Fato

Aqui, também vimos que a base de dados havia muitos valores nulos. Ex.:

Mês	Geral, grupo, subgrupo, item e subitem	Cód.	Brasil e Região Metropolitana até 2020	IPCA - Variação mensal (%)	IPCA - Variação acumulada no ano (%)	IPCA - Variação acumulada em 12 meses (%)	IPCA - Peso mensal (%)
8/1/2023	1103004.Inhame	1	Brasil	-6.72	-13.36	27.60	0.0021
8/1/2023	1103004.Inhame	1501	Belém (PA)	-	-	-	-
8/1/2023	1103004.Inhame	2301	Fortaleza (CE)	-	-	-	-
8/1/2023	1103004.Inhame	2601	Recife (PE)	-	-	-	-
8/1/2023	1103004.Inhame	2901	Salvador (BA)	-	-	-	-
8/1/2023	1103004.Inhame	3101	Belo Horizonte (MG)	-	-	-	-
8/1/2023	1103004.Inhame	3201	Grande Vitória (ES)	-1.00	-49.33	12.37	0.0268
8/1/2023	1103004.Inhame	3301	Rio de Janeiro (RJ)	-	-	-	-
8/1/2023	1103004.Inhame	3501	São Paulo (SP)	-	-	-	-
8/1/2023	1103004.Inhame	4101	Curitiba (PR)	-	-	-	-
8/1/2023	1103004.Inhame	4301	Porto Alegre (RS)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	1	Brasil	-3.28	-11.43	0.51	0.0188
8/1/2023	1103005.Mandioca (aipim)	1501	Belém (PA)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	2301	Fortaleza (CE)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	2601	Recife (PE)	-6.04	-15.72	3.35	0.1136
8/1/2023	1103005.Mandioca (aipim)	2901	Salvador (BA)	-0.04	3.77	14.09	0.0516
8/1/2023	1103005.Mandioca (aipim)	3101	Belo Horizonte (MG)	-3.29	-18.53	-11.58	0.0692
8/1/2023	1103005.Mandioca (aipim)	3201	Grande Vitória (ES)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	3301	Rio de Janeiro (RJ)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	3501	São Paulo (SP)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	4101	Curitiba (PR)	-	-	-	-
8/1/2023	1103005.Mandioca (aipim)	4301	Porto Alegre (RS)	-	-	-	-
8/1/2023	1103020.Abobrinha	1	Brasil	-0.09	9.64	-9.30	0.0077
8/1/2023	1103020.Abobrinha	1501	Belém (PA)	-	-	-	-
8/1/2023	1103020.Abobrinha	2301	Fortaleza (CE)	-	-	-	-
8/1/2023	1103020.Abobrinha	2601	Recife (PE)	-	-	-	-
8/1/2023	1103020.Abobrinha	2901	Salvador (BA)	-	-	-	-
8/1/2023	1103020.Abobrinha	3101	Belo Horizonte (MG)	-	-	-	-
8/1/2023	1103020.Abobrinha	3201	Grande Vitória (ES)	-	-	-	-
8/1/2023	1103020.Abobrinha	3301	Rio de Janeiro (RJ)	-	-	-	-
8/1/2023	1103020.Abobrinha	3501	São Paulo (SP)	-0.09	9.64	-9.30	0.0239

# 10. Modelo Dimensional

## F. Esquema Estrela



# 10. Modelo Dimensional

## G. Simulação de inserção de 10 “fatos”

dim_produto							
PK	nome_produto	grupo_produto	subgrupo_produto				
P1	Feijão	Alimentação no domicilio	Cereais, leguminosas e oleaginosas				
P2	Maçã	Alimentação no domicilio	Frutas				
P3	Picanha	Alimentação no domicilio	Carnes				
dim_tempo							
PK	ano	trimestre	mes	mes_ano	nome_mes		
T1	2024	1	2	02/2024	Fevereiro		
T2	2024	1	3	03/2024	Março		
T3	2024	1	4	04/2024	Abril		
dim_local							
PK	nome_capital	cod_capital					
L1	Rio de Janeiro	0					
L2	São Paulo	1					
L3	Recife	2					
fato_IPCA							
PK				Métricas			
FK_dim_produto	FK_dim_tempo	FK_dim_local		variacao_mensal_pct	variacao_acum_anual_pct	variacao_acum_12_meses_pct	peso_mensal_pct
P1	T1	L3		-1.7	1.45	-7.06	0.12
P1	T1	L1		-1.14	-3.09	0.39	0.29
P1	T2	L2		-8.27	-12.76	-14.98	0.14
P2	T1	L1		2.35	5.8	8.45	0.18
P2	T2	L2		1.85	4.2	6.75	0.16
P2	T3	L3		-0.95	2.1	4.3	0.17
P3	T1	L2		3.45	8.9	12.5	0.85
P3	T2	L1		2.8	7.25	10.8	0.92
P3	T3	L3		1.95	5.6	9.2	0.78
P1	T3	L2		-3.5	-5.8	-8.9	0.15



# 10. Modelo Dimensional do Data Mart (lógico)

## H. Estimativa de espaço

A estimativa foi realizada com base na tabela fato do CSV bruto, considerando o volume de dois anos de dados e o acréscimo de 20% para as dimensões, conforme a modelagem lógica do Data Mart.

Tabela fato\_IPCA

- Registros (2 anos de dados): 51.722
- Colunas: 7
  - 3 campos INTEGER  $\rightarrow 3 \times 4 = 12$  bytes
  - 4 campos NUMERIC  $\rightarrow 4 \times 6 = 24$  bytes
  - Total por linha: 36 bytes

## 10. Modelo Dimensional do Data Mart (lógico)

Cálculo:

$51.722 \times 36 = 1.862.000 \text{ bytes} \approx 1,77 \text{ MB}$

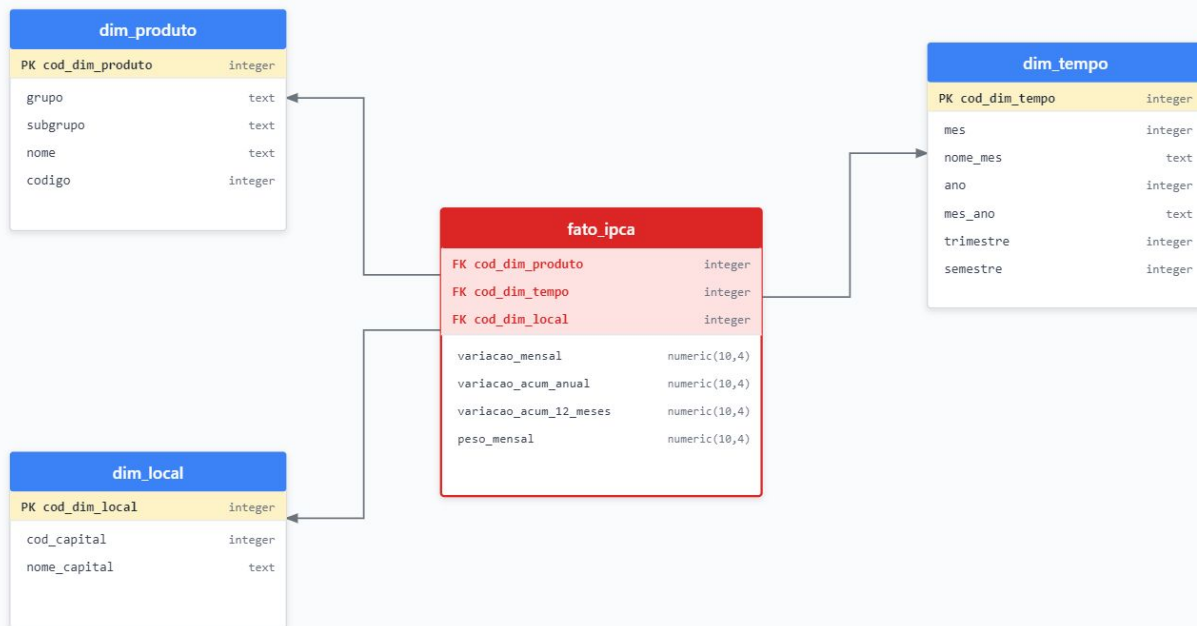
Adicionando 20% para dimensões:

$1,77 \text{ MB} \times 1,2 = \approx 2,12 \text{ MB}$

Espaço total estimado =  $\approx 2,12 \text{ MB}$

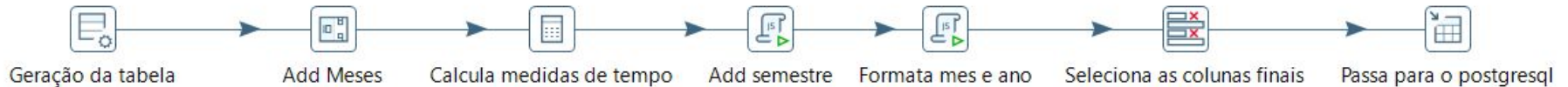
## ETAPA 4 - PROJETO FÍSICO DO BD

# 11. Modelo Relacional do Data Mart (físico)



# ETAPA 5 - EXTRAÇÃO, TRANSFORMAÇÃO E CARGA

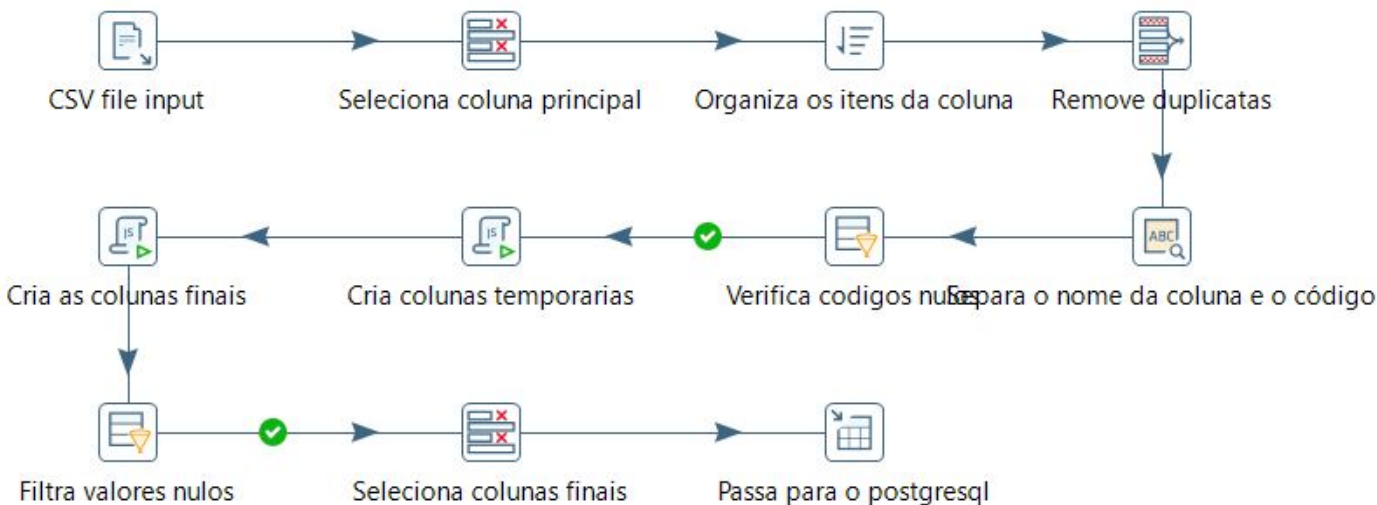
## 12. Plano de Carga da Dimensão Tempo



# 13. Plano de Carga da Dimensão Local

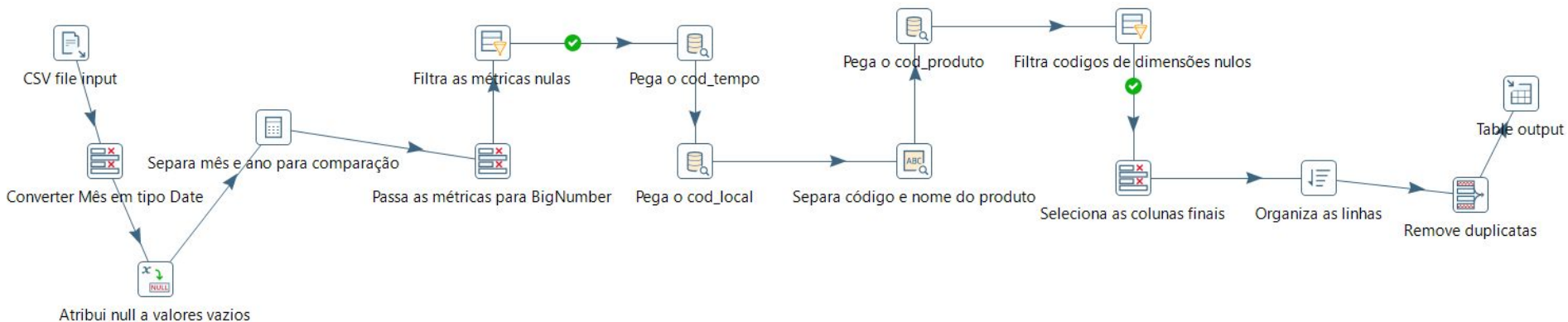


# 14. Plano de Carga da Dimensão Produto





# 15. Plano de Carga da Fato



# ETAPA 6 - APLICAÇÃO OLAP e PAINEL DE BORDO

# 16. Consulta OLAP 1

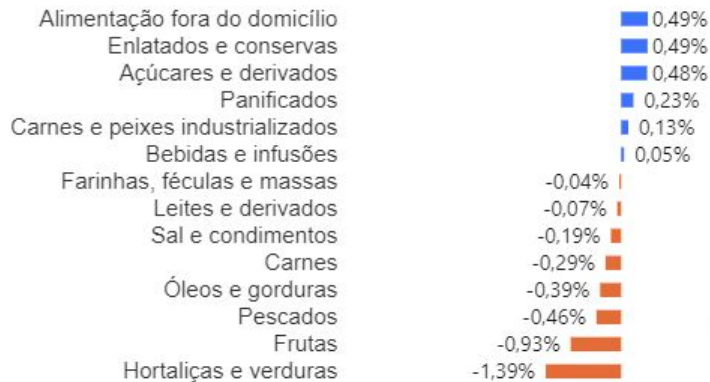
1) Qual categoria de alimentos (frutas, carnes, pescados) sofreu a maior variação de preço nos últimos 3 meses?

Ano/Mês

Seleções múltiplas

- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☒ 6
- ☒ 7
- ☒ 8

## Comparativo de Variação mensal de Preços por Subgrupo



## Maior Elevação de Preço

0,49%

Alimentação fora do domicílio

## Maior Redução de Preço

-3,76%

Tubérculos, raízes e legumes

# 17. Consulta OLAP 2

2) A inflação dos produtos é maior na região metropolitana de São Paulo ou do Rio de Janeiro?



# 18. Consulta OLAP 3

3) Dentre os 10 itens que mais subiram em determinado mês, quais deles têm o maior “Peso” no índice geral?

Ano/Mês

2025 (ano) + 1 (mes)

☒ Selecionar tudo

✓ ☐ 2023

✓ ☐ 2024

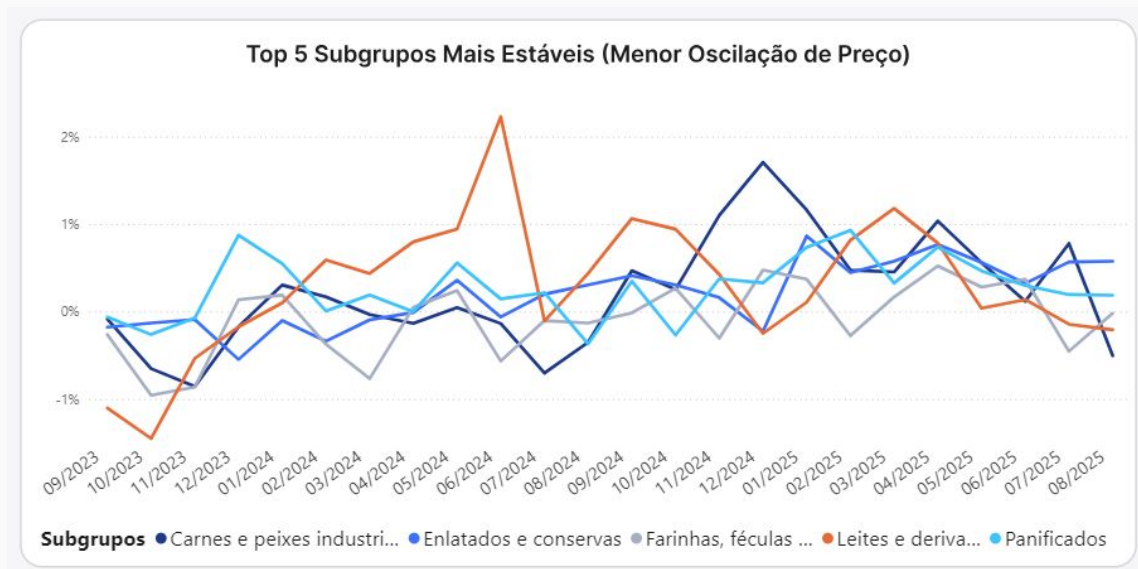
^ ☒ 2025

■ 1



# 19. Consulta OLAP 4

4) Quais subgrupos da alimentação no domicílio apresentam maior estabilidade de preços ao longo dos meses?



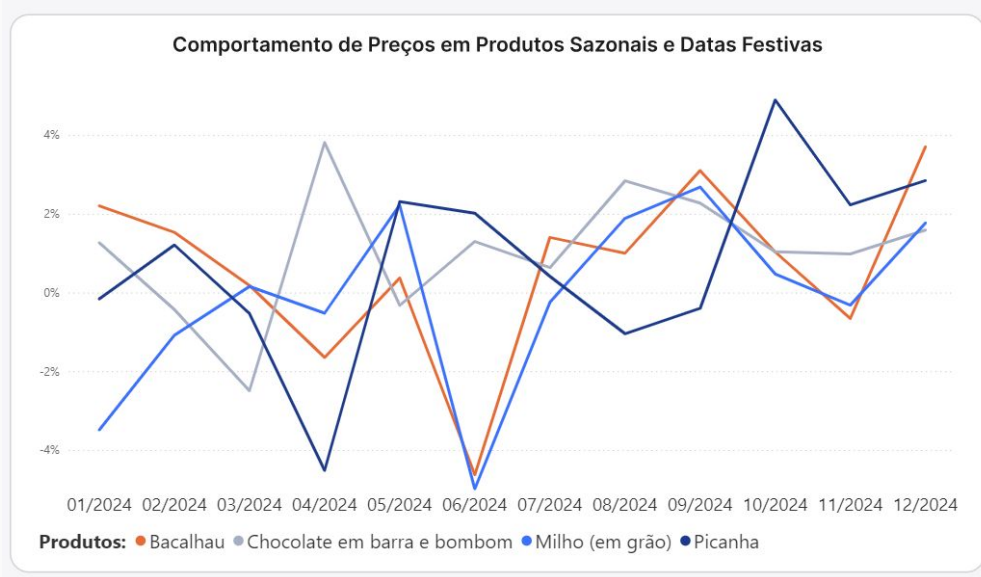
## 20. Consulta OLAP 4

4) Quais subgrupos da alimentação no domicílio apresentam maior estabilidade de preços ao longo dos meses?



# 21. Consulta OLAP 5

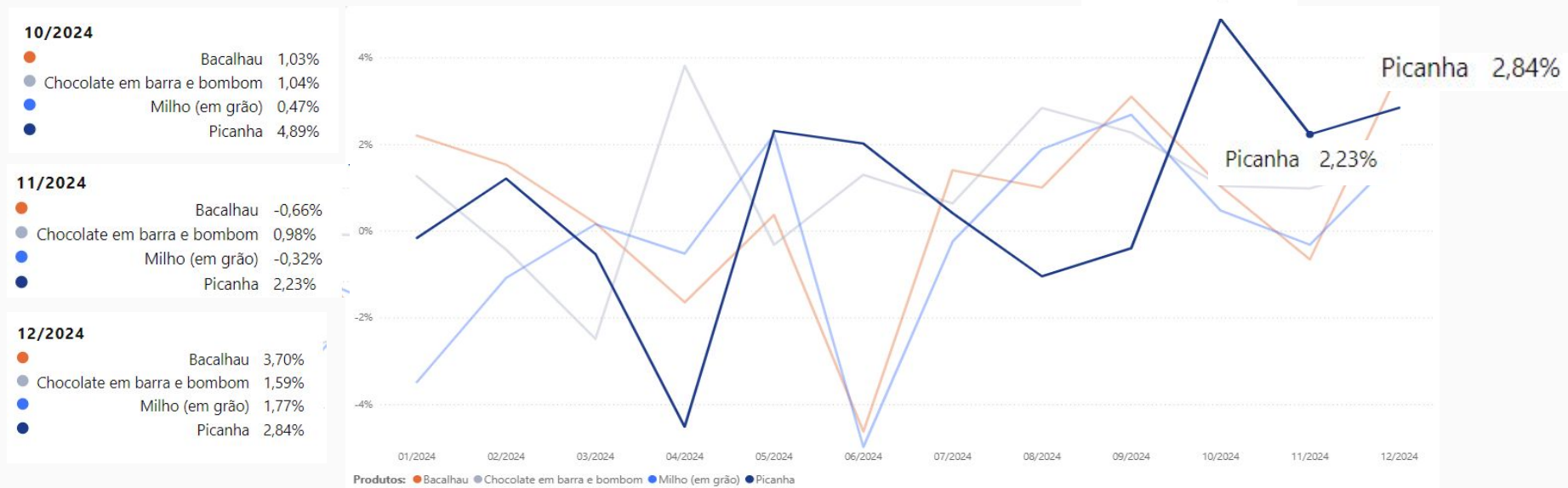
5) Produtos ligados a datas comemorativas mostram picos de preço claros nos meses anteriores ao evento? Com quanta antecedência os preços começam a subir?





## 22. Consulta OLAP 5

5) Produtos ligados a datas comemorativas mostram picos de preço claros nos meses anteriores ao evento? Com quanta antecedência os preços começam a subir?



## 23. Consulta OLAP 6

6) Qual região metropolitana apresentou a maior deflação em hortaliças no último trimestre?

Ano/Mês

2025

Trimestre

Seleções múltiplas

☒ Selecionar tudo

☐ 1

☐ 2

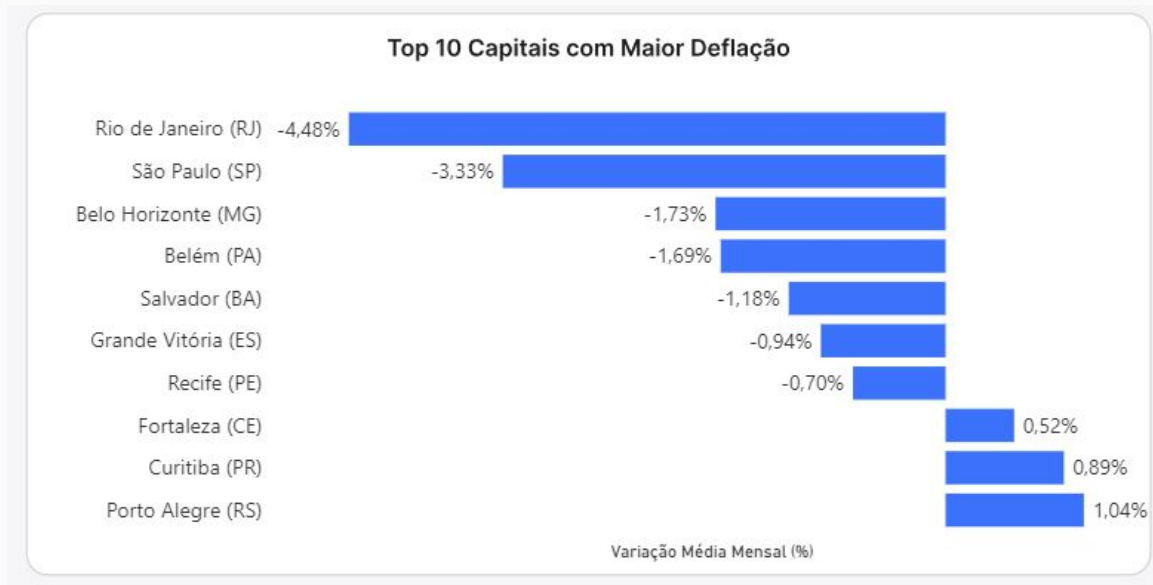
☒ 3

Grupo de Produto

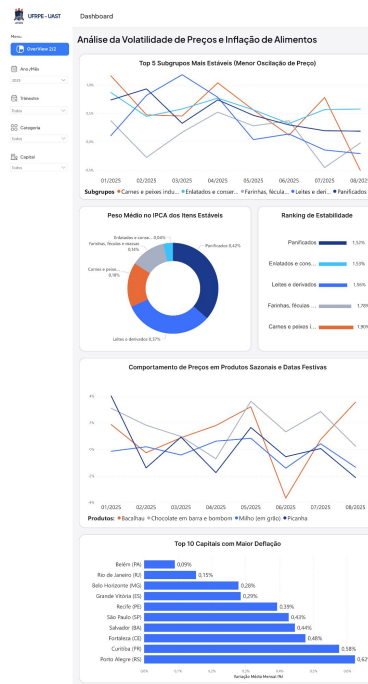
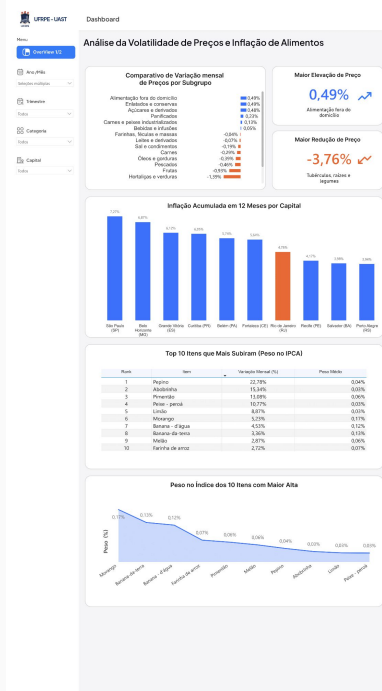
Hortaliças e verduras

Capital

Todos



## 24. Painel de Bordo



# Referências

1. Dados fontes: Os dados utilizados para a criação do Data Mart são do IBGE, disponibilizados no SIDRA.

E estão disponíveis no site: <https://sidra.ibge.gov.br/tabela/7060>