

Impacto da Reduflação nos indicadores de Shopper

Projeto da Disciplina de BI

Prof. Anderson Nascimento
prof.anderson@ica.ele.puc-rio.br

Componentes do Projeto:

Felipe Antônio Alberiz Loureiro- felipe.alberiz@gmail.com

Histórico de Versões

Data	Versão	Descrição	Autor	Aprovado por
10/02/2026	1	Versão atualizada com insights	Felipe	pendente

Sumário

1	INTRODUÇÃO	4
2	ESTUDO DE CASO	5
2.1	DESCRIÇÃO DO ESTUDO DE CASO	5
3	DESCRIÇÃO DO MODELO TRANSACIONAL.....	6
3.1	FONTE 1 - COMPOSTA POR ARQUIVOS NO FORMATO CSV QUE REPRESENTAM TABELAS ESTRUTURADAS DO NEGÓCIO.	6
4	PROPOSTA DE PROCESSO DE BI.....	8
5	MODELO MULTIDIMENSIONAL	9
6	ELABORAÇÃO DO DATA WAREHOUSE	10
6.1	DEFINIÇÃO DO DW	10
7	PROJETO DE ETL.....	12
7.1	DESCRIÇÃO DO PROJETO DE ETL	12
8	DASHBOARD	15
8.1	DESCRIÇÃO DA ELABORAÇÃO	15
8.2	TELAS DO DASHBOARD	16
9	CONCLUSÃO	17

1 Introdução

Este documento tem como objetivo estruturar, analisar e consolidar as principais necessidades do projeto “Impactos da Reduflação no Shopper”, desenvolvido a partir de um estudo de caso no mercado de suco em pó. O trabalho parte da observação de uma mudança estratégica relevante — a redução da gramatura da embalagem, mantendo o preço nominal — e busca compreender seus efeitos sobre o comportamento de compra, especialmente em indicadores como incidência e ticket médio.

A proposta do estudo é investigar de forma analítica e orientada a dados como a reduflação influencia a percepção de valor, a dinâmica competitiva e a performance da marca dentro da categoria. Para isso, o documento apresenta o contexto do mercado, os principais problemas identificados após a implementação da estratégia e o foco investigativo definido pelo cliente, com ênfase nos impactos sobre o shopper e nos possíveis desdobramentos estratégicos para a marca.

2 Estudo de Caso

2.1 Descrição do Estudo de Caso

Este estudo de caso analisa os impactos da redução no comportamento do shopper na categoria de suco em pó, a partir da redução da gramatura de uma marca líder, que passou de 25g para 18g. A mudança ocorreu em um contexto de pressão inflacionária e sensibilidade crescente ao preço, cenário em que diversas indústrias adotaram estratégias semelhantes para preservar margens. No entanto, a percepção do consumidor diante dessa alteração tornou-se um ponto central de investigação, especialmente considerando possíveis efeitos sobre incidência, ticket médio e desempenho competitivo da marca dentro da categoria.

Após a implementação da nova embalagem, observou-se variações relevantes nos principais indicadores de negócio, como queda de incidência e oscilações no ticket médio, levantando questionamentos sobre perda de valor percebido, migração para concorrentes e ajustes no comportamento de compra. O problema central do estudo reside em compreender se a redução cumpriu seu papel estratégico de proteção de rentabilidade ou se gerou impactos negativos mais amplos, como erosão de base de compradores e enfraquecimento do posicionamento da marca. O foco do negócio, portanto, está em mensurar e interpretar esses efeitos de forma estruturada, oferecendo ao cliente uma leitura clara sobre riscos, oportunidades e implicações estratégicas.

Diante desse contexto, o objetivo final do projeto é construir uma análise orientada a dados que permita identificar os impactos reais da redução no shopper, comparar o desempenho da marca com o da categoria e gerar insights acionáveis para tomada de decisão. O estudo busca transformar dados em evidências mais consistentes, apoiando recomendações estratégicas que equilibrem rentabilidade, percepção de valor e sustentabilidade de longo prazo da marca no mercado.

3 Descrição do Modelo Transacional

O modelo transacional deste projeto é composto por um conjunto de arquivos estruturados no formato CSV, representando tabelas relacionadas ao contexto de vendas e cadastro de produtos da categoria de suco em pó. Esses arquivos servem como base de dados bruta para o processo de ETL, realizado por meio da ferramenta Pentaho Data Integration.

As bases possuem estrutura relacional, com tabelas dimensionais que descrevem atributos de negócio, como canal de venda, fabricante, marca, categoria, empresa e produto. Cada tabela contém chaves identificadoras que permitem o relacionamento entre as entidades, formando um modelo lógico semelhante a um banco de dados transacional simplificado.

Os dados são extraídos dos arquivos CSV, tratados e organizados para posterior carga no ambiente analítico, onde serão utilizados para cálculo de indicadores como incidência e ticket médio, por exemplo.

3.1 Fonte 1 - Composta por arquivos no formato CSV que representam tabelas estruturadas do negócio

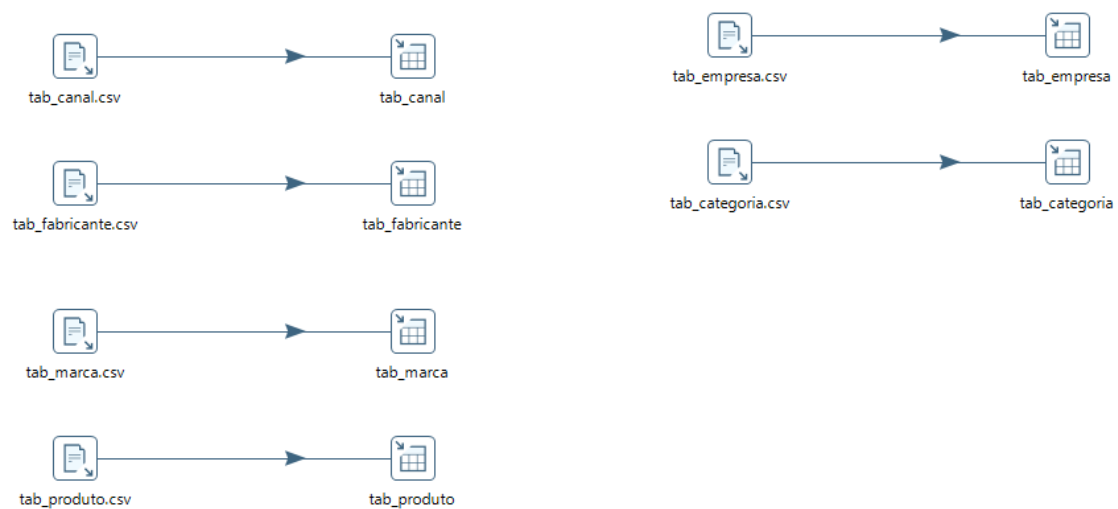
Os arquivos utilizados:

- tab_canal.csv
- tab_fabricante.csv
- tab_marca.csv
- tab_produto.csv
- tab_empresa.csv
- tab_categoria.csv

Cada arquivo contém registros organizados por chave primária, permitindo relacionamento entre entidades. A tabela de produtos atua como elemento central, conectando marca, fabricante, categoria e demais dimensões de negócio.

Esses dados são extraídos via Pentaho, conforme ilustrado no fluxo de ETL apresentado, onde cada arquivo CSV é importado e transformado em uma tabela estruturada no ambiente de destino.

Figura 1 – Fluxo de ETL no Pentaho Data Integration



O processo de extração e carga dos dados foi realizado por meio do Pentaho Data Integration, conforme ilustrado na Figura 1, onde cada arquivo CSV é importado e transformado em tabela estruturada no ambiente analítico.

4 Proposta de Processo de BI

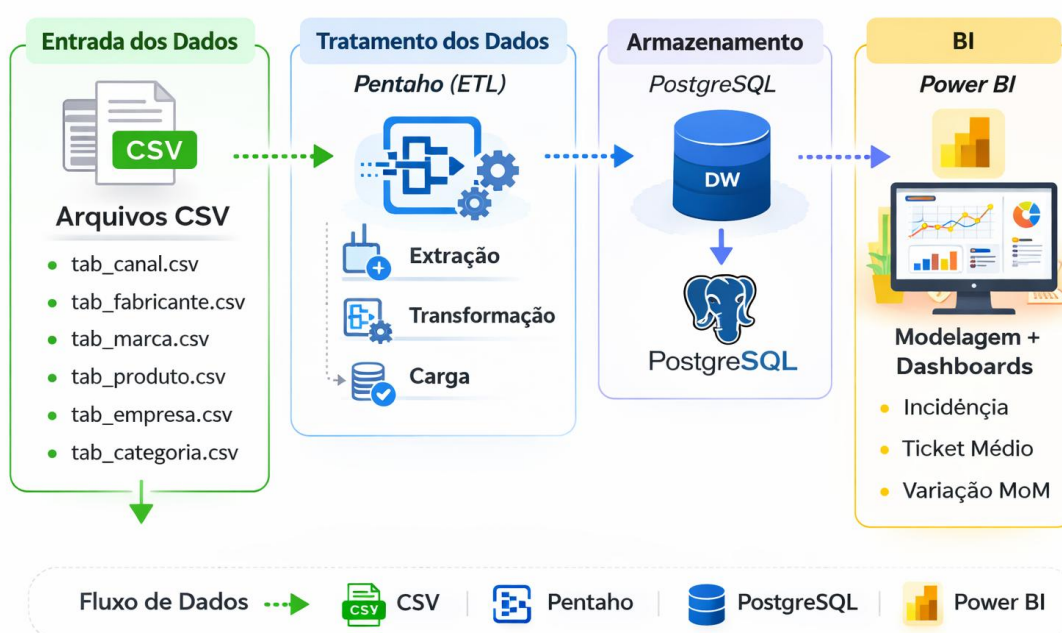
Esta seção apresenta o processo de BI proposto para o projeto, descrevendo o fluxo de dados desde a origem até a geração de análises e dashboards. A arquitetura foi estruturada com o objetivo de garantir organização, qualidade, integridade e disponibilidade das informações utilizadas no estudo sobre os impactos da reduflação no comportamento do shopper.

O processo inicia-se na camada de fontes de dados, composta por arquivos no formato CSV contendo informações estruturadas de canal, fabricante, marca, produto, empresa e categoria. Esses arquivos representam a base transacional do projeto e servem como insumo inicial para o tratamento analítico.

Na segunda etapa, os dados passam pelo processo de ETL, realizado por meio do Pentaho Data Integration. Nessa fase, são executadas as atividades de extração dos arquivos, tratamento de inconsistências, padronização de campos, validação de chaves e organização estrutural das tabelas. O ETL garante que os dados estejam limpos, organizados e prontos para armazenamento em ambiente analítico.

Após o tratamento, os dados são carregados em um banco de dados relacional PostgreSQL, que atua como repositório central (Data Warehouse simplificado). Essa camada é responsável por armazenar os dados de forma estruturada, permitindo consultas eficientes e garantindo maior governança e escalabilidade. A Figura 2 apresenta o croqui do processo de BI implementado no projeto, ilustrando o fluxo completo desde as fontes CSV até a geração de dashboards no Power BI.

Figura 2 – Croqui do Processo de BI



5 Modelo Multidimensional

Esta seção apresenta o modelo estrela (star schema) do estudo de caso Impactos da Redução no Shopper.

Figura 3 – Modelo Multidimensional



O modelo multidimensional desenvolvido para o estudo de caso foi estruturado no formato Star Schema (modelo estrela), amplamente utilizado em projetos de BI por sua simplicidade, desempenho e facilidade de análise.

No centro do modelo encontra-se a tabela fato fshopper, responsável por armazenar as métricas quantitativas do negócio. Essa tabela contém os registros transacionais de vendas e chaves estrangeiras para as dimensões. A tabela fato se relaciona com três dimensões principais: Dimensão Produto (dim_produto), responsável por armazenar os atributos descritivos dos produtos, permitindo análises por categoria, marca e fabricante; Dimensão Empresa (dim_empresa), com informações relacionadas à empresa e canal de venda e a Dimensão Calendário (dim_calendario), estruturada para permitir análises temporais.

6 Elaboração do Data Warehouse

O Data Warehouse desenvolvido tem como objetivo centralizar, organizar e integrar os dados provenientes de arquivos CSV tratados no processo de ETL realizado no Pentaho, armazenando-os em uma base estruturada no PostgreSQL. Essa base serve como fonte única e confiável para a camada analítica no Power BI, apoiando a tomada de decisão.

A estrutura do Data Warehouse foi organizada em três camadas principais:

- Camada de Origem de Dados – composta por arquivos CSV contendo os dados brutos do estudo de caso.
- Camada de Integração (ETL) – responsável pelo processo de extração, transformação e carga dos dados, realizado por meio do Pentaho. Nesta etapa foram executadas atividades como tratamento de inconsistências, padronização de formatos e criação de chaves substitutas.
- Camada de Armazenamento (Data Warehouse) – implementada em PostgreSQL, estruturada segundo modelagem dimensional no formato estrela, composta por uma tabela fato (fshopper) e tabelas dimensão (dim_produto, dim_empresa e dim_calendario), contendo os atributos descritivos utilizados nas análises.

A organização em modelo estrela garante melhor desempenho em consultas analíticas, simplicidade estrutural e maior facilidade na construção de relatórios no Power BI.

6.1 Definição do DW

6.1.1 Arquitetura: **Independente**

A arquitetura adotada para o Data Warehouse foi do tipo Independente, pois o ambiente foi construído especificamente para fins analíticos, sem integração direta com múltiplos sistemas transacionais. Os dados utilizados são provenientes de arquivos CSV externos, tratados via processo de ETL e carregados diretamente no DW, reduzindo a complexidade de integração e atendendo adequadamente ao escopo do projeto.

6.1.2 Abordagem de Construção: **Bottom-Up**

A abordagem de construção adotada foi Bottom-Up, onde inicialmente foi desenvolvido um modelo dimensional focado na área de análise de vendas, estruturado no formato estrela, sendo adequada para projetos acadêmicos pois permite entregar valor rapidamente.

6.1.3 Arquitetura Física: **On-Premises**

A arquitetura física escolhida foi a On-Premises, pois o Data Warehouse foi implementado em ambiente local utilizando PostgreSQL. Essa escolha foi adequada ao escopo do projeto, garantindo controle sobre o ambiente e simplicidade de implementação. Em um cenário corporativo de maior escala, poderia ser considerada a adoção de soluções em nuvem, visando escalabilidade e alta disponibilidade.

7 Projeto de ETL

7.1 Descrição do Projeto de ETL

O processo de Extração, Transformação e Carga (ETL) foi desenvolvido utilizando o Pentaho Data Integration (PDI), com o objetivo de preparar os dados provenientes de múltiplos arquivos CSV para armazenamento estruturado no Data Warehouse em PostgreSQL. O projeto de ETL foi dividido em fluxos específicos para carga das dimensões e da tabela fato, garantindo organização, padronização e integridade dos dados.

Extração

A etapa de extração consistiu na leitura de múltiplos arquivos no formato CSV, contendo dados brutos relacionados a:

- Produtos
- Fabricantes
- Marcas
- Categorias
- Empresas
- Canais
- Dados transacionais de vendas

Cada arquivo foi importado no Pentaho por meio do componente CSV File Input, permitindo a leitura estruturada dos campos e seus respectivos tipos de dados.

Transformação

A etapa de transformação foi responsável por aplicar regras de negócio, padronizações e integrações necessárias à modelagem dimensional.

Transformações das Dimensões

Para a construção da *dim_produto*, foram realizados:

- Ordenação dos dados (Sort Rows) para preparação dos joins
- Junções internas (Inner Join) entre as tabelas de produto, fabricante, marca e categoria

- Seleção e organização dos atributos finais (Select Values)
- Estruturação da dimensão com seus atributos descritivos

Para a construção da *dim_empresa*, foi realizada:

- Integração entre as tabelas de empresa e canal
- Junção interna (Inner Join)
- Seleção dos campos relevantes
- Consolidação do atributo canal diretamente na dimensão empresa

Essa decisão eliminou a necessidade de uma dimensão separada para canal, simplificando o modelo estrela.

Transformação da Tabela Fato

A tabela fato fshopper foi construída a partir da consolidação dos dados transacionais provenientes dos arquivos CSV. Diferentemente das dimensões, essa etapa não envolveu múltiplas junções complexas, uma vez que os dados já continham os identificadores necessários. Após a validação e organização dos campos, os dados foram carregados diretamente no Data Warehouse, mantendo a integridade referencial com as dimensões previamente carregadas

Carga

Após as transformações, os dados foram carregados no PostgreSQL utilizando os componentes de saída do Pentaho.

Foram criadas as seguintes estruturas no Data Warehouse:

- dim_produto
- dim_empresa
- dim_calendario
- fshopper

Figura 4 – ETL das Tabelas Dimensão

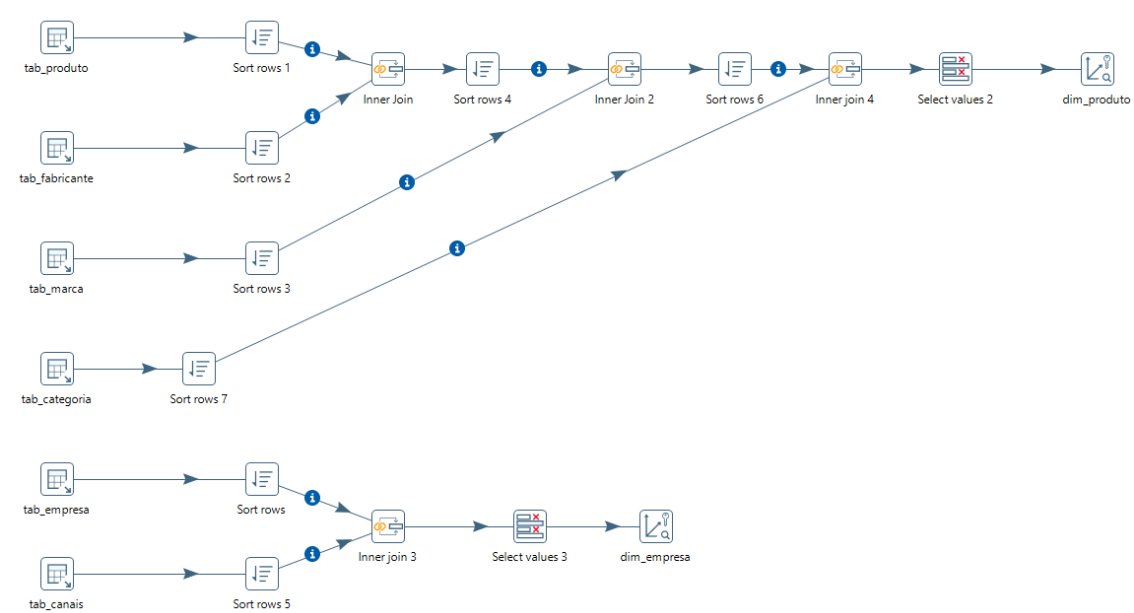
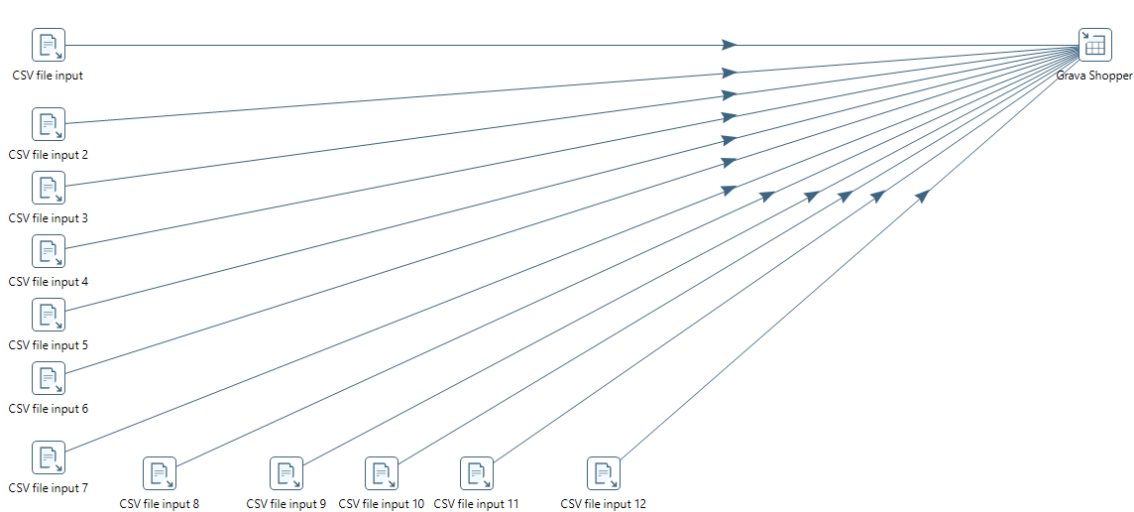


Figura 5 – ETL da Tabela Fato



8 Dashboard

8.1 Descrição da Elaboração

O dashboard foi desenvolvido utilizando a ferramenta Microsoft Power BI, escolhida por sua capacidade de fácil conexão com múltiplas fontes de dados e recursos avançados de visualização interativa.

A conexão foi realizada diretamente com o Data Warehouse implementado em PostgreSQL, onde os dados já se encontravam estruturados segundo o modelo dimensional no formato estrela. Após a importação dos dados, foram criadas medidas utilizando linguagem DAX para cálculo dos principais indicadores, como:

- Incidência de Notas
- Ticket Médio
- Análises por período (mês/ano)
- Análises por categoria e marca
- Análises por empresa e canal

A estrutura visual do dashboard foi organizada de forma a permitir análise interativa, com utilização de:

- Cartões de indicadores (KPIs)
- Gráficos de colunas e barras
- Segmentações de dados (filtros) por período, produto e empresa
- Tabelas detalhadas para análise granular

O objetivo principal do dashboard foi fornecer uma visão consolidada do comportamento da Marca, permitindo identificar padrões, variações ao longo do tempo e possíveis impactos relacionados ao comportamento de consumo.

A organização das informações buscou priorizar clareza visual, hierarquia de indicadores e facilidade de navegação, permitindo que o usuário final possa explorar os dados de forma intuitiva e eficiente para apoio à tomada de decisão.

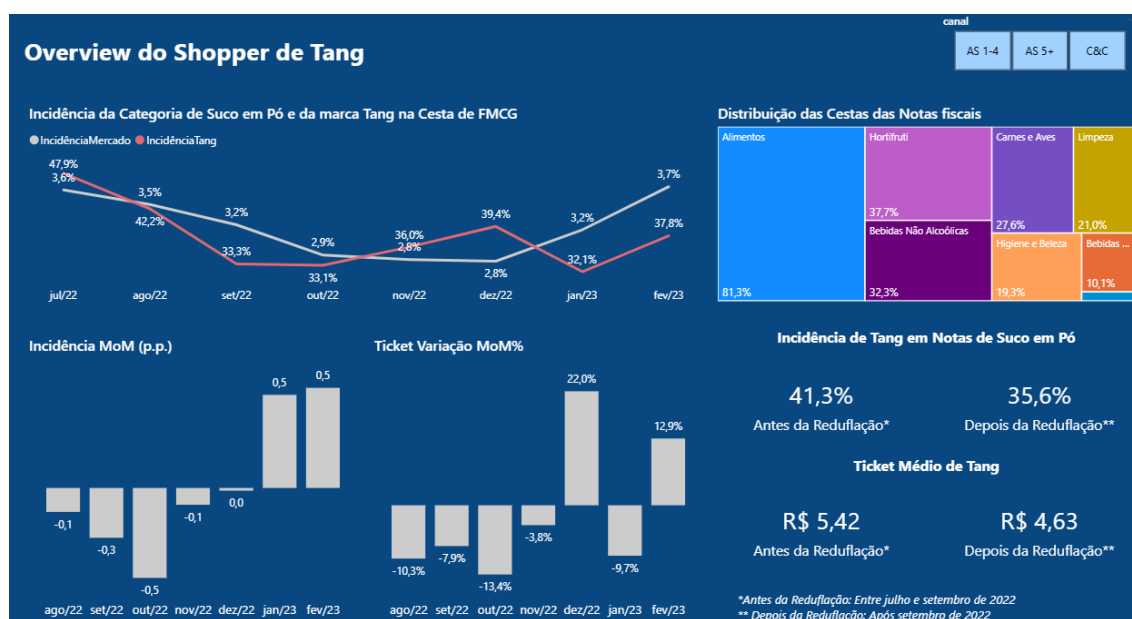
8.2 Telas do Dashboard

A estratégia de reduzir a quantidade sem alterar valores dos produtos teve mais "uma vítima". O suco em pó da Tang passou de 25g para 18g. A tentativa de "drible" na inflação não é ilegal, desde que o consumidor seja informado claramente sobre a redução. O objetivo aqui é analisar os impactos dessa tática na marca em específico.

A construção da tela teve como objetivo central consolidar, em um único painel visual, os principais indicadores necessários para avaliar o impacto estratégico da reduflação, permitindo:

- Comparação direta entre canais.
- Análise pré e pós-implementação.
- Avaliação de presença e valor por compra.
- Identificação de comportamento competitivo frente ao mercado.

Figura 6 – Dashboard



A estrutura do dashboard permite concluir que a reduflação impactou negativamente tanto a incidência quanto o ticket médio da marca nos dois canais analisados. O C&C apresentou maior sensibilidade em termos de incidência, enquanto o AS registrou queda proporcional mais relevante no ticket médio. A análise evolutiva reforça indícios de perda de competitividade relativa frente à categoria.

9 Conclusão

A elaboração deste projeto foi uma experiência bastante enriquecedora, pois permitiu aplicar na prática conceitos de Data Warehouse e Business Intelligence que muitas vezes ficam apenas no campo teórico. A construção do Data Warehouse foi uma etapa essencial para garantir consistência e confiabilidade das informações. Estruturar a tabela fato, definir corretamente as dimensões e organizar os relacionamentos permitiu criar uma base sólida para as análises. Esse processo reforçou a importância de um modelo bem estruturado para que os indicadores realmente representem a realidade do negócio.

Na etapa de visualização, utilizando o Power BI, foi possível transformar os dados em informações claras e objetivas. A consolidação de indicadores como incidência e ticket médio, antes e depois da redução, além da análise evolutiva da marca em comparação ao mercado, permitiu identificar impactos relevantes da estratégia adotada. Mais do que os números em si, o projeto reforçou a importância de analisar decisões estratégicas sob múltiplas perspectivas. Pequenas mudanças, como a redução do tamanho de embalagem, podem gerar impactos relevantes no comportamento do consumidor e no desempenho da marca.

Concluo este trabalho com a percepção de que a combinação entre modelagem de dados bem estruturada e ferramentas de visualização adequadas é fundamental para apoiar decisões mais assertivas nas organizações. A experiência foi desafiadora, mas extremamente valiosa para consolidar conhecimento técnico e desenvolver uma visão mais estratégica orientada por dados.