



PUC Minas

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS

NÚCLEO DE EDUCAÇÃO A DISTÂNCIA

Pós-graduação *Lato Sensu* em Analytics e Business Intelligence

RELATÓRIO TÉCNICO

ANÁLISE DE RECLAMAÇÕES NO PORTAL
CONSUMIDOR.GOV

Reginaldo Siqueira Carvalho

Belo Horizonte

2023

SUMÁRIO

1. Introdução	3
1.1. Contexto	3
1.2. Objetivos	3
1.3. Público alvo	4
2. Modelo de Dados	4
2.1. Modelo Dimensional	4
2.2. Fatos e Dimensões	5
3. Integração, Tratamento e Carga de Dados	9
3.1. Fontes de Dados.....	9
3.2. Processos de Integração e Carga (ETL).....	9
4. Camada de Apresentação	13
4.1 Dashboard	13
5. Registros de Homologação	19
5. Conclusões	22
6. Links	38
REFERÊNCIAS	38

1. Introdução

1.1. Contexto

A análise das reclamações dos consumidores é uma atividade de extrema importância para as empresas em geral, pois, ao utilizar técnicas avançadas de análise de dados e inteligência de negócios, essas empresas têm a capacidade de extrair insights valiosos a partir das reclamações dos consumidores.

Essa análise aprofundada permite identificar padrões, tendências e causas dos problemas enfrentados pelos clientes, fornecendo uma visão estratégica para aprimorar processos, produtos e serviços.

Além disso, a análise das reclamações dos consumidores permite às empresas antecipar tendências de mercado, compreender as preferências dos clientes e tomar decisões embasadas em dados concretos.

Dessa forma, a atuação das equipes de Business Intelligence na análise das reclamações contribui para que as empresas adotem melhores práticas, otimizem a experiência do cliente e fortaleçam sua posição competitiva no mercado.

1.2. Objetivos

O presente trabalho tem como objetivo realizar uma análise das reclamações direcionadas para a empresa Uber na plataforma Consumidor.gov.br (BRASIL, 2023) entre junho de 2021 e junho de 2023. Esta plataforma atua como um serviço público e gratuito, que possibilita a interação direta entre consumidores e empresas para a solução alternativa de conflitos de consumo pela internet.

Com foco na transparência e na busca por uma resolução ágil e desburocratizada, o serviço permite aos consumidores registrar suas reclamações e acompanhar as respostas das empresas, bem como avaliar o atendimento recebido.

Por meio dos dados disponibilizados na plataforma, será possível realizar uma análise do desempenho da empresa Uber, identificar as principais demandas dos consumidores e extrair insights acerca da satisfação do público em relação aos serviços prestados.

Essa análise contribuirá para que a empresa identifique áreas de melhoria, adote melhores práticas, aprimore a qualidade de seus produtos e serviços, fortaleça o relacionamento com os consumidores e, conseqüentemente, aumente sua competitividade no mercado.

1.3. Público alvo

O público-alvo que utilizará a solução a ser desenvolvida dentro da empresa será composto por gestores, analistas e tomadores de decisão. Esses profissionais têm a responsabilidade de conduzir e orientar as estratégias da organização, e a utilização da solução permitirá que eles tenham acesso rápido e fácil a informações relevantes e atualizadas sobre o desempenho do negócio.

A solução desenvolvida fornecerá análises consolidadas, gráficos intuitivos e métricas-chave, possibilitando uma visão holística das reclamações dos clientes, notas de avaliação, prazo de resposta dentre outros aspectos relevantes.

Essa ferramenta será essencial para a tomada de decisões embasadas em dados, facilitando a identificação de oportunidades, tendências emergentes e áreas que necessitam de melhorias.

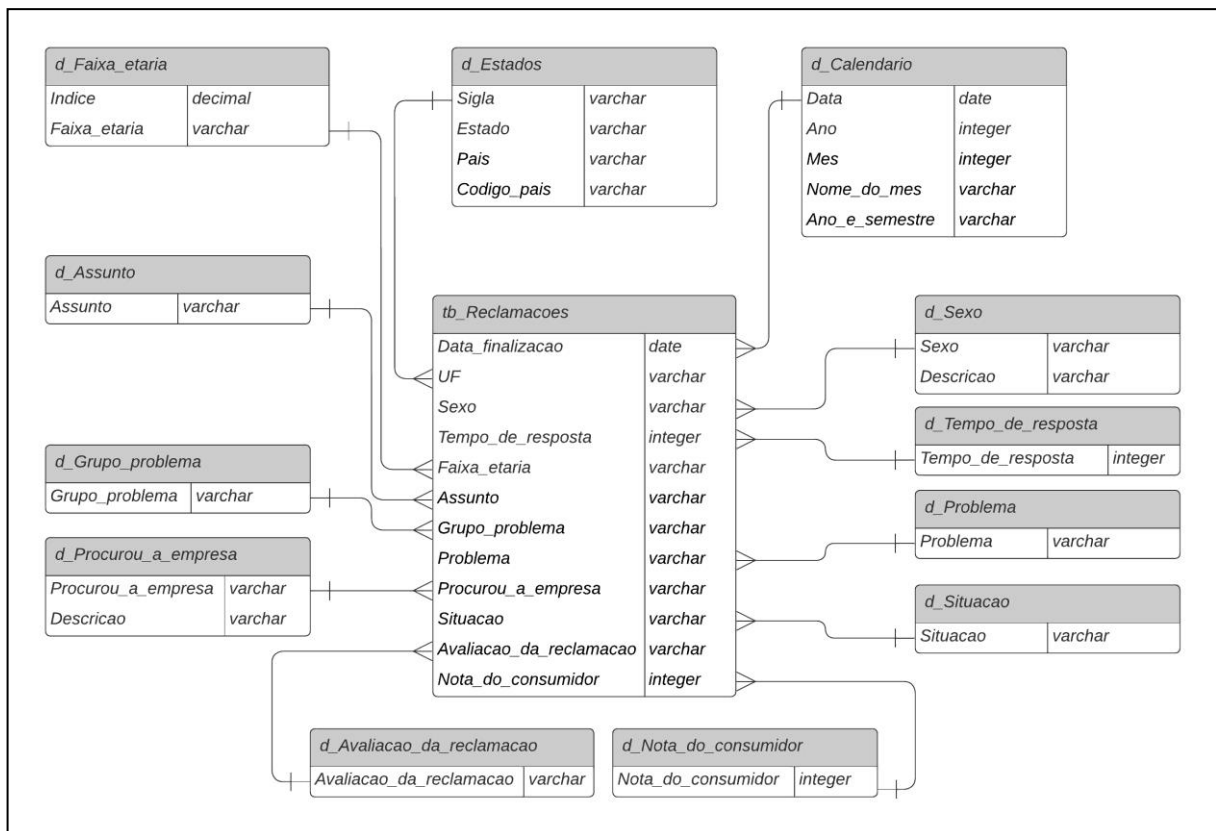
2. Modelo de Dados

2.1. Modelo Dimensional

Para o presente projeto, selecionamos o modelo dimensional Star Schema como estrutura central de armazenamento e análise de dados. A escolha do modelo Star Schema é fundamentada em sua eficácia comprovada na área de Business Intelligence e análise de dados. De acordo com Kimball (2002), o modelo Star Schema é amplamente reconhecido por sua simplicidade, facilidade de compreensão e capacidade de otimizar consultas complexas, tornando-o uma escolha ideal para a criação de um ambiente de análise de dados ágil e de alto desempenho.

A estrutura do modelo inclui uma tabela de fatos chamada "tb_Reclamacoes" e as tabelas de dimensão: "d_Assunto," "d_Avaliacao_da_reclamacao," "d_Calendario," "d_Estados," "d_Faixa_etaria," "d_Grupo_problema," "d_Nota_do_consumidor," "d_Problema," "d_Procurou_a_empresa," "d_Sexo," "d_Situacao," e "d_Tempo_de_resposta." A disposição dessas tabelas pode ser visualizada na figura a seguir.

Figura 1 – Modelo Dimensional



Fonte: elaborado pelo autor (2023)

2.2. Fatos e Dimensões

Aqui serão descritas as tabelas fato e dimensões que compõem o modelo dimensional do projeto.

Figura 2 – Descrição das Tabelas Fato e Dimensão

Tipo	Nome	Descrição	Exemplo
Fato	tb_Reclamacoes	Registros das reclamações	
Dimensão	d_Assunto	Assuntos	Aplicativo de mobilidade urbana, Serviços de entrega / delivery
Dimensão	d_Avaliacao_da_reclamacao	Reclamação resolvida	Resolvida, Não Resolvida
Dimensão	d_Calendario	Datas	01/06/2021, 18/09/2022
Dimensão	d_Estados	Unidades Federativas (UF)	SP, RJ, MG
Dimensão	d_Faixa_etaria	Faixas etárias	até 20 anos, entre 21 a 30 anos
Dimensão	d_Grupo_problema	Grupo do problema da reclamação	Cobrança / Contestação, Entrega do Produto
Dimensão	d_Nota_do_consumidor	Notas dos consumidores	1, 2, 3, 4, 5
Dimensão	d_Problema	Problemas	Reajuste - Discordância / dúvida, Cobrança após cancelamento do serviço
Dimensão	d_Procurou_a_empresa	Contato prévio com a empresa	Sim, Não
Dimensão	d_Sexo	Identidade de gênero	Feminino, Masculino, Outros
Dimensão	d_Situacao	Status da reclamação	Finalizada Avaliada, Finalizada não avaliada
Dimensão	d_Tempo_de_resposta	Dias para finalização da reclamação	0, 1, 2 ,3, 4

Fonte: elaborado pelo autor (2023)

2.2.1 Fato Reclamações

A tabela fato central deste projeto é a tabela Reclamações, representada pela tabela “tb_Reclamacoes” na figura 1, desempenha um papel fundamental ao armazenar e consolidar os dados de reclamações dos usuários.

A tabela "tb_Reclamacoes" mantém relacionamentos com as tabelas de dimensão do modelo e a granularidade dos dados nesta tabela é diária, permitindo acompanhar as tendências das reclamações ao longo do tempo.

2.2.2 Dimensão Assunto

A tabela dimensão Assunto, representada pela tabela “d_Assunto” na figura 1, apresenta os possíveis assuntos disponíveis ao usuário para seleção durante o cadastro da sua reclamação.

2.2.3 Dimensão Avaliação da Reclamação

A tabela dimensão Avaliação da Reclamação, representada pela tabela “d_Avaliacao_da_reclamacao” na figura 1, apresenta as opções de seleção disponíveis ao usuário que permitem indicar se a reclamação foi, ou não, resolvida após receber a resposta final da empresa, bem como a segmentação e análise dos dados posteriormente.

2.2.4 Dimensão Calendário

A tabela dimensão Calendário, representada pela tabela “d_Calendario” na figura 1, inclui valores em formato de data para melhor organizar a análise dos dados temporais.

2.2.5 Dimensão Estados

A tabela dimensão Estados, representada pela tabela “d_Estados” na figura 1, inclui valores para agrupamento das reclamações por estado no Brasil para melhor organizar a análise do perfil dos usuários.

2.2.6 Dimensão Faixa Etária

A tabela dimensão Faixa Etária, representada pela tabela “d_Faixa_etaria” na figura 1, inclui valores para agrupamento dos usuários por faixas de idade para melhor organizar a análise do perfil dos usuários.

2.2.7 Dimensão Grupo do Problema

A tabela dimensão Grupo do Problema, representada pela tabela “d_Grupo_problema” na figura 1, apresenta os possíveis grupamentos de problemas disponíveis ao usuário para seleção durante o cadastro da sua reclamação.

2.2.8 Dimensão Nota do Consumidor

A tabela dimensão Nota do Consumidor, representada pela tabela “d_Nota_do_consumidor” na figura 1, apresenta as opções de seleção disponíveis ao usuário para atribuir uma nota ao atendimento da empresa, após receber a resposta final da empresa em sua reclamação.

2.2.9 Dimensão Problema

A tabela dimensão Problema, representada pela tabela “d_Problema” na figura 1, apresenta os possíveis problemas disponíveis ao usuário para seleção durante o cadastro da sua reclamação.

2.2.10 Dimensão Procurou a Empresa

A tabela dimensão Procurou a Empresa, representada pela tabela “d_Procurou_a_empresa” na figura 1, apresenta as opções de seleção disponíveis ao usuário para informar se houve contato prévio com a empresa antes do registro da reclamação no portal.

2.2.11 Dimensão Sexo

A tabela dimensão Sexo, representada pela tabela “d_Sexo” na figura 1, inclui valores para agrupamento dos usuários por sexo para melhor organizar a análise do perfil dos usuários.

2.2.12 Dimensão Situação

A tabela dimensão Situação, representada pela tabela “d_Situacao” na figura 1, inclui valores para agrupamento das reclamações entre as que receberam avaliação dos usuários e as que não a receberam, ambas após receber a resposta final da empresa em sua reclamação.

2.2.13 Dimensão Tempo de Resposta

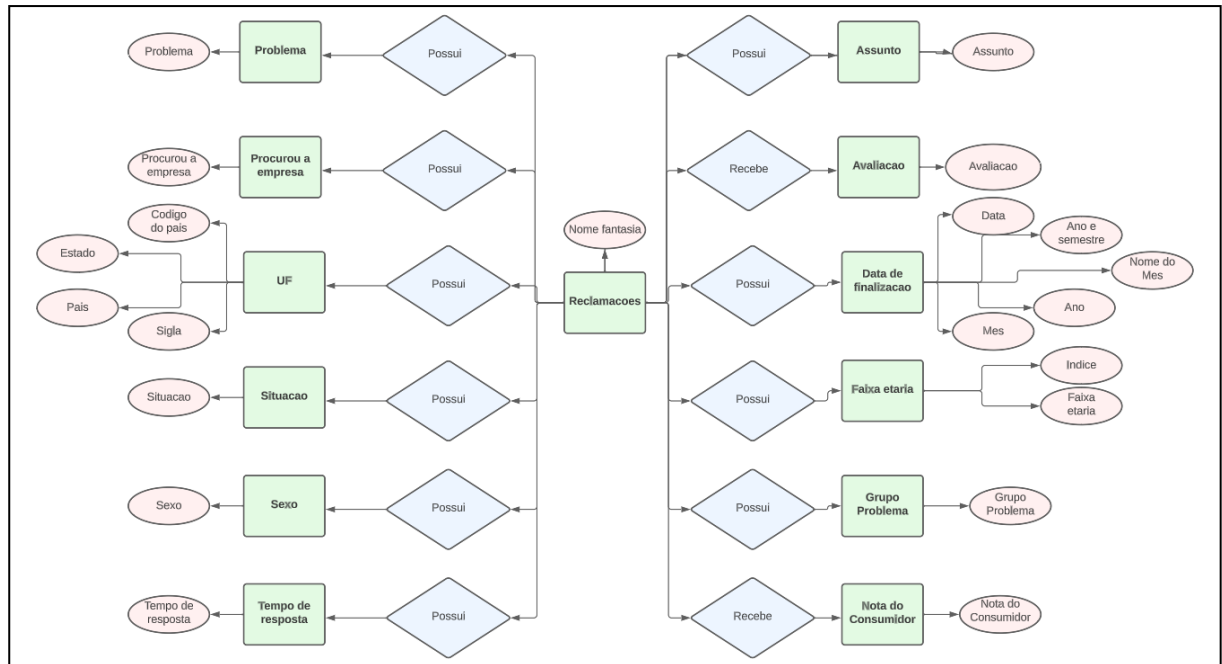
A tabela dimensão Tempo de Resposta, representada pela tabela “d_Tempo_de_resposta” na figura 1, apresenta os possíveis valores de tempo de resposta da reclamação por parte da empresa em cada registro.

3. Integração, Tratamento e Carga de Dados

3.1. Fontes de Dados

As fontes de dados utilizadas neste projeto são compostas por arquivos CSV disponibilizados mensalmente no site oficial do Consumidor.gov.br (BRASIL, 2023). Cada arquivo contém informações correspondentes a um mês/ano específico, permitindo-nos rastrear as reclamações ao longo do tempo. A partir dessas fontes de dados, foram extraídas a tabela de fato e as tabelas de dimensão, que desempenham papéis fundamentais em nosso modelo dimensional.

Para fornecer uma representação visual clara e estruturada das tabelas e seus relacionamentos em nosso modelo de dados, foi criado um Diagrama Entidade-Relacionamento (DER). Este diagrama desempenha um papel fundamental na compreensão da estrutura do nosso modelo dimensional e nas interações entre as tabelas.

Figura 3 – Diagrama Entidade Relacionamento (DER)

Fonte: elaborado pelo autor (2023)

No diagrama, podemos observar os objetos envolvidos, suas características e como se relacionam entre si.

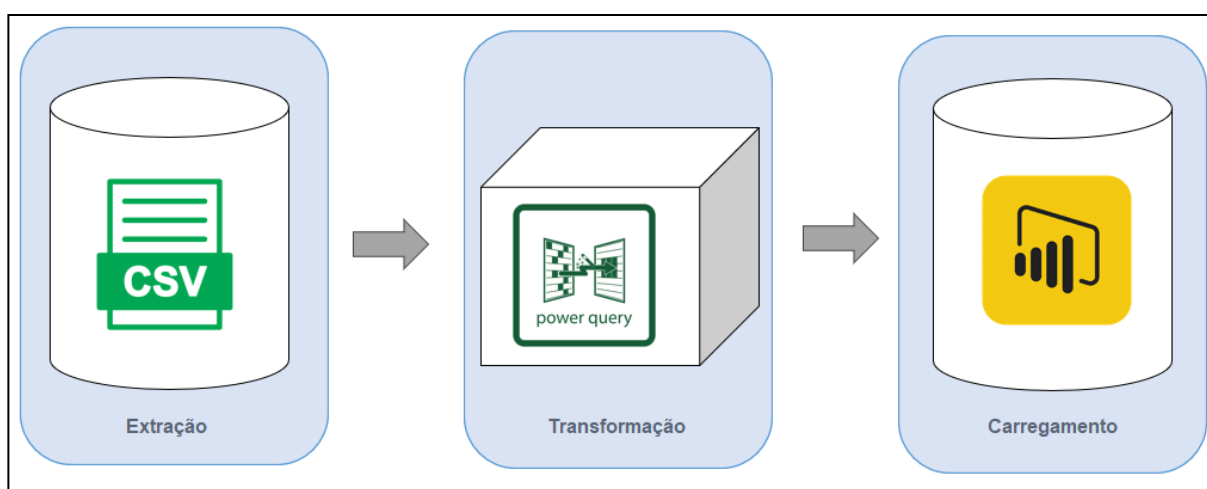
Os relacionamentos entre as tabelas são claramente representados no DER, permitindo-nos entender como as dimensões estão conectadas à tabela de fato e como podemos explorar os dados por meio de consultas e análises multidimensionais.

O Diagrama Entidade-Relacionamento apresentado fornece uma visão abrangente da estrutura do nosso modelo dimensional, destacando as tabelas de dimensão e a tabela de fato, bem como os relacionamentos que possibilitam análises ricas e detalhadas dos dados selecionados.

3.2. Processos de Integração e Carga (ETL)

O processo de Integração e Carga (ETL) foi realizado no ambiente do Power BI com apoio da ferramenta Power Query. Nesta seção, fornecemos uma visão geral das principais etapas realizadas durante o ETL, destacando os principais procedimentos adotados para transformar os dados brutos em um formato adequado para análise.

Figura 4 – Processo ETL



Fonte: elaborado pelo autor (2023)

3.2.1 Extração de Dados

A extração dos dados é realizada por meio do Power Query, ferramenta integrada ao ambiente do Power BI. Os dados são obtidos diretamente do site oficial Consumidor.gov.br (BRASIL, 2023), disponíveis em arquivos no formato CSV, e os conectores específicos do Power Query são utilizados para efetuar a leitura direta desses arquivos.

Os dados são automaticamente atualizados pelo Power BI ao detectar alterações nos dados de origem, quando novos arquivos são baixados para o diretório de destino.

Considerando que cada novo conjunto de dados é disponibilizado em um novo arquivo (para download), não há indicação de registro de versões ou log de alterações.

3.2.2 Tratamento e Limpeza de Dados

A transformação e limpeza dos dados foi realizada com uso do Power Query, tendo dentre as principais etapas:

- Remoção de colunas não relevantes para o projeto;
- Filtro da coluna 'nome fantasia' para exibição dos registros referentes a 'Uber';
- Filtro da coluna 'Avaliação Reclamação' para exibição dos registros referentes a 'Avaliada';
- Remoção de linhas que contivessem algum valor ausente 'null';
- Conversão do tipo de dados em colunas com valores numéricos;

Cabe comentar que a base de dados não exigiu outros tipos de tratamento como acentuação, maiúsculas/minúsculas, eliminação de espaços extras, dentre outros.

3.2.3 Carga de Dados (detalhar as etapas de carga dos dados)

Os dados tratados foram integrados no ambiente do Power BI, onde ocorreu a união dos arquivos CSV em uma consulta unificada.

Após essa etapa, foram criados relacionamentos entre as tabelas de dimensão e a tabela de fato central, permitindo análises multidimensionais e a exploração de informações contextuais no Dashboard.

4. Camada de Apresentação

4.1 Dashboard

O propósito dos dashboards desenvolvidos neste projeto é fornecer suporte às diversas equipes de tomada de decisão na empresa, sendo utilizados para a análise dos resultados e avaliação da evolução ao longo do tempo.

Durante a elaboração dos painéis, foi dada ênfase ao objetivo da empresa de elevar a nota média atribuída pelos consumidores ao final de cada atendimento realizado por meio da plataforma de reclamações.

Para atingir esse fim, diferentes combinações, análises e segmentações de dados foram empregadas, visando enriquecer a compreensão dos resultados alcançados pela empresa e identificar possíveis áreas de aprimoramento.

4.2 Painel Estratégico

Este painel apresenta os principais indicadores estratégicos monitorados pela empresa. Sua visualização prioriza a exibição dos resultados de forma global e acumulada, e conta com alguns visuais segmentados para acompanhamento da sua evolução.

4.2.1 *Período (ano e semestre)*

Este visual serve de apoio para realização de filtros de períodos para análise. Sua seleção permite escolher um ou mais semestres para visualização.

4.2.2 *Total de reclamações*

Este indicador exibe o total de reclamações no período selecionado. São consideradas para a análise apenas as reclamações na situação 'finalizada' e 'avaliada'.

4.2.3 Nota média

Este visual exibe a nota média no período, obtida através da média das notas informadas pelos consumidores ao final da avaliação, quando informado.

4.2.4 Tempo médio de resposta

Este visual exibe o tempo médio para o encerramento da reclamação (em dias) obtido através da média dos dias entre a abertura da reclamação e sua indicação como 'finalizada'.

4.2.5 Reclamações com contato prévio com a empresa

Este visual apresenta uma análise do número total de reclamações, categorizado com base na informação se os clientes procuraram ou não a empresa antes de registrar a reclamação na plataforma.

4.2.6 Top 3 reclamações por grupo problema

Este visual destaca as três principais reclamações mais recorrentes, categorizadas por Grupo Problema.

4.2.7 Total de reclamações segmentada por 'resolvida' e 'não resolvida'

Este visual fornece uma visão anual do total de reclamações, destacando a distribuição entre reclamações resolvidas e não resolvidas.

4.2.8 Índice de avaliação da reclamação como 'resolvida' e 'não resolvida'

Este visual exibe o índice percentual em relação ao total de reclamações, dividido entre aquelas que foram resolvidas e as não resolvidas.

4.3 Painel Tático

Este painel apresenta diversas segmentações de dados e combinações de informações a fim de obter insights sobre as fontes de avaliação das reclamações.

O objetivo deste painel é identificar, a partir da necessidade apresentada pela equipe estratégica, quais as origens e notas atribuídas a partir de diversas segmentações de consumidores, buscando apoiar o direcionamento dos esforços da empresa para a melhoria contínua das avaliações recebidas.

4.3.1 Período (ano e semestre)

Este visual serve de apoio para realização de filtros de períodos para análise. Sua seleção permite escolher um ou mais semestres para visualização

4.3.2 Mapa da nota média por estado

Este visual apresenta o mapa da nota média por estado (UF), oferecendo uma representação geográfica da satisfação média do cliente em diferentes regiões do Brasil.

4.3.3 Nota média por faixa etária

Este visual oferece uma visão abrangente da distribuição da nota média, segmentada por faixa etária.

4.3.4 Nota média por assunto

Este visual exibe a nota média atribuída pelos consumidores para diferentes assuntos relacionados a reclamações, sendo 'Aplicativo de mobilidade urbana' e 'Serviço de entrega / delivery'.

4.3.5 Nota média por grupo/problema e período

Este visual oferece uma análise abrangente da nota média atribuída pelos consumidores, segmentada por Grupo Problema e ordenados por períodos semestrais.

4.3.6 Nota média por sexo e período

Este visual fornece uma análise detalhada da nota média atribuída pelos consumidores, segmentada por sexo e dividida por períodos semestrais.

4.3.7 Nota e tempo médios segmentados por reclamações 'resolvidas' e 'não resolvidas'

Este visual oferece uma visão abrangente das reclamações, destacando a nota média e quantidade de reclamações distribuído em cada tempo de resolução (em dias) e segmentando os resultados para casos resolvidos e não resolvidos.

4.4 Painel Operacional

A proposta deste painel foi criar uma visualização das reclamações, e outras informações elencadas como de principal influência na nota final atribuída pelos consumidores, em uma visualização contínua de 30 dias contados a partir do último registro finalizado, sem a possibilidade de alteração do período.

Foram incluídos também alguns visuais estáticos, visando apoiar a comparação entre os valores atuais obtidos pela empresa e a evolução do seu desempenho histórico.

4.4.1 Total de reclamações (30 dias)

Este visual exibe o total de reclamações nos 30 últimos dias contados da data da última reclamação finalizada.

4.4.2 Nota média (30 dias)

Este visual exibe a nota média nos últimos 30 dias contados da data da última reclamação finalizada, bem como sua comparação com a nota sugerida como objetivo (meta).

4.4.3 Tempo médio de resposta (30 dias)

Este visual exibe o tempo médio de resposta (em dias) nos últimos 30 dias contados da data da última reclamação finalizada, bem como sua comparação com o tempo sugerido como objetivo (meta).

4.4.4 Avaliação da reclamação (30 dias)

Este visual fornece uma visão dos últimos 30 dias do total de reclamações, contados da data da última reclamação finalizada, destacando a distribuição entre reclamações resolvidas e não resolvidas.

4.4.5 Total de reclamações por semestre

Este visual exibe a evolução do total de reclamações na visão semestral. Também exibe uma linha de 'média' de todo o período para apoiar a comparação.

4.4.6 Nota média por semestre

Este visual exibe a evolução da nota média na visão semestral. Também exibe uma linha de 'média' de todo o período para apoiar a comparação.

4.4.7 Tempo médio de resposta por semestre

Este visual exibe a evolução do tempo médio de resposta na visão semestral. Também exibe uma linha de 'média' de todo o período para apoiar a comparação.

4.4.8 Registros por faixa etária

Este visual exibe o total de registros nos últimos 30 dias contados da data da última reclamação finalizada, segmentado por faixa etária.

5. Registros de Homologação

Aqui serão apresentados os testes da solução desenvolvida, confrontando os dados apresentados no dashboard com consultas independentes realizadas na base de dados de origem.

Para a realização dos testes fora utilizada a ferramenta Jupyter Notebook e a construção da estrutura foi realizada com uso da linguagem Python. Com vista a apoiar a continuidade da construção dos demais elementos deste trabalho, foram realizados quatorze testes em diferentes indicadores do Dashboard, conforme relacionados a seguir:

5.1. Configuração e tratamentos iniciais

Nesta etapa foram realizadas as configurações iniciais do algoritmo, a importação da base de dados na forma bruta a partir do mesmo diretório utilizado pelo Dashboard, e a eliminação de parte das colunas e linhas não utilizadas em ambos os projetos, a fim de nivelar a fonte de dados para comparação mais fidedigna possível.

Figura 5 – Configuração e tratamentos iniciais

Testes de Validação ¶

Aqui serão apresentados os testes da solução desenvolvida, confrontando os dados apresentados no dashboard com consultas independentes realizadas na base de dados de origem.

Imports:

```
In [1]: # Importando bibliotecas necessárias
import os
import pandas as pd
```

Configuração:

```
In [2]: # Diretório onde os arquivos CSV estão armazenados
diretorio = 'C:/Users/user/OneDrive/Documentos/Pós Graduação/TCC2/Base de dados'
```

```
In [3]: # Lista para armazenar todos os DataFrames carregados dos arquivos CSV
dataframes = []
```

Pré tratamento dos Dados:

```
In [4]: # Loop através dos arquivos no diretório
for filename in os.listdir(diretorio):
    if filename.endswith('.csv'):
        # Caminho completo do arquivo
        filepath = os.path.join(diretorio, filename)

        # Carregando o arquivo CSV em um DataFrame
        df = pd.read_csv(filepath, delimiter=';', header=0)

        # Removendo as colunas especificadas
        colunas_para_remover = ["Região", "Cidade", "Como Comprou Contratou", "Respondida", "Segmento de Mercado", "Área"]
        df = df.drop(colunas_para_remover, axis=1)

        # Filtrando as linhas com 'Uber' na coluna 'Nome Fantasia'
        df = df.loc[df['Nome Fantasia'] == 'Uber']

        # Filtrando as linhas com 'Avaliação Reclamação' diferente de 'Não Avaliada'
        df = df.loc[df['Avaliação Reclamação'] != 'Não Avaliada']

        # Removendo linhas que contenham algum valor em branco em qualquer coluna
        df = df.dropna(how='any')

        # Adicionando o DataFrame à lista
        dataframes.append(df)
```

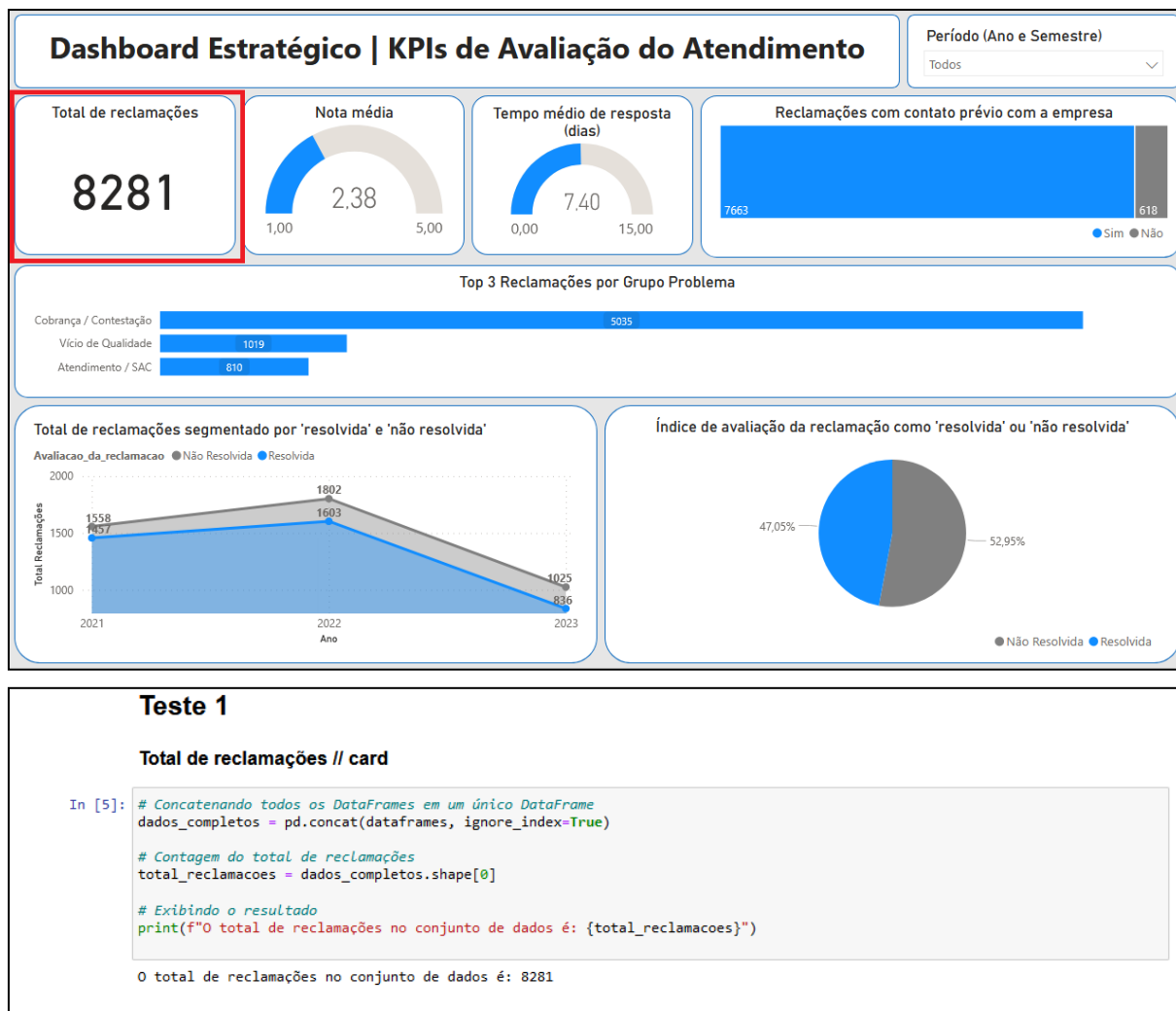
```
In [12]: dados_completos.columns
```

```
Out[12]: Index(['UF', 'Sexo', 'Faixa Etária', 'Data Finalização', 'Tempo Resposta',
               'Nome Fantasia', 'Assunto', 'Grupo Problema', 'Problema',
               'Procurou Empresa', 'Situação', 'Avaliação Reclamação',
               'Nota do Consumidor'],
              dtype='object')
```

Fonte: elaborado pelo autor (2023)

5.2. Teste: Total de reclamações

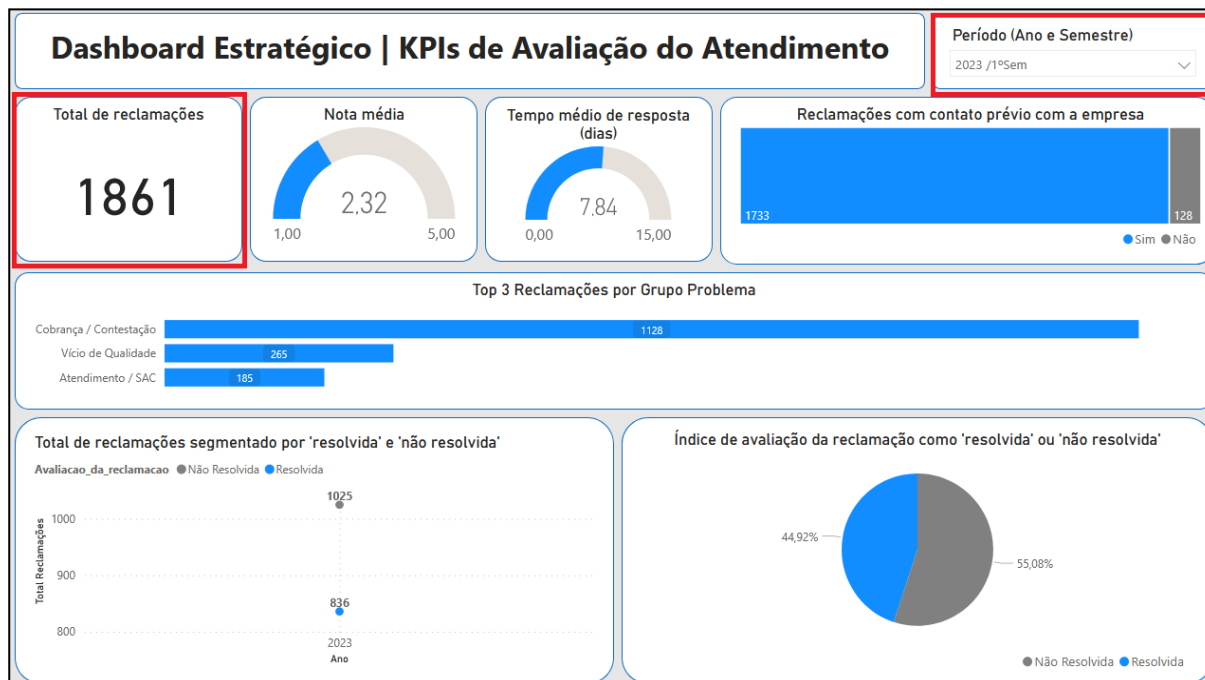
Figura 6 – Teste 1



Fonte: elaborado pelo autor (2023)

5.3. Teste: Total de reclamações por período

Figura 7 – Teste 2



Teste 2

Total de reclamações por período // card, segmentação de dados

```
In [11]: # Convertendo a coluna 'Data Finalização' para o formato de data
dados_completos['Data Finalização'] = pd.to_datetime(dados_completos['Data Finalização'], format='%d/%m/%Y', errors='coerce')

# Filtrando apenas os registros do primeiro semestre de 2023
dados_primeiro_semestre_2023 = dados_completos[(dados_completos['Data Finalização'].dt.year == 2023) &
(dados_completos['Data Finalização'].dt.month <= 6)]

# Contagem do total de reclamações no primeiro semestre de 2023
total_reclamacoes_primeiro_semestre_2023 = dados_primeiro_semestre_2023.shape[0]

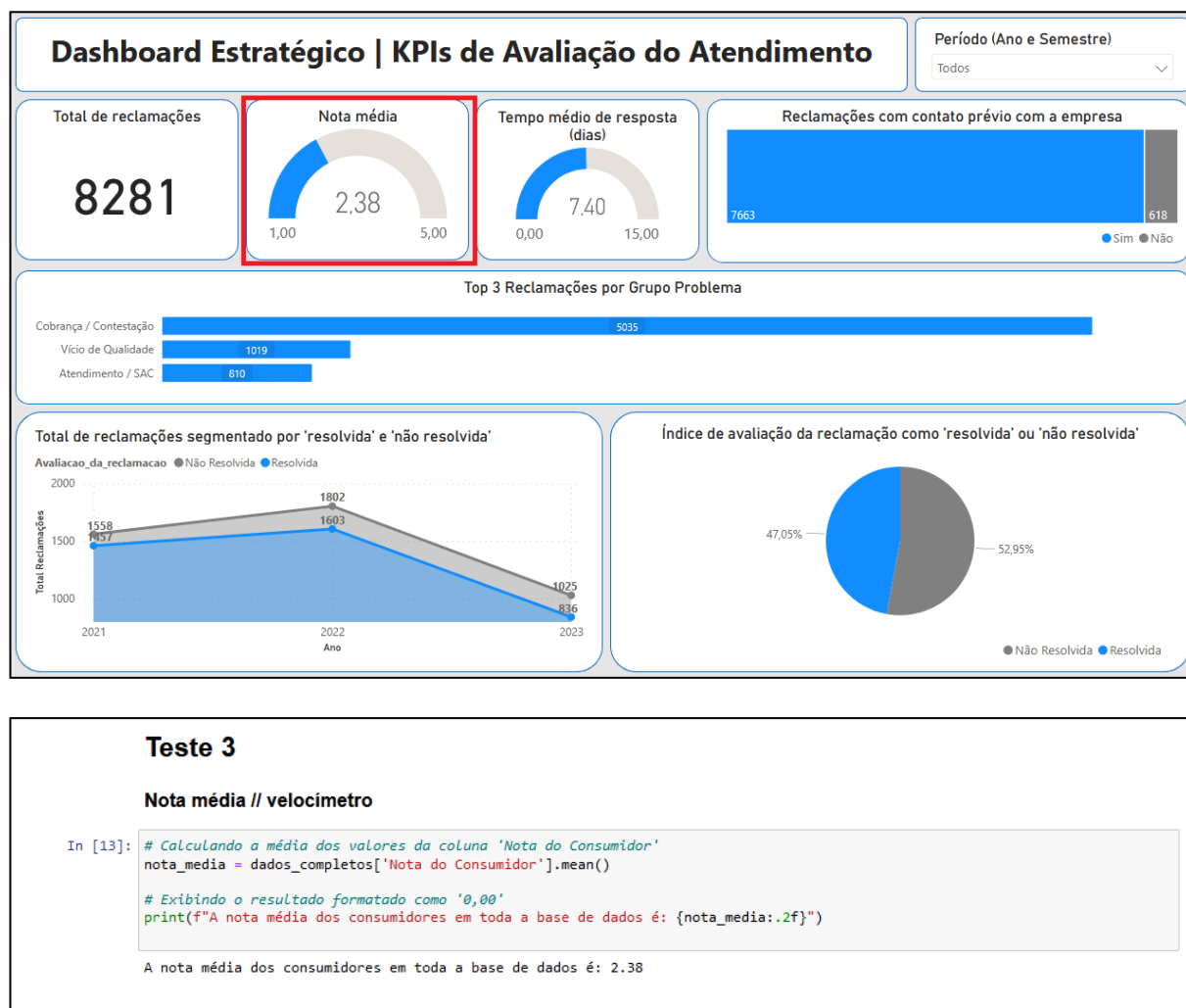
# Exibindo o resultado
print(f"O total de reclamações no primeiro semestre de 2023 é: {total_reclamacoes_primeiro_semestre_2023}")

O total de reclamações no primeiro semestre de 2023 é: 1861
```

Fonte: elaborado pelo autor (2023)

5.4. Teste: Nota média

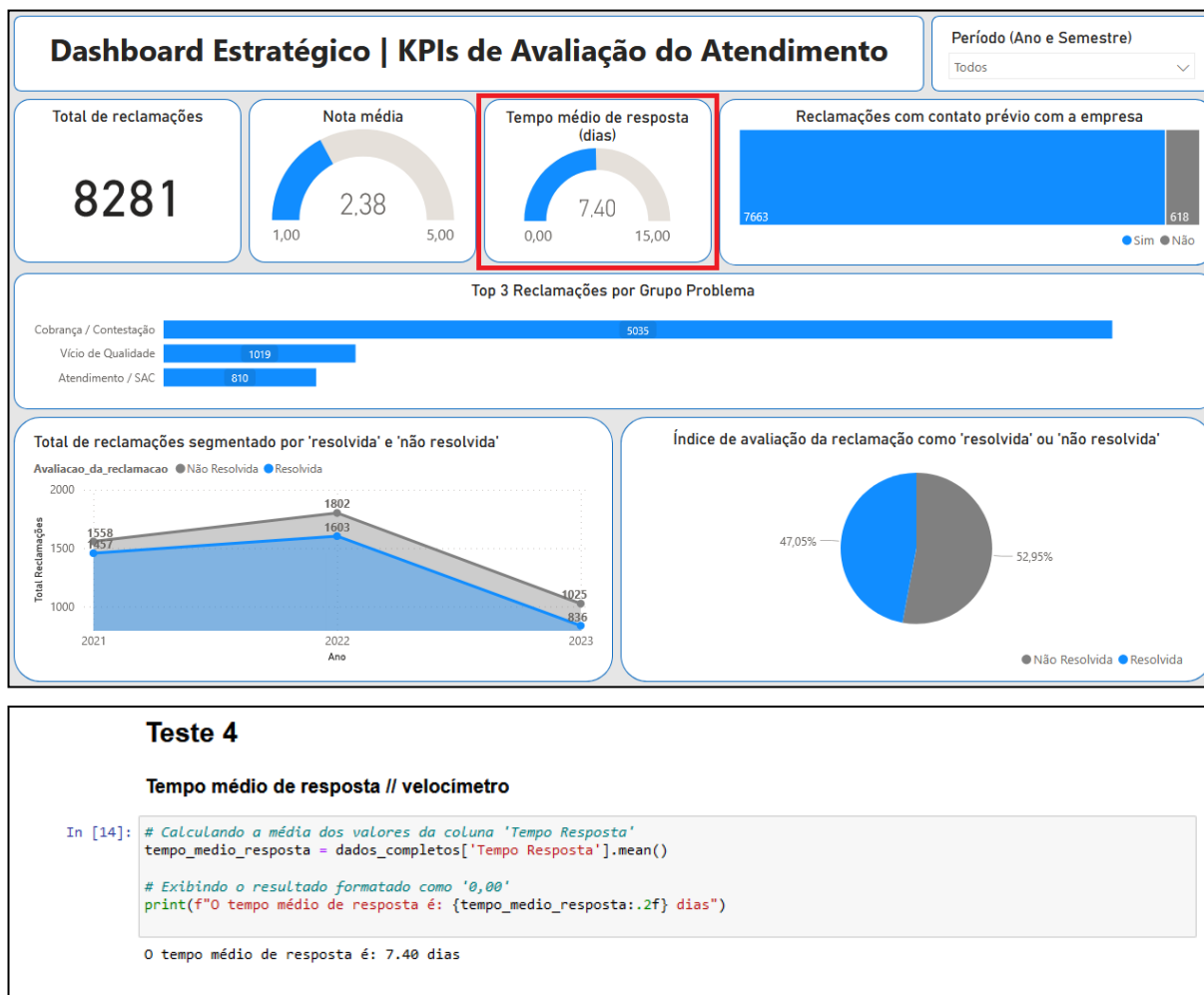
Figura 8 – Teste 3



Fonte: elaborado pelo autor (2023)

5.5. Teste: Tempo médio de resposta (dias)

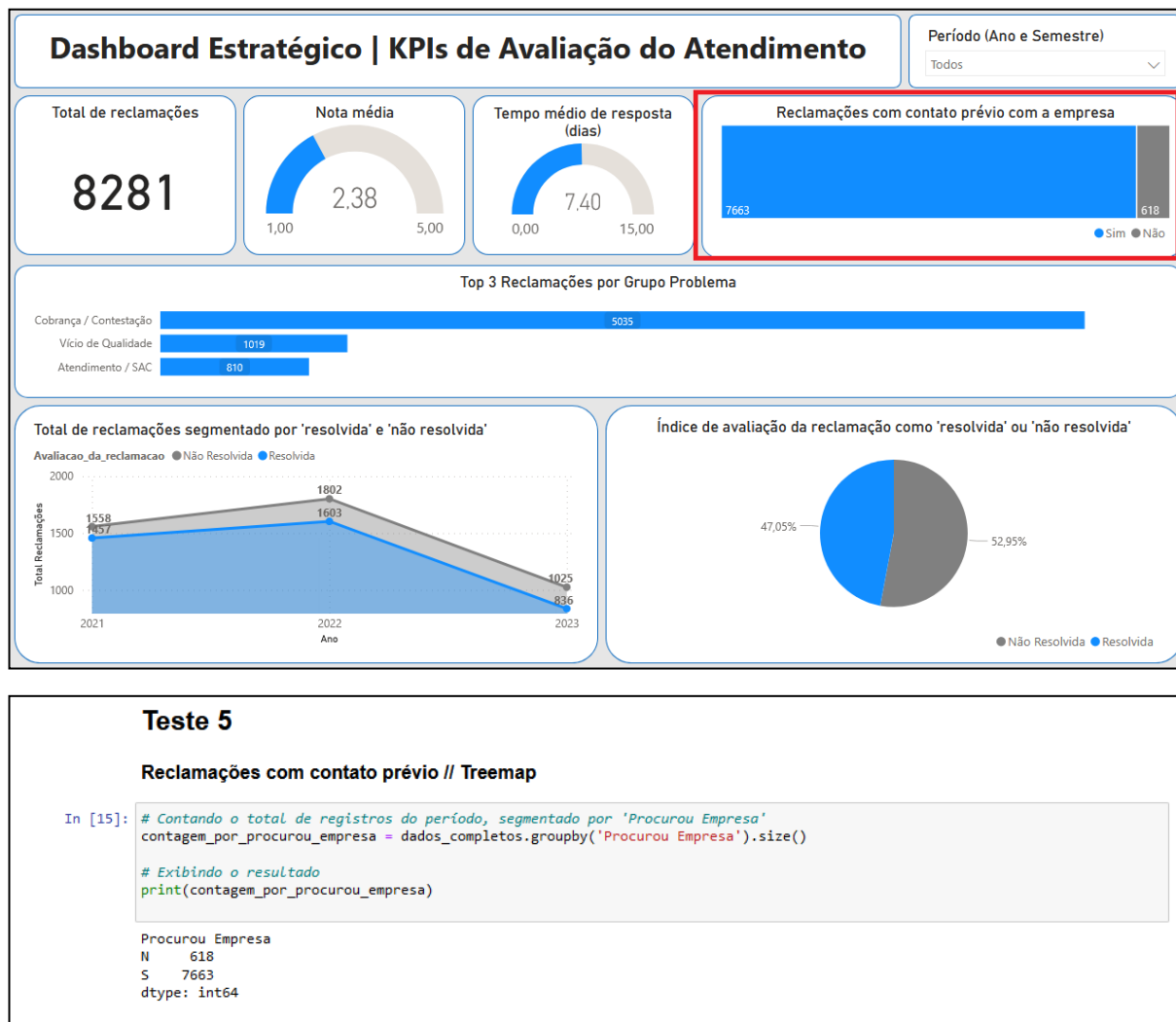
Figura 9 – Teste 4



Fonte: elaborado pelo autor (2023)

5.6. Teste: Reclamações com contato prévio com a empresa

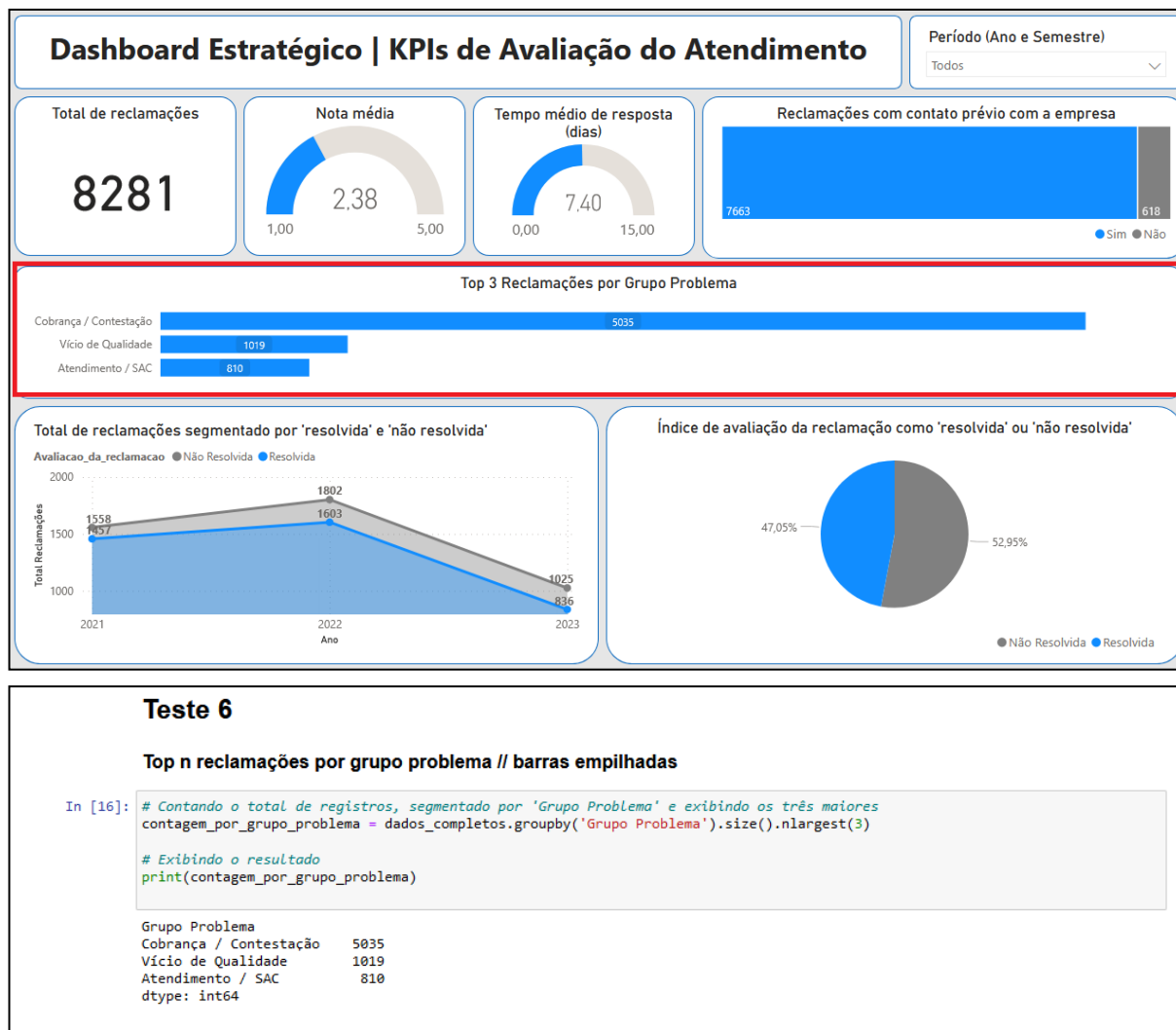
Figura 10 – Teste 5



Fonte: elaborado pelo autor (2023)

5.7. Teste: Top 3 Reclamações por Grupo Problema

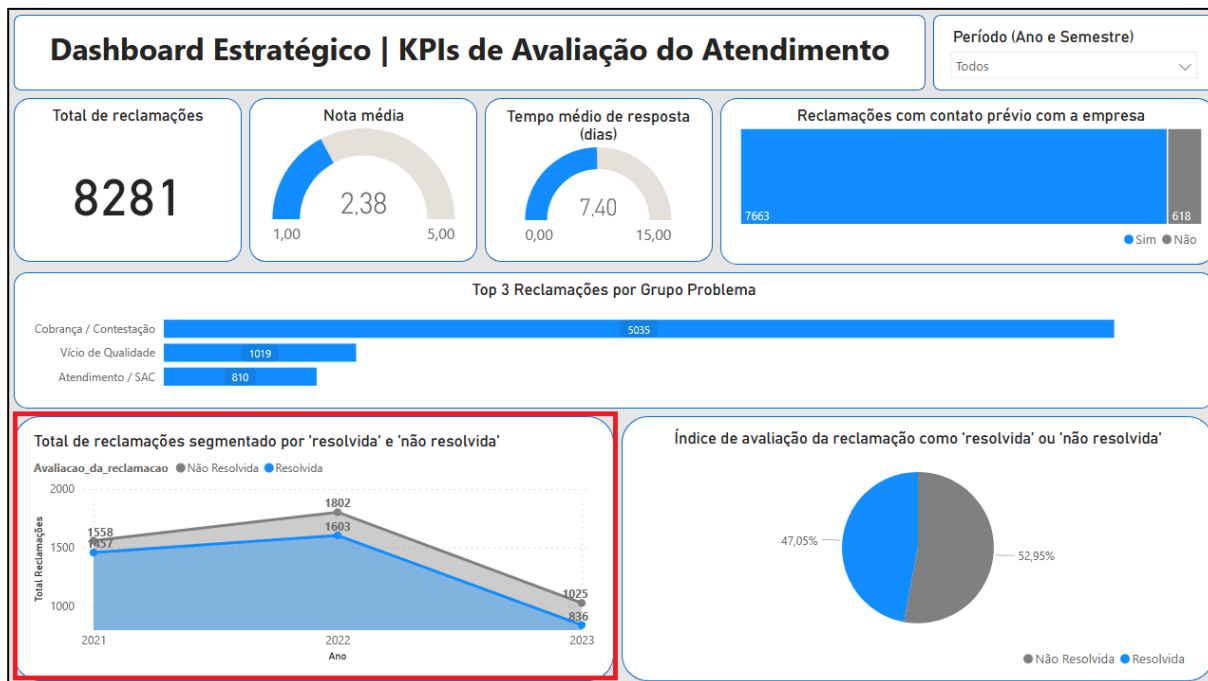
Figura 11 – Teste 6



Fonte: elaborado pelo autor (2023)

5.8. Teste: Total de reclamações segmentado por 'resolvida' e 'não resolvida'

Figura 12 – Teste 7



Teste 7

Total de reclamações segmentada por avaliação e ano // gráfico de área

```
In [17]: # Convertendo a coluna 'Data Finalização' para o formato de data
dados_completos['Data Finalização'] = pd.to_datetime(dados_completos['Data Finalização'], errors='coerce')

# Criando uma nova coluna para armazenar o ano da data de finalização
dados_completos['Ano'] = dados_completos['Data Finalização'].dt.year

# Contando o total de registros, segmentado por 'Avaliação Reclamação' e agrupado por ano
contagem_por_avaliacao_ano = dados_completos.groupby(['Avaliação Reclamação', 'Ano']).size()

# Exibindo o resultado
print(contagem_por_avaliacao_ano)
```

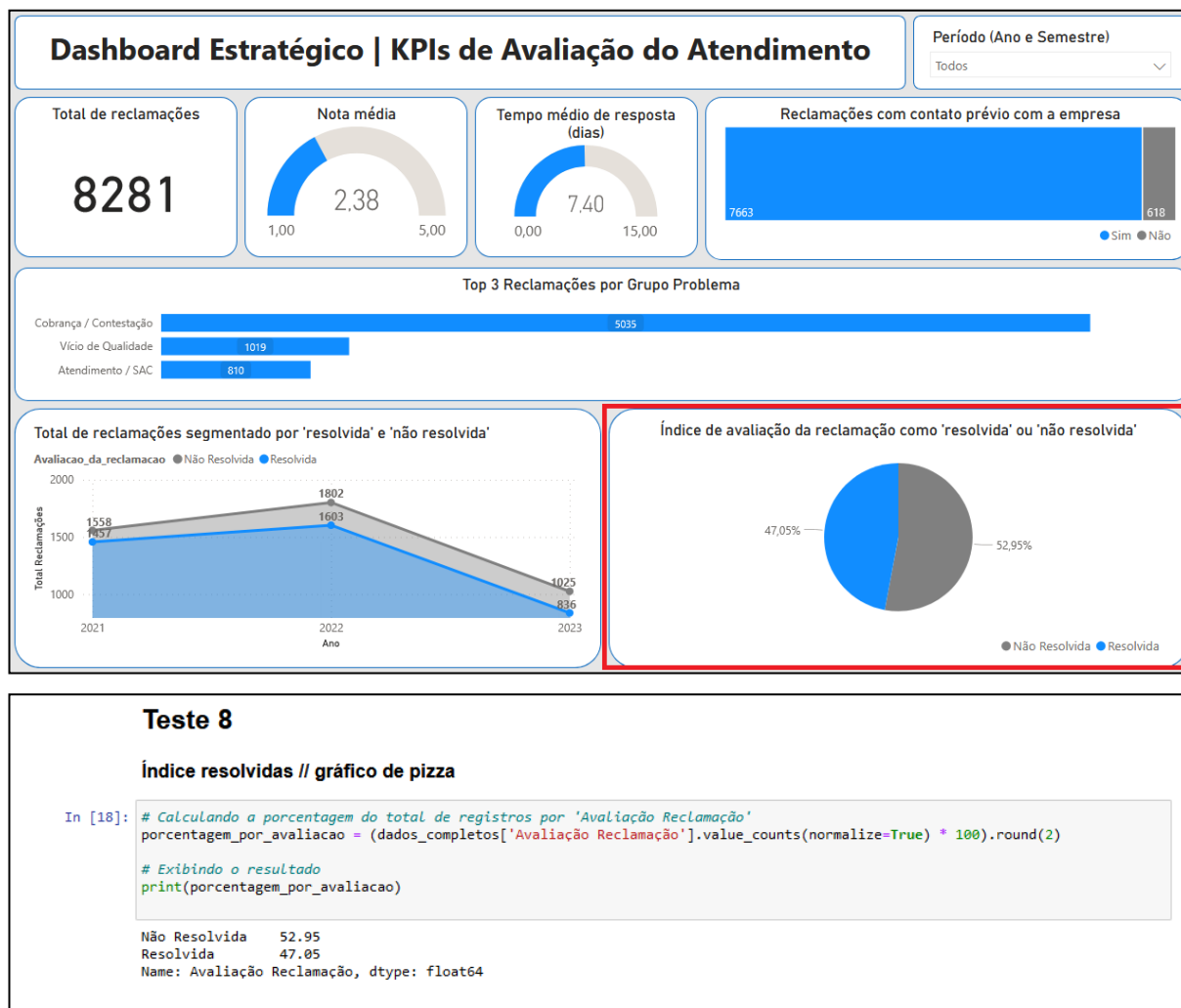
```
Avaliação Reclamação Ano
Não Resolvida      2021    1558
                  2022    1802
                  2023    1025
Resolvida          2021    1457
                  2022    1603
                  2023     836

dtype: int64
```

Fonte: elaborado pelo autor (2023)

5.9. Teste: Índice de avaliação da reclamação

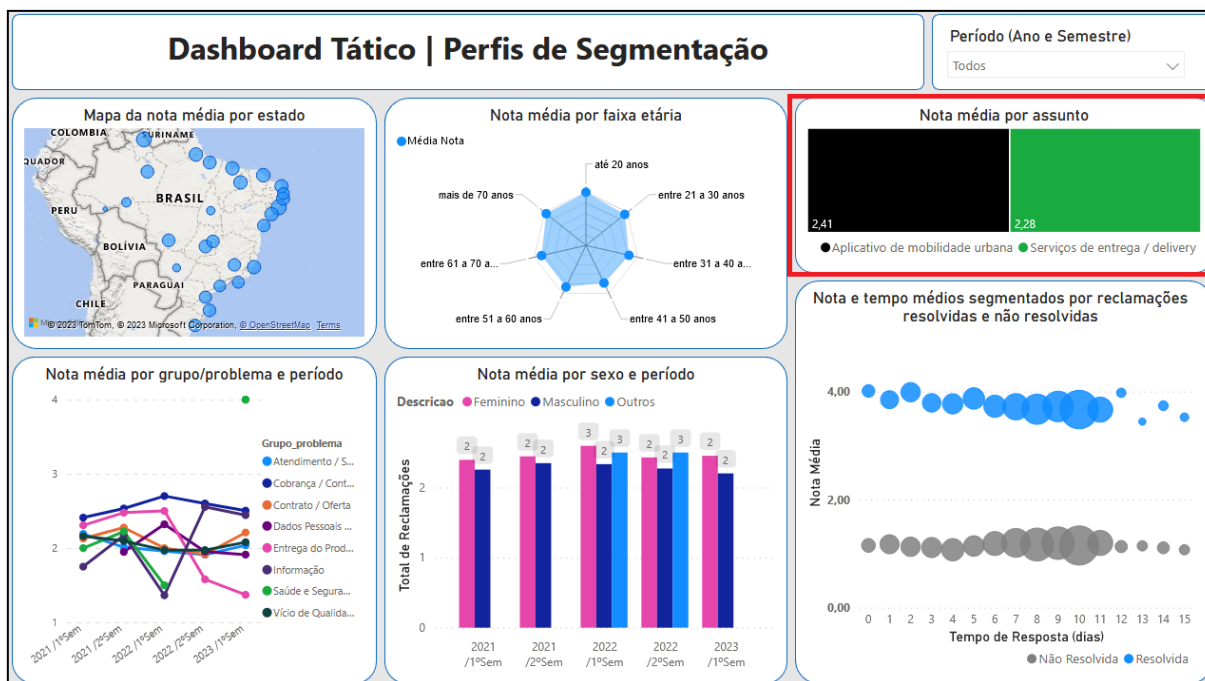
Figura 13 – Teste 8



Fonte: elaborado pelo autor (2023)

5.10. Teste: Nota média por assunto

Figura 14 – Teste 9

**Teste 9****Nota média por assunto // treemap**

```
In [20]: # Calculando a média da nota do consumidor por assunto
media_por_assunto = dados_completos.groupby('Assunto')['Nota do Consumidor'].mean()

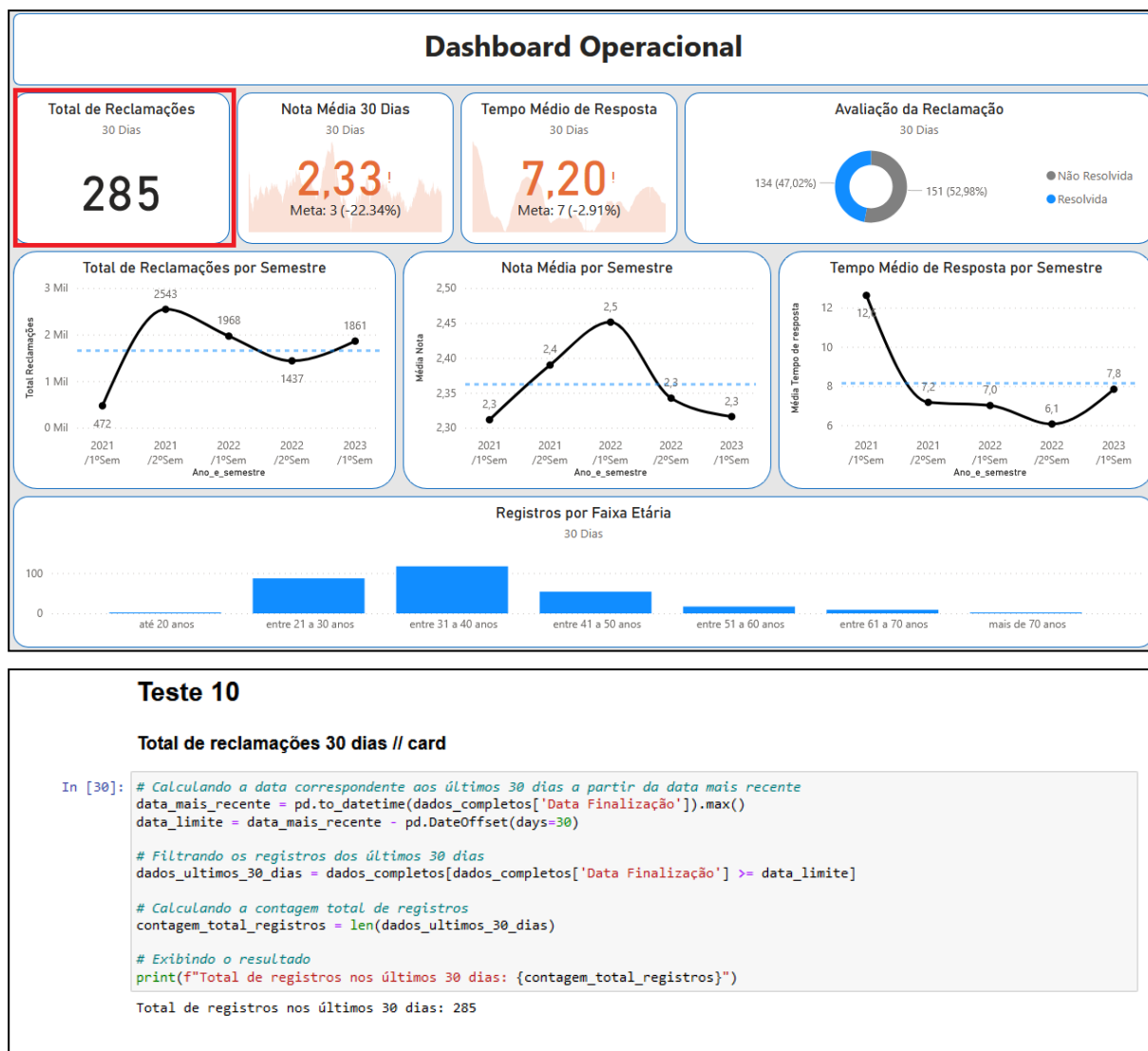
# Exibindo o resultado formatado com duas casas decimais
print(media_por_assunto.round(2))
```

```
Assunto
Aplicativo de mobilidade urbana    2.41
Serviços de entrega / delivery     2.28
Name: Nota do Consumidor, dtype: float64
```

Fonte: elaborado pelo autor (2023)

5.11. Teste: Total de reclamações (30 dias)

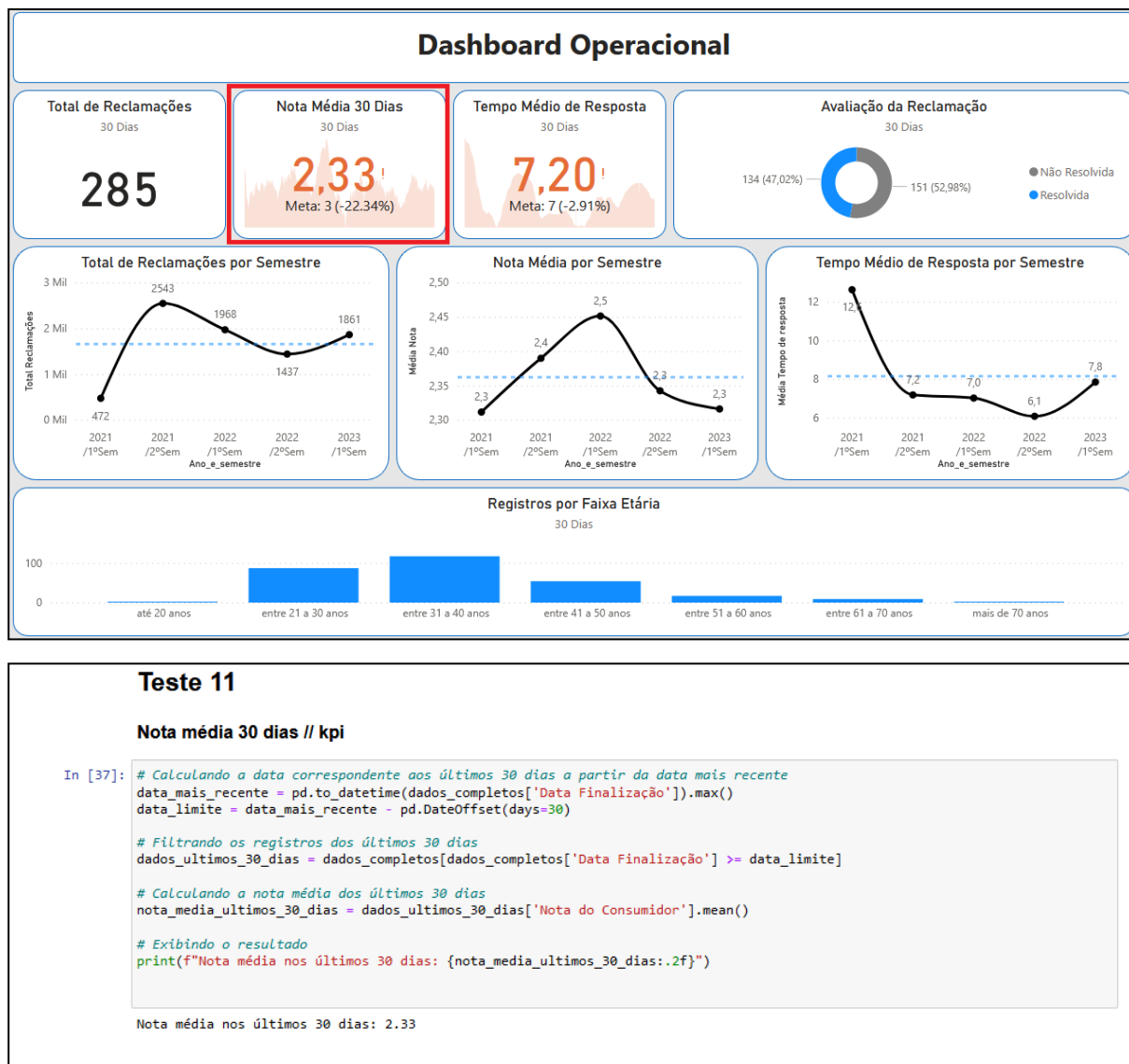
Figura 15 – Teste 10



Fonte: elaborado pelo autor (2023)

5.12. Teste: Nota média (30 dias)

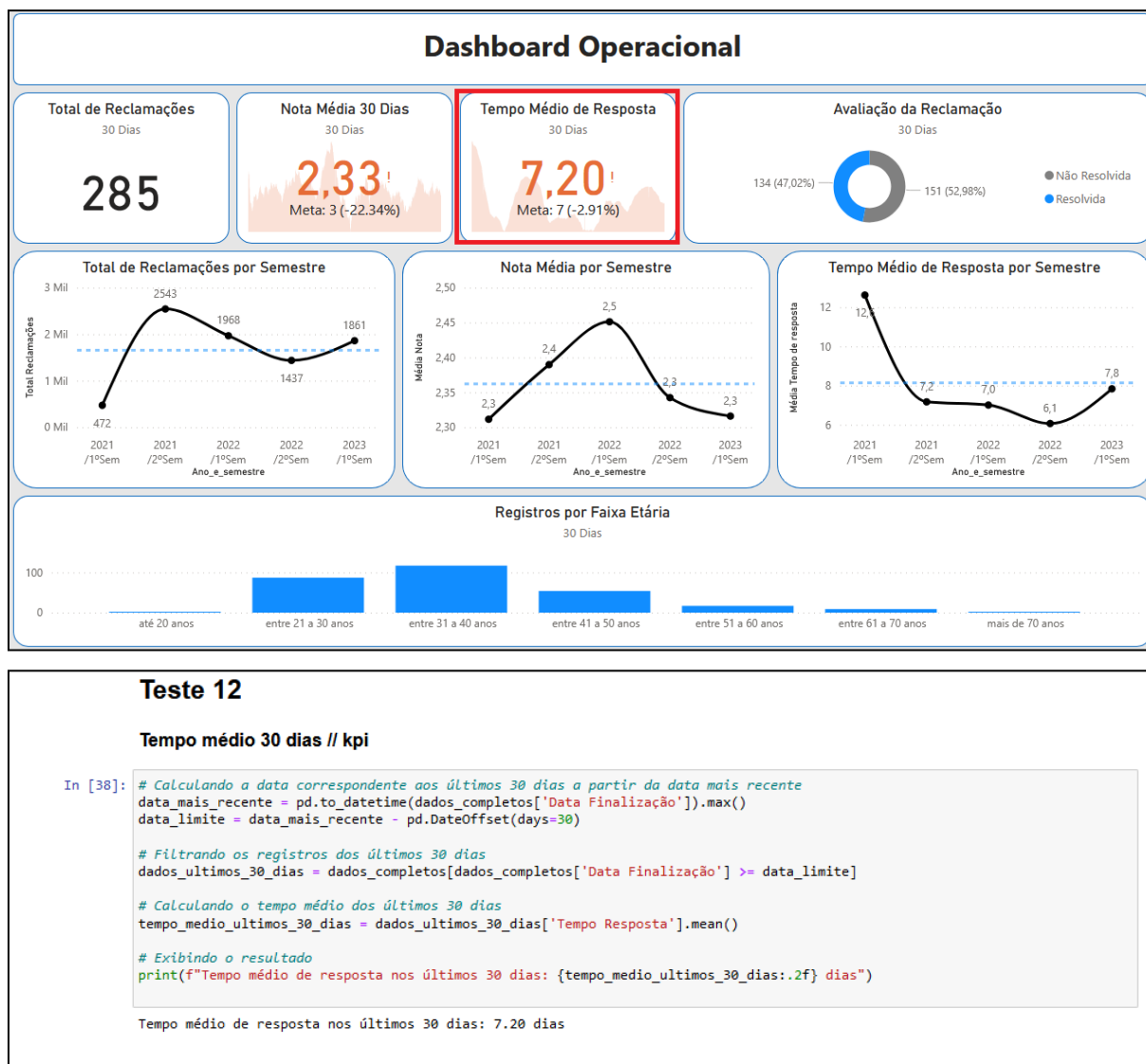
Figura 16 – Teste 11



Fonte: elaborado pelo autor (2023)

5.13. Teste: Tempo médio de resposta (30 dias)

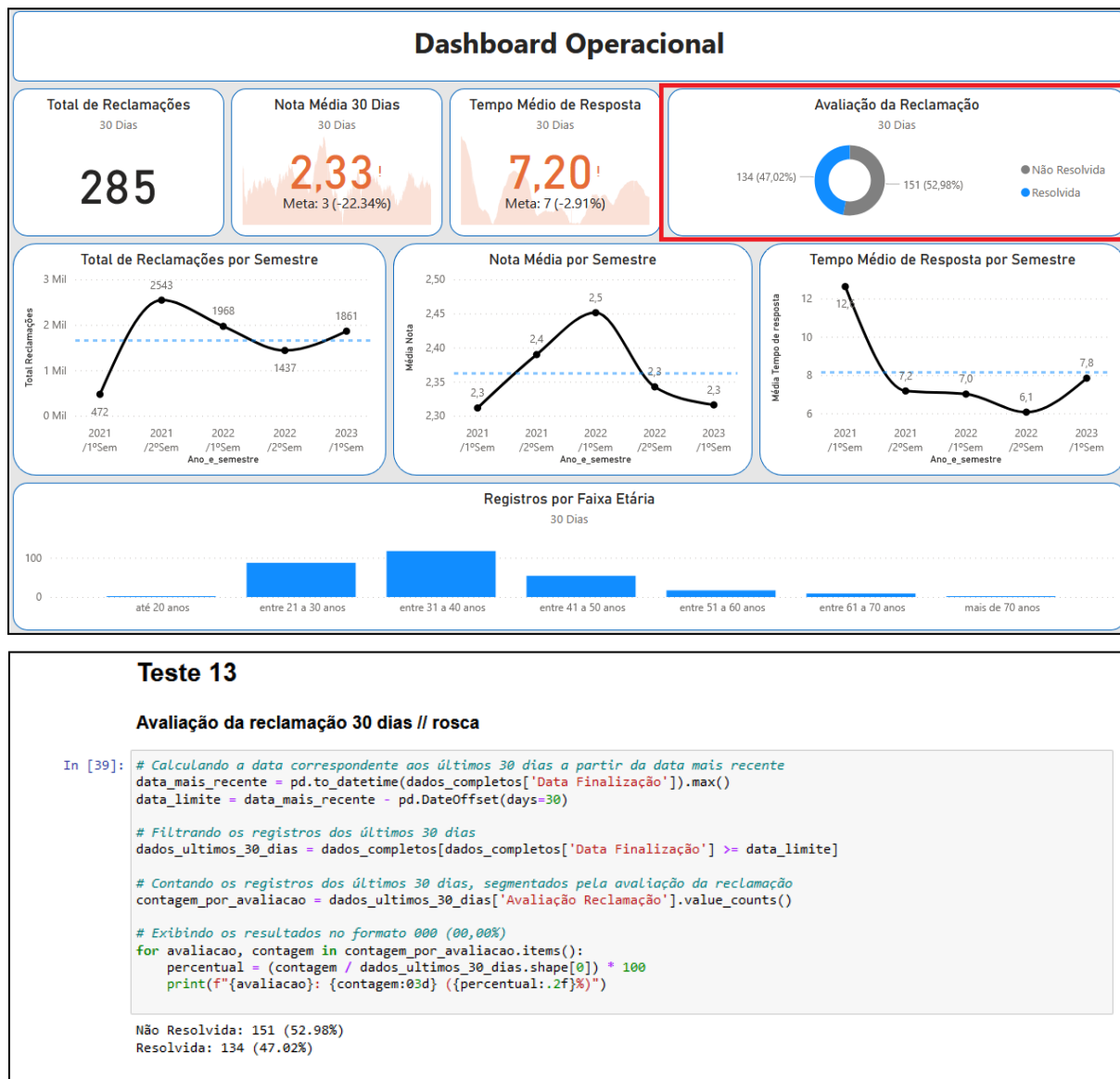
Figura 17 – Teste 12



Fonte: elaborado pelo autor (2023)

5.14. Teste: Avaliação da reclamação (30 dias)

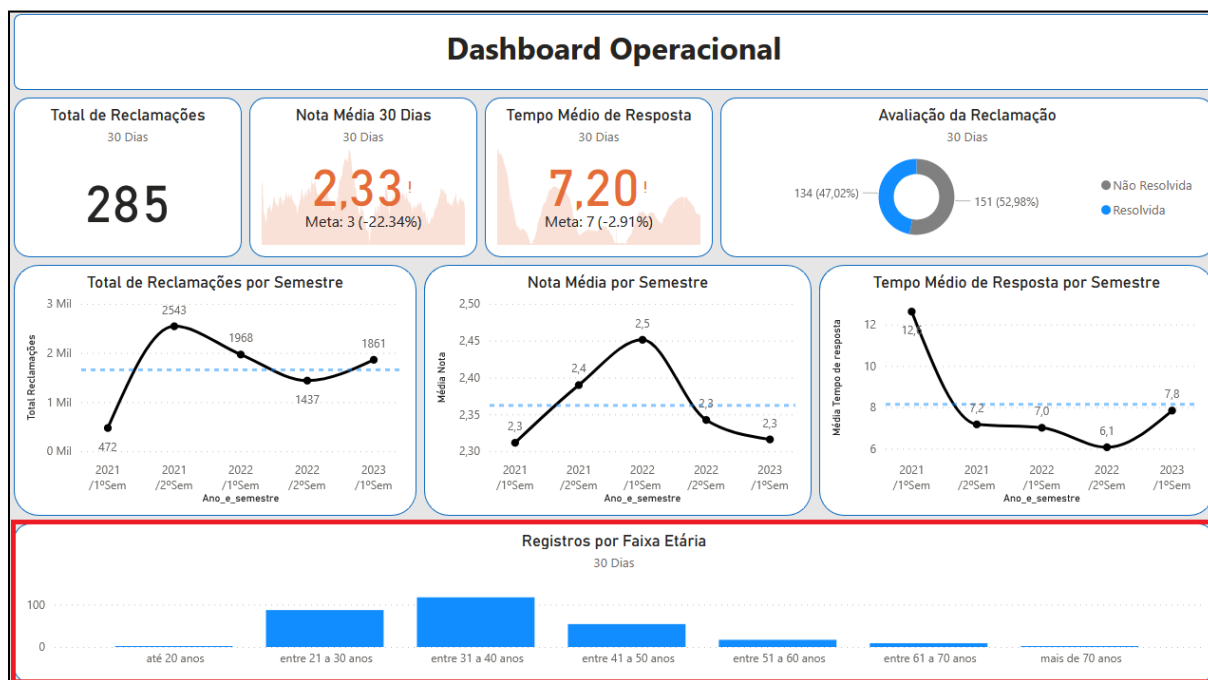
Figura 18 – Teste 13



Fonte: elaborado pelo autor (2023)

5.15. Teste: Registros por faixa etária (30 dias)

Figura 19 – Teste 14



Teste 14

Avaliação por faixa etária 30 dias // colunas empilhadas

```
In [47]: # Calculando a data mais recente
data_mais_recente = dados_completos['Data Finalização'].max()

# Calculando a data correspondente aos últimos 30 dias a partir da data mais recente
data_limite = data_mais_recente - pd.DateOffset(days=30)

# Filtrando os registros dos últimos 30 dias
dados_ultimos_30_dias = dados_completos[dados_completos['Data Finalização'] >= data_limite]

# Agrupando por faixa etária e contando os registros
contagem_por_faixa_etaria = dados_ultimos_30_dias.groupby('Faixa Etária').size()

# Exibindo o resultado
print(contagem_por_faixa_etaria)
print(f"Total de registros: {dados_ultimos_30_dias.shape[0]}")
```

```
Faixa Etária
até 20 anos      2
entre 21 a 30 anos  87
entre 31 a 40 anos 117
entre 41 a 50 anos  54
entre 51 a 60 anos  17
entre 61 a 70 anos   7
mais de 70 anos    1
dtype: int64
Total de registros: 285
```

Fonte: elaborado pelo autor (2023)

5. Conclusões

5.1 Análise

A seguir serão relacionados alguns pontos de destaque que puderam ser observados durante a elaboração do presente trabalho.

Foi possível observar que, durante o período avaliado, a empresa apresentou dificuldade em elevar a nota de avaliação dos consumidores. Em todos os recortes temporais visualizamos notas médias de avaliação com valores próximos, o que demonstra que a percepção dos consumidores não mostrou nenhuma evolução.

Também foi possível notar que, em todo o período, a empresa manteve mais de 50% das reclamações com status 'não resolvida'. Veremos a seguir que o fato de não resolver o problema do consumidor tem grande peso na nota de avaliação recebida posteriormente.

Como citado anteriormente, e talvez seja o principal achado obtido na avaliação dos dados, o fato de a empresa resolver, ou não, a demanda dos consumidores tem um peso decisivo na nota a ser recebida. Tal conclusão pode ser apresentada no painel Tático através do indicador 'Nota e tempo médio segmentados por reclamações resolvidas e não resolvidas'.

Este visual demonstrou que, independentemente do prazo que a empresa precisou para resolver a demanda, a nota seria maior caso o problema fosse solucionado, e menor caso não o fosse.

A análise dos dados também colocou o problema 'Cobrança/Contestação' como principal fonte de reclamações dos consumidores, com mais de 60% dos registros no período avaliado. Em um caso prático, este poderia ser um ponto de partida para a empresa avaliar suas operações e mitigar a origem dessas reclamações.

Importante mencionar também que em mais de 90% dos casos os clientes tentaram solucionar as demandas junto à empresa em 1ª instância. Nesse sentido é possível considerar que a empresa poderia estruturar as operações do SAC para manter a tratativa dentro da sua estrutura, evitando que os consumidores necessitassem procurar a plataforma do governo para mediação.

Com relação ao perfil dos consumidores, é possível mencionar que pessoas do sexo feminino tentem a dar uma nota pouco maior que os demais públicos, e que nos últimos 30 dias de análise, mais de 70% das reclamações estiveram concentradas em pessoas entre 20 e 40 anos. A consideração dessas informações pode apoiar as equipes de marketing a direcionar suas campanhas.

5.2 Proposta de Intervenção

Dentro do que fora apresentado anteriormente, bem como da análise dos dados divulgados no Dashboard, elencamos a seguir algumas propostas de intervenção que poderiam ser apresentadas aos gestores da empresa:

- Avaliar os motivos pelos quais maioria das reclamações não estar sendo resolvida, ou, se estão, como os consumidores estão percebendo a entrega de valor por parte da equipe de atendimento;
- Identificar o motivo de o problema 'Cobrança/Contestação' receber a maioria das reclamações, por exemplo cobranças em duplicidade, dificuldade de estorno junto às credenciadoras, mal funcionamento do aplicativo, usabilidade do cliente durante a jornada no aplicativo, ou outros;
- Revisar os procedimentos operacionais da equipe do SAC que acolhe as reclamações em 1ª instância e o motivo dessas demandas transbordarem para a plataforma Consumidor.Gov, pois, após a tratativa em 2ª instância, boa parte dessas demandas foram dadas como 'resolvidas';
- Compreender o motivo pelo qual o serviço de delivery da empresa estar regredindo nas notas de avaliação dos consumidores deste o 2º semestre de 2022;

5.3 Lições Aprendidas

Durante a construção do presente trabalho foi possível compreender, de forma prática, a importância de cada etapa necessária para a elaboração de uma análise de dados de forma robusta, em especial para as etapas de preparação e limpeza dos dados e a etapa dos testes da solução desenvolvidas, onde foi possível corrigir as informações antes da sua apresentação.

Também se pôde compreender a importância da análise dos dados e quão rica pode ser sua interpretação a partir do cruzamento de informações, validações de insights e novas descobertas.

Enquanto usuários, ao preencheremos um formulário que, por vezes, solicita informações do nosso perfil, como sexo, região, idade etc., ou detalhes da transação que está sendo realizada, é possível questionar o motivo do fornecimento de tais informações, porém tais dados, em algum momento, podem estar sendo utilizados de forma positiva para que as empresas aprimorem sua forma de atuação.

5.4 Limitações e Pontos de Extensão

Neste trabalho foi utilizada apenas uma fonte de dados, obtida em uma plataforma aberta na internet. Não foram contemplados dados ou outras informações de posse da empresa analisada a fim de otimizar o tempo em pesquisa. A ausência desse complemento não prejudicou a elaboração das análises apresentadas, entretanto poderia enriquecer ainda mais suas informações.

Também foram elencados KPIs julgados, de forma empírica, como importantes para o contexto das áreas relacionadas. Em um eventual caso prático, prestando este serviço para uma empresa, por exemplo, a construção dos painéis contemplaria informações que a própria julgasse relevantes.

Por fim, uma sugestão para a continuidade desta análise poderia ser a complementação do projeto com modelos de aprendizado de máquina, para o desenvolvimento de futuras soluções que possam apoiar a empresa a antever possíveis padrões e tendências nas análises das reclamações dos consumidores.

6. Links

Aqui estão disponibilizados os links para os artefatos criados, vídeo com a apresentação e para o repositório contendo os códigos fontes e demais arquivos do projeto.

- Link para o Dashboard:

<https://app.powerbi.com/view?r=eyJrljoiZGNiNWMxMTEtMmNmOC00OTAwLTg5YzEtYTliMWQzZGY1MWFmliwidCI6IjE0Y2JkNWE3LWVjOTQtNDZiYS1iMzE0LWNjMGZjOTcyYTE2MSIsImMiOi9>

- Repositório GitHub

<https://github.com/reginaldo-carvalho/TCC-PUCMG>

- Repositório Google Drive (alternativo)

https://drive.google.com/drive/folders/1tRweIQfijVgZx_cahBdf7z0WUjGUHc4G?usp=sharing

- Vídeo de apresentação:

<https://www.youtube.com/watch?v=2liUZR4BaF8>

REFERÊNCIAS

- KIMBALL, R.; ROSS, M. **The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling**. 2. ed. Canada: John Wiley & Sons, 2002.
- BRASIL. Secretaria Nacional do Consumidor - Senacon - Ministério da Justiça. Consumidor.gov.br. Disponível em: <https://consumidor.gov.br/>. Acesso em: agosto de 2023.