

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
COLEGIADO DO CURSO DE ENGENHARIA DE PRODUÇÃO**

BERNARDO PELIÇÃO PISSARRA

**DESENVOLVIMENTO DE UM DASHBOARD INFORMATIVO
PARA ANÁLISE DE DADOS DO RESSARCIMENTO AO SUS
POR OPERADORAS DE PLANOS DE SAÚDE UTILIZANDO
POWER BI**

VITÓRIA - ES

2022

BERNARDO PELIÇÃO PISSARRA

**DESENVOLVIMENTO DE UM DASHBOARD INFORMATIVO
PARA ANÁLISE DE DADOS DO RESSARCIMENTO AO SUS
POR OPERADORAS DE PLANOS DE SAÚDE UTILIZANDO
POWER BI**

Trabalho de Conclusão de Curso apresentado ao Colegiado do Curso de Engenharia de Produção do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Bacharel em Engenharia de Produção.

Orientador: Prof. Dr. Wilian Hiroshi Hisatugujm

Co-orientador: Prof. Dr. Renato Elias Nunes de Moraes

VITÓRIA, ES

2022

BERNARDO PELIÇÃO PISSARRA

**DESENVOLVIMENTO DE UM DASHBOARD INFORMATIVO
PARA ANÁLISE DE DADOS DO RESSARCIMENTO AO SUS
POR OPERADORAS DE PLANOS DE SAÚDE UTILIZANDO
POWER BI**

Trabalho de Conclusão de Curso apresentado ao Colegiado do Curso de Engenharia de Produção do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Bacharel em Engenharia de Produção.

Aprovada em 27 de janeiro de 2023.

COMISSÃO EXAMINADORA

Prof. Dr. Wilian Hiroshi Hisatugujm
Universidade Federal do Espírito Santo
Orientador

Prof. Dra. Luciana Lee
Universidade Federal do Espírito Santo

Prof. Me. Monique Rangel do Nascimento de Moraes
Faculdade Multivix Cariacica

AGRADECIMENTOS

Agradeço primeiramente a Deus, por me proporcionar a oportunidade de batalhar pela conquista de um grande sonho. Aos meus pais, Wanderson Martins Pissarra e Izabel Cristina Pelição por sempre acreditarem em mim e me garantirem todo apoio necessário para esta importante realização.

Aos professores Wilian Hiroshi Hisatugu e Renato Elias Nunes de Moraes, agradeço por não medirem esforços para orientar e ensinar um aluno com prazos desafiadores. Esta obra foi modelada pelo conhecimento de grandes orientadores que dedicaram seu tempo para me proporcionar uma importante experiência em minha graduação.

Deixo um enorme agradecimento a todos os colegas e professores que participaram de minha jornada, tanto da Universidade Federal de Viçosa quanto na Universidade Federal do Espírito Santo. Esses, se fizeram presentes em momentos difíceis de dúvidas e questionamentos e me acompanharam nas horas dedicadas aos estudos.

RESUMO

O Sistema Único de Saúde (SUS) é um direito de todos os cidadãos brasileiros e, portanto, é possível que os beneficiários de operadoras privadas de plano de saúde também usem os serviços do sistema público. Neste cenário, surge a necessidade de regulamentar o ressarcimento ao SUS, o qual é devido quando o atendimento realizado é suportado pela operadora privada e o beneficiário o realiza com recursos do SUS. Diante da inadimplência quanto ao ressarcimento pelas operadoras privadas, uma expressiva dívida foi gerada. Os dados coletados durante o processo de ressarcimento ao SUS são divulgados, de forma aberta, no portal de dados do governo, pela Agência Nacional de Saúde Suplementar - ANS. O objetivo deste trabalho é desenvolver um *dashboard* informativo utilizando a ferramenta Power BI que, a partir das bases de dados fornecidas pela ANS, permita ao usuário analisar o contexto da utilização do SUS por beneficiários de plano de saúde no período de janeiro de 2016 até setembro de 2021. Conclui-se, ao final, a entrega do *dashboard* e a exemplificação de possíveis análises construídas a partir deste.

Palavras-chave: **Análise de dados, Ressarcimento ao SUS, Business Intelligence, Power BI, Operadoras de plano de saúde, Dashboard, Sistema Único de Saúde, SUS.**

ABSTRACT

The Unified Health System (SUS) is a right of all Brazilian citizens, and, therefore, it is possible that the beneficiaries of private health plan operators also use the services of the public system. In this scenario, there is a need to regulate the SUS reimbursement, that is due when the service provided is supported by the private operator and the beneficiary uses SUS resource. Considering that there is a great shortage of reimbursement by private operators a significant debt has been generated. The data collected during the SUS reimbursement process is publicly disclosed on the government data portal, by the National Supplementary Health Agency (ANS). The goal of this work is to develop an informative dashboard using the Power BI tool that, from the databases provided by ANS, allows the user to analyze the context of the use of SUS by beneficiaries of health plans in the period from January 2016 to September 2021. Finally, the delivery of the dashboard and the example of possible analyzes built from it are concluded.

Keywords: Data Analysis, SUS reimbursement, Business Intelligence, Power BI, Private health plan operator, Dashboard, Unified Health System, SUS.

LISTA DE FIGURAS

Figura 1 - Fluxograma do processo administrativo de cobrança do ressarcimento ao SUS.....	17
Figura 1 - Níveis de análise de dados.....	19
Figura 2 - Sistema de Bancos de Dados.....	21
Figura 3 - Representação de <i>Data Warehouse</i>	23
Figura 4 - Arquitetura de sistema de Business Intelligence.....	24
Figura 5 - Tela inicial do Power Bi.....	25
Figura 7 - Importação de bibliotecas.....	31
Figura 8 - Definição de fonte de dados.....	31
Figura 9 - Trecho do código utilizado para importar tabelas referentes a Base 1 e Base 2.....	32
Figura 10 - Trecho do código que retira as colunas idênticas entre Base 1 e 2.....	33
Figura 11 - União ano a ano Bases 1 e 2.....	33
Figura 12 - Concatenação final dos seis anos de estudo.....	34
Figura 13 – Classificação de atendimentos com CID não especificada.....	38
Figura 14 - Modelo relacional Power BI.....	39
Figura 15 - <i>Menu</i> superior painel 1.....	40
Figura 16 - Situação de cobrança por modalidade de operadora.....	41
Figura 17 - Valor dos atendimentos e IEP.....	42
Figura 18- <i>Menu</i> superior painel 2.....	42
Figura 19 - Atendimentos por sexo e UF do beneficiário.....	43
Figura 20 - Estudo da variável Dias de permanência interna (DPI).....	44
Figura 21 - Número e Valor de atendimentos por Faixa etária e Sexo dos beneficiários.....	45
Figura 22 - <i>Menu</i> superior painel 3.....	46
Figura 23 - Análise de esfera dos estabelecimentos.....	46
Figura 24 - Análise de natureza do estabelecimento.....	47
Figura 25 - Tipo de atendimentos e localização dos estabelecimentos.....	48
Figura 26 - <i>Menu</i> superior painel 4.....	48
Figura 27 - Análise de atendimentos.....	49
Figura 28 - Frequência, especialidade e situação de cobrança dos atendimentos.....	50

Figura 29 - Anomalia detectada no ano de 2020.....	51
Figura 30 - Modelo de previsão Holter-Winters.....	51
Figura 31 - Análise modalidade de operadora.....	53
Figura 32 - Índice de Efetividade de Pagamento (IEP) 2019.....	54
Figura 33 - Média de Dias de permanência interna (DPI).....	55
Figura 34 - Número de atendimentos por Faixa etária e Sexo de beneficiários.....	56
Figura 35 - Valor de atendimentos por Faixa etária e Sexo de beneficiários.....	56
Figura 36 - Análise de Esfera do estabelecimento.....	57
Figura 37 - Análise de Natureza do estabelecimento.....	58
Figura 38 - Heatmap de frequência de atendimentos por Município do estabelecimento.....	58
Figura 39 - Análise de frequência de atendimentos por capítulo da CID-1.....	59
Figura 40 - Análise do capítulo dos atendimentos referentes ao capítulo de Neoplasia da CID-10.....	60
Figura 41 - Código Python referente à preparação de dados da Base 3: Dados de pagamento do Ressarcimento ao SUS por operadora.....	75
Figura 42 - Capa do dashboard.....	86
Figura 43 - Sumário do dashboard.....	87
Figura 44 - Painel 1: Operadoras de plano de saúde.....	88
Figura 45 - Painel 2: Perfil dos beneficiários de plano de saúde.....	89
Figura 46 - Painel 3: Localização dos atendimentos.....	90
Figura 47 - Painel 4: Atendimentos.....	91
Figura 48 - Caso de Uso.....	92

LISTA DE TABELAS

Tabela 1 - Variáveis utilizadas.....	28
Tabela 2 - Bases de dados disponíveis no Power BI.....	34
Tabela 3 - Cardinalidade entre tabelas.....	38
Tabela 4 - Comparação previsão x real para o ano de 2020.....	52
Tabela 5 - Dicionário de dados referente a Base 1: Dados do Ressarcimento ao SUS por Beneficiário Identificado e ABI.....	66
Tabela 6 - Dicionário de dados referentes à Base 2: Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde.....	70
Tabela 7 - Dicionário de dados referente a Base 3: Dados de pagamento do Ressarcimento ao SUS.....	72
Tabela 8 - Dicionário de dados referente à Base 4: CID-10, Classificação Internacional de Doenças.....	73
Tabela 9 - Renomeação de variáveis referentes à Base 1: Dados do Ressarcimento ao SUS por Beneficiários Identificados e ABI.....	75

LISTA DE SIGLAS

ABI	Aviso de Beneficiário Identificado
ANS	Agência Nacional de Saúde Suplementar
DATASUS	Departamento de Informática do SUS
DBMS	<i>DataBase Management System</i>
DIDES	Diretoria de Desenvolvimento Setorial
FNS	Fundo Nacional de Saúde
GRU	Guia de Recolhimento da União
IEP	Índice de Efetivo Pagamento
PSS	Planos e Seguros Privados de Assistência à Saúde
SUS	Sistema Único de Saúde
TCU	Tribunal de Contas da União
AIH	Autorização de Internação Hospitalar
APAC	Autorização de Procedimentos Ambulatoriais

SUMÁRIO

1 INTRODUÇÃO.....	13
1.1 JUSTIFICATIVA.....	14
1.2 OBJETIVO.....	15
1.2.1 Objetivo Geral.....	15
1.2.2 Objetivos Secundários.....	15
2 REFERENCIAL TEÓRICO.....	16
2.1 RESSARCIMENTO AO SUS.....	16
2.2 MANIPULAÇÃO DE DADOS.....	18
2.3 PYTHON.....	20
2.4 BANCO DE DADOS.....	20
2.5 DATA WAREHOUSE.....	22
2.6 BUSINESS INTELLIGENCE.....	23
2.7 VISUALIZAÇÃO DE DADOS.....	24
3 MATERIAIS E MÉTODOS.....	26
3.1 COLETA DE DADOS.....	26
3.1.1 Descrição das bases de dados.....	26
3.1.2 Descrição das variáveis.....	28
3.2 PREPARAÇÃO DE DADOS EM PYTHON.....	31
3.2.1 Criação do ambiente de trabalho.....	31
3.2.2 Preparação de dados: Base 1 Dados do Ressarcimento ao SUS por Beneficiário Identificado e ABI, e Base 2 Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde.....	32
3.2.3 Preparação de dados: Base 3 Dados de pagamento do Ressarcimento ao SUS por operadora.....	34
3.3 FERRAMENTA DE BUSINESS INTELLIGENCE – POWER BI.....	34
3.3.1 Importação de dados.....	34
3.3.2 Tratamento e transformação de dados.....	35
3.3.2.1 Base de dados: Ressarcimento ao SUS.....	35
3.3.2.2 Base de dados: Dados de pagamento do Ressarcimento ao SUS por operadora.....	36
3.3.2.3 Base de dados: CID-10 SUBCATEGORIAS.....	37
3.3.2.4 Base de dados: CID-10 CAPÍTULOS.....	37
3.3.3 Criação dos relacionamentos.....	38
3.3.4 Dashboard.....	40
3.3.4.1 Painel 1 – Operadoras de plano de saúde.....	40
3.3.4.2 Painel 2 – Perfil do beneficiário de plano de saúde.....	42
3.3.4.3 Painel 3 – Localização dos atendimentos.....	45
3.3.4.4 Painel 4 – Atendimentos.....	48
3.4 CASO DE USO.....	50
4 RESULTADOS E DISCUSSÕES.....	53
5 CONCLUSÃO.....	61
REFERÊNCIAS.....	62
APÊNDICES.....	65

1 INTRODUÇÃO

Em outubro de 1988, a organização da saúde pública no Brasil foi completamente remodelada com a publicação da Constituição Federal Brasileira, que determinou a implementação do estado de bem-estar social no Estado democrático de direito. Então, a partir da definição de saúde como direito de cidadania, houve o início da criação do sistema público de saúde brasileiro (PAIVA; TEIXEIRA, 2014). Dois anos depois, com a Lei Orgânica da Saúde, aprovada em 1990, foram estabelecidos os preceitos que regem o Sistema Único de Saúde (SUS), um dos maiores sistemas públicos de saúde do mundo, sendo o único a garantir assistência integral e completamente gratuita a todos os brasileiros (MINAS GERAIS, 2015).

Tendo em vista o cenário de reconhecimento da saúde como condição essencial à vida e ao exercício da cidadania, o artigo 199 da Constituição Federal de 1988 trata, de forma geral, a assistência à saúde como bem de consumo individual e de livre comercialização à iniciativa privada. Desta forma, além do acesso gratuito aos serviços públicos de saúde, também é direito do cidadão brasileiro a possibilidade de comprar serviços no mercado com capital próprio, caso desejar.

Posteriormente, a partir de 1991, com a aprovação da Lei 9.656 pelo Congresso Nacional, elaborou-se no Brasil a regulamentação dos Planos e Seguros Privados de Assistência à Saúde (PSS) (CARVALHO, 2003). Nesse âmbito, levando em consideração que o direito à saúde é dever do Estado, mas não é determinado como monopólio estatal, ficou permitida, então, a oferta de assistência suplementar à saúde por operadoras privadas (ANS, 2018).

Com a necessidade de defender o interesse público na assistência suplementar à saúde e regular as operadoras privadas, a publicação da Lei nº9.961/2000 fundou a Agência Nacional de Saúde Suplementar (ANS). A ANS, vinculada ao Ministério da Saúde, tem como propósito contribuir para o desenvolvimento da saúde no Brasil e melhorar as relações entre as operadoras privadas e seus beneficiários (ANS, 2018).

Devido ao fato de que o atendimento via SUS é direito de todo cidadão brasileiro, a possibilidade de uso simultâneo do sistema de saúde público e privado culminou na necessidade de regulamentar o ressarcimento ao SUS pelo artigo 32 da Lei 9.656/1998 (COSTA, 2018). O ressarcimento é devido quando ocorrem atendimentos a beneficiários de operadoras privadas realizados com recursos do SUS. Desta forma, busca-se equilibrar os gastos públicos e evitar que as operadoras ganhem com as mensalidades pagas pelos beneficiários sem terem gastos com os atendimentos, ou seja, quando o atendimento é

suportado pela operadora privada e o beneficiário o realiza com recursos do SUS (ANS, 2021a).

O Tribunal de Contas da União (TCU) decretou obrigação da ANS promover a notificação e cobrança regular dos custos de ressarcimento (ANS, 2019). Quando os beneficiários utilizam o sistema de saúde público, o ato é registrado pelo Departamento de Informática do SUS (DATASUS) e a ANS realiza o cruzamento desses dados com o Sistema de Informações de Beneficiários (SIB) (ANS, 2018). A partir da integração das bases de dados, estrutura-se o conhecimento a respeito do uso do SUS por beneficiários de operadoras privadas. O procedimento administrativo tem início após a notificação da ocorrência às operadoras por meio do Aviso de beneficiário identificado (ABI), e ocorre cerca de doze meses após o evento, tempo médio necessário para o faturamento do atendimento no SUS, envio e processamento da base de dados da ANS (ANS, 2021a). As operadoras privadas, por sua vez, ao receberem a notificação de ressarcimento pagam os custos gerados por seus beneficiários ou impugnam a constatação, dando início a um processo administrativo (MACHADO, 2019). Os recursos oriundos dos processos de ressarcimento ao SUS são destinados ao Fundo Nacional de Saúde (FNS), instituído pelo Decreto 64.867/1969, o órgão gestor que tem como finalidade financiar despesas de programas prioritários do Ministério da Saúde (ANS, 2021a).

Apesar de o processo de ressarcimento ter sido regulamentado há quase duas décadas, ainda existem obstáculos na devolução dos recursos, considerando que a carência do ressarcimento pelas operadoras privadas resultou em uma significativa dívida para o SUS. Diante desse contexto, este Trabalho de Conclusão de Curso tem como foco a análise do aprofundamento da dívida do ressarcimento ao SUS pelas operadoras privadas. O trabalho visa a desenvolver um *Dashboard* informativo que permita descrever a origem e desdobramento dos atendimentos realizados no período de janeiro de 2016 a setembro de 2021.

1.1 JUSTIFICATIVA

A causa que motivou a escolha do tema foi buscar entender a expressiva dívida das operadoras privadas com o Sistema Único de Saúde. Assim, por meio do estudo de dados referentes ao ressarcimento ao SUS pelas operadoras privadas, busca-se investigar características dos atendimentos realizados no período de janeiro de 2016 a setembro de 2021.

1.2 OBJETIVO

1.2.1 Objetivo Geral

- a) Desenvolver um *Dashboard* informativo, com base em dados públicos fornecidos pela Agência Nacional de Saúde Suplementar (ANS) que permita ao usuário analisar o contexto atual da dívida oriunda da carência do ressarcimento realizado ao SUS pelas operadoras privadas;
- b) Verificar a origem e o desdobramento dos custos identificados para ressarcimento no período de janeiro de 2016 até setembro de 2021.

1.2.2 Objetivos Secundários

- a) Identificar as operadoras privadas que utilizam os recursos do SUS e quantificar a efetividade de pagamento das dívidas referentes ao ressarcimento;
- b) Analisar as características dos beneficiários que geram despesas de ressarcimento ao SUS para as operadoras privadas;
- c) Observar o tipo de estabelecimento e a região geográfica onde ocorrem os atendimentos identificados;
- d) Investigar os atendimentos que geram custos de ressarcimento;
- e) Elaborar *Dashboard* em Power Bi que permita visualizar as informações coletadas;
- f) Demonstrar utilização do *Dashboard* por meio de desenvolvimento de caso de uso.

2 REFERENCIAL TEÓRICO

2.1 RESSARCIMENTO AO SUS

O ressarcimento ao SUS tem como fundamento constitucional a Lei nº 9.656, de 1998 (BRASIL, 1988), a qual estabelece no artigo 32 que as operadoras de planos de saúde devem reembolsar os atendimentos de saúde prestados a seus consumidores e a dependentes em instituições públicas ou privadas, conveniadas ou contratadas, que sejam integrantes do Sistema Único de Saúde (SUS). Antes de entender como ocorre o ressarcimento ao SUS, é importante compreender os tipos de atendimentos que são elegíveis ao ressarcimento:

- i. Autorização de Internação Hospitalar (AIH) é o documento que registra uma internação hospitalar realizada no SUS;
- ii. Autorização de Procedimentos Ambulatoriais (APAC) é o documento utilizado pelo SUS para autorizar procedimentos ambulatoriais de alta e média complexidade.

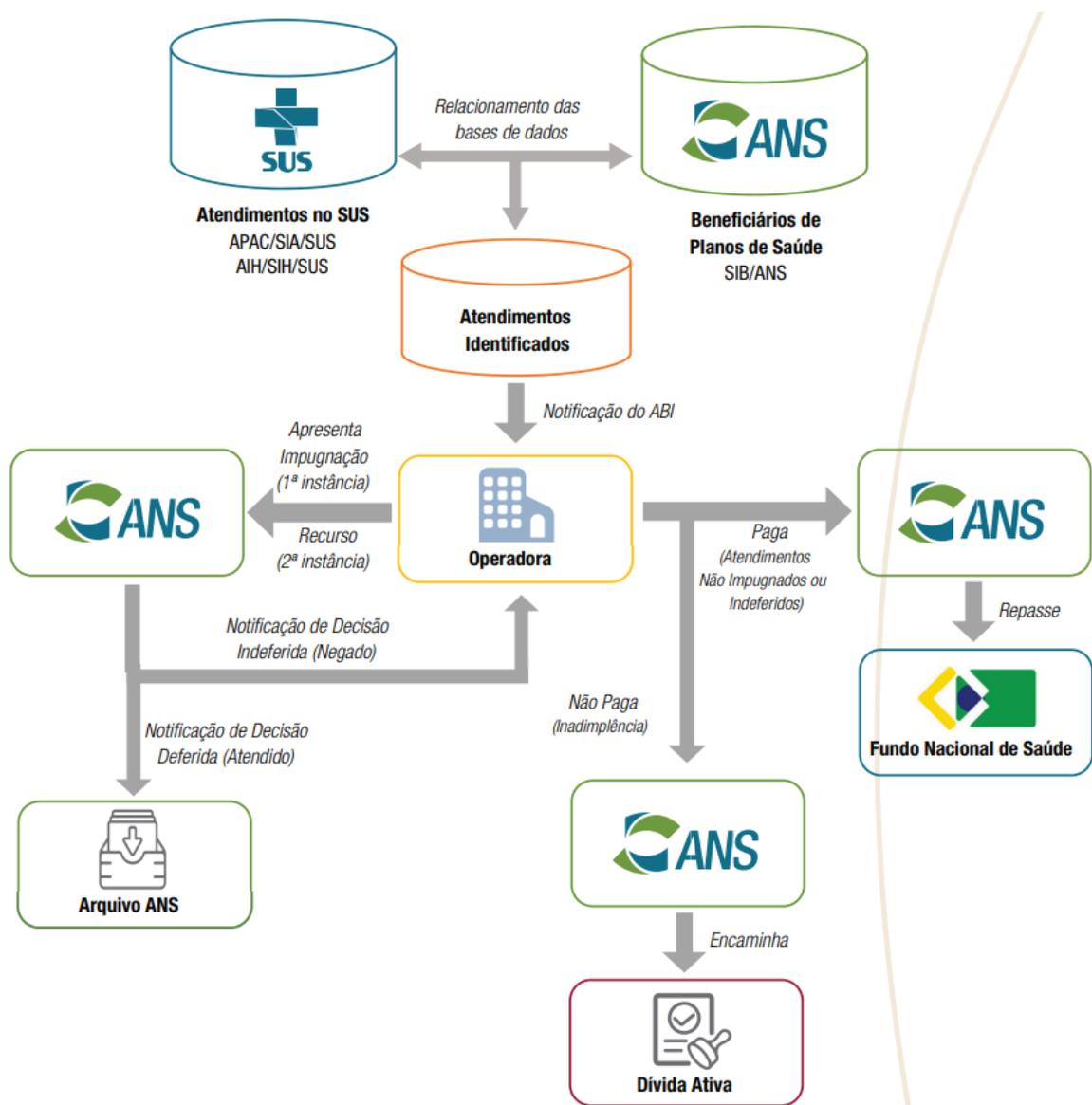
Segundo a ANS (2019), o início do processo de ressarcimento se dá pela identificação dos atendimentos que devem ser ressarcidos. Para isto, o Departamento de Informática do SUS fornece à ANS, regularmente, as bases de dados de atendimentos realizados e pagos pelo SUS. A ANS então cruza essas informações com o Sistema de Informações de Beneficiários, base de dados alimentada pelas operadoras privadas a respeito de seus beneficiários, para identificar os atendimentos realizados por beneficiários de operadoras privadas com recursos do SUS.

O processo administrativo é iniciado pela cobrança dos atendimentos identificados através do envio do ofício Aviso de Beneficiário Identificado (ABI), que agrupa trimestralmente os atendimentos com possível elegibilidade para ressarcimentos. As operadoras possuem trinta dias para impugnar esses atendimentos, caso entendam necessário. Se houver impugnações, elas são analisadas pela Diretoria de Desenvolvimento Setorial (DIDES - ANS) e uma decisão é tomada. A operadora pode, então, interpor recursos contra essa decisão, os quais são avaliados pela Diretoria Colegiada da ANS. Se a impugnação ou recurso for deferido, os valores relacionados às autorizações e atestados são cancelados e arquivados. Caso contrário, eles são aptos para cobrança por meio do documento Guia de Recolhimento da União (GRU). Os valores recolhidos pelo ressarcimento ao SUS são destinados ao Fundo Nacional de Saúde (FNS) e, caso não haja pagamento, são inscritos em

dívida ativa da ANS. A Figura 1 ilustra o fluxograma do processo administrativo de cobrança do ressarcimento ao SUS.

Os dados coletados durante o processo de ressarcimento ao SUS são divulgados de forma aberta no portal de dados do governo (dados.gov.br), por meio de bancos de dados consolidados e isentos de informações pessoais a respeito dos beneficiários. Além disso, o SUS elabora o Boletim de Ressarcimento ao SUS (ANS, 2021a) que é documento informativo a respeito dos dados consolidados de utilização dos recursos do SUS por beneficiários de operadoras privadas de plano de saúde. Esse boletim apresenta, geralmente, dados referentes aos cinco anos antecedentes à sua divulgação.

Figura 1 - Fluxograma do processo administrativo de cobrança do ressarcimento ao SUS



A etapa de impugnação das operadoras pode ocasionar alteração do valor cobrado a essas, seja por motivos de juros, descontos, cancelamentos, dentre outros. As bases de dados utilizadas neste trabalho não conseguem identificar com precisão o valor total pago pelas operadoras. Por exemplo, é impossível determinar o valor total pago em atendimentos que foram pagos de forma parcelada, porque não se pode diferenciar as parcelas já quitadas da dívida referente às parcelas futuras. Um atendimento cobrado de forma parcelada com valor de R\$500, e com R\$100 já quitados, é informado somente pela identificação de um atendimento de R\$500 parcelado.

Dessa forma, este estudo define como valor dos atendimentos aqueles identificados antes do período de impugnação. Tal valor é classificado quanto à situação de cobrança de sua GRU: paga, parcelada, pendente, suspensa ou não identificada para casos onde ainda não houve cobrança por parte da ANS às operadas privadas. Nesse sentido, é importante destacar que o valor informado nestas categorias se refere ao valor original da GRU e não ao efetivamente pago pelas operadoras após a impugnação.

Com o objetivo de mensurar a adimplência das operadoras, a RN (Resolução normativa) N° 521, de 29 de abril de 2022 (MINISTÉRIO DA SAÚDE, 2022), define o Índice de Efetivo Pagamento (IEP) como o indicador de adimplência administrativa do ressarcimento ao SUS. O IEP é calculado por meio da soma dos valores pagos e parcelados dos atendimentos dividida pelo total cobrado com GRU emitida, conforme Equação (1). Não são considerados débitos suspensos judicialmente, uma vez que são considerados pendências administrativas.

$$IEP = \frac{\text{Valores Pagos} + \text{Valores em Parcelamento}}{\text{Valores Cobrados com GRU emitida}} \quad (1)$$

As operadoras com valores de IEP próximos a 1 (um) indicam maior prontidão para o pagamento de dívidas, enquanto que, os próximos de 0 (zero) sugerem menor compromisso com seus débitos.

2.2 MANIPULAÇÃO DE DADOS

Segundo Santos *et al.* (2021) a evolução das tecnologias de comunicação, a criação da internet móvel, a maior oferta de acesso à internet para a população e o aumento da utilização da rede mundial de computadores foram fatores provedores do aumento exorbitante da

geração de dados nas últimas décadas. No entanto, para a utilização efetiva deste grande volume de dados, é necessário que ocorram processos que transformem esses dados brutos em informações.

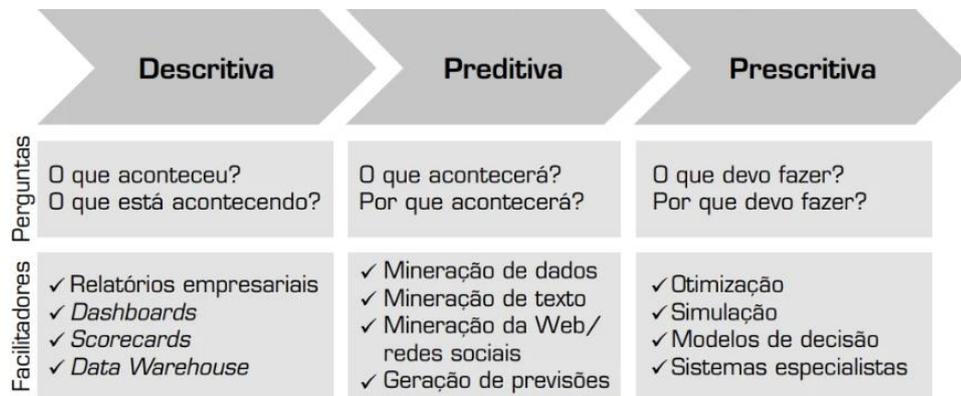
Conforme Goldschmidt (2015), existem etapas que devem ser seguidas para o manuseio e processamento de grandes volumes de dados. O autor explica que a primeira etapa é constituída pela coleta de dados, seguida pelo pré-processamento, que envolve seleção, tratamento, enriquecimento e transformação dos dados de forma a assegurar a qualidade (completude, veracidade e integridade) dos fatos representados. A próxima etapa consiste em armazenar os dados em local acessível para a etapas de análise e visualização de dados.

Os dados podem ser divididos em numéricos e categóricos: os numéricos representam valores quantitativos ou medidas, e podem ser discretos ou contínuos, ao passo que os dados categóricos representam valores qualitativos ou de atributos, e são classificados como nominais ou ordinais. Nesse contexto, o termo *DataType* indica qual é a classificação de determinada variável com base nos parâmetros descritos (MCCABE; CRAIG; MOORE, 2016).

Como esclarecem Sharda, Delen e Turban (2019), existem três níveis de análise de dados independentes, embora um tipo de aplicação leve ao outro, quais sejam: descritivas, preditivas ou prescritivas. Para a elaboração da *Dashboard* informativa a respeito do ressarcimento ao SUS foi utilizada, majoritariamente, a análise de dados descritiva, que procura responder perguntas como: “O que aconteceu?” ou “O que está acontecendo?”. A Figura 2, a seguir, ilustra os níveis de análise de dados.

Adiante, na Seção 3.4 – CASO DE USO, buscou-se demonstrar como o *Dashboard* pode ser utilizado para análises de dados preditivas com o auxílio do modelo de previsão de séries temporais de Holt-Winters.

Figura 26 - Níveis de análise de dados



Fonte: Sharda, Delen e Turban (2019), adaptada pelo autor.

De acordo com Hyndman e Athanasopoulos (2021), o modelo Holt-Winters é usado para prever valores futuros de uma variável em função do tempo, e é baseado no pressuposto de que a série temporal pode ser decomposta em três partes: tendência, sazonalidade e erro. A tendência se refere à tendência geral da série temporal, ao passo que a sazonalidade é a variação regular da série temporal (por exemplo, epidemias sazonais de gripe no período de inverno (DIZ, 2014)), e o erro é o componente aleatório da série temporal.

2.3 PYTHON

Python é uma linguagem de programação com sintaxe clara e interativa. Ela possui uma série de módulos e recursos que a configuram como uma linguagem versátil para diferentes tipos de projetos. Uma das principais vantagens do Python é sua grande comunidade e enorme quantidade de bibliotecas disponíveis, as quais compõem conjuntos de módulos que forneçam funcionalidades pré-implementadas que podem economizar tempo do programador, além de garantir confiabilidade dos códigos (PYTHON, 2021).

O Python é disponibilizado de forma gratuita para utilização. Neste trabalho, utilizou-se o produto Colaboratory, do Google Research, para escrever e executar os códigos. Apesar de a biblioteca padrão fornecida dentro do Python oferecer uma ampla gama de recursos para o programador, é possível importar novas bibliotecas para seu projeto (PYTHON, 2021). Neste trabalho, foram utilizadas três bibliotecas específicas para o manuseio de dados:

- Numpy: fornece suporte de alta performance para operações matemáticas, e é indicada para trabalhos que utilizam dados numéricos, especialmente para aplicações de ciência e análise de dados (NUMPY, 2022).
- Pandas: é uma biblioteca com recursos e ferramentas muito úteis para trabalhar com dados tabulares, como os encontrados em bases de dados (PANDAS, 2022).
- Google.colab: é um pacote de que fornece acesso a recursos específicos do Colaboratory, este oferece funções e métodos úteis para interação com o Google Drive (COLAB, s.d.).

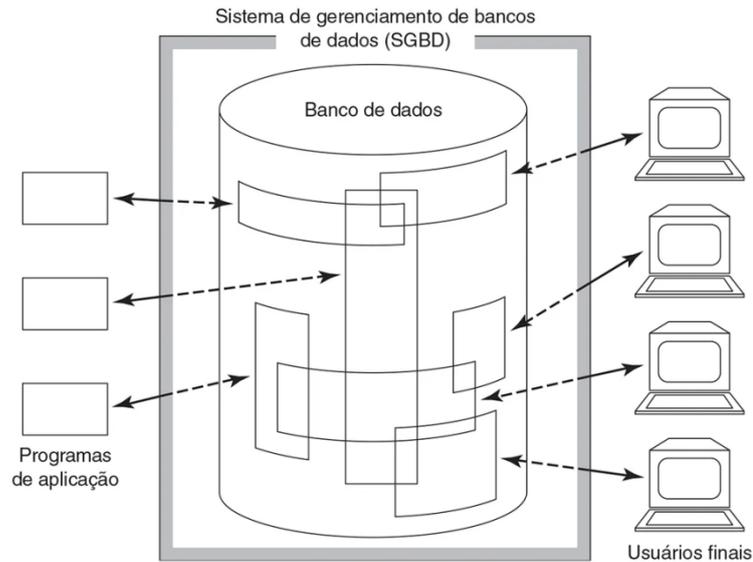
2.4 BANCO DE DADOS

De acordo com Machado (2020), um banco de dados é um conjunto de dados organizados de maneira lógica e coerente, com um significado inerente. Ele é criado com um propósito específico e é usado por um conjunto predefinido de usuários e aplicações. Além disso, um banco de dados representa algum aspecto do mundo real, chamado de "minimundo". Qualquer alteração no minimundo é automaticamente refletida no banco de dados.

É importante notar que um conjunto de dados desorganizado não pode ser considerado um banco de dados, pois ele é normalmente ordenado por *DataBase Management System* (DBMS), que é um sistema de gerenciamento de banco de dados. O conjunto de dados, o DBMS e os aplicativos relacionados são conhecidos como sistema de banco de dados, geralmente abreviado para apenas banco de dados (DATE, 2004).

Segundo Date (2004), um sistema de banco de dados, como representado na Figura 3, é uma ferramenta computacional que armazena e gerencia informações, permitindo aos usuários pesquisarem e atualizarem esses dados de acordo com suas necessidades. Essas informações podem ser qualquer tipo de dados relevantes para uma pessoa ou organização, e são conhecidas como "persistentes" porque só são removidas do banco de dados mediante uma solicitação explícita, ao invés de serem apagadas automaticamente após o uso. Em resumo, um banco de dados é uma coleção organizada de dados que são armazenados e mantidos de maneira acessível e atualizável.

Figura 37 - Sistema de Bancos de Dados



Fonte: Date (2004).

A cardinalidade é uma medida que descreve a forma como variáveis de diferentes bancos de dados se relacionam. É denominado como cardinalidade mínima o menor número de vezes que cada elemento de uma variável pode participar da relação, e, analogamente, como máxima o maior número de vezes que cada elemento de uma variável pode participar da relação. Com base nas cardinalidades mínima e máxima, e considerando n como qualquer número maior que 1 (um), o relacionamento entre variáveis de bancos de dados distintos pode ser classificado como:

- A. Um-para-um, se as cardinalidades mínima e máxima são iguais a 1;
- B. Um-para-muitos, se a cardinalidade mínima é igual a 1 e a cardinalidade máxima é igual a n ;
- C. Muitos-para-muitos, se ambas as cardinalidades são iguais a n .

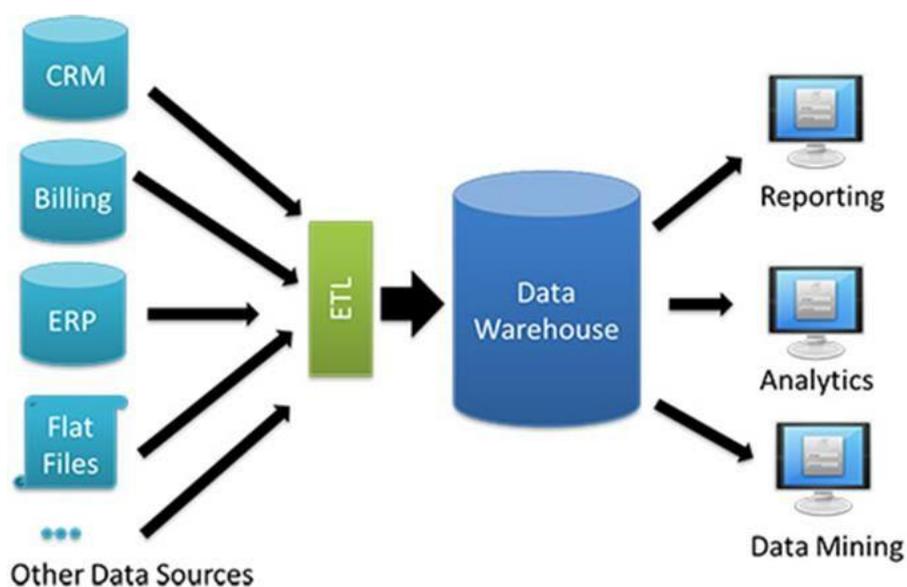
2.5 DATA WAREHOUSE

Como explica Date (2004), um *data warehouse* é um tipo de banco de dados que é usado para armazenar dados para fins de tomada de decisão, os quais foram criados para fornecer uma fonte única e limpa de dados para apoiar as decisões da gerência, sem afetar os sistemas operacionais. Nele, os dados armazenados já foram tratados e não há redundância de

informações. Após as etapas de limpeza e armazenamento, os dados sofrem apenas operações de consulta e exclusão, sem que possam ser alterados (VIDA *et al.*, 2021).

Existem quatro termos que descrevem as características de um *data warehouse*, os quais são orientados por assuntos, integrados, variáveis no tempo e não voláteis (VIDA *et al.*, 2021). Nesse sentido, ao ser orientado por assuntos, significa que o *data warehouse* é projetado para armazenar dados sobre um determinado tópico ou tema, enquanto que integrado define que os dados armazenados no *data warehouse* vêm de várias fontes diferentes e são combinados de uma forma que os torna consistentes e fáceis de usar (SHARDA; DELEN; TURBAN, 2019). Variável no tempo significa que os dados armazenados no *data warehouse* podem variar ao longo do tempo e não volátil significa que, uma vez que os dados são inseridos no *data warehouse*, eles não podem ser alterados, embora possam ser excluídos. A Figura 4 ilustra, adiante, o funcionamento de um *data warehouse*.

Figura 4 - Representação de *Data Warehouse*



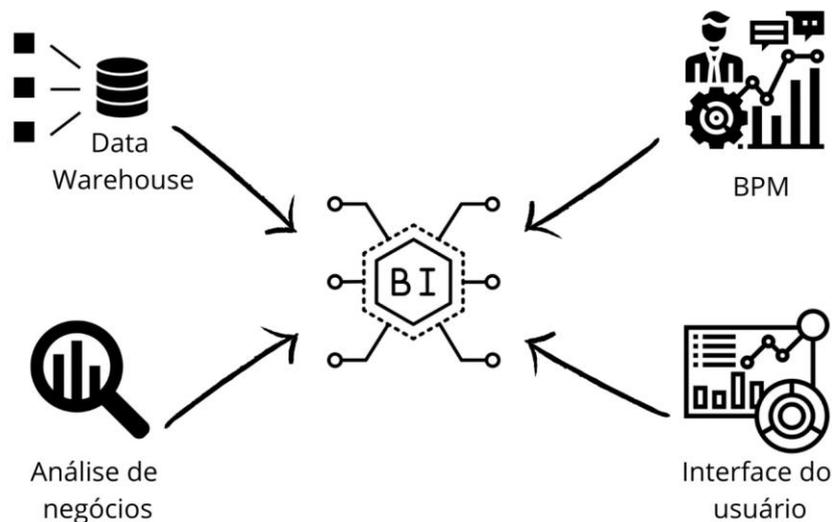
Fonte: Fátima (2019).

2.6 BUSINESS INTELLIGENCE

A inteligência de negócios, ou *Business Intelligence* (BI), combina arquiteturas, ferramentas, bases de dados, ferramentas analíticas, aplicativos e metodologias com o objetivo de auxiliar gestores e analistas na tomada de decisões. O processo de BI baseia-se em transformar dados em informações, para que estas embasem decisões e conseqüentemente ações (SHARDA; DELEN; TURBAN, 2019).

A arquitetura de um sistema de BI é composta, principalmente, por quatro componentes: *data warehouse*, análise de negócios, gestão de desempenho de negócios (BPM) e interface do usuário representada, muitas vezes, por um *Dashboard*. Os *Dashboards* são painéis que reúnem, em uma única tela, exibições visuais de importantes informações consolidadas e organizadas de forma agrupada e resumida. O ambiente de análise de negócios engloba ferramentas de manipulação, mineração e análise de dados e o termo BPM refere-se ao conjunto de processos, metodologias, parâmetros e tecnologias de negócios utilizadas para mensurar, monitorar e gerir desempenho de negócios (SHARDA; DELEN; TURBAN, 2019). A Figura 5, a seguir, ilustra a arquitetura de um sistema de BI

Figura 5 - Arquitetura de sistema de Business Intelligence



Fonte: Elaborado pelo autor.

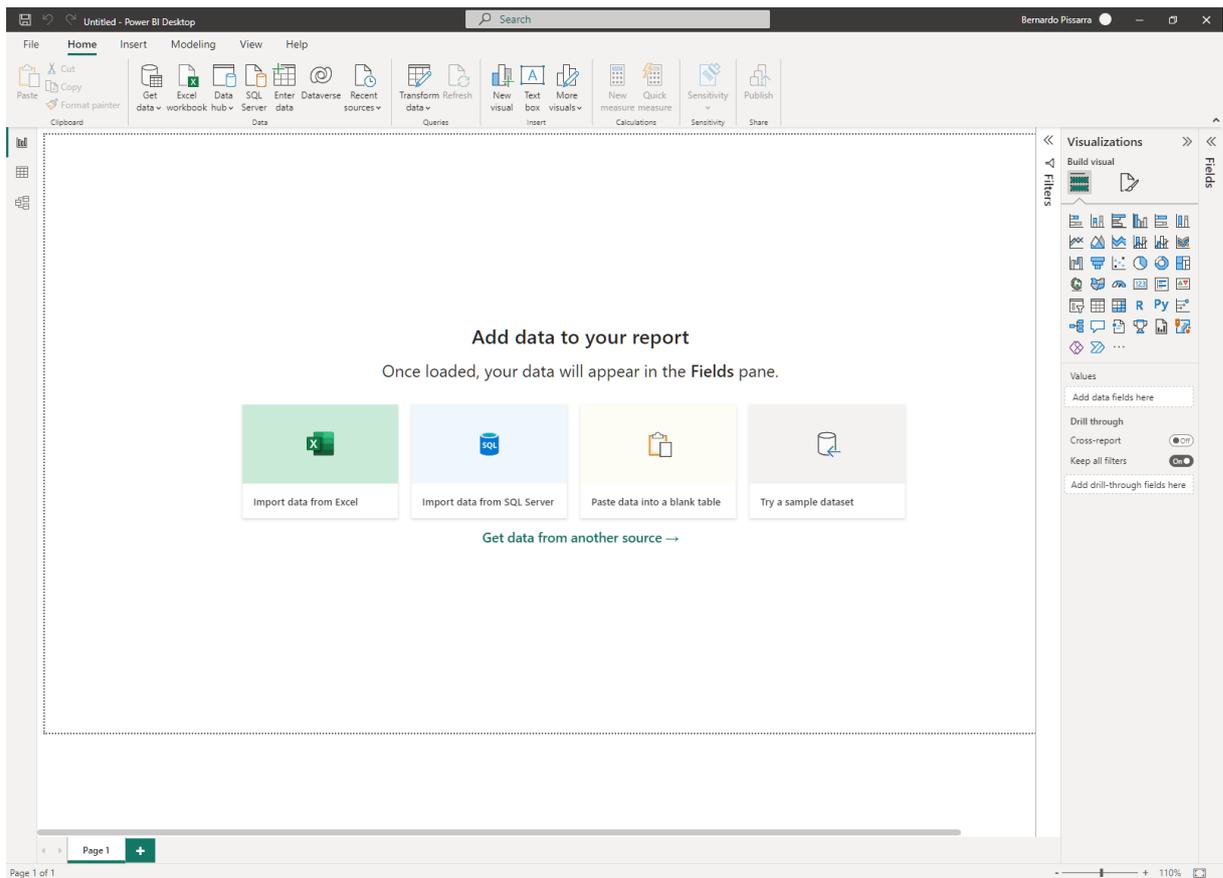
2.7 VISUALIZAÇÃO DE DADOS

As imagens têm sido utilizadas como forma de comunicação há muito tempo, sendo capazes de transmitir uma grande quantidade de informação de maneira rápida e sendo interpretáveis por pessoas de qualquer idioma. De acordo com Milani *et al.* (2020), a visualização de dados é um processo que utiliza tabelas, gráficos, mapas, diagramas, infográficos e painéis para facilitar a compreensão de dados e identificação de padrões e anomalias, assim como, auxiliar em comparações, decisões e obtenção de novos conhecimentos.

As empresas geram hoje uma quantidade enorme de dados em comparação com o que geravam anos atrás, e precisam de ferramentas para consolidar e exibir esses dados de forma mais clara. Para atender a essa demanda, surgiram várias ferramentas, como o Power BI, desenvolvido pela Microsoft (representado na Figura 6). O Power BI é uma coleção de serviços *on-line* e de recursos que permitem importar, transformar, visualizar e compartilhar dados (MILANI *et al.*, 2020).

O Power Query é um importante recurso dentro do Power BI, que permite transformar e preparar dados de várias fontes. Ele possui uma interface gráfica para acessar os dados e um editor para aplicar transformações. Utilizando a linguagem de programação M (Power Query Formula Language) é possível realizar o processo ETL (extrair, transformar e carregar dados) dentro do Power BI (MICROSOFT, 2022). Para a criação de cálculos utilizados em análises de dados é preferível utilizar a linguagem DAX (Data Analysis Expressions), também disponível na ferramenta.

Figura 6 - Tela inicial do Power Bi



Fonte: Elaborada pelo autor.

3 MATERIAIS E MÉTODOS

Este capítulo tem como objetivo descrever as técnicas e ferramentas utilizadas para o desenvolvimento do *dashboard* informativo. A primeira seção detalha a etapa de seleção e coleta das bases de dados utilizadas, seguida pela descrição das variáveis encontradas nestas. A segunda parte do capítulo representa a utilização de técnicas de programação em linguagem Python para manipular os dados de forma a prepará-los para o ambiente de *Business Intelligence*.

A terceira seção é constituída pelos processos de tratamento, transformação e enriquecimento de dados dentro da ferramenta Power BI, também é mostrado como as bases de dados foram relacionadas. Em seguida, é mostrado o desenvolvimento das visualizações e painéis que compõe o *dashboard* (melhor detalhado no APÊNDICE C).

A última seção do capítulo demonstra um caso de uso do *dashboard* para análises aprofundadas ao tema. Nesta ocasião, é analisado um cenário real percebido a partir do *dashboard* informativo, tendo como objetivo exemplificar a relevância de seu desenvolvimento.

3.1 COLETA DE DADOS

Primeiramente, foram selecionados dados referentes a cinco anos completos para as bases de dados utilizadas, de 2016 a 2020, acrescidos de três trimestres do ano de 2021, que se encontravam disponíveis para análise na data de extração. A escolha deste período se deu com base nos boletins informativos emitidos pelo SUS, que, ao serem publicados, consideram o período histórico dos últimos cinco anos. A utilização dos três trimestres divulgados permitiu ao trabalho manter as características de representatividade de cinco anos completos com acréscimo de atualização do último período divulgado.

3.1.1 Descrição das bases de dados:

Para o desenvolvimento do estudo, quatro bases de dados diferentes foram utilizadas, que podem ser descritas nos itens abaixo:

- **Base 1: Dados do Ressarcimento ao SUS por Beneficiário Identificado e ABI:**

Esta base de dados apresenta informações a respeito de ocorrências de utilização do SUS por beneficiários de operadoras privadas de planos de saúde (BRASIL, s.d.). Dela serão extraídas informações referentes aos beneficiários, atendimentos realizados, plano utilizado e local de ocorrência do procedimento. A base possui dimensão de 3.157.431 (três milhões cento e cinquenta e sete mil quatrocentos e trinta e uma) linhas e 26 (vinte e seis) colunas. O tipo de arquivo é CSV (*Comma-Separated Values*), com tamanho de 2,05 GB (*Gigabyte*).

- **Base 2: Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde:**

Dados relativos a pagamento, recursos e instâncias de julgamento de atendimentos realizados no SUS por beneficiários de operadoras privadas de plano de saúde (BRASIL, s.d.). Esta base de dados será utilizada de forma complementar às informações apresentadas a respeito dos atendimentos listados na Base 1. Ela possui dimensão de 3.157.431 (três milhões cento e cinquenta e sete mil quatrocentos e trinta e uma) linhas e 23 (vinte e três) colunas, seu tipo de arquivo é CSV, com tamanho de 1,89 GB.

- **Base 3: Dados de pagamento do Ressarcimento ao SUS por operadora:**

Dados referentes ao pagamento administrativo das dívidas das operadoras de planos de saúde relativas ao ressarcimento ao SUS (BRASIL, s.d.). Nesta base, é possível verificar os valores cobrados às operadoras, tal como o valor pago e parcelado por estas. A partir desta base de dados, é possível quantificar o IEP das operadoras de planos de saúde privado. A dimensão da base é de 128.320 (cento e vinte e oito mil trezentos e vinte) linhas, e 13 (treze) colunas. O tipo de arquivo é CSV, de tamanho 12,8 MB (*Megabyte*).

- **Base 4: CID – 10, Classificação Internacional de Doenças:**

A CID-10 é a décima revisão da Classificação Internacional de Doenças (DATASUS, 2008). Desenvolvido pela Organização Mundial da Saúde (OMS), ele é o padrão global para registro de informações de saúde e causas de morte. A utilização desta base tem como objetivo enriquecer as informações fornecidas pelas Bases 1 e 2, com sua utilização é possível identificar os atendimentos quanto aos capítulos e subcategorias pertencentes à CID. Dentre as seis tabelas constituintes desta base, foram utilizadas para a elaboração do estudo as tabelas:

- i. CID-10-SUBCATEGORIAS: a utilização desta tabela se mostra relevante para nomear a doença referente ao código CID indicado na Base 1. Sua dimensão é de 12.452 (doze mil

quatrocentos e cinquenta e duas) linhas e 8 (oito) colunas. Seu tipo de arquivo é CSV, com tamanho 1.322 KB (*Kilobyte*).

- ii. CID-10-CAPITULOS: com esta tabela é possível relacionar os atendimentos identificados aos capítulos da CID10 preestabelecidos. O arquivo, de tipo CSV, possui dimensão de 23 (vinte e três) linhas e 5 (cinco) colunas com tamanho de 3 KB.

A CID-10 é composta por 22 capítulos que agrupam doenças com características similares entre si. A lista é organizada em categorias, que são identificadas por uma letra e dois dígitos. As subcategorias são acrescentadas no nível seguinte da classificação, e são representadas por números de 0 a 9 (AMPLIMED, 2020).

3.1.2 Descrição das variáveis:

O Apêndice A contém a descrição de todas as variáveis disponíveis nas quatro bases de dados utilizadas. Para escopo deste trabalho, foram selecionadas trinta e uma variáveis, que estão descritas abaixo na Tabela 1:

Tabela 1 - Variáveis utilizadas

Variáveis referentes à Base 1: Dados do Ressarcimento ao SUS por Beneficiário Identificado e ABI		
Nome da Variável	Tipo de Variável	Descrição da Variável
COMPETENCIA	Data	Ano e mês do atendimento (AAA-MM).
NUM_ATENDIMENTO	Numérica, discreta	Identificador único do atendimento para os dados do ressarcimento ao SUS por operadora e Beneficiários Identificados SUS/ABI.
REGISTRO_OPERADOR A	Numérica, discreta	Número do registro da operadora na ANS.
TIPO_ATENDIMENTO	Numérica, discreta	Identificação entre AIH ou APAC.
UF_BENEFICIARIO	Catégorica, nominal	Sigla UF pertencente ao beneficiário do plano de saúde.
SEXO_BENEFICIARIO	Catégorica, nominal	Sexo do beneficiário.
FAIXA_ETARIA	Numérico, discreta	Faixa etária que o beneficiário pertence.
ESPECIALIDADE	Catégorica, nominal	Especificação da especialidade médica utilizada no atendimento AIH, caso existir.
COD_CID	Catégorica, nominal	Código CID (Classificação Internacional de Doenças e Problemas Relacionados à Saúde)
DIAS_PERM_INTERN	Numérica, discreta	Dias de permanência na internação para AIH; Intervalo entre datas de início e fim do tratamento para APAC.
UF_ESTAB	Catégorica, nominal	Sigla UF pertencente ao estabelecimento em que ocorreu o atendimento.

MUNICIPIO_ESTAB	Catagórica, nominal	Município do estabelecimento em que ocorreu o atendimento.
ESFERA_ESTAB	Numérica, discreta	Esfera Administrativa do estabelecimento em que ocorreu o atendimento (Federal, Municipal, Privada etc.)
NATUREZA_ESTAB	Catagórica, nominal	Natureza jurídica do estabelecimento em que ocorreu o atendimento.
VL_ATENDIMENTO	Numérica, contínua	Valor do atendimento realizado.
Variáveis referentes à Base 2: Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde		
Nome da Variável	Tipo de Variável	Descrição da Variável
VALOR_COBRADO	Numérica, contínua	Valor cobrado da operadora de plano saúde pelo SUS.
RAZAO_SOCIAL	Catagórica, nominal	Razão Social das operadoras de Plano de Saúde.
MODALIDADE_OPERADORA	Catagórica, nominal	Tipo de Operadora (Autogestão, Cooperativa, Filantropia etc.).
GRU_EMITIDA	Catagórica, nominal	Indica se a GRU de cobrança foi emitida, ou seja, se o procedimento foi cobrado pela ANS.
SITUACAO_GRU	Catagórica, nominal	Situação da GRU de cobrança (Pago, Parcelado, Pendente ou Suspenso Judicialmente).
SITUACAO_OPERADORA	Catagórica, nominal	Situação da operadora (Ativa ou Inativa).
Variáveis referentes à Base 3: Dados de pagamento do Ressarcimento ao SUS por operadora		
Nome da Variável	Tipo de Variável	Descrição da Variável
VALOR_COBRADO_ACUMULADO	Numérica, contínua	Valor total cobrado pela ANS à operadora a título de Ressarcimento ao SUS..
VALOR_PAGO_ACUMULADO	Numérica, contínua	Valor total pago pela operadora a título de Ressarcimento ao SUS.
VALOR_PARCELADO_ACUMULADO	Numérica, contínua	Valor total parcelado pela operadora referente a cobranças a título de Ressarcimento ao SUS.
Variáveis referentes à Base 4: CID – 10, Classificação Internacional de Doenças - CID-10-SUBCATEGORIAS		
Nome da Variável	Tipo de Variável	Descrição da Variável
DESCRICAÇÃO	Catagórica, nominal	Descrição (nome) da subcategoria da CID em que o procedimento realizado pertence.
SUBCAT	Catagórica, nominal	Código da subcategoria da CID-10
Variáveis referentes à Base 4: CID – 10, Classificação Internacional de Doenças - CID-10-CAPITULOS		
Nome da Variável	Tipo de Variável	Descrição da Variável
DESCRICAÇÃO	Catagórica, nominal	Descrição (nome) do capítulo da CID em que o atendimento realizado pertence.
CATINIC	Catagórica, nominal	Código da primeira categoria pertencente ao capítulo
CATFIM	Catagórica, nominal	Código da última categoria pertencente ao capítulo
DESCABREV	Catagórica, nominal	Descrição (nome) abreviada do capítulo da CID em que o atendimento realizado pertence
NUMCAP	Numérica, discreta	Número do capítulo da CID-10

Fonte: Elaborada pelo autor.

As variáveis *COMPETENCIA* e *NUM_ATENDIMENTO* serão utilizadas para identificar o atendimento, bem como para apontar quando este ocorreu. As variáveis *REGISTRO_OPERADORA* e *RAZAO_SOCIAL* serão usadas para identificar as operadoras privadas dos beneficiários que utilizaram os recursos do SUS, enquanto que *MODALIDADE_OPERADORA* e *SITUACAO_OPERADORA* irão descrever a modalidade da operadora e indicar se a mesma está ativa. A variável *UF_BENEFICIARIO* indicará a qual estado brasileiro o beneficiário pertence, ao passo que *SEXO_BENEFICIARIO*, *FAIXA_ETARIA* e *DIAS_PERM_INTERN* serão utilizados para determinar o sexo do beneficiário, a faixa de idade em que este se encontra e quantos dias permaneceu internado (caso de atendimento AIH) ou quantos dias durou o tratamento (caso de atendimento APAC), respectivamente.

Quanto aos atendimentos, a variável *TIPO_ATENDIMENTO* será utilizada com o objetivo de determinar se o atendimento realizado é referente a uma AIH ou APAC, e a variável *COD_CID* determina o código CID do atendimento com base na Classificação Internacional de Doenças (CID-10). As variáveis *DESCRICA0* (Base 4.i), *DESCRICA0* (Base 4.ii), *DESCABREV*, *SUBCAT*, *CATFIM*, *CATINIC* e *NUMCAP* serão utilizadas para descrever o atendimento com base nas subcategorias e capítulos descritos na CID-10. Para os atendimentos AIH, a variável *ESPECIALIDADE* descreve, caso existir, a especialidade médica utilizada no atendimento.

As variáveis *UF_ESTAB* e *MUNICIPIO_ESTAB* serão utilizadas para determinar o estado e município brasileiro em que está localizado o estabelecimento onde o atendimento ocorreu. Por sua vez, as variáveis *ESFERA_ESTAB* e *NATUREZA_ESTAB* serão usadas para determinar a esfera e a natureza deste estabelecimento, respectivamente.

A variável *VL_ATENDIMENTO* indica o valor identificado do atendimento, antes da etapa de Impugnação das operadoras privadas, e as variáveis *GRU_EMITIDA* e *SITUACAO_GRU* são usadas para determinar se a GRU de cobrança do atendimento foi emitida e a situação desta cobrança. Por seu turno, a variável *VALOR_COBRADO* relata o valor determinado para cobrança do atendimento após a fase de Impugnação e, por fim, as variáveis *VALOR_COBRADO_ACUMULADO*, *VALOR_PAGO_ACUMULADO* e *VALOR_PARCELADO_ACUMULADO* serão utilizadas para o cálculo do Índice de Efetividade de Pagamento (IEP) das operadoras privadas.

3.2 PREPARAÇÃO DE DADOS EM PYTHON

Antes de iniciar o manejo dos dados dentro do ambiente de *Business Intelligence* – Power BI, as bases 1, 2 e 3 passaram pela etapa de preparação de dados. Para isto, utilizou-se o ambiente Colaboratory fornecido pela Google Research. Os códigos foram escritos em linguagem de programação Python, e as manipulações realizadas em cada base podem ser descritas nos próximos itens.

3.2.1 Criação do ambiente de trabalho

Antes de trabalhar com as bases dentro do ambiente Google Colaboraty, foram necessários dois procedimentos: a importação das bibliotecas utilizadas para o desenvolvimento dos códigos e a definição da fonte de dados.

- Procedimento 1: O carregamento das bibliotecas utilizadas se deu pela execução do código ilustrado na Figura 7, por meio do qual foi importada a biblioteca Numpy, Pandas e o módulo Drive pertencente à biblioteca Google.Colab:

Figura 7 – Importação de bibliotecas

```
import numpy as np
import pandas as pd
from google.colab import drive
```

Fonte: Elaborada pelo autor.

- Procedimento 2: As bases de dados foram armazenadas na nuvem, dentro da plataforma Google Drive. Desta forma, definimos o caminho para a fonte de dados pela localização exata das bases dentro do Drive, Figura 8.

Figura 8 – Definição de fonte de dados

```
drive.mount('/content/drive', force_remount=True)
path = 'content/drive/My Drive/Basededados/'
```

Fonte: Elaborada pelo autor.

3.2.2 Preparação de dados: Base 1 Dados do Ressarcimento ao SUS por Beneficiário Identificado e ABI, e Base 2 Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde

Inicialmente, as duas bases de dados estavam dispostas separadas por ano de atendimento. Observou-se, também, que as bases de dados 1 e 2 continham informações diferentes e complementares sobre os mesmos atendimentos. Nesse sentido, foi oportuno construir uma única base de dados que reúne todas as informações relevantes para estas bases. Para isto, foram realizadas quatro operações:

- Operação 1: A primeira operação consiste na importação dos dados referentes aos seis anos, de cada uma das duas bases de dados, para o Google Collaboratory. Para isso, foi realizada a importação dos anos 2016 a 2021 para as bases 1 e 2, indicado pela Figura 9.

Figura 9 – Trecho do código utilizado para importar tabelas referentes a Base 1 e Base 2

```
base1.2016 = pd.read_csv(path+'benef_ident_SUS_ABI_2016.csv', sep =";")
base1.2017 = pd.read_csv(path+'benef_ident_SUS_ABI_2017.csv', sep =";")
base1.2018 = pd.read_csv(path+'benef_ident_SUS_ABI_2018.csv', sep =";")
base1.2019 = pd.read_csv(path+'benef_ident_SUS_ABI_2019.csv', sep =";")
base1.2020 = pd.read_csv(path+'benef_ident_SUS_ABI_2020.csv', sep =";")
base1.2021 = pd.read_csv(path+'benef_ident_SUS_ABI_2021.csv', sep =";")

base2.2016 = pd.read_csv(path+'ressarcimento_SUS_operadora_planos_2016.csv', sep =";", encoding='ISO-8859-1')
base2.2017 = pd.read_csv(path+'ressarcimento_SUS_operadora_planos_2017.csv', sep =";", encoding='ISO-8859-1')
base2.2018 = pd.read_csv(path+'ressarcimento_SUS_operadora_planos_2018.csv', sep =";", encoding='ISO-8859-1')
base2.2019 = pd.read_csv(path+'ressarcimento_SUS_operadora_planos_2019.csv', sep =";", encoding='ISO-8859-1')
base2.2020 = pd.read_csv(path+'ressarcimento_SUS_operadora_planos_2020.csv', sep =";", encoding='ISO-8859-1')
base2.2021 = pd.read_csv(path+'ressarcimento_SUS_operadora_planos_2021.csv', sep =";", encoding='ISO-8859-1')
```

Fonte: Elaborada pelo autor.

- Operação 2: A segunda parte do código, Figura 10, realiza a primeira redução de variáveis do trabalho. Uma vez que não é necessária a existência de variáveis duplicadas, o código utiliza a função `pandas.DataFrame.drop` (PANDAS, 2022b) para remover da Base 2 as colunas que existem de forma idêntica na Base 1.

Figura 10 - Trecho do código que retira as colunas idênticas entre Base 1 e 2

```
base2.2016.drop(['COMPETENCIA','ABI','TIPO_ATENDIMENTO','CD_CNES',
               'DATA_ATUALIZACAO_ARQUIVO','REGISTRO_OPERADORA'],axis=1,inplace=True)
base2.2017.drop(['COMPETENCIA','ABI','TIPO_ATENDIMENTO','CD_CNES',
               'DATA_ATUALIZACAO_ARQUIVO','REGISTRO_OPERADORA'],axis=1,inplace=True)
base2.2018.drop(['COMPETENCIA','ABI','TIPO_ATENDIMENTO','CD_CNES',
               'DATA_ATUALIZACAO_ARQUIVO','REGISTRO_OPERADORA'],axis=1,inplace=True)
base2.2019.drop(['COMPETENCIA','ABI','TIPO_ATENDIMENTO','CD_CNES',
               'DATA_ATUALIZACAO_ARQUIVO','REGISTRO_OPERADORA'],axis=1,inplace=True)
base2.2020.drop(['COMPETENCIA','ABI','TIPO_ATENDIMENTO','CD_CNES',
               'DATA_ATUALIZACAO_ARQUIVO','REGISTRO_OPERADORA'],axis=1,inplace=True)
base2.2021.drop(['COMPETENCIA','ABI','TIPO_ATENDIMENTO','CD_CNES',
               'DATA_ATUALIZACAO_ARQUIVO','REGISTRO_OPERADORA'],axis=1,inplace=True)
```

Fonte: Elaborada pelo autor.

- Operação 3: É importante destacar que o número de atendimento é único, e, por isso, a variável NUM_ATENDIMENTO é um identificador único presente nas duas bases. A função pandas.DataFrame.merge, disponível na biblioteca Pandas, permite adicionar as variáveis da Base 2 à Base 1 utilizando o número de atendimento como identificador. Desta forma, garante-se que a base final não misture informações sobre atendimentos diferentes (PANDAS, 2022a). O recurso foi aplicado ano a ano, como ilustrado na Figura 11.

Figura 11 - União ano a ano Bases 1 e 2

```
basefinal.2016 = base1.2016.merge(base2.2016, on='NUM_ATENDIMENTO')
basefinal.2017 = base1.2017.merge(base2.2017, on='NUM_ATENDIMENTO')
basefinal.2018 = base1.2018.merge(base2.2018, on='NUM_ATENDIMENTO')
basefinal.2019 = base1.2019.merge(base2.2019, on='NUM_ATENDIMENTO')
basefinal.2020 = base1.2020.merge(base2.2020, on='NUM_ATENDIMENTO')
basefinal.2021 = base1.2021.merge(base2.2021, on='NUM_ATENDIMENTO')
```

Fonte: Elaborada pelo autor.

- Operação 4: Para concatenar (juntar) as bases de dados geradas, utilizou-se a função pandas.concat, presente na biblioteca Pandas. Este recurso permite concatenar as bases de dados de 2016 a 2021 ao longo de um eixo. Assim, escolhendo o eixo x como referência, as bases foram concatenadas como linhas de uma mesma base, e não como colunas (PANDAS, 2022b). Com isso, concatenou-se as bases de dados referentes aos anos 2016 a 2021 em somente uma única base de dados final. A Figura 12 ilustra o código utilizado.

Figura 12 - Concatenação final dos seis anos de estudo

```
frames = [basefinal.2016, basefinal.2017, basefinal.2018,
          basefinal.2019, basefinal.2020, basefinal.2021]
basefinal = pd.concat(frames)
```

Fonte: Elaborada pelo autor.

3.2.3 Preparação de dados: Base 3 Dados de pagamento do Ressarcimento ao SUS por operadora:

Analogamente, a Base 3 também estava disponível separada por ano das informações. Para ela, criou-se uma única base de dados utilizando as funções `pandas.DataFrame.drop` e `pandas.concat`. O código utilizado pode ser visualizado no Apêndice B.

3.3 FERRAMENTA DE BUSINESS INTELLIGENCE – POWER BI

Após a etapa de preparação das bases, estas foram carregadas no *software* de *Business Intelligence* da Microsoft, o Power BI. Neste ambiente, foram realizadas as etapas de tratamento de dados, relacionamento entre as bases de dados, criação da *Dashboard* e desenvolvimento de caso de uso.

3.3.1 Importação de dados

O primeiro passo realizado para a utilização da ferramenta foi a importação das bases geradas na Seção 3.2, seguida pela criação de uma base adicional denominada Faixa Etária. Na Tabela 2 é possível visualizar todas as bases de dados utilizadas para a execução do estudo.

Tabela 2 - Bases de dados disponíveis no Power BI

Nome da base de dados	Fonte da base de dados
Ressarcimento SUS	Merge da Base 1 e Base 2
Dados de pagamento do Ressarcimento ao SUS por operadora	Base 3

CID-10-SUBCATEGORIAS	Base 4
CID-10-CAPITULOS	Base 4
Faixa Etária	Base de criação local

Fonte: Elaborada pelo autor.

A base Faixa Etária tem como objetivo complementar informações da faixa de idade dos beneficiários, que são provenientes da base Ressarcimento SUS. Ao utilizá-la, o identificador da faixa etária dos beneficiários pode ser explicitado como a idade (intervalo de idade no qual o beneficiário se encontra) correspondente. A Tabela foi criada a partir da entrada de dados no Power Query, no formato JSON, e seu código pode ser visualizado no Apêndice B.

3.3.2 Tratamento e transformação de dados

O tratamento e a transformação dos dados para cada tabela é descrito nos itens a seguir. Nesses, foi realizada uma série de operações importantes para a construção das análises. Algumas das operações comuns incluem a limpeza de dados, a transformação de variáveis e a criação de novas variáveis. O objetivo dessa etapa é preparar os dados para análise, tornando-os de mais fácil interpretação. Além disso, essas operações também podem ajudar a detectar e a corrigir erros e inconsistências nos dados.

3.3.2.1 Base de dados: Ressarcimento ao SUS

O código escrito em linguagem M para o tratamento dos dados da base Ressarcimento ao SUS pode ser visualizado no Apêndice B. Para esta base, a primeira operação realizada foi a de ajuste do DataType correto para cada variável. Em seguida, as variáveis ainda existentes na base e que fugiam do escopo do projeto foram removidas, enquanto que as restantes foram renomeadas de forma a tornar a leitura dos dados mais intuitiva, disponível no APÊNDICE A. A variável *FAIXA_ETARIA* foi tratada com intuito de exibir um identificador único para cada categoria dos segmentos de idades, pois antes dessa operação existiam números diferentes para a mesma faixa etária.

A variável *SEXO_BENEFICIARIO* continha valores que foram preenchidos com “F” e “M”, para sexo feminino e masculino, respectivamente. Ademais, foram substituídos os

valores desses campos de forma a tornar esta variável unicamente preenchida pelos termos “Feminino” ou “Masculino”. A variável *TIPO_ATENDIMENTO* era preenchida com valor “1” para simbolizar atendimento AIH e “2” para APAC, mas, após o tratamento, os valores foram substituídos para “AIH” e “APAC”, respectivamente. A variável *ESFERA_ESTAB* era preenchida com os identificadores “1”, “2”, “3” e “4”, contudo, para facilitar o entendimento da leitura de dados, esses foram substituídos por “Não informado”, “Federal”, “Estadual” e “Municipal”, respectivamente.

Foram realizados dois ajustes para a variável *SITUACAO_GRU*. Os campos nulos que indicavam ausência de dados foram preenchidos com o termo “Não especificado” e os valores denominados “pendente de pagamento (inclui suspensão SCI)” foram anexados à classificação “Pendente”. Com base na divisão apresentada no Panorama de Ressarcimento ao SUS (ANS, 2021b) das categorias de GRU emitidas, a segregação da subcategoria “suspensa SCI” não é relevante para o escopo do trabalho.

Criou-se a variável *País*, que permitiu determinar o país Brasil para todas as visualizações referentes aos municípios dos estabelecimentos onde ocorreram os atendimentos. Antes do ajuste, a existência de cidades com o mesmo nome ao redor do mundo causava dificuldades para o Power BI identificar o local correto do atendimento. Com o objetivo de classificar os atendimentos com CID não especificada, o código CID “#999” foi inserido nos atendimentos com código CID vazios por representar um código inexistente.

3.3.2.2 Base de dados: Dados de pagamento do Ressarcimento ao SUS por operadora

É possível consultar o código escrito em linguagem M para o tratamento dos dados da base Dados de pagamento do ressarcimento ao SUS por operadora no Apêndice B.

Com o objetivo de somar as variáveis *VALOR_PAGO_ACUMULADO* e *VALOR_PARCELADO_ACUMULADO*, criou-se a variável *Valor pago+parcelado*. Essa variável será utilizada de forma a comparar os valores pagos e parcelados pelas operadoras privadas a respeito dos atendimentos com o valor efetivamente cobrado pela ANS. Com base na Equação 1, utilizou-se a linguagem DAX para calcular o valor do IEP:

$$\text{IEP} = \text{divide}(\text{CALCULATE}(\text{SUM}('Dados de pagamento do Ressarcimento ao SUS por operadora'[Valor pago+parcelado])), \text{CALCULATE}(\text{SUM}('Dados de pagamento do Ressarcimento ao SUS por operadora'[VALOR_COBRADO_ACUMULADO])))$$

3.3.2.3 Base de dados: CID-10 SUBCATEGORIAS

O código escrito em linguagem M, referente ao tratamento dos dados da base CID-10 SUBCATEGORIAS, pode ser visualizado no Apêndice B. O primeiro passo realizado foi a remoção de variáveis que não seriam utilizadas com base no escopo do projeto. Com o objetivo de estabelecer um relacionamento entre esta base e a base CID-10 CAPÍTULOS, foram criadas três novas variáveis.

A variável *SUBCAT* armazena o código da subcategoria da CID, o qual é formado por uma letra e uma série de números. A partir desta variável, foram criadas as variáveis *Letra categoria* e *Num categoria* que fragmentam a variável *SUBCAT* e retornam a letra e o número referentes à categoria da CID. Com a utilização das duas novas variáveis, e o intervalo dos capítulos da CID definidos pelas variáveis *CATINIC* e *CATFIM* da base de dados CID-10 CAPÍTULOS, criou-se a variável *numerocap*. Esta retorna o número do capítulo em que a subcategoria da doença identificada na CID está contida. O código DAX referente à criação da variável *numerocap* pode ser visualizado no Apêndice B.

Com o objetivo de indicar os atendimentos com CID não especificada, foi adicionada uma subcategoria nesta base de dados. Para isto, adicionou-se uma linha com valores para as variáveis de *SUBCAT* = “#999” e *DESCRICAÇÃO* = “CID Não especificada”. A partir da subcategoria criada de código “#999”, a variável *numerocap* retorna a um capítulo fictício de número 99, que será descrito no item 4.3.2.4.

3.3.2.4 Base de dados: CID-10 CAPITULOS

O código escrito em linguagem M referente ao tratamento dos dados da base CID-10 CAPITULOS pode ser visualizado no Apêndice B. O último passo necessário para a indicação dos atendimentos com CID não especificada é a criação do capítulo com número 99 citada no item 4.3.2.3.

Na base de dados CID-10 CAPITULOS foi inserida uma nova linha, a qual traz os seguintes valores para as variáveis: *NUMCAP* = “99”, *CATINIC*=”#0” *CATFIM*=”#9999” *DESCRICAÇÃO*=“ CID Não Especificada” e *DESCABREV*=“CID Não Especificada”. Desta forma, a partir do relacionamento entre as bases de dados CID-10 SUBCATEGORIAS e CID-10 CAPITULOS por meio das variáveis *numerocap* e *NUMCAP*, foi possível classificar atendimentos com código CID vazios como pertencentes ao capítulo “CID Não Especificada”

e subcategoria “CID Não Especificada”. A Figura 13, indica à esquerda os atendimentos com capítulo da CID não especificados, e à direita os atendimentos com código CID = #999.

Figura 13 - Classificação de atendimentos com CID não especificada

Capítulo CID	% Atendimentos	CID	N° Atendimentos	%	Descrição
II. Neoplasias (tumores)	37,87%	#999	388562	12,31%	CID Não Especificada
CID Não Especificada	12,31%	Z940	174591	5,53%	Rim transplantado
XXI. Contatos com serviços de saúde	9,43%	C504	171730	5,44%	Neoplasia maligna do quadrante superior externo da mama
XV. Gravidez parto e puerpério	6,70%	C500	166573	5,28%	Neoplasia maligna do
VIII. Doenças do ouvido e da apófise mastóide	5,21%	Total	3157431	100,00%	
XIX. Lesões enven e alg out conseq causas externas	4,64%				
I. Algumas doenças infecciosas e parasitárias	3,23%				

Fonte: Elaborada pelo autor.

3.3.3 Criação dos relacionamentos

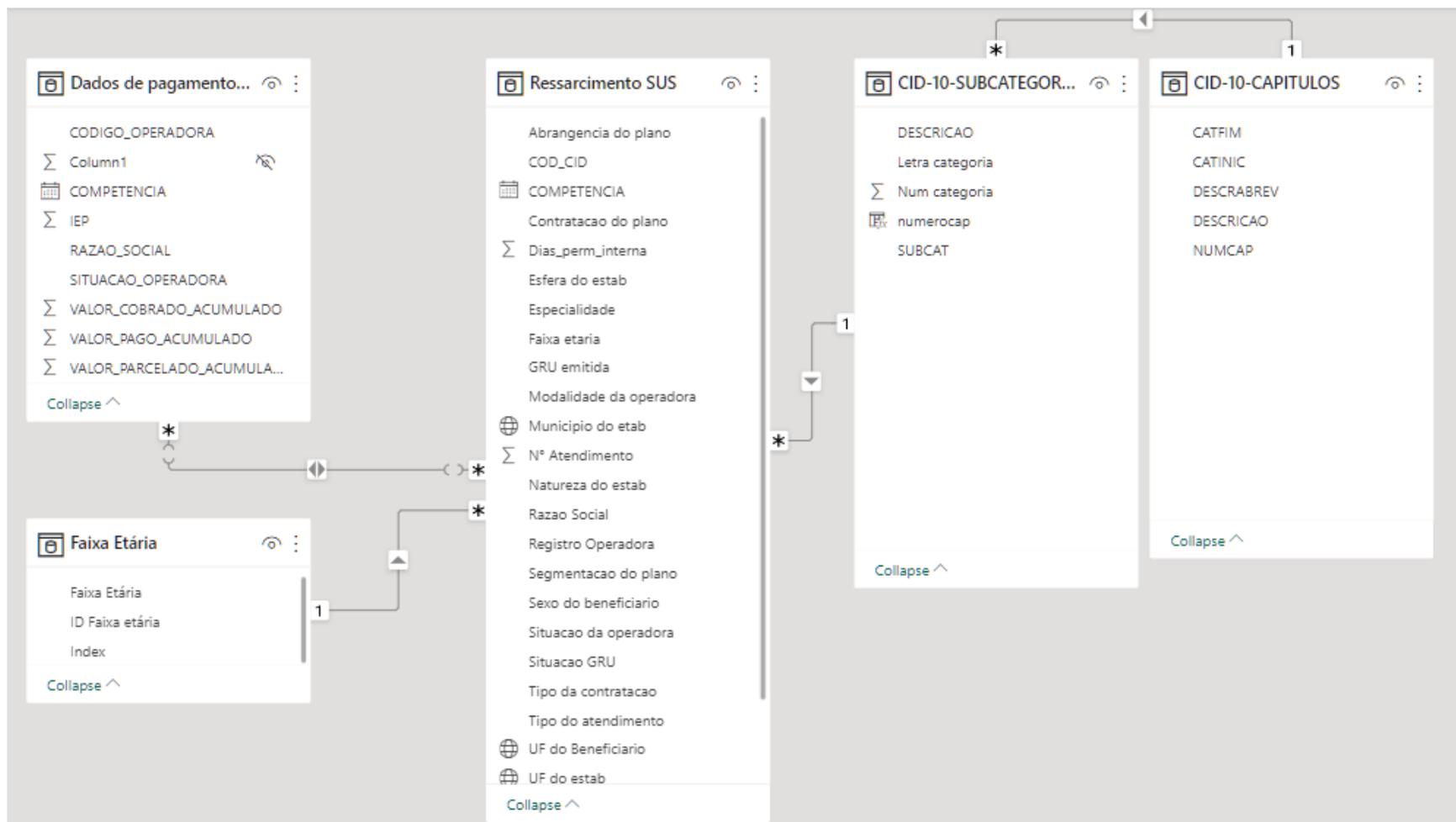
O relacionamento entre as tabelas foi realizado com base nas cardinalidades descritas na Tabela 3, a seguir, podendo ser visualizado dentro do Power BI na Figura 14, adiante.

Tabela 3 - Cardinalidade entre tabelas

Base X	Base Y	Cardinalidade
Ressarcimento ao SUS	Dados de pagamento do Ressarcimento ao SUS por operadora	Muitos para muitos
Faixa Etária	Ressarcimento ao SUS	1 para muitos
CID-10 SUBCATEGORIAS	Ressarcimento ao SUS	1 para muitos
CID-10 CAPITULOS	CID-10 SUBCATEGORIAS	1 para muitos

Fonte: Elaborada pelo autor.

Figura 14 - Modelo relacional Power BI



Fonte: Elaborada pelo autor.

3.3.4 Dashboard

Com os dados devidamente tratados e relacionados, iniciou-se a elaboração das visualizações pertinentes para o escopo do projeto. As informações coletadas originaram a criação quatro painéis dentro do Power BI: Operadoras de plano de saúde, Perfil do beneficiário, Localização dos atendimentos e Atendimentos que podem ser vistos no Apêndice C.

Para cada painel está disponível, no *menu* superior, filtros que permitem ao usuário realizar análises mais precisas com base em seus interesses. É possível visualizar, também, um *card* com valores consolidados sobre os atendimentos. A Figura 15 ilustra a disposição do *menu* superior onde, para o Painel 1, os seis filtros disponíveis estão na parte superior da figura, e na parte inferior é apresentado o *card* contendo três valores relevantes. Com o intuito de contextualizar o projeto e exibir a descrição das variáveis utilizadas, elaborou-se a capa e sumário do *dashboard* APÊNDICE C.

3.3.4.1 Painel 1 – Operadoras de plano de saúde

O primeiro painel apresenta o conhecimento adquirido a respeito das operadoras que utilizam os recursos do SUS. O *menu* superior deste é constituído pelos filtros: Data do atendimento, Código da operadora, Razão Social, Modalidade da operadora, Situação da operadora e Procedimento cobrado (difere atendimentos com GRU emitida ou não). Do mesmo modo ocorre com o *card* de valores composto por: Número de atendimentos, Valor dos atendimentos e Valor cobrado com GRU, conforme a Figura 15.

Figura 15 - Menu superior painel 1

Data do atendimento	Registro Operadora	Razao Social	Modalidade da operadora	Situacao da operadora	Procedimento cobrado
01/01/2016 01/09/2021	All	All	All	All	All
		3157431 Número de atendimentos	R\$ 5.245.516.788 Valor dos atendimentos	R\$ 3.467.239.190 Valor cobrado com GRU	

Fonte: Elaborada pelo autor.

As visualizações que compõem a Figura 16 descrevem a situação da cobrança dos atendimentos por modalidade de operadora. A linha em vermelho no gráfico superior representa o percentual relativo ao montante total do valor dos atendimentos para cada modalidade. Com o objetivo de melhor visualizar a diferença entre as modalidades, o gráfico

inferior traz o visual 100% escalonado dos dados. Por exemplo, a Medicina de Grupo é responsável por 49,82% dos atendimentos sendo que R\$672.697.991 (25,74%) em dívidas não estão especificadas, R\$532.459.285 (20,38%) foram pagas, R\$494.473.616 (18,92%) foram parcelados, R\$546.885.632 (20,93%) estão pendentes e R\$366.776.104 (14,03%) estão suspensos judicialmente.

Figura 16 - Situação de cobrança por modalidade de operadora

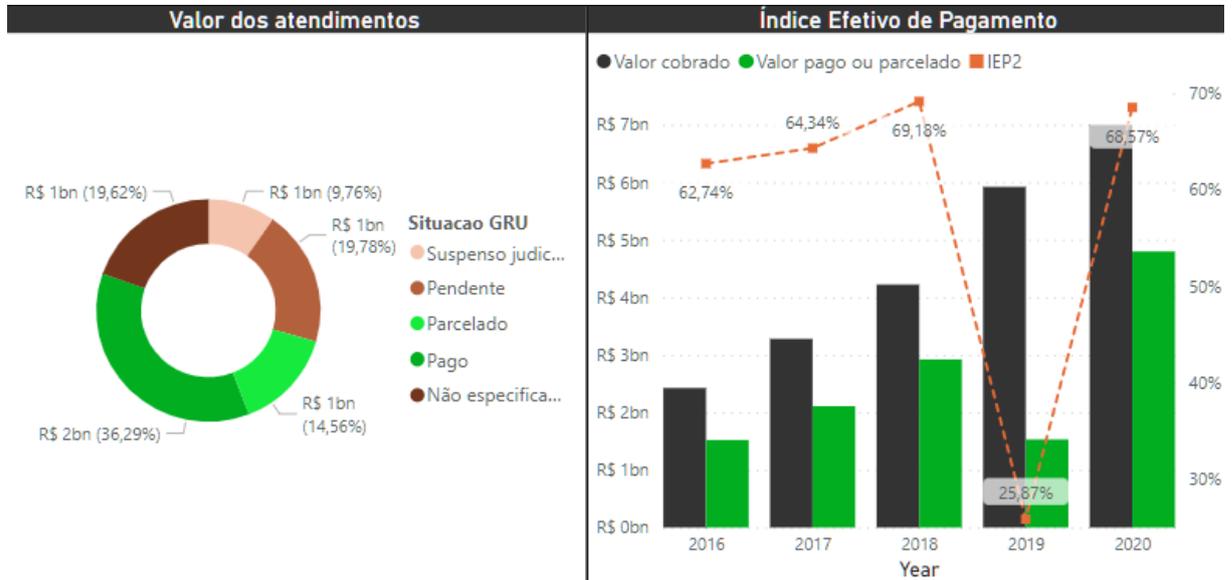


Fonte: Elaborada pelo autor.

A Figura 17 ilustra, à direita, o índice efetivo de pagamento acumulado das operadoras de plano de saúde. A coluna verde-claro representa o valor total pago e parcelado e a coluna cinza representa o valor total cobrado. A linha pontilhada mostra o IEP durante o período em questão. A visualização à esquerda mostra a distribuição do valor dos atendimentos de acordo com a situação de cobrança. Na Figura 17 é possível ver, por exemplo, que 14,56% dos

atendimentos foram parcelados, e que no ano de 2016, R\$2.425.637.264,29 foram cobrados e R\$1.521.730.496,57 foram pagos ou parcelados, resultando no IEP de 62,74%.

Figura 17 - Valor dos atendimentos e IEP



Fonte: Elaborada pelo autor.

3.3.4.2 Painel 2 – Perfil do beneficiário de plano de saúde

O objetivo deste painel é auxiliar na compreensão das variáveis relacionadas às características dos beneficiários de planos de saúde que utilizam o SUS. Com isto, é possível entender de melhor como os custos são gerados sob a perspectiva do usuário. Para este painel, o *menu* superior é composto pelos filtros: Data do atendimento, Código da operadora, Razão Social, Faixa etária, Sexo do beneficiário, Tipo do atendimento e Procedimento cobrado. O *card* de valores informa, adicionalmente, a Média de dias de permanência interna (DPS) Figura 18.

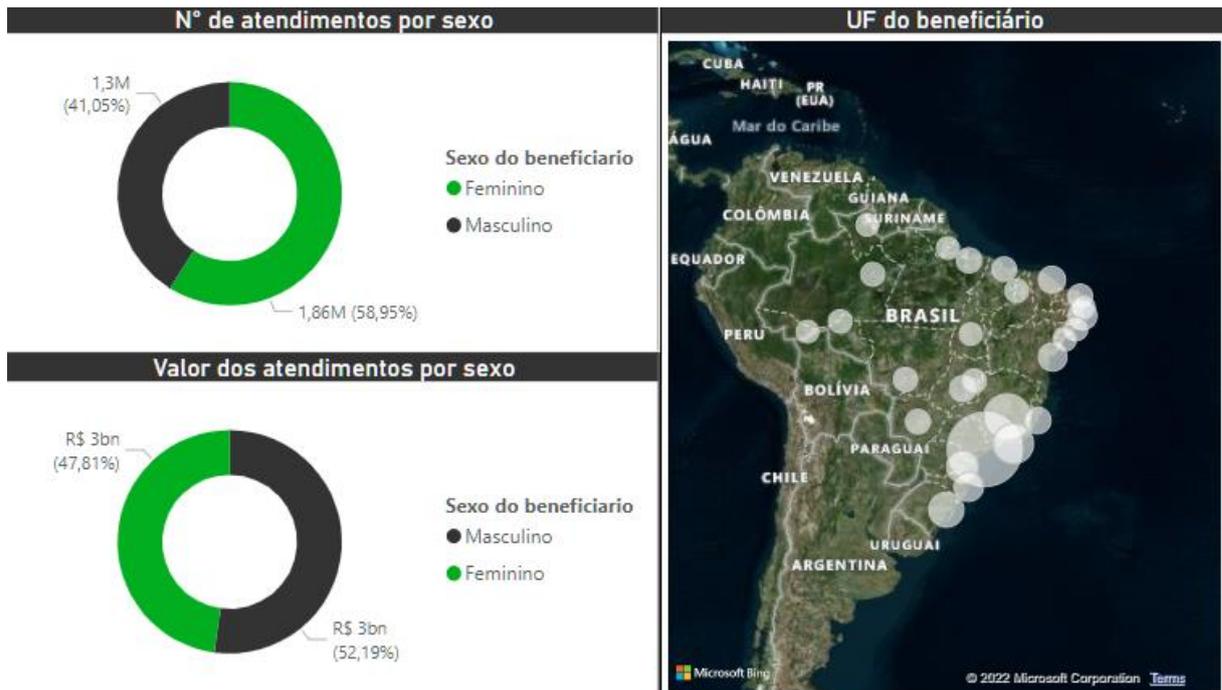
Figura 18- Menu superior painel 2



Fonte: Elaborada pelo autor.

As três visualizações que compõem a Figura 19 descrevem o número de atendimentos por sexo, valor dos atendimentos por sexo e frequência por UF dos beneficiários que utilizam o SUS. É possível perceber, por exemplo, que o sexo masculino é responsável por 1,3 milhões de atendimentos e R\$3 bilhões em dívidas. O estado com maior número de atendimentos identificados é São Paulo.

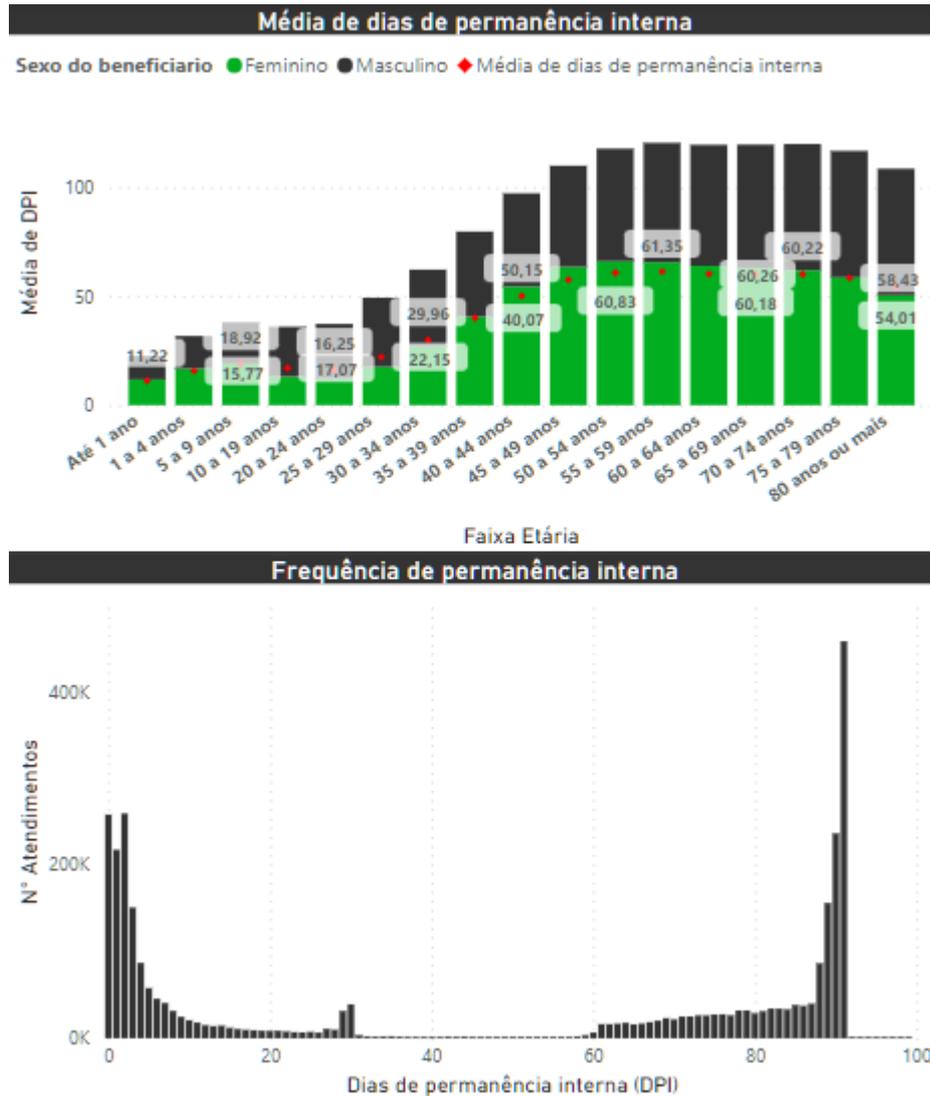
Figura 19 - Atendimentos por sexo e UF do beneficiário



Fonte: Elaborada pelo autor.

Para o estudo da variável Dias de permanência interna (DPI), foram elaboradas duas visualizações, conforme a Figura 20. A primeira demonstra a média de DPI por Faixa etária, segmentada também pelo sexo do beneficiário. Com o intuito de analisar a frequência de permanência interna, a segunda visualização plota o N° de atendimentos que ocorrem para cada quantidade de dias de permanência. É perceptível que a faixa etária com maior média de dias de permanência interna (61,35 dias) é a de 55 a 59 anos, e que as pessoas que permanecem de 31 a 59 dias são menos frequentes

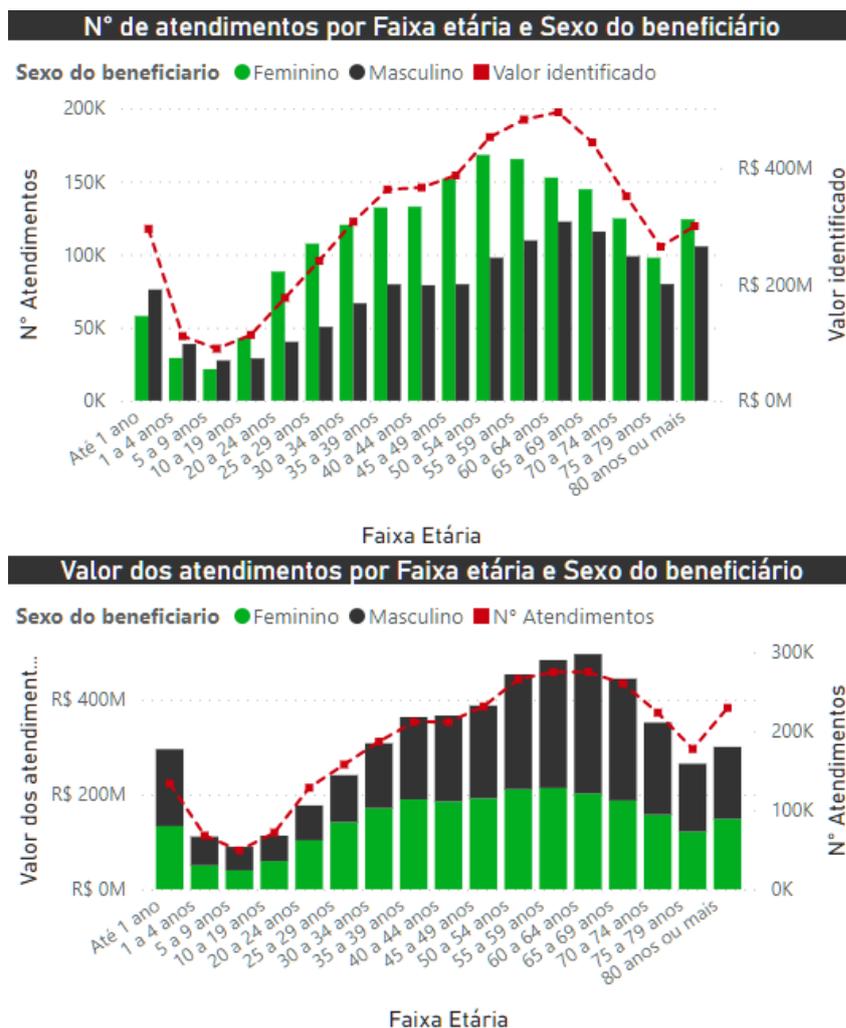
Figura 20 - Estudo da variável Dias de permanência interna (DPI)



Fonte: Elaborada pelo autor.

Os gráficos da Figura 21 foram gerados visando a entender a influência das variáveis Faixa etária e sexo dos beneficiários nos custos gerados. O primeiro demonstra o N° de atendimentos por Faixa etária e Sexo do beneficiário, com a linha em vermelho representando o Valor dos atendimentos gerado por cada faixa etária. Já o segundo gráfico traz analogamente o somatório dos Valores dos atendimentos por Faixa etária e Sexo do beneficiário. Esse, por sua vez, representa o N° de atendimentos pela linha em vermelho. Na Figura 21 é possível visualizar, por exemplo, que a faixa etária de 60 a 64 anos detém o maior número de atendimentos (275.072 atendimentos), além disso, é responsável pela maior parte dos custos gerados (R\$496.119.822).

Figura 21 - Número e Valor de atendimentos por Faixa etária e Sexo dos beneficiários



Fonte: Elaborada pelo autor.

3.3.4.3 Painel 3 – Localização dos atendimentos

O terceiro painel da *Dashboard* tem como foco relacionar características do local onde ocorrem os atendimentos e o valor de ressarcimento gerados por esses. Para isto, são utilizadas as informações a respeito da Natureza, Esfera e localização dos estabelecimentos. Os filtros disponíveis são: Data do atendimento, Procedimento cobrado, UF do estabelecimento, Esfera do estabelecimento, Município do estabelecimento e Natureza do estabelecimento, conforme ilustrados na Figura 22.

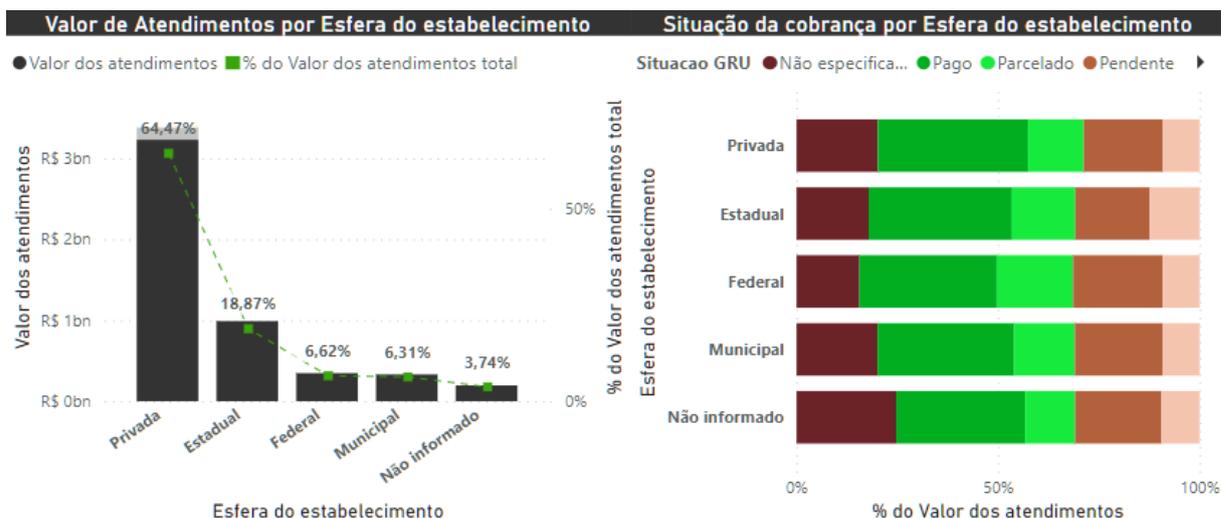
Figura 22 - Menu superior painel 3

Data do atendimento	Tipo do atendim...	Procedimento cobr...	Esfera do estab.	Natureza do est...	UF do estab.	Município do etab.
01/01/2016 - 01/09/2021	All	All	All	All	All	All
3157431 Nº de atendimentos		R\$ 5.245.516.788 Valor dos atendimentos	R\$ 3.467.239.190 Valor cobrado com GRU			

Fonte: Elaborada pelo autor.

Quanto à Esfera do estabelecimento, a Figura 23 ilustra, à esquerda, o somatório do Valor dos atendimentos para cada esfera existente, e à direita traz a visão 100% escalonada que reflete a situação da cobrança referente ao valor dos atendimentos para cada esfera de estabelecimento. Percebe-se que a esfera de estabelecimento privada é responsável por R\$3.381.525.665 (64,47%) dos custos gerados, no qual 20,19% não estão especificados, 37,22% foram pagos, 13,84% foram parcelados, 19,65% estão pendentes e 9,1% estão suspensos judicialmente.

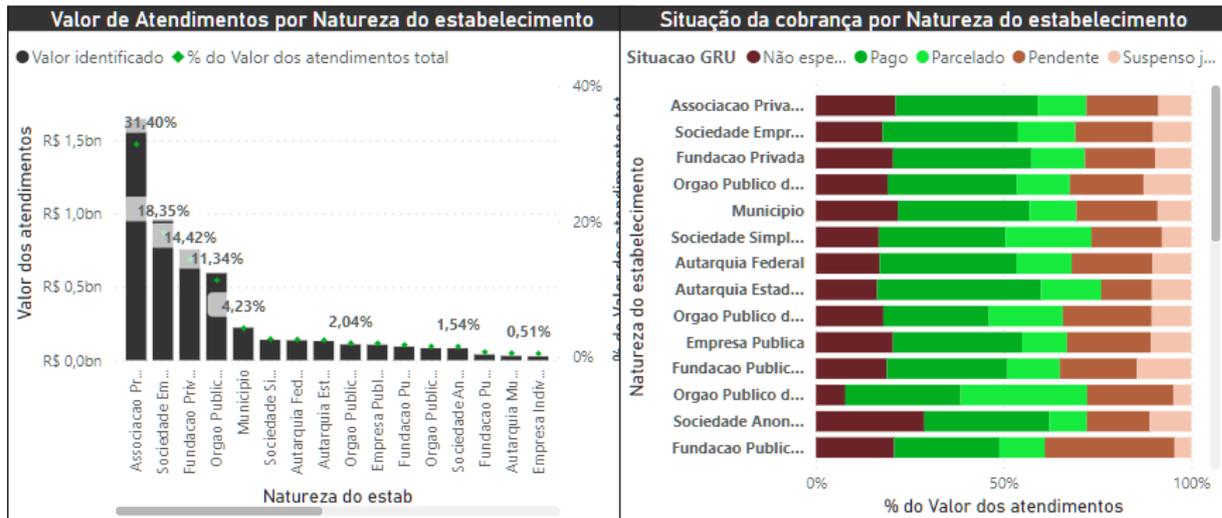
Figura 23 - Análise de esfera dos estabelecimentos



Fonte: Elaborada pelo autor.

Analogamente, a Figura 24 representa a relação entre o valor dos atendimentos e a Natureza do estabelecimento, tal como a visualização escalonada da situação da cobrança referente ao valor dos atendimentos. Por exemplo, os estabelecimentos com natureza associação privada são responsáveis por R\$1.647.228.825 (31,4%) dos custos gerados, sendo que 21,21% não estão especificados, 37,92% foram pagos, 12,91% foram parcelados, 19,26% estão pendentes e 8,71% estão suspensos judicialmente.

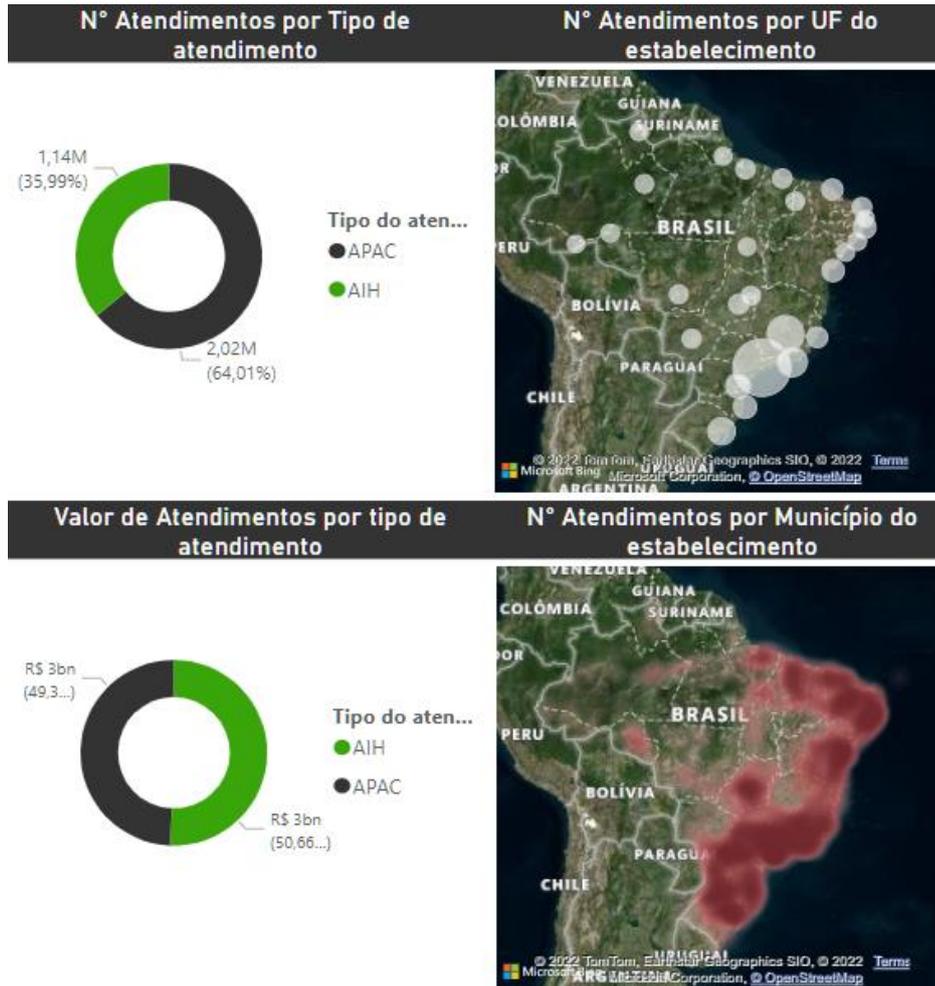
Figura 24 - Análise de natureza do estabelecimento



Fonte: Elaborada pelo autor.

A distribuição do número de atendimentos e valor gerado por esses, com base no tipo do atendimento, pode ser visualizada nos gráficos à esquerda na Figura 25. Ao lado direito, registra-se a frequência de atendimentos por UF do estabelecimento, assim como o Heatmap dos atendimentos por município do estabelecimento. É possível perceber, por exemplo, que os atendimentos APAC somam 2,02 milhões (64,01%) dos atendimentos e R\$ 3 bilhões em dívidas.

Figura 25 - Tipo de atendimentos e localização dos estabelecimentos



Fonte: Elaborada pelo autor.

3.3.4.4 Painel 4 - Atendimentos

O painel Atendimentos tem como finalidade evidenciar a quantidade de atendimentos para cada CID que ocorreu durante o período de escopo do trabalho. Os filtros disponíveis são: Data do atendimento, Registro da operadora, Razão Social, Tipo do atendimento, Procedimento cobrado, Capítulo CID, Sexo do beneficiário e Faixa etária, conforme Figura 26.

Figura 26 - Menu superior painel 4

Data do atendimento	Registro Operadora	Razao Social	Tipo do atendimento	Procedimento cob...	Capítulo CID	Sexo do benefici...	Faixa Etária
01/01/2016	All	All	All	All	All	All	All
		3157431	R\$ 5.245.516.788	R\$ 3.467.239.190			
		Número de atendimentos	Valor dos atendimentos	Valor cobrado com GRU			

Fonte: Elaborada pelo autor.

No painel, estão disponíveis quatro tabelas: à esquerda, as tabelas de número de atendimentos por operadora de planos de saúde e número de atendimentos por ano; à direita, as tabelas de percentual de atendimentos por capítulo e frequência de atendimentos por subcategoria da CID. As tabelas podem ser visualizadas na Figura 27, e partir destas é possível visualizar, por exemplo, que o capítulo CID de Neoplasia corresponde a 37,87% dos atendimentos. Além disso, o atendimento que ocorre com maior frequência (174.591 atendimentos) é o de rim transplantado. Percebe-se que o ano com maior número de atendimentos foi o de 2016 e as operadoras que mais possuem atendimentos a serem ressarcidos são a Hapvida, Notre Dame e Amil.

Figura 27 - Análise de atendimentos

Razao Social	N° Atendimentos	%
Hapvida Assistencia Medica S.A.	233584	7,40%
Notre Dame Intermédica Saúde S.A.	160626	5,09%
Amil Assistência Médica Internacional S.A.	160533	5,08%
Bradesco Saúde S.A.	106378	3,37%
Unimed Belo Horizonte Cooperativa De Trabalho Médico	78109	2,47%
Central Nacional Unimed - Cooperativa Central	76379	2,42%
Total	315743	100,00%

Year	N° Atendimentos	%
2021	388226	12,30%
2020	465302	14,74%
2019	535070	16,95%
2018	567590	17,98%
2017	586077	18,56%
2016	615166	19,48%
Total	3157431	100,00%

Capítulo CID	% Atendimentos
II. Neoplasias (tumores)	37,87%
CID Não Especificada	12,31%
XXI. Contatos com serviços de saúde	9,43%
XV. Gravidez parto e puerpério	6,70%
VIII. Doenças do ouvido e da apófise mastóide	5,21%
XIX. Lesões enven e alg out conseq causas externas	4,64%
I. Algumas doenças infecciosas e parasitárias	3,23%
X. Doenças do aparelho respiratório	3,03%
Total	100,00%

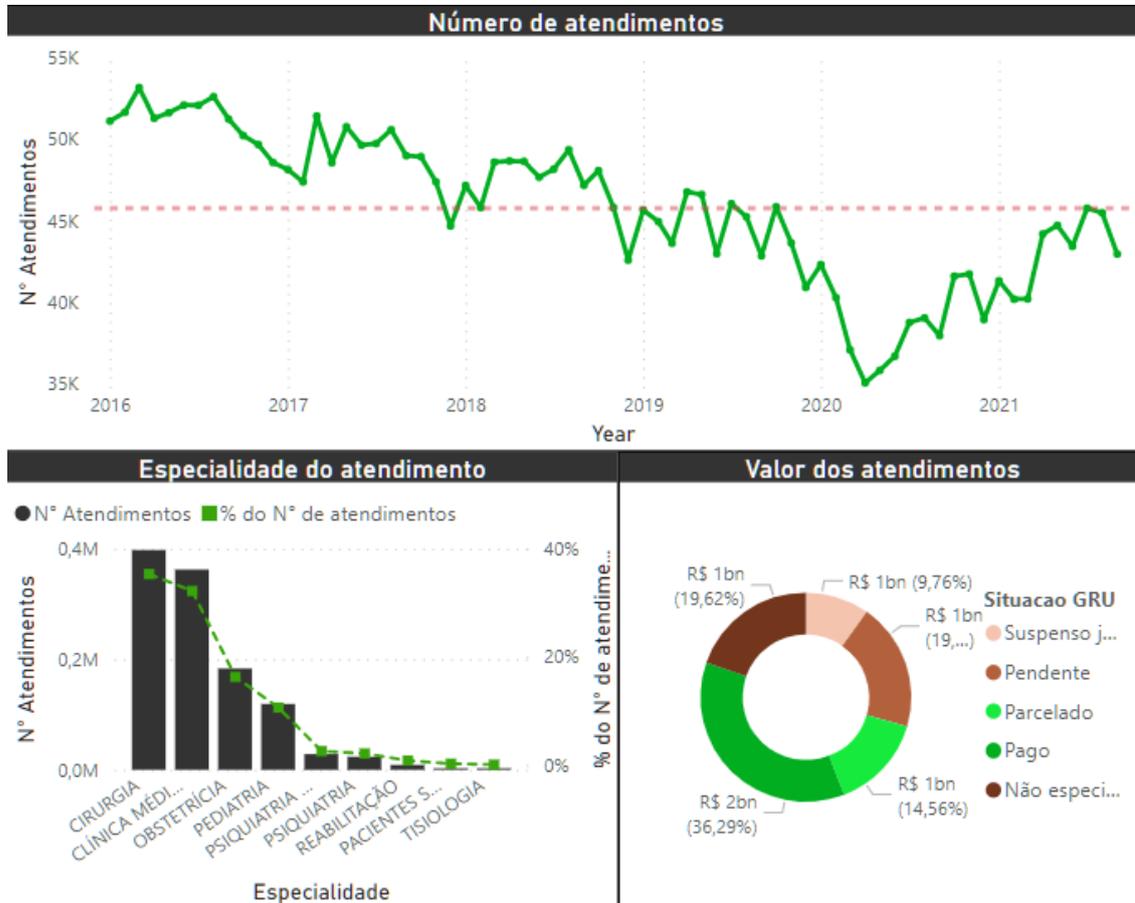
CID	N° Atendimentos	%	Descrição
#999	388562	12,31%	CID Não Especificada
Z940	174591	5,53%	Rim transplantado
C504	171730	5,44%	Neoplasia maligna do quadrante superior externo da mama
C500	166573	5,28%	Neoplasia maligna do mamilo e aréola
C61	145692	4,61%	Neoplasia maligna da próstata
C509	116456	3,69%	Neoplasia maligna da mama, não especificada
Total	3157431	100,00%	

Fonte: Elaborada pelo autor.

A região central é composta por três visualizações: a superior expressa o número de atendimentos em função do tempo. As inferiores demonstram a frequência de atendimentos para cada especialidade de atendimentos AIH e a situação da cobrança dos atendimentos, respectivamente, conforme Figura 28. É possível observar, por exemplo, que a especialidade

do atendimento mais frequente é a cirurgia, representando 35,31% dos atendimentos com especialidade.

Figura 28 - Frequência, especialidade e situação de cobrança dos atendimentos

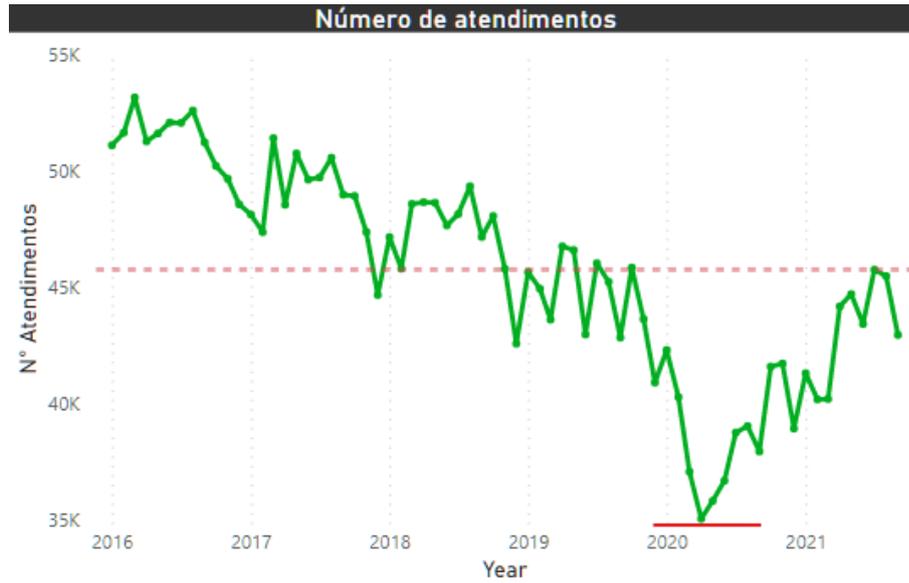


Fonte: Elaborada pelo autor.

3.4 CASO DE USO

Utilizando as informações coletadas durante a elaboração desta obra, este capítulo busca exemplificar o potencial analítico gerado pelo *Dashboard* desenvolvido. A partir do painel Atendimentos, constatou-se uma possível anomalia no comportamento do número de atendimentos identificados pela ANS para ressarcimento ao SUS no ano de 2020, como indicado em vermelho na Figura 29.

Figura 29 - Anomalia detectada no ano de 2020



Fonte: Elaborada pelo autor.

Com base nos dados históricos de 2017 a 2019, utilizou-se o modelo de Holter-Winters para o desenvolvimento de uma análise preditiva para o ano de 2020. A previsão levou em consideração a existência de tendência negativa do número de atendimentos (representada pela linha pontilhada) e a sazonalidade identificada por meio de quedas bruscas nos meses de dezembro e picos no decorrer do ano. Os limites inferiores e superiores da previsão representam um intervalo de confiança de 95%. Os valores reais para o ano de 2020 foram plotados pela linha vermelha sólida, como se vê na Figura 30.

Figura 30 - Modelo de previsão Holter-Winters



Fonte: Elaborada pelo autor.

Com o objetivo de comparar o valor previsto para número de atendimentos no ano de 2020 com a quantidade real, elaborou-se a Tabela 4.

Tabela 4 - Comparação previsão x real para o ano de 2020

Mês	Nº de atendimentos previsto	Nº de atendimentos real	Nº Real / Nº Previsto
Janeiro	42630	42308	99,2%
Fevereiro	41432	40284	97,2%
Março	44002	37087	84,3%
Abril	44006	35067	79,7%
Maio	43950	35830	81,5%
Junho	42934	36698	85,5%
Julho	43168	38766	89,8%
Agosto	44027	39043	88,7%
Setembro	42392	37956	89,5%
Outubro	42756	41596	97,3%
Novembro	41042	41728	101,7%
Dezembro	38423	38939	101,3%
Total	510762	465302	91,1%

Fonte: Elaborada pelo autor.

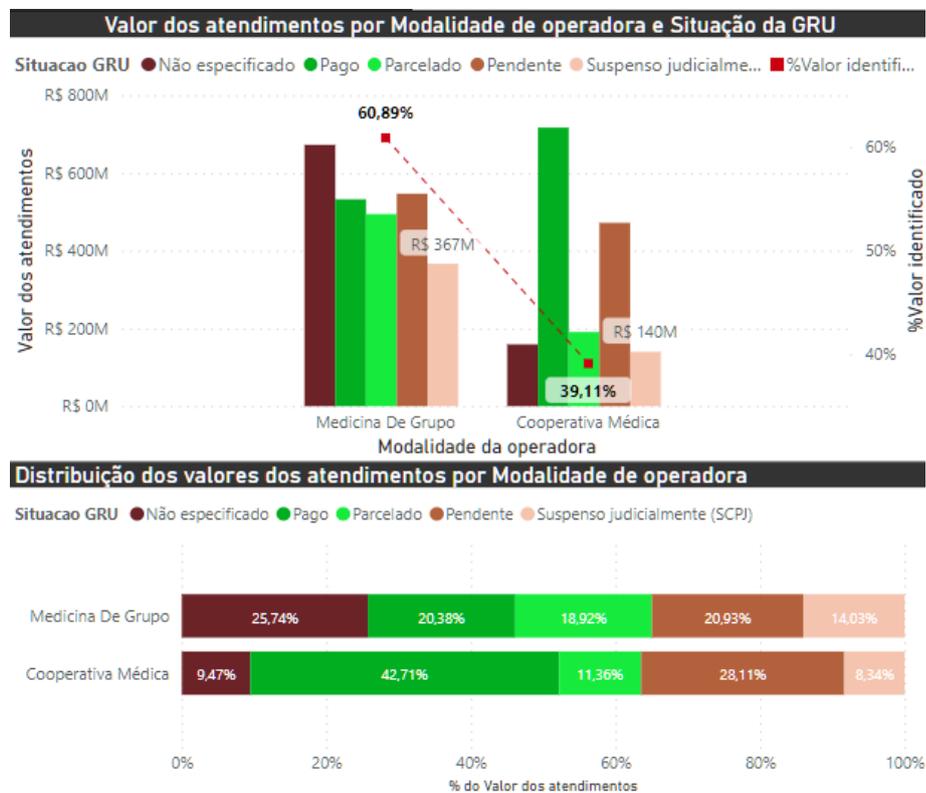
É possível observar que, no ano de 2020, o número de atendimentos foi 8,9% menor que o esperado, com destaque no mês de abril atingindo apenas 79,7% das expectativas. Levando em consideração o escopo do capítulo como uma demonstração de utilização do *Dashboard* e dos dados coletados, iniciou-se a busca por fontes secundárias de informação. A pandemia de Covid-19, decretada pela Organização Mundial da Saúde em março de 2020, é apontada como fator crucial para o desvio de previsão dos atendimentos (JOHNS HOPKINS, 2022). A demanda de recursos causou um sobrecarregamento de utilização do SUS, além do cancelamento ou adiamento de atendimentos então considerados não prioritários. Uma das áreas médicas fortemente afetadas foi a de tratamento de casos de câncer, pois as estimativas apontam que 43% de pacientes oncológicos sofreram algum impacto por conta da pandemia, destacando-se, dentre os diferentes tipos de impacto, o adiamento e cancelamento de tratamentos (AVELAR *et al.*, 2021).

4 RESULTADOS E DISCUSSÕES

A partir da elaboração deste trabalho, obteve-se como resultado o desenvolvimento de um painel descritivo a respeito dos atendimentos que geram custos de ressarcimento ao SUS. O uso do *dashboard* poderá auxiliar em pesquisas e análises, e a capacidade informativa desse recurso não se limita aos resultados e discussões apontados neste capítulo. Cabe a esta seção descrever exemplos de análises que podem ser realizadas com base na utilização dos painéis. A utilização do *dashboard* elaborado abarca uma nova visão de informações referentes ao uso do SUS por beneficiários de planos de saúde privados, e visa a trazer, de forma facilitada e transparente, dados para quaisquer usuários de interesse.

Os valores identificados pelos atendimentos são expressivos e tornaram relevante o entendimento da origem dos custos. Quanto às operadoras de plano de saúde, constatou-se que 81,82% dos custos são referentes a operadoras de modalidades Medicina de Grupo e Cooperativa Médica (Figura 16). Estas possuem 39,3% e 54,07% dos custos identificados em situação paga ou parcelada respectivamente (Figura 31).

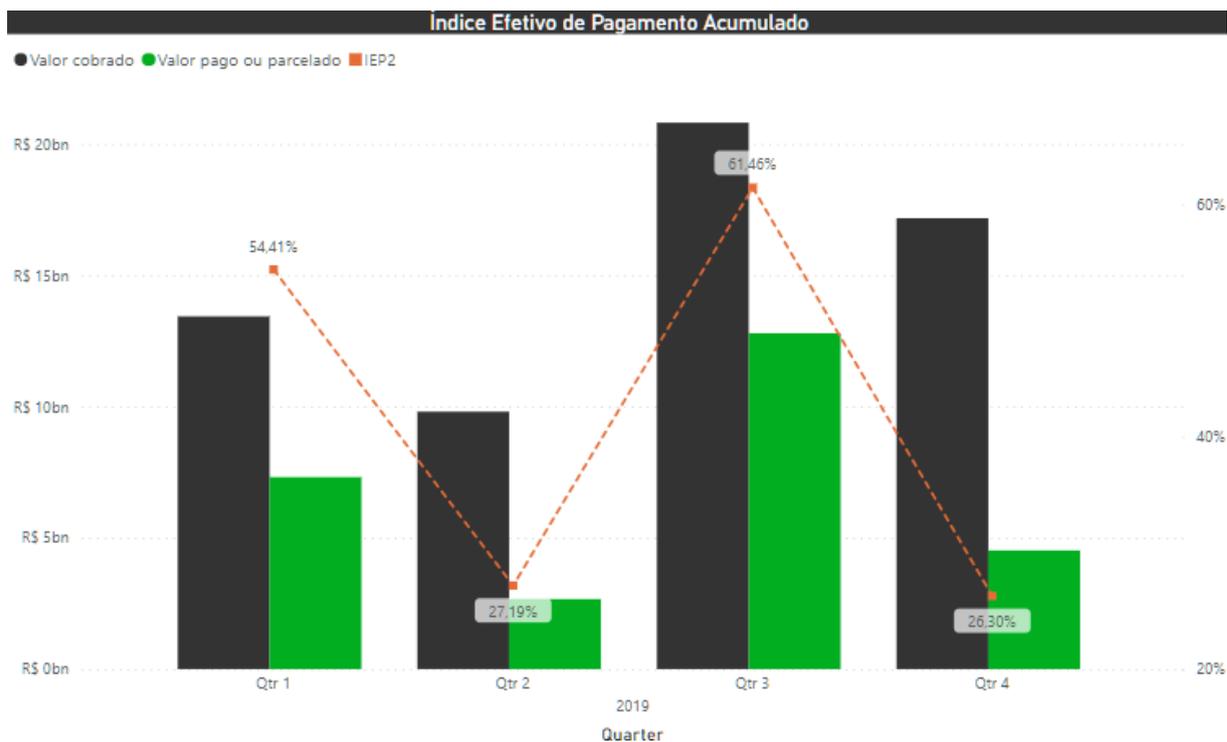
Figura 31 - Análise modalidade de operadora



Fonte: Elaborada pelo autor.

O Índice Efetivo de Pagamento Acumulado para os anos em análise oscila entre 62,74% e 69,18% (Figura 17). O valor registrado em 2019, de 25,87%, foi considerado anômalo devido ao baixo valor de atendimentos pagos nos segundo e quarto trimestres de 2019, identificado na base de dados 3 por motivo desconhecido, vide Figura 32.

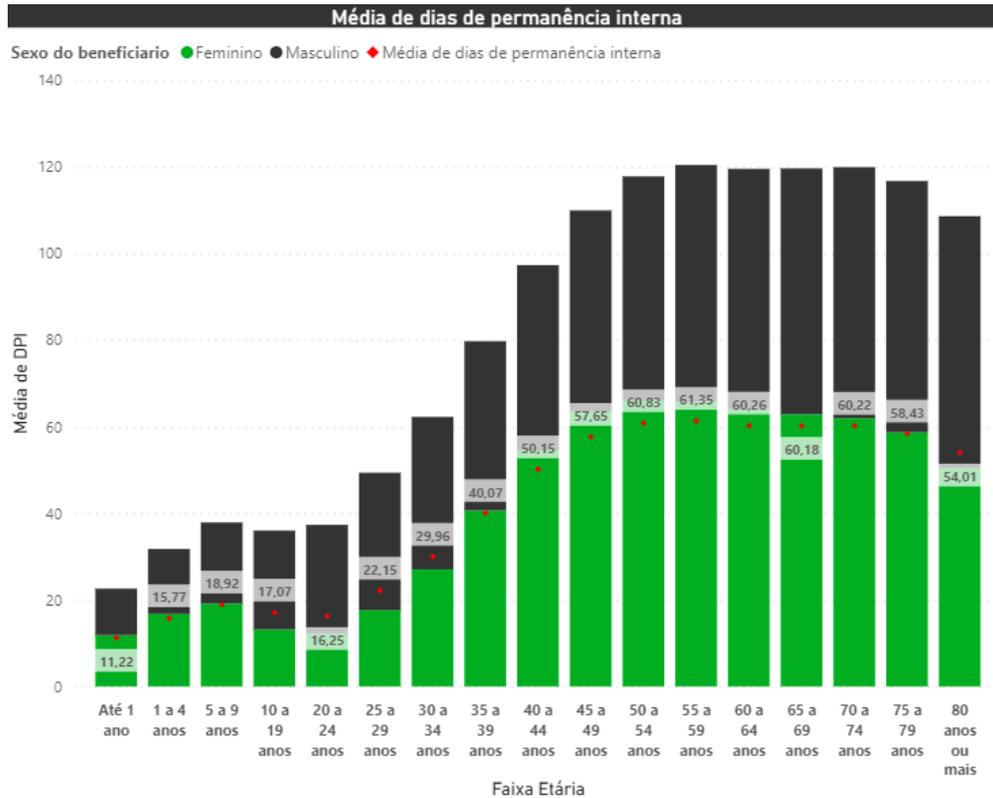
Figura 32 - Índice de Efetividade de Pagamento (IEP) 2019



Fonte: Elaborada pelo autor.

Analisando o perfil dos beneficiários, observou-se que a média de dias de permanência interna para mulheres é geralmente maior do que a de homens, com exceção dos grupos etários entre 10 e 34 anos, e acima de 80 anos. A distribuição média de dias de permanência apresenta um comportamento crescente até atingir seu pico entre 50 e 74 anos, com uma média aproximada de 60 dias. Após os 75 anos, há uma tendência de queda para essa variável, como visto na Figura 33.

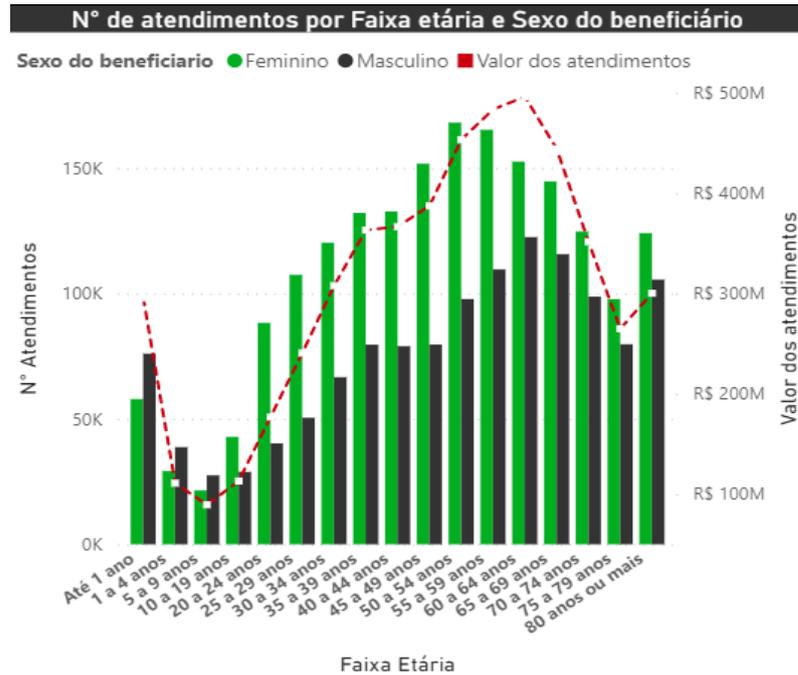
Figura 33 - Média de Dias de permanência interna (DPI)



Fonte: Elaborada pelo autor.

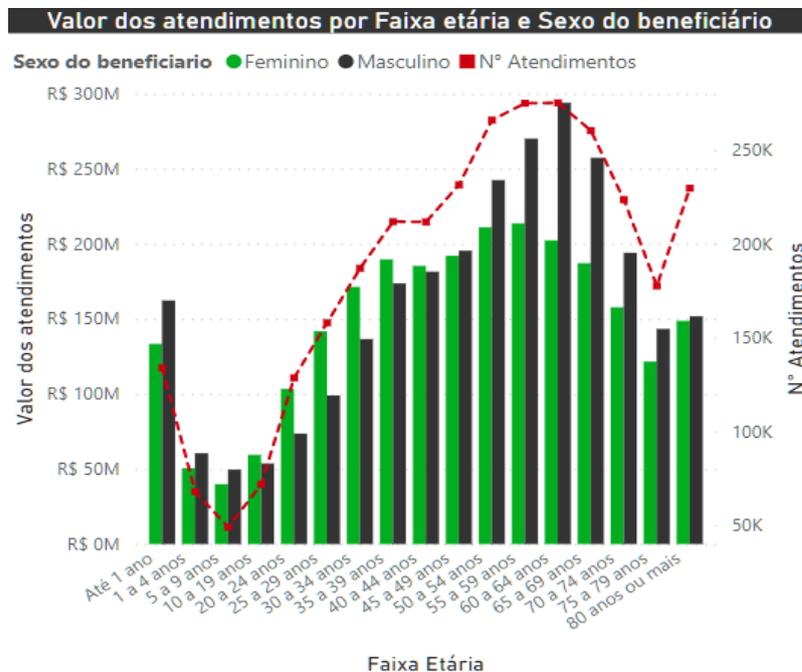
Observando os dados, podemos visualizar que, após a idade de 10 anos, o número de atendimentos realizados para o sexo feminino é maior do que o número de atendimentos para o sexo masculino. No entanto, embora haja menos atendimentos para o sexo masculino, o custo total desses atendimentos é maior do que o custo dos atendimentos realizados para o sexo feminino, representando 52,19% do valor total (Figura 19). Isso sugere que, em média, os atendimentos realizados para o sexo masculino são mais caros do que os atendimentos realizados para o sexo feminino (Figura 34). Ao examinar Figura 35 é possível perceber que a diferença entre o custo dos atendimentos masculinos e femininos é mais acentuada na faixa etária entre 50 a 69 anos.

Figura 34 - Número de atendimentos por Faixa etária e Sexo de beneficiários



Fonte: Elaborada pelo autor.

Figura 35 - Valor de atendimentos por Faixa etária e Sexo de beneficiários

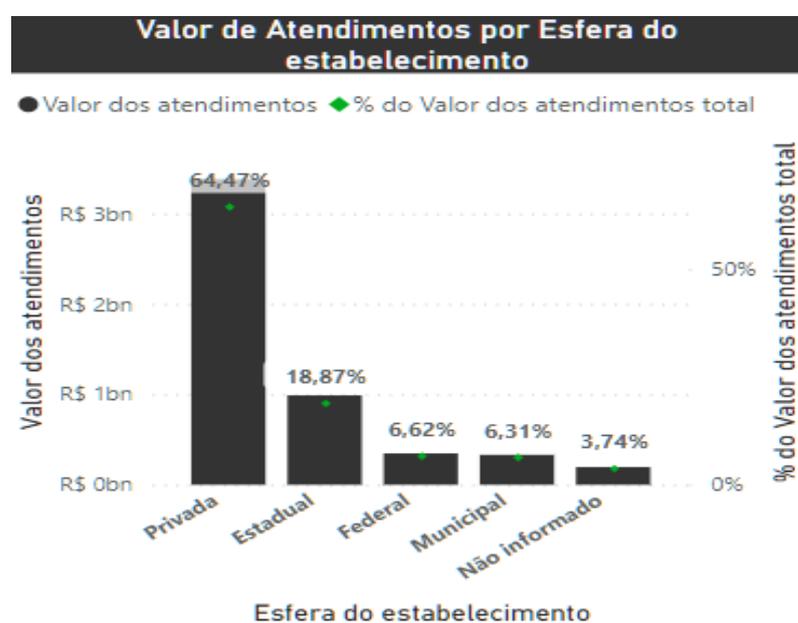


Fonte: Elaborada pelo autor.

Ao analisar as informações sobre os estabelecimentos onde os atendimentos ocorrem, pode-se observar que os estabelecimentos privados e estaduais, juntos, são responsáveis por mais de 83,34% do valor gerado com os atendimentos, de acordo com a Figura 36, adiante.

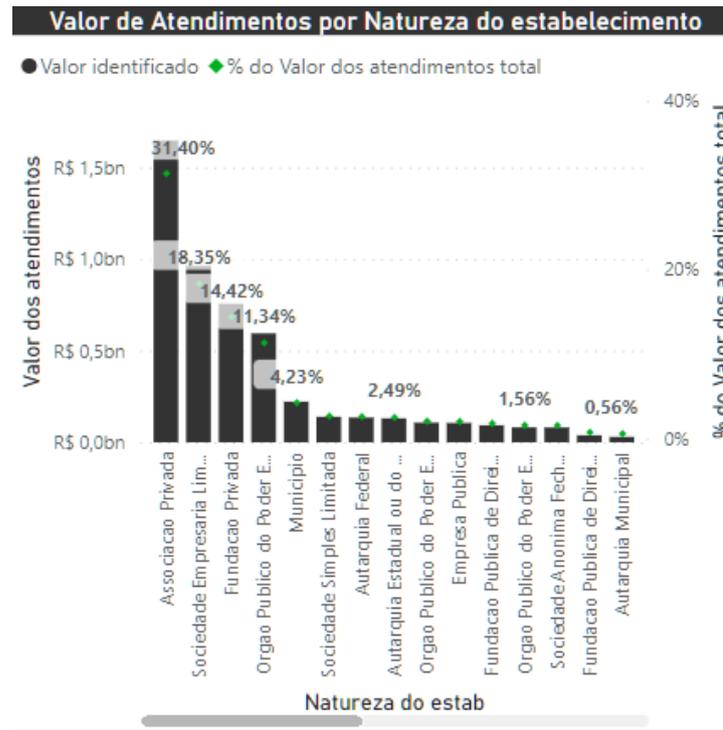
Os estabelecimentos de natureza: Associação privada, Sociedade empresarial limitada, Fundação privada e Órgão público do poder executivo estadual ou do Distrito Federal correspondem a 75,51% do valor total dos atendimentos, como indicado na Figura 37. Esses atendimentos estão concentrados principalmente nas regiões Nordeste, Sul e Sudeste, além disso, há uma incidência significativa no estado de Goiás e no Distrito Federal (Figura 38). Estas informações se tornam relevantes para entender a distribuição de atendimentos no país e, conseqüentemente, para o planejamento de estratégias de distribuição de recursos.

Figura 36 - Análise de Esfera do estabelecimento



Fonte: Elaborada pelo autor.

Figura 37 - Análise de Natureza do estabelecimento



Fonte: Elaborada pelo autor.

Figura 38 – Heatmap de frequência de atendimentos por Município do estabelecimento



Fonte: Elaborada pelo autor.

Observou-se que os atendimentos de Autorização de Internação Hospitalar (AIH) constituem, aproximadamente, 36% do número total de atendimentos e 50,7% do valor total dos atendimentos (Figura 25). Isso indica que os atendimentos de AIH são mais custosos para o Sistema Único de Saúde (SUS).

A partir do enriquecimento da base de ressarcimento fornecida pela ANS com a base de identificação de CID (Classificação Internacional de Doenças), foi possível verificar que o capítulo da CID com o maior número de atendimentos é o de neoplasias (tumores), representando 37,87% da base. A Figura 39, a seguir, informa o percentual de atendimentos para os capítulos mais frequentes.

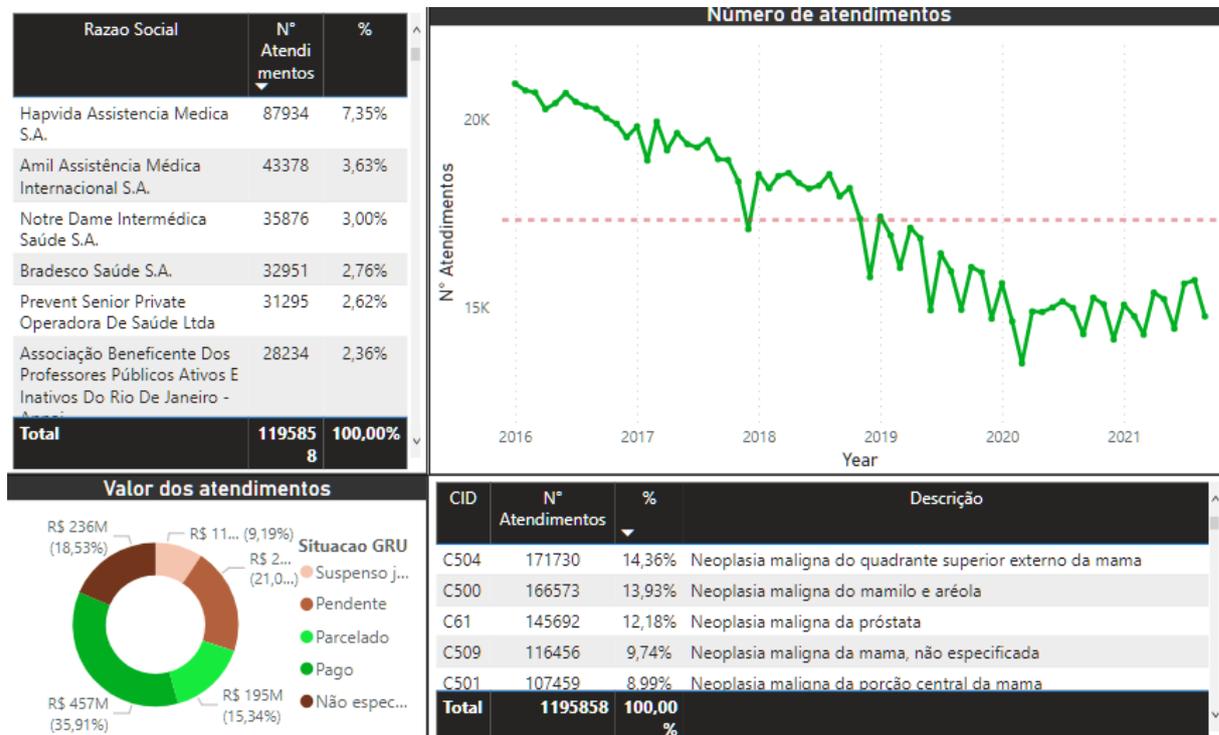
Figura 39 - Análise de frequência de atendimentos por capítulo da CID-10

DESCRICAÇÃO	% Atendimentos
Capítulo II - Neoplasias [tumores]	37,87%
CID Não Especificada	12,31%
Capítulo XXI - Fatores que influenciam o estado de saúde e o contato com os serviços de saúde	9,43%
Capítulo XV - Gravidez, parto e puerpério	6,70%
Capítulo VIII - Doenças do ouvido e da apófise mastóide	5,21%
Capítulo XIX - Lesões, envenenamento e algumas outras conseqüências de causas externas	4,64%
Capítulo I - Algumas doenças infecciosas e parasitárias	3,23%
Capítulo X - Doenças do aparelho respiratório	3,03%
Capítulo IX - Doenças do aparelho circulatório	3,01%
Capítulo XIV - Doenças do aparelho geniturinário	2,89%
Capítulo XI - Doenças do aparelho digestivo	2,72%
Capítulo VII - Doenças do olho e anexos	2,61%
Capítulo VI - Doenças do sistema nervoso	1,09%
Capítulo V - Transtornos mentais e comportamentais	1,04%
Total	100,00%

Fonte: Elaborada pelo autor.

O painel Atendimentos fornece um potencial analítico muito valioso para os usuários, o qual permite a identificação de padrões e tendências nos atendimentos realizados. É possível visualizar, para cada descrição de procedimento fornecida pelo CID, o número de atendimentos realizados por cada operadora de plano de saúde e a situação de cobrança desses. A Figura 40 demonstra para o capítulo de neoplasia a frequência de cada procedimento por CID e por operadora de plano de saúde. Além disso, permite ver a evolução desses atendimentos ao longo do período de análise.

Figura 40 - Análise do capítulo dos atendimentos referentes ao capítulo de Neoplasia da CID-10



Fonte: Elaborada pelo autor.

Por meio da análise deste painel, é possível identificar que os atendimentos de neoplasia maligna para o quadrante superior externo da mama, mamilo e aréola e próstata são os mais frequentes. É possível perceber, também, que as operadoras de planos privados que mais realizam atendimentos de Neoplasia são a Hapvida Assistência Médica S.A e Amil Assistência Médica Internacional S.A, compreendendo, juntas, 10,98% do número total de atendimentos deste capítulo da CID.

5 CONCLUSÃO

Os objetivos estabelecidos para este trabalho foram o desenvolvimento de um *dashboard* informativo, com base em dados públicos fornecidos pela Agência Nacional de Saúde Suplementar (ANS), que permita ao usuário analisar o contexto atual da dívida oriunda da carência do ressarcimento realizado ao SUS pelas operadoras privadas e a verificação da origem e desdobramento dos custos identificados para ressarcimento. Este trabalho atingiu seus objetivos, pois elaborou-se com sucesso o *dashboard* proposto, tornando possível as análises realizadas na Seção 4. Com a utilização deste, é possível entender quais operadoras têm beneficiários que utilizam o SUS, quais procedimentos são realizados por esses beneficiários e onde esses procedimentos ocorrem. Além disso, segmenta os beneficiários quanto a atributos de idade, sexo e faixa etária.

O *dashboard* traz uma nova visão sobre o contexto de utilização do SUS por beneficiários no período de janeiro de 2016 até setembro de 2021. Exemplificou-se o potencial analítico do *dashboard* a partir da demonstração do uso de um modelo preditivo aplicado aos resultados gerados pela obra. O enriquecimento de dados, provenientes da base de dados CID-10, foi crucial para os resultados alcançados. Dessa forma, foi possível identificar o motivo que leva a um beneficiário recorrer a um atendimento no SUS.

A limitação de informações a respeito da quantia efetivamente paga pelas operadoras privadas tornou impossível determinar a quantia exata devida pelas operadoras. Ademais, se mostrou intrigante o lapso de pagamentos indicado no segundo e quarto trimestre de 2019 pela base de dados referentes ao pagamento das operadoras privadas.

Os resultados encontrados neste trabalho poderão ser utilizados para melhor visualizar a utilização de recursos de saúde pública no Brasil. Espera-se que com o *dashboard* desenvolvido, beneficiários possam também identificar o cometimento de suas operadoras privadas para com suas responsabilidades.

Como sugestão para trabalhos futuros, indica-se explorar de forma mais aprofundada as bases de dados utilizadas neste trabalho, recomenda-se também, a busca por fontes de dados secundários que permitam determinar os valores efetivamente pagos pelas operadoras de planos privados. Assim, será possível diferenciar as parcelas pagas e pendentes de um procedimento pago de forma parcelada.

REFERÊNCIAS

- AMPLIMED – SISTEMA PARA CLÍNICAS. **O que é CID-10?** Guia completo para o dia a dia da clínica médica. 2020. Disponível em: <https://www.amplimed.com.br/cid-10/>. Acesso em: 23. DEZ. 2023.
- ANS – Agência Nacional de Saúde Suplementar. **Aspectos financeiros do ressarcimento ao SUS:** Orientações sobre os procedimentos de cobrança. Rio de Janeiro: Agência Nacional de Saúde Suplementar, 2018. Disponível em: https://www.gov.br/ans/pt-br/arquivos/aceso-a-informacao/perfil-do-setor/dados-e-indicadores-do-setor/dados-e-publicacoes-do-ressarcimento-ao-sus/cartilha_aspectos_financeiros_do_ressarcimento_ao_sus.pdf. Acesso em 12 dez. 2022.
- ANS – Agência Nacional de Saúde Suplementar. **Boletim Informativo Utilização do Sistema Único de Saúde por Beneficiários de Planos de Saúde e Ressarcimento ao SUS.** 13 ed. Rio de Janeiro: Agência Nacional de Saúde Suplementar, 2021a. Disponível em: https://www.gov.br/ans/pt-br/arquivos/aceso-a-informacao/perfil-do-setor/dados-e-indicadores-do-setor/dados-e-publicacoes-do-ressarcimento-ao-sus/Boletim_ressarcimento_SUS_13a_Edicao_Rev04.pdf. Acesso em: 10 dez. 2022.
- ANS – Agência Nacional de Saúde Suplementar. **Guia do Ressarcimento ao SUS – Impugnações e recursos.** Rio de Janeiro: Agência Nacional de Saúde Suplementar, 2019. Disponível em: <https://www.gov.br/ans/pt-br/arquivos/aceso-a-informacao/perfil-do-setor/dados-e-indicadores-do-setor/dados-e-publicacoes-do-ressarcimento-ao-sus/guia-ressarcimento-ao-sus-2019-versao-2a-pdf#:~:text=A%20organização%20do%20Guia%20está,a%20cobrança%20dos%20valores%20devidos>. Acesso em: 10 dez. 2022.
- ANS – Agência Nacional de Saúde Suplementar. **Panorama do ressarcimento ao SUS.** 2021b. Disponível em: <https://app.powerbi.com/view?r=eyJrIjoiY2Y4MDFmYWUtdNDdjZi00ZjkzLTkyYmEtZjA2MzNjZTk4MTJjIiwidCI6IjlkYmE0ODBlTRmYTctNDJmNC1iYmEzLTBmYjEzNzVmYmU1ZiJ9>. Acesso em: 22 dez. 2022.
- AVELAR, F. G. *et al.* Complicações da Covid-19: desdobramentos para o Sistema Único de Saúde. **Physis: Revista de Saúde Coletiva**, Rio de Janeiro, v. 31, n. 1, e310133, 2021.
- BRASIL. Governo Federal. **Dados Do Ressarcimento Ao Sus Por Beneficiário Identificado E Abi.** S.d. Disponível Em: <https://Dados.Gov.Br/Dados/Conjuntos-Dados/Beneficiarios-Identificados-Sus-Abi>. ACESSO EM: 23 DEZ. 2022.
- BRASIL. Governo Federal. **Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde.** S.d. Disponível em: <https://dados.gov.br/dados/conjuntos-dados/dados-do-ressarcimento-ao-sus-por-operadora-de-planos-de-saude>. Acesso em: 23 dez. 2022.
- BRASIL. **Constituição da República Federativa do Brasil de 1988.** Brasília, DF: Presidência da República, 1988.

CARVALHO, E. B. **A Regulamentação Dos Planos E Seguros Privados De Assistência A Saúde No Brasil: A Reconstrução De Uma Historia De Disputas**. 2003. 247f. Dissertação (Mestrado). Universidade Estadual De Campinas, Campinas, 2003.

COSTA, J. C. S. **Os impactos econômico-financeiros do ressarcimento ao SUS para as operadoras de planos de saúde**. 2018. 60f. Monografia. Prêmio ANS: concurso de Monografia sobre saúde suplementar. Belo Horizonte, 2018. Disponível em: [file:///C:/Users/berna/Desktop/TCC/Bibliografia/Artigos/2ª%20colocada%20\(T1\)%20-%20Jéssica%20Coeiro%20de%20Souza%20Costa%20\(monografia%20025\).pdf](file:///C:/Users/berna/Desktop/TCC/Bibliografia/Artigos/2ª%20colocada%20(T1)%20-%20Jéssica%20Coeiro%20de%20Souza%20Costa%20(monografia%20025).pdf). Acessado em: 20 dez. 2022.

DATASUS. **Arquivos em formato CSV**. 2008. Disponível em: <http://www2.datasus.gov.br/cid10/V2008/descrcsv.htm>. Acesso em: 22. Dez. 2022.

DATE, C. J. **Introdução a Sistemas de Bancos de Dados**. São Paulo: Elsevier editora, 2004.

DIZ, A. S. T. **Monitorização de epidemias sazonais de gripe com base em dados de pesquisa de motores de busca**. 2014. 122f. Dissertação (Mestrado Integrado em Ciências Farmacêuticas). Instituto Superior de Ciências da Saúde Egas Moniz, Portugal, 2015. Disponível em: <https://comum.rcaap.pt/handle/10400.26/13024>. Acesso em: 13 dez. 2022.

FÁTIMA, N. **O que é um data warehouse - definição, exemplo e benefícios**. 2019. Disponível em: <https://www.astera.com/pt/tipo/blog/definição-de-data-warehouse/>. Acesso em: 15 de dez. 2022.

GOLDSCHMIDT, R. **Data Mining**. Grupo GEN, 2015. E-book. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788595156395/>. Acesso em: 22 dez. 2022.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. Melbourne, Austrália: OTexts, 2021.

JOHNS HOPKINS. **COVID-19 Data repository by the center for systems science and engineering (CSSE)**. 2022. Disponível em: <https://github.com/CSSEGISandData/COVID-19>. Acesso em: 22 dez. 2022.

MACHADO, F. N. R. **Banco De Dados – Projeto e Implementação**. 4. Ed. São Paulo: Érica, 2020.

MACHADO, M. C. **Estudo descritivo do padrão do ressarcimento ao SUS e a negativa de procedimento pela operadora como possível causa do fato gerador**. 2019. 76f. Dissertação (Mestrado em Ciências). Universidade de São Paulo, Ribeirão Preto, 2019.

MCCABE, G. P.; CRAIG, B. A.; MOORE, D. S. **Introduction to the Practice of Statistics**. New York: W.H. Freeman and Company, 2016.

MICROSOFT. **O que é o Power Query?** 2022. Disponível em: <https://learn.microsoft.com/pt-br/power-query/power-query-what-is-power-query>. Acesso em: 15 de dez. 2022.

MILANI, A. M. P. *et al.* **Visualização de Dados**. 1. ed. Porto Alegre: Sagah, 2020.

MINAS GERAIS. SUS – 27 anos transformando a história da saúde no Brasil. 2015. Disponível: [https://www.saude.mg.gov.br/component/gmg/story/7152-sus-27-anos-transformando-a-historia-da-saude-no-brasil#:~:text=O%20SUS%20foi%20criado%20em,a%20toda%20a%20popula%C3%A7%C3%A3o%20brasileira](https://www.saude.mg.gov.br/component/gmg/story/7152-sus-27-anos-transformando-a-historia-da-saude-no-brasil#:~:text=O%20SUS%20foi%20criado%20em,a%20toda%20a%20popula%C3%A7%C3%A3o%20brasileira.). Acessado em: 10 dez. 2022.

MINISTÉRIO DA SAÚDE. Agência Nacional De Saúde Suplementar. **Resolução Normativa - RN nº 521, de 29 de abril de 2022**. Disponível Em: https://Bvms.Saude.Gov.Br/Bvs/Saudelegis/Ans/2022/Res0521_04_05_2022.Html. Acesso Em 23 Dez. 2022.

NUMPY. **Numpy1.24.0released**. 2022. Disponível em: <https://numpy.org/>. Acesso em: 23 dez. 2022.

PAIVA, C. H. A.; TEIXEIRA, L. A. Reforma sanitária e a criação do sistema único de saúde: notas sobre contextos e autores. **Hist. Cienc. Saúde**, Manguinhos, v. 21, n. 1, p. 15-35, 2014. <https://doi.org/10.1590/s0104-59702014000100002>

PANDAS. 2022. Disponível em: <https://pandas.pydata.org/>. Acesso em: 23 dez. 2022.

PANDAS. **pandas.concat**. 2022b. Disponível em: <https://pandas.pydata.org/docs/reference/api/pandas.concat.html?highlight=concat#pandas.concat>. Acesso em: 22 dez. 2022.

PANDAS. **pandas.DataFrame.merge**. 2022a. Disponível em: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.merge.html>. Acesso em: 22 dez. 2022.

PYTHON. **A Biblioteca virtual do Python**. 2021. Disponível em: <https://docs.python.org/pt-br/3/library/>. Acesso em: 23 dez. 2022.

PYTHON. **Guide**. 2021. Disponível em: <https://wiki.python.org/moin/BeginnersGuide/Overview>. Acesso em: 23 dez. 2022.

SANTOS, R. R. et al. **Fundamentos de Big Data**. 2021. E-book. ISBN 9786556901749. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786556901749/>. Acesso em: 23 dez. 2022.

SHARDA, R.; DELEN, D.; TURBAN, E. **Business Intelligence e Análise de Dados para Gestão do Negócio**. 2019. E-book. ISBN 9788582605202. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9788582605202/>. Acesso em: 22 dez. 2022.

VIDA, E. da S. *et al.* **Data warehouse**. 2021. E-book. Disponível em: <https://integrada.minhabiblioteca.com.br/#/books/9786556901916/>. Acesso em: 22 dez. 2022.

APÊNDICES

APÊNDICE A – TABELAS

Nesta seção estão listados os dicionários de dados fornecidos pelas bases de dados descritas na Seção 3.1.1 – Descrição das bases de dados. Esses documentos são disponibilizados juntamente às bases de dados, assim como a tabela de referência para renomeação da Base 1: Dados do Ressarcimento o SUS por Beneficiários Identificados e ABI.

Tabela 5 - Dicionário de dados referente a Base 1: Dados do Ressarcimento ao SUS por Beneficiário Identificado e ABI

Nome da Variável	Tipo	Descrição
COMPETENCIA	Date	Ano e mês do atendimento (AAA-MM)
ABI	Number	Aviso de Beneficiário Identificado
ANO_ABI	Number	Data do ABI
NUM_ATENDIMENTO	Number	Identificador do atendimento para os dados do ressarcimento ao SUS por operadora e Beneficiários Identificados SUS/ABI
REGISTRO_OPERADORA	Number	Número do registro da operadora na ANS
TIPO_ATENDIMENTO	Number	Identificação entre AIH ou APAC 1 - AIH 2- APAC
UF_BENEFICIARIO	Varchar	SIGLA UF AC Acre AL Alagoas AP Amapá AM Amazonas BA Bahia CE Ceará DF Distrito Federal ES Espírito Santo GO Goiás MA Maranhão MT Mato Grosso MS Mato Grosso do Sul MG Minas Gerais PA Pará PB Paraíba PR Paraná PE Pernambuco PI Piauí RJ Rio de Janeiro RN Rio Grande do Norte RS Rio Grande do Sul RO Rondônia RR Roraima SC Santa Catarina SP São Paulo SE Sergipe

		TO Tocantins EX Exterior XXDesconhecido
CONTRATAÇÃO_PLANO	Varchar	Época do contrato Anterior à Lei - Anterior à Lei 9.656/98 Posterior à Lei - Posterior à Lei 9.656/98 Vazio - Não Informado
ABRANGÊNCIA_PLANO	Varchar	Abrangência Geográfica do plano do beneficiário Nacional Grupo de Estados Estadual Grupo de Municípios Municipal Outra Vazio - Não Informado
TIPO_CONTRATAÇÃO	Varchar	Tipo de Contratação Individual ou Familiar Coletivo Empresarial Coletivo por adesão Coletivo não identificado Vazio - Não Informado
SEGMENTAÇÃO_PLANO	Varchar	Segmentação Grupo Ambulatorial – inclui Ambulatorial / Ambulatorial + odonto Hospitalar - inclui Hosp. c/ Obstetrícia + Odonto / Hosp. c/ Obstetrícia / Hosp. s/ Obstetrícia + Odonto / Hosp. s/ Obstetrícia Hospitalar e Ambulatorial - inclui Hosp. c/ Obstetrícia + Ambulatorial + Odonto / Hosp. c/ Obstetrícia + Ambulatorial / Hosp. c/s Obstetrícia + Ambulatorial / Hosp. s/ Obstetrícia + Ambulatorial + Odonto / Hosp. s/ Obstetrícia + Ambulatorial Referência - inclui apenas Referência Odontológico - inclui apenas Odontológico Vazio - Não Informado

FATOR_MODERADOR	Varchar	Tipo de fator moderador (mecanismo de regulação de utilização) do plano: Ausente Coparticipação Franquia Franquia + coparticipação Não identificado
SEXO_BENEFICIARIO	Varchar	Sexo M - Masculino F - Feminino Vazio - Não informado
FAIXA_ETARIA	Number	Faixa Etária 01 até 1 ano 02, 1 a 4 anos 03, 5 a 9 anos 04, 10 a 14 anos 04,05 15 a 19 anos 06,07 20 a 24 anos 08,09 25 a 29 anos 10,11 30 a 34 anos 12,13 35 a 39 anos 14,15 40 a 44 anos 16,17 45 a 49 anos 18,19 50 a 54 anos 20,21 55 a 59 anos 22,60 a 64 anos 23,65 a 69 anos 24, 70 a 74 anos 25, 75 a 79 anos 26, 80 anos ou mais 99 Inconsistente
ESPECIALIDADE	Varchar	Especialidade CIRURGIA OBSTETRÍCIA CLÍNICA MÉDICA PACIENTES SOB CUIDADOS PROLONGADOS PSIQUIATRIA FISILOGIA PEDIATRIA REABILITAÇÃO PSIQUIATRIA EM HOSPITAL DIA

COD_PROCEDIMENTO	Number	Código do procedimento principal no SIGTAP (Sistema de Gerenciamento da Tabela de Procedimentos, Medicamentos e OPM do SUS)
COD_CID	Varchar	Código CID (Classificação Internacional de Doenças e Problemas Relacionados à Saúde)
DIAS_PERM_INTERN	Number	Dias de permanência na internação - TIPO_ATENDIMENTO = 2 (APACs), o valor apresentado neste campo diz respeito ao intervalo entre as datas de início e fim do tratament
CD_CNES	Number	Código do Cadastro Nacional de Estabelecimentos de Saúde do prestador do atendimento
UF_ESTAB	Varchar	UF do estabelecimento AC Acre AL Alagoas AP Amapá AM Amazonas BA Bahia CE Ceará DF Distrito Federal ES Espírito Santo GO Goiás MA Maranhão MT Mato Grosso MS Mato Grosso do Sul MG Minas Gerais PA Pará PB Paraíba PR Paraná PE Pernambuco PI Piauí RJ Rio de Janeiro RN Rio Grande do Norte RS Rio Grande do Sul RO Rondônia RR Roraima SC Santa Catarina SP São Paulo SE Sergipe

		TO Tocantins; EX Exterior
MUNICIPIO_ESTAB	Varchar	Município do estabelecimento
ESFERA_ESTAB	Number	Esfera Administrativa do estabelecimento 1 Federal 2 Estadual 3 Municipal 4 Privada 0 Não informado
NATUREZA_ESTAB	Varchar	Natureza jurídica do estabelecimento - Até mai/12 informações referem-se aos campos "Natureza" e "Esfera Administrativa" / De jun/12 a out/15, aos campos "Natureza" e "Esfera Administrativa" e também aos campos "Natureza Jurídica" e "Esfera Jurídica" / A partir de nov/15, aos campos "Natureza Jurídica" e "Esfera Jurídica"
VL_ATENDIMENTO	Number	Valor do atendimento (vírgula como delimitador de decimais)
VL_RESSARC	Number	Valor cobrado com GRU (vírgula como delimitador de decimais)
VL_PAGO_RESSARC	Number	Valor pago no ressarcimento (vírgula como delimitador de decimais)
DT_ATUALIZACAO	Date	Data de atualização do arquivo (AAA-MM-DD)

Fonte: ANS (2022)

Tabela 6 - Dicionário de dados referentes à Base 2: Dados do Ressarcimento ao SUS por Operadora de Planos de Saúde

Nome do Campo	Tipo	Descrição
#COMPETENCIA	Varchar	Ano e mês do atendimento (formato ISO 8601: AAAA-MM)
ABI	Varchar	Aviso de Beneficiário Identificado
NUM_ATENDIMENTO	Number	Identificador do atendimento para os dados do conjunto “ ressarcimento ao SUS por operadora e Beneficiários Identificados SUS/ABI”
TIPO_ATENDIMENTO	Number	Identificação entre AIH ou APAC

		1- AIH 2- APAC
REGISTRO_OPERADORA	Number	Número do Registro da operadora na ANS
RAZAO_SOCIAL	Varchar	Razão Social das operadoras de Plano de Saúde
MODALIDADE_OPERADORA	Varchar	Tipo de Operadora AUTOGESTÃO COOPERATIVA MÉDICA FILANTROPIA MEDICINA DE GRUPO SEGURADORA ESPECIALIZADA EM SAÚDE
SITUACAO_OPERADORA	Varchar	Situação Ativa Inativa
CD_CNES	Number	Código do Cadastro Nacional de Estabelecimentos de Saúde do prestador do atendimento
STATUS_ANALISE_ATENDIMENTO	Varchar	Situação da análise do atendimento AVALIANDO EM 1º INSTÂNCIA AVALIANDO EM 2º INSTÂNCIA DEFERIDA EM 1º INSTÂNCIA DEFERIDA EM 2º INSTÂNCIA IMPUGNADA INTEMPESTIVAMENTE EM 1º INSTÂNCIA IMPUGNADA INTEMPESTIVAMENTE EM 2º INSTÂNCIA INDEFERIDA EM 1º INSTÂNCIA, SEM RECURSO INDEFERIDA EM 2º INSTÂNCIA NÃO IMPUGNADA
VALOR_ATENDIMENTO	Number	Valor do atendimento
VALOR_COBRADO	Number	Valor cobrado com GRU
GRU_EMITIDA	Varchar	GRU foi emitida? SIM NÃO

NUMERO_GRU	Number	Número da GR
SITUACAO_GRU	Varchar	SITUAÇÃO DA GRU pago; parcelado; pendente de pagamento (inclui suspensão SCI); suspenso judicialmente (SCPJ).
VALOR_ORIGINAL_GRU	Number	Valor original da GRU
VALOR_ATUALIZADO_GRU	Number	Valor atualizado da GRU
DATA_EMISSAO_GRU	Date	Data de emissão da GRU no formato ISO 8601: AAAA-MM-DD
DATA_VENCIMENTO_GRU	Date	Data de vencimento da GRU no formato ISO 8601: AAAA-MM-DD
DATA_PAGAMENTO_GRU	Date	Data de pagamento da GRU no formato ISO 8601: AAAA-MM-DD
DATA_DIVIDA_ATIVA	Date	Data de inscrição em Dívida Ativa no formato ISO 8601: AAAA-MM-DD
STATUS_PARCELAMENTO	Varchar	Situação do parcelamento AVALIANDO DEFERIDO DEFERIDO-CANCELADO INDEFERIDO PROPOSTO PROPOSTO-CANCELADO QUITADO REACTUADA
DATA_ATUALIZACAO_ARQUIVO	Date	Data de atualização do arquivo no formato ISO 8601: AAA-MM-DD

Fonte: ANS (2022)

Tabela 7 - Dicionário de dados referente a Base 3: Dados de pagamento do Ressarcimento ao SUS

Nome do Campo	Tipo	Tamanho	Descrição
COMPETENCIA	Varchar	10	Competência dos dados no formato ISO 8601 (AAAA/MM)
CODIGO_OPERADORA	Varchar	10	Número de Registro da Operadora na ANS
RAZAO_SOCIAL	Varchar	150	Razão Social da Operadora

SITUACAO_OPERADORA	Varchar	50	Situação do registro da operadora na ANS (Ativa ou Inativa)
VALOR_PAGO_ACUMULADO	Number	15	Valor total pago pela operadora a título de Ressarcimento ao SUS (vírgula como delimitador de decimais)
VALOR_PARCELADO_ACUMULADO	Number	15	Valor total parcelado pela operadora referente a cobranças a título de Ressarcimento ao SUS (vírgula como delimitador de decimais)
VALOR_COBRADO_ACUMULADO	Number	15	Valor total cobrado pela ANS à operadora a título de Ressarcimento ao SUS (vírgula como delimitador de decimais)
INDICE_ADIMPLENCIA	Number	10	Percentual de valores totais pagos ou parcelados pela operadora, em relação ao total cobrado pela ANS à operadora a título de Ressarcimento ao SUS
VALOR_PAGO_ACUMULADO_5ANOS	Number	15	Valor pago pela operadora a título de Ressarcimento ao SUS, nos últimos 5 anos (vírgula como delimitador de decimais)
VALOR_PARCELADO_ACUMULADO_5ANOS	Number	15	Valor parcelado pela operadora referente a cobranças a título de Ressarcimento ao SUS, nos últimos 5 anos (vírgula como delimitador de decimais)
VALOR_COBRADO_ACUMULADO_5ANOS	Number	15	Valor cobrado pela ANS à operadora a título de Ressarcimento ao SUS, nos últimos 5 anos (vírgula como delimitador de decimais)
INDICE_ADIMPLENCIA_5ANOS	Number	10	Percentual de valores pagos ou parcelados pela operadora, em relação ao cobrado pela ANS à operadora a título de Ressarcimento ao SUS, nos últimos 5 anos
DATA_ATUALIZACAO	Date	10	Data de atualização do arquivo no formato ISO 8601 (AAAA/MM/DD)

Fonte: ANS (2022)

Tabela 8 - Dicionário de dados referente à Base 4: CID019, Classificação Internacional de Doenças

Nome do Campo	Descrição
CID-10-CAPITULOS	
NUMCAP	Número do capítulo (em numeração arábica); se igual a zero, indica o capítulo que contém os códigos não oficialmente pertinentes à CID-10
CATINIC	Código da primeira categoria do capítulo

CATFIM	código da última categoria do capítulo
DESCRICA0	Descrição (nome) do capítulo
DESCABREV	Descrição (nome) abreviado do capítulo, com até 50 caracteres
CID-10-SUBCATEGORIAS	
SUBCAT	Código da subcategoria (sem incluir ponto); para categorias sem subcategorias, o quarto caractere está em branco
CLASSIF	Indica se a situação da subcategoria em relação à classificação cruz/asterisco: em branco: não tem dupla classificação; +: classificação por etiologia; e *: classificação por manifestação.
RESTRSEXO	Indica se a subcategoria só pode ser usada para homens ou mulheres: Em branco: pode ser utilizada em qualquer situação; f: só deve ser utilizada para o sexo feminino; e m: só deve ser utilizada para o sexo masculino.
CAUSAOBITO	Indica se a subcategoria pode causar óbito: em branco: não há restrição; e n: a subcategoria tem pouca probabilidade de causar óbito. Observação: além disto, deve-se atentar para o fato de que as subcategorias da classificação asterisco não devem ser utilizadas na classificação de causas de óbitos, assim como as subcategorias do capítulo xix e do capítulo xxi.
DESCRICA0	Descrição (nome) da subcategoria
DESCABREV	Descrição (nome) abreviado da subcategoria, com até 50 caracteres (inclui o código da subcategoria)
REFER	Contém, quando a subcategoria tiver dupla classificação, o código da subcategoria segundo a outra classificação (nem todos os casos de dupla classificação contém este campo)
EXCLUIDOS	Contém o(s) código(s) de subcategorias excluídas que agora fazem parte desta subcategoria

Fonte: Elaborada pelo autor

Tabela 9 - Renomeação de variáveis referentes à Base 1: Dados do Ressarcimento ao SUS por Beneficiários Identificados e ABI

Variável	Nome após renomeação
#COMPETENCIA	COMPETENCIA
NUM_ATENDIMENTO	Nº Atendimento
REGISTRO_OPERADORA	Registro Operadora
TIPO_ATENDIMENTO	Tipo do atendimento
UF_BENEFICIARIO	UF do Beneficiario
SEXO_BENEFICIARIO	Sexo do beneficiario
FAIXA_ETARIA	Faixa etaria
ESPECIALIDADE	Especialidade
DIAS_PERM_INTERN	Dias_perm_interna
UF_ESTAB	UF do estab
MUNICIPIO_ESTAB	Municipio do etab
ESFERA_ESTAB	Esfera do estab
NATUREZA_ESTAB	Natureza do estab
VL_ATENDIMENTO	Valor do atendimento
VL_PAGO_RESSARC	Valor pago ressarcimento
RAZAO_SOCIAL	Razao Social
MODALIDADE_OPERADOR A	Modalidade da operadora
SITUACAO_OPERADORA	Situacao da operadora
VALOR_COBRADO	Valor cobrado
GRU_EMITIDA	GRU emitida
SITUACAO_GRU	Situacao GRU

Fonte: Elaborada pelo autor.

APÊNDICE B – CÓDIGOS DE PROGRAMAÇÃO

- Código escrito em linguagem Python referente ao item 3.2.3 Preparação de dados: Base 3 Dados de pagamento do Ressarcimento ao SUS por operadora:

Figura 41 - Código Python referente à preparação de dados da Base 3: Dados de pagamento do Ressarcimento ao SUS por operadora:

```
#Importação da Base 3 Dados de pagamento do Ressarcimento ao SUS por operadora (janeiro de 2016 a setembro de 2021)
base1 = pd.read_csv(path+'2016_Indice_de_Efetivo_Pagamento(2).csv', sep=";", encoding='ISO-8859-1')
base2 = pd.read_csv(path+'2017_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base3 = pd.read_csv(path+'2018_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base4 = pd.read_csv(path+'2019_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base5 = pd.read_csv(path+'202001_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base6 = pd.read_csv(path+'202002_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base7 = pd.read_csv(path+'202003_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base8 = pd.read_csv(path+'202004_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base9 = pd.read_csv(path+'202005_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base10 = pd.read_csv(path+'202006_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base11 = pd.read_csv(path+'202007_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base12 = pd.read_csv(path+'202008_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base13 = pd.read_csv(path+'202009_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base14 = pd.read_csv(path+'202010_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base15 = pd.read_csv(path+'202011_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base16 = pd.read_csv(path+'202012_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base17 = pd.read_csv(path+'202101_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base18 = pd.read_csv(path+'202102_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base19 = pd.read_csv(path+'202103_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base20 = pd.read_csv(path+'202104_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base21 = pd.read_csv(path+'202105_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base22 = pd.read_csv(path+'202106_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base23 = pd.read_csv(path+'202107_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base24 = pd.read_csv(path+'202108_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base25 = pd.read_csv(path+'202109_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base26 = pd.read_csv(path+'202110_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base27 = pd.read_csv(path+'202111_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base28 = pd.read_csv(path+'202112_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base29 = pd.read_csv(path+'202201_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base30 = pd.read_csv(path+'202202_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base31 = pd.read_csv(path+'202203_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base32 = pd.read_csv(path+'202204_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base33 = pd.read_csv(path+'202205_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base34 = pd.read_csv(path+'202206_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base35 = pd.read_csv(path+'202207_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base36 = pd.read_csv(path+'202208_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')
base37 = pd.read_csv(path+'202209_Indice_de_Efetivo_Pagamento.csv', sep=";", encoding='ISO-8859-1')

#concatenação das bases
frames = [base1, base2, base3, base4, base5, base6, base7, base8, base9, base10, base11,
          base12, base13, base14, base15, base16, base17, base18, base19, base20, base21,
          base22, base23, base24, base25, base26, base27, base28, base29, base30, base31,
          base32, base33, base34, base35, base36, base37]
base = pd.concat(frames)

#Remoção de colunas indesejadas
base.drop(['VALOR_PAGO_ACUMULADO_SANOS', 'VALOR_PARCELADO_ACUMULADO_SANOS', 'VALOR_COBRADO_ACUMULADO_SANOS',
          'INDICE_ADIMPLENCIA_SANOS', 'INDICE_ADIMPLENCIA', 'DT_ATUALIZACAO'], axis=1, inplace=True)
```

Fonte: Elaborada pelo autor.

- **Criação da base de dados Faixa Etária, gerada no formato JSON dentro do Power Query, referente ao item:**

Source=

```
Table.FromRows(Json.Document(Binary.Decompress(Binary.FromText("Tc89EsIgfEXhrTD
UKfh5QCxdB5MipYWmUBflOtyY76Txljn5mAtzxhyXeH19PyGH/XHEbZmxePKvYJTnm
aqn5unyT4ZK3rLE7rEQixxeiZwuInPyWqFVaGa7YqtatgxrahkzrKllrWGb2MJawzaxhbWO7
Wp5bMd2tdxhYifaRsUOtdxsTWcIxzvc95v/2H4=",BinaryEncoding.Base64),
Compression.Deflate)), let _t = ((type nullable text) meta [Serialized.Text = true]) in type
table [#"ID Faixa etária" = _t, #"Faixa Etária" = _t]),
```

- **Código em linguagem M utilizado para o tratamento de dados referente ao item 3.3.2.1 Base de dados: Ressarcimento ao SUS**

let

Source =

```
Csv.Document(File.Contents("C:\Users\berna\Downloads\basefinal.csv"),[Delimiter=";",
Columns=44, Encoding=65001, QuoteStyle=QuoteStyle.None]),
```

```
#"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
```

```
#"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"",
Int64.Type}, {"#COMPETENCIA", type date}, {"ABI", type text}, {"ANO_ABI",
Int64.Type}, {"NUM_ATENDIMENTO", Int64.Type}, {"REGISTRO_OPERADORA",
Int64.Type}, {"TIPO_ATENDIMENTO", Int64.Type}, {"UF_BENEFICIARIO", type text},
{"CONTRATACAO_PLANO", type text}, {"ABRANGENCIA_PLANO", type text},
{"TIPO_CONTRATACAO", type text}, {"SEGMENTACAO_PLANO", type text},
{"SEXO_BENEFICIARIO", type text}, {"FAIXA_ETARIA", Int64.Type},
{"ESPECIALIDADE", type text}, {"COD_PROCEDIMENTO", Int64.Type}, {"COD_CID",
type text}, {"DIAS_PERM_INTERN", Int64.Type}, {"CD_CNES", Int64.Type},
{"UF_ESTAB", type text}, {"MUNICIPIO_ESTAB", type text}, {"ESFERA_ESTAB",
Int64.Type}, {"NATUREZA_ESTAB", type text}, {"VL_ATENDIMENTO", Int64.Type},
{"VL_RESSARC", Int64.Type}, {"VL_PAGO_RESSARC", Int64.Type},
{"DT_ATUALIZACAO", type date}, {"RAZAO_SOCIAL", type text},
{"MODALIDADE_OPERADORA", type text}, {"SITUACAO_OPERADORA", type text},
```

```
{ "STATUS_ANALISE_ATENDIMENTO", type text}, {"VALOR_ATENDIMENTO",
Int64.Type}, {"VALOR_COBRADO", Int64.Type}, {"GRU_EMITIDA", type text},
{"NUMERO_GRU", type number}, {"SITUACAO_GRU", type text},
{"VALOR_ORIGINAL_GRU", Int64.Type}, {"VALOR_ATUALIZADO_GRU",
Int64.Type}, {"DATA_EMISSAO_GRU", type date}, {"DATA_VENCIMENTO_GRU",
type date}, {"DATA_PAGAMENTO_GRU", type date}, {"DATA_DIVIDA_ATIVA", type
date}, {"STATUS_PARCELAMENTO", type text}, {"F", Int64.Type}}),
```

```
#"Removed Columns" = Table.RemoveColumns("#Changed Type",{ "", "ABI",
"ANO_ABI", "COD_PROCEDIMENTO", "CD_CNES", "VL_RESSARC",
"DT_ATUALIZACAO", "STATUS_ANALISE_ATENDIMENTO",
"VALOR_ATENDIMENTO", "NUMERO_GRU", "VALOR_ORIGINAL_GRU",
"VALOR_ATUALIZADO_GRU", "DATA_EMISSAO_GRU",
"DATA_VENCIMENTO_GRU", "DATA_PAGAMENTO_GRU",
"DATA_DIVIDA_ATIVA", "STATUS_PARCELAMENTO", "F",
"CONTRATAcao_PLANO", "ABRANGENCIA_PLANO", "TIPO_CONTRATAcao",
"SEGMENTACAO_PLANO"}),
```

```
#"Renamed Columns" = Table.RenameColumns("#Removed
Columns",{ {"#COMPETENCIA", "COMPETENCIA"}, {"NUM_ATENDIMENTO", "N°
Atendimento"}, {"REGISTRO_OPERADORA", "Registro Operadora"},
{"TIPO_ATENDIMENTO", "Tipo do atendimento"}, {"UF_BENEFICIARIO", "UF do
Beneficiario"}, {"SEXO_BENEFICIARIO", "Sexo do beneficiario"}, {"FAIXA_ETARIA",
"Faixa etaria"}, {"ESPECIALIDADE", "Especialidade"}, {"DIAS_PERM_INTERN",
"Dias_perm_interna"}, {"UF_ESTAB", "UF do estab"}, {"MUNICIPIO_ESTAB",
"Municipio do etab"}, {"ESFERA_ESTAB", "Esfera do estab"}, {"NATUREZA_ESTAB",
"Natureza do estab"}, {"VL_ATENDIMENTO", "Valor do atendimento"},
{"VL_PAGO_RESSARC", "Valor pago ressarcimento"}, {"RAZAO_SOCIAL", "Razao
Social"}, {"MODALIDADE_OPERADORA", "Modalidade da operadora"},
{"SITUACAO_OPERADORA", "Situacao da operadora"}, {"VALOR_COBRADO", "Valor
cobrado"}, {"GRU_EMITIDA", "GRU emitida"}, {"SITUACAO_GRU", "Situacao
GRU"})),
```

```
#"Replaced Value" = Table.ReplaceValue("#Renamed
Columns",5,4,Replacer.ReplaceValue,{"Faixa etaria"}),
```

```

#"Replaced Value1" = Table.ReplaceValue("#Replaced
Value",7,6,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value2" = Table.ReplaceValue("#Replaced
Value1",9,8,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value3" = Table.ReplaceValue("#Replaced
Value2",11,10,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value4" = Table.ReplaceValue("#Replaced
Value3",13,12,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value5" = Table.ReplaceValue("#Replaced
Value4",15,14,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value6" = Table.ReplaceValue("#Replaced
Value5",17,16,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value7" = Table.ReplaceValue("#Replaced
Value6",19,18,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value8" = Table.ReplaceValue("#Replaced
Value7",21,20,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value9" = Table.ReplaceValue("#Replaced
Value8",27,26,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value10" = Table.ReplaceValue("#Replaced
Value9",0,1,Replacer.ReplaceValue,{"Faixa etaria"}),

#"Replaced Value11" = Table.ReplaceValue("#Replaced Value10","pendente de
pagamento (inclui suspensão SCI)","Pendente",Replacer.ReplaceValue,{"Situacao GRU"}),

#"Replaced Value12" = Table.ReplaceValue("#Replaced Value11","", "Não
especificado",Replacer.ReplaceValue,{"Situacao GRU"}),

#"Replaced Value13" = Table.ReplaceValue("#Replaced
Value12","F","Feminino",Replacer.ReplaceValue,{"Sexo do beneficiario"}),

#"Replaced Value14" = Table.ReplaceValue("#Replaced
Value13","M","Masculino",Replacer.ReplaceValue,{"Sexo do beneficiario"}),

#"Changed Type1" = Table.TransformColumnTypes("#Replaced Value14",{{"Tipo do
atendimento", type text}}),

```

```

#"Replaced Value15" = Table.ReplaceValue("#"Changed
Type1","1","AIH",Replacer.ReplaceValue,{"Tipo do atendimento"}),

#"Replaced Value16" = Table.ReplaceValue("#"Replaced
Value15","2","APAC",Replacer.ReplaceValue,{"Tipo do atendimento"}),

#"Changed Type2" = Table.TransformColumnTypes("#"Replaced Value16",{ "Esfera do
estab", type text })),

#"Replaced Value17" = Table.ReplaceValue("#"Changed Type2","0","Não
informado",Replacer.ReplaceValue,{"Esfera do estab"}),

#"Replaced Value18" = Table.ReplaceValue("#"Replaced
Value17","1","Federal",Replacer.ReplaceValue,{"Esfera do estab"}),

#"Replaced Value19" = Table.ReplaceValue("#"Replaced
Value18","2","Estadual",Replacer.ReplaceValue,{"Esfera do estab"}),

#"Replaced Value20" = Table.ReplaceValue("#"Replaced
Value19","3","Municipal",Replacer.ReplaceValue,{"Esfera do estab"}),

#"Replaced Value21" = Table.ReplaceValue("#"Replaced
Value20","4","Privada",Replacer.ReplaceValue,{"Esfera do estab"}),

#"Added Custom" = Table.AddColumn("#"Replaced Value21", "País", each "Brazil"),

#"Replaced Value22" = Table.ReplaceValue("#"Added Custom","suspensao judicialmente
(SCPJ)","Suspensao judicialmente (SCPJ)",Replacer.ReplaceValue,{"Situacao GRU"}),

#"Replaced Value23" = Table.ReplaceValue("#"Replaced
Value22","","#999",Replacer.ReplaceValue,{"COD_CID"})

```

in

```
#"Replaced Value23"
```

- **Código em linguagem M utilizado para o tratamento de dados referente ao item 3.3.2.2 Base de dados: Dados de pagamento do Ressarcimento ao SUS por operadora**

let

```
Source = Csv.Document(File.Contents("C:\Users\berna\Downloads\baseieprefix
(2).csv"),[Delimiter=";", Columns=8, Encoding=65001, QuoteStyle=QuoteStyle.None]),
```

```

#"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),

#"Changed Type" = Table.TransformColumnTypes("#Promoted Headers",{{"",
Int64.Type}, {"COMPETENCIA", type date}, {"CODIGO_OPERADORA", Int64.Type},
{"RAZAO_SOCIAL", type text}, {"SITUACAO_OPERADORA", type text},
{"VALOR_PAGO_ACUMULADO", type number},
{"VALOR_PARCELADO_ACUMULADO", type number},
{"VALOR_COBRADO_ACUMULADO", type number}}),

#"Added Custom1" = Table.AddColumn("#Changed Type", "Valor pago+parcelado", each
[VALOR_PAGO_ACUMULADO]+[VALOR_PARCELADO_ACUMULADO]),

#"Changed Type2" = Table.TransformColumnTypes("#Added Custom1",{{"Valor
pago+parcelado", Currency.Type}, {"VALOR_COBRADO_ACUMULADO",
Currency.Type}, {"VALOR_PARCELADO_ACUMULADO", Currency.Type},
{"VALOR_PAGO_ACUMULADO", Currency.Type}})

```

in

```
#"Changed Type2"
```

- **Código em linguagem M utilizado para o tratamento de dados referente ao item 3.3.2.3 Base de dados: CID-10 SUBCATEGORIAS**

let

```

Source = Csv.Document(File.Contents("C:\Users\berna\Desktop\TCC\bases\Extração
oficial - 31-10-2022\CID10\CID-10-SUBCATEGORIAS.CSV"),[Delimiter=";", Columns=9,
Encoding=1252, QuoteStyle=QuoteStyle.None]),

#"Changed Type" = Table.TransformColumnTypes(Source,{{"Column1", type text},
{"Column2", type text}, {"Column3", type text}, {"Column4", type text}, {"Column5", type
text}, {"Column6", type text}, {"Column7", type text}, {"Column8", type text}, {"Column9",
type text}}),

#"Promoted Headers" = Table.PromoteHeaders("#Changed Type",
[PromoteAllScalars=true]),

```

```
#"Changed Type1" = Table.TransformColumnTypes("#Promoted Headers",{{"SUBCAT",
type text}, {"CLASSIF", type text}, {"RESTRSEXO", type text}, {"CAUSAOBITO", type
text}, {"DESCRICA0", type text}, {"DESCRABREV", type text}, {"REFER", type text},
{"EXCLUIDOS", type text}, {"", type text}}),
```

```
#"Removed Columns1" = Table.RemoveColumns("#Changed Type1",{"CLASSIF", "",
"EXCLUIDOS", "REFER", "CAUSAOBITO", "DESCRABREV", "RESTRSEXO"}),
```

```
NewRows = Table.InsertRows("#Removed Columns1", 0, {[SUBCAT="#999",
DESCRICA0="CID Não Especificada"]}),
```

```
#"Duplicated Column" = Table.DuplicateColumn(NewRows, "SUBCAT", "SUBCAT -
Copy"),
```

```
#"Split Column by Position" = Table.SplitColumn("#Duplicated Column", "SUBCAT -
Copy", Splitter.SplitTextByPositions({0, 3}, false), {"SUBCAT - Copy.1", "SUBCAT -
Copy.2"}),
```

```
#"Changed Type2" = Table.TransformColumnTypes("#Split Column by
Position",{{"SUBCAT - Copy.1", type text}, {"SUBCAT - Copy.2", Int64.Type}}),
```

```
#"Removed Columns" = Table.RemoveColumns("#Changed Type2",{"SUBCAT -
Copy.2"}),
```

```
#"Split Column by Position1" = Table.SplitColumn("#Removed Columns", "SUBCAT -
Copy.1", Splitter.SplitTextByPositions({0, 1}, false), {"SUBCAT - Copy.1.1", "SUBCAT -
Copy.1.2"}),
```

```
#"Renamed Columns" = Table.RenameColumns("#Split Column by
Position1",{{"SUBCAT - Copy.1.1", "Letra categoria"}, {"SUBCAT - Copy.1.2", "Num
categoria"}})
```

in

```
#"Renamed Columns"
```

- **Código em linguagem DAX utilizado para a criação da variável numerocap referente ao item 3.3.2.3 Base de dados: CID-10 SUBCATEGORIAS**

numerocap =

```

if('CID-10-SUBCATEGORIAS'[Letra categoria]="A" || 'CID-10-SUBCATEGORIAS'[Letra
categoria]="B" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "1",
if(('CID-10-SUBCATEGORIAS'[Letra categoria]="C" && 'CID-10-
SUBCATEGORIAS'[Num categoria]>=0 &&'CID-10-SUBCATEGORIAS'[Num catego-
ria]<=99) || ('CID-10-SUBCATEGORIAS'[Letra categoria]="D" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=48), "2",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="D" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=50
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=89, "3",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="E" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=90, "4",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="F" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "5",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="G" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "6",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="H" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=59, "7",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="H" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=60
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=95, "8",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="I" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "9",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="J" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "10",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="K" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=93, "11",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="L" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "12",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="M" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "13",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="N" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "14",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="O" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "15",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="P" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0

```

```

&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=96, "16",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="Q" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "17",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="R" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "18",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="S" || 'CID-10-SUBCATEGORIAS'[Letra
categoria]="T" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "19",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="V" || 'CID-10-SUBCATEGORIAS'[Letra
categoria]="W"|| 'CID-10-SUBCATEGORIAS'[Letra categoria]="X" || 'CID-10-
SUBCATEGORIAS'[Letra categoria]="Y" && 'CID-10-SUBCATEGORIAS'[Num catego-
ria]>=0 && 'CID-10-SUBCATEGORIAS'[Num categoria]<=99,"20",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="Z" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=0
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "21",
if('CID-10-SUBCATEGORIAS'[Letra categoria]="U" &&
'CID-10-SUBCATEGORIAS'[Num categoria]>=4
&& 'CID-10-SUBCATEGORIAS'[Num categoria]<=99, "22", "falso"))))))))))))))))))))

```

- **Código em linguagem M utilizado para o tratamento de dados referente ao item**

3.3.2.4 Base de dados: CID-10 CAPÍTULOS

let

```

Source = Csv.Document(File.Contents("C:\Users\berna\Desktop\TCC\bases\Extração
oficial - 31-10-2022\CID10\CID-10-CAPITULOS.CSV"),[Delimiter=";", Columns=6,
Encoding=1252, QuoteStyle=QuoteStyle.None]),

```

```

#"Promoted Headers" = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),

```

```

#"Changed Type" = Table.TransformColumnTypes("#Promoted Headers",{{"NUMCAP",
Int64.Type}, {"CATINIC", type text}, {"CATFIM", type text}, {"DESCRICA0", type text},
{"DESCRABREV", type text}, {"", type text}}),

```

```

#"Removed Columns" = Table.RemoveColumns("#Changed Type",{""}),

```

```

NewRows2 = Table.InsertRows("#Removed Columns", 0, {[NUMCAP=99,
CATINIC="#0", CATFIM="#9999",DESCRICA0="CID Não
Especificada",DESCRABREV="CID Não Especificada"]})

```

in

```

NewRows2

```

APÊNDICE C – DASHBOARD

Figura 42 - Capa do *dashboard*



Ressarcimento ao SUS por operadoras de planos de saúde

(2016 - setembro 2021)

Fonte: Elaborada pelo autor.

Figura 43 - Sumário do *dashboard*

Glossário

Data do atendimento: data de início do atendimento no SUS por beneficiário de plano de saúde.

Registro Operadora: identificador das operadoras quanto ao seu registro na ANS.

Razão social: razão social das operadoras de plano de saúde.

Modalidade da operadora: classificação das operadoras de planos privados de assistência à saúde.

Situação da operadora: classificação da operadora conforme sua situação cadastral perante à ANS.

Procedimento cobrado: indica se o procedimento foi cobrado via Guia de Receita da União - GRU.

Valor dos atendimentos: valor referente ao atendimento realizado no SUS que a ANS identificou como tendo sido prestado a um beneficiário de plano de saúde e cujo valor é passível de cobrança à respectiva operadora, conforme a regra de valoração estabelecida pela agência reguladora.

Valor cobrado com GRU: valor referente a cobrança, via Guia de Receita da União (GRU), do atendimento identificado como passível de ressarcimento. Situação da GRU: Situação da GRU de cobrança (pago, parcelado, pendente ou suspenso judicialmente).

Índice Efetivo de pagamento Acumulado: indicador de adimplência administrativa do Ressarcimento ao SUS, constituído pela soma dos valores pagos e parcelados administrativamente sobre o total cobrado com Guia de Recolhimento da União, GRU, emitida. Não considera valores pagos judicialmente e trata-se de um índice cumulativo, em que valores pagos em períodos pretéritos refletem nos indicadores futuros.

Valor cobrado: Atendimentos não impugnados ou com valor devido após análise administrativa, notificados por meio de GRU

Valor pago ou parcelado: Atendimentos cobrados e pagos administrativamente à vista ou parcelados através de Requerimento de Parcelamento de Débitos - RPD, deferido.

Fonte: Elaborada pelo autor.

Faixa etária: classificação dos beneficiários quanto à sua idade.

Sexo beneficiário: classificação dos beneficiários quanto ao seu sexo.

Tipo do atendimento: classificação dos procedimentos em relação ao seu tipo de atendimento, AIH ou APAC.

Dias de permanência interna(DPI): Dias de permanência na internação para AIH, intervalo entre datas de início e fim do tratamento para APAC.

UF do beneficiário: classificação do beneficiário quanto a sua Unidade Federativa do Brasil (UF).

Esfera do estab.: classificação do estabelecimento onde ocorreu o atendimento quanto a sua esfera administrativa.

Natureza do estab.: classificação do estabelecimento onde ocorreu o atendimento quanto a sua natureza jurídica.

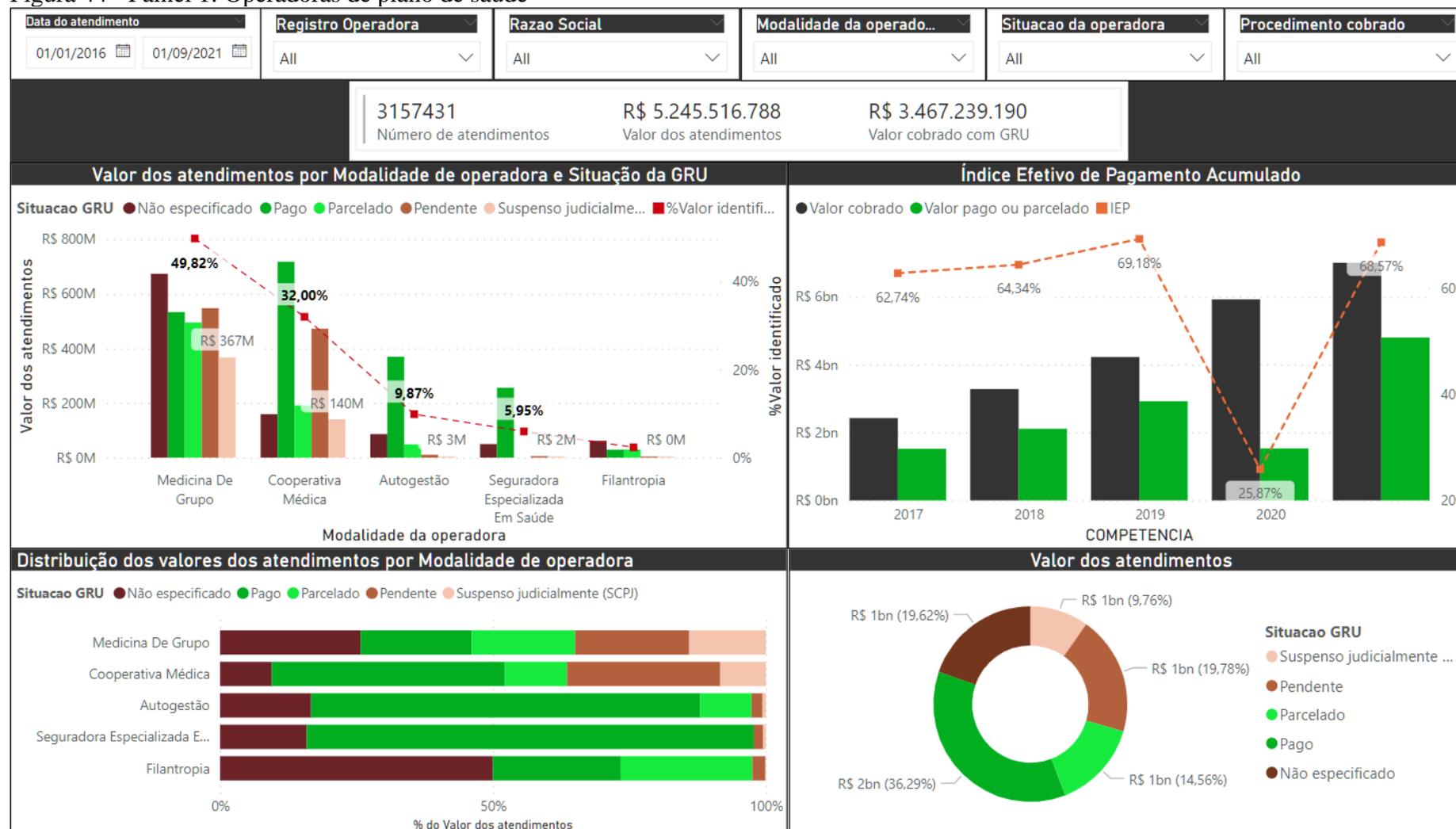
Uf do estab.: classificação do estabelecimento onde ocorreu o atendimento quanto a sua Unidade Federativa do Brasil (UF).

Município do estab.: classificação do estabelecimento onde ocorreu o atendimento quanto ao município que pertence.

Capítulo CID: classificação do atendimento quanto ao capítulo da CID que este pertence.

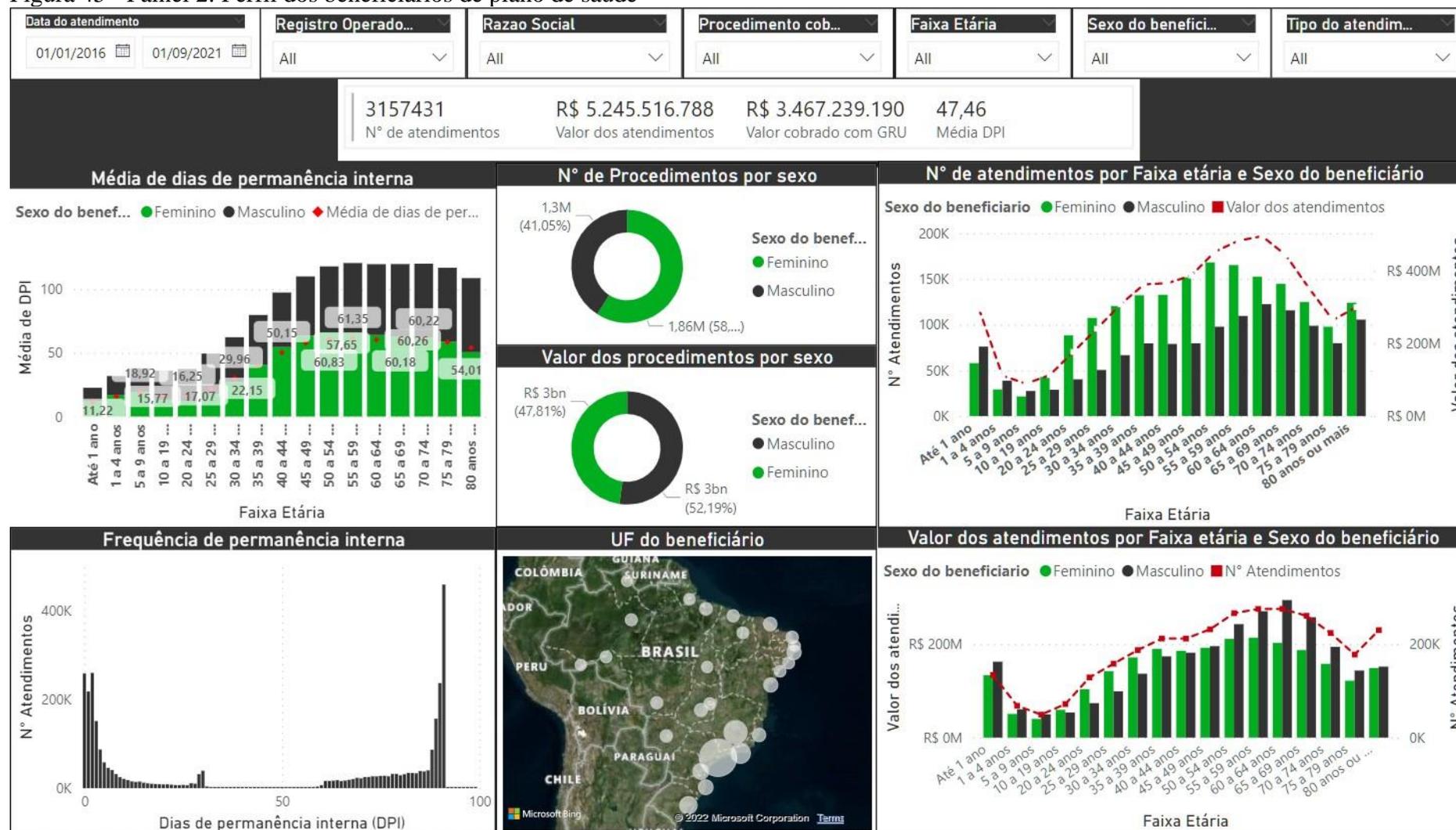
Código CID: classificação do atendimento quanto ao código CID (Classificação Internacional de Doenças e Problemas Relacionados à Saúde)

Figura 44 - Painel 1: Operadoras de plano de saúde



Fonte: Elaborada pelo autor.

Figura 45 - Painel 2: Perfil dos beneficiários de plano de saúde



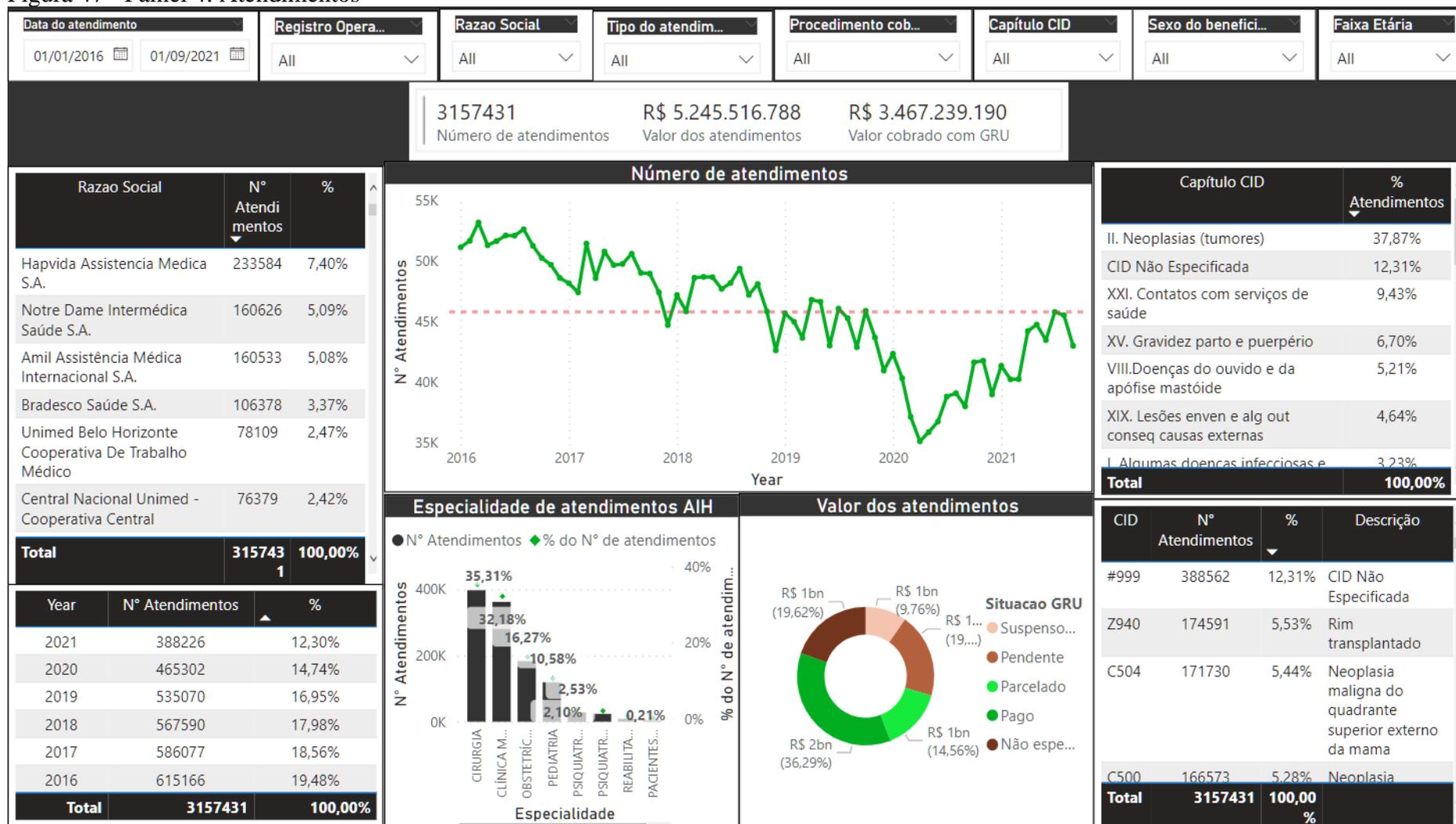
Fonte: Elaborada pelo autor.

Figura 46 - Painel 3: Localização dos atendimentos



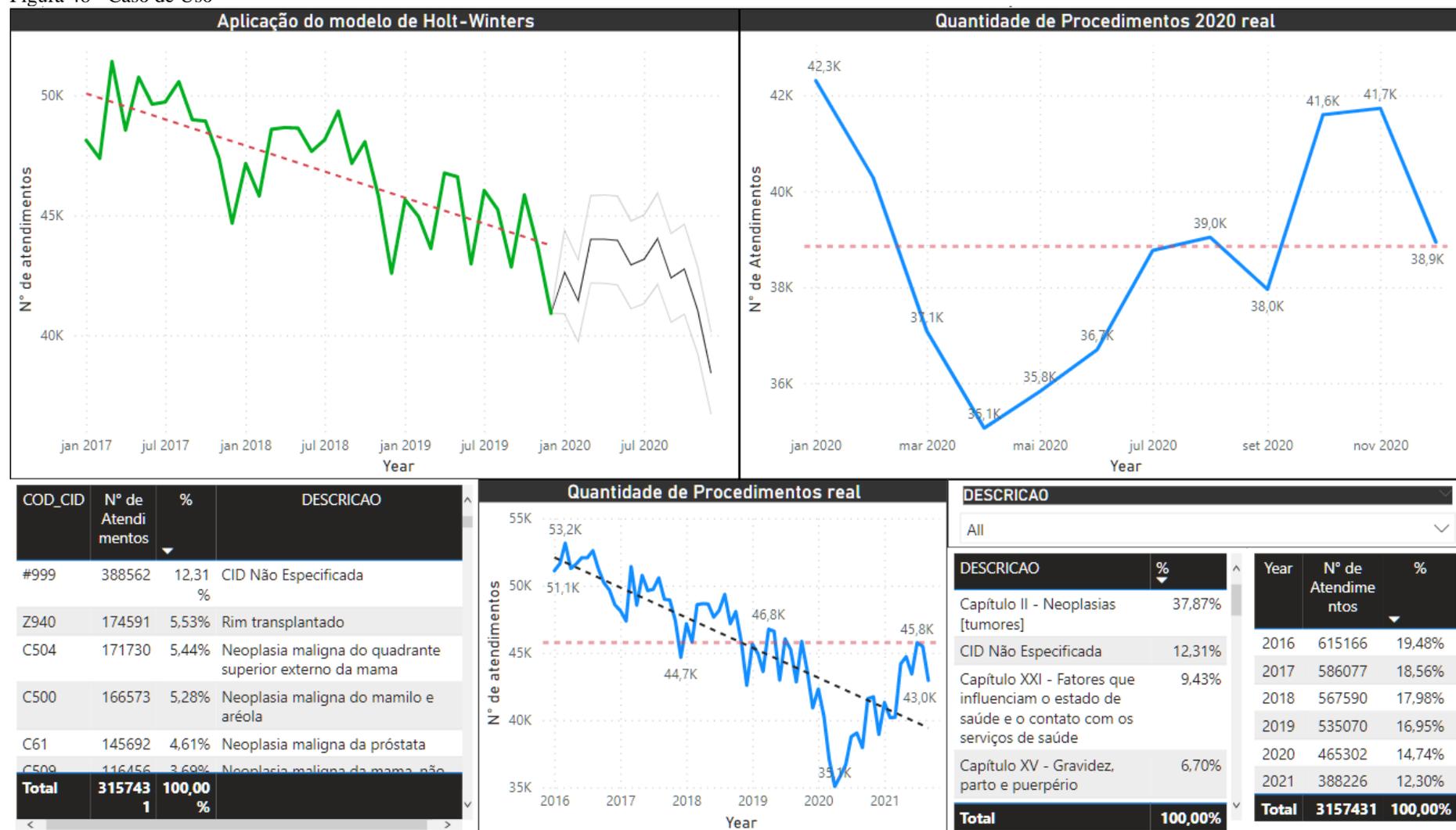
Fonte: Elaborada pelo autor.

Figura 47 - Painel 4: Atendimentos



Fonte: Elaborada pelo autor.

Figura 48 - Caso de Uso



Fonte: Elaborada pelo autor.