

Análise do impacto da pandemia do covid-19 para o ramo do seguro de vida

Projeto Final do curso Pós-graduação BI Master

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

Componente do Projeto:

<Grasiela da Silva Machado –
grasielasmachado@gmail.com >

Histórico de Versões

Data	Versão	Descrição	Autor	Aprovado por
13.05.2023	1.0	Criação do template	Grasiela Machado	Grasiela Machado

Sumário

1	INTRODUÇÃO	4
2	ESTUDO DE CASO	5
2.1	DESCRIÇÃO DO ESTUDO DE CASO	5
3	DESCRIÇÃO DO MODELO TRANSACIONAL.....	6
3.1	FONTE 1 – ARQUIVOS CSV E XLSX.....	6
3.2	FONTE 2 – DIAGRAMA DE CLASSES	6
4	PROPOSTA DE PROCESSO DE BI.....	7
5	MODELO MULTIDIMENSIONAL	8
6	ELABORAÇÃO DO DATA WAREHOUSE.....	9
6.1	DEFINIÇÃO DO DW	9
7	PROJETO DE ETL.....	10
7.1	DESCRIÇÃO DO PROJETO DE ETL	10
8	DASHBOARD	13
8.1	DESCRIÇÃO DA ELABORAÇÃO	13
8.2	TELAS DO DASHBOARD	13
9	CONCLUSÃO	15
10	ARQUIVOS	16
10.1	FONTE DE DADOS	16
10.2	ETL – PDI.....	17
10.3	DW – POSTGRES	18
10.4	RELATÓRIOS – POWER BI.....	18

1 Introdução

Este documento tem por finalidade coletar, analisar e definir as principais necessidades do projeto do estudo de caso ***Análise do impacto da pandemia do covid-19 para o ramo do seguro de vida.*** O documento procura demonstrar os principais problemas atuais e o foco investigativo desejado pelo cliente.

2 Estudo de Caso

2.1 Descrição do Estudo de Caso

O escopo do projeto envolve análise do impacto da pandemia do covid-19 para o ramo de seguro de vida no Brasil através de dados disponibilizados pela SUSEP (órgão fiscalizador e regulador do ramo de seguro), levando em consideração as dimensões de empresas, ramos, região geográfica do Brasil dos últimos seis anos.

Busca-se verificar a evolução do prêmio direto e sinistros de seguros, principalmente do segmento vida no período anterior, durante e logo após o período crítico da pandemia do covid-19.

Ao final do projeto, além da construção do Data Warehouse objetiva-se disponibilizar relatórios com gráficos e alguns indicadores (KPI) que permitam verificar como foi a evolução do prêmio e sinistro no período.

3 Descrição do Modelo Transacional

3.1 Fonte 1 – Arquivos CSV

A fonte de dados utilizada: arquivos no formato csv representados pelo esquema de dados na figura 1.

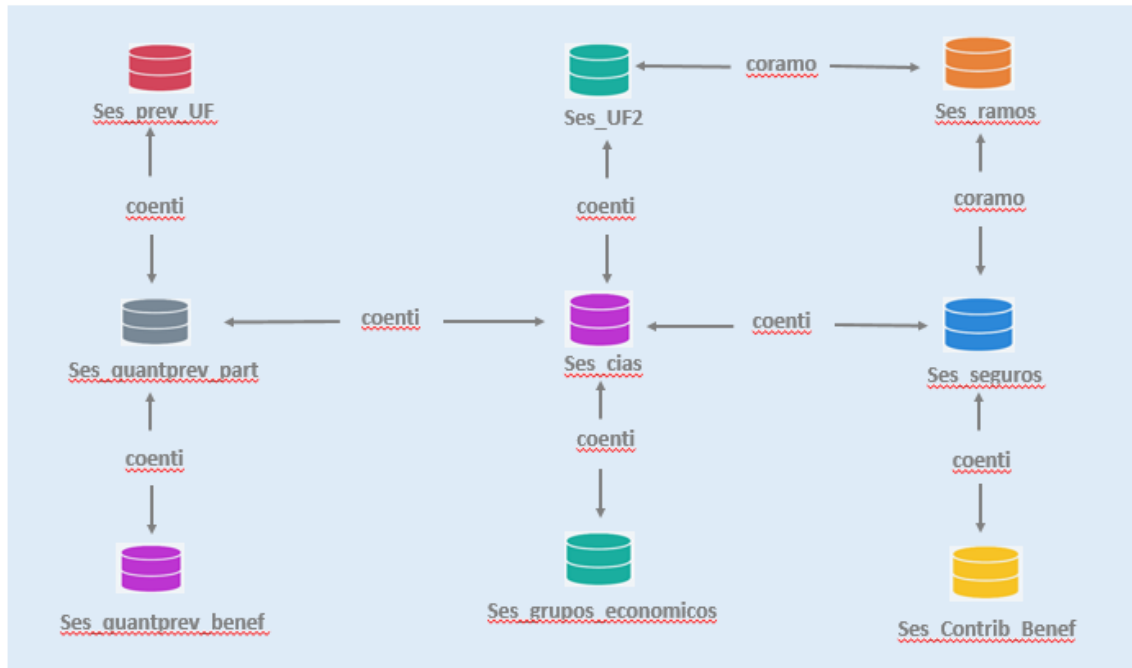
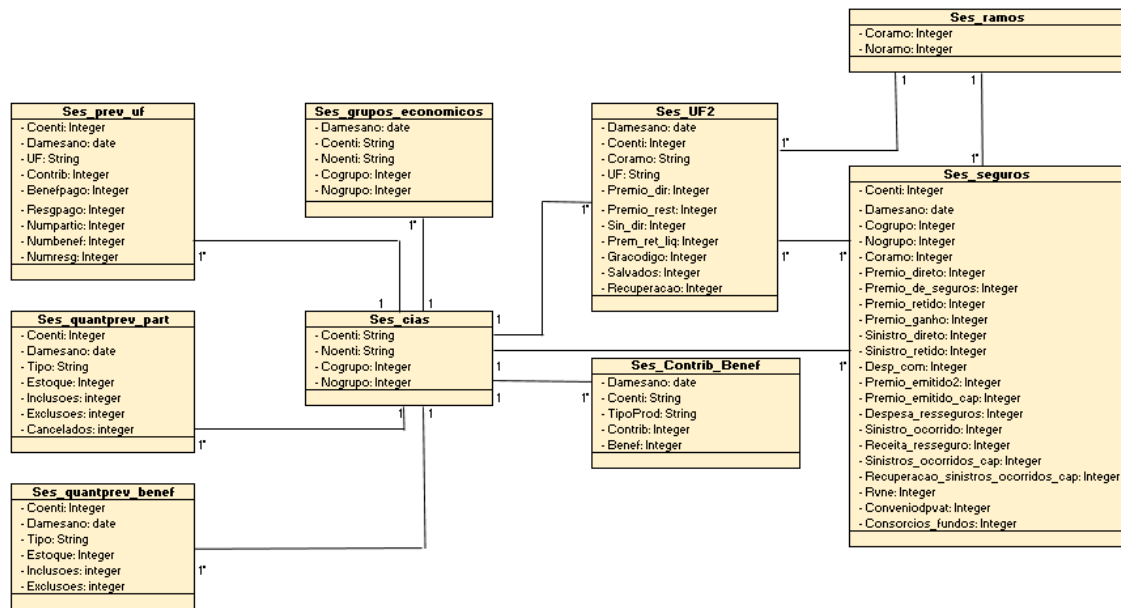


Figura 1

3.2 Fonte 2 – Diagrama de Classes

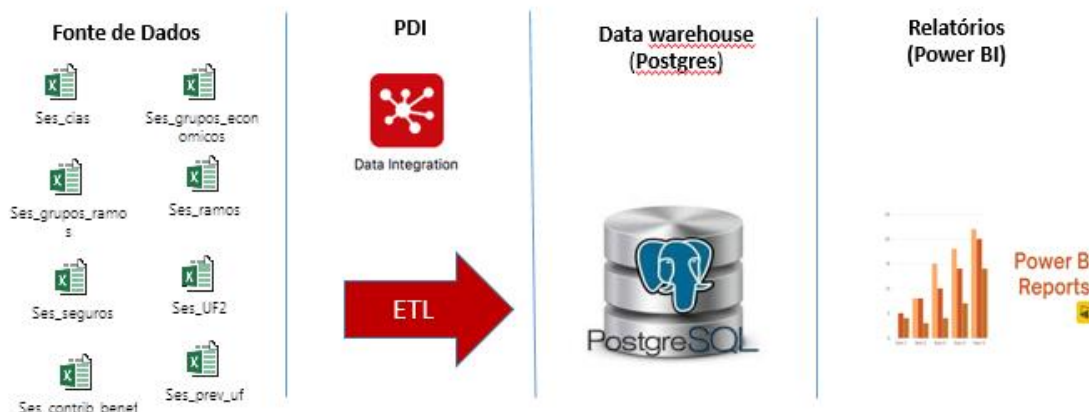
Com base no esquema de dados e dos arquivos disponibilizados pela Susep, foi possível criar um diagrama de classes e para uma melhor visualização e entendimento, está representado na figura abaixo:



4 Proposta de Processo de BI

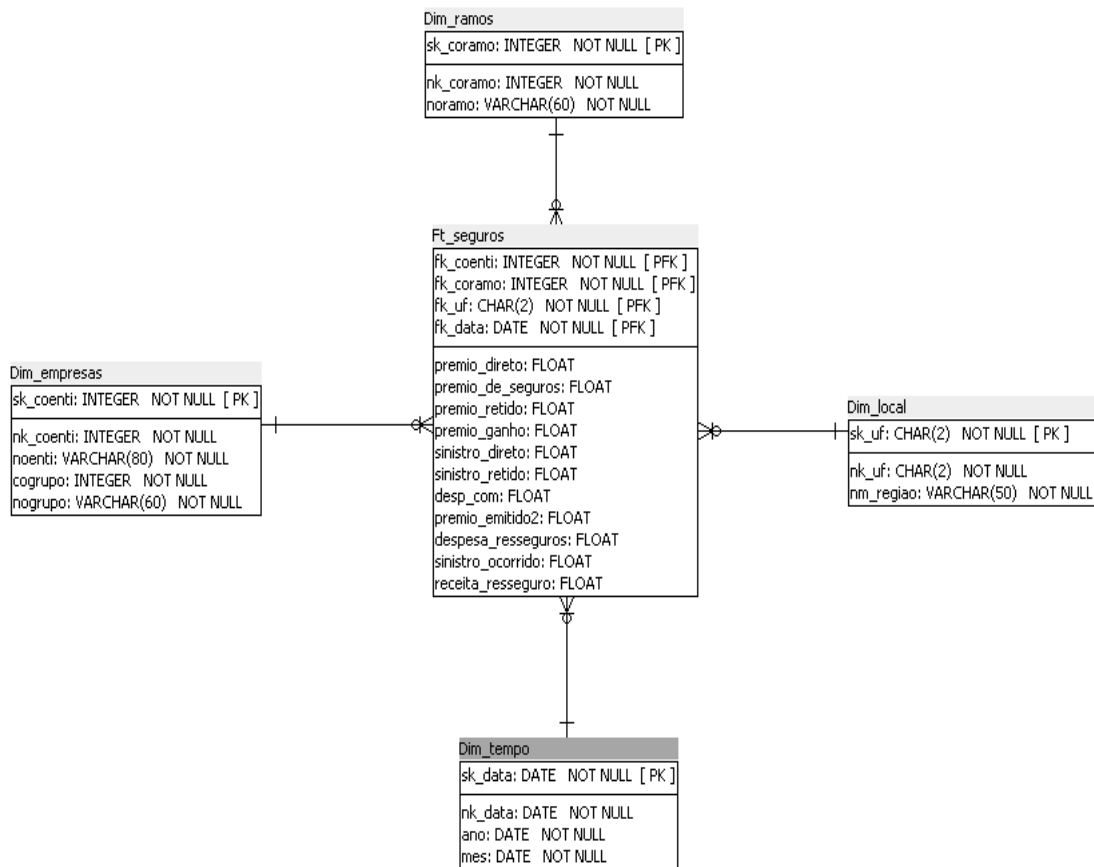
O processo de BI proposto para o projeto tem como fontes de dados os arquivos csv, que através da ferramenta Pentaho Data Integration (PDI) serão extraídos e após as devidas transformações (etapa ETL), serão posteriormente armazenados no *DW Seguros* criado no SGBD Postgres.

Os relatórios serão desenvolvidos através da ferramenta Microsoft Power BI que acessará os dados diretamente do *DW Seguros*.



5 Modelo Multidimensional

Esta seção apresenta o modelo estrela (star schema) do estudo de caso ***Avaliação do impacto da pandemia do covid-19 no ramo de seguro de vida.***



6 Elaboração do Data Warehouse

O Data Warehouse será a fonte integradora de informações, a tecnologia será utilizada com o intuito de servir de base, sendo responsável por armazenar e fornecer os dados que serão utilizados na análise desse estudo.

6.1 Definição do DW

6.1.1 Arquitetura

Independente – foi escolhido esse tipo de arquitetura por possibilitar uma implementação mais rápida.

6.1.2 Abordagem de Construção

Botton-Up – o motivo da escolha dessa abordagem é devido a vantagem de permitir uma implementação mais rápida, e assim permitir acesso mais rápido às informações.

6.1.3 Arquitetura Física

Nuvem – foi escolhida por permitir implementação mais rápida e mais barata sem a necessidade de altos investimentos em tecnologia (hardware e software), mão de obra e infraestrutura (local, segurança, manutenção,etc).

7 Projeto de ETL

7.1 Descrição do Projeto de ETL

Para realizar o processo de ETL foi utilizada a ferramenta PDI, e as cargas realizadas nas tabelas dimensão e fato seguiram o modelo multidimensional apresentado no item 5.

As cargas nas tabelas dimensão seguiram o seguinte processo:

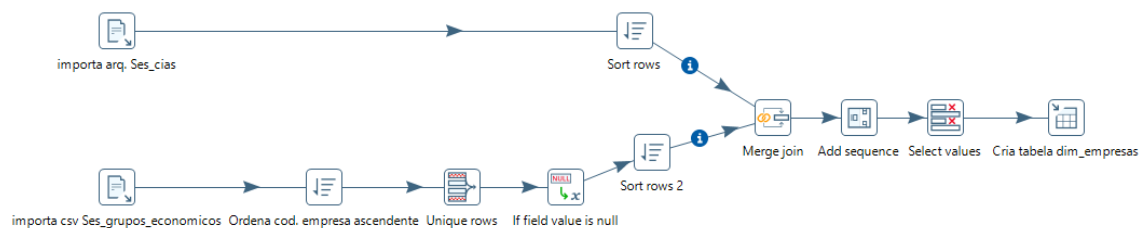
7.1.1. Tabela Dimensão Ramo – os dados foram extraídos do arquivo *Ses_ramos.csv*. Foi feito ajuste no nome do ramo, excluindo-se o código do ramo, para evitar redundância de informação.

Processo de ETL – PDI:



7.1.2. Tabela Dimensão Empresas - os dados foram extraídos do arquivo *Ses_cias.csv* e *Ses_grupos_economicos.csv*. Foi feito join entre os arquivos para inclusão do grupo econômico, com tratamento para exclusão de duplicidade de linhas.

Processo de ETL – PDI:



7.1.3. Tabela Dimensão Local – Essa tabela foi criada no PDI utilizando como referência o campo “UF” do arquivo *Ses_UF2.csv*. Foi feito tratamento para excluir linhas duplicadas, e foi criada a coluna “Região”.

Processo de ETL – PDI:



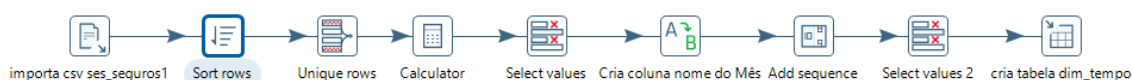
7.1.4. Tabela Dimensão Tempo – Essa tabela foi criada no PDI em duas etapas. Na primeira etapa, foi criado o arquivo *Ses_seguros1.csv* com dados dos últimos seis anos e na segunda etapa foi criada a tabela *dim_tempo* utilizando como referência o campo “damesano” do arquivo *Ses_seguros1.csv*. Foi feito tratamento para exclusão de linhas duplicadas. Foram criadas as colunas “ano”, “mês” e *nome_mes*”.

Processo de ETL – PDI:

Etapa 1 – Cria arquivo *Ses_seguros1.csv*

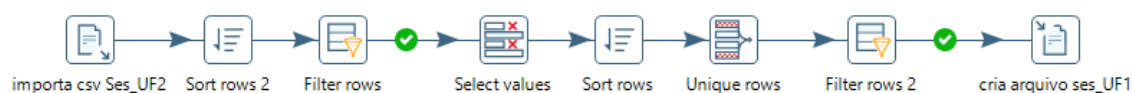


Etapa 2 – Cria tabela Dimensão tempo

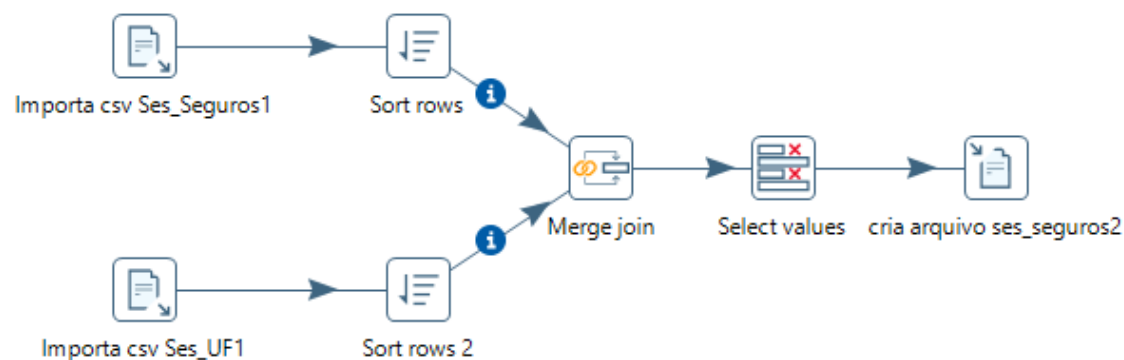


7.1.5. Tabela Fato Seguros – Para criar a tabela Fato foi necessário realizar duas etapas antes:

Etapa 1 – Criação do arquivo *Ses_UF1.csv*, com dados por UF dos últimos seis anos.



Etapa 2 – foi feito join dos arquivos *Ses_seguros1.csv* e *Ses_UF1.csv* para acrescentar a UF e foi criado o arquivo *Ses_seguros2.csv*.



Etapa 3 - Criação da tabela Fato utilizando o arquivo *Ses_seguros2.csv* criado na etapa 1. A tabela contempla os dados do ramo Seguros do período 2017 a 2022.



8 Dashboard

8.1 Descrição da Elaboração

Os dashboards foram criados utilizando como ferramenta o Power BI, acessando os dados diretamente do *DW Seguros*.

8.2 Telas do Dashboard

Foram desenvolvidos dois dashboards: um apresentando um **Panorama Geral** (figura 1) com os principais indicadores do mercado de Seguros e um segundo **Evolução Segmento Vida**, com o foco mais específico no segmento Vida (figura 2).

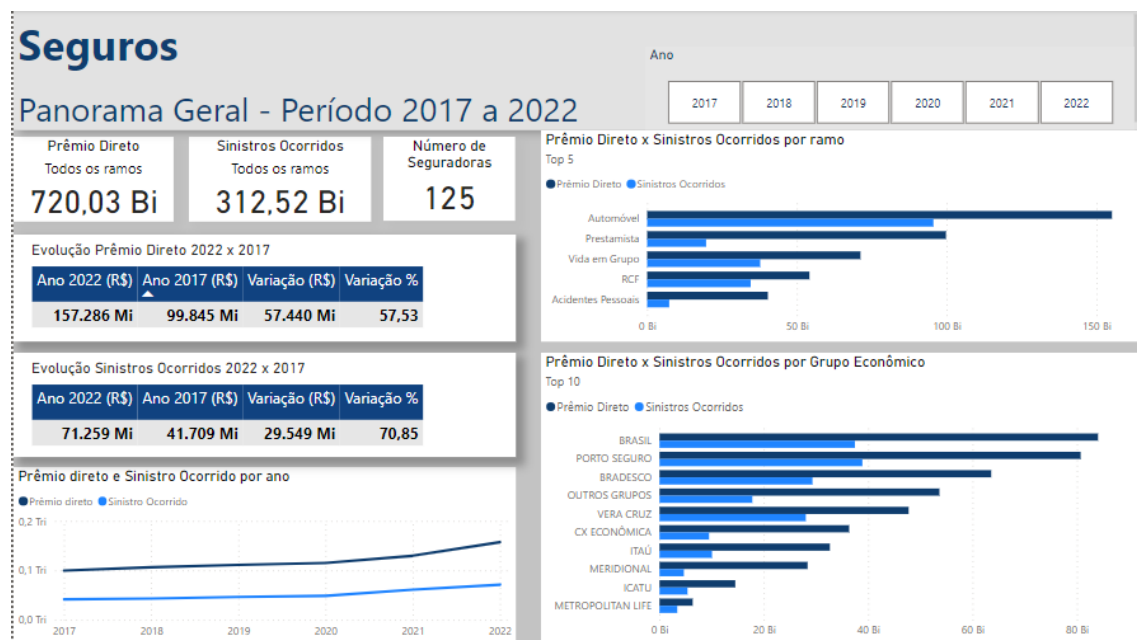


Figura 1 – Panorama Geral

Seguros

Evolução Segmento Vida - Período 2017 a 2022

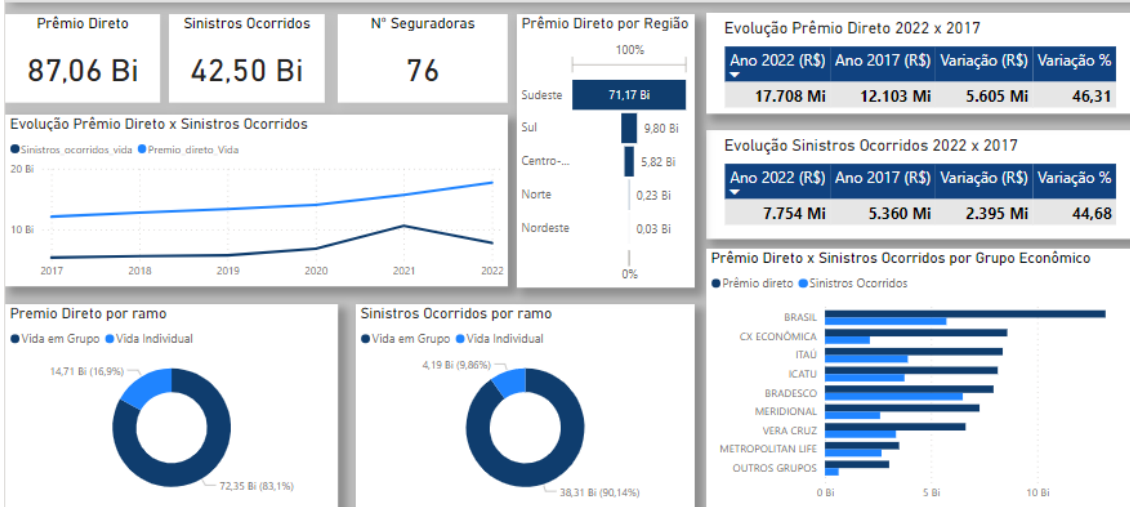


Figura 2 – Evolução Segmento Vida

9 Conclusão

O objetivo desse projeto foi através dos dados disponibilizados pela Susep (superintendência de Seguros Privados), e utilizando o processo de BI, verificar se realmente houve um impacto significativo no mercado de Seguros, mais especificamente no segmento de seguro de Vida em decorrência da pandemia da Covid-19, conforme noticiado por diversas fontes de comunicação.

Com a criação dos dashboards, foi possível realizar diversas análises, tais como:

- Verificar que o mercado de seguros no Brasil como um todo, no período de 2017 a 2022, teve um crescimento bem expressivo nos prêmios diretos, na ordem de 57% como também nos sinistros ocorridos, que foi na ordem de 70,8%. Mas o saldo (prêmio direto – sinistros ocorridos) foi bem positivo em termos de valores, R\$ 27.89 Milhões.

- Analisando especificamente o segmento de vida é possível verificar a mesma tendência de crescimento, com um aumento no prêmio direto de 46,31% e os sinistros ocorridos em 44,68%. E da mesma forma, em termos financeiros, o saldo foi bem positivo. O prêmio direto cresceu R\$ 5,60 Milhões, enquanto o sinistro ocorrido cresceu R\$ 2,40 Milhões, deixando um saldo positivo de R\$ 3,20 Milhões.

- Pelo gráfico, é possível identificar que houve um pico de pagamentos de sinistros no ano de 2021, provavelmente devido a pandemia, mas em 2022 volta para o patamar de antes da pandemia, mesmo com o aumento significativo dos prêmios diretos.

- Pode-se verificar também, que a venda de seguro de vida ainda está muito concentrado na região Sudeste.

- E, também pode-se verificar uma tendência de crescimento nos prêmios diretos em 2022, podendo ser uma sinalização de que esse comportamento com relação ao seguro de vida não só permaneça, como cresça daqui por diante.

Essas foram algumas das análises realizadas, com apenas dois dashboards criados. O que demonstra o quão valioso pode ser a realização de um processo de BI, organizando os dados, facilitando a obtenção de informações relevantes para qualquer tipo de análise, em qualquer tipo de negócio, inclusive um estudo de caso como esse.

10 Arquivos

10.1 Fonte de dados



Ses_ramos.csv

- ↳ Dados dos ramos de seguros



Ses_cias.csv

- ↳ Dados das empresas do mercado Segurador no Brasil



Ses_grupos_economicos.csv

- ↳ Dados dos grupos econômicos no mercado Segurador no Brasil



Ses_seguros.csv

- ↳ Dados históricos de todos os tipos de prêmio, sinistros, despesas e receitas de seguro, resseguro, capitalização do mercado brasileiro



ses_seguros1.csv

- ↳ Dados dos últimos seis anos de todos os tipos de prêmio, sinistros, despesas e receitas de seguro, resseguro, capitalização do mercado brasileiro



ses_seguros2.csv

- ↳ Dados dos últimos seis anos de todos os tipos de prêmio, sinistros, despesas e receitas de seguro, resseguro, capitalização do mercado brasileiro por UF



SES_UF2.csv

- ↳ Dados históricos de todos os tipos de prêmio, sinistros, despesas e receitas de seguro, resseguro, capitalização do mercado brasileiro por UF

10.2 ETL – PDI



etl01-cria-dim-ramo
s.ktr

⇒ cria a tabela Dimensão Ramos



etl02-cria-dim-empr
esas.ktr

⇒ cria tabela Dimensão Empresa



etl03-cria-dim-local.
ktr

⇒ cria tabela Dimensão Local



etl04-cria-dim-temp
o.ktr

⇒ cria tabela Dimensão Tempo



etl05-cria-arquivo_s
es_seguros1.ktr

⇒ cria arquivo de dados de seguro dos últimos seis anos



etl06-cria-arquivo_s
es_UF1.ktr

⇒ cria arquivo de dados de seguro dos últimos seis anos por UF



etl07-Cria-arquivo_s
es_seguros2.ktr

⇒ cria arquivo único de dados de seguros dos últimos seis anos por UF



etl08-Cria-fato_seg
uros.ktr

⇒ cria tabela Fato de Seguros

10.3 DW – Postgres



dw_seguros.sql

10.4 Relatórios – Power BI



Seguros.pbix

⇒ Dashboards criados: Panorama Geral e Evolução do Segmento Vida