

## 2. Análise Exploratória com Dados da SUSEP

O **setor de seguros** desempenha um papel essencial na economia, fornecendo proteção financeira contra riscos e incertezas. A análise dos prêmios arrecadados e dos sinistros pagos é crucial para compreender a dinâmica do mercado, avaliar a sustentabilidade das operações das seguradoras e identificar tendências que influenciam a gestão de riscos e a formulação de estratégias.

Este relatório utiliza dados reais da [Superintendência de Seguros Privados \(SUSEP\)](#) para examinar o desempenho do setor segurador no Brasil, com foco na evolução de prêmios e sinistros em diferentes tipos de seguros e unidades federativas no intervalo de dez anos (2014-2024). Além das comparações entre os diversos produtos de seguros, o estudo inclui análises avançadas, como projeções para os próximos meses.

Os resultados são apresentados por meio de gráficos e tabelas que ilustram as tendências observadas, fornecendo uma visão detalhada dos desafios e oportunidades do setor. Com essas informações, o relatório oferece insights valiosos para tomadores de decisão, reguladores e demais agentes envolvidos no mercado de seguros, ajudando a orientar futuras ações e estratégias no setor.

Dados utilizados na análise:

SES_UF2.csv	Seguros: Prêmios e Sinistros ( UF )
damesano	Ano e mês da informação
ramos	Codigo do Ramo
UF	Unidade Federativa
premio_dir	Premios Diretos
sin_dir	Sinistros Diretos
Ses_ramos.csv	Ramos de seguros
coramo	Código do Grupo/Ramo no FIP
noramo	Nome do Ramo

### 2.1 Análise Temporal

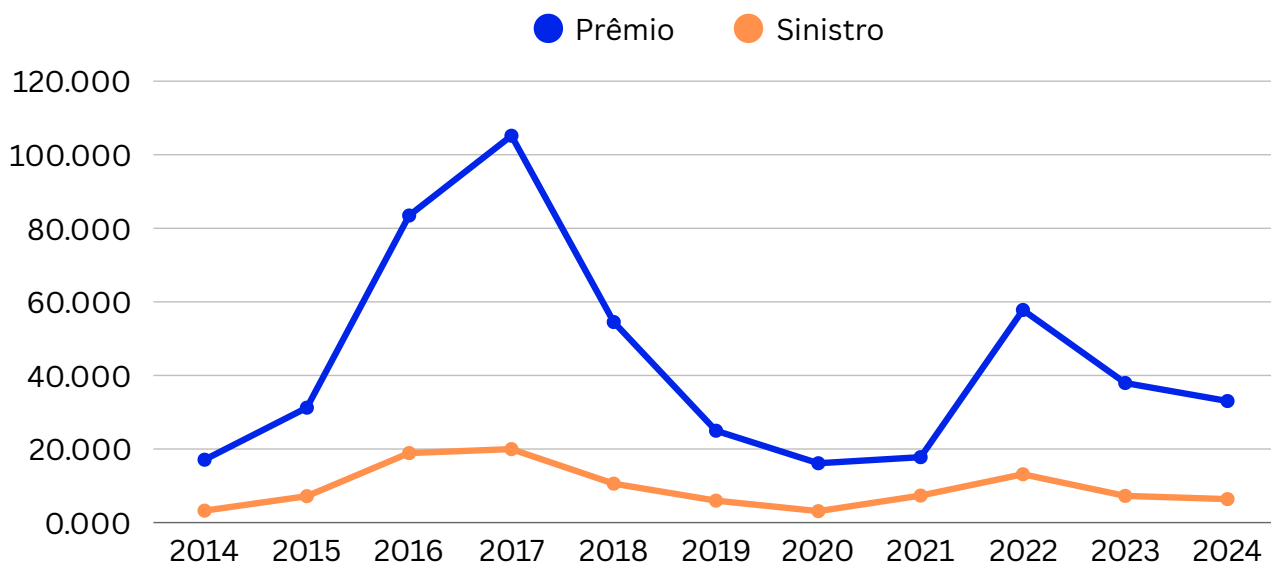
#### 2.1.1 Prêmios e Sinistros

De acordo com a SUSEP (Superintendência de Seguros Privados), o mercado de seguros brasileiro apresentou um crescimento no valor dos prêmios arrecadados em 2017 em comparação com 2016. Esse aumento foi observado em diversos segmentos, especialmente nos **seguros de vida e automóveis**, impulsionados pela retomada do crédito e pelo aumento da demanda por cobertura. Em 2017, o mercado de seguros gerou R\$ 105 bilhões em prêmios, um crescimento de 26,5% em relação aos R\$ 83 bilhões registrados em 2016. Esse aumento não foi causado apenas pela **inflação** no custo dos prêmios, mas também pelo **crescimento da base de segurados e pela melhora do cenário econômico**.

Já em 2022, a recuperação econômica pós-pandemia fortaleceu a confiança de consumidores e empresas, resultando em um aumento na contratação de seguros para proteção financeira e segurança. Além disso, o país enfrentou desafios climáticos, como secas e chuvas intensas, que elevaram a demanda por seguros, especialmente nos segmentos rural e automotivo, devido ao crescimento dos sinistros.

## Evolução dos Prêmios arrecadados e Sinistros pagos (2014-2024)

(em 1 milhão)

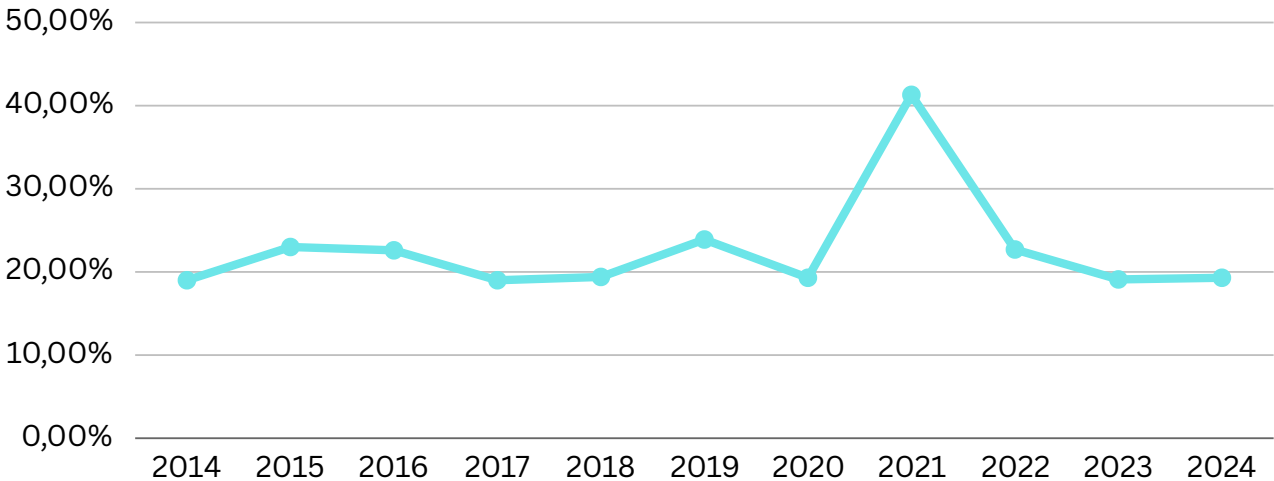


### 2.1.2 Sinistralidade

Em 2021, o Brasil ainda enfrentava as consequências da crise sanitária, o que levou a um aumento nas demandas por seguros de saúde, especialmente devido ao prolongamento da pandemia e ao crescimento das despesas médicas. Além disso, o impacto econômico resultante das medidas restritivas e da instabilidade financeira também fez com que mais pessoas e empresas buscassem proteção adicional, refletindo em um aumento no volume de sinistros, especialmente no seguro de vida e saúde.

Outro fator importante foi o agravamento das condições climáticas no Brasil em 2021, com secas prolongadas e chuvas excessivas em várias regiões do país. Esses eventos climáticos aumentaram a frequência de sinistros em setores como o seguro rural e o seguro de automóveis, devido a danos causados por desastres naturais, como enchentes, deslizamentos de terra e perdas agrícolas. O aumento nos sinistros associados a esses fatores contribuiu para o pico da sinistralidade no setor de seguros, impactando diretamente o valor pago pelas seguradoras e elevando os custos operacionais do mercado.

## Evolução da Sinistralidade (2014-2024)



## 2.2 Comparação entre Tipos de Seguros

### 2.2.2 Prêmios por produto

O seguro VGBL/VAGP/VRGP/VRSA/VRID/VDR (código 1392 e 0994) tem caráter de longo prazo e maior complexidade, o que geralmente reflete um maior valor de prêmio. Esse tipo de seguro é mais voltado ao planejamento financeiro, pensões e previdência, sendo geralmente escolhido por pessoas com uma maior capacidade financeira, ou como parte de uma estratégia de investimento.

O seguro de Automóvel - Casco (código 0531) reflete a alta demanda por proteção de veículos. Esse tipo de seguro é muito comum no Brasil, especialmente considerando o grande número de veículos em circulação e a busca por cobertura contra danos causados a veículos próprios em acidentes, roubos ou furtos. Em termos de volume de prêmios, é um dos produtos mais relevantes, visto que a frota de veículos no país é uma das maiores do mundo.

Esses dados refletem as diferentes necessidades do mercado de seguros brasileiro, onde a maior parte da arrecadação está em produtos voltados para proteção patrimonial, como o seguro de automóveis e vida, mas também há uma presença significativa de seguros que visam segurança financeira e planejamento de longo prazo, como o VGBL. O DPVAT (Danos Pessoais Causados por Veículos Automotores de Via Terrestre) e os riscos nomeados e operacionais, por sua vez, indicam a prevalência de produtos obrigatórios ou de nicho, com prêmios mais baixos e com menos flexibilidade.

## Prêmios arrecadados por Produto (2014-2024)

(em 1 milhão)



### 2.2.2 Sinistros por produto

Ao analisar os sinistros pagos por diferentes tipos de seguros, observa-se uma clara variação no valor dos sinistros 0531 e o restante:

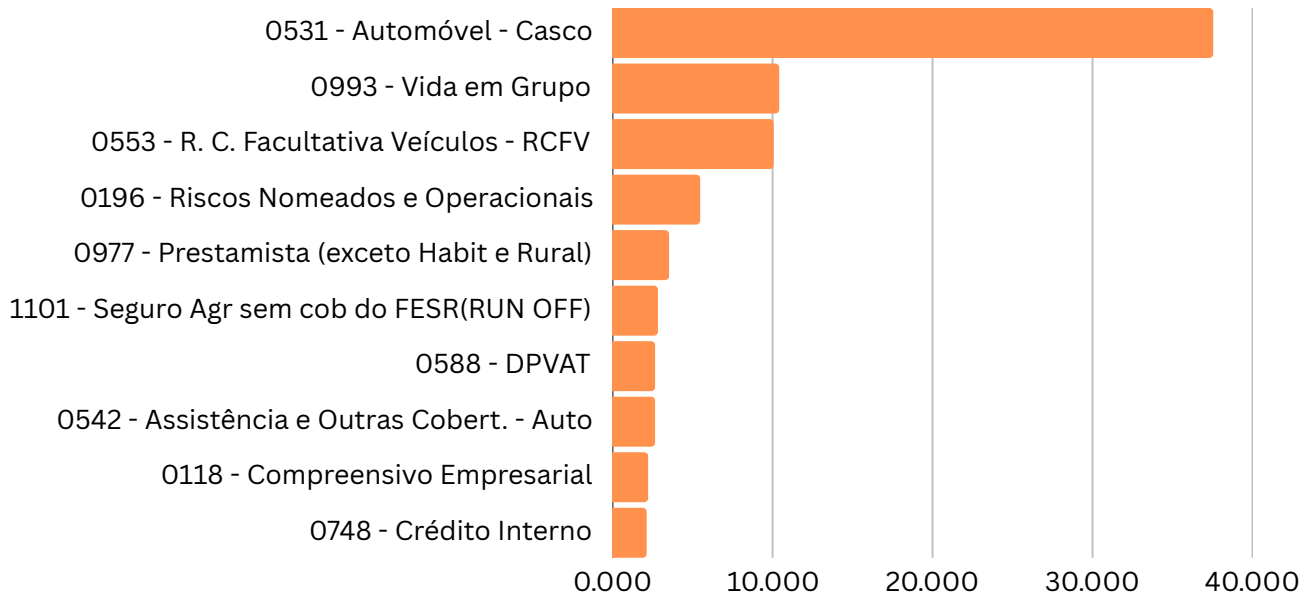
O seguro Automóvel - Casco (código 0531), com o maior valor de sinistros pagos, atingindo R\$ 37.546, reflete a alta frequência de ocorrências de **acidentes de trânsito**, que geram um número significativo de sinistros. Esse tipo de seguro, que cobre danos ao veículo, é diretamente impactado por fatores como aumento da frota de veículos e a sinistralidade resultante de acidentes, roubos ou furtos.

E o seguro Vida em Grupo (código 0993), com sinistros pagos no valor de R\$ 10.416, também apresenta um valor considerável, considerando que cobre principalmente **mortes e invalidez de trabalhadores, geralmente associados a grandes grupos corporativos**. Esse seguro tende a ter um número elevado de sinistros pagos, uma vez que é um dos seguros mais contratados pelas empresas para seus funcionários.

O código 0553 refere-se ao ramo de seguro "Responsabilidade Civil Facultativa de Veículos (RCFV)". Este seguro oferece cobertura para **danos materiais e corporais causados a terceiros em acidentes envolvendo o veículo segurado**, sendo uma proteção adicional ao seguro obrigatório (DPVAT).

## Sinistros pagos por Produto (2014-2024)

(em 1 milhão)



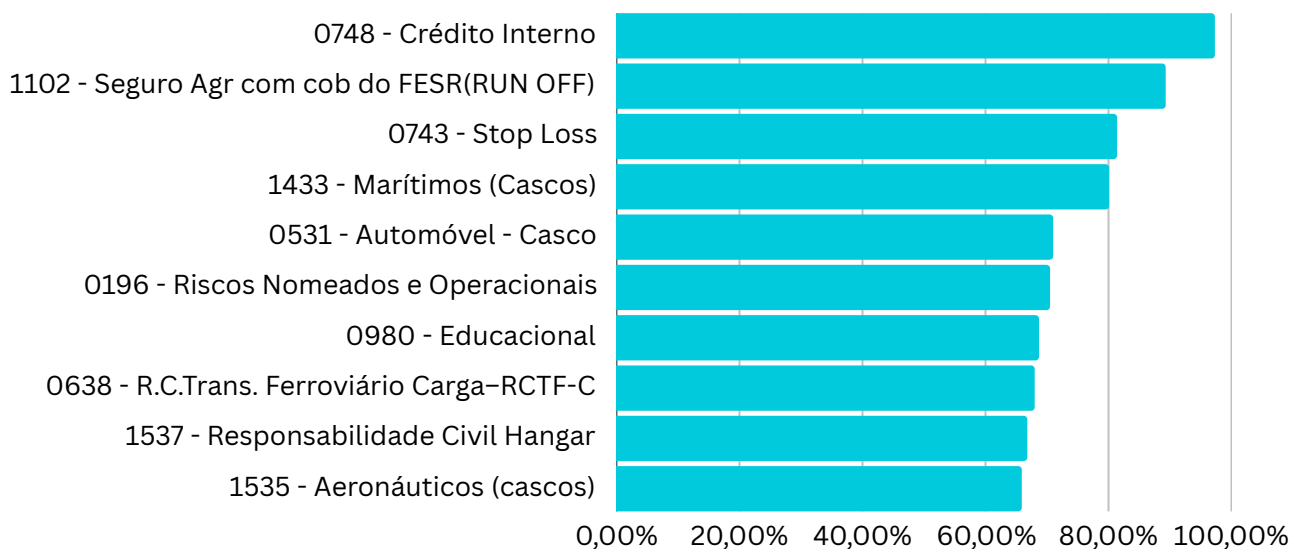
### 2.2.3 Sinistralidade por produto

O ramo de **Crédito Interno** refere-se a seguros que protegem empresas contra inadimplência de clientes em operações de crédito realizadas dentro do país. A sinistralidade elevada (97,3%) indica que as seguradoras estão pagando quase o mesmo valor que arrecadam em prêmios, o que pode ser reflexo de um ambiente econômico adverso, com alta inadimplência., seguido por **Seguro Agrícola com cobertura do FESR (RUN OFF)** que oferece proteção contra perdas na produção agrícola devido a eventos climáticos extremos, como secas e chuvas excessivas., e **"Stop Loss"** com 81,4%, é um tipo de proteção utilizada por empresas e seguradoras para limitar perdas em carteiras de risco, funcionando como uma espécie de "seguro para seguros".

Esses valores indicam que a quantidade de sinistros pagos nesses segmentos é consideravelmente próxima ou até superior aos prêmios arrecadados, o que pode sugerir uma gestão de risco mais desafiadora ou o impacto de eventos adversos específicos.

## Sinistralidade por Produto (2014-2024)

(em %)



### 2.3 Análise por Unidade Federativa

A análise por Unidade Federativa (UF) revela variações significativas entre os estados, refletindo as características regionais, os níveis de cobertura e os riscos associados a cada localidade. A sinistralidade é calculada pela relação entre os sinistros pagos (sin\_dir) e os prêmios diretos arrecadados (premio\_dir), sendo expressa em porcentagem.

O estado com **maior sinistralidade** é **Rondônia (RO)**, com 30,95%, seguido por **Sergipe (SE)** (30,81%) e **Mato Grosso (MT)** (27,03%). Esses números indicam que nessas regiões há um maior comprometimento dos prêmios arrecadados com o pagamento de sinistros, podendo estar relacionado a fatores como maior exposição a riscos, frequências de eventos adversos ou especificidades do mercado local.

Em termos de volume absoluto, **São Paulo (SP)** se destaca com R\$ 201,3 bilhões em prêmios arrecadados e uma sinistralidade de 21,36%, o que indica um mercado amplo e diversificado, capaz de absorver grandes volumes de sinistros sem comprometer sua estabilidade financeira. Já estados com prêmios significativamente menores, como **Roraima (RR)** e **Tocantins (TO)**, também possuem sinistralidades na faixa de 23%, mas com um impacto proporcionalmente maior devido ao menor volume de arrecadação.

A sinistralidade média entre os estados gira em torno de **21,8%**, com algumas exceções acima ou abaixo desse intervalo. Esse comportamento reflete as diferenças regionais no consumo de seguros, no perfil dos segurados e na frequência de eventos que resultam em sinistros. A tendência é que estados com maior atividade econômica e uma maior diversificação de produtos apresentem sinistralidades mais estáveis, enquanto regiões com menor volume de prêmios podem ter variações mais significativas.

UF	Prêmio (em 1 milhão)	Sinistro (em 1 milhão)	Sinistralidade
RO	1.563	487	31,1%
SE	1.773	551	31,1%
MT	6.231	1.697	27,2%
PI	1.668	451	27,0%
RS	32.435	8.627	26,6%

## 2.4.1. Regressão Linear

### 2.4.1. Regressão Linear

A regressão realizada busca estimar a relação entre a variável dependente (sinistros diretos pagos) com as variáveis explicativas (prêmios diretos arrecadados, tempo).

Os resultados indicam que para cada aumento de R\$ 1 milhão no prêmio arrecadado, os sinistros aumentam, em média, R\$ 29.850. Esse coeficiente sugere uma relação positiva entre prêmio arrecadado e sinistros pagos, o que faz sentido, já que um maior volume de arrecadação geralmente reflete um mercado maior, com mais apólices e, conseqüentemente, mais sinistros. Já o coeficiente tempo indica que, ao longo do tempo, há um crescimento médio de R\$ 33,40 mil nos sinistros diretos por mês. Esse valor parece pequeno, mas pode indicar uma tendência crescente ao longo dos anos.

### Limitações

Pelo  $R^2$  (0.005197), o modelo explica apenas cerca de 0,52% da variação no valor dos sinistros diretos. Isso significa que há muitos outros fatores influenciando os sinistros que não estão incluídos no modelo.

Para tornar a análise mais robusta, pode-se recomendar a inclusão de novas variáveis explicativas, como número de apólices ativas, severidade dos sinistros, eventos climáticos extremos e mudanças regulatórias. Além disso, testar outros modelos estatísticos, como modelos não lineares ou séries temporais avançadas, pode ajudar a melhorar as previsões.

### Por que usar Regressão Linear?

Este modelo serve como uma boa abordagem exploratória para entender relações básicas entre variáveis. Ele ajuda a testar se o crescimento nos prêmios arrecadados realmente impacta os sinistros e se há uma tendência temporal relevante.

#### Código em R

```
library(readr) # Carrega o pacote para ler arquivos CSV
library(dplyr) # Carrega o pacote para manipulação de dados
library(ggplot2) # Carrega o pacote para visualização de dados
# Lê os dados da base CSV
dados <- read_csv2("BaseCompleta/SES_UF2.csv")
# Exibe a estrutura do dataframe 'dados'
str(dados)
# Verifica a quantidade de valores ausentes (NA) em cada coluna do dataframe
colSums(is.na(dados))
# Cria um modelo de regressão linear (lm)
modelo <- lm(sin_dir ~ premio_dir + as.numeric(damesano), data = dados)
# Exibe o resumo do modelo de regressão
summary(modelo)
```

### 2.4.2. Projeção de Série Temporal: ETS (Error, Trend, Seasonality)

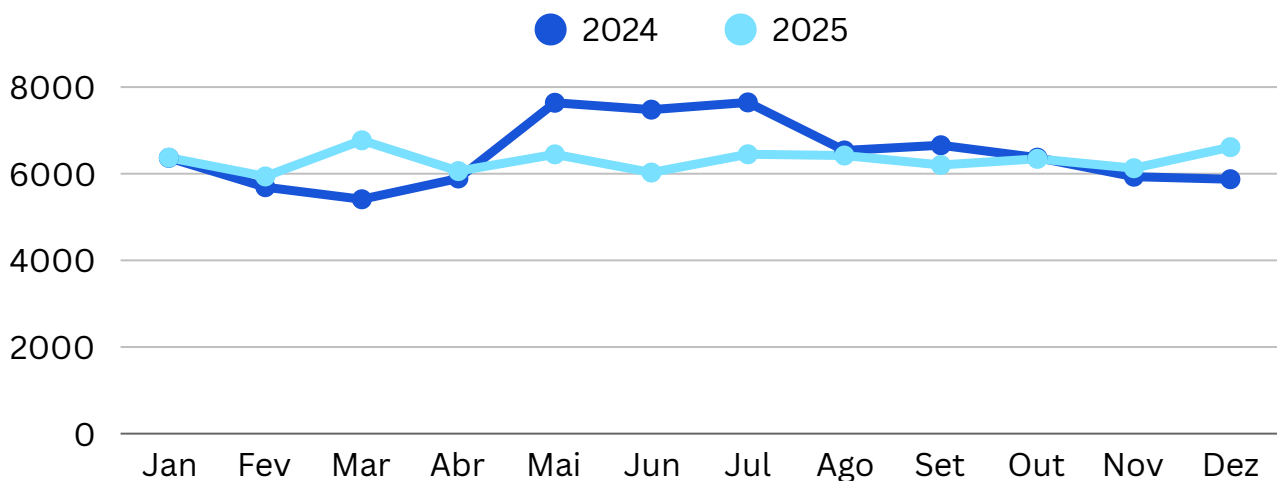
Os valores projetados a seguir refletem a tendência esperada para este ano de 2025, levando em consideração as flutuações sazonais e o comportamento histórico dos sinistros. Em comparação com o ano anterior, observa-se uma certa estabilidade, com uma leve variação ao longo do ano. A projeção pode servir como base para planejar ações de mitigação de riscos e otimizar o gerenciamento de recursos ao longo do ano de 2025.

A maior previsão ocorre em março (6.771,17), enquanto o menor valor projetado é para o mês de fevereiro (5.937,26).

O modelo de projeção ETS é uma solução simples e eficiente para previsões de curto prazo, especialmente quando há padrões bem definidos de tendência e sazonalidade.

#### Sinistro Pago em 2024 x Sinistro Projetado 2025

(em 1 milhão)



#### Código em R

```
library(forecast) # Carrega a biblioteca para modelagem de séries temporais
library(tidyverse) # Carrega o pacote para manipulação de dados
library(ggplot2) # Carrega o pacote para visualização de dados
library(scales) # Carrega o pacote para ajustes de escalas nos gráficos
# Lê os dados da base CSV
dados <- read_csv2("BaseCompleta/SES_UF2.csv")
# Agrupa os dados por dmesano e soma os sinistros
dados_ts <- dados %>%
  group_by(dmesano) %>%
  summarise(sin_dir = sum(sin_dir, na.rm = TRUE)) %>%
  arrange(dmesano)
# Cria uma série temporal dos sinistros mensais, convertendo os valores para
milhões
de reais
ts_sin <- ts(dados_ts$sin_dir / 1000000, start = c(2014, 1), frequency = 12)
```



### Código em R (continuação)

```
# Cria um vetor com os nomes dos meses
meses <- c("Jan", "Fev", "Mar", "Abr", "Mai", "Jun", "Jul", "Ago", "Set", "Out", "Nov", "Dez")
# Extrai o ano e o mês da variável damesano
dados_ts <- dados_ts %>%
  mutate(
    Ano = as.integer(substr(damesano, 1, 4)), # Extrai os primeiros 4 dígitos como ano
    Mes = as.integer(substr(damesano, 5, 6)) # Extrai os últimos 2 dígitos como mês
  )
# Filtra os dados do ano de 2024 e ajusta o formato do mês
dados_2024 <- dados_ts %>%
  filter(Ano == 2024) %>%
  mutate(Ano = "2024", Mes = factor(meses[Mes], levels = meses),
    sin_dir = sin_dir / 1000000) # Convertendo para milhões
# Cria um modelo de previsão usando o método ETS
modelo_ets <- ets(ts_sin)
previsao_ets <- forecast(modelo_ets, h = 12) # Gera previsão para os próximos 12 meses
previsao <- previsao_ets$mean # Obtém os valores previstos para 2025
# Cria um dataframe com os valores previstos para 2025
dados_2025 <- data.frame(
  Mes = meses, # Atribui os meses ao dataframe
  Sinistro = as.numeric(previsao), # Converte os valores previstos para numérico
  Ano = "2025"
)
# Junta os dados reais de 2024 com os dados projetados de 2025
dados_plot <- bind_rows(
  dados_2024 %>% select(Mes, Sinistro = sin_dir, Ano), # Seleciona e renomeia os dados
  de 2024
  dados_2025 # Adiciona os dados previstos de 2025
)
# Cria o gráfico de linhas comparando os sinistros reais de 2024 com os projetados
de 2025
ggplot(dados_plot, aes(x = Mes, y = Sinistro, group = Ano, color = Ano)) +
  geom_line(size = 1.2) + # Adiciona as linhas conectando os pontos
  geom_point(size = 2) + # Adiciona os pontos nos meses
  ggtitle("Sinistro de 2024 e projeção para 2025") + # Adiciona o título do gráfico
  xlab("Mês") + # Define o rótulo do eixo X
  ylab("Sinistros (em R$ 1 milhão)") + # Define o rótulo do eixo Y
  scale_y_continuous(labels = scales::label_number(big.mark = ".", decimal.mark = ",")) +
  # Ajusta o formato dos números
  scale_x_discrete(limits = meses) + # Mantém a ordem correta dos meses no eixo X
  theme_minimal() + # Define um tema limpo para o gráfico
  theme(axis.text.x = element_text(angle = 45, hjust = 1), # Rotaciona os rótulos do eixo X
  legend.position = "bottom", # Posiciona a legenda abaixo do gráfico
  panel.grid.major = element_blank(), # Remove as grades maiores
  panel.grid.minor = element_blank()) # Remove as grades menores
```

### 2.4.3 Análise de Precisão das Projeções

Indicadores	Resultado	Interpretação
MPE (Erro Percentual Médio)	0.26%	Calcula o erro percentual médio, e é uma métrica útil para entender a magnitude do erro em relação ao valor real.
ACF1 (Autocorrelação do Erro de Previsão)	0.23	Mede a relação entre os erros de previsão de períodos consecutivos. Valores próximos de zero indicam que os erros não têm correlação entre si, o que é desejável, pois sugere que os erros são independentes.

O MPE mostra que as previsões estão bem próximas dos valores reais, com um erro muito pequeno. O ACF1 indica que os erros de previsão de um mês estão um pouco relacionados com os erros do mês seguinte, o que significa que o modelo pode não estar capturando completamente como os dados se comportam ao longo do tempo. Isso sugere que o modelo pode ser melhorado para ajustar melhor essas relações e melhorar as previsões futuras.

Para melhorar o modelo, uma abordagem seria a inclusão de variáveis explicativas adicionais que possam influenciar os sinistros, como fatores econômicos, climáticos ou sazonais, poderia também ajudar a aprimorar a previsão.

### Recomendações

Aplicar modelos de regressão por serem mais eficazes quando houver dados disponíveis de fatores externos (como taxa de juros) que influenciam a série.  
Exemplo: Regressão Linear Multivariada.

**Código em R**  
`accuracy(previsao_ets)`