



UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE ECONOMIA



WILLIAN HIDEKI TENGUAN

**Métodos e modelos de previsão de séries temporais para índices  
de criminalidade**

CAMPINAS  
2021

WILLIAN HIDEKI TENGUAN

**Métodos e modelos de previsão de séries temporais para índices  
de criminalidade**

Monografia apresentada ao Instituto de  
Economia da Universidade Estadual de  
Campinas como parte dos requisitos  
exigidos para a obtenção do título de  
Bacharel em Ciências Econômicas.

Orientador: Prof. Dr. Marcelo Justus dos  
Santos

CAMPINAS  
2021

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Economia  
Mirian Clavico Alves - CRB 8/8708

T253m Tengan, Willian Hideki, 1995-  
Métodos e modelos de previsão de séries temporais para índices de criminalidade / Willian Hideki Tengan. – Campinas, SP : [s.n.], 2021.

Orientador: Marcelo Justus dos Santos.  
Trabalho de Conclusão de Curso (graduação) – Universidade Estadual de Campinas, Instituto de Economia.

1. Crime - Aspectos econômicos. 2. Séries temporais. 3. Previsão estatística. 4. Modelagem de dados. I. Santos, Marcelo Justus dos, 1977-. II. Universidade Estadual de Campinas. Instituto de Economia. III. Título.

Informações adicionais, complementares

**Palavras-chave em inglês:**

Crime - Economic aspects

Time series

Statistical forecasting

Data modeling

**Titulação:** Bacharel em Ciências Econômicas

**Banca examinadora:**

Marcelo Justus dos Santos [Orientador]

Paulo Ricardo Prates Boitrigo

**Data de entrega do trabalho definitivo:** 03-12-2021

WILLIAN HIDEKI TENGUAN

**Métodos e modelos de previsão de séries temporais para índices de criminalidade**

Monografia apresentada ao Instituto de Economia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Bacharel em Ciências Econômicas.

Data de aprovação: 03 / 12 / 2021

Banca Examinadora

---

Prof. Dr. Marcelo Justus dos Santos – Presidente da banca  
Universidade de Campinas - UNICAMP

---

Mestre Paulo Ricardo Prates Boitrigo  
Universidade de Campinas - UNICAMP

## **AGRADECIMENTOS**

Agradeço, primeiramente, ao professor e meu orientador Marcelo Justus, por todo apoio na execução deste trabalho.

A minha família e amigos, que sempre me apoiaram. Em especial a meus pais, Neide e Edson, que batalharam muito para que este momento chegasse.

## RESUMO

TENGUAN, Willian Hideki. Métodos e Modelos de Previsão de séries temporais para índices de criminalidade. Orientador: Dr. Marcelo Justus dos Santos. Ano. 2021. Monografia (Graduação em Ciências Econômicas) – Instituto de Economia, Universidade Estadual de Campinas, Campinas, 2021.

Neste trabalho foram avaliados diferentes métodos e modelos de previsão para projetar três séries de criminalidade no Estado de Minas Gerais, para o ano de 2021: roubo de veículos, furto de veículos e vítimas de homicídios consumados. Foram implementados 9 modelos diferentes, tomando como referência o erro de previsão. Os modelos implantados, em sua maioria, apresentam previsões mais acuradas frente a um método ingênuo de previsão.

**Palavras-chave:** Economia do Crime; Séries Temporais; Métodos de Previsão; Modelos Híbridos; R Studio.

## **ABSTRACT**

TENGUAN, Willian Hideki. Time series forecast methods and models for criminal rate. Orientador: Dr. Marcelo Justus dos Santos. Year. 2021. Monografia (Graduação em Ciências Econômicas) – Instituto de Economia, Universidade Estadual de Campinas, Campinas, 2021.

In this study, different forecasting methods and models were evaluated to predict three criminal time series from the State of Minas Gerais for the year of 2021: vehicles robbery, vehicles theft and victims of consummate murders. Nine different models were implemented, based on the forecast error. Most of the implemented models present more accurate forecasts compared to a naive forecasting method.

**Keywords:** Economics of Crime; Time Series; Forecast Methods; Hybrid Models; R Studio.

## LISTA DE ILUSTRAÇÕES

### FIGURAS

|                                                                                                                                                                                                                                    |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| FIGURA 1 - Total de registros reportados de roubos de veículos no Estado de Minas Gerais no período de janeiro de 2015 a dezembro de 2020 .....                                                                                    | 15 |
| FIGURA 2 - Total de registros reportados de furtos de veículos no Estado de Minas Gerais no período de janeiro de 2015 a dezembro de 2020 .....                                                                                    | 15 |
| FIGURA 3 - Total de registros reportados de vítimas de homicídios consumados no Estado de Minas Gerais no período de janeiro de 2015 a dezembro de 2020 .....                                                                      | 16 |
| FIGURA 4 - Decomposição clássica aditiva entre tendência, sazonalidade e componente aleatório da série Furto de Veículos .....                                                                                                     | 17 |
| FIGURA 5 - Decomposição clássica aditiva entre tendência, sazonalidade e componente aleatório da série Roubo de Veículos .....                                                                                                     | 17 |
| FIGURA 6 - Decomposição clássica aditiva entre tendência, sazonalidade e componente aleatório da série Vítimas de Homicídios Consumados .....                                                                                      | 18 |
| FIGURA 7 - Exemplo de arquitetura de uma rede neural para séries temporais, contendo três inputs, uma camada invisível com dois neurônios e um output (a previsão) .....                                                           | 23 |
| FIGURA 8 - Previsões para 2021 da série Vítimas de Homicídios Consumados utilizando modelos Suavização Exponencial (ETS), Holt-Winters aditivo (HW_Add) e Holt-Winters multiplicativo (HW_Mult) .....                              | 27 |
| FIGURA 9 - Previsões para 2021 da série Vítimas de Homicídios Consumados utilizando modelos ARIMA, Prophet e Regressão Linear .....                                                                                                | 28 |
| FIGURA 10 - Previsões para 2021 da série Vítimas de Homicídios Consumados utilizando modelos de Redes Neurais Artificiais de Camada Única (ANN Single Layer), Redes Neurais de Múltiplas Camadas (ANN Multi Layer) e Híbrido ..... | 29 |
| FIGURA 11 - Previsões para 2021 da série Furto de Veículos utilizando modelos Suavização Exponencial (ETS), Holt-Winters aditivo (HW_Add) e Holt-Winters multiplicativo (HW_Mult) .....                                            | 30 |
| FIGURA 12 - Previsões para 2021 da série Furto de Veículos utilizando modelos ARIMA, Prophet e Regressão Linear .....                                                                                                              | 31 |
| FIGURA 13 - Previsões para 2021 da série Furto de Veículos utilizando modelos de Redes Neurais Artificiais de Camada Única (ANN Single Layer), Redes Neurais de Múltiplas Camadas (ANN Multi Layer) e Híbrido .....                | 32 |



|                                                                                                                                                                                                                     |    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| FIGURA 14 - Previsões para 2021 da série Roubo de Veículos utilizando modelos Suavização Exponencial (ETS), Holt-Winters aditivo (HW_Add) e Holt-Winters multiplicativo (HW_Mult) .....                             | 33 |
| FIGURA 15 - Previsões para 2021 da série Roubo de Veículos utilizando modelos ARIMA, Prophet e Regressão Linear .....                                                                                               | 34 |
| FIGURA 16 - Previsões para 2021 da série Roubo de Veículos utilizando modelos de Redes Neurais Artificiais de Camada Única (ANN Single Layer), Redes Neurais de Múltiplas Camadas (ANN Multi Layer) e Híbrido ..... | 35 |
| FIGURA 17- Previsões para 2022 da série Furto de Veículos utilizando modelo de Holt-Winters Multiplicativo .....                                                                                                    | 39 |
| FIGURA 18- Previsões para 2022 da série Roubo de Veículos utilizando modelo de Holt-Winters Multiplicativo .....                                                                                                    | 40 |
| FIGURA 19- Previsões para 2022 da série Vítimas de Homicídios Consumados utilizando modelo de Holt-Winters Multiplicativo .....                                                                                     | 41 |

## QUADROS

|                                                                                                                                 |    |
|---------------------------------------------------------------------------------------------------------------------------------|----|
| QUADRO 1- Exemplo de matriz para sazonalidade .....                                                                             | 24 |
| QUADRO 2- Parâmetros dos modelos escolhidos .....                                                                               | 26 |
| QUADRO 3 - Resumo da performance dos modelos de previsão para 12 passos à frente .....                                          | 36 |
| QUADRO 4 - Previsões para 2022 da série Furto de Veículos utilizando modelo de Holt-Winters Multiplicativo .....                | 40 |
| QUADRO 5- Previsões para 2022 da série Roubo de Veículos utilizando modelo de Holt-Winters Multiplicativo .....                 | 41 |
| QUADRO 6 - Previsões para 2022 da série Vítimas de Homicídios Consumados utilizando modelo de Holt-Winters Multiplicativo ..... | 42 |

## SUMÁRIO

|          |                                                                      |           |
|----------|----------------------------------------------------------------------|-----------|
| <b>1</b> | <b>INTRODUÇÃO .....</b>                                              | <b>11</b> |
|          | 1.1 O USO DE MÉTODOS E MODELOS DE PREVISÃO E REVISÃO LITERÁRIA ..... | 11        |
| <b>2</b> | <b>METODOLOGIA.....</b>                                              | <b>14</b> |
|          | 2.1 FONTE DE DADOS, VARIÁVEIS E PRÉ-TRATAMENTO .....                 | 14        |
|          | 2.2 APRESENTAÇÃO DAS SÉRIES .....                                    | 14        |
|          | 2.3 DECOMPOSIÇÃO DAS SÉRIES.....                                     | 16        |
|          | 2.4 AVALIAÇÃO DOS MODELOS E COMPARAÇÃO .....                         | 18        |
| <b>3</b> | <b>MÉTODOS DE PREVISÃO IMPLANTADOS.....</b>                          | <b>20</b> |
|          | 3.1 SUAVIZAÇÃO EXPONENCIAL .....                                     | 20        |
|          | 3.2 SUAVIZAÇÃO EXPONENCIAL SAZONAL DE HOLT-WINTERS .....             | 20        |
|          | 3.3 SAZONAL AUTORREGRESSIVO INTEGRADO COM MÉDIAS MÓVES (SARIMA)..... | 21        |
|          | 3.4 REDES NEURAIS ARTIFICIAIS (ANN).....                             | 22        |
|          | 3.5 REGRESSÃO LINEAR PARA SÉRIES TEMPORAIS.....                      | 23        |
|          | 3.6 PROPHET .....                                                    | 24        |
|          | 3.7 MODELOS HÍBRIDOS.....                                            | 25        |
| <b>4</b> | <b>PARÂMETROS DOS MODELOS .....</b>                                  | <b>26</b> |
| <b>5</b> | <b>RESULTADOS DOS MODELOS .....</b>                                  | <b>27</b> |
|          | 5.1 SÉRIE VÍTIMA DE HOMICÍDIOS CONSUMADOS:.....                      | 27        |
|          | 5.2 SÉRIE FURTO DE VEÍCULOS .....                                    | 30        |
|          | 5.3 SÉRIE ROUBO DE VEÍCULOS: .....                                   | 33        |
|          | 5.4 QUADRO RESUMO DA PERFORMANCE DOS MODELOS .....                   | 35        |
| <b>6</b> | <b>CONCLUSÃO E CONSIDERAÇÕES FINAIS.....</b>                         | <b>37</b> |
| <b>7</b> | <b>PREVISÃO PARA 2022 .....</b>                                      | <b>39</b> |
|          | <b>REFERÊNCIAS.....</b>                                              | <b>43</b> |
|          | <b>APÊNDICE A-CÓDIGOS R STUDIO: SÉRIE HOMICÍDIOS.....</b>            | <b>46</b> |
|          | <b>APÊNDICE B-CÓDIGOS R STUDIO: SÉRIE DE FURTOS.....</b>             | <b>52</b> |
|          | <b>APÊNDICE C-CÓDIGOS R STUDIO: SÉRIE DE ROUBOS.....</b>             | <b>58</b> |

## 1 INTRODUÇÃO

Atualmente, a criminalidade é um enorme problema enfrentado pela sociedade do Brasil. O alto índice de violência afeta o comportamento da população brasileira em geral, ou seja, indivíduos de todas as classes e regiões, gerando uma crescente sensação de insegurança. O custo social da criminalidade atinge o campo econômico, auferindo grandes gastos com segurança, promovendo prejuízos materiais, reduzindo investimentos em regiões violentas e diminuindo o capital humano.

Através da análise dos dados é possível estimar, através de métodos quantitativos, os índices futuros a fim de se fornecer informações importantes e ajustar expectativas dos agentes econômicos, indicando cenários prováveis e plausíveis, suportando a tomada de decisão e auxiliando no planejamento. Assim, o uso de modelos de previsão estatísticos é um ferramental importante, utilizando-se de dados do passado para estimar o futuro, identificando tendências e padrões com os dados. A partir disso, com o avanço do processamento dos computadores e de técnicas como o aprendizado das máquinas (*Machine Learning - ML*), diversos novos modelos de alta acurácia são criados.

### 1.1 O uso de métodos e modelos de previsão e revisão literária

Conforme apontado por Box, Jenkins e Reinsel (2008), a utilização das informações observadas em uma série temporal no momento  $t$  para previsão de algum tempo  $t + l$  providencia uma base para planejamentos econômicos e de negócios, planejamentos de produção, controles de produção e inventário, além de controles e otimizações de processos industriais. Sendo assim, são diversos os temas em que os métodos e modelos de previsão podem ter utilidade para os agentes, sejam eles públicos ou privados.

No trabalho de Naveena *et al.* (2017), foi utilizado um modelo híbrido dos modelos ARIMA (Autorregressivo Integrado de Médias Móveis - para a parte linear) e ANN (Redes Neurais Artificiais – para a parte não linear) para prever o preço de café da espécie Robusta na Índia, apresentando um erro absoluto percentual médio de 2,96, enquanto o modelo puramente ARIMA e puramente ANN apresentaram 4,02 e 3,65, respectivamente.

Nos estudos realizados por Chaudhurri e Ghosh (2016) também foi utilizado o modelo de Redes Neurais Artificiais, além de outros modelos como o Autorregressivo Geral de Heterocedasticidade Condicional (GARCH) para prever a taxa de câmbio. O estudo concluiu que os modelos de Redes Neurais Artificiais tiveram melhor performance.

No trabalho de Azevedo (2017), foram analisados ativos negociados na Bolsa de Valores brasileira (BOVESPA) em que, utilizando-se também do modelo GARCH, entre outros, buscou-se entender empiricamente as propriedades das diferentes metodologias na previsão de séries temporais.

Já Rodrigues e Rodrigues (2020), buscaram trabalhar diversos modelos de previsão para aplicar no consumo brasileiro de combustíveis leves a fim de, através da análise dos erros de previsão, comparar e encontrar o melhor modelo que atende aos seus objetivos.

O estudo de Boone *et al.* (2019) traz conexões entre as novas informações geradas na era da *big data* e suas tecnologias associadas com os modelos de previsão de vendas dentro da cadeia de oferta, possibilitando aumentar a acurácia destes modelos.

Steinbuks (2019) ressalta a importância da previsão de demanda de eletricidade futura nos países em desenvolvimento. Além disso, mostra que os modelos econométricos de previsão possuem erros muito menores que métodos simplesmente heurísticos, que assumem que a produção de eletricidade cresce a uma taxa exógena ou proporcional ao crescimento do produto.

O trabalho de Athanasopoulos *et al.* (2011) avalia a performance de diversos métodos univariados e multivariados, tanto puramente de séries temporais, quanto de modelos com variáveis explanatórias, para a previsão de dados de turismo.

Rostami-Tabar *et al.* (2021) destaca a importância das práticas de previsões para o bem social, buscando-se o avanço em objetivos como igualdade, justiça social e sustentabilidade. Tais práticas são referidas como “Previsão para o bem social” (*Forecasting for Social good – FSG*), onde os benefícios para a sociedade e o ambiente são predominantes sobre os resultados econômicos e financeiros.

Métodos baseados em Redes Neurais Artificiais (ANN) foram utilizados por Corcoran, Wilson e Ware (2003) para prever incidentes criminosos em determinadas áreas geográficas, identificando espaços de relativos índices elevados (*hot spots*).

Já os autores Chen, Yuan e Shu (2008) compararam o método de Suavização Exponencial com o modelo ARIMA para realizar uma previsão de curto prazo de crimes contra a propriedade numa cidade da China, em que chegaram à conclusão que o modelo ARIMA foi o mais acurado.

No estudo de Singh (2018) foram comparados diversos métodos e modelos de previsão para incidentes em crimes, como o ARIMA, as redes neurais artificiais (ANN), as redes neurais recorrentes e a memória de curto prazo longa (LSTM). Mediante a isso, o autor conclui que o modelo que teve maior acurácia de previsão foi o de redes neurais recorrentes, apresentando a menor diferença entre os modelos de teste e treino.

Portanto, nota-se que há espaço para a utilização de métodos e modelos de previsão em diversas áreas, inclusive na análise de séries de criminalidade, como pretendemos executar neste trabalho, cujo objetivo é o de fornecer cenários plausíveis em um horizonte de previsão que sirva de fonte de informação à tomada de decisão dos agentes públicos e privados.

## **2 METODOLOGIA**

Para o desenvolvimento do presente estudo, serão analisadas três séries de criminalidade do estado de Minas Gerais: roubo de veículos, furto de veículos e vítimas de homicídios consumados. O processo de escolha de Minas Gerais foi pautado na disponibilidade e qualidade dos dados.

No Decreto Lei nº 2848, roubo é descrito no Artigo 157 como “Subtrair coisa móvel alheia, para si ou para outrem, mediante grave ameaça ou violência a pessoa, ou depois de havê-la, por qualquer meio, reduzido à impossibilidade de resistência”; furto é descrito no Artigo 157 como “Subtrair, para si ou para outrem, coisa alheia móvel” e homicídio é descrito no Artigo 121 como “Matar alguém”. Portanto, as duas primeiras séries (roubo e furto de veículos) se tratam de Crimes contra o Patrimônio, ou seja, toda ação que atente contra bens de uma pessoa ou organização, que tem como objetivo qualquer coisa que tenha valor patrimonial. Já a terceira série, homicídio, enquadra-se como Crime contra a Vida.

### **2.1 Fonte de dados, variáveis e pré-tratamento**

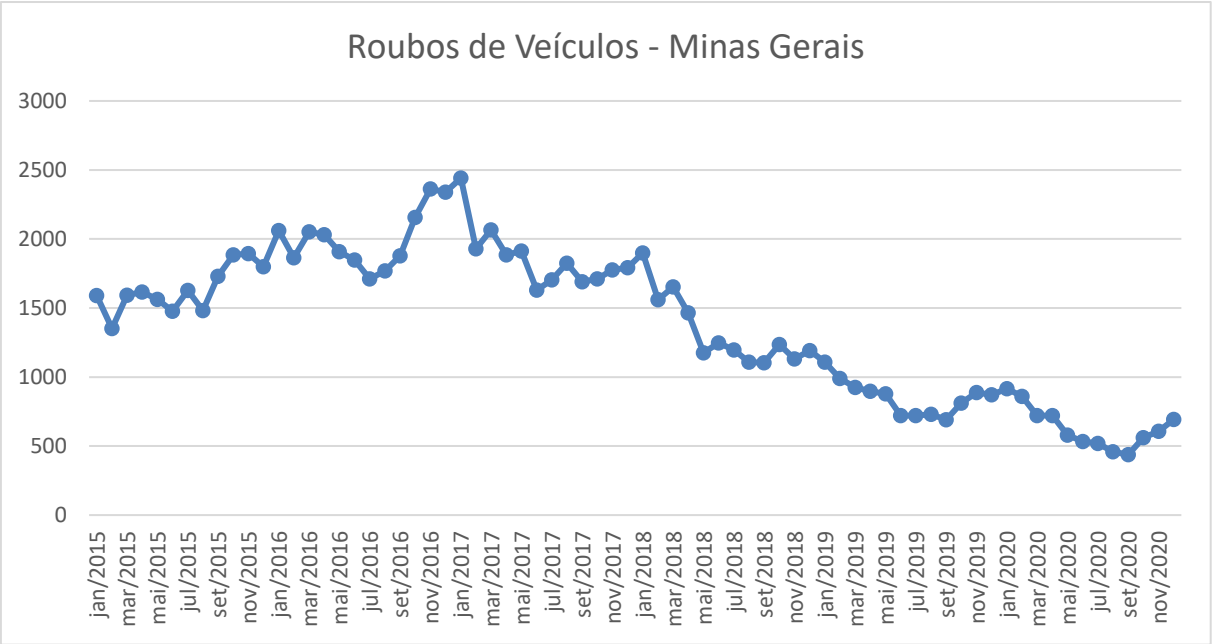
Os dados abertos foram extraídos a partir do portal da Secretaria de Estado de Justiça e Segurança Pública de Minas Gerais – SEJUSP – MG, sendo divulgados mensalmente. A agregação é ao nível de cada um dos 853 municípios componentes do Estado para cada série apresentada.

Foi realizado um processo de extração dos dados do portal da SUJESP, consolidação em um banco de dados Microsoft Access, agrupamento por Estado e mês para cada série, carregamento no software R Studio e conversão em objetos de séries temporais (timeseries).

### **2.2 Apresentação das séries**

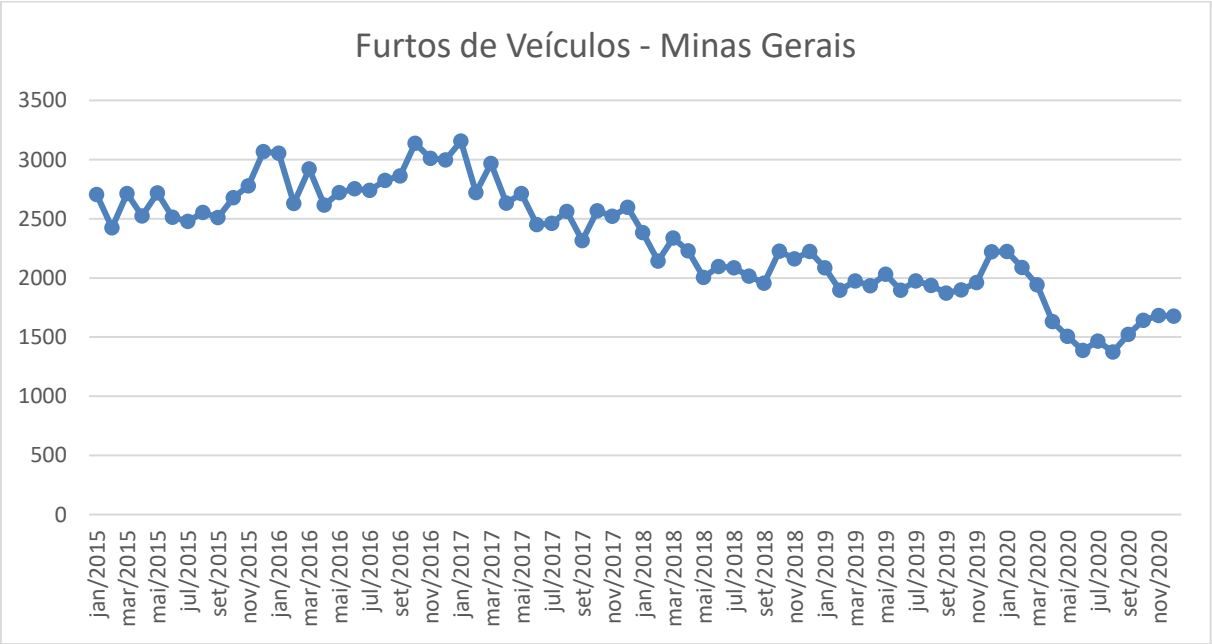
Nas figuras abaixo, podem ser visualizadas as séries que se iniciam em janeiro de 2015 e se encerram em dezembro de 2020 (representados pelo eixo horizontal) e o total de registros (representado pelo eixo vertical):

FIGURA 1 - Total de registros reportados de roubos de veículos no Estado de Minas Gerais no período de janeiro de 2015 a dezembro de 2020



Fonte: SEJUSP

FIGURA 2 - Total de registros reportados de furtos de veículos no Estado de Minas Gerais no período de janeiro de 2015 a dezembro de 2020



Fonte: SEJUSP

**FIGURA 3 - Total de registros reportados de vítimas de homicídios consumados no Estado de Minas Gerais no período de janeiro de 2015 a dezembro de 2020**



Fonte: SEJUSP

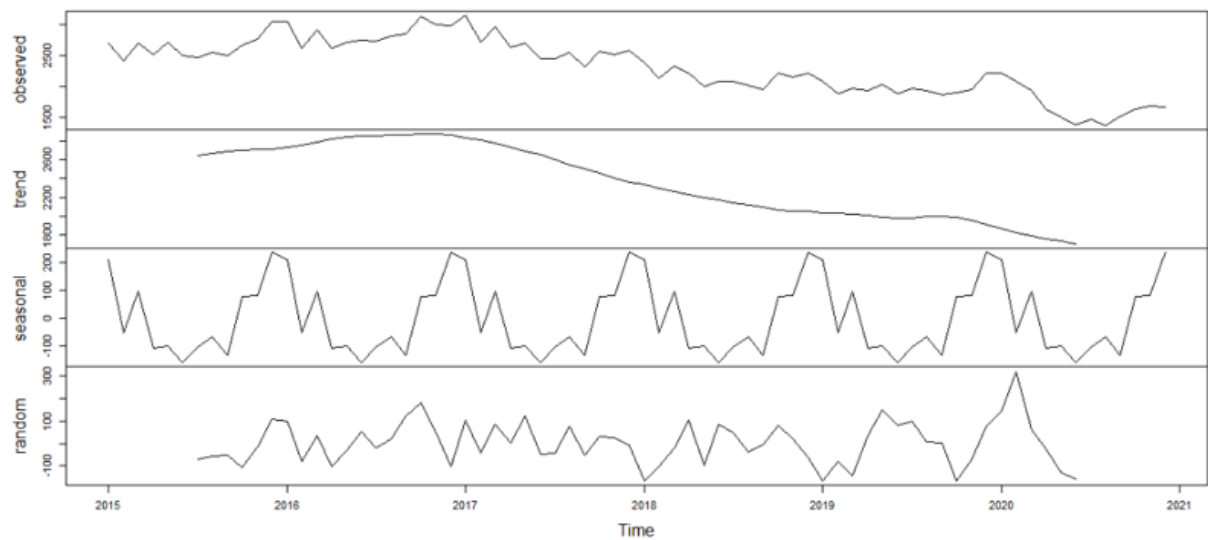
## 2.3 Decomposição das séries

A partir do carregamento dos dados, foi realizado a decomposição aditiva das séries em tendência, sazonalidade e componente aleatório, que permite visualizar melhor os dados. Além disso, conforme apontado por Souza, Samohyl e Meurer (2004), a decomposição clássica é uma ferramenta útil pois, além de permitir previsões, auxilia na tomada de decisão acerca do método de previsão mais adequado às características dos dados disponíveis.

As séries decompostas são apresentadas a seguir:

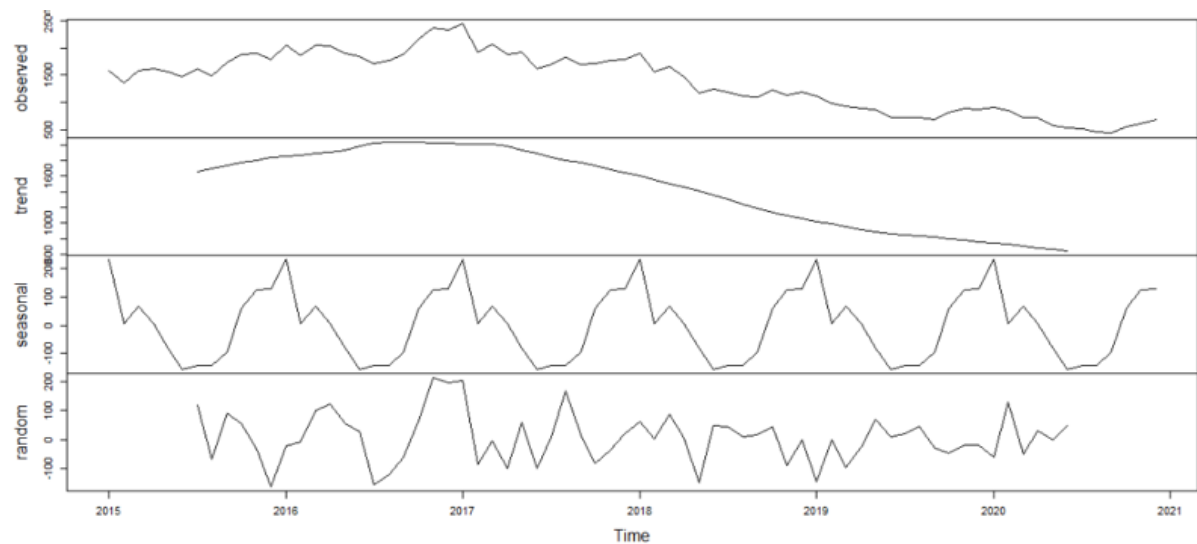


FIGURA 4 - Decomposição clássica aditiva entre tendência, sazonalidade e componente aleatório da série Furto de Veículos



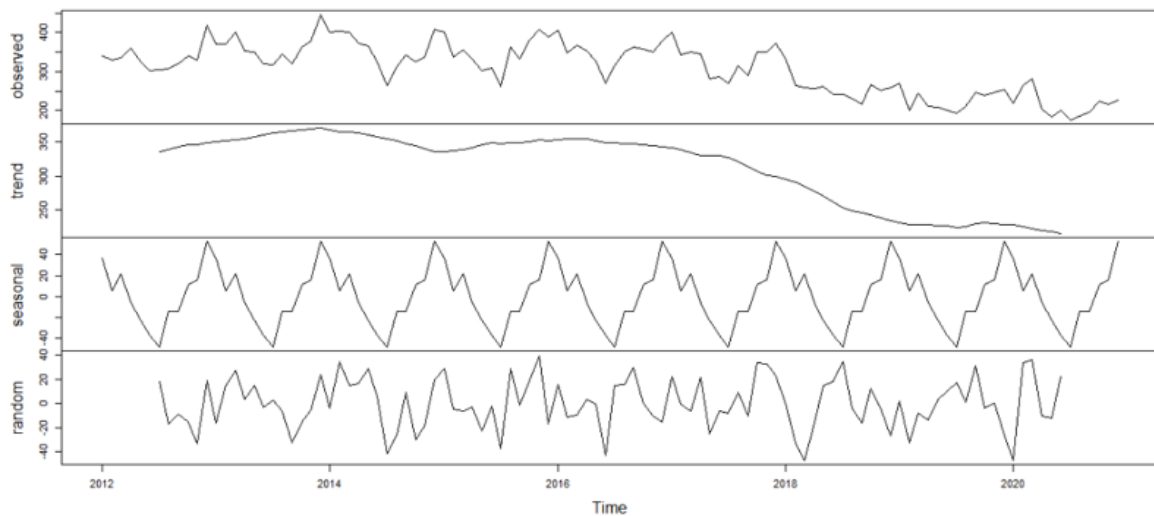
Fonte: SEJUSP

FIGURA 5 - Decomposição clássica aditiva entre tendência, sazonalidade e componente aleatório da série Roubo de Veículos



Fonte: SEJUSP

FIGURA 6 - Decomposição clássica aditiva entre tendência, sazonalidade e componente aleatório da série Vítimas de Homicídios Consumados



Fonte: SEJUSP

## 2.4 Avaliação dos modelos e comparação

Em cada série, foi separado dois tipos de dados: os dados de treino, demarcado pelo período de 2015 a 2020, e os dados de teste, demarcado pelo período posterior a 2020, até a última publicação dos dados. Assim, foi realizado a previsão com horizonte  $h = 12$  para cada uma das séries. A cada publicação de novos dados realizados pela SUJESP, que costumam ocorrer em 1 mês após o fechamento do mês anterior, foi feita uma comparação entre os resultados previstos por cada um dos modelos, com o que foi de fato realizado, determinando assim as métricas de erros dos diferentes modelos, das distintas séries, para cada mês de 2021.

Ao final de 2021, é possível verificar, através de um quadro comparativo, todos os valores previstos e realizados, calculando assim a raiz quadrada do erro quadrático médio (RMSE) e o erro absoluto percentual médio (MAPE) que serão índices fundamentais para a comparação entre os modelos, além de ser parâmetro de escolha do melhor modelo.

O erro é entendido como a diferença entre o valor observado e sua previsão, e pode ser formalizado na seguinte equação proposta por Hydman e Athanasopoulos (2018):

$$e_{T+h} = y_{T+h} - \hat{y}_{T+h}|_T$$

A raiz quadrada do erro quadrático médio (RMSE) é definida por:

$$RMSE = \sqrt{\frac{1}{h} \sum_{j=1}^h e^2}$$

O erro absoluto percentual médio (MAPE) é definido por:

$$MAPE = \frac{100\%}{n} \cdot \sum_{j=1}^h \frac{e^t}{y^t}$$

### 3 MÉTODOS DE PREVISÃO IMPLANTADOS

#### 3.1 Suavização Exponencial

Segundo Hyndmann e Athanasopoulos (2018), o método de Suavização Exponencial foi proposto ao final da década de 1950 e serviu de motivação para grande parte dos métodos de previsão. Suas previsões se baseiam nas médias ponderadas de observações passadas, com seus pesos decaindo exponencialmente conforme as observações ficam mais longes do tempo  $t$ .

O método de Suavização Exponencial pode ser representado da seguinte forma:

$$\hat{y}_{T+1}|T = \alpha y_T + \alpha(1 - \alpha)y_{T-1} + \alpha(1 - \alpha)^2 y_{T-2} + \dots$$

Onde  $\alpha$  é o parâmetro de suavização com valor entre 0 e 1.

O valor previsto um passo à frente  $T+1$  é uma média ponderada de todas as observações anteriores da série. A taxa que o peso diminui é controlada pelo parâmetro  $\alpha$ .

#### 3.2 Suavização Exponencial Sazonal de Holt-Winters

O modelo proposto por Holt (1957) e Winters (1960) compreende uma equação de previsão e três equações de suavização – uma para o nível  $\ell_t$ , uma para a tendência  $b_t$ , e uma para o componente sazonal  $s_t$ , com os seus parâmetros de suavização  $\alpha$ ,  $\beta^*$  e  $\gamma$ .

A forma aditiva do método de Holt-Winters é dada por:

$$\hat{y}_{t+h}|t = \ell_t + hb_t + s_{t+h-m(k+1)}$$

Sendo:

$$\begin{aligned}\ell_t &= \alpha(y_t - s_{t-m}) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\ s_t &= \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m}\end{aligned}$$

Onde  $m$  representa a frequência da sazonalidade. Em caso de dados mensais, “ $m$ ” é igual a 12. Além disso, “ $k$ ” é a parte inteira de  $(h-1)/m$ , que assegura que a estimativa do índice sazonal usada na previsão venha do último ano da amostra.

A forma multiplicativa do método é dada por:

$$\hat{y}_{t+h}|_t = (l_t + hb_t)s_{t+h-m(k+1)}$$

Sendo:

$$\begin{aligned} \ell_t &= \alpha \left( \frac{y_t}{s_{t-m}} \right) + (1 - \alpha)(\ell_{t-1} + b_{t-1}) \\ b_t &= \beta * (\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1} \\ s_t &= \gamma \left( \frac{y_t}{\ell_{t-1} + b_{t-1}} \right) + (1 - \gamma)s_{t-m} \end{aligned}$$

### 3.3 Sazonal Autorregressivo Integrado com Médias Móveis (SARIMA)

ARIMA é um acrônimo para modelo Autorregressivo Integrado de Médias Móveis que, ao se incorporar as relações sazonais, passa a ser nomeado como SARIMA.

Os modelos SARIMA são bastante populares para previsão de séries temporais, enquanto a abordagem ARIMA permite que valores futuros de uma série sejam previstos tomando como base apenas seus valores anteriores.

Um modelo SARIMA ( $p, d, q$ ) ( $P, D, Q$ ) é representado da seguinte forma, segundo Barros e Ferreira (2018):

$$\phi(L)\Phi(L)\Delta^d\Delta^D y_t = \theta(L)\Theta(L)\varepsilon_t$$

Onde:

$p$  é a ordem do polinômio autorregressivo não sazonal  $\phi(L)$ ;

$P$  é a ordem do polinômio autorregressivo sazonal  $\Phi(L)$ ;

$q$  é a ordem do polinômio de médias móveis não sazonal  $\theta(L)$ ;

$Q$  é a ordem do polinômio de médias móveis sazonal  $\Theta(L)$ ;

$d$  é a ordem de diferença não sazonal;

$D$  é a ordem de diferença sazonal;

$$\phi(L) = (1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p);$$

$$\Phi(L) = (1 - \phi_1 L^s - \phi_2 L^{2s} - \dots - \phi_p L^{Ps});$$

$$\theta(L) = (1 - \theta_1 L - 1 - \theta_2 L^2 - \dots - \theta_p L^q);$$

$$\Theta(L) = (1 - \Theta_1 L^s - \Theta_2 L^{2s} - \dots - \Theta_p L^{Qs});$$

$$\Delta = 1 - L;$$

$L$  é o operador de defasagem tal que  $L^n y_t = y_{t-n}$ .

### 3.4 Redes Neurais Artificiais (ANN)

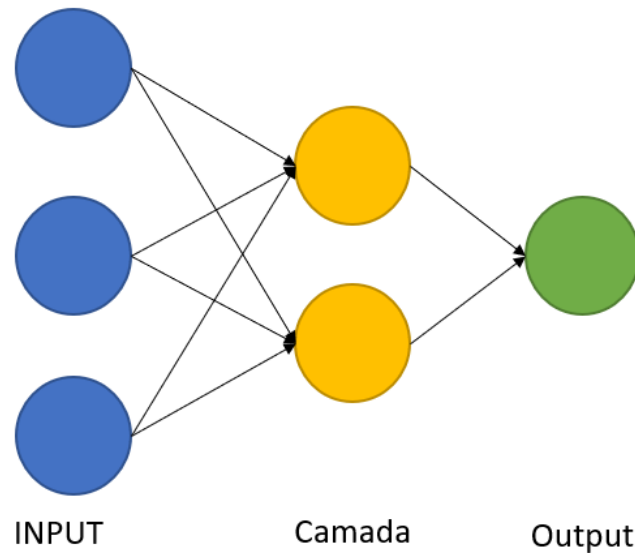
Atualmente, os modelos de Redes Neurais ganharam bastante relevância, junto com seus temas relacionados, como: aprendizado das máquinas (*Machine Learning – ML*), *Deep Learning* e Inteligência Artificial (IA). Sua estrutura é pensada para imitar o *design* do cérebro humano, sendo aplicada em diversos problemas científicos, tendo aumentado o seu uso em estatística, especialmente no reconhecimento de padrões.

Os modelos ANN podem ser pensados como um sistema composto de uma série de *inputs* e de *outputs*, em uma interação possivelmente não linear. No contexto de séries temporais, o *output* pode ser entendido como o valor a ser previsto e os *inputs* como os valores nos tempos anteriores ou outras variáveis explanatórias. As conexões entre essas séries são feitas por uma ou mais camadas invisíveis de neurônios ou nós.

A estrutura de uma Rede Neural Artificial é comumente chamada de arquitetura. Desse modo, escolher a arquitetura envolve a determinação da quantidade de camadas, o número de neurônios de cada camada e como os *inputs*, *outputs* e as camadas invisíveis se conectam.

Um exemplo de Rede Neural típica contendo três *inputs* e uma camada com dois neurônios pode ser representada da seguinte forma:

FIGURA 7 - Exemplo de arquitetura de uma rede neural para séries temporais, contendo três inputs, uma camada invisível com dois neurônios e um output (a previsão).



Fonte: Autor

### 3.5 Regressão Linear para Séries Temporais

O conceito básico de um modelo de regressão consiste na possibilidade de prever uma série temporal de interesse  $\underline{y}$  (variável de previsão), assumindo que ela possui uma relação linear com outra série temporal  $\underline{x}$  (variável “preditora”).

Segundo Hyndmann e Athanasopoulos (2018), uma previsão de  $\underline{y}$  pode ser obtido pela equação:

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 x_{1,t} + \hat{\beta}_2 x_{2,t} + \dots + \hat{\beta}_k x_{k,t}$$

Nos modelos de Regressão Linear para Séries Temporais, há algumas variáveis preditoras que ocorrem frequentemente, sendo que, duas delas se destacam pela importância: a tendência e a sazonalidade.

A tendência pode ser modelada usando  $x_{1,t} = t$  como preditor:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

Já a sazonalidade pode ser realizada utilizando-se um quadro, assumindo valor 1 para o mês representado e 0 para os outros meses, conforme exemplo:

QUADRO 1- Exemplo de matriz para sazonalidade

|           | m1 | m2 | m3 | m4 | m5 | m6 |
|-----------|----|----|----|----|----|----|
| Janeiro   | 1  | 0  | 0  | 0  | 0  | 0  |
| Fevereiro | 0  | 1  | 0  | 0  | 0  | 0  |
| Março     | 0  | 0  | 1  | 0  | 0  | 0  |
| Abril     | 0  | 0  | 0  | 1  | 0  | 0  |
| Mai       | 0  | 0  | 0  | 0  | 1  | 0  |
| Junho     | 0  | 0  | 0  | 0  | 0  | 1  |

Fonte: Autor

A escolha dos coeficientes ( $\beta$ ) se dá através da minimização dos erros quadráticos, conforme a minimização da seguinte fórmula:

$$\sum_{t=1}^T \varepsilon_t^2 = \sum_{t=1}^T (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_k x_{k,t})^2.$$

### 3.6 Prophet

O modelo Prophet foi introduzido pelo Facebook originalmente para prever dados diários com sazonalidade semanal e anual, com efeitos de feriados. Posteriormente, foi estendido para cobrir maiores tipos de dados sazonais.

O modelo pode ser considerado como um Modelo de Regressão não-linear, na forma:

$$y_t = g(t) + s(t) + h(t) + \varepsilon_t$$

Onde  $g(t)$  representa a tendência,  $s(t)$  descreve os padrões sazonais,  $h(t)$  captura os efeitos de feriados e  $\varepsilon_t$  é o termo de erro do tipo ruído branco.

O modelo Prophet utiliza abordagem Bayesiana para selecionar automaticamente os nós e outras características do modelo.



### 3.7 Modelos Híbridos

Modelos híbridos combinam paradigmas de diferentes algoritmos de previsão para a formulação de um modelo único. Seu uso permite capturar diferentes padrões dentro de uma mesma série, podendo ser utilizado para previsão de uma série  $Y_t$  dividindo-a em dois componentes: um linear ( $L_t$ ) e outro não-linear ( $N_t$ ).

Para tal, é comum utilizar a decomposição da série em tendência, assumida como linear, sazonalidade e componente aleatório ou resíduo, assumidos como não lineares. Neste trabalho, foi utilizado o método de Suavização Exponencial para a parte linear e Redes Neurais Artificiais de Camada Única para a parte não-linear.

#### 4 PARÂMETROS DOS MODELOS

O quadro a seguir mostra os resultados da escolha dos parâmetros dos modelos:

QUADRO 2- Parâmetros dos modelos escolhidos

| Modelo                                      | Homicídios                                                                                                                                                                                                                                                                                                                                                                              | Furto de Veículos                                                                                                                                                                                                                                                                                                                                                                                                             | Roubo de Veículos                                                                                                                                                                                                                                                                                                                                                                                |
|---------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Modelo Linear                               | $\alpha = 16,03;$<br>$\beta_{trend} = -0,46;$<br>$\beta_{season2} = -0,64;$<br>$\beta_{season3} = -0,31;$<br>$\beta_{season4} = -0,73;$<br>$\beta_{season5} = -1,07;$<br>$\beta_{season6} = -1,18;$<br>$\beta_{season7} = -1,31;$<br>$\beta_{season8} = 0,71;$<br>$\beta_{season9} = -0,68;$<br>$\beta_{season10} = -0,13;$<br>$\beta_{season11} = -0,08;$<br>$\beta_{season12} = 0,13$ | $\alpha = 71713,7;$<br>$\beta_{trend} = -533,66;$<br>$\beta_{season2} = -7996,48;$<br>$\beta_{season3} = -2639,29;$<br>$\beta_{season4} = -8383,95;$<br>$\beta_{season5} = -7026,03;$<br>$\beta_{season6} = -9327,98;$<br>$\beta_{season7} = -8375,06;$<br>$\beta_{season8} = -7381,39;$<br>$\beta_{season9} = -8037,94;$<br>$\beta_{season10} = -2126,36;$<br>$\beta_{season11} = -1830,58;$<br>$\beta_{season12} = 1971,96$ | $\alpha = 212,09;$<br>$\beta_{trend} = -1,53;$<br>$\beta_{season2} = -13,68;$<br>$\beta_{season3} = -8,13;$<br>$\beta_{season4} = -10,74;$<br>$\beta_{season5} = -16,43;$<br>$\beta_{season6} = -21,41;$<br>$\beta_{season7} = -19,68;$<br>$\beta_{season8} = -19,86;$<br>$\beta_{season9} = -16,92;$<br>$\beta_{season10} = -5,49;$<br>$\beta_{season11} = -0,77;$<br>$\beta_{season12} = 1,56$ |
| Suavização Exponencial                      | $\alpha=0,43$                                                                                                                                                                                                                                                                                                                                                                           | $\alpha=0,76$                                                                                                                                                                                                                                                                                                                                                                                                                 | $\alpha=0,84$                                                                                                                                                                                                                                                                                                                                                                                    |
| Holt-Winters Aditivo                        | $\alpha=0,38;$<br>$\beta=0,02$                                                                                                                                                                                                                                                                                                                                                          | $\alpha=0,73;$<br>$\beta=0,04$                                                                                                                                                                                                                                                                                                                                                                                                | $\alpha=0,65;$<br>$\beta=0,08$                                                                                                                                                                                                                                                                                                                                                                   |
| Holt-Winters Multiplicativo                 | $\alpha=0,48;$<br>$\beta=0,03$                                                                                                                                                                                                                                                                                                                                                          | $\alpha=0,84;$<br>$\beta=0,05$                                                                                                                                                                                                                                                                                                                                                                                                | $\alpha=0,74;$<br>$\beta=0,08$                                                                                                                                                                                                                                                                                                                                                                   |
| ARIMA                                       | $p=0; d=1; q=1;$<br>$P=0; D=1; Q=1$                                                                                                                                                                                                                                                                                                                                                     | $p=1; d=1; q=0;$<br>$P=0; D=1; Q=1$                                                                                                                                                                                                                                                                                                                                                                                           | $p=1; d=1; q=0;$<br>$P=0; D=1; Q=1$                                                                                                                                                                                                                                                                                                                                                              |
| Redes Neurais Artificiais - Camada única    | $p=2; P=1; k=2;$<br>$m=1$                                                                                                                                                                                                                                                                                                                                                               | $p=1; P=1; k=2;$<br>$m=1$                                                                                                                                                                                                                                                                                                                                                                                                     | $p=1; P=1; k=2;$<br>$m=1$                                                                                                                                                                                                                                                                                                                                                                        |
| Redes Neurais Artificiais - Camada Múltipla | $n=3;$<br>$r=20$                                                                                                                                                                                                                                                                                                                                                                        | $n=5;$<br>$r=20$                                                                                                                                                                                                                                                                                                                                                                                                              | $n=5;$<br>$r=20$                                                                                                                                                                                                                                                                                                                                                                                 |
| Prophet                                     | $k=0,30; m=0,92;$<br>$growth = linear$                                                                                                                                                                                                                                                                                                                                                  | $k=0,28; m=0,83,$<br>$growth = linear$                                                                                                                                                                                                                                                                                                                                                                                        | $k=0,79; m=0,63,$<br>$growth = linear$                                                                                                                                                                                                                                                                                                                                                           |
| Híbrido                                     | $\alpha=0,99;$<br>$p=4; P=1; k=3;$<br>$m=1$                                                                                                                                                                                                                                                                                                                                             | $\alpha=0,99;$<br>$p=7; P=1; k=4; m=1$                                                                                                                                                                                                                                                                                                                                                                                        | $\alpha=0,99;$<br>$p=5; P=1; k=4; m=1$                                                                                                                                                                                                                                                                                                                                                           |

Fonte: Autor

## 5 RESULTADOS DOS MODELOS

Os resultados dos modelos são apresentados nas figuras e nos quadros a seguir:

### 5.1 Série Vítima de Homicídios Consumados:

FIGURA 8 - Previsões para 2021 da série Vítimas de Homicídios Consumados utilizando modelos Suavização Exponencial (ETS), Holt-Winters aditivo (HW\_Add) e Holt-Winters multiplicativo (HW\_Mult)



Fonte: SEJUSP

FIGURA 9 - Previsões para 2021 da série Vítimas de Homicídios Consumados utilizando modelos ARIMA, Prophet e Regressão Linear

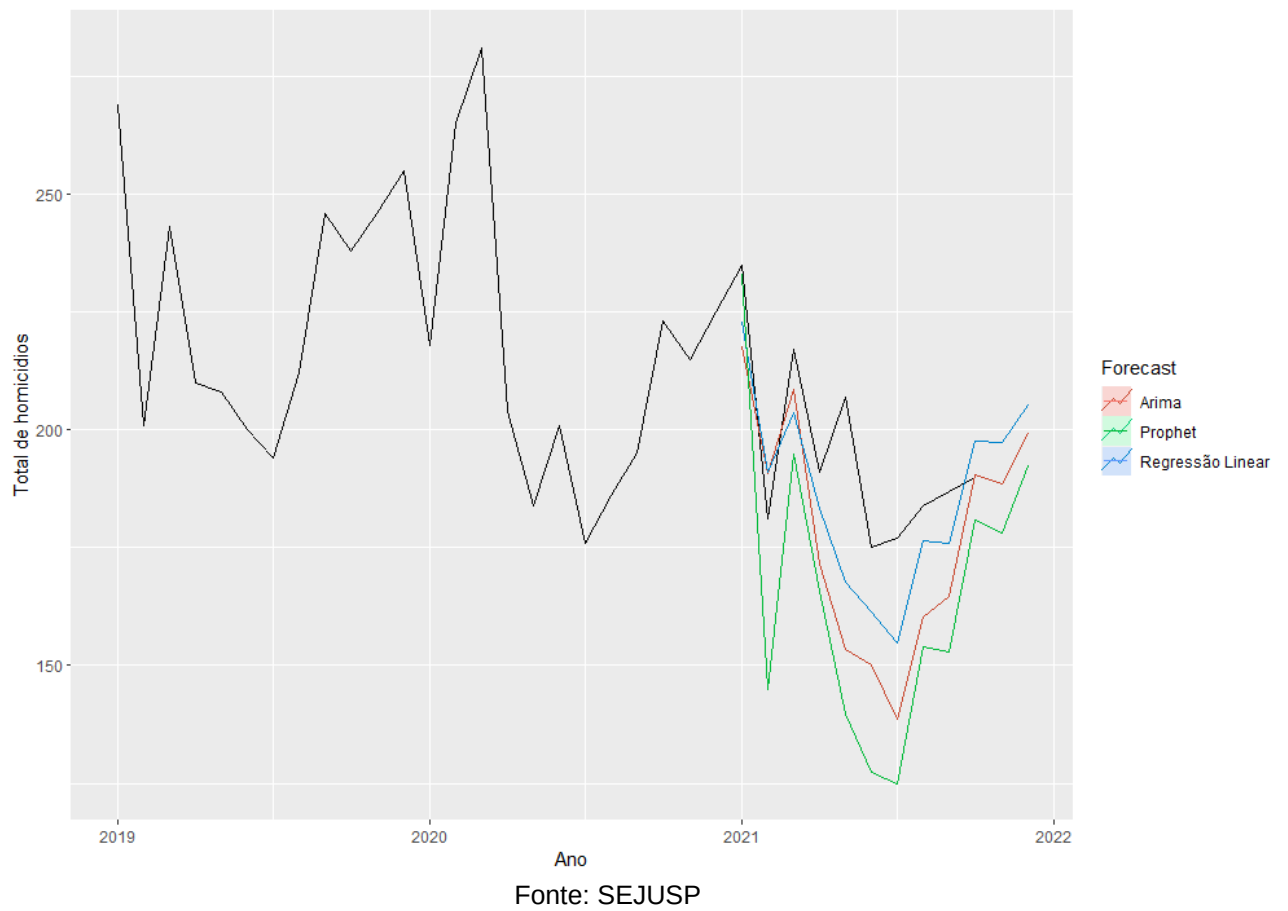


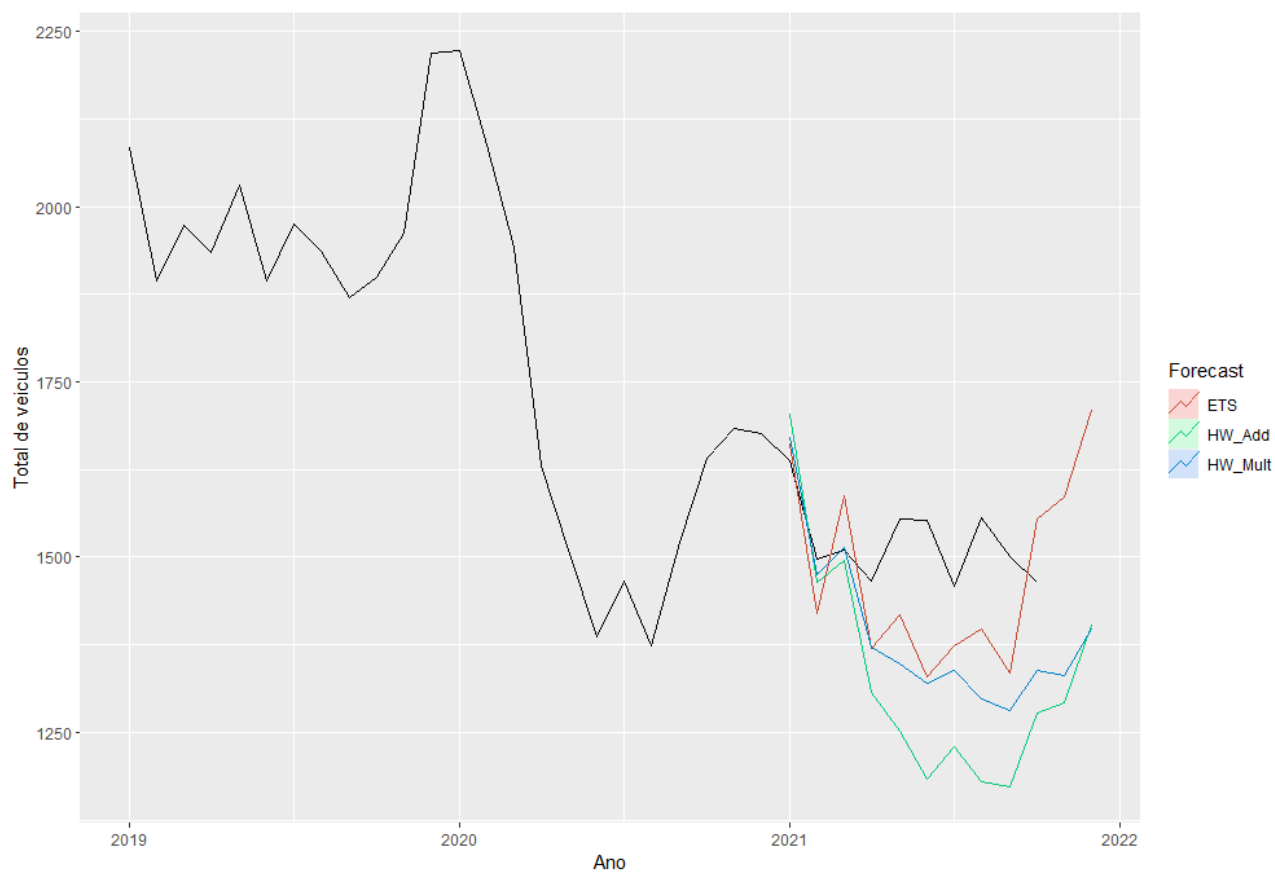
FIGURA 10 - Previsões para 2021 da série Vítimas de Homicídios Consumados utilizando modelos de Redes Neurais Artificiais de Camada Única (ANN Single Layer), Redes Neurais de Múltiplas Camadas (ANN Multi Layer) e Híbrido



Fonte: SEJUSP

## 5.2 Série Furto de Veículos

FIGURA 11 - Previsões para 2021 da série Furto de Veículos utilizando modelos Suavização Exponencial (ETS), Holt-Winters aditivo (HW\_Add) e Holt-Winters multiplicativo (HW\_Mult)



Fonte: SEJUSP

FIGURA 12 - Previsões para 2021 da série Furto de Veículos utilizando modelos ARIMA, Prophet e Regressão Linear



Fonte: SEJUSP

FIGURA 13 - Previsões para 2021 da série Furto de Veículos utilizando modelos de Redes Neurais Artificiais de Camada Única (ANN Single Layer), Redes Neurais de Múltiplas Camadas (ANN Multi Layer) e Híbrido

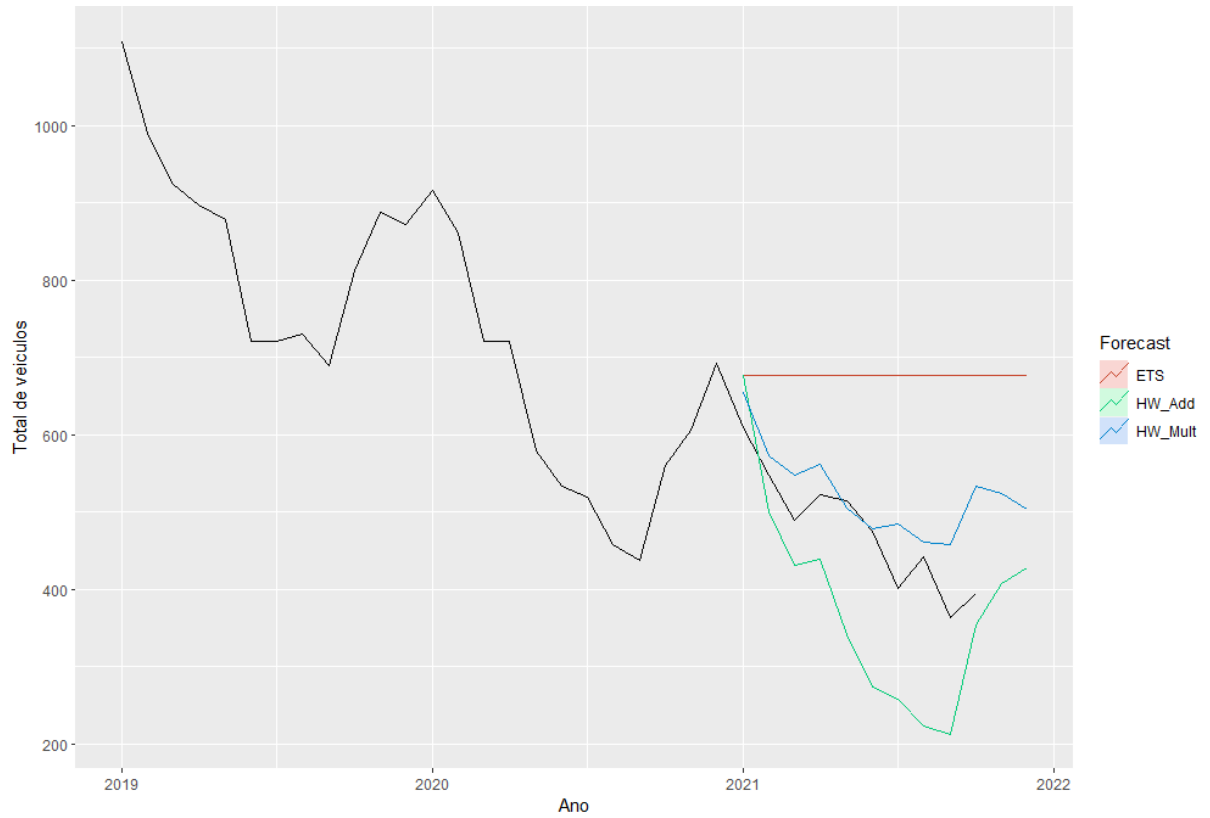


Fonte: SEJUSP



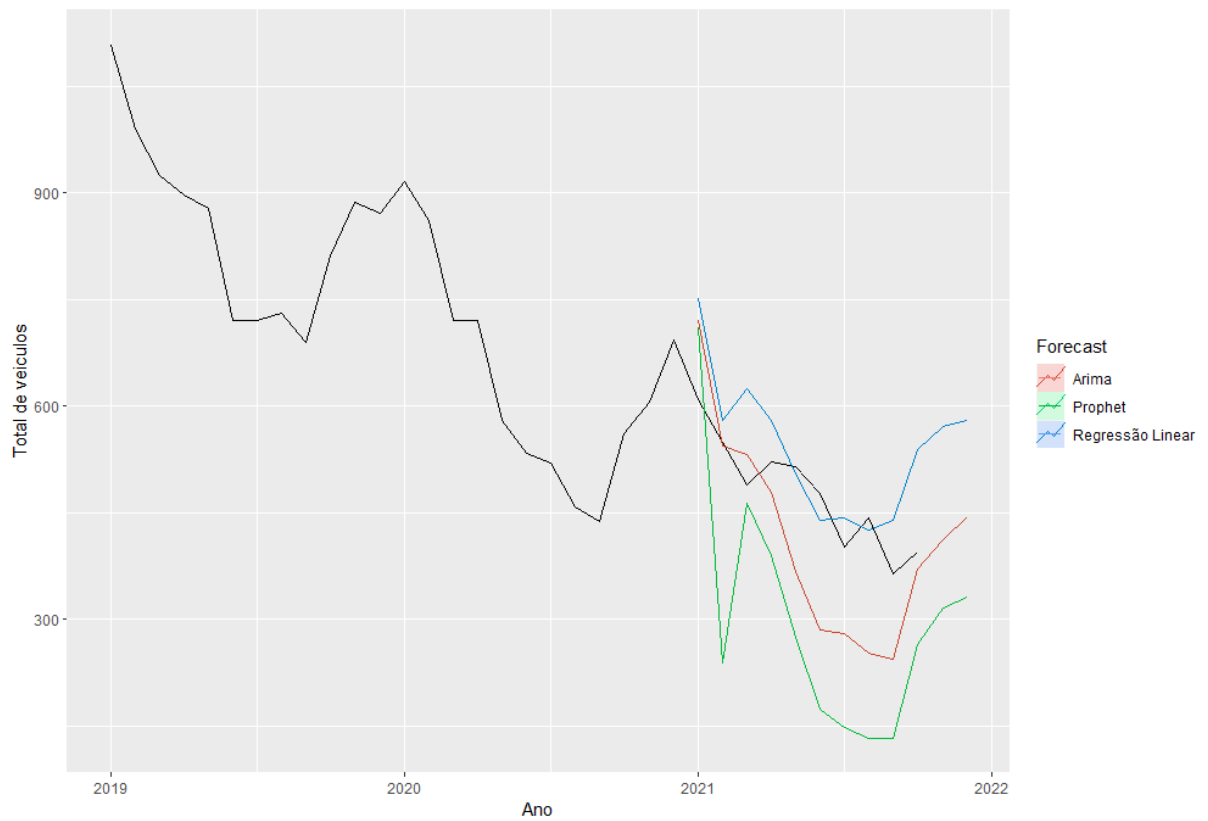
### 5.3 Série Roubo de Veículos:

FIGURA 14 - Previsões para 2021 da série Roubo de Veículos utilizando modelos Suavização Exponencial (ETS), Holt-Winters aditivo (HW\_Add) e Holt-Winters multiplicativo (HW\_Mult)



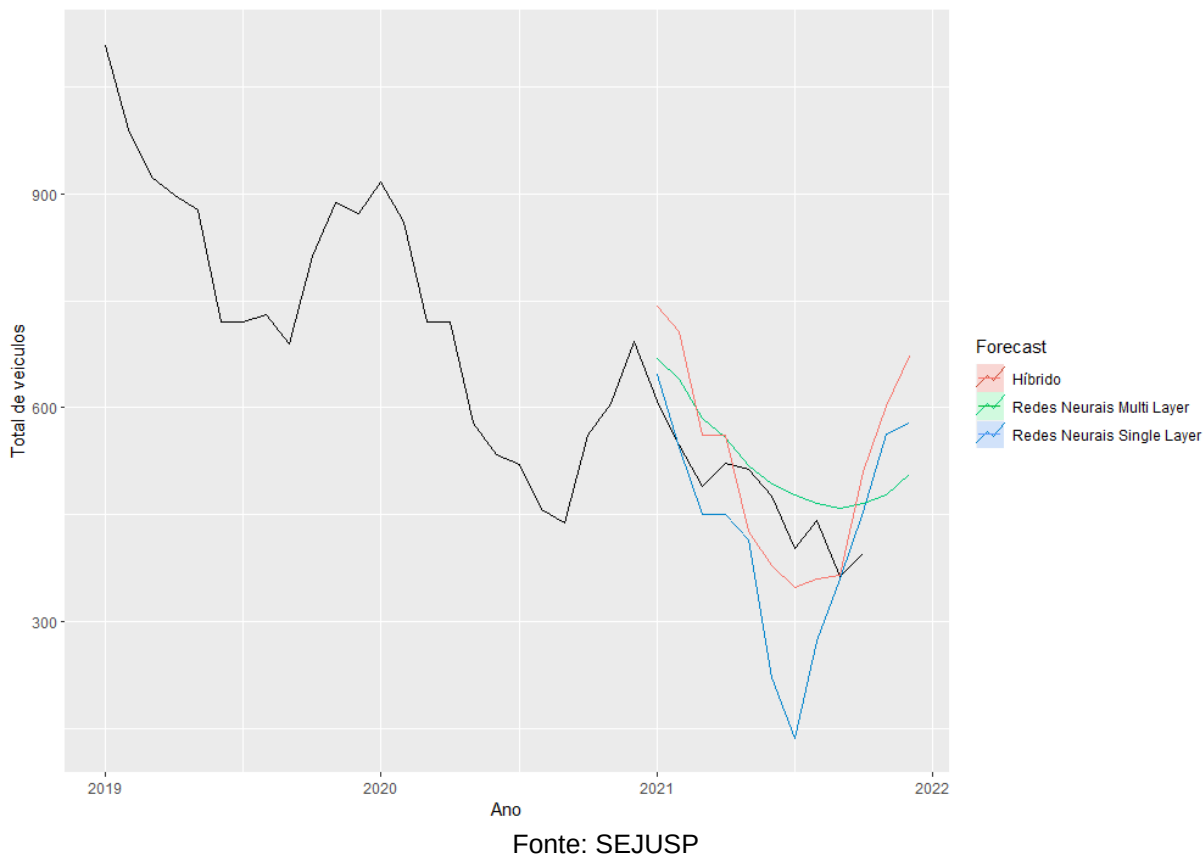
Fonte: SEJUSP

FIGURA 15 - Previsões para 2021 da série Roubo de Veículos utilizando modelos ARIMA, Prophet e Regressão Linear



Fonte: SEJUSP

FIGURA 16 - Previsões para 2021 da série Roubo de Veículos utilizando modelos de Redes Neurais Artificiais de Camada Única (ANN Single Layer), Redes Neurais de Múltiplas Camadas (ANN Multi Layer) e Híbrido



#### 5.4 Quadro Resumo da Performance dos Modelos

Para uma comparação objetiva entre os modelos, no QUAD. 3 são apresentados a raiz quadrada do erro quadrático médio (RMSE) e o erro absoluto percentual médio (MAPE) para cada modelo de cada série:

QUADRO 3 - Resumo da performance dos modelos de previsão para 12 passos à frente

| Modelo                                      | Homicídios |       | Furto de Veículos |       | Roubo de Veículos |       |
|---------------------------------------------|------------|-------|-------------------|-------|-------------------|-------|
|                                             | RMSE       | MAPE  | RMSE              | MAPE  | RMSE              | MAPE  |
| Modelo Linear                               | 17,2       | 7,43  | 148,83            | 8,76  | 84,78             | 14,78 |
| Suavização Exponencial                      | 19,99      | 7,97  | 125,86            | 7,48  | 213,98            | 45,7  |
| Holt-Winters Aditivo                        | 26,7       | 11,46 | 243,83            | 13,57 | 133,97            | 25,98 |
| Holt-Winters Multiplicativo                 | 21,05      | 8,65  | 158,72            | 8,64  | 65,45             | 11,97 |
| ARIMA                                       | 26,17      | 11,35 | 230,02            | 12,68 | 118,3             | 21,7  |
| Redes Neurais Artificiais - Camada única    | 24,36      | 11,68 | 302,28            | 17,7  | 52,45             | 8,8   |
| Redes Neurais Artificiais - Camada Múltipla | 21,31      | 9,14  | 570,79            | 34,9  | 153,73            | 29,13 |
| Prophet                                     | 37,52      | 17,27 | 215,05            | 12,43 | 224,59            | 44,41 |
| Híbrido                                     | 16,7       | 7,22  | 186,88            | 9,05  | 89,55             | 16,54 |
| Seasonal Naïve                              | 37,38      | 13,92 | 316,35            | 15,5  | 184,32            | 31,32 |

Fonte: Autor

## 6 CONCLUSÃO E CONSIDERAÇÕES FINAIS

Assim como pode ser observado nos dados apresentados no presente trabalho, todas as três séries escolhidas – Vítimas de Homicídios Consumados, Furto e Roubo de Veículos – apresentavam tendência de queda no período analisado, com níveis diferentes entre elas.

Durante a pandemia de COVID-19, uma das medidas aplicadas para evitar e frear o contágio de novas pessoas pela doença foi o distanciamento social associado a redução da circulação de pessoas. A partir de 2021, as medidas de distanciamento social tornavam-se mais ou menos rígidas, dependendo do cenário e andamento da doença e vacinação. Esse fato impactou, ainda que de forma heterogênea, as séries de criminalidade analisadas.

Os dados sugerem que as séries de crimes contra o patrimônio, ou seja, Roubo e Furto de Veículos, foram fortemente impactadas, com redução acentuada na primeira metade de 2020, especialmente, no que diz respeito aos furtos. Para a série de crime contra a vida, com vítimas de homicídios consumados, o impacto é menor. Isso pode ser visto através do componente aleatório na decomposição das séries.

Apesar deste cenário de grande incerteza, os modelos apresentados, em sua maioria, tiveram desempenho superior ao modelo ingênuo (Season Naïve).

Para homicídios, somente o modelo Prophet obteve índices de erro maiores que o modelo ingênuo. Nos demais modelos, destaque para o modelo Híbrido, sendo de Suavização Exponencial para a parte linear e de Redes Neurais Artificiais para a parte não-linear, obteve-se a raiz do erro quadrático médio (RMSE) de 16,7 e erro absoluto percentual médio de 7,22%, sendo estes, os menores entre todos os modelos apresentados.

Quanto ao furto de veículos, somente os modelos de Redes Neurais Artificiais obtiveram índices piores que o modelo ingênuo, sendo que, o de camadas múltiplas, resultaram em um pior desempenho. Do restante, todos apresentaram métricas menores, destacando-se o modelo de Suavização Exponencial, com MAPE de 7,48% e RMSE de 125,86.

Além disso, no que diz respeito ao roubo de veículos, os modelos de Suavização Exponencial e Prophet apresentaram índices piores se comparado ao modelo ingênuo. O melhor desempenho foi obtido através das Redes Neurais com

Camada Única que apresentou um RMSE de apenas 52,45 e MAPE de 8,8%; frente aos 184,32 e 31,32%, respectivamente, do modelo ingênuo.

Na comparação geral entre as séries, vale notar que os seguintes modelos sempre apresentaram desempenho melhor que o modelo ingênuo: Modelo Linear, Holt-Winters aditivo, Holt-Winters multiplicativo, ARIMA e Híbrido. Dentre eles, destaca-se o modelo de Holt-Winters multiplicativo que, embora não tenha sido o modelo de melhor performance nas séries individuais, obteve índices baixos de erros em todas as séries trabalhadas.

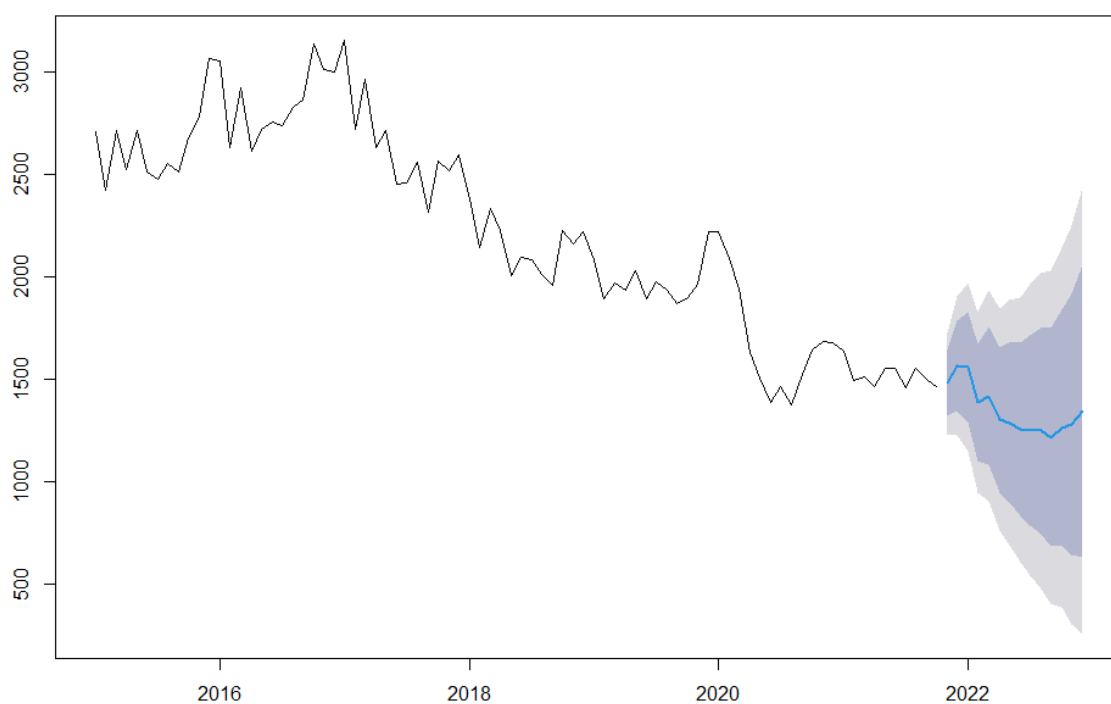
Assim, pode-se concluir que os modelos de previsão avaliados fornecem projeções para séries de criminalidade com um bom nível de acurácia, mesmo em ambientes de incerteza, como é o caso do período analisado. Isso possibilita fornecer aos agentes públicos e privados insumos para a tomada de decisão, indicando cenários possíveis e auxiliando no planejamento.

A análise realizada no presente trabalho não descarta a necessidade de estudos relacionados ao tema. Refinar os modelos aqui mencionados, adicionar variáveis explicativas, incluir novos modelos e avaliá-los em locais e períodos diferentes são exemplos de novos trabalhos a serem executados.

## 7 PREVISÃO PARA 2022

Por conta dos bons resultados obtidos, foi escolhido o modelo de Holt-Winters multiplicativo para realizar previsões para os meses de 2022 das três séries escolhidas. Para aumentar a acurácia do modelo, foram incluídos os dados disponíveis de 2021 (Jan-Out).

FIGURA 17- Previsões para 2022 da série Furto de Veículos utilizando modelo de Holt-Winters Multiplicativo



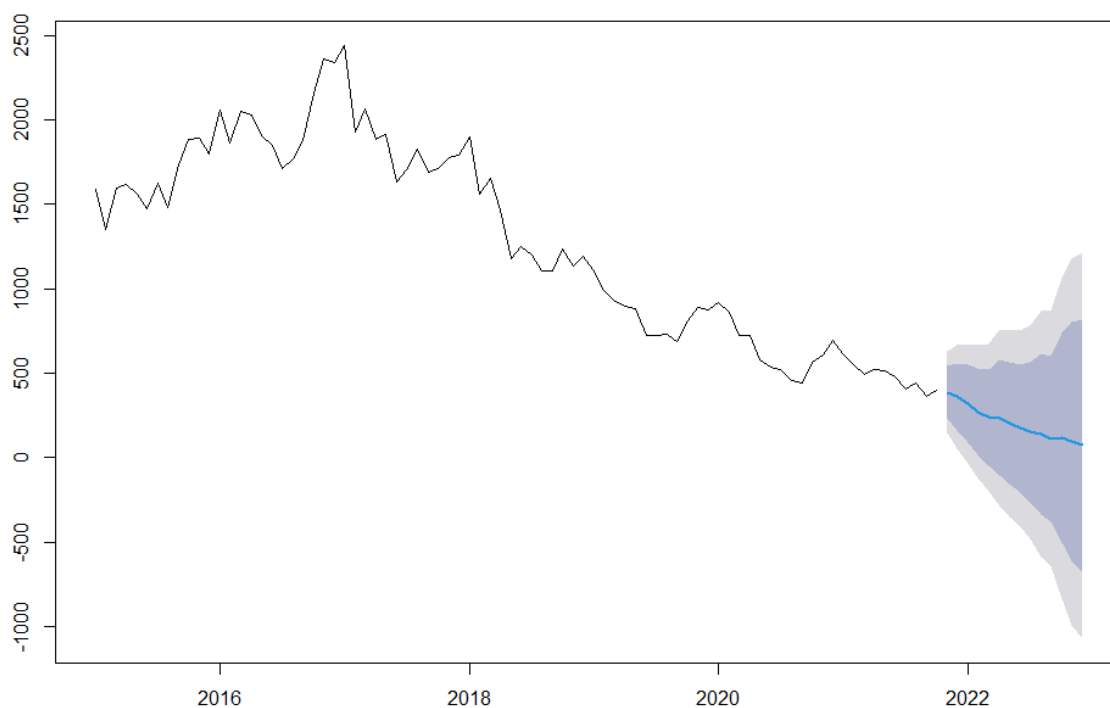
Fonte: SEJUSP

QUADRO 4 - Previsões para 2022 da série Furto de Veículos utilizando modelo de Holt-Winters Multiplicativo

| Período  | Forecast |
|----------|----------|
| jan/2022 | 1558     |
| fev/2022 | 1387     |
| mar/2022 | 1420     |
| abr/2022 | 1304     |
| mai/2022 | 1288     |
| jun/2022 | 1255     |
| jul/2022 | 1253     |
| ago/2022 | 1250     |
| set/2022 | 1218     |
| out/2022 | 1264     |
| nov/2022 | 1278     |
| dez/2022 | 1348     |

Fonte: Autor

FIGURA 18- Previsões para 2022 da série Roubo de Veículos utilizando modelo de Holt-Winters Multiplicativo



Fonte: SEJUSP

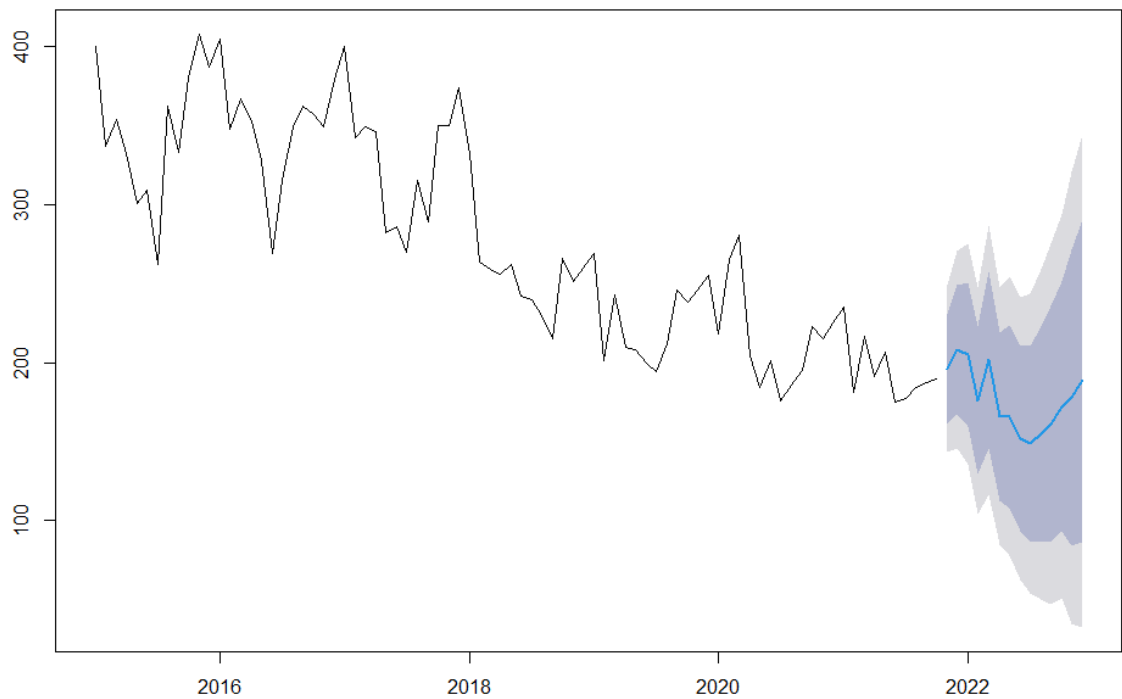


QUADRO 5- Previsões para 2022 da série Roubo de Veículos utilizando modelo de Holt-Winters Multiplicativo

| Período  | Forecast |
|----------|----------|
| jan/2022 | 317      |
| fev/2022 | 270      |
| mar/2022 | 239      |
| abr/2022 | 235      |
| mai/2022 | 201      |
| jun/2022 | 171      |
| jul/2022 | 150      |
| ago/2022 | 139      |
| set/2022 | 113      |
| out/2022 | 115      |
| nov/2022 | 95       |
| dez/2022 | 70       |

Fonte: Autor

FIGURA 19- Previsões para 2022 da série Vítimas de Homicídios Consumados utilizando modelo de Holt-Winters Multiplicativo



Fonte: SEJUSP

QUADRO 6 - Previsões para 2022 da série Vítimas de Homicídios Consumados utilizando modelo de Holt-Winters Multiplicativo

| Período  | Forecast |
|----------|----------|
| jan/2022 | 205      |
| fev/2022 | 176      |
| mar/2022 | 202      |
| abr/2022 | 166      |
| mai/2022 | 166      |
| jun/2022 | 152      |
| jul/2022 | 149      |
| ago/2022 | 155      |
| set/2022 | 161      |
| out/2022 | 172      |
| nov/2022 | 178      |
| dez/2022 | 189      |

Fonte: Autor

## REFERÊNCIAS

- ALMEIDA, Lucas Butzer. **Criminalidade no Brasil: Um Olhar Econômico sobre suas Causas e Consequências**. 2016. Monografia (Graduação em Economia) - Universidade de Campinas, Instituto de Economia, Campinas, 2016.
- ATHANASOPOULOUS, George *et al.* **The Tourism forecasting competition**. International Journal of Forecasting, vol. 27, n. 3, p. 822-844, 2011.
- AZEVEDO, Sérgio Henrique Andrade de. **Avaliação dos modelos preditivos de risco financeiros aplicados à realidade brasileira**. 2019. Monografia (Graduação em Economia) - Universidade de Campinas, Instituto de Economia, Campinas, 2017.
- BARROS, Anna Carolina *et al.* **Análise de séries temporais em R: curso introdutório**. Rio de Janeiro: Elsevier, 2018.
- BECKER, Gary S. Crime and punishment: An economic approach. *In*: Fielding, Nigel G.; Clarke, Alan; Witt, Robert. **The economic dimensions of crime**. London: Palgrave Macmillan, 1968, p. 13-68.
- BECKER, Kalinca Léia. **Uma análise econômica da relação entre a educação e a violência**. 2012. Tese (Doutorado em Economia Aplicada) - Universidade de São Paulo, Escola Superior de Agricultura Luiz de Queiroz, 2012.
- BOONE, Tonya *et al.* **Forecasting sales in the supply chain: Consumer analytics in the big data era**. International Journal of Forecasting, v. 35, n. 1, p. 170-180, 2019.
- BOX, G. E. P.; JENKINS, G.M.; REINSEL, G.C. **Time Series Analysis: Forecasting and Control**. Wiley, Oxford. 2008.
- CARVALHO, Alexandre Xavier YWATA de *et al.* **Custos das mortes por causas externas no Brasil**. Rev. Bras. Biom, v. 26, n. 3, p. 23-47, 2008.
- CHATFIELD, C. **Time-series Forecasting**. London: Chapman & Hall/CRC, 2000.
- CHAUDHURI, Tamal Datta; GHOSH, Indranil. **Artificial Neural Network and Time Series Modeling Bases Approach to Forecasting the Exchange Rate in a Multivariate Framework**. Journal of Insurance and Financial Management, v. 1, p. 92-123. 2016.
- CHEN, Peng; YUAN, Hongyong; SHU, Xueming. Forecasting crime using the arima model. *In*: **Fifth International Conference on Fuzzy Systems and Knowledge Discovery**. IEEE, 2008. p. 627-630.
- COGHLAN, Avril. **A little book of R for time series**. Published under Creative Commons Attribution, v. 3, 2015.
- DHIRI, Sanjay *et al.* **Modelling and predicting property crime trends in England and Wales**. London: Home Office, 1999.

EHRLICH, Isaac. **Participation in illegitimate activities: A theoretical and empirical investigation**. Chicago: Journal of Political Economy, v. 81, n. 3, p.526-536, 1973.

FAJNZYLBER, Pablo; LEDERMAN, Daniel; LOAYZA, Norman. **Determinants of Crime Rates in Latin America and the World: Viewpoints**. Washington: The World Bank, 1998.

GORR, Wilpen; HARRIES, Richard. **Introduction to crime forecasting**. International Journal of Forecasting, v. 19, n. 4, p. 551-555, 2003.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning**. [S.l.]: Springer, 2009.

HYNDMAN, Rob J.; ATHANASOPOULOS, George. **Forecasting: principles and practice**. Melbourne: OTexts, ed. 3, 2018.

MOREIRA, Gustavo Carvalho; KASSOUF, Ana Lucia; JUSTUS, Marcelo. **The role of social capital in the victimization risk against property: evidence from Brazil**. Economia e Sociedade, v. 28, p. 563-585, 2019.

MURRAY, Joseph; DE CASTRO CERQUEIRA, Daniel Ricardo; KAHN, Tulio. **Crime and violence in Brazil: Systematic review of time trends, prevalence rates and risk factors**. Aggression and Violent Behavior, v. 18, p. 471-483, 2013.

NAVEENA, K. *et al.* **Hybrid ARIMA-ANN modelling for forecasting the price of Robusta coffee in India**. Int. J. Curr. Microbiol. Appl. Sci, v. 6, n. 7, p. 1721-1726, 2017.

OLIVEIRA, Henrique Fontaneda. **A Renda como uma das Determinantes do Crime**. 2017. Monografia (Graduação em Economia) - Universidade de Campinas, Instituto de Economia, Campinas, 2017.

PYNDYCK, R. S.; RUBENFIELD, D. L. **Econometric Models and Economic Forecasts**. New York: McGrawHill, ed. 3, 1991.

RODRIGUES, Lucas; RODRIGUES, Luciano. **Métodos de Previsão Aplicados ao Consumo Brasileiro de Combustíveis Leves**. Foz do Iguaçu: Cooperativismo, Inovação e Sustentabilidade para o Desenvolvimento Rural, 2020.

ROSTAMI-TABAR, Bahman *et al.* **Forecasting for social good**. International Journal of Forecasting, 2021.

SANTOS, Marcelo Justus dos. **Dinâmica temporal da criminalidade: mais evidências sobre o “efeito inércia” nas taxas de crimes letais nos estados brasileiros**. Revista EconomiA, v. 10, p. 169-194, 2009.

SINGH, Prakash. **Time series Forecasting on crime data in Amsterdam for a Software Company**. 2018. Dissertação (Mestrado em Métodos Analíticos Avançados) - Information Management School, Universidade Nova de Lisboa, 2018.

SOUZA, Gueibi Peres; SAMOBYL, Robert Wayne; MEURER, Roberto. Previsão do consumo de energia elétrica do setor industrial em Santa Catarina—um estudo comparativo entre diferentes métodos de previsão através de suas discrepâncias. **XXXVI Simpósio Brasileiro de Pesquisa Operacional-O Impacto da Pesquisa Operacional nas Novas Tendências Multidisciplinares**, São João Del Rei, MG, 2004.

STEFANI, Kenny Zoia. **Educação, Trabalho e Violência: Análise da Relação entre proporção de jovens “nem-nem” e taxa de homicídios no Brasil**. 2017. Monografia (Graduação em Ciências Econômicas) - Universidade Estadual de Campinas, 2017.

STEINBUKS, Jevgenijs. **Assessing the accuracy of electricity production forecasts in developing countries**. International Journal of Forecasting, v. 35, n. 3, p. 1175-1185, 2019.

TAYLOR, Sean J.; LETHAM, Benjamin. **Forecasting at scale**. The American Statistician, v. 72, n. 1, p. 37-45, 2018.

## APÊNDICE A-CÓDIGOS R STUDIO: Série Homicídios

```
#### MONOGRAFIA - METODOS E MODELOS DE PREVISAO PARA SERIES TEMPORAIS DE
CRIMINALIDADE EM MINAS GERAIS ####
#### ALUNO: WILLIAN HIDEKI TENGUAN ####
#### RA: 178735 ####
#### ORIENTADOR: PROF. DR. MARCELO JUSTUS DOS SANTOS ####
#### PARTE 1: VITIMAS DE HOMICIDIOS CONSUMADOS ####
#### CRIADO EM: 21/11/2021 ####
```

```
# CARREGA OS PACOTES NECESSARIOS AOS MODELOS #
```

```
library(RODBC)
library(readxl)
library(forecast)
library(TSstudio)
library(tidyverse)
library(nnfor)
library(prophet)
library(lubridate)
library(zoo)
```

```
# DEFINE O BANCO DE DADOS JA NORMALIZADO #
```

```
con <- odbcConnect("ToyDB")
setwd("D:/UNICAMP/MONOGRAFIA/TECNICAS/DATABASES")
```

```
# CARREGA OS DADOS RELATIVOS AS VITIMAS DE HOMICIDIOS CONSUMADOS #
```

```
qry_homic <- "SELECT ANO, MES, SUM(Registros) AS QtdRegistros
              FROM MG_HOMICIDIOS_1
              WHERE ANO >= 2015
              GROUP BY ANO, MES
              ORDER BY ANO, MES"
homic_tbl <- sqlQuery(con, qry_homic)
str(homic_tbl)
list(homic_tbl)
```

```
# CRIA OS OBJETOS TIME-SERIES NECESSARIOS
```

```
qry_homic <- "SELECT ANO, MES, SUM(Registros) AS QtdRegistros
              FROM MG_HOMICIDIOS_1
              WHERE ANO >= 2015
              GROUP BY ANO, MES
              ORDER BY ANO, MES"
homic_tbl <- sqlQuery(con, qry_homic)
str(homic_tbl)
list(homic_tbl)
```

```
homic_ts <- ts(homic_tbl$QtdRegistros, start = c(2015,1), frequency = 12)
homic_ts_diff <- diff(homic_ts, 1)
```

```
plot(homic_ts)
```

```
# CRIA AS TIMESERIES DE TREINO E TESTE
```

```

homic_train_ts <- window(homic_ts, end = c(2020, 12))
homic_test_ts <- window(homic_ts, start = c(2021,1))

# CRIA AS TS DE TREINO E TESTE DIFERENCIADAS
homic_train_diff <- diff(homic_train, 1)
homic_test_diff <- diff(homic_test, 1)

# PLOTA A SERIE PARA ANALISE PREVIA

ts.plot(homic_ts, type = "o", lwd="2", col="2")
lines(lowess(homic_ts, f=1/12), col=4, lwd="2", type="l")
plot(decompose(homic_ts))

plot(homic_ts_diff)

## Tendencia de queda pronunciada a partir de 2018.
## Apareta haver forte sazonalidade, principalmente nos comecos e fins de anos.
## A tendencia de queda se mantem, visualmente, ao mesmo nivel antes e depois da pandemia de
COVID-19.

# GERA DATAFRAME PARA O MODELO PROPHET #
ds <- as.Date(paste(homic_tbl$ANO, month=homic_tbl$MES, 1, sep="-"), "%Y-%m-%d")
y <- homic_tbl$QtdRegistros
homic_train_df <- subset(data.frame(ds, y), ds <= '2020-12-01')
homic_test_df <- subset(data.frame(ds, y), ds > '2020-12-01')

# GERA SERIES PARA O MODELO HIBRIDO #

homic_linear <- decompose(homic_ts)$trend
homic_linear_train <- window(homic_linear, end=c(2020, 12))
homic_linear_test <- window(homic_linear, start=c(2021, 1))
plot(homic_trend)
homic_naolinear <- decompose(homic_ts)$seasonal+decompose(homic_ts)$random
homic_naolinear_train <- window(homic_naolinear, end=c(2020,12))
homic_naolinear_test <- window(homic_naolinear, start=c(2021,1))
plot(homic_naolinear)

### MODELOS ###

# MODELO LINEAR APLICADO A SERIES TEMPORAIS #

homic_forecast_tslm <- tslm(homic_train_ts ~ trend + season, lambda="auto")
homic_forecast_tslm_2 <- forecast(homic_forecast_tslm, h=12)

homic_forecast_tslm$coefficients

?tslm

autoplot(homic_ts) +
  autolayer(homic_forecast_tslm_2, series="Modelo de Regressao Linear", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de homicidios") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

```

```
summary(homic_forecast_tslm_2)
a <- accuracy(homic_forecast_tslm_2, x = homic_test_ts)
```

```
# MODELO DE SUAVIZACAO EXPONENCIAL #
homic_forecast_ets <- ets(homic_train_ts, model="ZZZ", opt.crit = "mse")
homic_forecast_ets_2 <- forecast(homic_forecast_ets, h=12)
```

```
autoplot(homic_ts) +
  autolayer(homic_forecast_ets_2, series="Suavização Exponencial", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de homicídios") +
  ggtitle("Homicídios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(homic_forecast_ets_2)
accuracy(homic_forecast_ets_2, x=homic_test_ts)
```

```
# HOLT-WINTERS - ADDITIVE #
```

```
homic_forecast_hw_add <- HoltWinters(homic_train_ts, beta=NULL, gamma=TRUE, seasonal =
"additive")
homic_forecast_hw_add_2 <- forecast(homic_forecast_hw_add, h=12)
```

```
autoplot(homic_ts) +
  autolayer(homic_forecast_hw_add_2, series="HW additive forecasts", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de homicídios") +
  ggtitle("Homicídios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(homic_forecast_hw_add_2)
accuracy(homic_forecast_hw_add_2, x=homic_test_ts)
```

```
# HOLT-WINTERS - MULTIPLICATIVE #
```

```
homic_forecast_hw_mult <- HoltWinters(homic_train_ts, beta=NULL, gamma=TRUE, seasonal =
"multiplicative")
homic_forecast_hw_mult_2 <- forecast(homic_forecast_hw_mult, h=12)
```

```
autoplot(homic_ts) +
  autolayer(homic_forecast_hw_mult_2, series="HW multiplicative forecasts", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de homicídios") +
  ggtitle("Homicídios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(homic_forecast_hw_mult_2)
accuracy(homic_forecast_hw_mult_2, x=homic_test_ts)
```

```
# ARIMA #
```

```
homic_forecast_arima <- auto.arima(homic_train_ts)
homic_forecast_arima_2 <- forecast(homic_forecast_arima, h=12)
```

```
autoplot(homic_ts) +
```



```

autolayer(homic_forecast_arima_2, series="Arima", PI=FALSE) +
xlab("Ano") +
ylab("Total de homicídios") +
ggtitle("Homicídios consumados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(homic_forecast_arima_2)
accuracy(homic_forecast_arima_2, x=homic_test_ts)

# Redes Neurais Artificiais - ANN - Single Layer #

homic_forecast_nnetar <- nnetar(homic_train_ts)
homic_forecast_nnetar_2 <- forecast(homic_forecast_nnetar, h=12)

autoplot(homic_ts) +
autolayer(homic_forecast_nnetar_2, series="Redes Neurais Artificiais - Uma camada", PI=FALSE) +
xlab("Ano") +
ylab("Total de homicídios") +
ggtitle("Homicídios consumados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(homic_forecast_nnetar_2)
accuracy(homic_forecast_nnetar_2, x=homic_test_ts)

# Redes Neurais Artificiais - ANN - Multi Layer #

homic_forecast_ann <- mlp(homic_train_ts)
homic_forecast_ann_2 <- forecast(homic_forecast_ann, h=12)

autoplot(homic_ts) +
autolayer(homic_forecast_ann_2, series="Redes Neurais Artificiais", PI=FALSE) +
xlab("Ano") +
ylab("Total de homicídios") +
ggtitle("Homicídios consumados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(homic_forecast_ann_2)
accuracy(homic_forecast_ann_2, x=homic_test_ts)

# Prophet #

homic_forecast_prophet <- prophet(homic_train_df)
Future1 <- make_future_dataframe(homic_forecast_prophet, periods = 12, freq='month')
Forecast1 <- predict(homic_forecast_prophet, Future1)
tail(Forecast1[c('ds', 'yhat')], n=12)
summary(Forecast1)

homic_forecast_prophet_2 <- ts(tail(Forecast1$yhat, 12), start = c(2021,1), frequency=12)

autoplot(homic_ts) +
autolayer(homic_forecast_prophet_2, series="Prophet") +
xlab("Ano") +
ylab("Total de homicídios") +
ggtitle("Homicídios consumados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

```

```
summary(homic_forecast_prophet_2)
accuracy(homic_forecast_prophet_2, x=homic_test_ts)
```

```
dyplot.prophet(homic_forecast_prophet, Forecast1)
summary(homic_forecast_prophet)
```

```
# Modelo Hibrido: ETS para linear, ANN Sigle Layer para nao-linear #
```

```
homic_forecast_hib_linear <- ets(na.omit(homic_linear_train))
homic_forecast_hib_linear_2 <- forecast(homic_forecast_hib_linear, h=12)
accuracy(homic_forecast_hib_linear_2, x=homic_linear_test)
```

```
autoplot(na.omit(homic_linear)) +
  autolayer(homic_forecast_hib_linear_2, series="a") +
  xlab("Ano") +
  ylab("Total de homicidios") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
homic_forecast_hib_nlinear <- nnetar(na.omit(homic_naolinear_train))
homic_forecast_hib_nlinear_2 <- forecast(homic_forecast_hib_nlinear, h=12)
accuracy(homic_forecast_hib_nlinear_2, x=homic_naolinear_test)
```

```
autoplot(na.omit(homic_naolinear)) +
  autolayer(homic_forecast_hib_nlinear_2, series="a") +
  xlab("Ano") +
  ylab("Total de homicidios") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
homic_forecast_hibrido <- homic_forecast_hib_linear_2$mean + homic_forecast_hib_nlinear_2$mean
```

```
accuracy(homic_forecast_hibrido, x=homic_test_ts)
```

```
autoplot(homic_ts) +
  autolayer(homic_forecast_hibrido, series="Modelo Hibrido") +
  xlab("Ano") +
  ylab("Total de homicidios") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
# Modelo snaive de comparacao
```

```
homic_forecast_snaive <- snaive(homic_train_ts)
homic_forecast_snaive_2 <- forecast(homic_forecast_snaive, h=12)
accuracy(homic_forecast_snaive_2, x=homic_test_ts)
```

```
# Plotando todos os modelos
```

```
plot_homic <- window(homic_ts, start = c(2019,1))
```

```
# Modelos ETS e HW
```

```

autoplot(plot_homic) +
  autolayer(homic_forecast_ets_2, series="ETS", PI=FALSE) +
  autolayer(homic_forecast_hw_add_2, series="HW_Add", PI=FALSE) +
  autolayer(homic_forecast_hw_mult_2, series="HW_Mult", PI=FALSE) +

  xlab("Ano") +
  ylab("Total de homicídios") +
  ggtitle("Homicídios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

```

# Modelos Arima, Linear e Prophet

```

autoplot(plot_homic) +
  autolayer(homic_forecast_arima_2, series="Arima", PI=FALSE) +
  autolayer(homic_forecast_tslm_2, series="Regressão Linear", PI=FALSE) +
  autolayer(homic_forecast_prophet_2, series="Prophet", PI=FALSE) +

  xlab("Ano") +
  ylab("Total de homicídios") +
  ggtitle("Homicídios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

```

# Modelos ANN e Híbrido

```

autoplot(plot_homic) +
  autolayer(homic_forecast_ann_2, series="ANN Single Layer", PI=FALSE) +
  autolayer(homic_forecast_nnetar_2, series="ANN Multi Layer", PI=FALSE) +
  autolayer(homic_forecast_hibrido, series="Híbrido", PI=FALSE) +

  xlab("Ano") +
  ylab("Total de homicídios") +
  ggtitle("Homicídios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

```

## PREVISAO DE 2022 USANDO HW MULT ##

```

homic_forecast_2022 <- forecast(HoltWinters(homic_ts, beta=NULL, gamma=TRUE, seasonal =
"multiplicative"), h=14)

```

```

plot(homic_forecast_2022,
  main = 'Previsão para 2022: Vítimas de Homicídios Consumados')

```

## APÊNDICE B-CÓDIGOS R STUDIO: Série Furtos

```
#### MONOGRAFIA - METODOS E MODELOS DE PREVISAO PARA SERIES TEMPORAIS DE
CRIMINALIDADE EM MINAS GERAIS ####
#### ALUNO: WILLIAN HIDEKI TENGUAN ####
#### RA: 178735 ####
#### ORIENTADOR: PROF. DR. MARCELO JUSTUS DOS SANTOS ####
#### PARTE 2: FURTO DE VEICULOS ####
#### CRIADO EM: 21/11/2021 ####
```

```
# CARREGA OS PACOTES NECESSARIOS AOS MODELOS #
```

```
library(RODBC)
library(readxl)
library(forecast)
library(TSstudio)
library(tidyverse)
library(nnfor)
library(prophet)
library(lubridate)
library(zoo)
```

```
# DEFINE O BANCO DE DADOS JA NORMALIZADO #
```

```
con <- odbcConnect("ToyDB")
setwd("D:/UNICAMP/MONOGRAFIA/TECNICAS/DATABASES")
```

```
# CARREGA OS DADOS RELATIVOS AS VITIMAS DE HOMICIDIOS CONSUMADOS #
```

```
qry_furto <- "SELECT ANO, MES, SUM(Registros) AS QtdRegistros
              FROM MG_VEICULOSFURTADOS
              WHERE ANO >= 2015
              GROUP BY ANO, MES
              ORDER BY ANO, MES"
furto_tbl <- sqlQuery(con, qry_furto)
str(furto_tbl)
list(furto_tbl)
```

```
furto_ts <- ts(furto_tbl$QtdRegistros, start = c(2015,1), frequency = 12)
furto_ts_diff <- diff(furto_ts, 1)
```

```
plot(furto_ts)
```

```
# CRIA AS TIMESERIES DE TREINO E TESTE #
```

```
furto_train_ts <- window(furto_ts, end = c(2020, 12))
furto_test_ts <- window(furto_ts, start = c(2021,1))
```

```
# CRIA AS TS DE TREINO E TESTE DIFERENCIADAS #
```

```
furto_train_diff <- diff(furto_train_ts, 1)
furto_test_diff <- diff(furto_test_ts, 1)
```

```
# PLOTA A SERIE PARA ANALISE PREVIA #
```

```
ts.plot(furto_ts, type = "o", lwd="2", col="2")
```

```
lines(lowess(furto_ts, f=1/12), col=4, lwd="2", type="l")
plot(decompose(furto_ts))
```

```
plot(furto_ts_diff)
```

```
## Tendência de queda pronunciada a partir de 2018.
## Aparenta haver forte sazonalidade, principalmente nos comecos e fins de anos.
## A tendência de queda se mantem, visualmente, ao mesmo nivel antes e depois da pandemia de COVID-19.
```

```
# GERA DATAFRAME PARA O MODELO PROPHET #
ds <- as.Date(paste(furto_tbl$ANO, month=furto_tbl$MES, 1, sep="-"), "%Y-%m-%d")
y <- furto_tbl$QtdRegistros
furto_train_df <- subset(data.frame(ds, y), ds <= '2020-12-01')
furto_test_df <- subset(data.frame(ds, y), ds > '2020-12-01')
```

```
# GERA SERIES PARA O MODELO HIBRIDO #
```

```
furto_linear <- decompose(furto_ts)$trend
furto_linear_train <- window(furto_linear, end=c(2020, 12))
furto_linear_test <- window(furto_linear, start=c(2021, 1))
plot(furto_linear)
furto_naolinear <- decompose(furto_ts)$seasonal+decompose(furto_ts)$random
furto_naolinear_train <- window(furto_naolinear, end=c(2020,12))
furto_naolinear_test <- window(furto_naolinear, start=c(2021,1))
plot(furto_naolinear)
```

```
## MODELOS ###
```

```
# MODELO LINEAR APLICADO A SERIES TEMPORAIS #
```

```
furto_forecast_tslm <- tslm(furto_train_ts ~ trend + season, lambda="auto")
furto_forecast_tslm_2 <- forecast(furto_forecast_tslm, h=12)
```

```
furto_forecast_tslm$coefficients
```

```
autoplot(furto_ts) +
  autolayer(furto_forecast_tslm_2, series="Modelo de Regressao Linear", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Furtados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(furto_forecast_tslm_2)
accuracy(furto_forecast_tslm_2, x = furto_test_ts)
```

```
# MODELO DE SUAVIZACAO EXPONENCIAL #
```

```
furto_forecast_ets <- ets(furto_train_ts, model="ZZZ", opt.crit = "mse")
furto_forecast_ets_2 <- forecast(furto_forecast_ets, h=12)
```

```
autoplot(furto_ts) +
  autolayer(furto_forecast_ets_2, series="Suavização Exponencial", PI=FALSE) +
```

```

xlab("Ano") +
ylab("Total de veiculos") +
ggtitle("Veiculos Furtados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_ets_2)
accuracy(furto_forecast_ets_2, x=furto_test_ts)

# HOLT-WINTERS - ADDITIVE #

furto_forecast_hw_add <- HoltWinters(furto_train_ts, beta=NULL, gamma=TRUE, seasonal =
"additive")
furto_forecast_hw_add_2 <- forecast(furto_forecast_hw_add, h=12)

autoplot(furto_ts) +
autolayer(furto_forecast_hw_add_2, series="HW additive forecasts", PI=FALSE) +
xlab("Ano") +
ylab("Total de veiculos") +
ggtitle("Homicidios consumados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_hw_add_2)
accuracy(furto_forecast_hw_add_2, x=furto_test_ts)

# HOLT-WINTERS - MULTIPLICATIVE #
furto_forecast_hw_mult <- HoltWinters(furto_train_ts, beta=NULL, gamma=TRUE, seasonal =
"multiplicative")
furto_forecast_hw_mult_2 <- forecast(furto_forecast_hw_mult, h=12)

autoplot(furto_ts) +
autolayer(furto_forecast_hw_mult_2, series="HW multiplicative forecasts", PI=FALSE) +
xlab("Ano") +
ylab("Total de veiculos") +
ggtitle("Veiculos Furtados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_hw_mult_2)
accuracy(furto_forecast_hw_mult_2, x=furto_test_ts)

# ARIMA #

furto_forecast_arima <- auto.arima(furto_train_ts)
furto_forecast_arima_2 <- forecast(furto_forecast_arima, h=12)

autoplot(furto_ts) +
autolayer(furto_forecast_arima_2, series="Arima", PI=FALSE) +
xlab("Ano") +
ylab("Total de veiculos") +
ggtitle("Veiculos Furtados em Minas Gerais") +
guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_arima_2)
accuracy(furto_forecast_arima_2, x=furto_test_ts)

# Redes Neurais Artificiais - ANN - Single Layer #

furto_forecast_nnetar <- nnetar(furto_train_ts)

```

```

furto_forecast_nnetar_2 <- forecast(furto_forecast_nnetar, h=12)

autoplot(furto_ts) +
  autolayer(furto_forecast_nnetar_2, series="Redes Neurais Artificiais - Uma camada", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Furtados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_nnetar_2)
accuracy(furto_forecast_nnetar_2, x=furto_test_ts)

# Redes Neurais Artificiais - ANN - Multi Layer #

furto_forecast_ann <- mlp(furto_train_ts)
furto_forecast_ann_2 <- forecast(furto_forecast_ann, h=12)

autoplot(furto_ts) +
  autolayer(furto_forecast_ann_2, series="Redes Neurais Artificiais", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Furtados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_ann_2)
accuracy(furto_forecast_ann_2, x=furto_test_ts)

# Prophet #

furto_forecast_prophet <- prophet(furto_train_df)
Future1_furto <- make_future_dataframe(furto_forecast_prophet, periods = 12, freq='month')
Forecast1_furto <- predict(furto_forecast_prophet, Future1_furto)
tail(Forecast1_furto[c('ds', 'yhat')], n=12)
summary(Forecast1_furto)

furto_forecast_prophet_2 <- ts(tail(Forecast1_furto$yhat, 12), start = c(2021,1), frequency=12)

autoplot(furto_ts) +
  autolayer(furto_forecast_prophet_2, series="Prophet") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

summary(furto_forecast_prophet_2)
accuracy(furto_forecast_prophet_2, x=furto_test_ts)

dyplot.prophet(furto_forecast_prophet, Forecast1_furto)
summary(furto_forecast_prophet)

# Modelo Hibrido: ETS para linear, ANN Sigle Layer para nao-linear #

furto_forecast_hib_linear <- ets(na.omit(furto_linear_train))

```

```
furto_forecast_hib_linear_2 <- forecast(furto_forecast_hib_linear, h=12)
accuracy(furto_forecast_hib_linear_2, x=furto_linear_test)
```

```
autoplot(na.omit(furto_linear)) +
  autolayer(furto_forecast_hib_linear_2, series="a") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
furto_forecast_hib_nlinear <- nnetar(na.omit(furto_naolinear_train))
furto_forecast_hib_nlinear_2 <- forecast(furto_forecast_hib_nlinear, h=12)
accuracy(furto_forecast_hib_nlinear_2, x=furto_naolinear_test)
```

```
autoplot(na.omit(furto_naolinear)) +
  autolayer(furto_forecast_hib_nlinear_2, series="a") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
furto_forecast_hibrido <- furto_forecast_hib_linear_2$mean + furto_forecast_hib_nlinear_2$mean
accuracy(furto_forecast_hibrido, x=furto_test_ts)
```

```
# Comparativo: Snaive #
```

```
furto_forecast_snaive <- snaive(furto_train_ts)
furto_forecast_snaive_2 <- forecast(furto_forecast_snaive, h=12)
accuracy(furto_forecast_snaive_2, x=furto_test_ts)
```

```
autoplot(furto_ts) +
  autolayer(furto_forecast_hibrido, series="Modelo Hibrido") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
# Plotando todos os modelos
plot <- window(furto_ts, start=c(2019, 1))
```

```
# Modelos ETS e HW
```

```
autoplot(plot) +
  autolayer(furto_forecast_ets_2, series="ETS", PI=FALSE) +
  autolayer(furto_forecast_hw_add_2, series="HW_Add", PI=FALSE) +
  autolayer(furto_forecast_hw_mult_2, series="HW_Mult", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Furtados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
# Modelos Arima, Linear e Prophet
```



```

autoplot(plot) +
  autolayer(furto_forecast_arima_2, series="Arima", PI=FALSE) +
  autolayer(furto_forecast_tslm_2, series="Regressão Linear", PI=FALSE) +
  autolayer(furto_forecast_prophet_2, series="Prophet", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Furtados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

```

# Modelos ANN e Híbrido

```

autoplot(plot) +
  autolayer(furto_forecast_ann_2, series="Redes Neurais Single Layer") +
  autolayer(furto_forecast_nnetar_2, series="Redes Neurais Multi Layer") +
  autolayer(furto_forecast_hibrido, series="Modelo Híbrido") +

  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Furtados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))

```

## PREVISAO DE 2022 USANDO HW MULT ##

```

furto_forecast_2022 <- forecast(HoltWinters(furto_ts, beta=NULL, gamma=TRUE, seasonal =
"multiplicative"), h=14)

```

```

plot(furto_forecast_2022,
  main = 'Previsão para 2022: Furto de Veículos')
?plot

```

## APÊNDICE C-CÓDIGOS R STUDIO: Série Roubos

```
#### MONOGRAFIA - METODOS E MODELOS DE PREVISAO PARA SERIES TEMPORAIS DE
CRIMINALIDADE EM MINAS GERAIS ####
#### ALUNO: WILLIAN HIDEKI TENGUAN ####
#### RA: 178735 ####
#### ORIENTADOR: PROF. DR. MARCELO JUSTUS DOS SANTOS ####
#### PARTE 3: ROUBO DE VEICULOS ####
#### CRIADO EM: 21/11/2021 ####

# CARREGA OS PACOTES NECESSARIOS AOS MODELOS #

library(RODBC)
library(readxl)
library(forecast)
library(TSstudio)
library(tidyverse)
library(nnfor)
library(prophet)
library(lubridate)
library(zoo)

# DEFINE O BANCO DE DADOS JA NORMALIZADO #

con <- odbcConnect("ToyDB")
setwd("D:/UNICAMP/MONOGRAFIA/TECNICAS/DATABASES")

# CARREGA OS DADOS RELATIVOS AS VITIMAS DE HOMICIDIOS CONSUMADOS #

qry_roubo <- "SELECT ANO, MES, SUM(Registros) AS QtdRegistros
              FROM MG_VEICULOSROUBADOS
              WHERE ANO >= 2015
              GROUP BY ANO, MES
              ORDER BY ANO, MES"
roubo_tbl <- sqlQuery(con, qry_roubo)
str(roubo_tbl)
list(roubo_tbl)

roubo_ts <- ts(roubo_tbl$QtdRegistros, start = c(2015,1), frequency = 12)
roubo_ts_diff <- diff(roubo_ts, 1)

plot(roubo_ts)

# CRIA AS TIMESERIES DE TREINO E TESTE #
roubo_train_ts <- window(roubo_ts, end = c(2020, 12))
roubo_test_ts <- window(roubo_ts, start = c(2021,1))

# CRIA AS TS DE TREINO E TESTE DIFERENCIADAS #
roubo_train_diff <- diff(roubo_train_ts, 1)
roubo_test_diff <- diff(roubo_test_ts, 1)

# PLOTA A SERIE PARA ANALISE PREVIA #

ts.plot(roubo_ts, type = "o", lwd="2", col="2")
lines(lowess(roubo_ts, f=1/12), col=4, lwd="2", type="l")
plot(decompose(roubo_ts))
```

```
plot(roubo_ts_diff)
```

```
## Tendencia de queda pronunciada a partir de 2018.
## Aparenta haver forte sazonalidade, principalmente nos comecos e fins de anos.
## A tendencia de queda se mantem, visualmente, ao mesmo nivel antes e depois da pandemia de COVID-19.
```

```
# GERA DATAFRAME PARA O MODELO PROPHET #
ds <- as.Date(paste(roubo_tbl$ANO, month=roubo_tbl$MES, 1, sep="-"), "%Y-%m-%d")
y <- roubo_tbl$QtdRegistros
roubo_train_df <- subset(data.frame(ds, y), ds <= '2020-12-01')
roubo_test_df <- subset(data.frame(ds, y), ds > '2020-12-01')
```

```
# GERA SERIES PARA O MODELO HIBRIDO #
```

```
roubo_linear <- decompose(roubo_ts)$trend
roubo_linear_train <- window(roubo_linear, end=c(2020, 12))
roubo_linear_test <- window(roubo_linear, start=c(2021, 1))
plot(roubo_linear)
roubo_naolinear <- decompose(roubo_ts)$seasonal+decompose(roubo_ts)$random
roubo_naolinear_train <- window(roubo_naolinear, end=c(2020,12))
roubo_naolinear_test <- window(roubo_naolinear, start=c(2021,1))
plot(roubo_naolinear)
```

```
## MODELOS ###
```

```
# MODELO LINEAR APLICADO A SERIES TEMPORAIS #
```

```
roubo_forecast_tslm <- tslm(roubo_train_ts ~ trend + season, lambda="auto")
roubo_forecast_tslm_2 <- forecast(roubo_forecast_tslm, h=12)
```

```
roubo_forecast_tslm$coefficients
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_tslm_2, series="Modelo de Regressao Linear", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_tslm_2)
accuracy(roubo_forecast_tslm_2, x = roubo_test_ts)
```

```
# MODELO DE SUAVIZACAO EXPONENCIAL #
```

```
roubo_forecast_ets <- ets(roubo_train_ts, model="ZZZ", opt.crit = "mse")
roubo_forecast_ets_2 <- forecast(roubo_forecast_ets, h=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_ets_2, series="Suavização Exponencial", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_ets_2)
accuracy(roubo_forecast_ets_2, x=roubo_test_ts)
```

```
# HOLT-WINTERS - ADDITIVE #
```

```
roubo_forecast_hw_add <- HoltWinters(roubo_train_ts, beta=NULL, gamma=TRUE, seasonal =
"additive")
roubo_forecast_hw_add_2 <- forecast(roubo_forecast_hw_add, h=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_hw_add_2, series="HW additive forecasts", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_hw_add_2)
accuracy(roubo_forecast_hw_add_2, x=roubo_test_ts)
```

```
# HOLT-WINTERS - MULTIPLICATIVE #
```

```
roubo_forecast_hw_mult <- HoltWinters(roubo_train_ts, beta=NULL, gamma=TRUE, seasonal =
"multiplicative")
roubo_forecast_hw_mult_2 <- forecast(roubo_forecast_hw_mult, h=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_hw_mult_2, series="HW multiplicative forecasts", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_hw_mult_2)
accuracy(roubo_forecast_hw_mult_2, x=roubo_test_ts)
```

```
# ARIMA #
```

```
roubo_forecast_arima <- auto.arima(roubo_train_ts)
roubo_forecast_arima_2 <- forecast(roubo_forecast_arima, h=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_arima_2, series="Arima", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_arima_2)
accuracy(roubo_forecast_arima_2, x=roubo_test_ts)
```

```
# Redes Neurais Artificiais - ANN - Single Layer #
```

```
roubo_forecast_nnetar <- nnetar(roubo_train_ts)
roubo_forecast_nnetar_2 <- forecast(roubo_forecast_nnetar, h=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_nnetar_2, series="Redes Neurais Artificiais - Uma camada", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_nnetar_2)
accuracy(roubo_forecast_nnetar_2, x=roubo_test_ts)
```

```
# Redes Neurais Artificiais - ANN - Multi Layer #
```

```
roubo_forecast_ann <- mlp(roubo_train_ts)
roubo_forecast_ann_2 <- forecast(roubo_forecast_ann, h=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_ann_2, series="Redes Neurais Artificiais", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_ann_2)
accuracy(roubo_forecast_ann_2, x=roubo_test_ts)
```

```
# Prophet #
```

```
roubo_forecast_prophet <- prophet(roubo_train_df)
Future1_roubo <- make_future_dataframe(roubo_forecast_prophet, periods = 12, freq='month')
Forecast1_roubo <- predict(roubo_forecast_prophet, Future1_roubo)
tail(Forecast1_roubo[c('ds', 'yhat')], n=12)
summary(Forecast1_roubo)
```

```
roubo_forecast_prophet_2 <- ts(tail(Forecast1_roubo$yhat, 12), start = c(2021,1), frequency=12)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_prophet_2, series="Prophet") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Homicidios Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
summary(roubo_forecast_prophet_2)
accuracy(roubo_forecast_prophet_2, x=roubo_test_ts)
```

```
dyplot.prophet(roubo_forecast_prophet, Forecast1_roubo)
summary(roubo_forecast_prophet)
```

```
# Modelo Hibrido: ETS para linear, ANN Sigle Layer para nao-linear #
```

```
roubo_forecast_hib_linear <- ets(na.omit(roubo_linear_train))
roubo_forecast_hib_linear_2 <- forecast(roubo_forecast_hib_linear, h=12)
accuracy(roubo_forecast_hib_linear_2, x=roubo_linear_test)
```

```
autoplot(na.omit(roubo_linear)) +
  autolayer(roubo_forecast_hib_linear_2, series="a") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
roubo_forecast_hib_nlinear <- nnetar(na.omit(roubo_naolinear_train))
roubo_forecast_hib_nlinear_2 <- forecast(roubo_forecast_hib_nlinear, h=12)
accuracy(roubo_forecast_hib_nlinear_2, x=roubo_naolinear_test)
```

```
autoplot(na.omit(roubo_naolinear)) +
  autolayer(roubo_forecast_hib_nlinear_2, series="a") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
roubo_forecast_hibrido <- roubo_forecast_hib_linear_2$mean + roubo_forecast_hib_nlinear_2$mean
accuracy(roubo_forecast_hibrido, x=roubo_test_ts)
```

```
# Comparativo: Snaive #
```

```
roubo_forecast_snaive <- snaive(roubo_train_ts)
roubo_forecast_snaive_2 <- forecast(roubo_forecast_snaive, h=12)
accuracy(roubo_forecast_snaive_2, x=roubo_test_ts)
```

```
autoplot(roubo_ts) +
  autolayer(roubo_forecast_hibrido, series="Modelo Hibrido") +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Homicidios consumados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
# Plotando todos os modelos
plot_roubo <- window(roubo_ts, start=c(2019, 1))
```

```
# Modelos ETS e HW
```

```
autoplot(plot_roubo) +
  autolayer(roubo_forecast_ets_2, series="ETS", PI=FALSE) +
  autolayer(roubo_forecast_hw_add_2, series="HW_Add", PI=FALSE) +
  autolayer(roubo_forecast_hw_mult_2, series="HW_Mult", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
# Modelos Arima, tSLM e Prophet
```

```
autoplot(plot_roubo) +
  autolayer(roubo_forecast_arima_2, series="Arima", PI=FALSE) +
  autolayer(roubo_forecast_tslm_2, series="Regressão Linear", PI=FALSE) +
  autolayer(roubo_forecast_prophet_2, series="Prophet", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
# Modelos ANN e hibrido
```

```
autoplot(plot_roubo) +
  autolayer(roubo_forecast_ann_2, series="Redes Neurais Single Layer", PI=FALSE) +
  autolayer(roubo_forecast_nnetar_2, series="Redes Neurais Multi Layer", PI=FALSE) +
  autolayer(roubo_forecast_hibrido, series="Híbrido", PI=FALSE) +
  xlab("Ano") +
  ylab("Total de veiculos") +
  ggtitle("Veiculos Roubados em Minas Gerais") +
  guides(colour=guide_legend(title="Forecast"))
```

```
## PREVISAO DE 2022 USANDO HW MULT ##
```

```
roubo_forecast_2022 <- forecast(HoltWinters(roubo_ts, beta=NULL, gamma=TRUE, seasonal =
"multiplicative"), h=14)
```

```
plot(roubo_forecast_2022,
     main = 'Previsão para 2022: Roubo de Veículos')
```