

Análise de séries temporais

Disciplina: Métodos Estatísticos de Previsão

Rafael Sebastião Arocho e Márcio Antonio Vieira

21/11/2025

1 Visão Geral dos Dados

Para esta análise, utilizamos o dataset *House Property Sales Time Series*.

- **Fonte dos dados:** [Kaggle](#).
- **Período:** 2007 a 2019
- **Frequência:** Trimestral ($n = 347$ observações)
- **Variável Resposta (Y_t):**
 - *MA*: Preço (\$) mediano de casas e unidades habitacionais

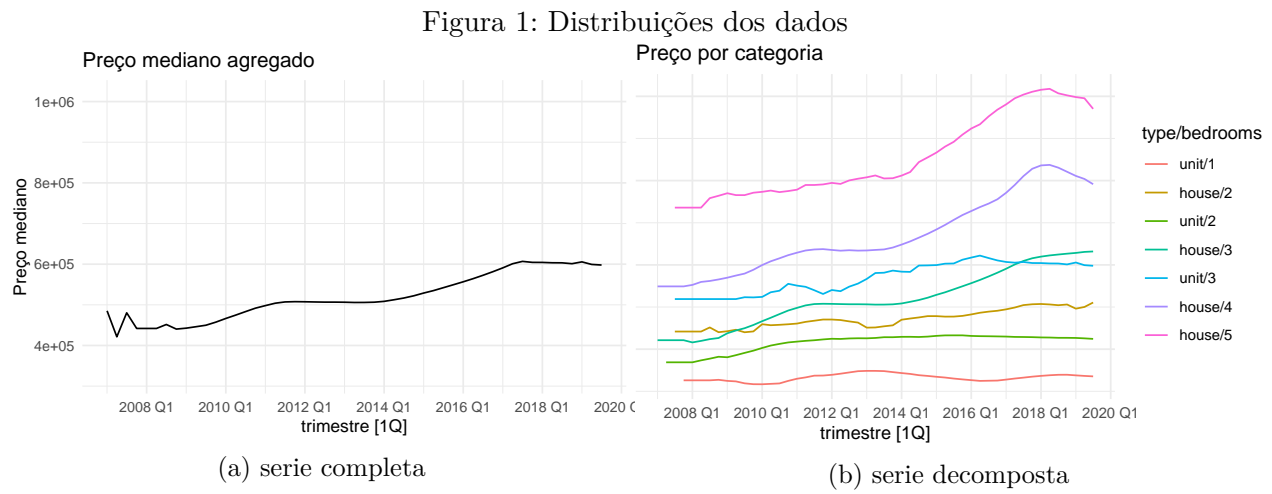
Tabela 1

(a) sumário dos dados

Characteristic	N	Overall N = 347	house N = 200	unit N = 147
MA	347	507,744 (427,623, 628,423)	584,932 (484,806, 771,248)	425,922 (339,125, 535,063)
bedrooms	347			
1		48 (14%)	0 (0%)	48 (33%)
2		99 (29%)	49 (25%)	50 (34%)
3		100 (29%)	51 (26%)	49 (33%)
4		51 (15%)	51 (26%)	0 (0%)
5		49 (14%)	49 (25%)	0 (0%)

¹ Median (Q1, Q3); n (%)

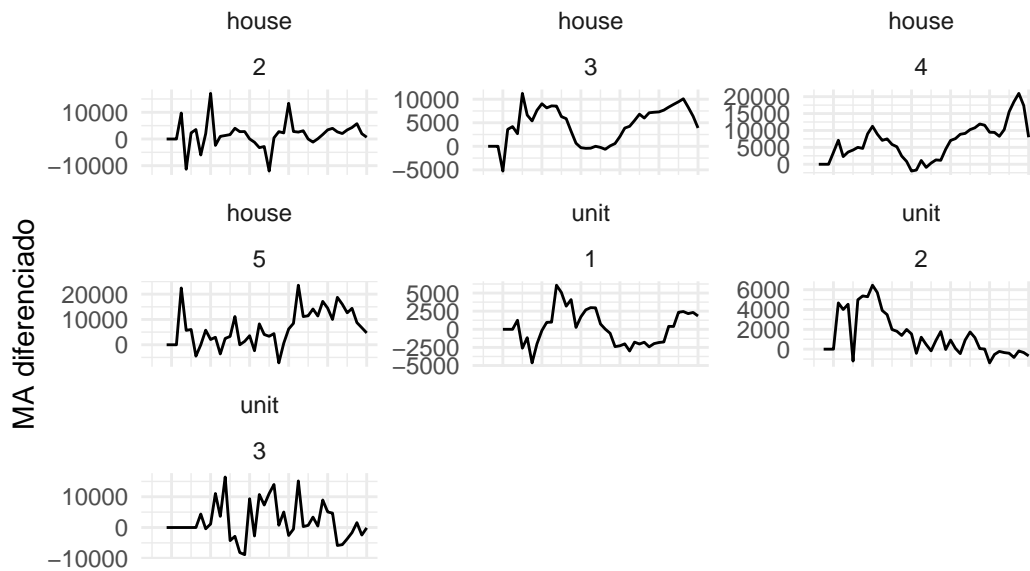
2 Visualização da Série



Observamos que será necessário separar as series em tipo de moradia e número de quartos para simplificar o ajuste.

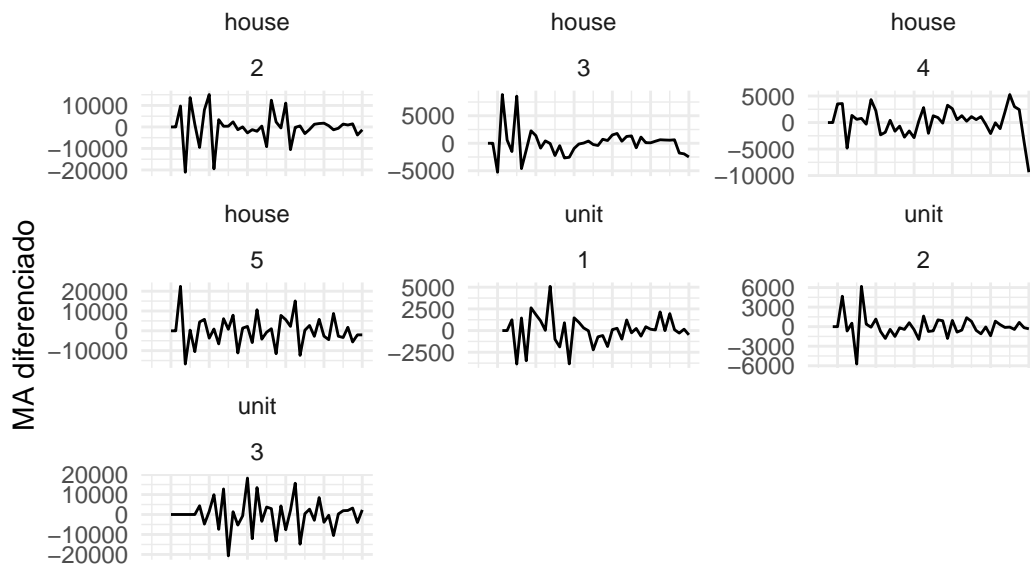
Parece haver tendência em praticamente todas as séries decompostas. Portanto, vamos utilizar diferenciação de nível um e dois para ter outro olhar sobre as séries.

Figura 2: comportamentos diferenciados
Diferenciação com $d = 1$



(a) diferenciação de primeiro nível

Diferenciação com $d = 2$



(b) diferenciação de segundo nível

Observando as séries com as diferenciações aplicadas vamos seguir a análise com os preços das **casas de 5 quartos**.

3 Ajuste de modelos

3.1 Modelos ARIMA

Vamos observar o ACF e o PACF para os dados filtrados com o critério anterior (casas de 5 quartos). Verificaremos tanto para a série não diferenciada quanto diferenciada para um e dois retardos.

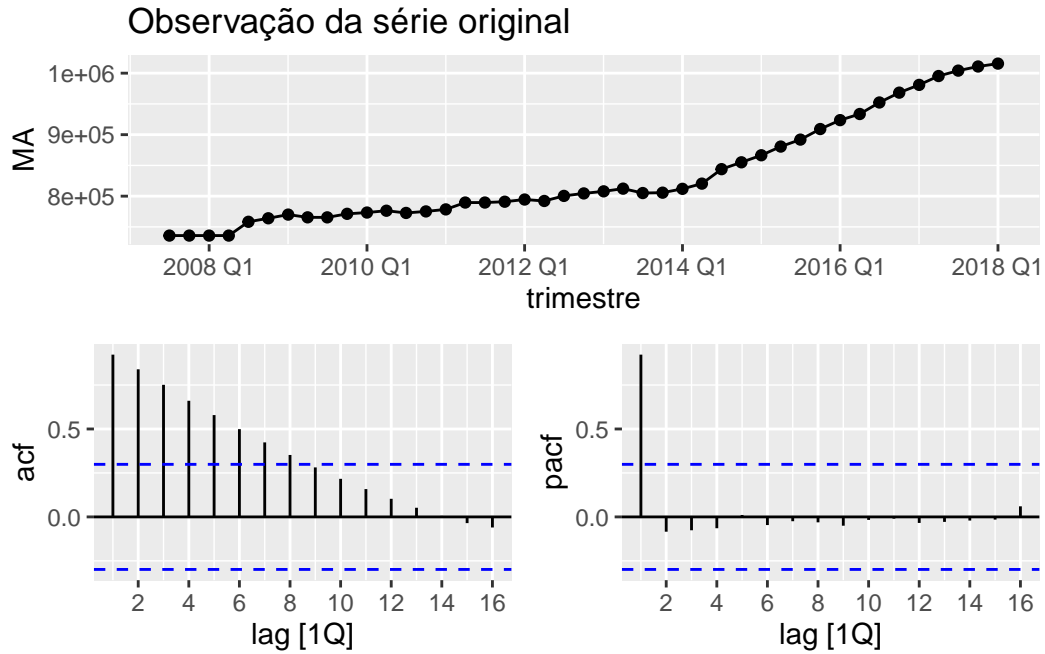


Figura 3: Serie original

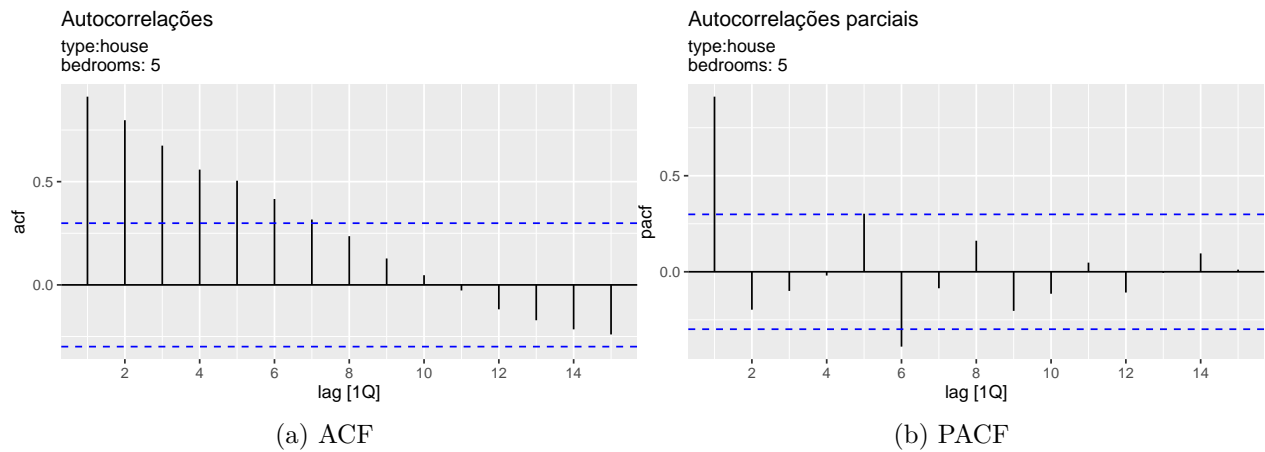


Figura 4: primeira diferenciação

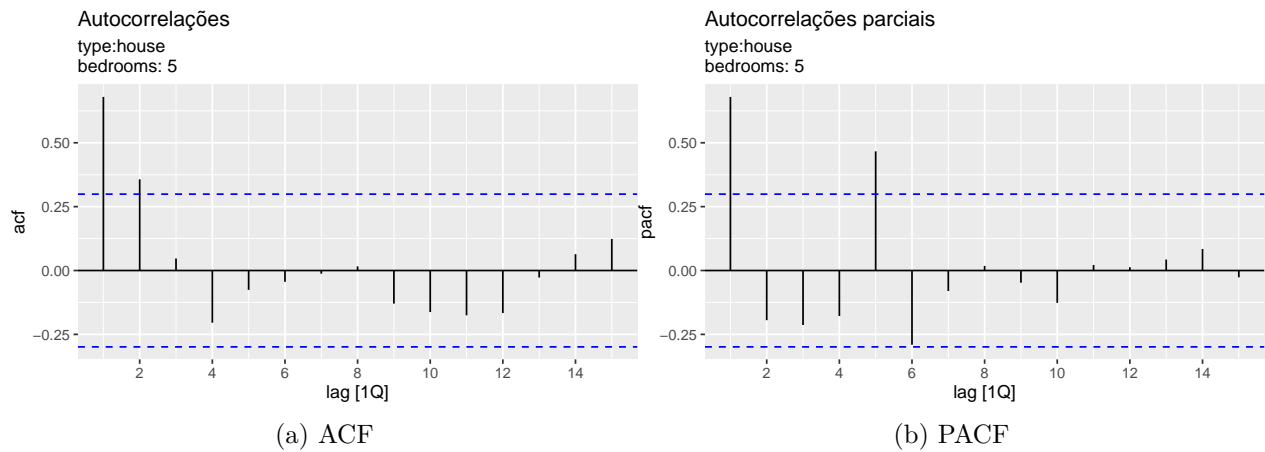


Figura 5: segunda diferenciação

Parece haver um comportamento de ondas senóides no gráfico de ACF e um comportamento irregular no PACF que pode indicar sazonalidade. Vamos ajustar de início um modelo $ARIMA(2, 1, 0)$. Após o ajuste inicial faremos a sobrefixação dos modelos para identificar o melhor modelo através da avaliação das métricas de AIC

3.1.1 Ajustes sequenciais

Tabela 2: Estimativas dos modelos não sazonais (Parte 1)

(a) ARIMA(2,1,0)						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ar1	0.36	0.15	2.40	0.02
house	5	ar2	0.22	0.15	1.45	0.15
house	5	constant	2,706.81	907.04	2.98	0.00
(b) ARIMA(1,1,0)						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ar1	0.45	0.14	3.34	0.00
house	5	constant	3,544.18	948.83	3.74	0.00
(c) ARIMA(1,1,1)						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ar1	0.81	0.19	4.30	0.00
house	5	ma1	-0.49	0.31	-1.56	0.13
house	5	constant	1,191.68	452.11	2.64	0.01
(d) ARIMA(1,2,0)						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ar1	-0.42	0.14	-3.07	0.00

Tabela 3: Estimativas dos modelos complexos e sazonais (Parte 2)

(a) ARIMA(2,2,0)						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ar1	-0.49	0.15	-3.17	0.00
house	5	ar2	-0.14	0.15	-0.90	0.37
(b) Sazonal ARIMA(0,2,1)(0,0,1)[4]						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ma1	-0.45	0.15	-2.91	0.01
house	5	sma1	-0.76	0.22	-3.49	0.00
(c) Sazonal ARIMA(1,2,1)(0,0,1)[4]						
type	bedrooms	term	estimate	std.error	statistic	p.value
house	5	ar1	0.00	0.30	0.00	1.00
house	5	ma1	-0.45	0.28	-1.62	0.11
house	5	sma1	-0.76	0.22	-3.41	0.00

Podemos, também, analisar os modelos com base no AIC:

Tabela 4: Comparação de valores de AIC

.model	AIC
ARIMA(0,2,1)(0,0,1)[4]	832.9159
ARIMA(1,2,1)(0,0,1)[4]	834.9159
ARIMA(1,2,0)	842.2771
ARIMA(2,2,0)	843.4800
ARIMA(2,1,0)	859.7227
ARIMA(1,1,1)	859.7261
ARIMA(1,1,0)	859.7691

3.1.2 Análise de resíduos

Ajustados os modelos daremos sequência na análise com o modelo ARIMA(0,2,1)(0,0,1)[4] por ter apresentado significância nos parâmetros e o menor dos valores de AIC.

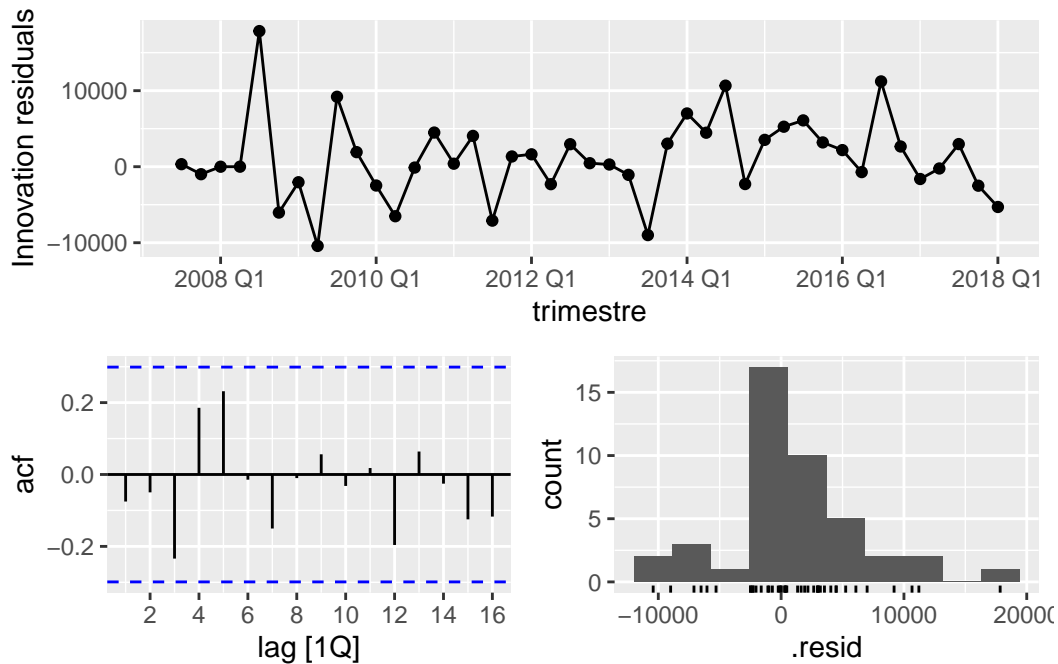


Figura 6: Análise de resíduos do modelo ARIMA

Tabela 5: Testes de residuos

(a) Teste de Shapiro Wilk para normalidade dos residuos

statistic	p.value	method
0.9636118	0.1877009	Shapiro-Wilk normality test

(b) Teste de Box-Pierce para normalidade dos residuos

statistic	p.value	parameter	method
9.326706	0.6748011	12	Box-Pierce test

Os testes nos mostram que os residuos apresentam normalidade e o teste de box-pierce não rejeita a hipótese de independencia, o que é um indicador que o ajuste está adequado.

3.1.3 Previsão

Vamos utilizar o modelo escolhido para fazer uma previsão 6 passos a frente

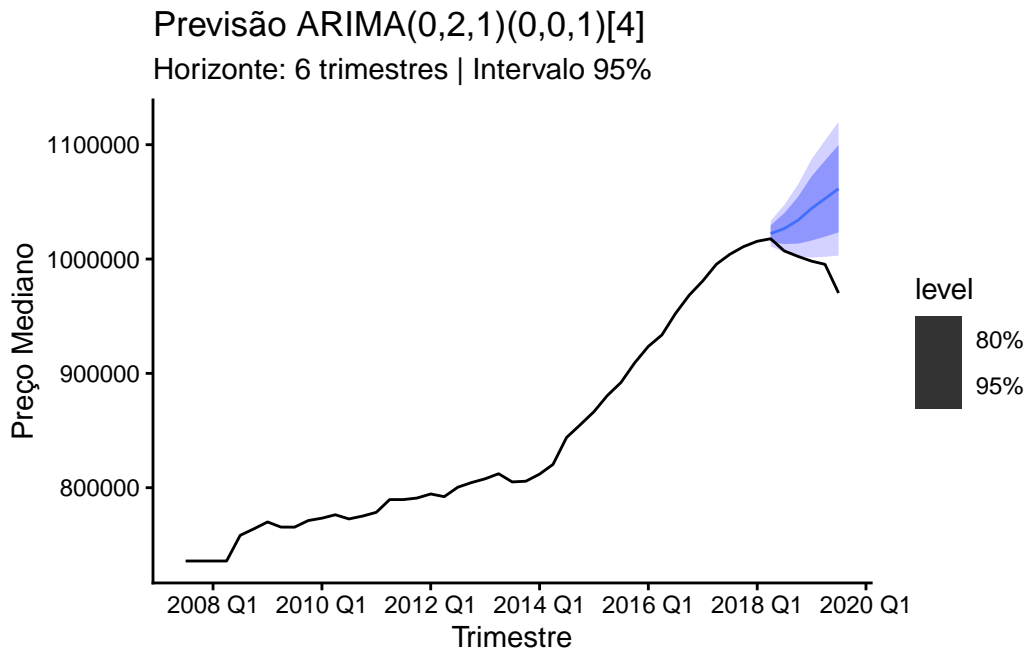


Figura 7: previsão 6 passos a frente

Observamos que a previsão manteve a tendência que é apresentada durante o tempo de treinamento, mas acaba errando os valores reais pois no fim da série existe uma mudança brusca de comportamento que não foi ajustada no modelo.

3.2 Modelo de alisamento exponencial

Seguindo a mesma lógica dos ajustes anteriores faremos o ajuste de modelo de alisamento exponencial. Utilizaremos alisamento exponencial de Holt-Winters aditivo por haver tendência e sazonalidade constante na série.

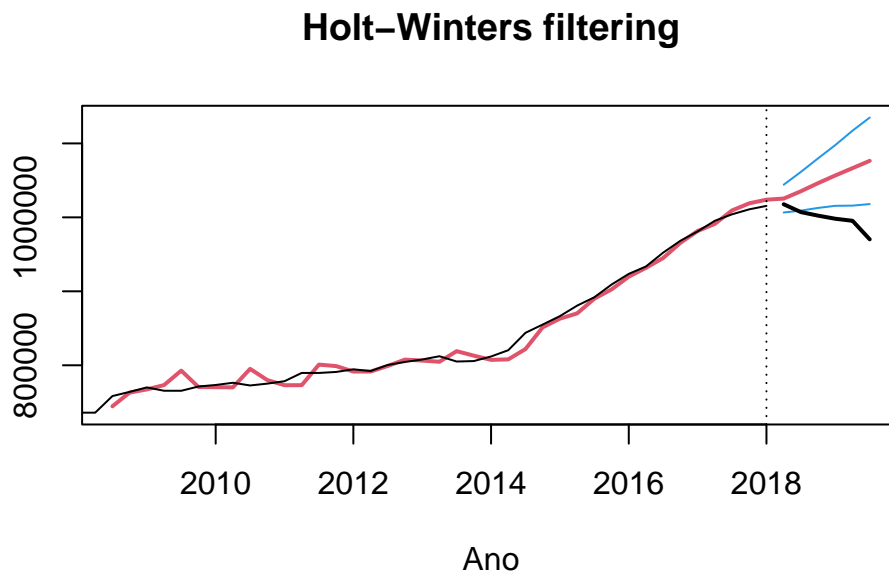


Figura 8: Previsão ajuste de Holt-Winters aditivo

Observamos a mesma dificuldade de acertar a previsão devido a queda abrupta dos preços no fim do período.

4 Comparação dos modelos

Vamos comparar os dois modelos em relação ao erro quadrático médio:

Previsões Pontuais HW:

	Qtr1	Qtr2	Qtr3	Qtr4
2018		1025381	1035293	1046315
2019	1056702	1066553	1076466	

Previsões Pontuais ARIMA:

```
[1] 1022220 1026582 1033842 1044259 1052874 1061490
```

Valores Reais (Base de Teste):

```
[1] 1017752 1007114 1002323 998136 995363 970268
```

Erro Quadrático Médio (MSE) para os 6 passos HW:

```
[1] 3760607450
```

Erro Quadrático Médio (MSE) para os 6 passos ARIMA:

```
[1] 2524795264
```

Observamos um EQM menor no ajuste com o modelo ARIMA, então podemos concluir que se trata do melhor modelo para ajustar a série de preços medianos de casas com 5 quartos.

5 Conclusão

Escolhemos uma das séries decompostas para realizar o processo de ajustes de séries temporais devido à dificuldade de ajuste na série inteira ou com diversas séries diferentes. A sobrefixação dos modelos ARIMA nos permitiu identificar o melhor modelo, um ARIMA com duas diferenciações, uma componente de média móvel simples e outra sazonal ($ARIMA(0,2,1)(0,0,1)[4]$), que não havíamos identificado de início através da análise de ACF e PACF. O modelo se mostrou bem ajustado em relação as suposições de normalidade dos resíduos e independência. Ajustamos, também, um modelo de alisamento exponencial de Holt-Winters, mas ao compara-lo com o modelo ARIMA verificamos um erro quadrático médio maior. As previsões 6 passos a frente se mostraram ruins devido à um fenómeno de decrescimento repentino no fim da série, esse comportamento não incluído no treinamento do modelo faz com que nossa previsão siga a tendência anterior de crescimento e desvie de forma significativa dos resultados reais.