

Today's Agenda

Analyzing Time Series Data

1. Fitting the models
2. Interpreting the models
3. Evaluating the models

Justin Leinaweaiver (Summer 2023)

What is the "best" model of bachelor's degree completion in the Session 1 data?

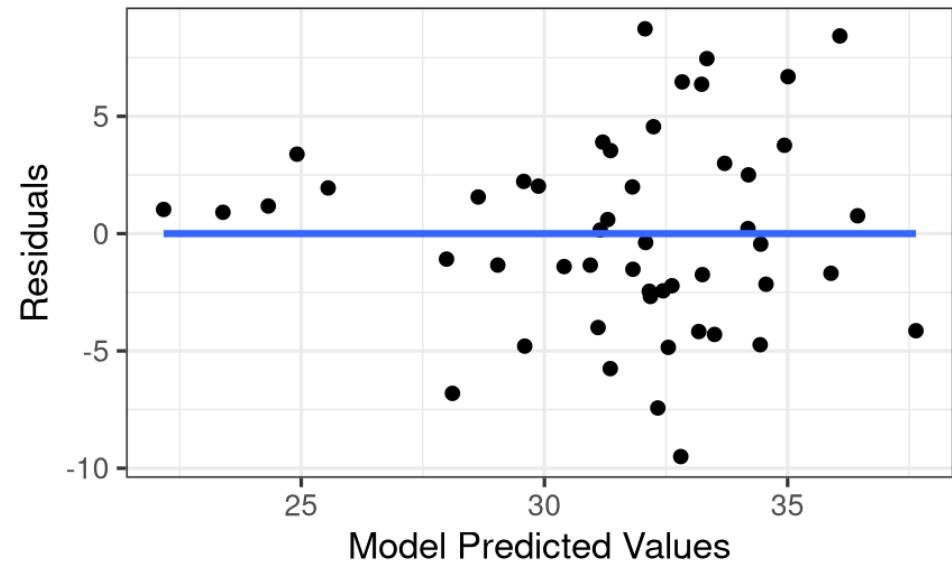
- Outcome:
 - Bachelors' Degrees
- Predictors to consider:
 - GDP (Rate), Homeownership, Minimum wage, State Tax Rate on Wages, Unemployment

		Bachelors' Degrees (%)				
		(1)	(2)	(3)	(4)	(5)
GDP Rate		-79.42				
		(61.13)				
Homeownership			-0.35*			
			(0.15)			
Minimum Wage				0.96*		
				(0.22)		
State Taxes					0.45	
					(0.24)	
Unemployment						-1.83*
						(0.90)
Constant	35.66*	55.23*	24.22*	29.30*	38.53*	
	(3.22)	(10.32)	(1.79)	(1.41)	(3.49)	
Observations	50	50	50	50	50	
Adjusted R ²	0.01	0.08	0.27	0.05	0.06	
<i>Note:</i>		*p<0.05				

	Bachelors' Degrees (%)		
	(1)	(2)	(3)
Homeownership	-0.35*		
	(0.15)		
Minimum Wage		0.96*	
		(0.22)	
Unemployment			-1.83*
			(0.90)
Constant	55.23*	24.22*	38.53*
	(10.32)	(1.79)	(3.49)
Observations	50	50	50
Adjusted R ²	0.08	0.27	0.06
<i>Note:</i>	*p<0.05		

	Bachelors' Degrees (%)			
	(1)	(2)	(3)	(4)
Homeownership	-0.35*			-0.21
	(0.15)			(0.14)
Minimum Wage		0.96*		0.82*
		(0.22)		(0.22)
Unemployment			-1.83*	-1.89*
			(0.90)	(0.76)
Constant	55.23*	24.22*	38.53*	46.23*
	(10.32)	(1.79)	(3.49)	(11.09)
Observations	50	50	50	50
Adjusted R ²	0.08	0.27	0.06	0.34
Residual Std. Error	5.09 (df = 48)	4.53 (df = 48)	5.15 (df = 48)	4.30 (df = 46)
F Statistic	5.27* (df = 1; 48)	19.36* (df = 1; 48)	4.12* (df = 1; 48)	9.58* (df = 3; 46)
Note:	*p<0.05			

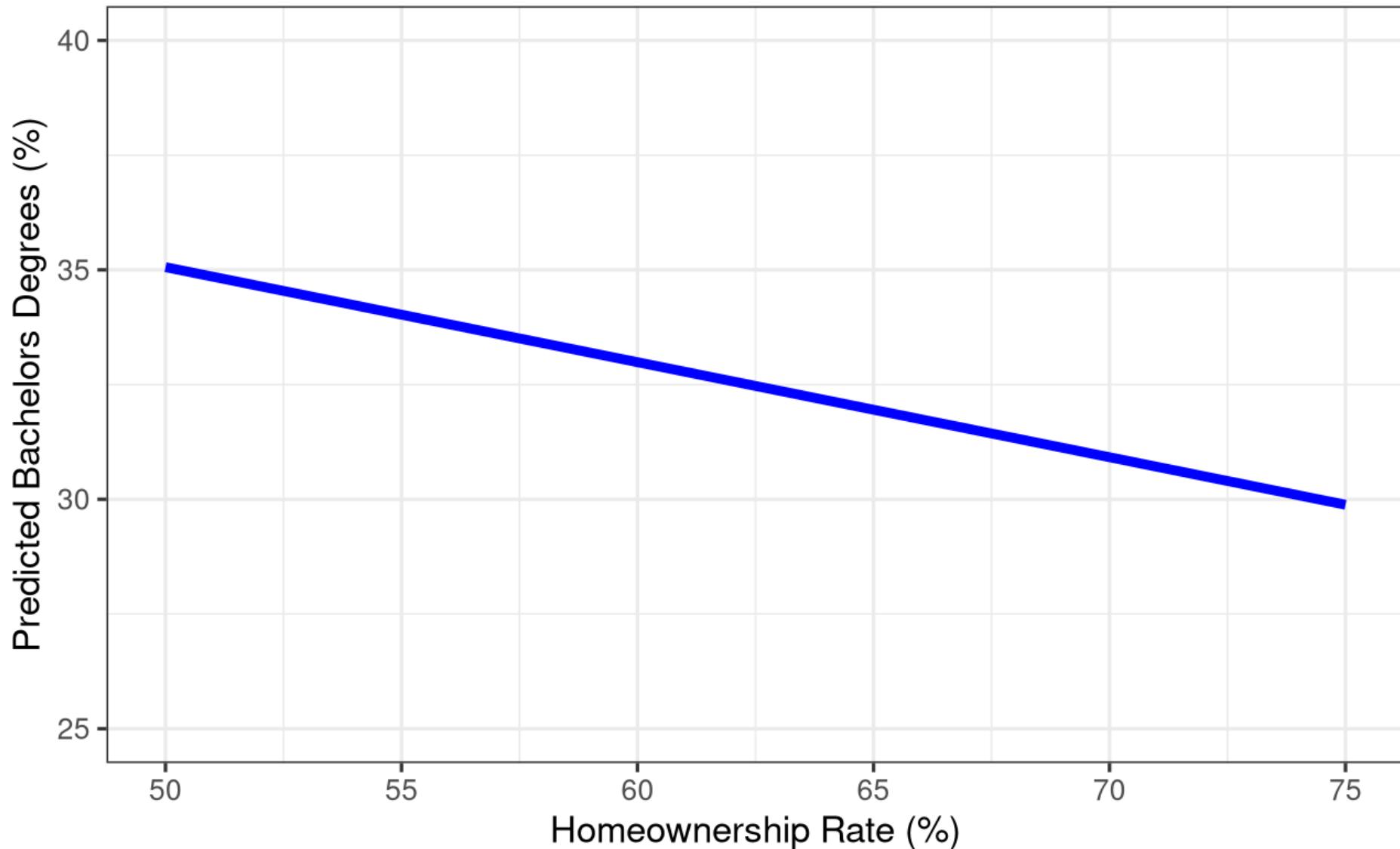
	Bachelors' Degrees (%)
Homeownership	-0.21 (0.14)
Minimum Wage	0.82* (0.22)
Unemployment	-1.89* (0.76)
Constant	46.23* (11.09)
Observations	50
Adjusted R ²	0.34
Residual Std. Error	4.30 (df = 46)
F Statistic	9.58* (df = 3; 46)
Note:	*p<0.05



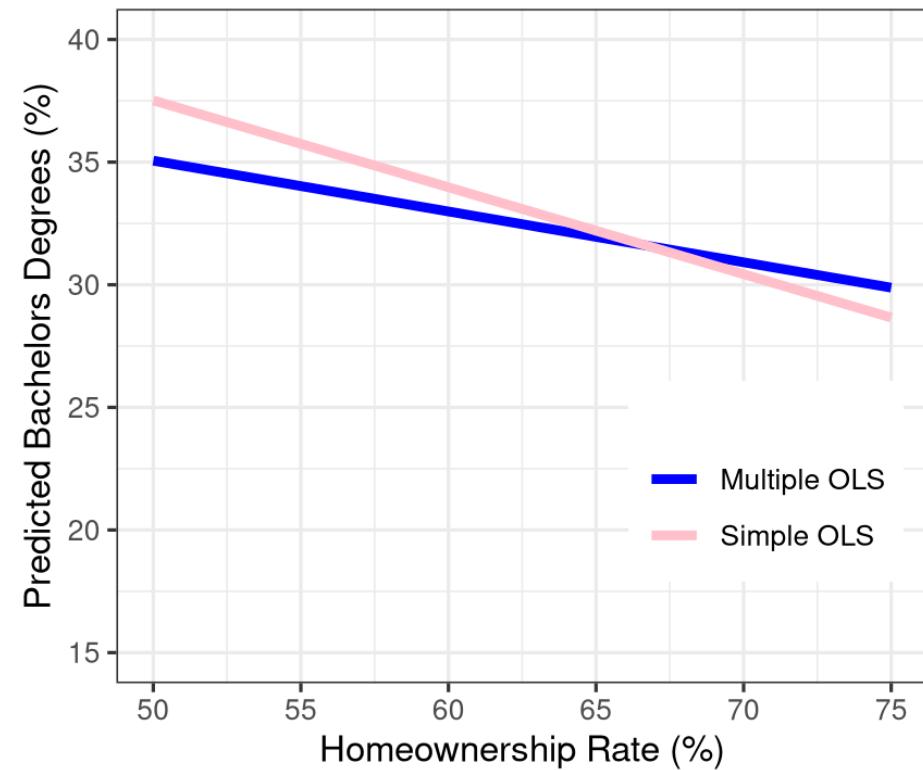
	Home	Wage	Unempl
Home	1.00	-0.37	-0.13
Wage	-0.37	1.00	-0.04
Unempl	-0.13	-0.04	1.00

	Bachelors' Degrees (%)
Homeownership	-0.21 (0.14)
Minimum Wage	0.82* (0.22)
Unemployment	-1.89* (0.76)
Constant	46.23* (11.09)
Observations	50
Adjusted R ²	0.34
F Statistic	9.58* (df = 3; 46)
Note:	*p<0.05

Make a marginal effects plot of homeownership using our multiple regression results



	Bachelors' Degrees (%)	
	(1)	(2)
homeowner_rate	-0.35*	-0.21
	(0.15)	(0.14)
min_wage		0.82*
		(0.22)
unemployment		-1.89*
		(0.76)
Constant	55.23*	46.23*
	(10.32)	(11.09)
Observations	50	50
Adjusted R ²	0.08	0.34
F Statistic	5.27* (df = 1; 48)	9.58* (df = 3; 46)
Note:	*p<0.05	



	Bachelors' Degrees (%)			
	(1)	(2)	(3)	(4)
Homeownership	-0.35*			-0.21
	(0.15)			(0.14)
Minimum Wage		0.96*		0.82*
		(0.22)		(0.22)
Unemployment			-1.83*	-1.89*
			(0.90)	(0.76)
Constant	55.23*	24.22*	38.53*	46.23*
	(10.32)	(1.79)	(3.49)	(11.09)
Observations	50	50	50	50
Adjusted R ²	0.08	0.27	0.06	0.34
F Statistic	5.27* (df = 1; 48)	19.36* (df = 1; 48)	4.12* (df = 1; 48)	9.58* (df = 3; 46)
Note:	*p<0.05			

Reviewing Sessions 1, 2 & 3

1. Descriptive statistics & visualizations
2. Simple OLS Regression
3. Multiple OLS Regression

The Components of Time Series Data

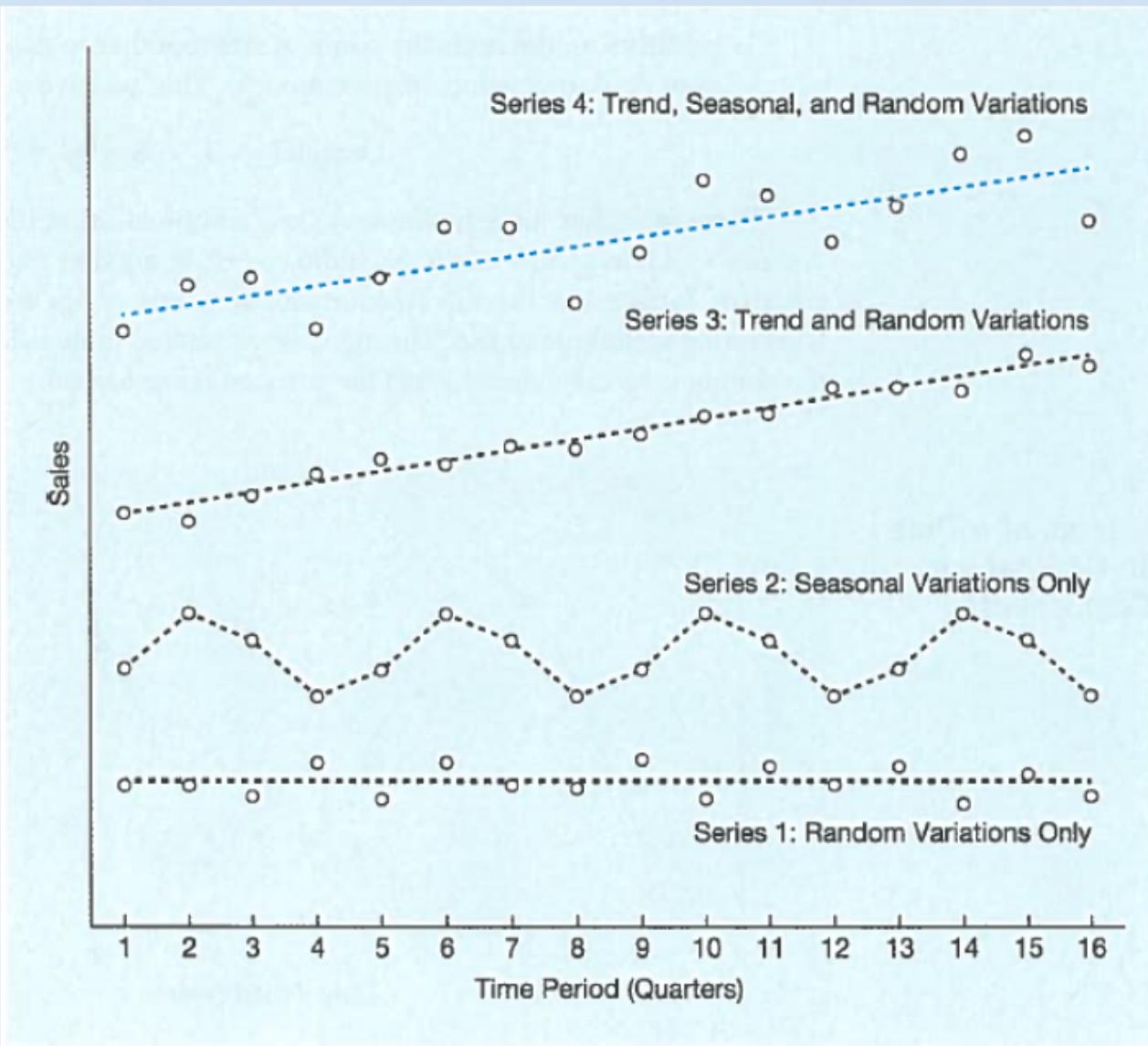


Figure 5.2
(Render, Stair, Jr.,
Hanna and Hale 2018)

Session 4 Data: Missouri Across Time

state	year	gdp_millions	gdp_rate	unemployment	population	bachelors	homeowner	manufacturing
MO	2014	287507.5	2.47	6.2	6059.13	27.5	70.5	256.8
MO	2015	296929.3	3.28	5.1	6075.41	27.8	68.5	261.9
MO	2016	300914.5	1.34	4.6	6091.38	28.5	66.7	264.0
MO	2017	308002.3	2.36	3.7	6111.38	29.1	67.6	266.7
MO	2018	318153.5	3.30	3.2	6125.99	29.5	69.7	272.9
MO	2019	332485.8	4.50	3.2	6140.48	30.2	69.1	277.4
MO	2020	330249.9	-0.67	6.2	6154.00	31.9	71.1	266.6
MO	2021	358572.0	8.58	4.1	6169.82	31.7	72.8	271.4
MO	2022	389931.2	8.75	2.5	6177.96	NA	70.6	282.6

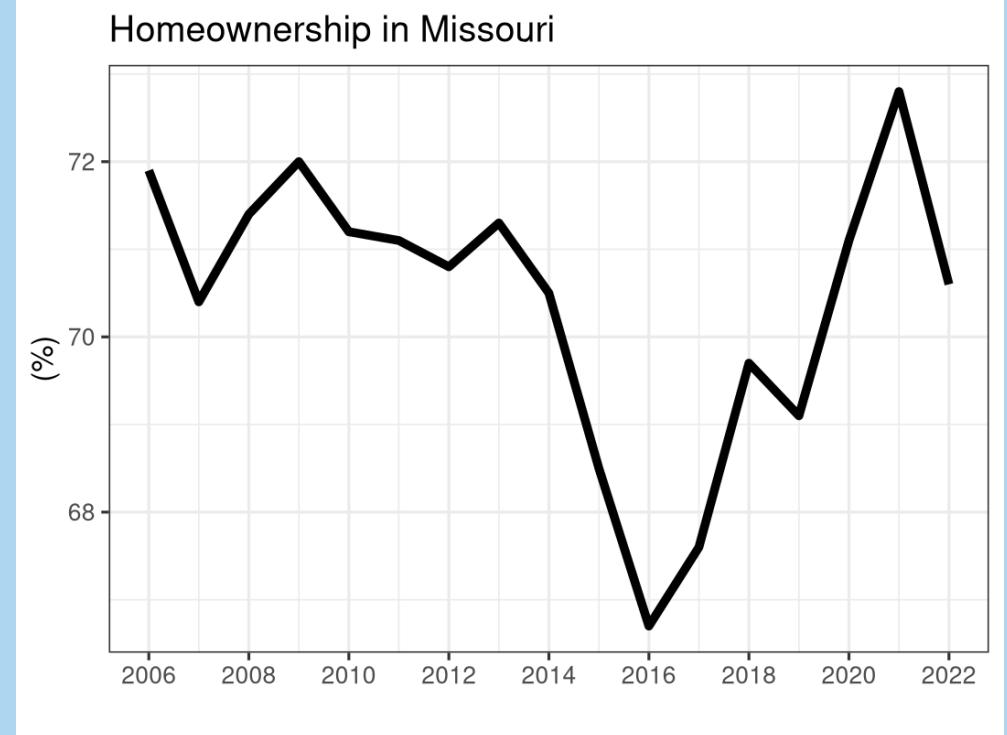
Homeownership in Missouri



Random Variation

Forecasting Tools

1. Naive Forecasts
2. Moving Averages (MA)
3. Weighted Moving Averages (WMA)



Forecast 1: Naïve Forecast

Set forecast to the last observation

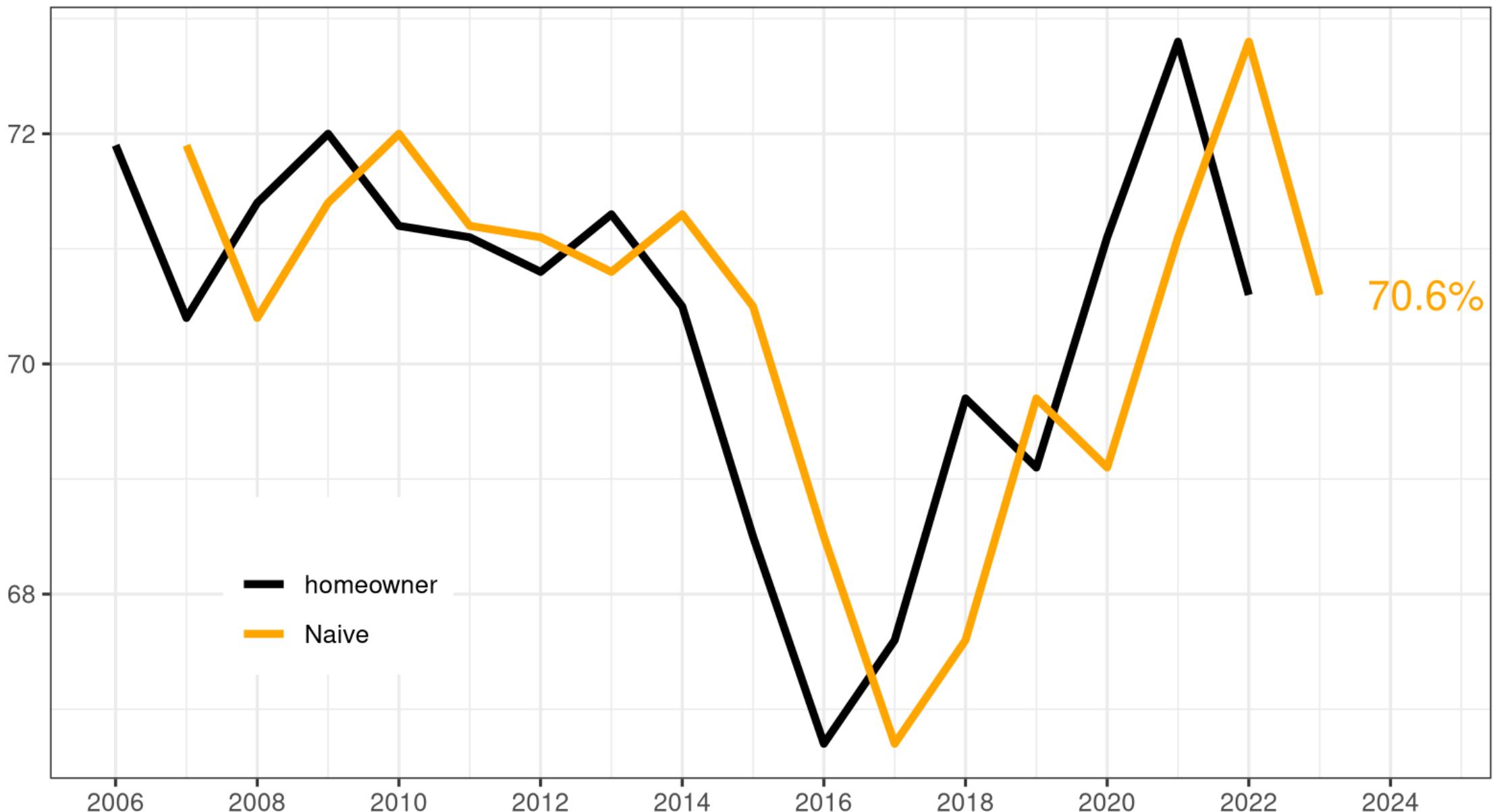
$$\text{Forecast}_t = \text{Actual}_{t-1}$$

Forecast 1: Naïve Forecast

	A	B	C
1	Time	Actual	Forecast
2	1998	X_{1998}	---
3	1999	X_{1999}	$= B2$
4	2000	X_{2000}	$= B3$

Calculate the forecast, extend one year & visualize

Homeownership Rates (Missouri)



Forecast Accuracy: Mean Squared Error (MSE)

1. Calculate the forecast error
 - a. Forecast Error = Actual Value - Forecast Value
2. Square each forecast error
3. Calculate the mean of the squared errors

$$MSE = \frac{\sum(\text{Error})^2}{n}$$

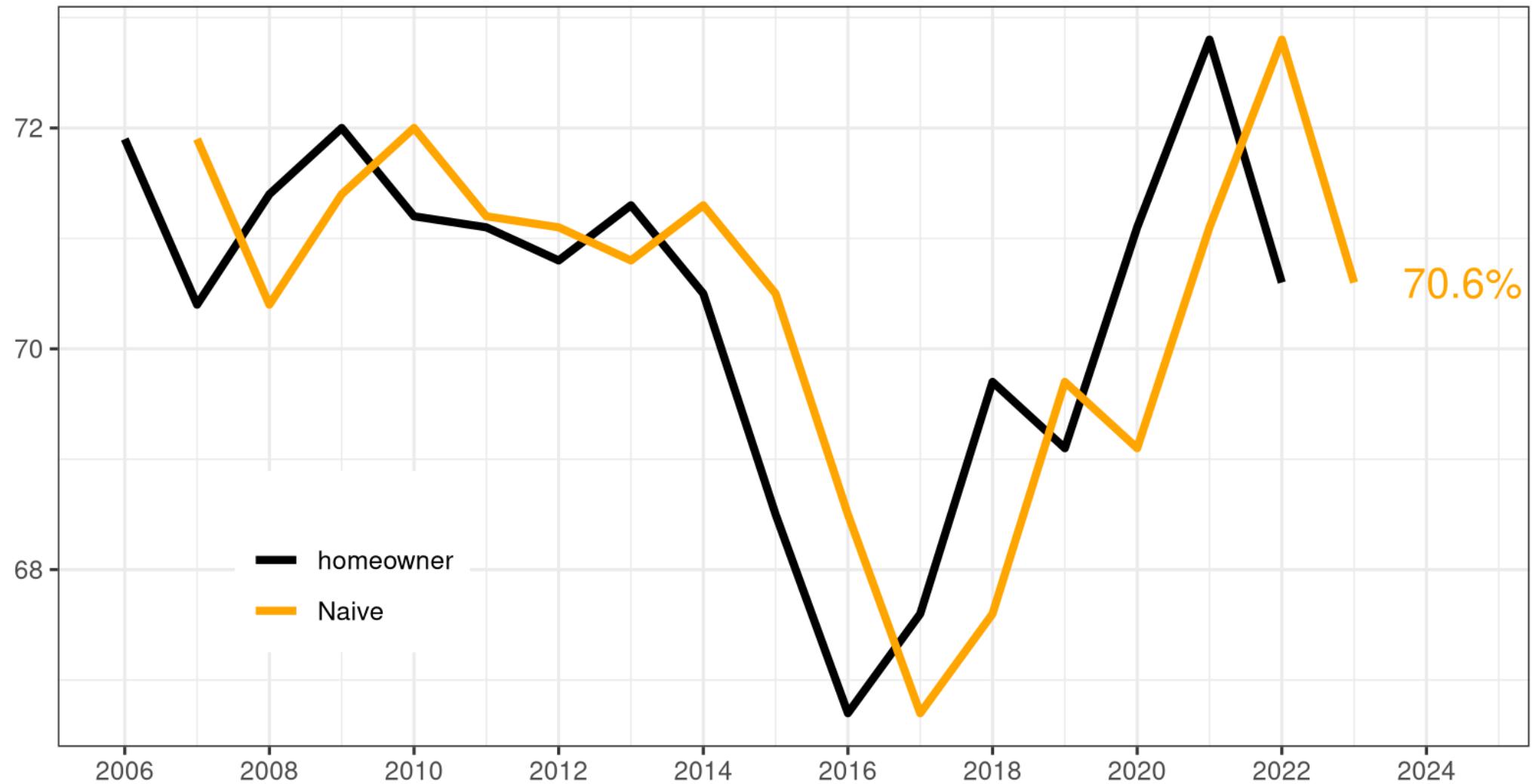
$$\text{MSE} = \frac{\sum(\text{Error})^2}{n}$$

	A	B	C	D
1	Time	Actual	Forecast	Error ²
2	1998	X ₁₉₉₈	---	---
3	1999	X ₁₉₉₉	= B2	= (B3 - C3) ²
4	2000	X ₂₀₀₀	= B3	= (B4 - C4) ²

MSE = AVERAGE(D3:D4)

Homeownership Rates (Missouri)

MSE = 1.8619



Forecast 2: Moving Average Forecast

$$F_{t+1} = \frac{Y_t + Y_{t-1} + \cdots + Y_{t-n+1}}{n}$$

F_{t+1} = forecast for time period $t + 1$

Y_t = actual value in time period t

n = number of periods to average

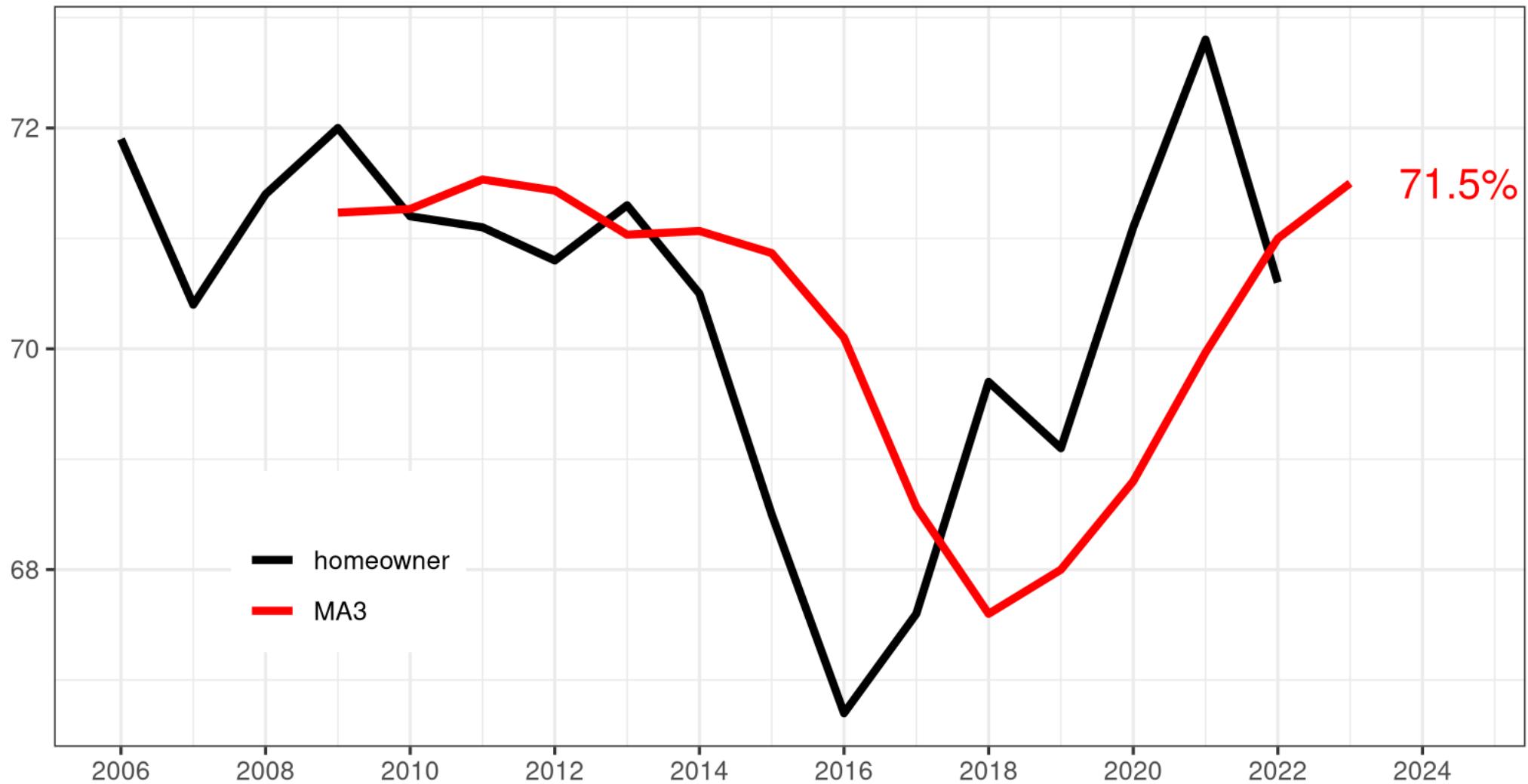
Forecast 2: Moving Average Forecast (3)

	A	B	C
1	Time	Actual	Forecast
2	1998	X_{1998}	---
3	1999	X_{1999}	---
4	2000	X_{2000}	---
5	2001	X_{2001}	$= (B2 + B3 + B4) / 3$
6	2002	X_{2002}	$= (B3 + B4 + B5) / 3$
7	2003	X_{2003}	$= (B4 + B5 + B6) / 3$

To do: Forecast, MSE, extend one year & visualize

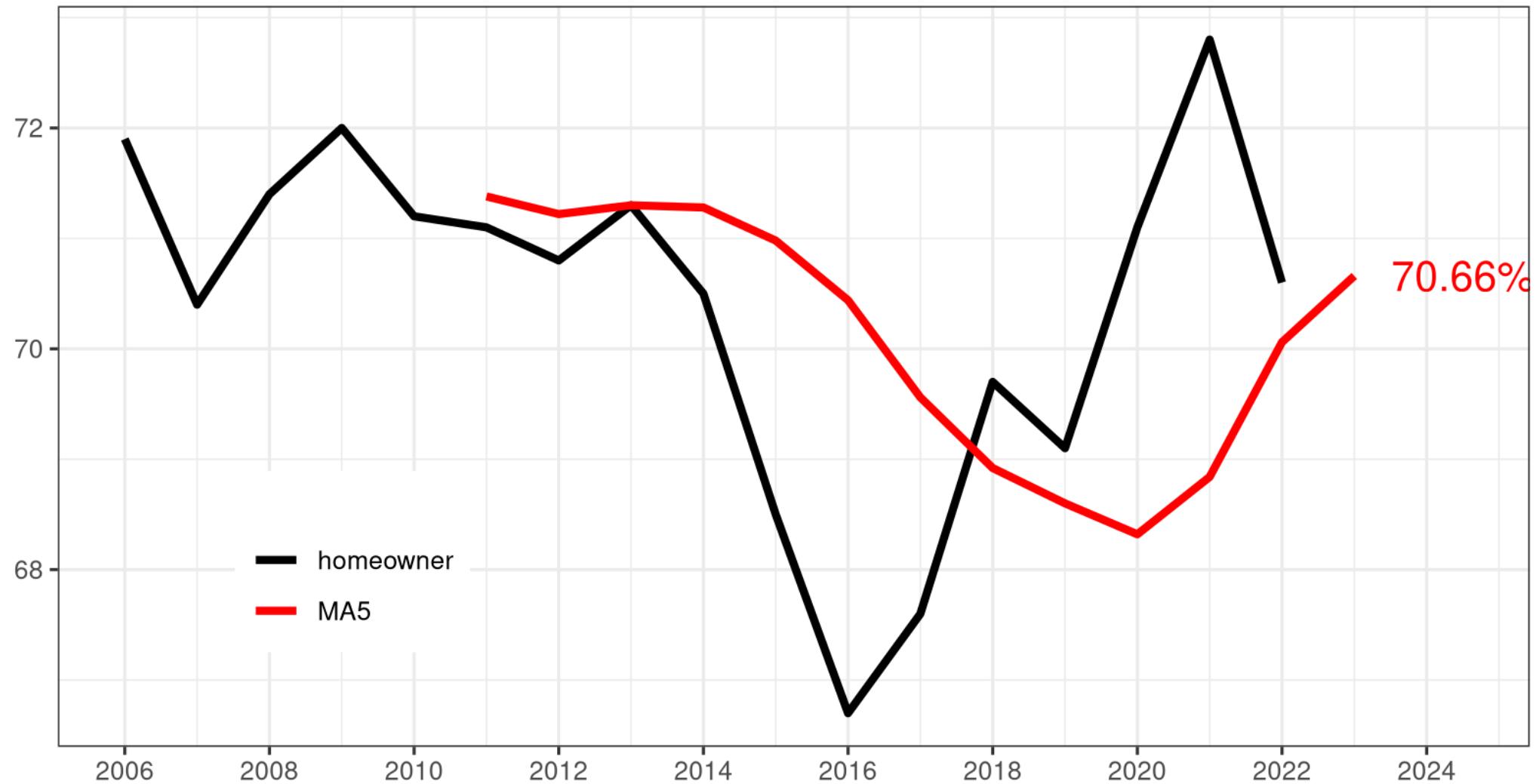
Homeownership Rates (Missouri)

MSE = 2.769



Homeownership Rates (Missouri)

MSE = 4.1169



Forecast 3: Weighted Moving Average Forecast

$$F_{t+1} = \frac{\sum (\text{Weight in period } i)(\text{Actual value in period } i)}{\sum (\text{Weights})}$$

Forecast 3: Weighted MA-3 Forecast

$$F_{t+1} = \frac{\sum (\text{Weight in period } i)(\text{Actual value in period } i)}{\sum (\text{Weights})}$$

$$Forecast_t = \frac{(Actual_{t-1} * 2 + Actual_{t-2} * 2 + Actual_{t-3} * 1)}{5}$$

Forecast 3: Weighted MA-3 Forecast

$$F_{t+1} = \frac{\sum (\text{Weight in period } i)(\text{Actual value in period } i)}{\sum (\text{Weights})}$$

$$\text{Forecast}_t = \frac{(\text{Actual}_{t-1} * 3 + \text{Actual}_{t-2} * 2 + \text{Actual}_{t-3} * 1)}{6}$$

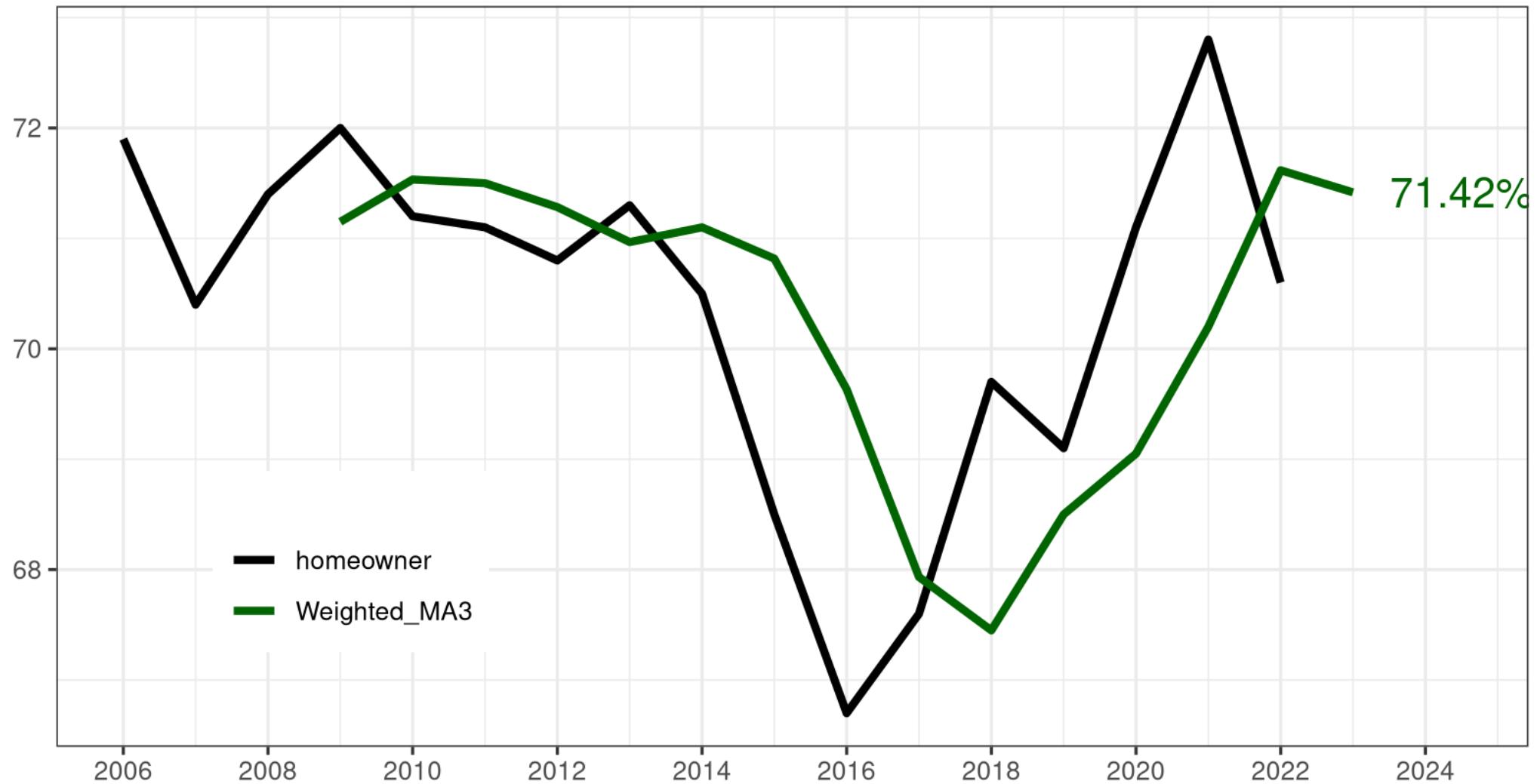
Forecast 3: Weighted MA-3 Forecast

	A	B	C
1	Time	Actual	Forecast
2	1998	X_{1998}	---
3	1999	X_{1999}	---
4	2000	X_{2000}	---
5	2001	X_{2001}	$= (B2*1 + B3*2 + B4*3) / 6$
6	2002	X_{2002}	$= (B3*1 + B4*2 + B5*3) / 6$
7	2003	X_{2003}	$= (B4*1 + B5*2 + B6*3) / 6$

To do: Forecast, MSE, extend one year & visualize

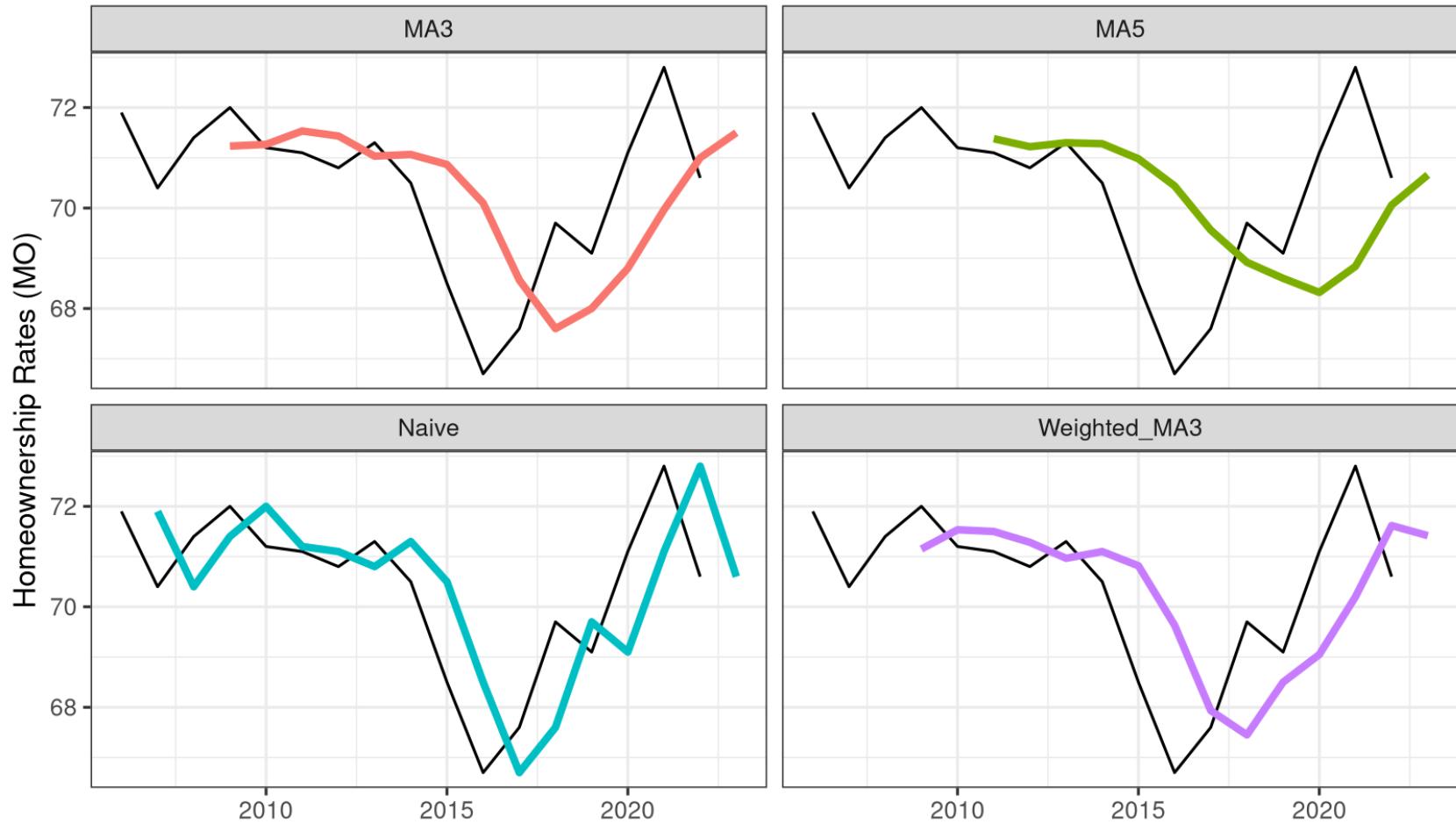
Homeownership Rates (Missouri)

MSE = 2.3714



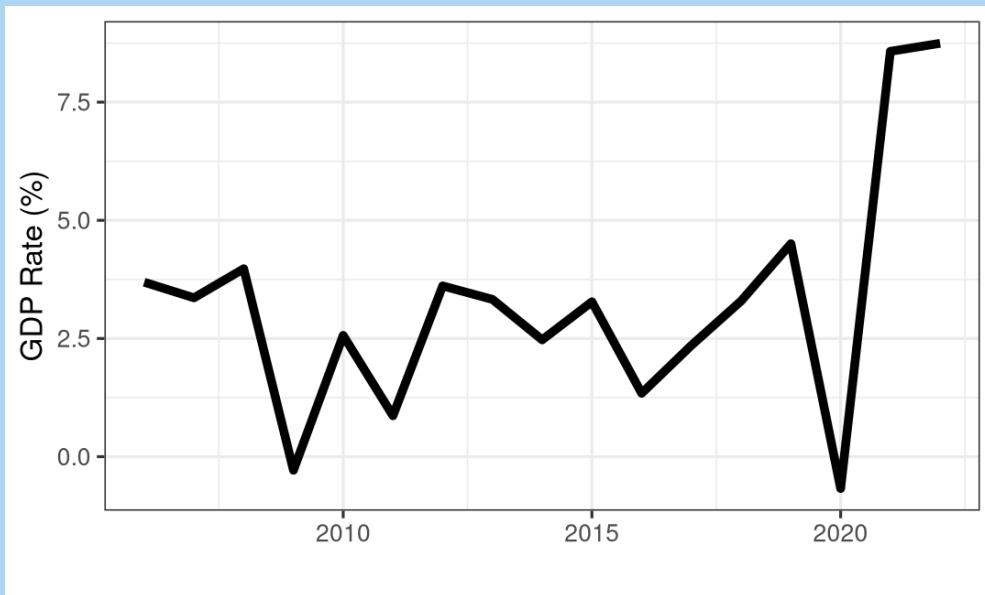
Comparing our Forecasts

Method	MSE	Prediction
Naive	1.88	70.6%
Weighted-MA3	2.85	71.4%
MA-3	3.46	71.5%
MA-5	4.46	70.7%



Method	Naive	Weighted-MA3	MA-3	MA-5
MSE	1.882	2.848	3.456	4.465

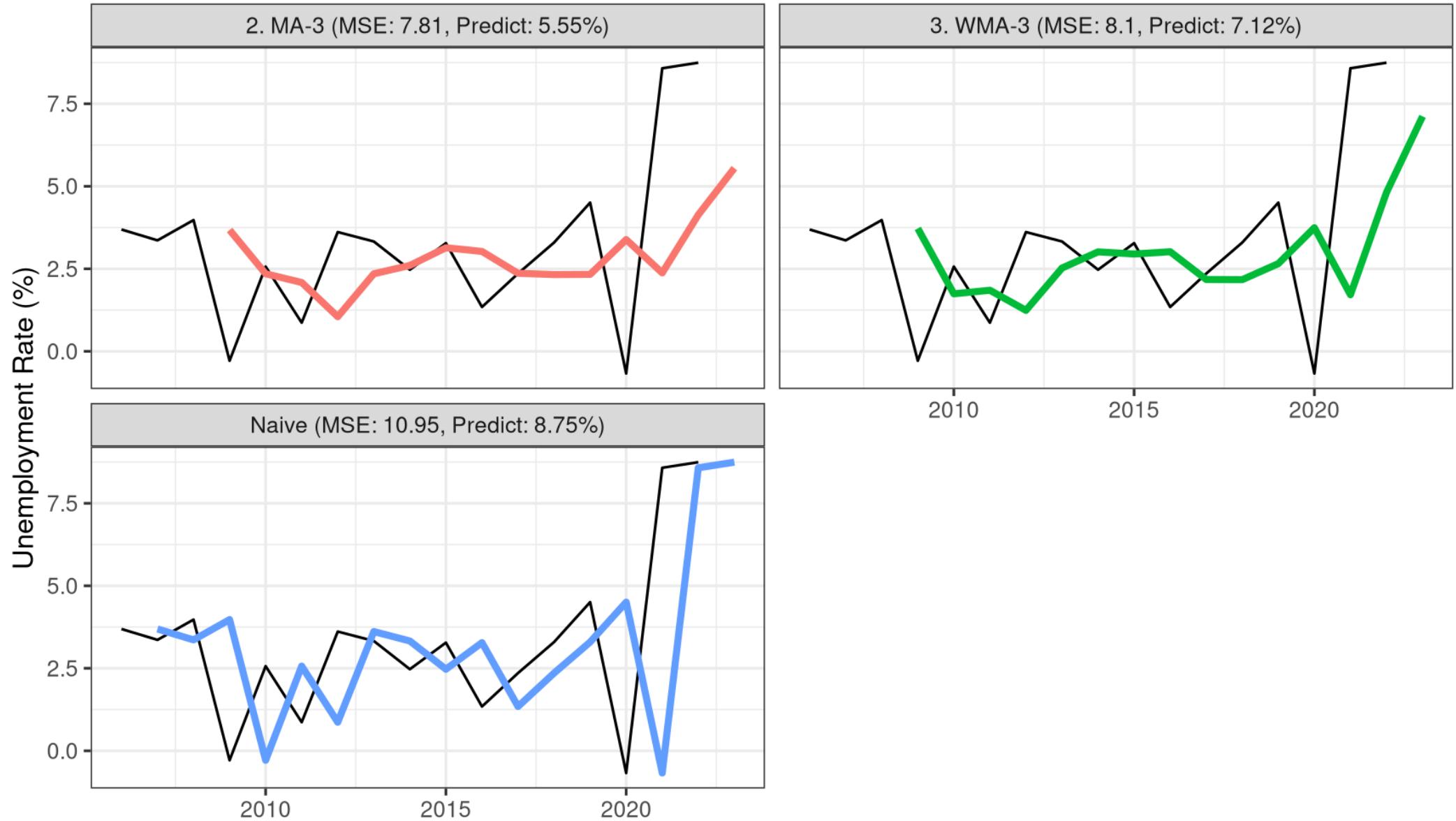
Practice Using the Tools



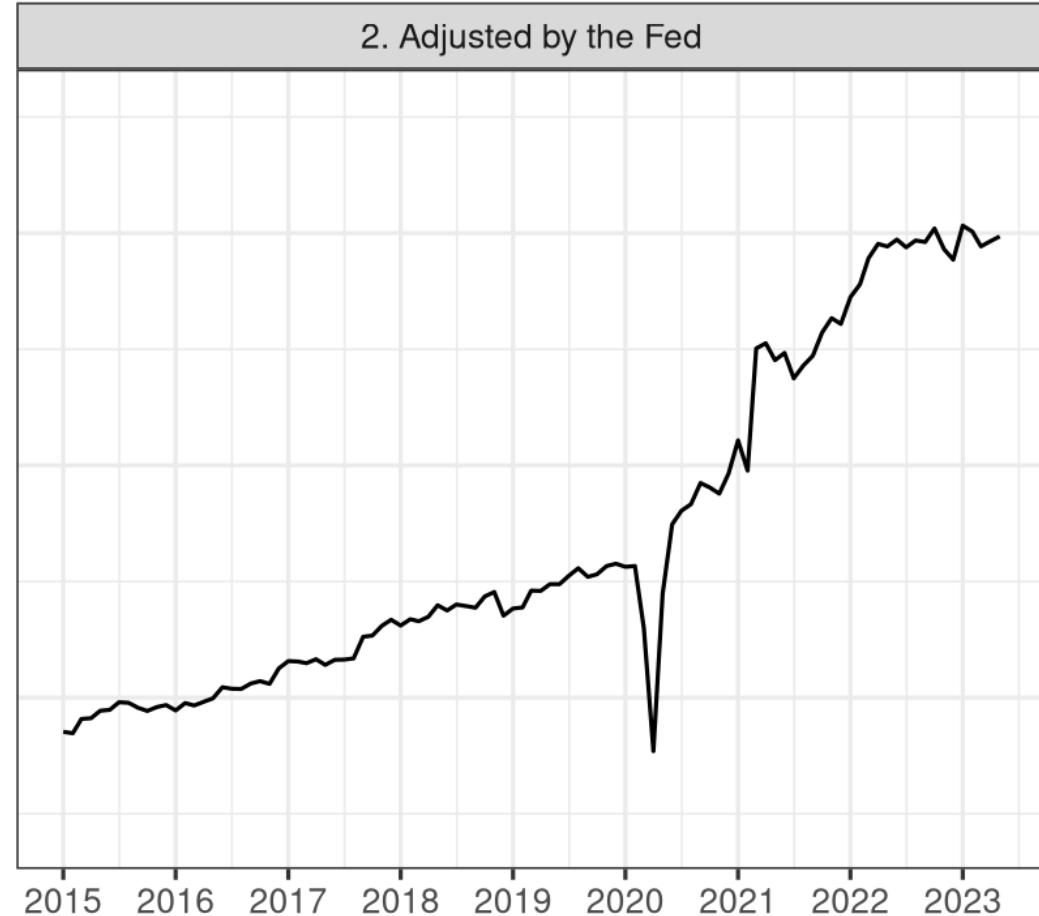
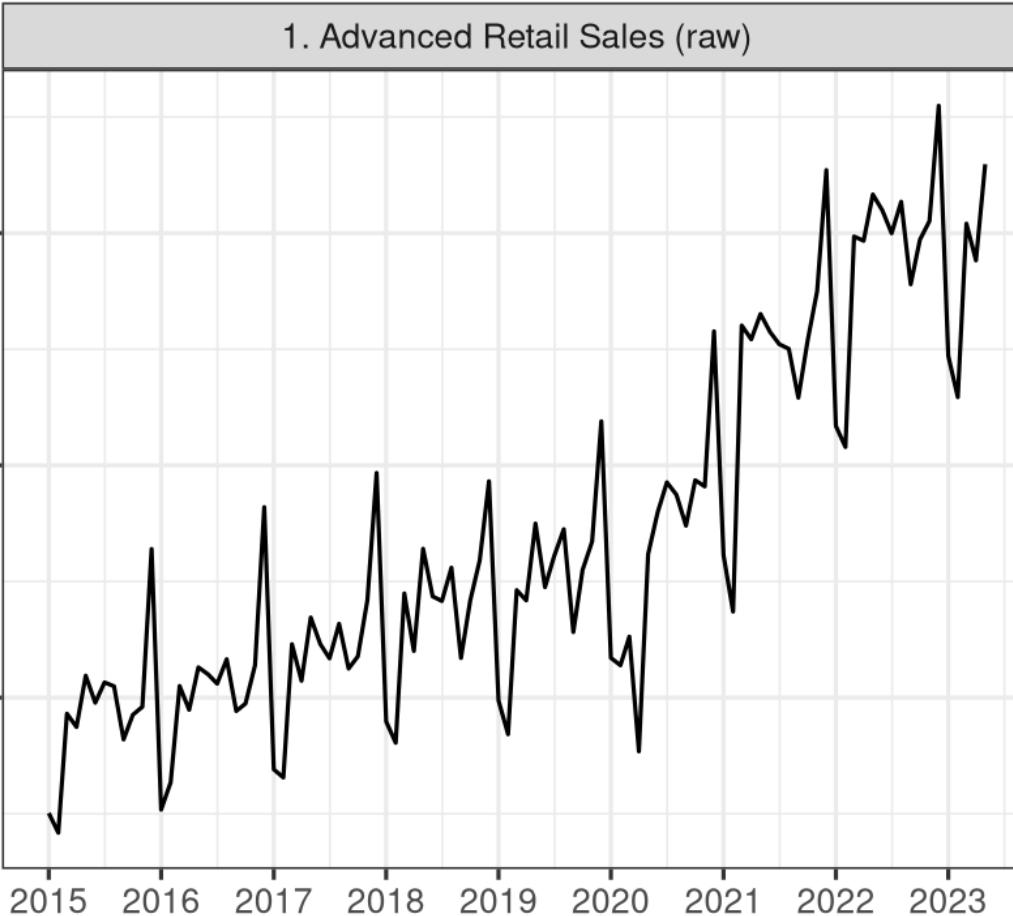
What is the best forecast model of GDP rates in MO?

- Naïve
- MA (3), or
- Weighted MA (3)

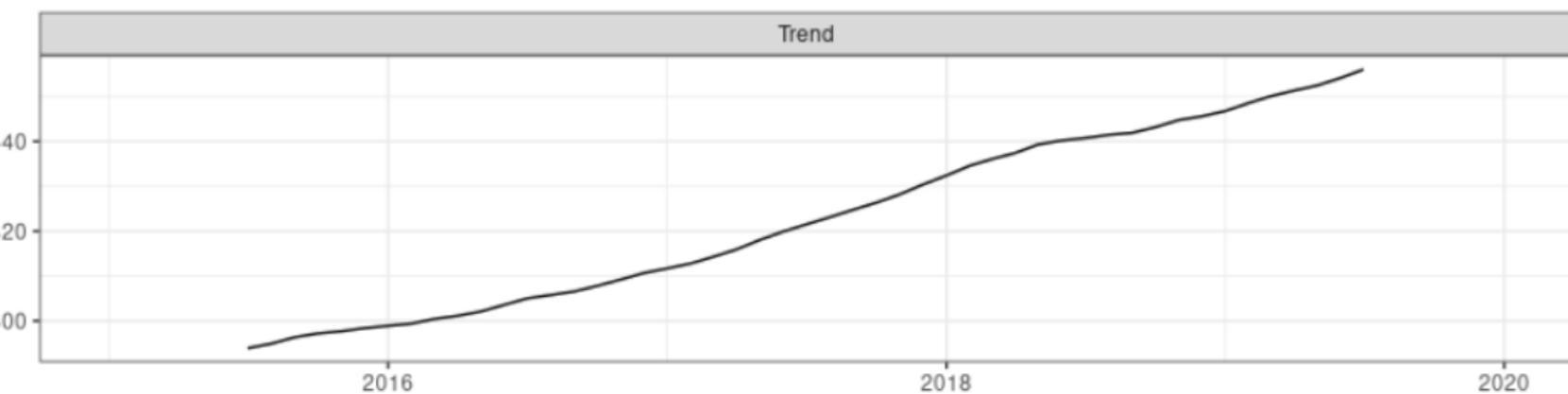
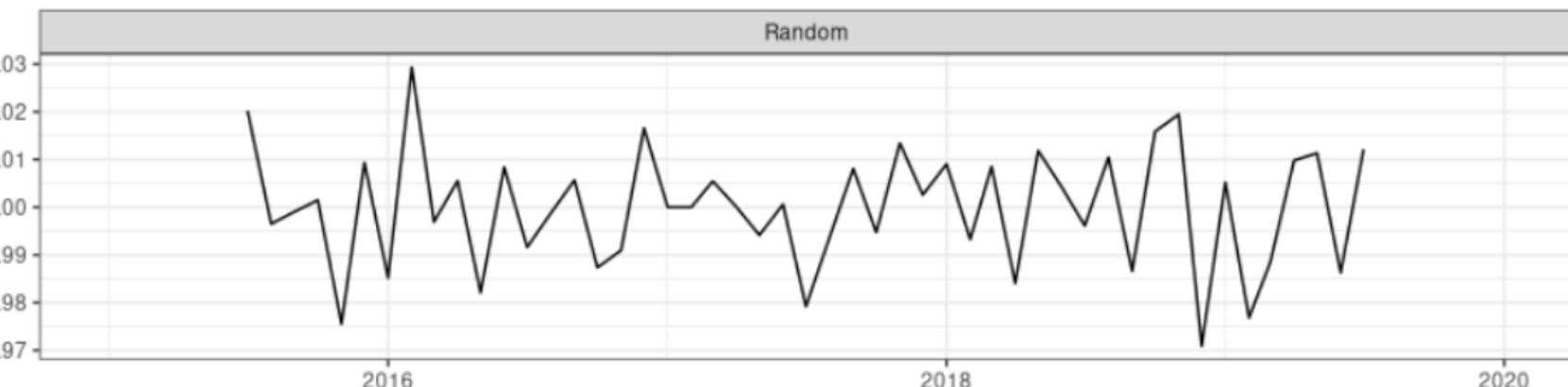
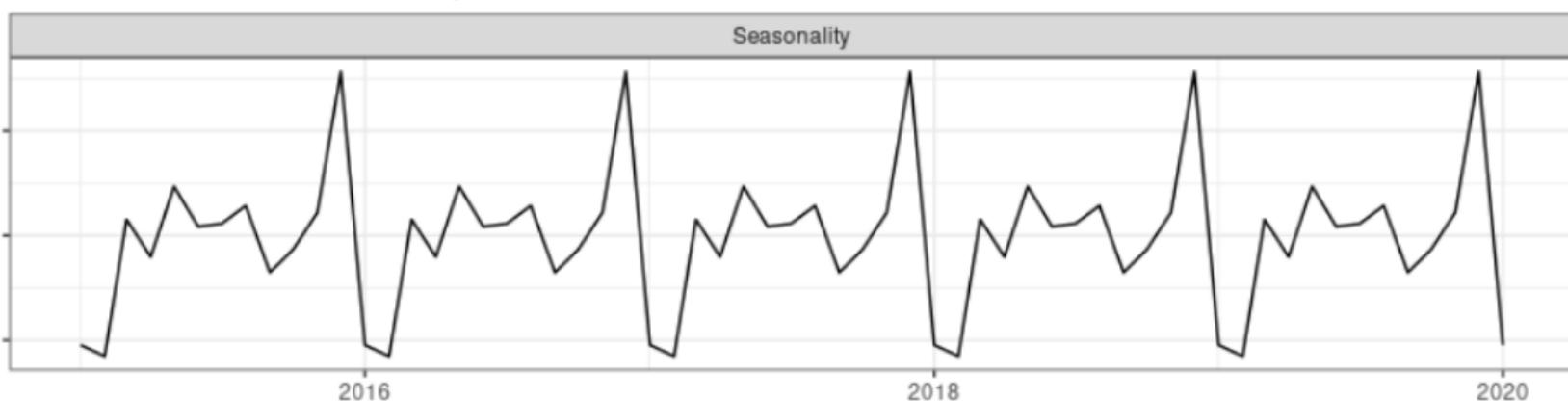
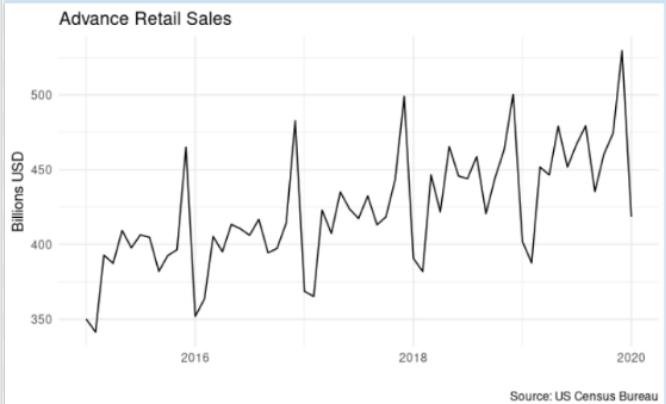
Predict the next year and calculate the MSE

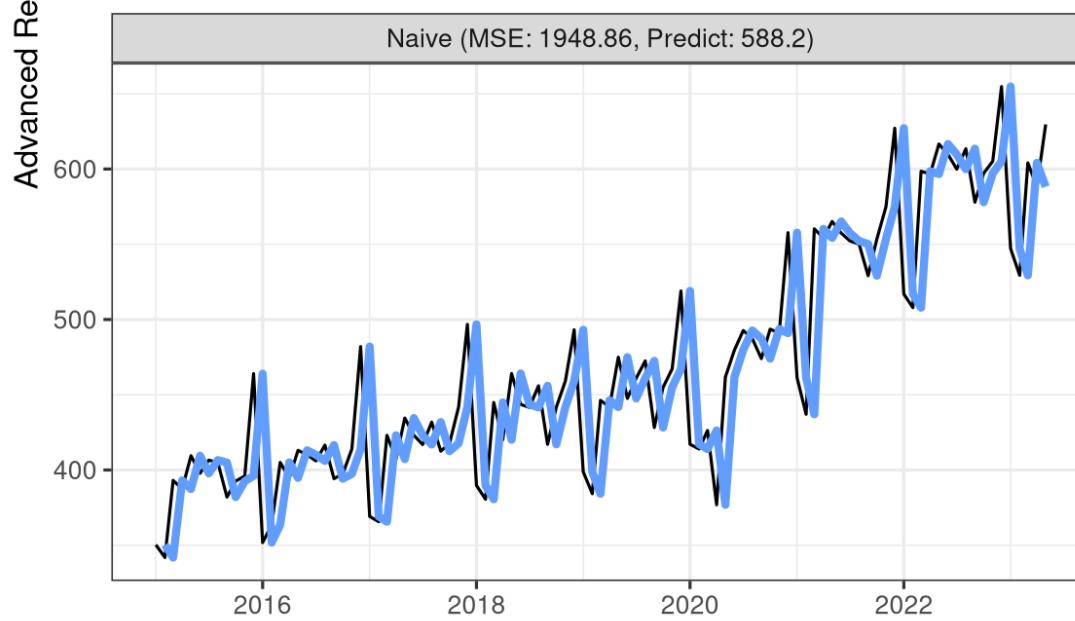
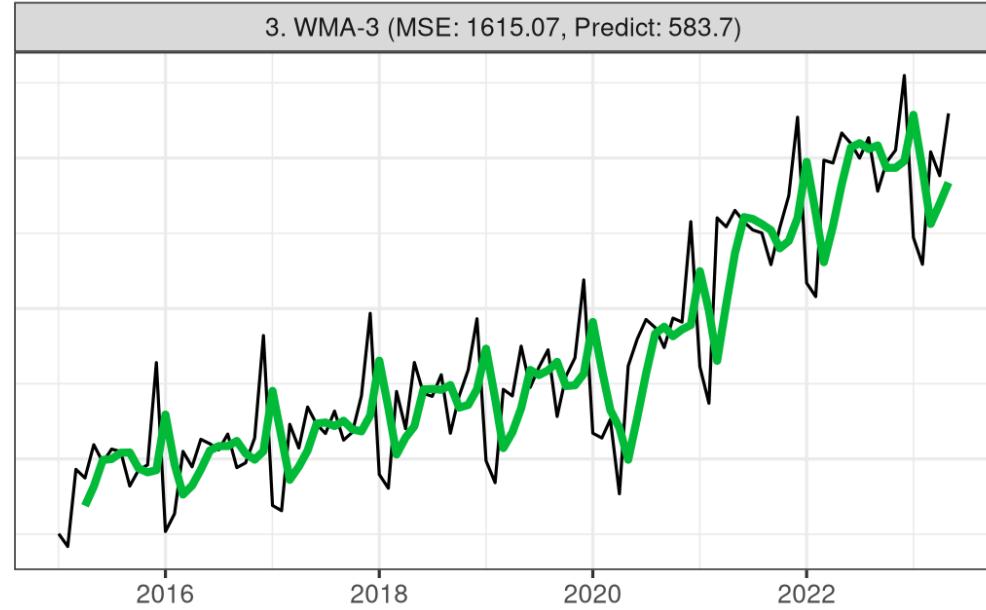
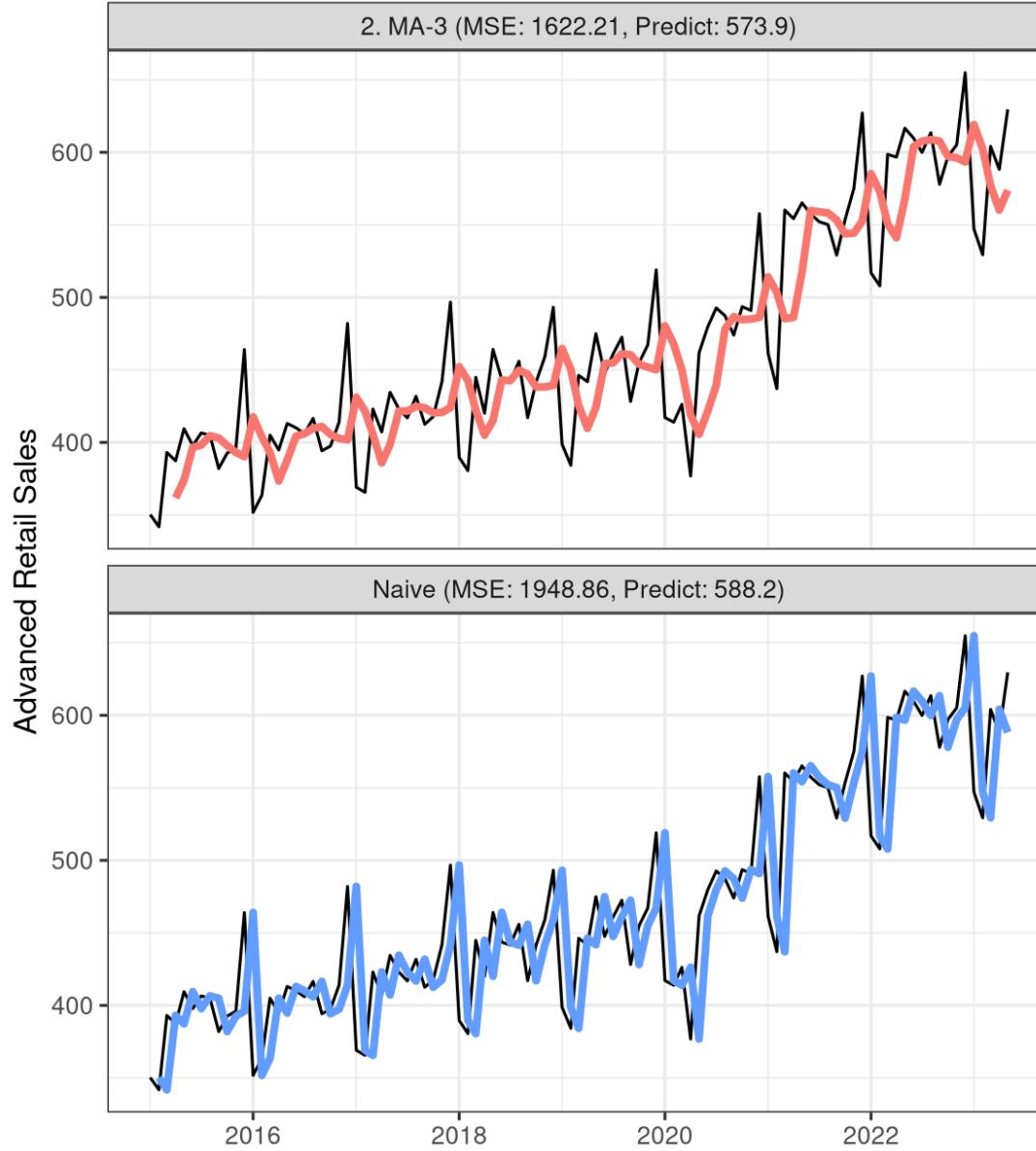


Billions USD



Advance Retail Sales: Decomposition





Linear Model 1 of Advanced Retail Sales

date	Retail Sales	Time
2015-01-01	350.29	1
2015-02-01	341.83	2
2015-03-01	393.09	3
2015-04-01	387.36	4
2015-05-01	409.45	5
2015-06-01	397.81	6
2015-07-01	406.55	7
2015-08-01	404.92	8

Create a "time" variable and regress advanced retail sales on it

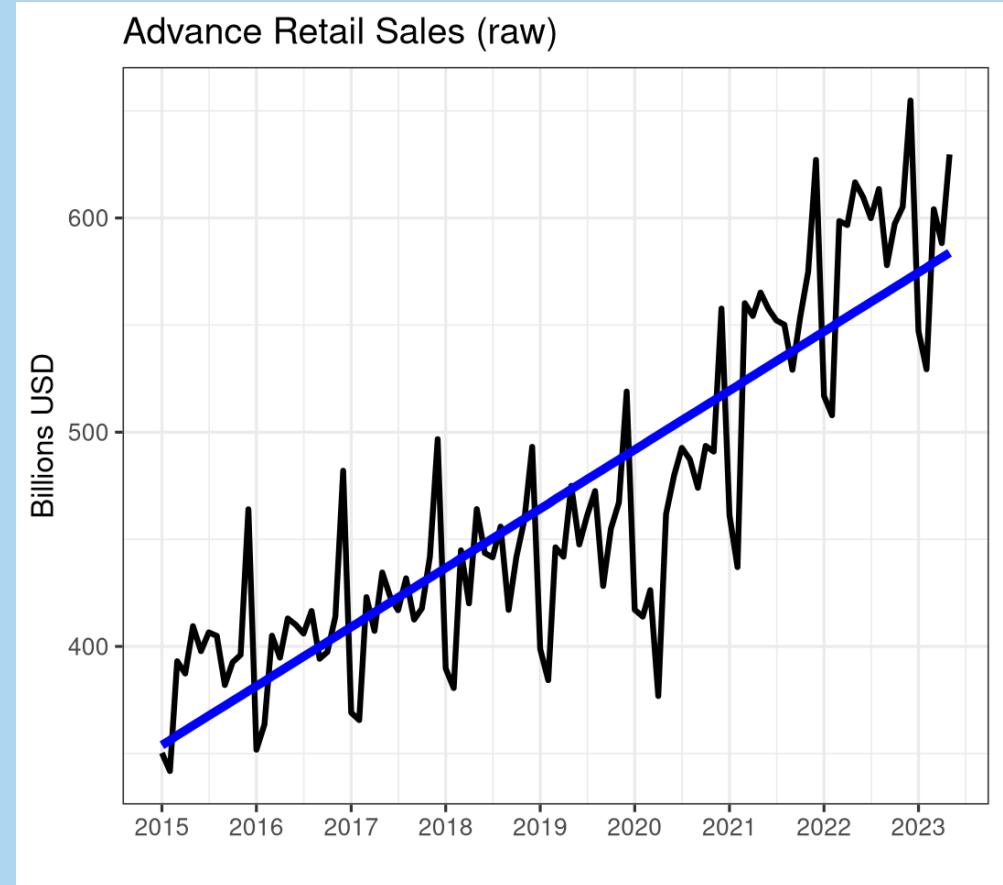
Linear Model 1 of Advanced Retail Sales

Retail Sales	
Time	2.30*
	(0.14)
Constant	351.60*
	(7.93)
Observations	101
Adjusted R ²	0.74
Residual Std. Error	39.56 (df = 99)
F Statistic	289.80* (df = 1; 99)
Note:	*p<0.05

Retail Sales	
Time	2.30*
	(0.14)
Constant	351.60*
	(7.93)
Observations	101
Adjusted R ²	0.74
Residual Std. Error	39.56 (df = 99)
F Statistic	289.80* (df = 1; 99)
<i>Note:</i>	*p<0.05

$$\text{Retail Sales} = 351.6 + 2.3 * (\text{Time})$$

Retail Sales	
Time	2.30*
	(0.14)
Constant	351.60*
	(7.93)
Observations	101
Adjusted R ²	0.74
Residual Std. Error	39.56 (df = 99)
F Statistic	289.80* (df = 1; 99)
Note:	*p<0.05



RESIDUAL OUTPUT

<i>Observation</i>	<i>Predicted advance_retail_sales</i>	<i>Residuals</i>
1	377.3008973	-27.23389727
2	378.8458612	-37.3868612
3	380.3908251	12.45717486
4	381.9357891	5.416210929
5	383.480753	25.89524699
6	385.0257169	12.72628306
7	386.5706809	19.82231913
8	388.1156448	16.61335519
9	389.6606087	-7.640608743
10	391.2055727	1.339427322
11	392.7505366	3.739463388
12	394.2955005	70.66649945

✓ fx =J25^2

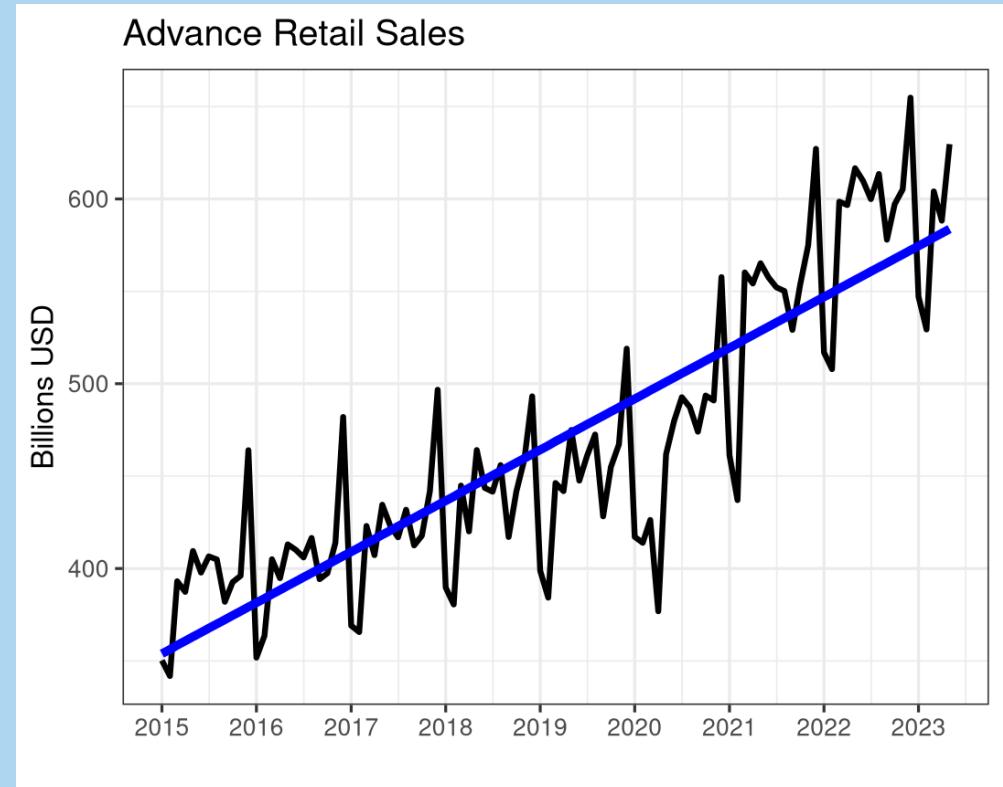
H	I	J	K
<i>Observation</i>	<i>Predicted advance_retail_sales</i>	<i>Residuals</i>	<i>Error^2</i>
1	377.3008973	-27.23389727	741.6851604
2	378.8458612	-37.3868612	1397.777391
3	380.3908251	12.45717486	155.1812056
4	381.9357891	5.416210929	29.33534083
5	383.480753	25.89524699	670.5638169
6	385.0257169	12.72628306	161.9582805
7	386.5706809	19.82231913	392.9243355
8	388.1156448	16.61335519	276.0035707
9	389.6606087	-7.640608743	58.37890197
10	391.2055727	1.339427322	1.794065552
11	392.7505366	3.739463388	13.98358643
12	394.2955005	70.66649945	4993.754145

fx =AVERAGE(K25:K84)

H	I	J	K	L
Observation	Predicted_advance_retail_sales	Residuals	Error^2	MSE
1	377.3008973	-27.23389727	741.6851604	820.403628
2	378.8458612	-37.3868612	1397.777391	
3	380.3908251	12.45717486	155.1812056	
4	381.9357891	5.416210929	29.33534083	
5	383.480753	25.89524699	670.5638169	
6	385.0257169	12.72628306	161.9582805	
7	386.5706809	19.82231913	392.9243355	
8	388.1156448	16.61335519	276.0035707	
9	389.6606087	-7.640608743	58.37890197	
10	391.2055727	1.339427322	1.794065552	
11	392.7505366	3.739463388	13.98358643	
12	394.2955005	70.66649945	4993.754145	

Summer 2023 MSE = 1533.89

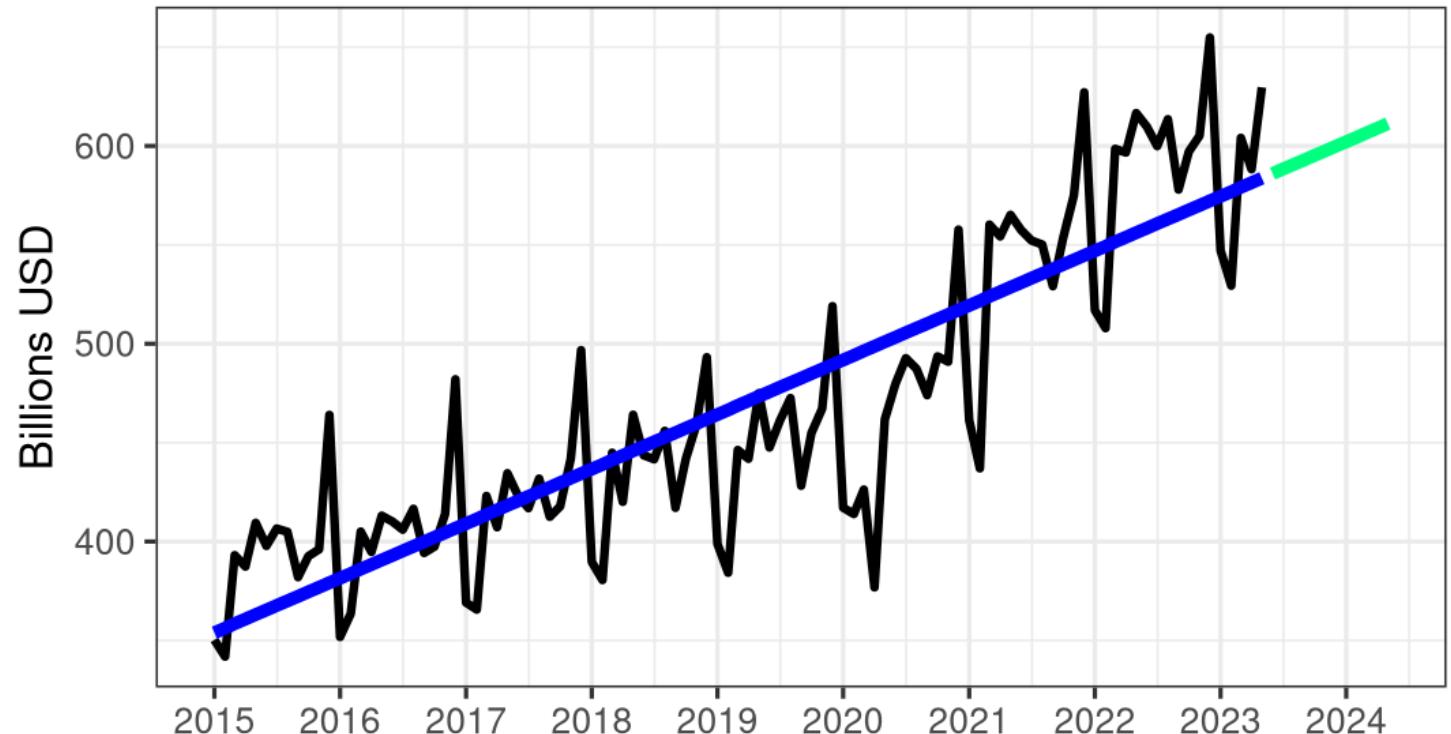
Method	MSE
Naive	1948.9
MA-3	1622.2
Weighted-MA3	1615.1
Linear OLS	1533.9



Retail Sales	
Time	2.30*
	(0.14)
Constant	351.60*
	(7.93)
Observations	101
Adjusted R ²	0.74
Residual Std. Error	39.56 (df = 99)
F Statistic	289.80* (df = 1; 99)
<i>Note:</i>	*p<0.05

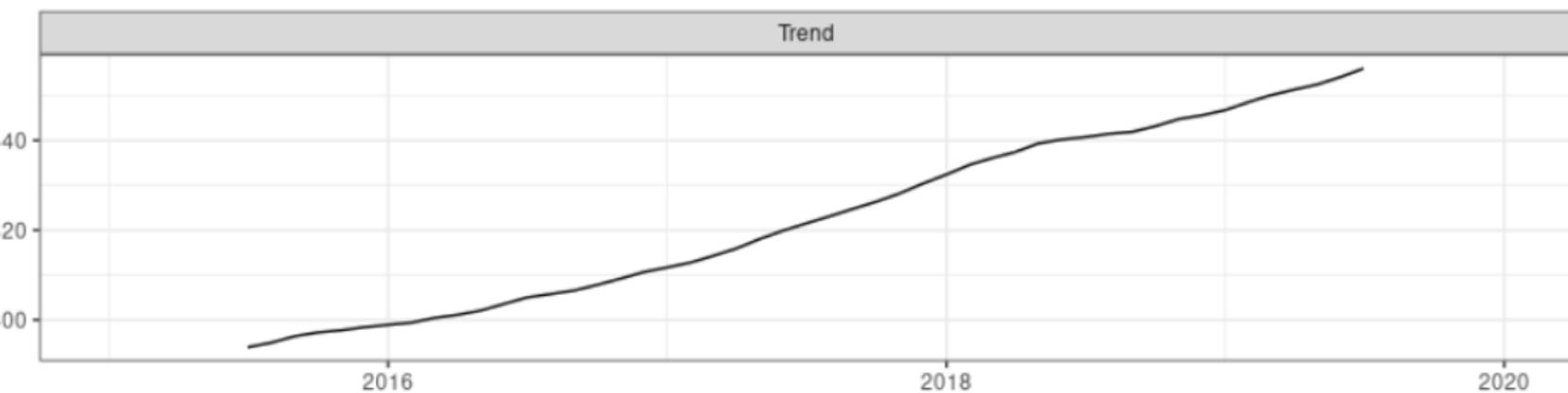
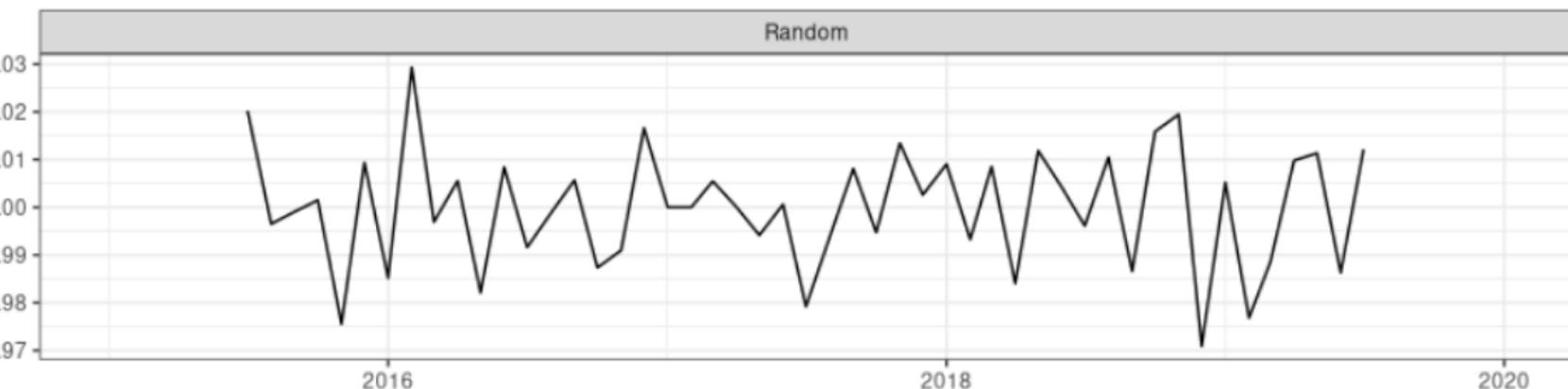
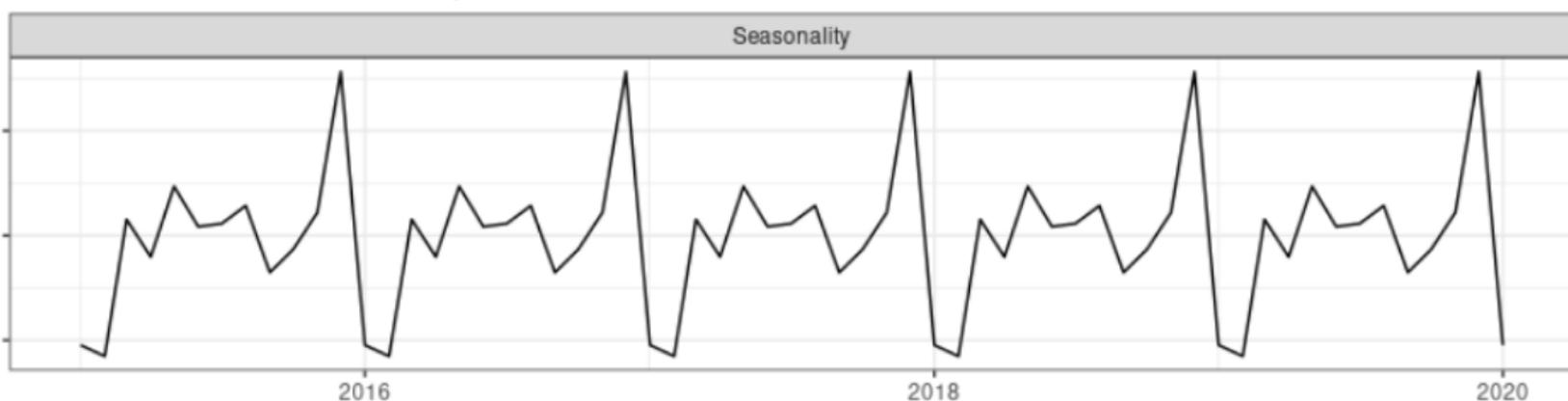
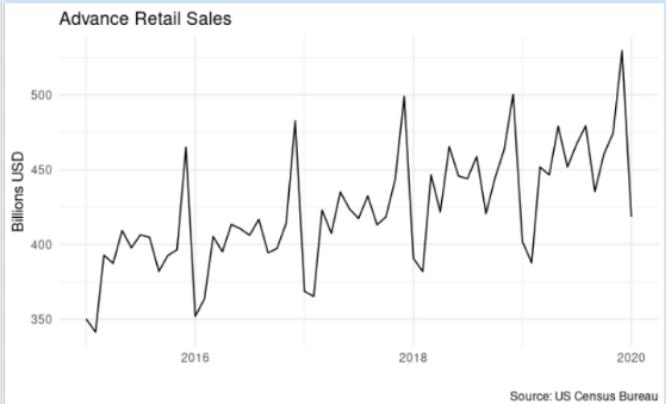
$$\text{Retail Sales} = 351.6 + 2.3 * (\text{Time})$$

Advance Retail Sales

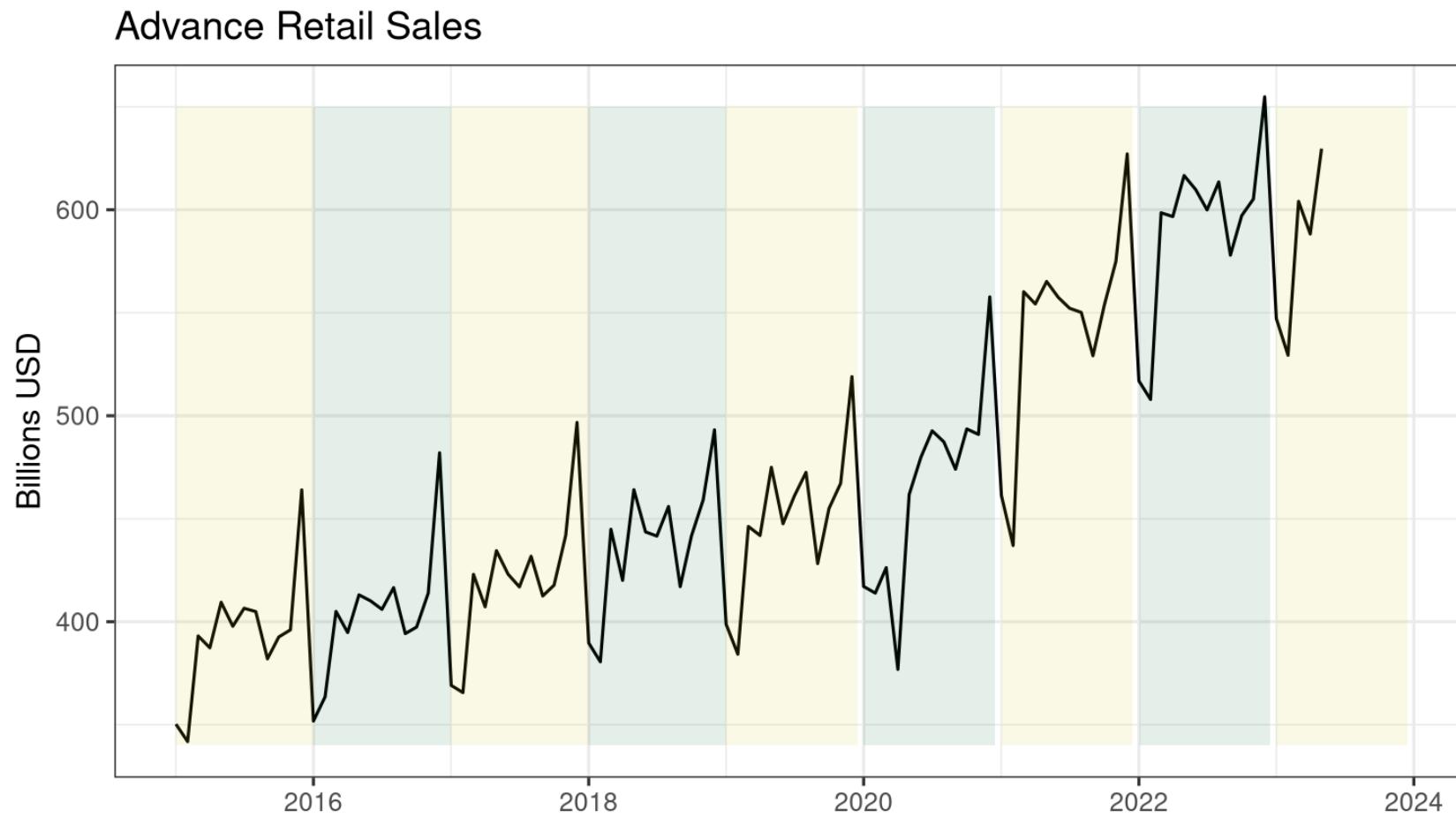


Time	102	103	104	105	106	107	108	109	110	111	112	113
Predictions	586	588	591	593	595	598	600	602	604	607	609	611

Advance Retail Sales: Decomposition

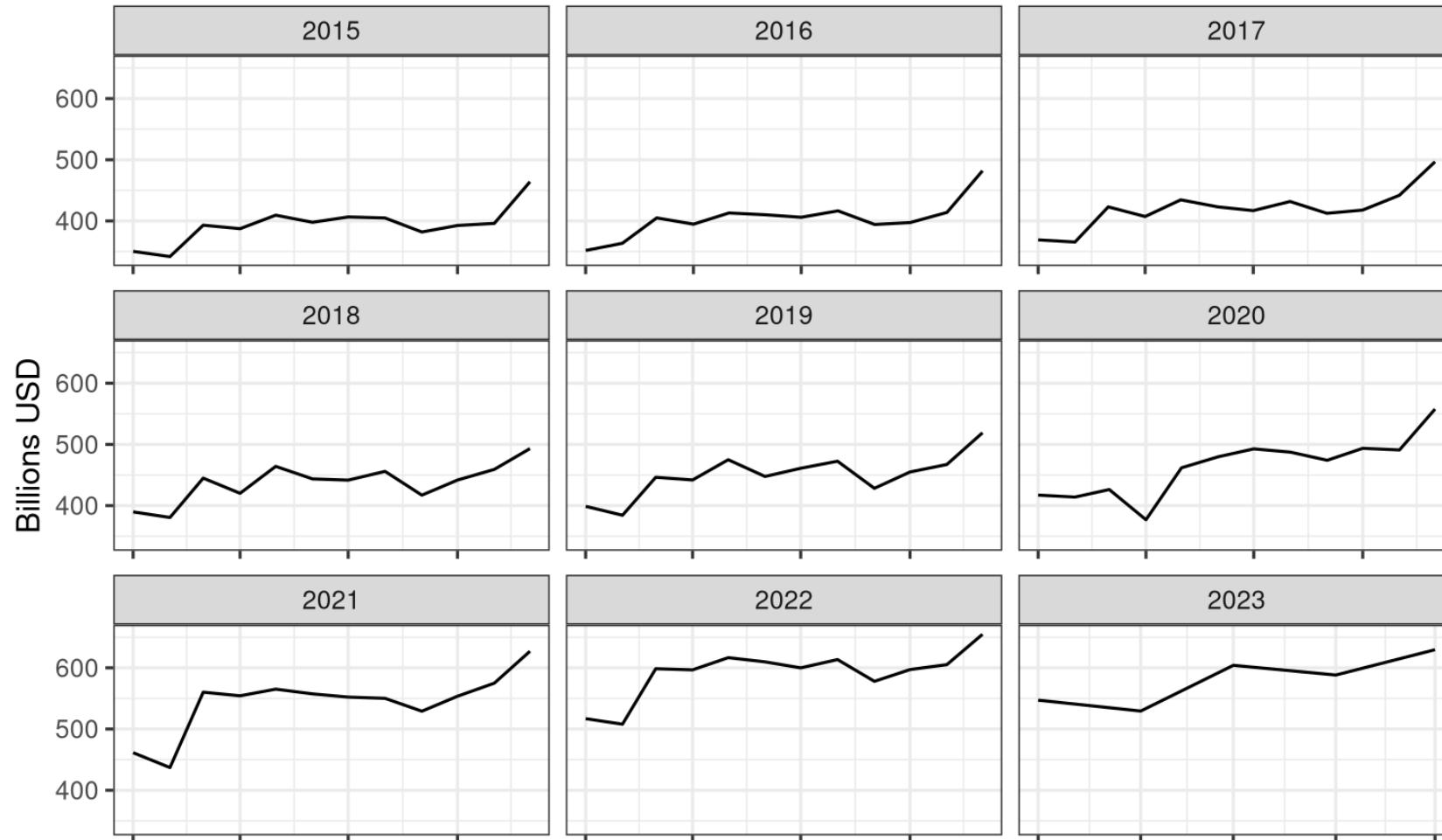


"Seasonality is a characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year" (Investopedia 2020).

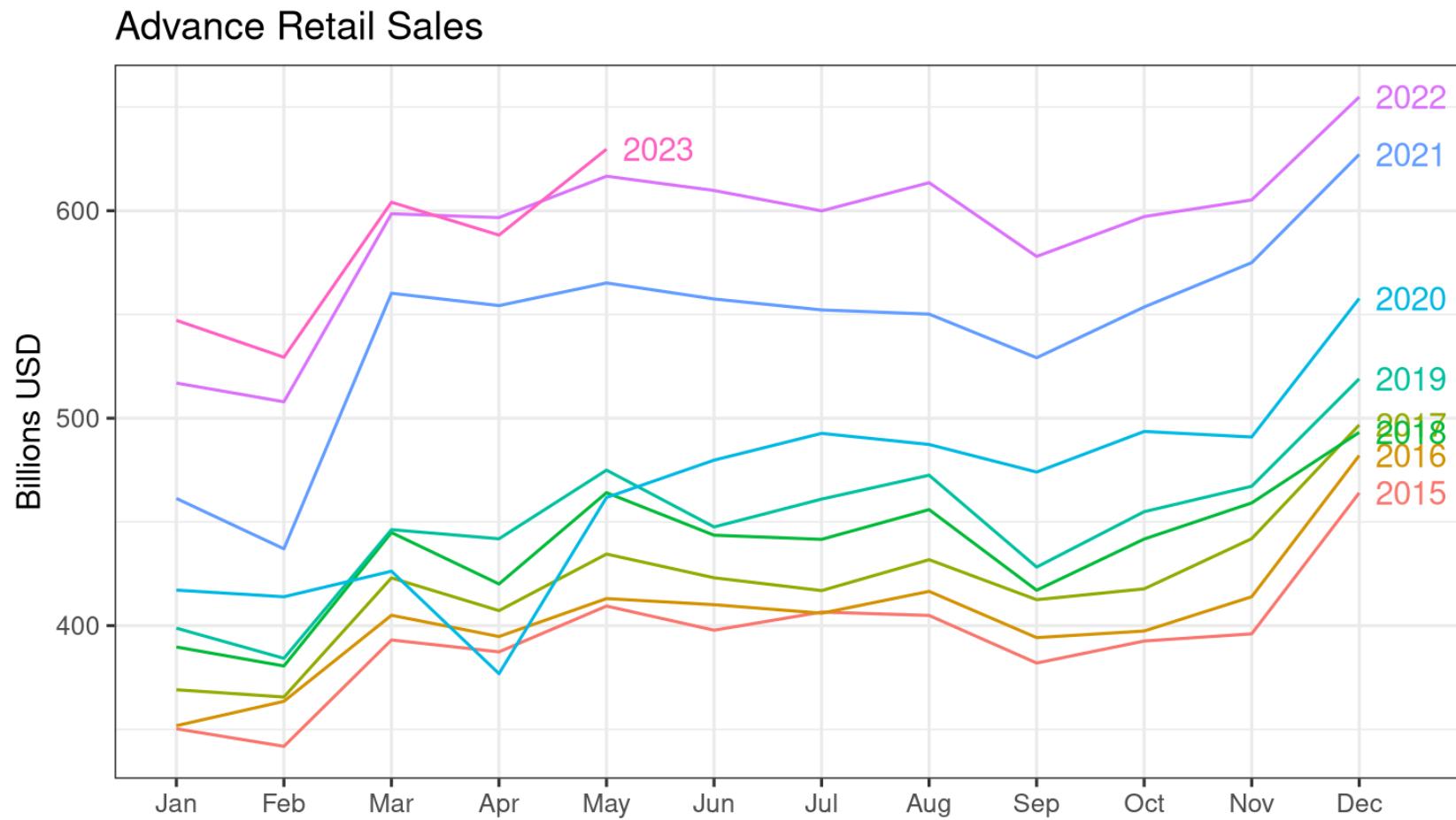


Source: U.S. Census Bureau

"Seasonality is a characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year" (Investopedia 2020).



"Seasonality is a characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year" (Investopedia 2020).



Source: U.S. Census Bureau

Model 2: Linear with Seasonal Dummies

1. Fit the model

- Time period = 1:101
- Spring = '1' if Apr, May, Jun
- Summer = '1' if Jul, Aug, Sep
- Fall = '1' if Oct, Nov, Dec

2. Visualize the model (line plot)

3. Predict the next 12 months

Fitting Linear Trend Models with OLS

1	date	year	month	advance_retail_sales	Time	Spring	Summer	Fall
2	2015-01-01	2015	1	350.067		1		
3	2015-02-01	2015	2	341.459		2		
4	2015-03-01	2015	3	392.848		3		
5	2015-04-01	2015	4	387.352		4		
6	2015-05-01	2015	5	409.376		5		
7	2015-06-01	2015	6	397.752		6		
8	2015-07-01	2015	7	406.393		7		
9	2015-08-01	2015	8	404.729		8		
10	2015-09-01	2015	9	382.02		9		
11	2015-10-01	2015	10	392.545		10		
12	2015-11-01	2015	11	396.49		11		
13	2015-12-01	2015	12	464.962		12		

Fitting Linear Trend Models with OLS

1	date	year	month	advance_retail_sales	Time	Spring	Summer	Fall
2	2015-01-01	2015	1	350.067	1	1	0	
3	2015-02-01	2015	2	341.459	2	0	0	
4	2015-03-01	2015	3	392.848	3	0		
5	2015-04-01	2015	4	387.352	4	1		
6	2015-05-01	2015	5	409.376	5	1		
7	2015-06-01	2015	6	397.752	6	1		
8	2015-07-01	2015	7	406.393	7	0		
9	2015-08-01	2015	8	404.729	8	0		
10	2015-09-01	2015	9	382.02	9	0		
11	2015-10-01	2015	10	392.545	10	0		
12	2015-11-01	2015	11	396.49	11	0		
13	2015-12-01	2015	12	464.962	12	0		
	2016-01-01	2016	1	351.89	13			

Fitting Linear Trend Models with OLS

	A	B	C	D	E	F	G	H
1	date	year	month	advance_retail_sales	Time	Spring	Summer	Fall
2	2015-01-01	2015	1	350.067	1	0	0	0
3	2015-02-01	2015	2	341.459	2	0	0	0
4	2015-03-01	2015	3	392.848	3	0	0	0
5	2015-04-01	2015	4	387.352	4	1	0	0
6	2015-05-01	2015	5	409.376	5	1	0	0
7	2015-06-01	2015	6	397.752	6	1	0	0
8	2015-07-01	2015	7	406.393	7	0	0	1
9	2015-08-01	2015	8	404.729	8	0	0	1
10	2015-09-01	2015	9	382.02	9	0	0	1
11	2015-10-01	2015	10	392.545	10	0	0	0
12	2015-11-01	2015	11	396.49	11	0	0	0
13	2015-12-01	2015	12	464.962	12	0	0	0
14	2016-01-01	2016	1	351.89	13			

Fitting Linear Trend Models with OLS

1	date	year	month	advance_retail_sales	Time	Spring	Summer	Fall	Winter
2	2015-01-01	2015	1	350.067	1	0	0	0	0
3	2015-02-01	2015	2	341.459	2	0	0	0	0
4	2015-03-01	2015	3	392.848	3	0	0	0	0
5	2015-04-01	2015	4	387.352	4	1	0	0	0
6	2015-05-01	2015	5	409.376	5	1	0	0	0
7	2015-06-01	2015	6	397.752	6	1	0	0	0
8	2015-07-01	2015	7	406.393	7	0	1	0	0
9	2015-08-01	2015	8	404.729	8	0	1	0	0
10	2015-09-01	2015	9	382.02	9	0	1	0	0
11	2015-10-01	2015	10	392.545	10	0	0	0	1
12	2015-11-01	2015	11	396.49	11	0	0	0	1
13	2015-12-01	2015	12	464.962	12	0	0	0	1
14	2016-01-01	2016	1	351.89	13				

Model 2: Linear with Seasonal Dummies

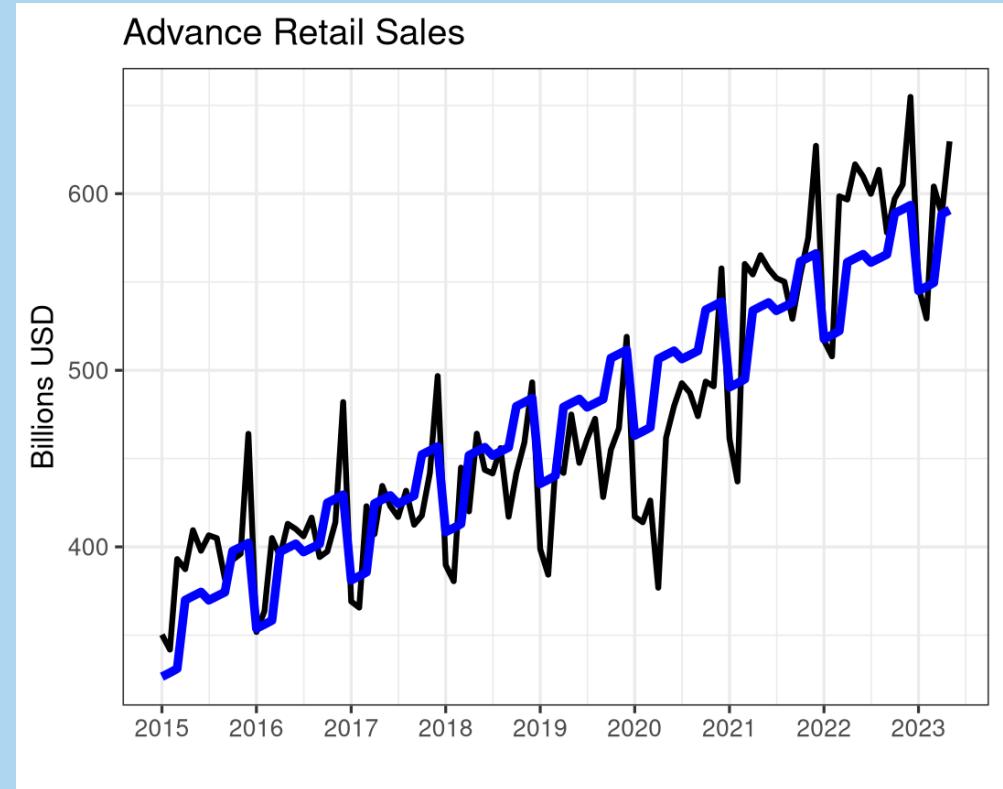
1. Fit the model

- Time period = 1:101
- Spring = '1' if Apr, May, Jun
- Summer = '1' if Jul, Aug, Sep
- Fall = '1' if Oct, Nov, Dec

2. Visualize the model (line plot)

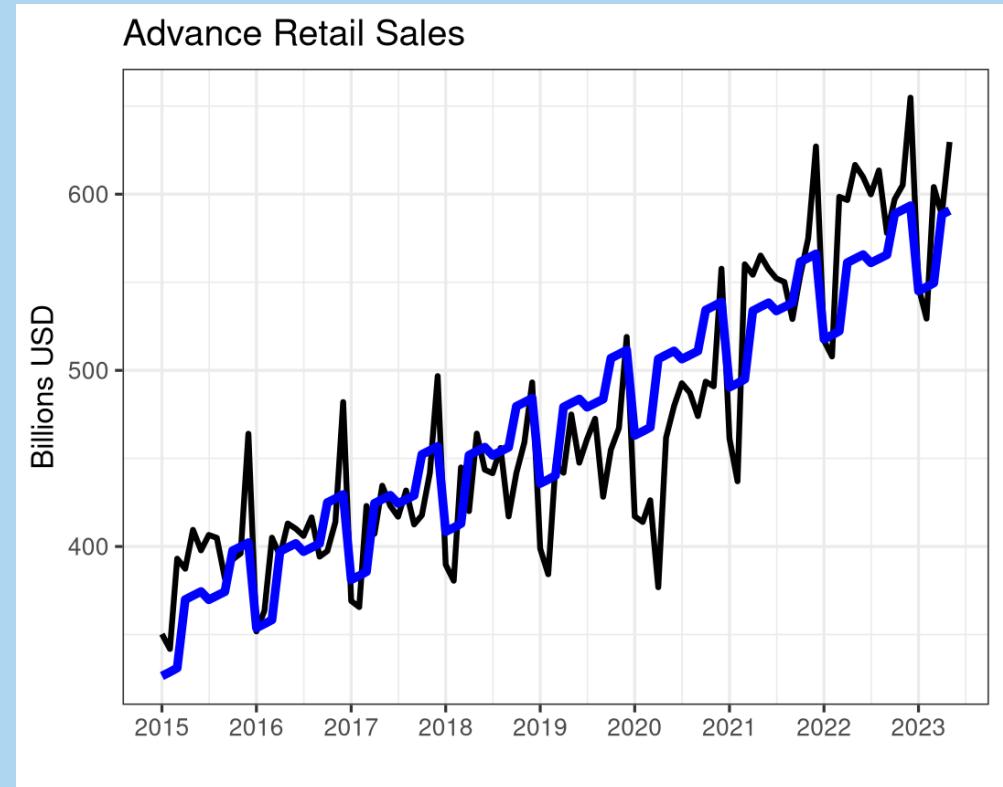
3. Predict the next 12 months

	Retail Sales
time	2.28*
	(0.12)
Spring	36.55*
	(9.70)
Summer	29.64*
	(9.90)
Fall	50.64*
	(9.91)
Constant	324.20*
	(9.08)
Observations	101
Adjusted R ²	0.80
Note:	*p<0.05



Time	102	103	104	105	106	107
Predictions	593	588	591	593	616	618
Time	108	109	110	111	112	113
Predictions	621	572	575	577	616	618

Method	MSE
Naive	1948.9
MA-3	1622.2
Weighted-MA3	1615.1
Linear OLS	1533.9
Seasonal OLS	1183.6



Model 3: Linear with Monthly Dummies

1. Fit the model

- Time period = 1:101
- Month dummies (x 11)

2. Visualize the model (line plot)

3. Predict the next 12 months

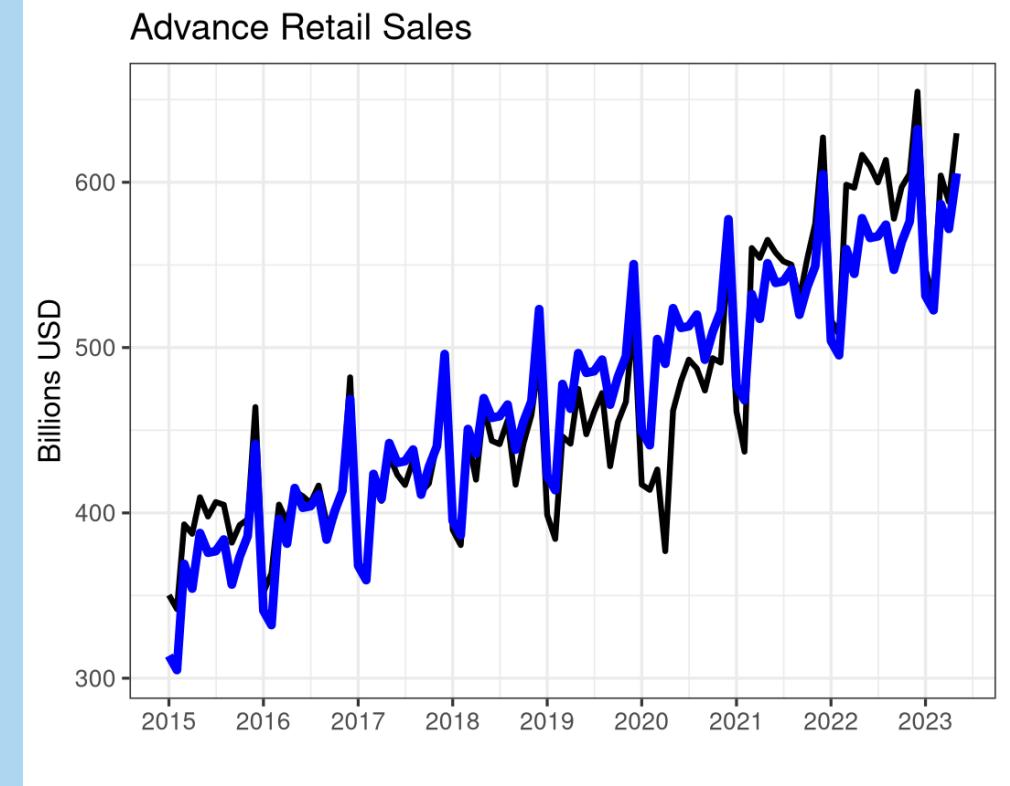
Fitting Linear Trend Models with OLS

	A	B	C	D	E	Feb	Mar	Apr	May	I
1	date	year	month	advance_retail_sales	Time	1	0	0	0	0
2	2015-01-01	2015	1	350.067		1	0	0	0	0
3	2015-02-01	2015	2	341.459		2	1	0	0	0
4	2015-03-01	2015	3	392.848		3	0	1	0	0
5	2015-04-01	2015	4	387.352		4	0	0	1	0
6	2015-05-01	2015	5	409.376		5	0	0	0	1
7	2015-06-01	2015	6	397.752		6	0	0	0	0
8	2015-07-01	2015	7	406.393		7	0	0	0	0
9	2015-08-01	2015	8	404.729		8	0	0	0	0
10	2015-09-01	2015	9	382.02		9	0	0	0	0
11	2015-10-01	2015	10	392.545		10	0	0	0	0
12	2015-11-01	2015	11	396.49		11	0	0	0	0
13	2015-12-01	2015	12	464.962		12	0	0	0	0
..	2016-01-01	2016	1	351.80		13				

=====	
advance_retail_sales	

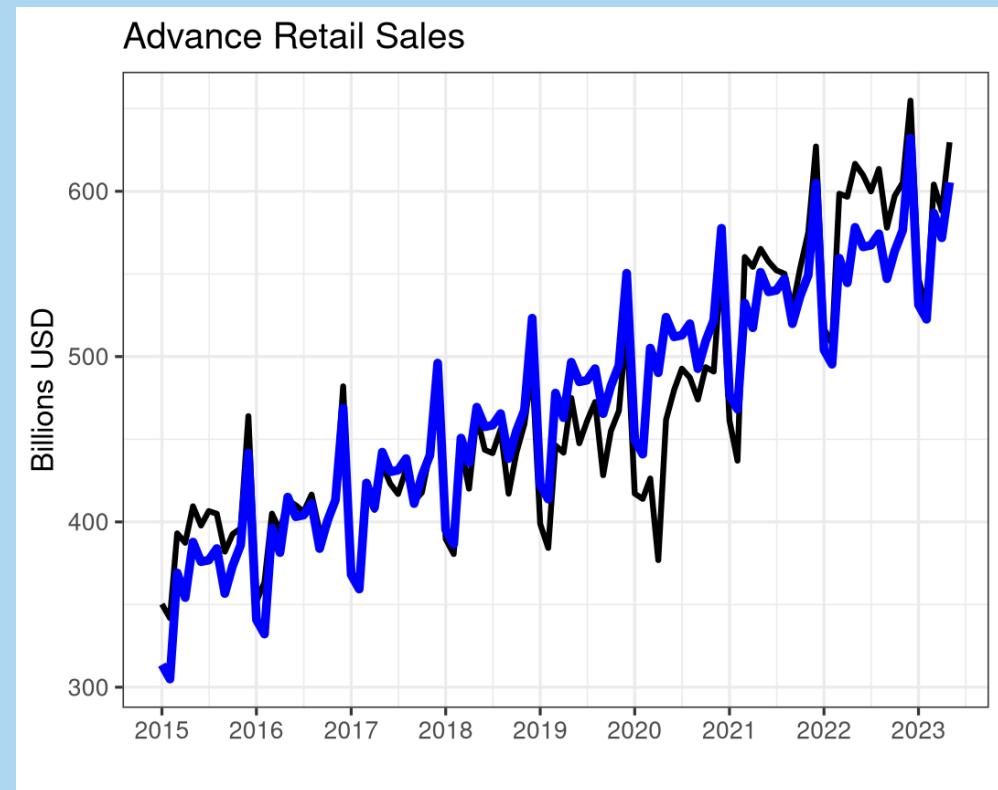
time	1.33* (0.05)
month_nameFebruary	-6.01 (4.31)
month_nameMarch	48.53* (4.31)
month_nameApril	34.95* (4.31)
...Other months omitted...	
Constant	339.55* (3.31)

Observations	60
R2	0.98
Adjusted R2	0.97
Residual Std. Error	6.81 (df = 47)
F Statistic	161.59* (df = 12; 47)
=====	
Note:	*p<0.05

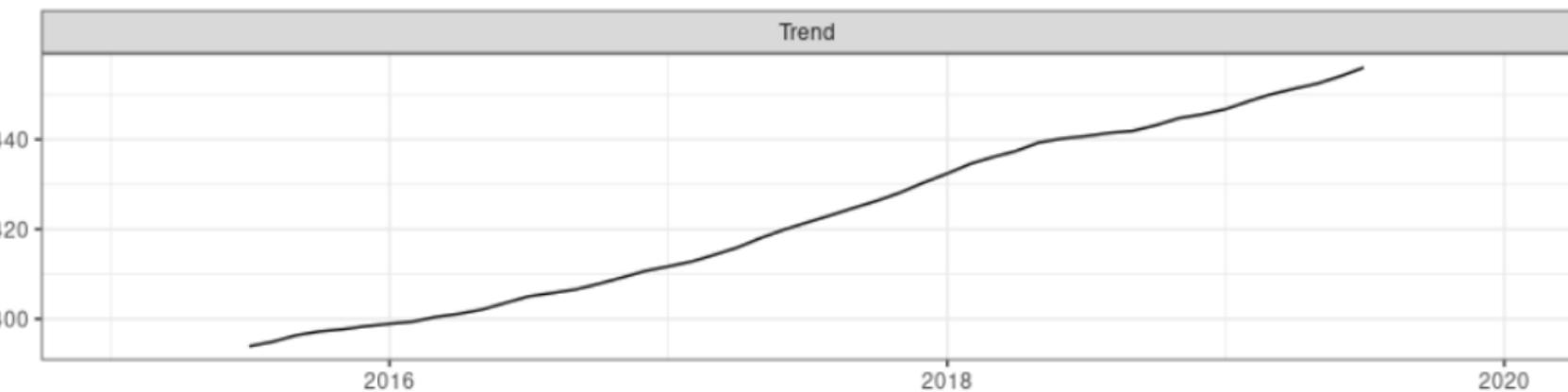
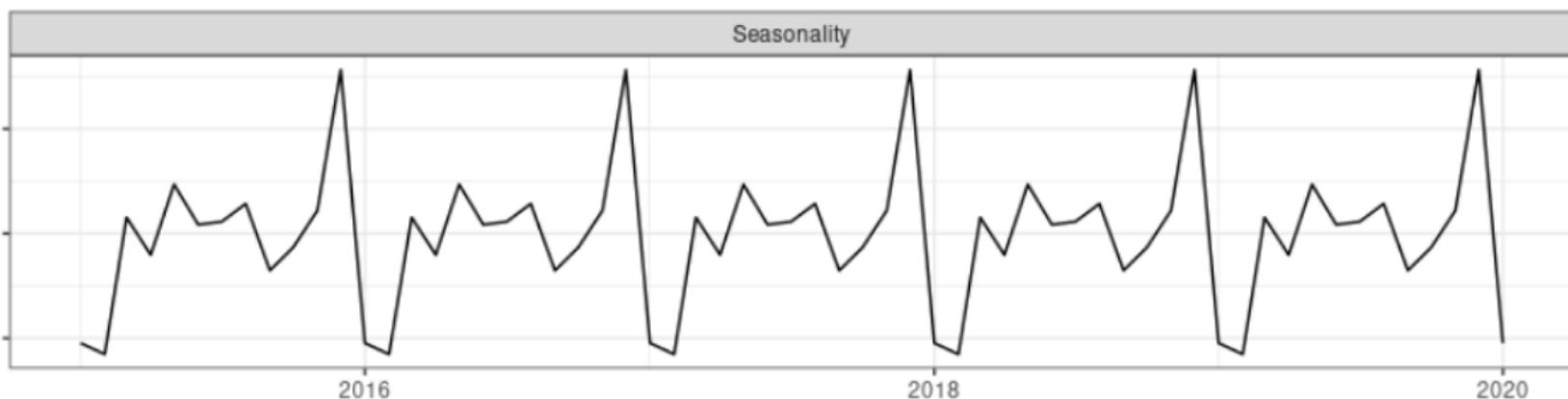
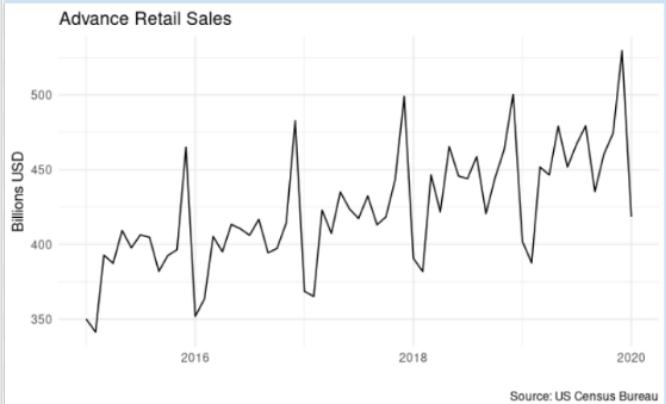


Time	102	103	104	105	106	107
Predictions	594	595	602	574	591	604
Time	108	109	110	111	112	113
Predictions	659	558	550	614	599	633

Method	MSE
Naive	1948.9
MA-3	1622.2
Weighted-MA3	1615.1
Linear OLS	1533.9
Seasonal OLS	1183.6
Monthly OLS	720.9



Advance Retail Sales: Decomposition





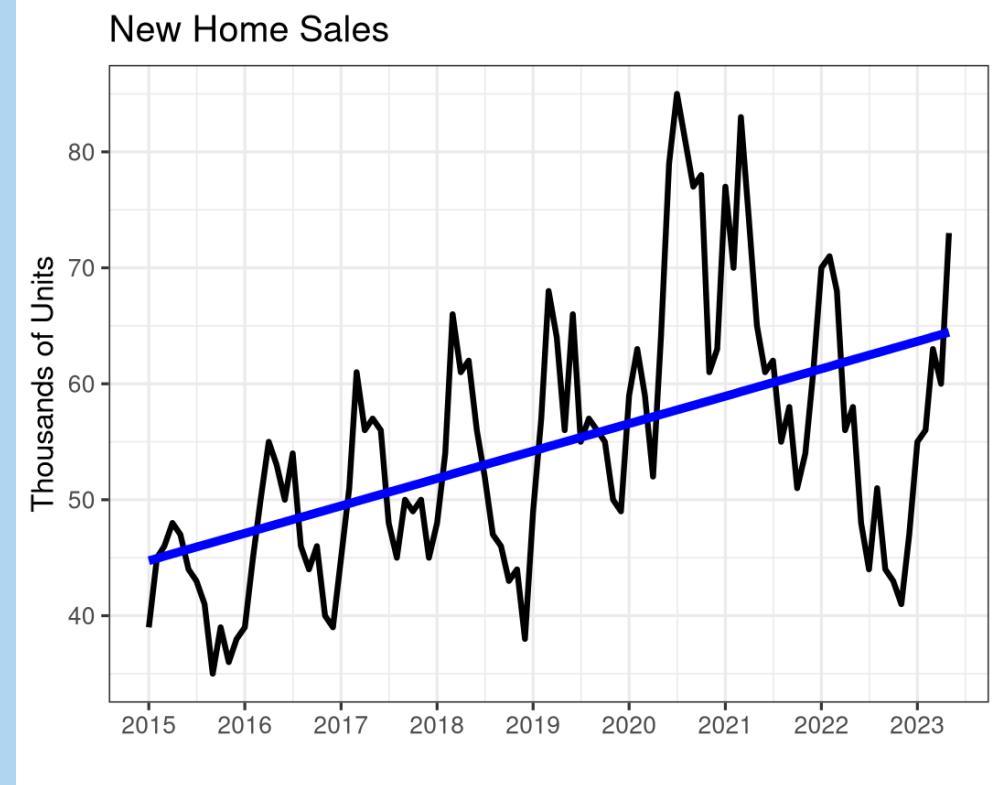
U.S. recessions are shaded; the most recent end date is undecided.

Sources: Census; HUD

fred.stlouisfed.org

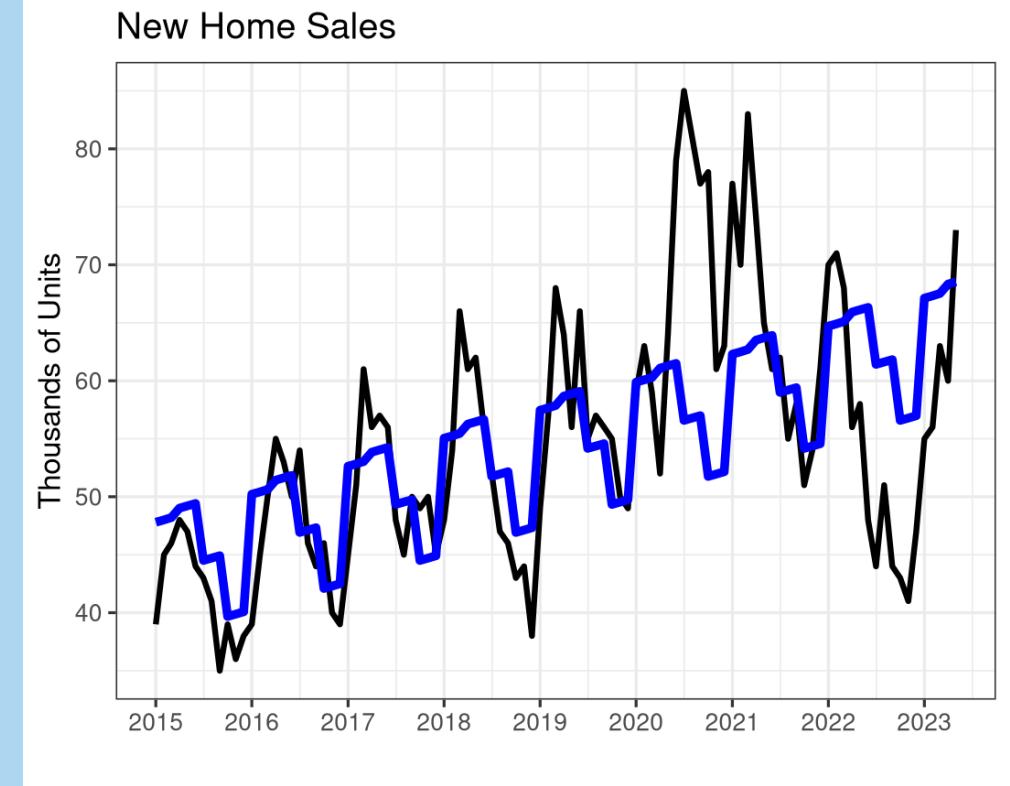
Regress new home sales on:
- Approach 1: Time
- Approach 2: Time and Season

New Home Sales	
time	0.20*
	(0.03)
Constant	44.55*
	(1.96)
Observations	101
Adjusted R ²	0.25
Note:	*p<0.05



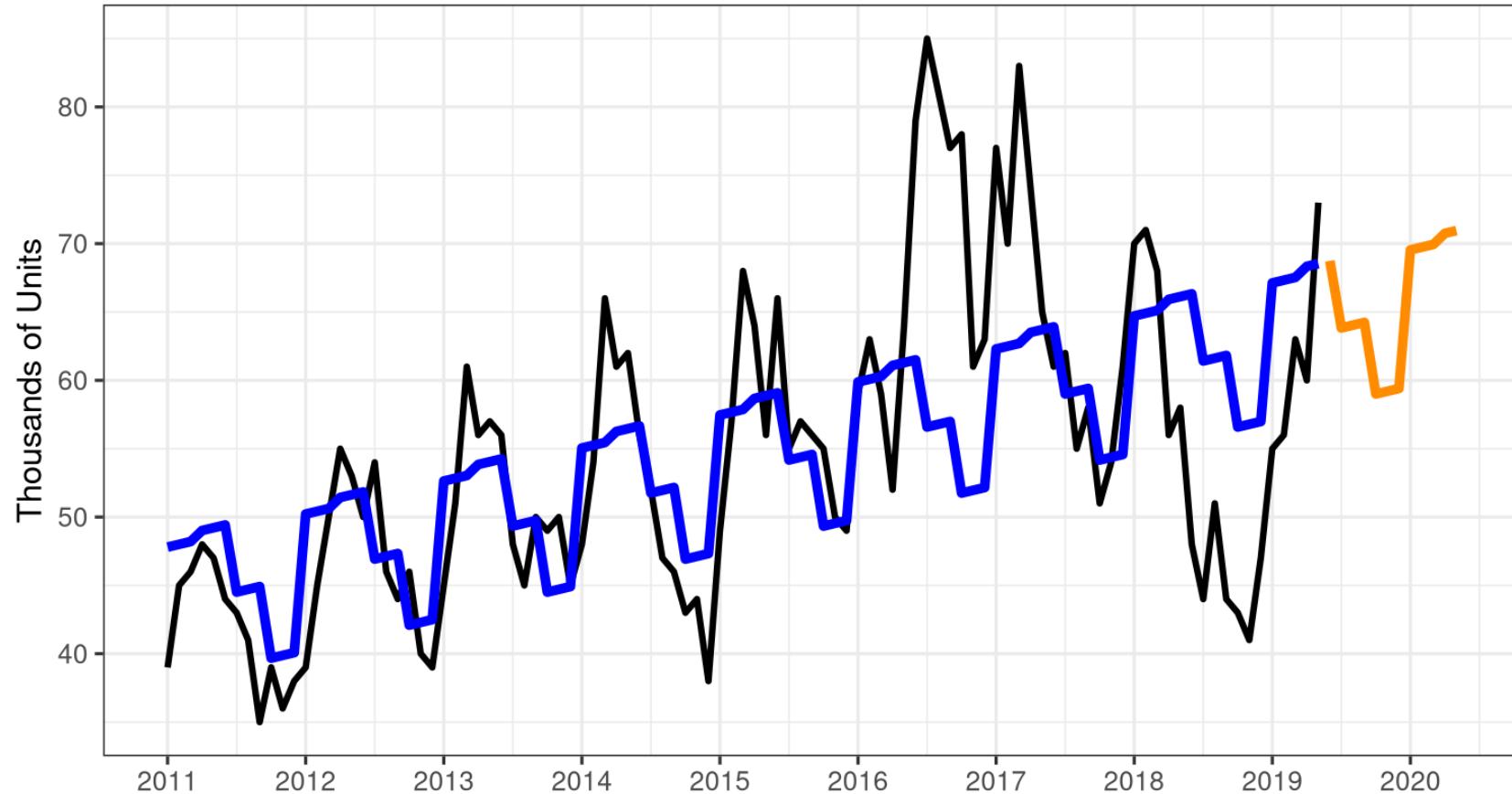
Method	MSE
Linear OLS	93.2

	New Home Sales
time	0.20*
	(0.03)
Spring	0.61
	(2.45)
Summer	-4.50
	(2.50)
Fall	-9.94*
	(2.50)
Constant	47.60*
	(2.29)
Observations	101
Adjusted R ²	0.38
Note:	*p<0.05



Method	MSE
Linear OLS	93.2
Seasonal OLS	75.6

New Home Sales



Time	102.0	103.0	104	105.0	106	107.0	108.0	109.0	110.0	111.0	112.0	113
Predictions	68.7	63.8	64	64.2	59	59.2	59.4	69.5	69.7	69.9	70.8	71

Wrapping Up

This intensive course is designed to help students learn essential aspects of applied data analysis (e.g. visualization, descriptive statistics and modeling) using Excel.