Dozentur
Produktion und Logistik
Priv.-Doz. Dr. Rainer Kleber

Project Report

Lecture Business Forecasting

Mini Case Solution

<u>Team 09</u>

Md Saddam Hossain

ID- 220614

Salamah Akhtar

ID- 231933

Md. Asaduzzaman

ID- 226255

Tahsin Nakib Talukder

ID- 221221

I Table of Contents

1. Introduction

The objective of this case is to perform forecast on German retail sales from April 2020 to December 2020. As part of forecasting, we aim to develop an appropriate forecasting method (additive and multiplicative Holt-Winters and ARIMA) and compare with a benchmark model. Also included are point forecasts and prediction intervals. We also check if chosen method performs better using different estimation and hold-out samples.

1.1 Primary data analysis

We performed decomposition on the data series from Jan 2002 - Mar 2020 to check the components. Figure 1 shows a prominent seasonal component after decomposition. The consistent peaks and troughs at regular time intervals confirming presence of seasonality in the data.   However, the trend remains almost constant and rises from 2015 to 2020.

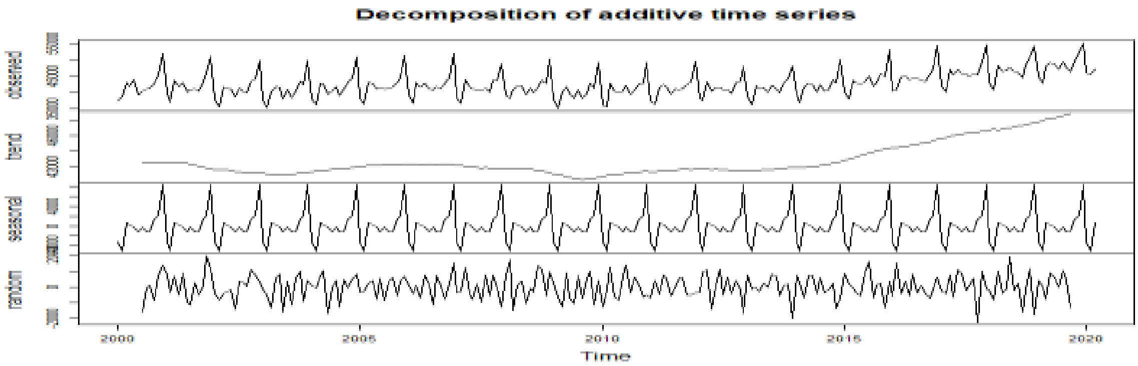

*Figure 1: Decomposition of data serie*s

1.2 Forecasting method selection

Since we found seasonality in our data, we focus on methods which take seasonality in consideration. We did not consider Simple Exponential Smoothing (SES) and Linear Exponential Smoothing (LES) because the produced forecasts fail to predict the seasonality. Therefore, we proceeded with the following forecasting methods:

- Seasonal Naïve (as benchmark)
- Holt-Winters Additive and Multiplicative (with and without trend)
- ARIMA

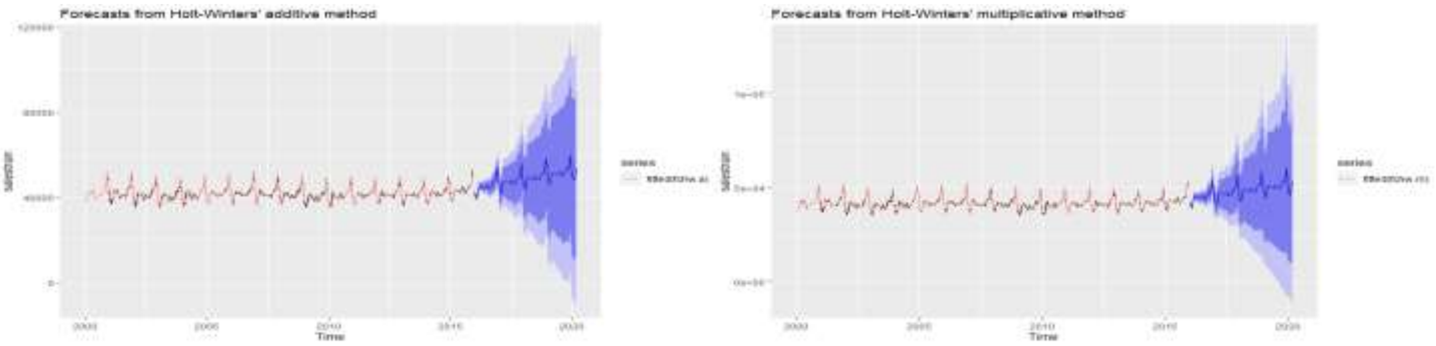

Figure 2: H-W Additive and Multiplicative

We performed Holt’s damped method, Holt-Winters Additive and Multiplicative, both with and without trend. The Holt’s damped method fails to capture the seasonality of the data (Figure A1 in Appendix). The produced plots from H-W (Holt-Winters) additive and multiplicative methods are shown in Figure 2. Both methods produce forecasts that clearly align with the seasonality.

ARIMA: To check the stationarity of the data, we did ndiffs/ nsdiffs and found that one differencing is required. After seasonal differencing, we performed visual diagnostics on ACF and PACF plots and suggested following seasonal ARIMA models:

Models	AIC	AICc	BIC
ARIMA (2,1,0)(0,1,1)	2,964	2,965	2,977
ARIMA (2,1,1)(0,1,1)	2,962	2,962	2,978
ARIMA (2,1,2)(0,1,1)	2,956	**2,957**	2,975

Table 1: Probable ARIMA models with Information Criteria using R

Among the three Information Criteria measures, we chose AICc for performance measure (Table 1). Based on the lowest AICc, we selected ARIMA (2,1,2)(0,1,1).

The residuals of ARIMA (2,1,2) (0,1,1) show bell-shaped histogram which represents normally distributed data. We see very few spikes (crossing the blue line in ACF plot) indicating additional non-seasonal terms could be added in the model (see Figure A2 in Appendix). But this could make the model complex which we do not prefer within our current scope of study.

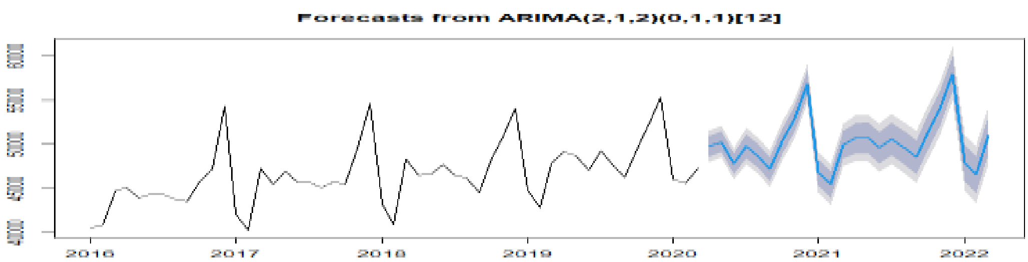

Figure 3: Forecasting with ARIMA (2,1,2) (0,1,1)

1.3 Comparison of performance measures

We chose seasonal naïve as the benchmark method since the data has seasonality. We used hold-out sample to perform rolling forecasting error and found the best-performed model by considering lowest root mean squared error (RMSE).

Forecasting Method	RMSE	MAE	MAPE	Point Forecast
Seasonal Naïve	1,309	1,057	2.5	49,225
SES	4,343	3,353	6.9	43,938
LES (Holt's local)	14,448	12,679	27.2	45,484
LES Damped	5,849	4,930	10.2	41,865
Holt-Winters Additive	2,635	2,195	4.7	41,951
H-W Additive (without trend)	2,647	2,171	4.6	40,718
H-W Multiplicative	2,202	1,706	3.6	40,203
H-W Multiplicative (without trend)	3,467	3,020	6.4	40,204
ARIMA (2,1,2) (0,1,1)	851			49,781

Among all the methods, ARIMA (2,1,2) (0,1,1) performs better in terms of lowest RMSE.

Month/ Year	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Apr-20	49,781	48,636	50,927	48,030	51,533
May-20	50,183	49,035	51,332	48,427	51,940
Jun-20	47,705	46,557	48,854	45,949	49,462
Jul-20	49,972	48,654	51,289	47,957	51,987
Aug-20	48,483	47,161	49,804	46,462	50,503
Sep-20	47,319	45,969	48,668	45,255	49,383
Oct-20	50,600	49,191	52,008	48,446	52,754
Nov-20	52,481	51,060	53,902	50,307	54,655
Dec-20	56,853	55,396	58,310	54,625	59,081

Table 2: Point forecasts and prediction intervals for ARIMA (2,1,2) (0,1,1)

1.4 Using different estimation samples to check forecast performance

We used different estimation sample size in ARIMA (2,1,2) (0,1,1) model and found the results tabulated below:

Estimation Sample	Prediction Interval
2005 Jan - 2010 Dec	[47778.20, 51497.19]
2004 Jan - 2012 Dec	[48009.61, 51438.00]
2008 Jan - 2016 Dec	[48059.82, 51763.15]

2. Conclusion

Based on the above results, we found ARIMA (2,1,2) (0,1,1) model to be the best-fit for the provided data. We consider this result a reasonable one because adding more terms to the ARIMA model would make it complex one which is not preferred.

Appendix

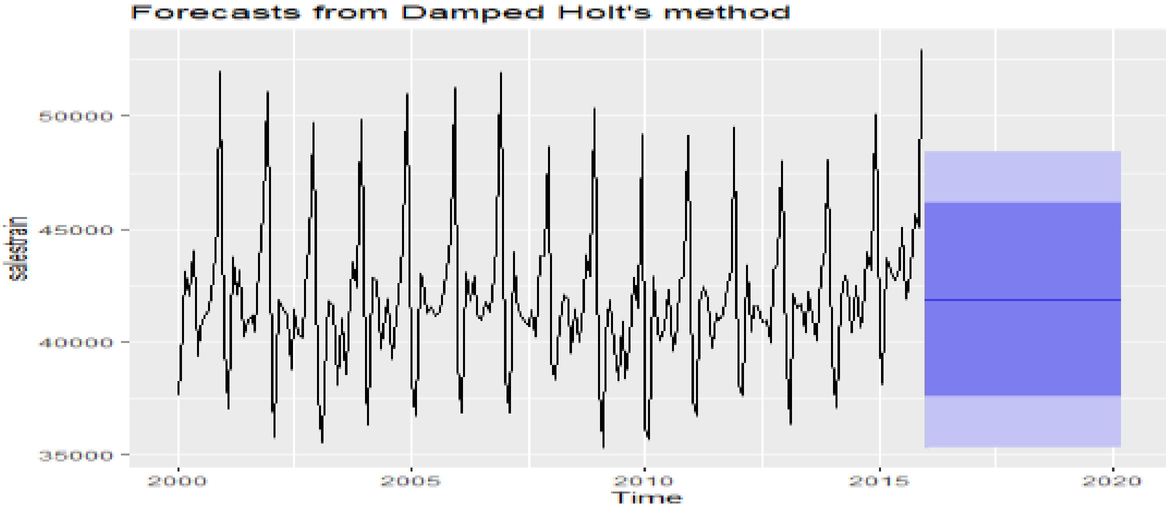

Figure A1: Forecasting with Holt's local method

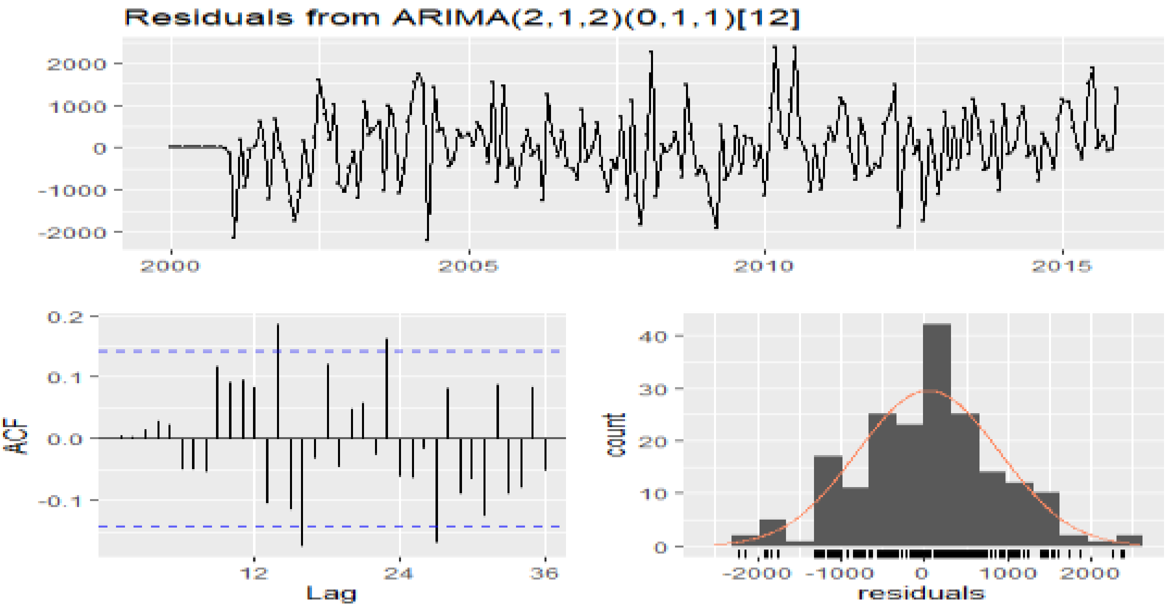

Figure A2: Residuals from selected model ARIMA(2,1,2)(0,1,1)