

**Student name:** Khoi Pham

**Student number:** 574883

**Supervisor:** Dr. Robert Rooderkerk

**Co-reader:** Dr. Clint Pennings

**Company coach:** Joep Breurkens

**Improving forecasting performance: The case of Philips**

MSc Supply Chain Management

Rotterdam School of Management

Erasmus University Rotterdam

**Master Thesis**

July 2021

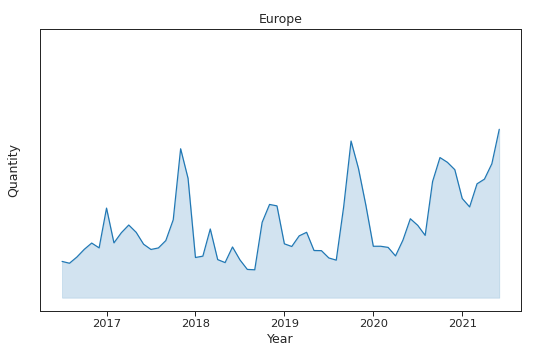
# Analysis of the models

This chapter performs the analysis of all the models, starting with an overview, followed by the analysis on the performance of each model.

## Overview

The modelling to forecast the quantity of air fryers is executed in Python. As discussed in chapter 3, different methods, including single one and hybrid one, are tested in order to find out which model has the best performance. Before the modelling, the first step is to look at the data plot of historical data, which is depicted in Figure 4‑1:

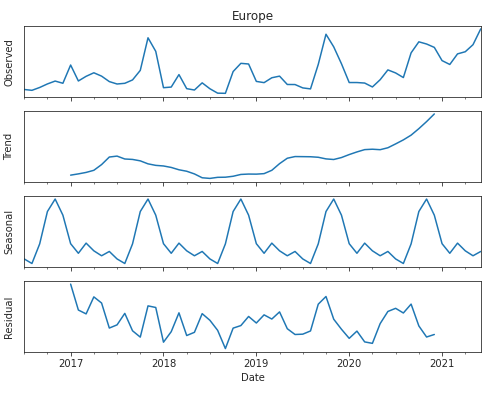
Figure 4‑1 Quantity of air fryers sold in EU7 through time



As can be seen in Figure 4‑1, the air fryer historical sales shows a strong seasonality and an increasing trend through years. Overall, throughout the given period, Philips air fryers sales data shows an upward trend interweaving with seasonal cycles that repeat within the year. The latest sales for 2021 are four times the average sales of 2016. The author notices that the growth in sales for the last three months of the given years seem to be the highest compared to the figures for other months of the year. The three outstanding peaks of Philips air fryer took place in October and November of years of 2017, 2019, and 2020.

Moreover, in order to disprove the dummy assumptions of patterns perceived, the study conducts to decompose data set to check the trend or seasonality phenomenon. The decomposition of data splits data into trend and seasonal patterns which are shown in Figure 4‑2 below:

Figure 4‑2 Decomposition of EU data



As observed in Figure 4‑2, the data set was decomposed in four parts: observed, trend, seasonal, and residual. It is apparent that air fryer sales witnessed an upward trend and strong seasonality.

Preliminary analysis is executed to inspect features including the trend pattern, and the seasonality component, then the statistical method is applied to test whether the data is non-stationary. From the observed values, we can seethat there is no clear hint of non-stationary within the time series as well as no clear trend. Thus non-seasonal differencing was not necessary. To confirm this the Augmented Dicky-Fuller (ADF) test was performed to test the null hypothesis of a unit root. With a ADF p-value less than 0.05 this resulted in not rejecting the null hypothesis of non-stationary. On the other hand, the trend values are currently being tested and their stationary characteristics are being verified by KPSS. Based upon the p-value of KPSS test smaller than 0.05, there is a clear evidence for rejecting the hypothesis of stationary series. In short, the data series is non-stationary.

Table 4‑1 Stationary Tests for Europe

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Results of Dickey-Fuller Test | | | | |
|  | Market | Test Statistic | p-value | Lags Used |
| 0 | KPSS | 0.683205 | 0.015072 | 3.0 |
| 1 | ADF | -2.140208 | 0.228657 | 2.0 |

Let’s plot the Partial Autocorrelation (PACF) and Autocorrelation (ACF) in the Figure 3‑4 to test whether is a sign of heteroscedasticity.

Figure 3‑4 Europe-Partial Autocorrelation and Autocorrelation

Chart

Description automatically generated

Chart, box and whisker chart

Description automatically generated

Looking at the plots of ACF and PACF, the orders indicate the possibility of seasonal components. The significant peak of both ACF and PACF plots is at lag 1, and have following peak at lag 12, then repeat the cycle of seasons. It suggests that both seasonal and non-seasonal autoregressive order could start with 1. Most of ACF and PACF residual values of seasonal orders are within significant interval at 95% confidence level. However, the existence of the ACF and PACF orders beyond the tolerant confidence limits suggests the presence of autocorrelation between residuals.

## Performance of Holt-Winter model

The seasonal and trend components of the data are not fixed in absolute numbers but are relatively constant variations in percentage terms around the underlying level of the data. The trend components will be owing to variables such as monthly production being impacted by population demand or market expansion, while the seasonal components will be due to the impacts of the season on sales. Thus, the study suggests that seasonal patterns and trend components are seen to interact in multiplicative manners to generate the finest predicted value in Holt-Winter model. The level constant is determined using a weighted average of the current seasonal component estimation and the preceding seasonal component estimate over a 12-period period. In the instance of Winter Exponential Smoothing, the random components affect the seasonal influence.

The level index shows the discrepancy from current data to future changes while the trend index offers an assessment of the movement shifting from time to time. The seasonality index evaluates the deviation in seasonality from a local mean. The coefficients of level terms, trend terms, and seasonal terms are 0.005; 0.005; and 0.639 respectively. Obviously, the smoothing parameters of level and trend patterns are relatively close to zero. In contrast, the coefficient of seasonal terms of over 0.6 indicates the seasonal component gives much more weight to the time-series data as opposed to the level and trend smoothing parameters. The Holt-Winter outcome emphasizes my expectation that the data of Philips air fryer is strongly dominated by the seasonality of orders.

Table 4‑2: Holt-Winter summary output

Exponential Smoothing Model Results

================================================================================

Dep. Variable: Quantity No. Observations: 48

Model: Exponential Smoothing SSE 22851904836.148

Optimized: True AIC 991.093

Trend: Multiplicative BIC 1021.032

Seasonal: Multiplicative AICC 1014.679

Seasonal Periods: 12 Date: Sat, 17 Jul 2021

Box-Cox: False Time: 12:33:05

Box-Cox Coeff.: None

=================================================================================

coeff code optimized

---------------------------------------------------------------------------------

smoothing\_level 0.0050000 alpha True

smoothing\_trend 0.0050000 beta True

smoothing\_seasonal 0.6396429 gamma True

initial\_level 31418.500 l.0 True

initial\_trend 1.0128930 b.0 True

initial\_seasons.0 0.8953005 s.0 True

initial\_seasons.1 0.8501679 s.1 True

initial\_seasons.2 1.0031351 s.2 True

initial\_seasons.3 1.1936598 s.3 True

initial\_seasons.4 1.3481229 s.4 True

initial\_seasons.5 1.2293076 s.5 True

initial\_seasons.6 2.2104174 s.6 True

initial\_seasons.7 1.3526107 s.7 True

initial\_seasons.8 1.6027818 s.8 True

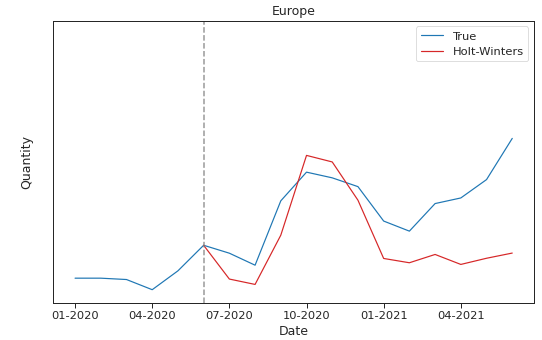
initial\_seasons.9 1.7931473 s.9 True

initial\_seasons.10 1.6128077 s.10 True

initial\_seasons.11 1.3195729 s.11 True

After calculating the smoothing constants for level, trend, and seasonality to optimise the errors of demand forecasting of Philips, the forecasting line is shown in the Figure 4‑3. Given the application of exponential smoothing model considering with the components of trend, seasonality, and level, the Holt-Winter prediction model tends to move consistent with the real data, yet has a significant margin of error between forecasting and actual data.

Figure 4‑3 Holt-Winter forecasting model



## **Performance of SARIMA model**

The seasonality of the 12 months presented in the given time series, the SARIMA method decomposes the components of trend, seasonality and residuals. The application of a 12-period moving average attempts to separate the seasonal effects from the time series, and then calculates the seasonal index. Looking at the trend index, it can be seen that Philips’ sales have steadily increased towards the end of the year. It is clear that the addition of seasonal, and trending components allows us to better predict overall long-term sales trends. To separate random from a trend after deseasonalisation, trend equations often use regression techniques based on the characteristic decreasing series in question and the time index.

Since the data is not a stationary time-series data, there is no constancy over times. In order to avoid the heteroscedasticity errors obtained from the model, the dataset needs to be rendered to a stationary series before using for the modelling. The typical way of converting a non-stationary into a stationary series is to subtract the previous observation from the current observation. However, with the support of function SARIMAX, this process can be simply done by adjusting the non-seasonal parameter *d* from 0 to 1. When the non-seasonal differencing *d* equals to 1, the data set is automatically converted as the order-1 differencing time series. The non-stationary of the observed data is identified with the non-seasonal differencing d is 1. Besides, the seasonal differencing parameter D is also equal to 1 in order to ensure the seasonal stationary of time-series data. Hence, SARIMA model (p, 1, q) x (P, 1, Q)12 is proposed. Drawing from the Table 4‑3, the SARIMA(1, 1, 3) x (0, 1, 1)12 performs best with the lowest AIC value of 802.77. With the order of p of 1, the present time-series data  relies on the prior data . With the order of P of 0, the current year is likely to not be reliant on its earlier year’s data. With the order of q of 3 and the order of Q of 1, the current time-series is affected by the forgoing random shocks.

Table 4‑3 SARIMA result tables

Order Seasonal Order AIC In-sample RMSE In-sample MAPE

0 (0, 1, 0) (0, 1, 0, 12) 815.219430 17364.375748 0.383769

1 (0, 1, 0) (0, 1, 1, 12) 809.325732 15219.786981 0.328222

2 (0, 1, 0) (0, 1, 2, 12) 811.294121 15177.428581 0.326979

3 (0, 1, 0) (1, 1, 0, 12) 810.738141 15431.604717 0.331464

4 (0, 1, 0) (1, 1, 1, 12) 811.296084 15179.091529 0.326924

.. ... ... ... ... ...

139 (3, 1, 3) (1, 1, 1, 12) 817.651955 16238.168615 0.379239

140 (3, 1, 3) (1, 1, 2, 12) 812.006037 16041.061514 0.388898

141 (3, 1, 3) (2, 1, 0, 12) 811.205626 16340.665544 0.393510

142 (3, 1, 3) (2, 1, 1, 12) 816.059033 16255.736499 0.376433

143 (3, 1, 3) (2, 1, 2, 12) 813.991338 16071.924977 0.389548

[144 rows x 5 columns]

Best Model: (1, 1, 3) (0, 1, 1, 12)], AIC=802.7684524606885

The summary of SARIMA output is shown in Table 4‑4. Based on p-value of parameters smaller than 0.05, all coefficients of SARIMA model are statistically significant. The p-value of Ljung-Box of 0.64 shows that there is no autocorrelation in the residuals, so they are independently distributed over time.

Table 4‑4 SARIMA summary output

SARIMAX Results

============================================================================================

Dep. Variable: Quantity No. Observations: 48

Model: SARIMAX(1, 1, 3)x(0, 1, [1], 12) Log Likelihood -393.175

Date: Sun, 18 Jul 2021 AIC 798.349

Time: 03:03:48 BIC 807.681

Sample: 0 HQIC 801.571

- 48

Covariance Type: opg

==================================================================================

coef std err z P>|z| [0.025 0.975]

------------------------------------------------------------------------------

ar.L1 -0.9109 0.172 -5.285 0.000 -1.249 -0.573

ma.L1 0.9597 0.187 5.128 0.000 0.593 1.327

ma.L2 -0.6286 0.288 -2.183 0.029 -1.193 -0.064

ma.L3 -0.7205 0.170 -4.248 0.000 -1.053 -0.388

ma.S.L12 -0.7264 0.437 -1.663 0.026 -1.582 0.130

sigma2 3.726e+08 9.03e-10 4.13e+17 0.000 3.73e+08 3.73e+08

===================================================================================

Ljung-Box (L1) (Q): 0.22 Jarque-Bera (JB): 2.06

Prob(Q): 0.64 Prob(JB): 0.36

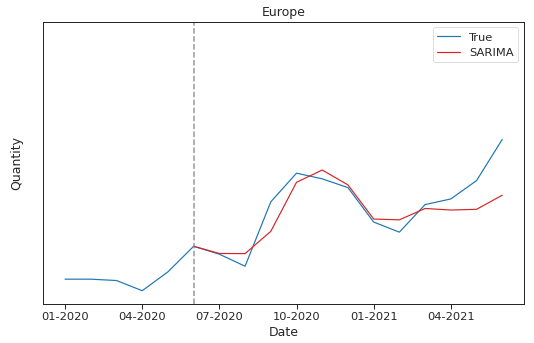
Heteroskedasticity (H): 0.56 Skew: 0.36

Prob(H) (two-sided): 0.33 Kurtosis: 3.95

===================================================================================

After estimating the non-seasonal parameters and seasonal auto-aggressive parameters in SARIMA model to optimise the errors of demand forecasting of Philips, the forecasting line is shown in the Figure 4‑4. The SARIMA forecasting model tends to move in accordance with the movement of real data with smaller error margin as opposed to Holt-Winter model. The SARIMA forecasting methods is likely to give the most consistent performance among all methods.

Figure 4‑4 SARIMA forecasting model



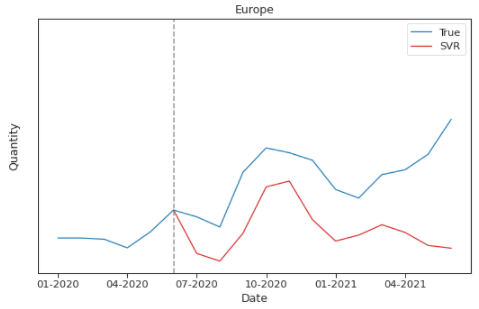
## Performance of SVR model

The SVR is the linear model that can categorize data not statistically separated, and classify them into linear problem. SVR is a predictor of neural network based on the idea of learning algorithms and the structural risk reduction principle. SVR may simplify complicated non-linear regression issues in a high-dimensional function into linear regression issues.

Given the volatile series, the radial basis function kernel (RBF) or Gaussian Kernel is adopted rather than the linear default kernel function in SVR method. The study is performed using the cross-validation technique and RBF classification, with optimum parameter settings being achieved for SVR when the distance among points in the origin space is lowest. There is no standard measurement to identify the free parameters C, the gamma kernel function and the loss function for the SVR model. After running a cross validation technique, the value of free parameters C is varied from the range 0.001 to 100, while, and the gamma kernel function run from the range 0.0001 to 10. The search procedure with optimal parameters (C=1; gamma = 1) is performed best to cover the spread of data among generations. The accuracy of model is greater with the high value of C parameters and gamma parameter. The entire process of performance has been documented.

After estimating the machine learning function and the free parameter in SVR model to optimise the errors of demand forecasting of Philips, the forecasting line is shown in the Figure 4‑5. The SVR forecasting model mostly move consistent with the movement of real data, yet with larger error margin as opposed to Holt-Winter model. The orders following April 2021 of SVR forecasting model tend to go opposite to turning point with the actual direction of real data. Given the standard SVR algorithm adopts fixed values of regularization constant (Cao & Gu, 2002), SVR method is not likely to perform well when the non-stationary series has heterogeneous noises.

Figure 4‑5: SVR forecasting model

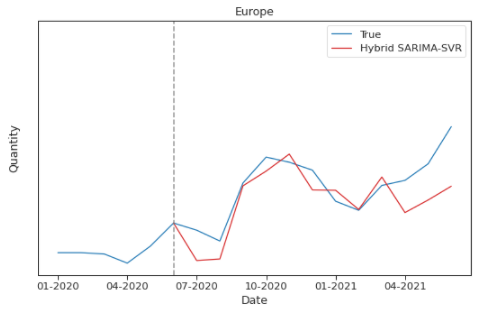


## Performance of hybrid SARIMA-SVR model

The hybrid model used in the study integrates the seasonal auto-regressive integrated moving average method (SARIMA) and the support vector regression method (SVR). In order to estimate and analyse linear component of the air fryer sales, two steps are applied to the methodology used for determine the hybrid model: firstly a SARIMA model is developed, and then an SVM model is developed to determine non-linear patterns in a series of residual generated with the SARIMA model. In the hybrid model, SARIMA model is first applied to filter out the residual components contained in forecasting variables. The filtered forecasting variables are then used random forest approach in SVR for constructing a forecasting model based on permutation performance.

After estimating parameters in the hybrid SARIMA-SVR model to optimise the errors of demand forecasting of Philips, the forecasting line is shown in the Figure 4‑6. The SARIMA-SVR forecasting model tends to move in accordance with the movement of real data with smaller error margin. The performance of hybrid SARIMA-SVR model outperforms the benchmark model as Holt-Winter model.

Figure 4‑6 Hybrid SARIMA-SVR forecasting model



## The comparative performance of forecasting models

Figure 4‑7 shows the comparison of actual sales data, Philips LTSP forecast values, and prediction results of time-series models and machine learning models. We find that the forecast results of LTSP are much lower than the actual data. The first part of LTSP's forecast is quite close to the actual data, but the tail is calculated from the end of 2020, the data of LTSP is a bit flat and the tail is quite far from the real data. At the tail end of the time model, from February 2021, the actual sales data of Philips air fryers marks a sharp growth spurt after a period of sales decline. The SARIMA model produces the forecast results that seem to match the real data best, when the forecast data starts to bounce back from February 2021, but does not record a strong growth like the actual demand patterns. The Holt-Winter forecast model and the Philips LTSP model also predict the same direction, but the predicted data is too low compared to the real data. The prediction data of the Hybrid SARIMA-SVR model is not too far from the actual data, but the model often produces noisy forecast results at some moments when the prediction results deviate from the actual trend. The weird thing is that the SVR predictive model at the tail end is going down while the real line is going up. In Python code, after overfitting data with Random Forest, SVR prediction model is likely to be very resistant, and generates the poor performance. For SVR, an important attribute for successful training of the SVR model is the appearance of a constant relationship between the independent and dependent variables (Cao & Gu, 2002), however the Philips sales non-stationary time-series model dampens the accuracy of the prediction model of SVR.

Figure 4‑7: The comparison of actual data and prediction results

The study assesses the forecasting performance in the sample, which the author utilizes to provide suggestions on how to use the sample prediction of our approach. The performance of the aforementioned time-series forecasting models, as well as the neural network machine learning algorithm methods, are measured based on the accuracy measurements. Among performance indicators, Root Mean Squared Error (RMSE) and Mean Absolute Percentage Errors (MAPE) are good approaches to assess the forecast accuracy. MAPE tells us how the average discrepancy between the predicted value and the actual. However, the measure could be misleading due to the abnormal value, hence RMSE provides the error measure from the perspective of standard deviation of data to give the accuracy of prediction. There is no standard of good MAPE and RMSE error measures to evaluate good forecast model. The forecast accuracy could be determined by comparing the performance of benchmark model, singular models, and hybrid one. In the study, Holt-Winters serves as a comparison benchmark towards SARIMA, SVR, and hybrid models.

Table 4‑4 shows the performance of the models. Clearly, it is demonstrated that hybrid SARIMA-SVR model has also significant low RMSE value considering Europe market. Taking a closer look at the Table 4‑4, the author sees the same pattern for MAPE value of hybrid SARIMA-SVR method. Surprisingly, in contrast to the literature review that neural networks and machine learning method have proven to have better accuracy, SVR model has the poorest forecasting performance among methods. That could be probably given the non-stationary series has noises and the standard SVR algorithm adopts fixed values of regularization constant. Thus, SVR performs badly in predict Philip air fryer sales. The integration of time-series analysis into SVR machine learning method could improve the performance of several aspects of non-stationary dataset in comparison with the accuracy of SVR model. Thus, it is not surprising that the hybrid model to lead to a better outcome. By contrast, the SARIMA method indeed deliver the best performance, based on both RMSE and MAPE. This is understandable because of the strong indication of the seasonal component yet non-stationary phenomenon in the series. Based on the concept of non-seasonal and seasonal auto aggressive moving average, SARIMA model is the best forecasting model among all methods to predict the long-term sale of Philips air fryers. With SARIMA model, it is expected that if the sale data of Philips air fryer has been performed with strong seasonal patterns in the long-term, the improvement with SARIMA could have been smoothed and achieved the closest prediction.

Table 4‑5 Performance of forecasting models according to RMSE and MAPE

|  |  |  |
| --- | --- | --- |
| Model | MAPE | RMSE |
| LTSP | 0.415 | 43519.4 |
| Holt-Winters | 0.300 | 33322.5 |
| SARIMA | 0.102 | 13804.7 |
| SVR | 0.482 | 47990.1 |
| Hybrid SARIMA-SVR | 0.165 | 18554.9 |

According to SARIMA model's MAPE index, the forecast model is only about 10% away from reality and is considered a good predictive model. With the best performance of SARIMA, we have the data forecast table of the SARIMA model, the LTSP forecasting of Philips, and the actual model over time in Table 4‑6 to make comparison between true and predicted values.

Table 4‑6: The predicted data of SARIMA model

|  |  |  |  |
| --- | --- | --- | --- |
|  | Actual data | LTSP | SARIMA |
| 07-20 | 56,065 | 86752.00 | 56531.99 |
| 08-20 | 48,284 | 68701.00 | 56448.00 |
| 09-20 | 90,023 | 96162.00 | 70900.73 |
| 10-20 | 108,622 | 109993.00 | 102706.36 |
| 11-20 | 104,936 | 88983.00 | 110643.48 |
| 12-20 | 99,233 | 74613.00 | 101006.98 |
| 01-21 | 76,908 | 38344.19 | 78877.30 |
| 02-21 | 70,371 | 34470.35 | 78345.43 |
| 03-21 | 88,255 | 36422.14 | 85711.06 |
| 04-21 | 91,890 | 28683.68 | 84707.87 |
| 05-21 | 103,798 | 42814.45 | 85163.21 |
| 06-21 | 130,519 | 45126.71 | 94313.71 |

# Conclusion and Discussion

## Conclusion

Forecasting is a very complex process relating to multiple variables and scenarios. Forecast also has a strong impact in an organization operation. Good forecasting helps improve supply chain performance, facilitates the evaluation of supply chain management's effectiveness, better reflects customer needs, reduces reaction times, and supports risk assessment. To improve the forecast performance at Philips, the first stage is to clearly define the stages of sale process. In the case of Philips air fryer, the author focuses on the long-term sale plan with high expectation based on the historical sales. After defining the forecasting target, the author carries out to clean up data, and elevate data with additional training before modelling. The next stage is to choose the suitable forecasting method. Within the scope of study, this thesis compares the performance of time-series statistical, machine learning and hybrid methods to forecast the volume of air fryer in order to better support the company in business planning. According to the accuracy results of all the methods, it can be concluded that Seasonal Autoregressive Integrated Moving Average (SARIMA) model acquires the most accuracy outcome for quantity prediction.

The researcher estimates the predicted data based on the monthly historical sale data ranged from 2016 to 2019. Looking at the summary of historical sale, the author notices that the time-series data has clear seasonality over time. After running preliminary analysis, it is found that there is no clear hint of non-stationary within the time series as well as no clear trend. Testing model with ACF and PACF plots, the existence of seasonal orders beyond the tolerant confidence limits suggests the presence of autocorrelation between residuals. Considering the Holt-Winter model, the outcome emphasizes the expectation that the data of Philips air fryer is strongly dominated by the seasonality of orders. Given the lack of controlling seasonal patterns, the prediction of Holt-Winter model has a significant margin of error between forecasting and actual data. Fixing the issue by adding a seasonal term, the SARIMA forecasting model tends to move closest in accordance with the movement of real data. Considering the machine learning algorithm, the prediction of SVR method offers the larger error margin between forecasting and actual data. SVR method is likely to be the poorest forecasting performance among forecasting models due to the heterogeneous noise. That is also probably given the data is not linearly separable. In order to minimize the noises in data set, the study conduct the hybrid approach combined SARIMA model and SVR technique. The hybrid approach generates better prediction as compared to SVR method, yet not as good as SARIMA model.

In conclusion, based on the analysis, the author can conclude that even though the machine learning methods are proved for their high accuracy, but that is not applicable for every cases. In this case study, the research question “*Which forecasting model is being used in LTSP for air fryer at Philips KA and how accurate is it?”* is clearly answered with four main forecasting models, namely Holt-Winter exponential smoothing model, Seasonal Autoregressive Integrated Moving Average model (SARIMA), and machine learning algorithm as support vector regression (SVR) compared to the hybrid approach combining SARIMA and SVR model.

The second research question *“To what extent can a single approach or a combination of forecasting methods outperform the judgmental forecast in producing more accurate forecast for air fryers product at Philips Kitchen Appliances?”* can be answered that the time-series SARIMA forecasting model demonstrates the most promising result comparing to simple exponential smoothing Holt-Winters or machine learning method such as SVR, and even outperforms the hybrid method SARIMA-SVR.

The third research question “*How much is the improvement in forecasting performance that the proposed solution generates?”* could be suggested. In comparison to the benchmark model as Holt-Winter exponential smoothing model, the prediction of SARIMA model performs clearly better with RMSE three time lower than RMSE of Holt-Winter. Therefore, in order to improve forecasting performance, the non-seasonal and seasonal component in Philips’s sale data should be considered to add into the time-series model. It explains better the data pattern and offers better prediction with the same patterns in the continuous period. Given the model-based forecast is only based on historical data, the prediction is also influenced by other exogenous factors. Therefore, the author proposes to apply the regression model to enhance the accuracy of demand forecasting models in the next study.

## Limitation and future research

This thesis aims at proposing improvement solution for forecasting air fryer products at Philips KA. Those proposal was built and tested base on the historical data of five year, from July 2016 to June 2020. However, the dataset is only available in monthly scale, there are only 48 data points for the training set, which is relatively small even though the data was collected in 5 years horizon. This is a trade-off between the relevant of the data source and the number of data points used in the training process, as too old data is hardly applicable to predict the current growth and trend of the product and too little data reduces the accuracy of the models.

Secondly, when choosing the scope of this research, the author only took into account the forecast of the CMMs, which is only the forecast from the Marketing department. The data limitation of cross-function, such as demand forecasting, may limit the application of the forecast. For future research, it is recommended that one consider the cross-function information to increase the applicability of the forecast. This also matches with Philips KA’s goal to use USP as the starting point for LTSP from October 2021.

Another improvement area is that machine learning models can learn and improve from itself. In this thesis, however, the solution is proposed but has not been applied to current practice at Philips yet. Therefore, the models have not had the chances to compare the output with the actual situation nor received feedback on how well the models work and how it can be improved. Deployment plan does not lay within the scope of this study but can be further investigated in future research. Fildes et al. (2009) have found that up to 80% statistical forecast are amended judgmentally. Therefore, combining statistical forecast methods with judgmental forecast can lead to a better result as it can make use of both the accuracy of the former method and the human knowledge of the latter method. This thesis, however, focuses on the computation method but the combining approach can be a good suggestion for future research.