Pharma Drug Sales – Analysis and Forecasting

Time Series and Forecasting 2024/2025

Group 18

Athos Freitas – up202108792

Luís Du – up202105385

Master in Artificial Intelligence

Professor: Maria Eduarda Silva

10th January, 2025

1. Introduction

This project aims to analyze and forecast pharmaceutical drug sales using time series data collected from the Point-of-Sale system of a single pharmacy over six years (2024–2019) [1]. The primary objective is to uncover underlying patterns in the data, such as trends and seasonality, and to develop models capable of generating accurate and reliable sales forecasts.

Trend and seasonality patterns of the time series will be explored, along with the evaluation of different models. Various forecasting models and strategies will be evaluated, with an emphasis on identifying the most effective and computationally efficient approaches. Accuracy measures will be applied to assess the quality of the results. Finally, the study will conclude with a discussion of the benefits and limitations of the chosen approaches.

2. Methodology

The analysis involves several key steps to ensure a comprehensive understanding and accurate forecasting of pharmaceutical drug sales.

Firstly, **Exploratory Data Analysis (EDA)** will be conducted to visually inspect the time series and detect patterns such as trends and seasonality. Variance will be analyzed to identify any instability, and different transformations will be applied as needed to stabilize it. Statistical tests, including the Augmented Dickey-Fuller (ADF) test, will be used to confirm stationarity. If the time series is not stationary, transformations or differencing techniques will be applied to stabilize its mean and variance, ensuring that it meets the assumptions for time series modeling.

The next step is **Data Partitioning**, where the dataset will be divided into training and test sets. This separation is crucial for model development and evaluation, enabling the assessment of the models' performance on unseen data.

During **Model Development**, ARIMA and SARIMA models [2] will be constructed to capture the underlying patterns in the data. The process will involve diagnostic checks such as autocorrelation and partial autocorrelation analysis, unit-root testing, and the Ljung-Box test to ensure the models' adequacy. Models will be shortlisted based on their Akaike Information Criterion (AIC) scores. Additionally, an alternative modeling approach using STL decomposition will be applied to each time series component – trend, seasonality, and residuals – to enhance forecasting accuracy.

In the **Forecasting and Evaluation** phase, the selected models will be employed to generate forecasts for the test set. Prediction accuracy will be assessed using cross-validation and comparisons with actual values.

Both rolling forecasts and long-term step-ahead forecasting strategies will be explored to evaluate the models' performance in different scenarios. To quantify the uncertainty associated with the forecasts, 95% prediction intervals will be calculated.

The **Quality Measures** phase will involve applying accuracy metrics to evaluate the reliability and precision of the forecasting results, ensuring a robust assessment of the models' effectiveness.

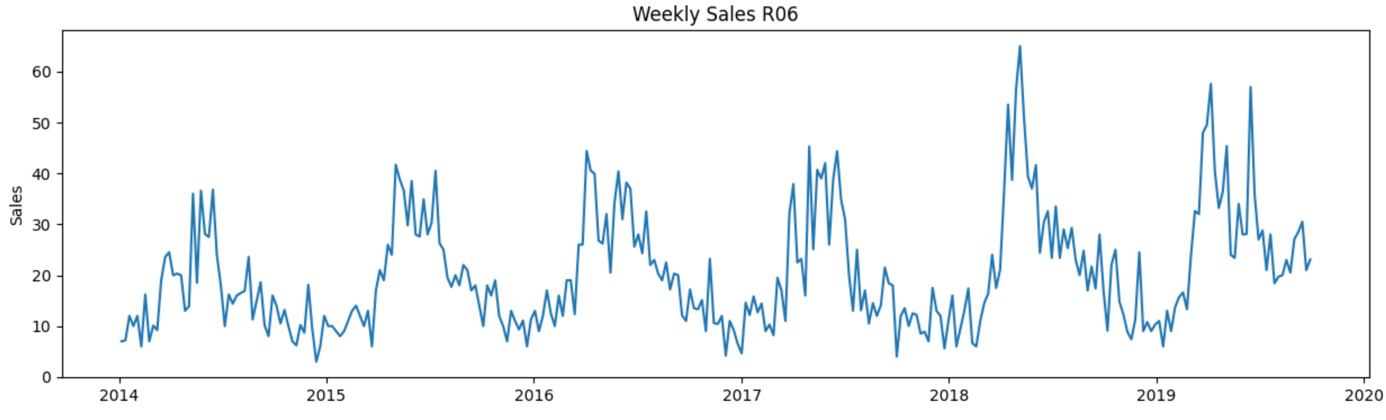
Finally, the study will conclude with a **Discussion** of the strengths and limitations of the ARIMA and SARIMA models in forecasting pharmaceutical sales. This section will provide insights into the models' suitability for this domain, highlight areas for improvement, and offer recommendations for future research.

3. Exploratory Data Analysis

3.1. Time series plot

The chosen dataset for this study is derived from a Point-of-Sale (POS) system of a single pharmacy, covering a period of six years. The research underlying this project considers eight distinct time series, each summarizing the sales of a specific group of pharmaceutical products. These series exhibit varying statistical features, offering a diverse foundation for analysis and forecasting.

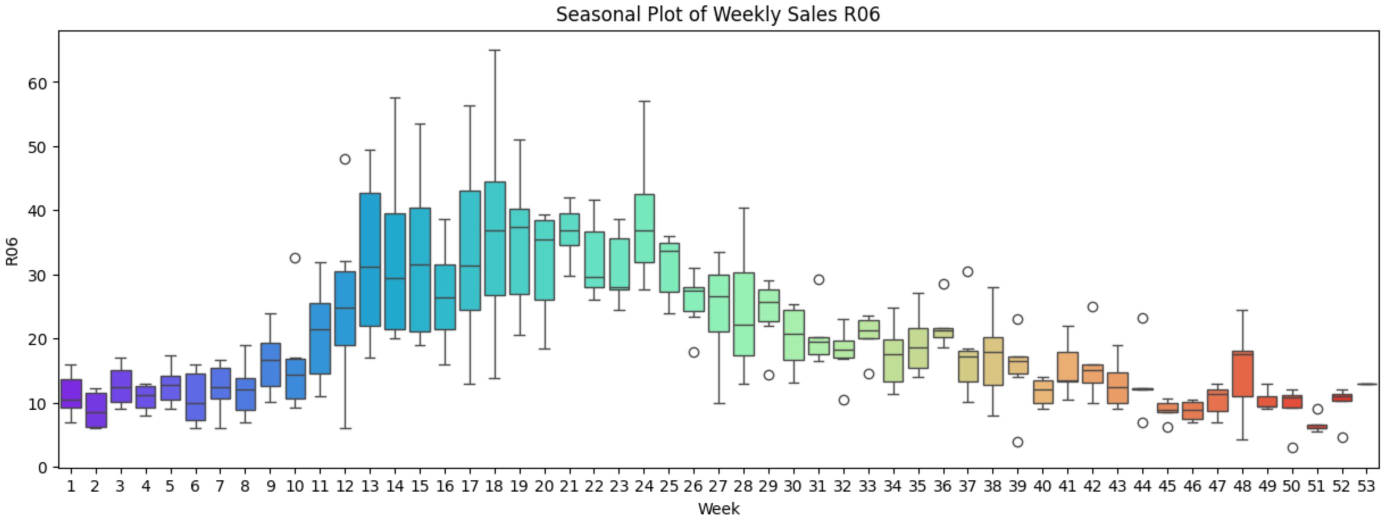
In this study, we will focus on the analysis of the **R06 category**, which includes **antihistamines for systemic use.** This category is particularly relevant due to its consistent demand and significance within the pharmaceutical sector, making it an ideal candidate for exploring time series patterns.



By visualizing the time series data, we observe several key characteristics. First, there is no clear evidence of a trend over time, indicating that the data does not exhibit a consistent upward or downward movement. Second, the data displays pronounced seasonality, with recurring patterns that suggest regular fluctuations at specific intervals. Lastly, the variance is relatively stable throughout the series, except for a noticeable spike during the early weeks of 2018.

3.2. Seasonal plot

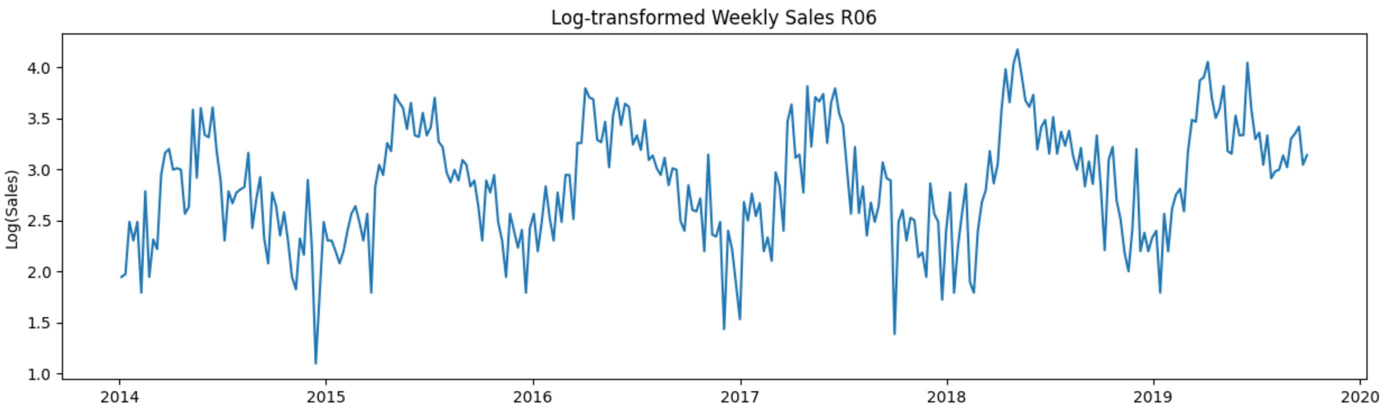
For a clearer representation of the seasonal patterns, we may use the seasonal plots.



The seasonal plot reveals that sales consistently peak during the spring, followed by a gradual decline throughout the remainder of the year. This pattern aligns with expectations, as it reflects the increased demand for antihistamines during allergy seasons, when symptoms are most prevalent.

3.3. Variance stabilization

To stabilize the variability over the series, the Box-Cox transformations can be applied. One specific case of this transformation is to take the **logarithm of the data**.

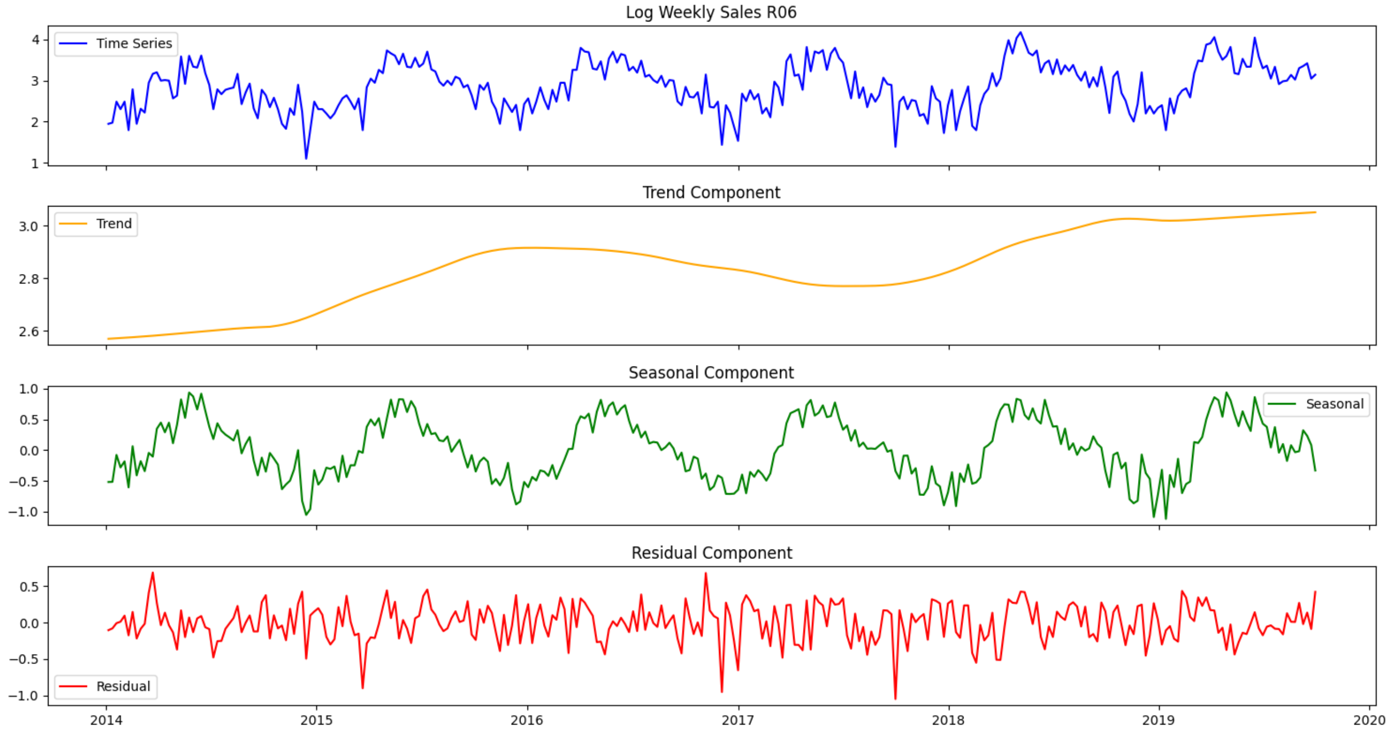


After applying a log transformation to the time series, the variance appears significantly reduced, with the previously noticeable spike in early 2018 no longer evident. This indicates that the transformation effectively stabilized the data, making it more suitable for modeling.

The best **Box-Cox transformation** was also explored by identifying the lambda value that minimizes variance (λ=0.10). Since this value is close to zero, the result of the Box-Cox transformation is very similar to the log transformation. To maintain simplicity and enhance interpretability, we will use the log-transformed time series in the subsequent sections.

3.4. Seasonal-Trend decomposition using Loess (STL)

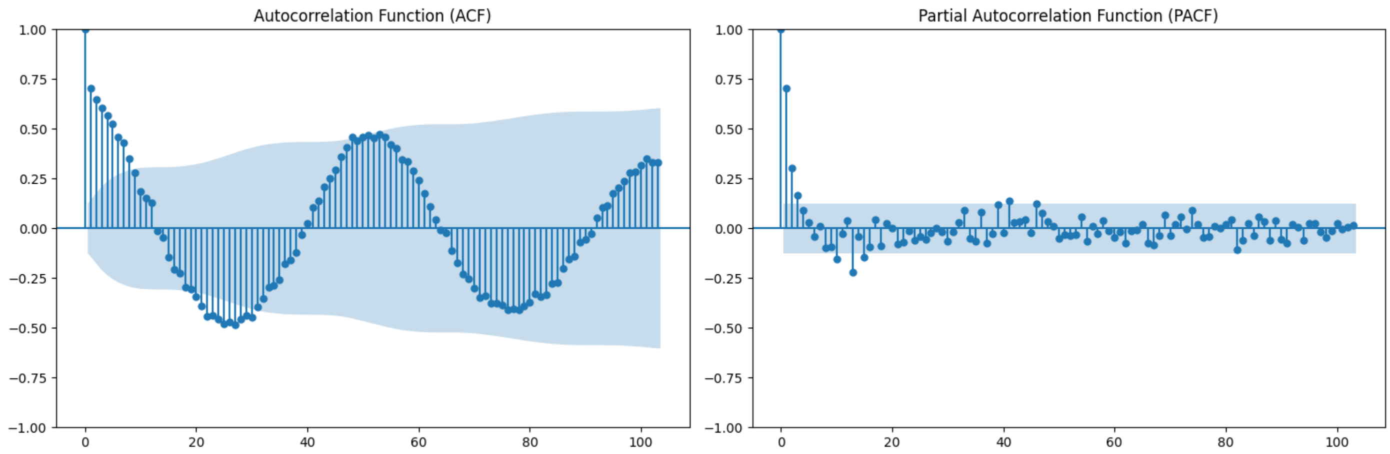
The Seasonal Decomposition of Time Series by Loess [3] decomposes a time series into seasonal, trend and irregular components using Loess.



From the decomposition, we can confirm some of the earlier observations about the time series. Firstly, there is a lack of an evident trend, as the trend component does not exhibit a clear directional movement over time. Secondly, the seasonal component reveals a recurring pattern that aligns with the previously observed seasonality, reinforcing the periodic nature of the data.

3.5. Autocorrelation and Partial Autocorrelation Functions

By examining the Autocorrelation and Partial Autocorrelation functions of the original time series, we observe several key patterns. The **ACF** clearly exhibits cyclic patterns, gradually approaching zero, which indicates seasonality and diminishing correlations at higher lags. On the other hand, the **PACF** shows distinct spikes at lags 1, 2, and 3, suggesting strong direct correlations at these lags and highlighting their potential relevance in modeling autoregressive components of the time series.



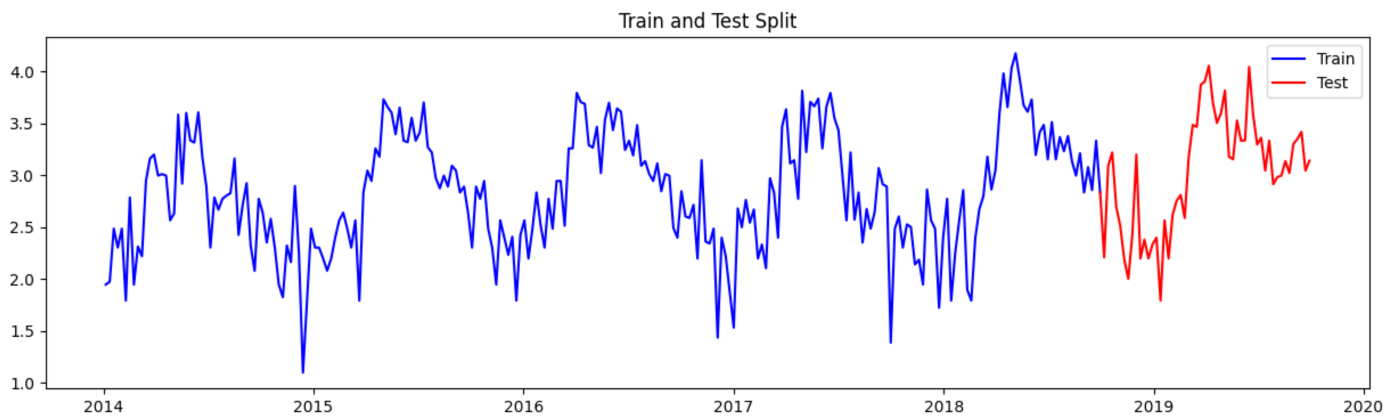
3.6. Stationarity tests

To assess the stationarity of the time series, both the **Augmented Dickey-Fuller (ADF)** and **Kwiatkowski-Phillips-Schmidt-Shin (KPSS)** tests were applied. The ADF test examines the null hypothesis that a unit root is present in the time series, indicating non-stationarity. In this case, the ADF statistic is lower than the critical values, and the p-value is less than 0.05, confirming that **the log-transformed time series is stationary.**

The KPSS test, on the other hand, evaluates the null hypothesis that the time series is stationary around a deterministic trend. The KPSS test indicates that 296 differences would be required to achieve a relatively stationary time series. However, applying such a high number of differences is impractical in real-world scenarios, suggesting that alternative approaches should be considered for addressing stationarity.

4. Data partitioning

The most recent 52 weeks of data will be designated as the testing set, providing a realistic forecast horizon for evaluating model performance and prediction accuracy.



5. Modelling

5.1. Autoregressive models

Based on the observed ACF and PACF during the data analysis phase, which suggest the presence of autoregressive patterns, we will explore **AR(2)** and **AR(3)** models to capture the temporal dependencies in the data.

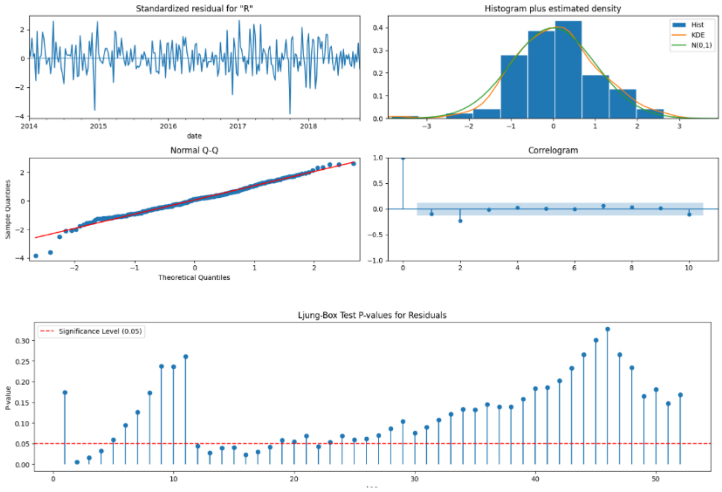
- **AR(2)**

SARIMAX Results

Dep. Variable:	R06	No. Observations:	248
Model:	SARIMAX(2, 0, 0)	Log Likelihood	-119.475
Date:	Thu, 09 Jan 2025	AIC	244.950
Time:	09:26:59	BIC	255.490
Sample:	01-05-2014	HQIC	249.193
	- 09-30-2018		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.5835	0.054	10.873	0.000	0.478	0.689
ar.L2	0.4093	0.053	7.713	0.000	0.305	0.513
sigma2	0.1509	0.011	13.443	0.000	0.129	0.173

Ljung-Box (L1) (Q):	2.28	Jarque-Bera (JB):	11.92
Prob(Q):	0.13	Prob(JB):	0.00
Heteroskedasticity (H):	1.17	Skew:	-0.19
Prob(H) (two-sided):	0.47	Kurtosis:	4.01



For the **AR(2)** model, all the **parameters are statistically significant**; however, there is still **some correlation between the residuals.** This suggests that the model does not fully capture all the underlying patterns in the data, indicating the need for a more complex model to better explain the observed behavior.

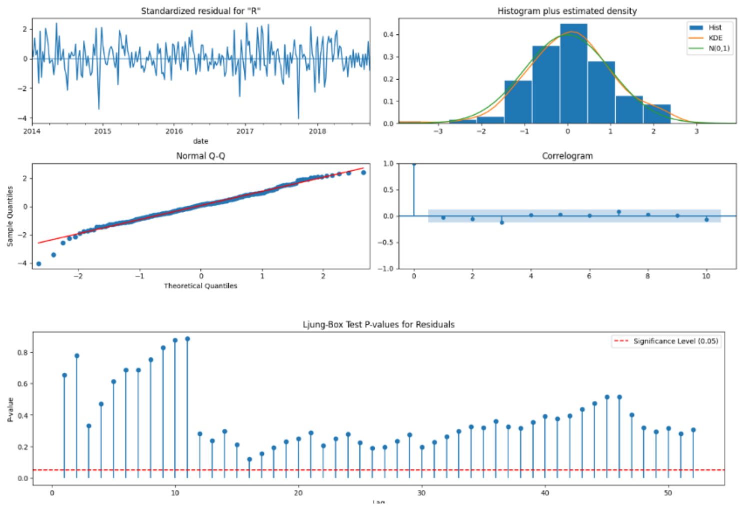
- **AR(3)**

SARIMAX Results

Dep. Variable:	R06	No. Observations:	248
Model:	SARIMAX(3, 0, 0)	Log Likelihood	-113.195
Date:	Thu, 09 Jan 2025	AIC	234.389
Time:	09:27:00	BIC	248.443
Sample:	01-05-2014	HQIC	240.047
	- 09-30-2018		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.4924	0.062	7.960	0.000	0.371	0.614
ar.L2	0.2785	0.064	4.331	0.000	0.152	0.405
ar.L3	0.2233	0.063	3.571	0.000	0.101	0.346
sigma2	0.1434	0.011	13.471	0.000	0.123	0.164

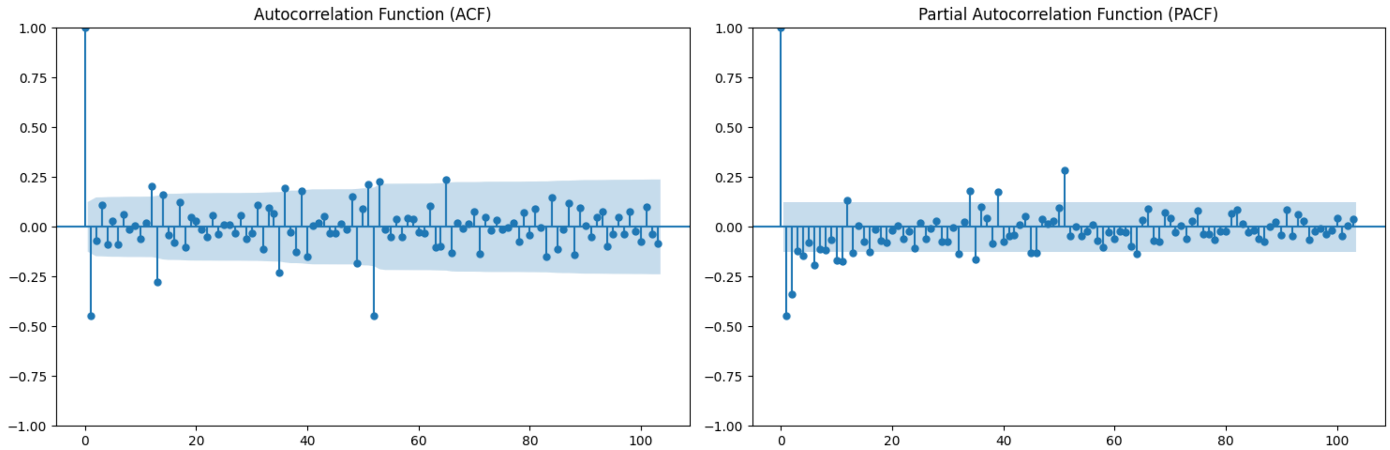
Ljung-Box (L1) (Q):	0.27	Jarque-Bera (JB):	14.82
Prob(Q):	0.60	Prob(JB):	0.00
Heteroskedasticity (H):	1.15	Skew:	-0.28
Prob(H) (two-sided):	0.54	Kurtosis:	4.06



Using the AR(3) model, we observe that all the **parameters remain statistically significant**, and the **residuals are now independent**. This indicates that the AR(3) model provides a better fit to the data compared to the AR(2) model, effectively capturing the underlying patterns and reducing residual autocorrelation.

5.2. SARIMA models

To address the seasonality, we have decided to apply a **52-step difference** to the time series, which corresponds to the seasonal cycle observed over the course of a year. The **ADF test** was then applied to assess the stationarity of the differenced series. Based on the results of the ADF test, we conclude that only **one difference (d=1)** is sufficient to achieve stationarity in the time series, ensuring that the data is appropriately transformed for modeling.



Looking at the ACF and PACF, we observe clear spikes at both **lag 1** and **lag 52** in the plots. This indicates strong correlations at these specific lags, which likely correspond to **immediate** and **seasonal dependencies** in the time series, suggesting that both short-term and yearly cyclical patterns are important factors in modeling the data.

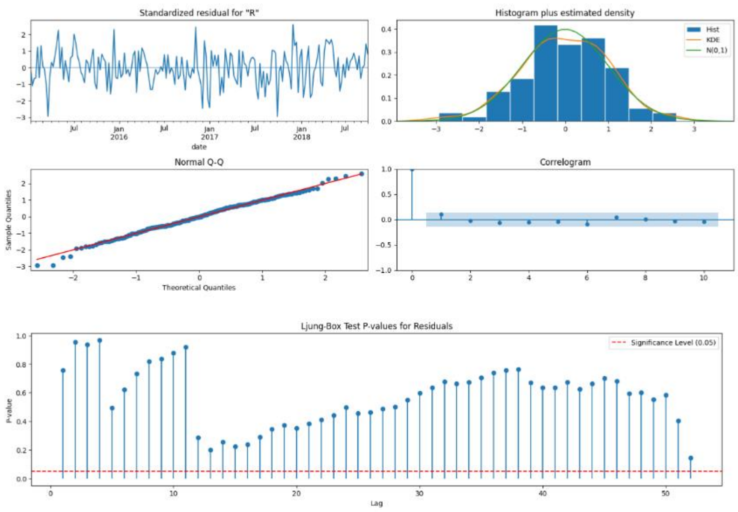
- **SARIMA(0,1,1)(1,1,1)$_{52}$**

SARIMAX Results

Dep. Variable:	R06	No. Observations:	248
Model:	SARIMAX(0, 1, 1)x(1, 1, 1, 52)	Log Likelihood	-105.063
Date:	Thu, 09 Jan 2025	AIC	218.127
Time:	09:27:36	BIC	231.219
Sample:	01-05-2014	HQIC	223.427
	- 09-30-2018		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ma.L1	-0.9254	0.032	-29.141	0.000	-0.988	-0.863
ar.S.L52	-0.3076	0.164	-1.873	0.061	-0.629	0.014
ma.S.L52	-0.4197	0.195	-2.150	0.032	-0.802	-0.037
sigma2	0.1476	0.016	9.017	0.000	0.116	0.180

Ljung-Box (L1) (Q):	2.32	Jarque-Bera (JB):	0.97
Prob(Q):	0.13	Prob(JB):	0.61
Heteroskedasticity (H):	1.32	Skew:	-0.17
Prob(H) (two-sided):	0.26	Kurtosis:	3.09



In this SARIMA model, **all the parameters are statistically significant, and the residuals are uncorrelated**, indicating a good fit. Let's now check the model with p=1 to see if it further improves the performance.

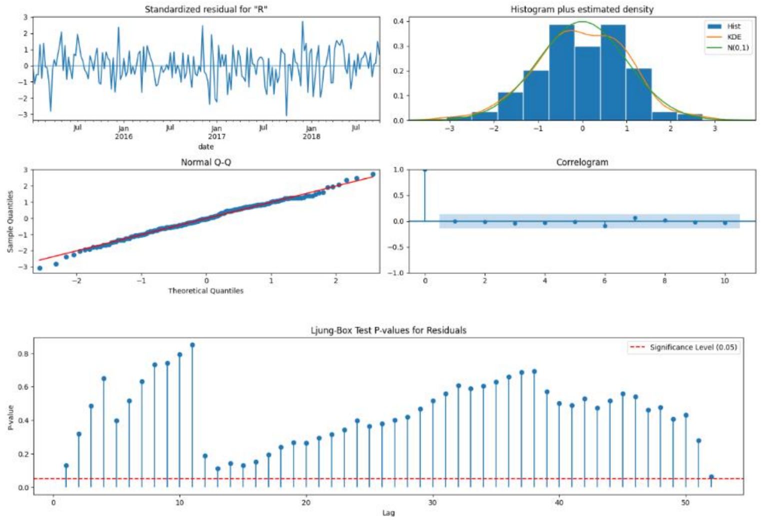
- **SARIMA(1,1,1)(1,1,1)$_{52}$**

SARIMAX Results

Dep. Variable:	R06	No. Observations:	248
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 52)	Log Likelihood	-103.633
Date:	Thu, 09 Jan 2025	AIC	217.266
Time:	09:28:02	BIC	233.631
Sample:	01-05-2014	HQIC	223.892
	- 09-30-2018		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.1293	0.080	1.614	0.106	-0.028	0.286
ma.L1	-0.9395	0.033	-28.153	0.000	-1.005	-0.874
ar.S.L52	-0.3455	0.164	-2.101	0.036	-0.668	-0.023
ma.S.L52	-0.3694	0.202	-1.830	0.067	-0.765	0.026
sigma2	0.1463	0.016	9.196	0.000	0.115	0.178

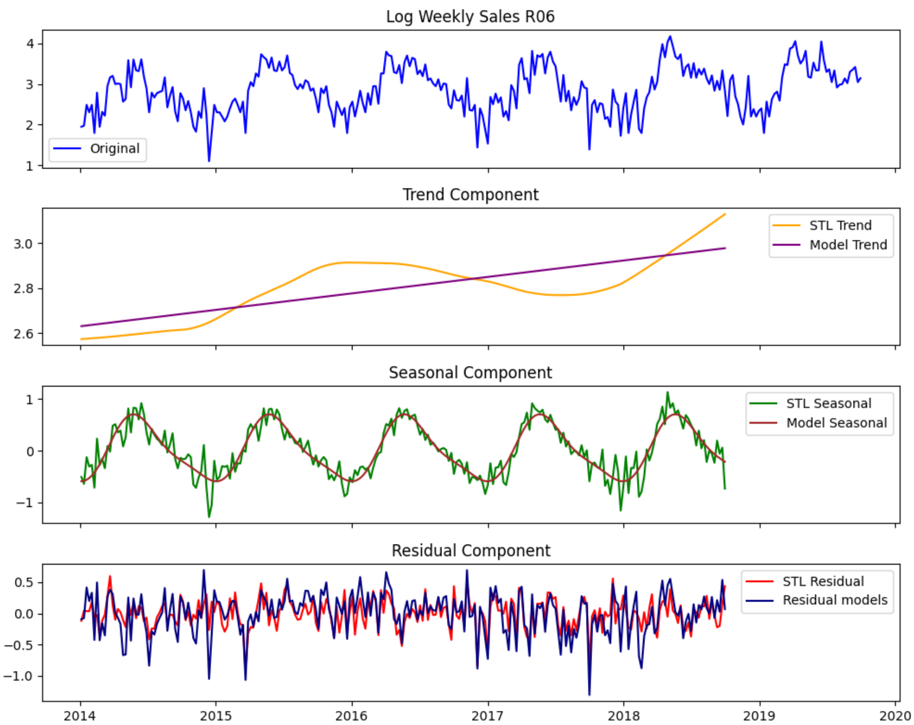
Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	1.02
Prob(Q):	0.98	Prob(JB):	0.60
Heteroskedasticity (H):	1.30	Skew:	-0.17
Prob(H) (two-sided):	0.29	Kurtosis:	3.12



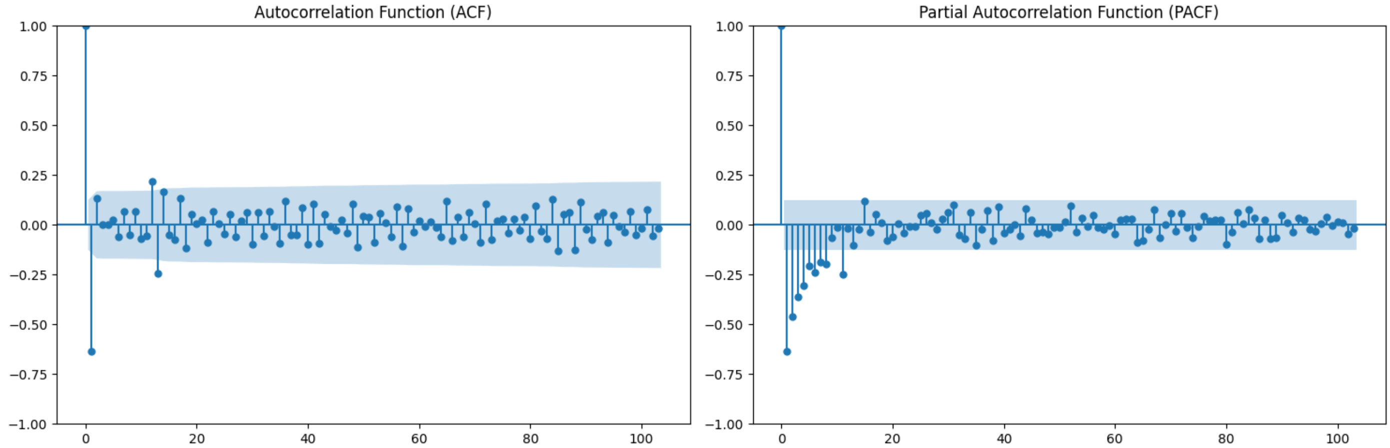
In this model, both requirements are also fulfilled, with **all parameters being statistically significant and the residuals uncorrelated**, making it also a strong candidate for forecasting.

5.3. Alternative modelling using STL decomposition

STL decomposition breaks the time series into three components: **trend, seasonality**, and **residuals**. The trend is approximately linear, making it suitable for a **linear model**. The seasonal component can be represented as a **Fourier series** [4], simplifying its cyclical behaviour. After removing the trend and seasonality, the residuals can be predicted using an **ARIMA model**, as the seasonality is already accounted for, eliminating the need for SARIMA. This approach offers a clear and efficient method for forecasting.



After applying the difference operator twice (d=2), the residual portion of the training time series was tested for stationarity using the ADF test. The ADF statistic value was -8.45, which is well below the critical value at the 1% significance level (-3.46). This indicates that the residuals "pass" the ADF test, demonstrating signs of **stationarity**.



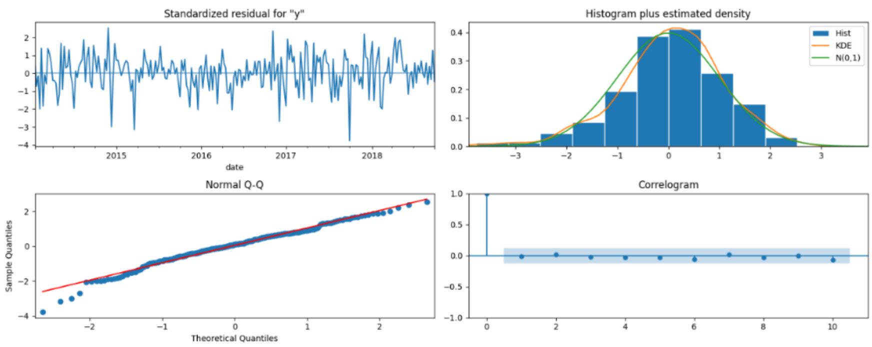
The ACF exhibits a sharp drop towards zero from lag 1 to lag 2, suggesting the presence of a **moving average (MA) component**. Meanwhile, the PACF gradually converges to zero over the first 10 lags, with lag 1 being significantly higher than subsequent lags, indicating a possible autoregressive (AR) component. Together, these patterns point to an ARIMA(1,2,1) model as a suitable candidate. However, after testing this model, some residual correlation was observed at lag 2, prompting a refinement to an ARIMA(1,2,2) model to better capture the underlying dynamics.

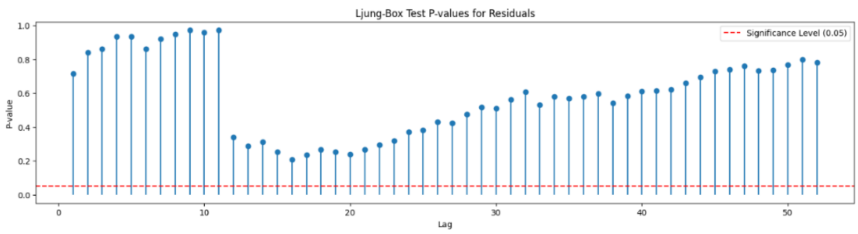
SARIMAX Results

Dep. Variable:	y	No. Observations:	248
Model:	SARIMAX(1, 2, 2)	Log Likelihood	-76.820
Date:	Thu, 09 Jan 2025	AIC	161.640
Time:	09:28:04	BIC	175.661
Sample:	01-05-2014	HQIC	167.286
	- 09-30-2018		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.1428	0.068	2.096	0.036	0.009	0.276
ma.L1	-1.9790	0.027	-73.272	0.000	-2.032	-1.926
ma.L2	0.9839	0.027	36.271	0.000	0.931	1.037
sigma2	0.1043	0.008	12.616	0.000	0.088	0.121

Ljung-Box (L1) (Q):	0.02	Jarque-Bera (JB):	18.73
Prob(Q):	0.89	Prob(JB):	0.00
Heteroskedasticity (H):	1.11	Skew:	-0.51
Prob(H) (two-sided):	0.64	Kurtosis:	3.89





The ARIMA(1,2,2) model demonstrates **strong statistical performance** based on the model summary and diagnostic tests. The **AR and MA parameters are statistically significant**, confirming their relevance in capturing the underlying structure of the time series. The model's AIC and BIC values are 161.64 and 175.66, respectively, indicating a good balance between model complexity and fit. The Ljung-Box test further validates the model, as none of the lags exhibit significant residual correlation. This suggests that the ARIMA(1,2,2) model effectively captures the time series structure and adequately accounts for the patterns in the data.

5.4. Choosing the best model

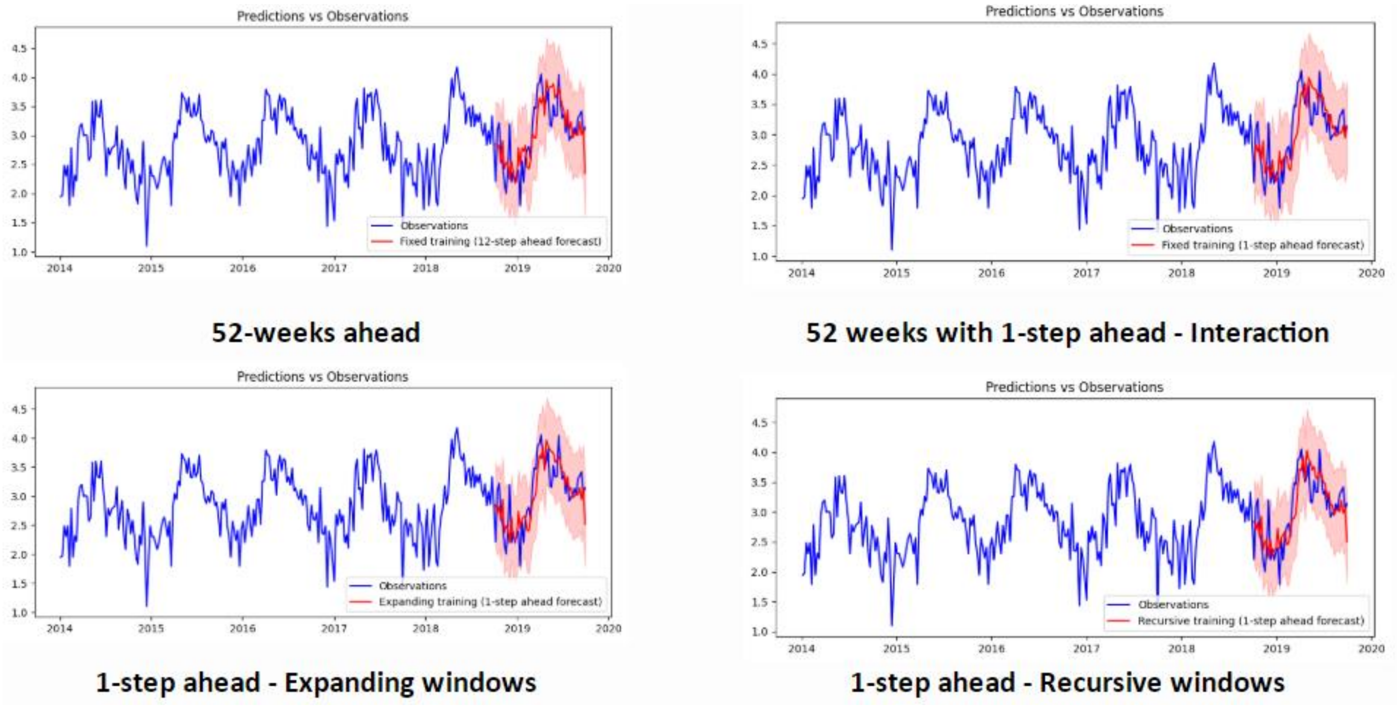
In selecting the best models for forecasting, two different approaches were considered. For the traditional modeling method, various models were evaluated, and the **SARIMA(1,1,1)(1,1,1)$_{52}$ model** was chosen due to its lowest AIC [5] value, indicating the best balance between model fit and complexity. In the alternative decomposition approach, the **ARIMA(1,2,2) model** was selected for forecasting the residuals, as it effectively captures the remaining structure after accounting for the trend and seasonal components.

6. Forecast strategies

6.1. SARIMA

To **forecast 52 weeks ahead**, a single model is fitted to the training dataset and used to generate predictions. Several strategies can be employed for **1-step ahead forecasting**:

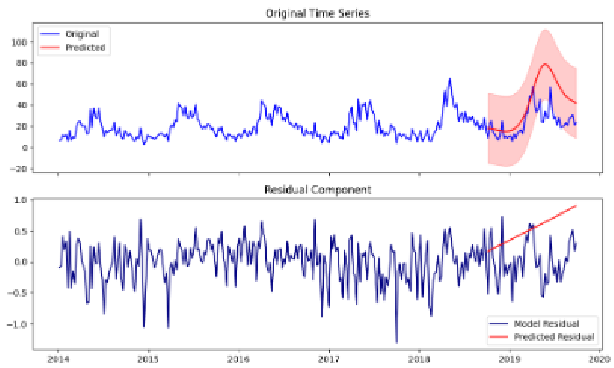
- In the **1-step ahead - Interaction strategy**, new data points become available incrementally, and the initially fitted model interacts with the growing dataset to iteratively generate forecasts, reflecting the evolving time series without refitting.
- The **1-step ahead - Expanding windows** strategy increases the training dataset by one unit at each step. The model is refitted with the additional data, generating forecasts based on progressively larger datasets that adapt to new information.
- Finally, the **1-step ahead - Recursive windows** strategy uses a fixed training sample size, sliding the window forward by one unit with each new data point. This method generates forecasts from models fitted to the most recent observations, preserving a rolling window of data.



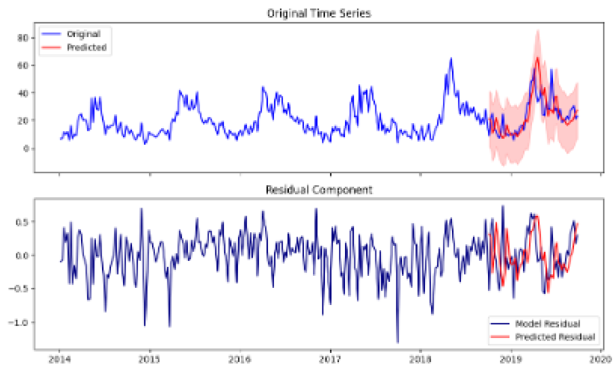
6.2. Decomposition forecast

In this section, we explore forecasting strategies that are **applied specifically to the residuals**, following the decomposition method where separate models are used for the trend, seasonality, and residual components. The final predictions are generated by combining the forecasts from the different models applied to each individual component. For all strategies discussed, the trend and seasonality components were fitted and predicted only once, based on the training dataset, ensuring that these components remain consistent across all forecasting approaches. The focus is on how well the residuals can be modelled and integrated with the predictions from the trend and seasonal models to provide accurate forecasts.

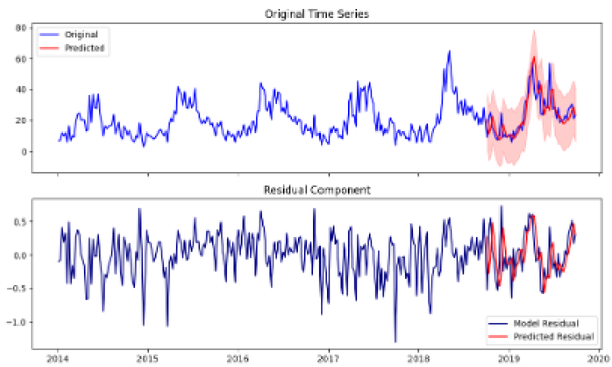
Among the four previously discussed strategies, the **long-term forecast** is particularly noteworthy. Analyzing predictions versus actual observations reveals that while the ARIMA model performs reasonably well in the initial weeks, it gradually diverges into a linear upward trend over time. This behavior highlights the model's inability to fully capture the underlying dynamics beyond the short term. Consequently, when predictions from the trend, seasonality, and residual models are combined, the final forecast's accuracy diminishes, as the residual model's limitations skew the overall trajectory.



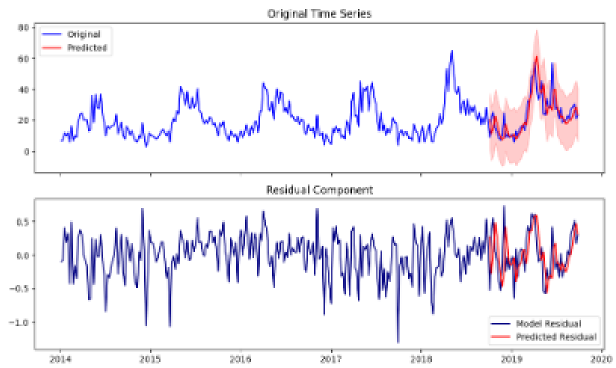
52-weeks ahead



52 weeks with 1-step ahead - Interaction



1-step ahead - Expanding windows

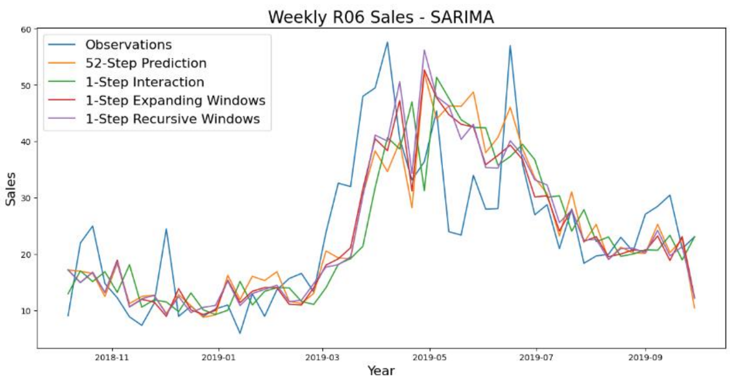


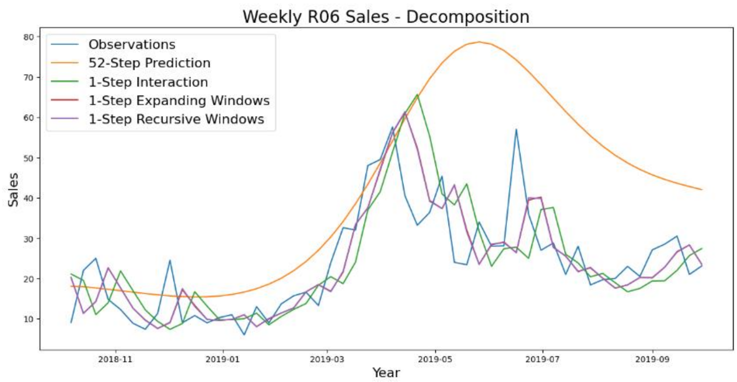
1-step ahead - Recursive windows

7. Assessing the quality of produced forecasts

7.1. Forecast plots

The analysis of the SARIMA forecasts reveals that **all predictions exhibit a similar pattern**, following the general trend of the time series. However, none of the models were able to capture key fluctuations in the data, particularly the early rise in sales between March and April, as well as the decline in sales during May and June. These gaps in prediction suggest that while the SARIMA model performs adequately in capturing the overall trend and seasonality, it **struggles with sudden or short-term shifts in the data**, highlighting a potential limitation in forecasting more dynamic or volatile periods.



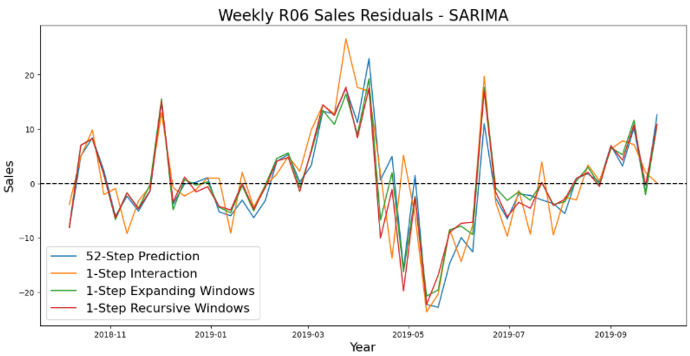


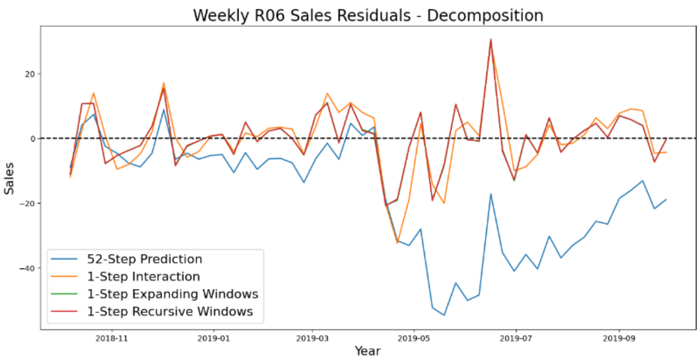
On the other hand, the Decomposition forecasts show that the 52-step prediction was reasonably accurate up until April but deviated significantly in the following months. However, the other forecasting strategies, which utilized new data points iteratively, adapted better to changes in the time series. The expanding windows and recursive windows forecasts produced similar results to the previous week's predictions, with only slight variations, reflecting the **models' reliance on progressively updated data**. In contrast, the interaction forecast seemed to lag by approximately two weeks instead of just one, which suggests a slight delay in adapting to the most recent trends.

7.2. Forecast residuals plot

The SARIMA model's forecasts were most accurate between December and February, and July to October, with small residuals indicating a good fit. However, between March and June, the model's accuracy decreased, as larger residuals showed that the forecasts did not capture fluctuations in sales during these months. This suggests the model struggled with seasonal variations during the spring and early summer.

The residuals from the Decomposition forecasts were generally close to the actual values, except for the 52-step prediction, which started well but became significantly inaccurate later. Overall, the residuals were somewhat unstable, indicating variability in forecast accuracy. Similar to the SARIMA model, the forecasts were most accurate between December and February, with smaller residuals during these months, suggesting better model performance during this period.





7.3. Accuracy measures

To evaluate the accuracy of the forecasts, we used several error metrics, including *Mean Error* (ME), *Root Mean Squared Error* (RMSE), *Mean Absolute Error* (MAE), *Mean Percentage Error* (MPE), *Mean Absolute Percentage Error* (MAPE), and *Auto-correlation at lag 1* (ACF1) of the residuals. Among these, **MAPE** and **ACF1** are particularly significant. **MAPE provides the absolute error**, offering a clear indication of forecast accuracy, unlike MPE, where positive and negative errors can cancel each other out.

ACF1, on the other hand, **measures the correlation in the residuals at lag 1**, helping to assess whether the model has effectively captured the underlying structure of the time series. If significant autocorrelation remains in the residuals, it suggests that the model has not fully accounted for all patterns in the data.

Model	Forecast Method	ME	RMSE	MAE	MPE	MAPE	ACF1
SARIMA(1, 1, 1)(1, 1, 1)[52]	52-weeks ahead	-0.012	9.077	6.787	-8.767	30.872	0.449
	1-step ahead - Interaction	0.468	9.579	7.078	-7.067	30.84	0.486
	1-step ahead - Expanding windows	0.424	8.499	6.319	-6.127	28.222	0.334
	1-step ahead - Recursive windows	0.13	8.617	6.463	-6.995	28.618	0.359
Decomposition with STL	52-weeks ahead	21.325	24.565	21.325	86.634	86.634	0.733
	1-step ahead - Interaction	-0.564	10.898	7.827	-7.815	33.995	0.496
	1-step ahead - Expanding windows	-0.141	8.973	6.461	-5.977	29.263	0.108
	1-step ahead - Recursive windows	-0.136	8.962	6.462	-5.979	29.31	0.097

The analysis of the accuracy measures shows that the **expanding windows method** provided the best forecasts for both models, followed closely by the recursive windows approach. The MAPE values were similar for both models, with SARIMA at 28.22% and Decomposition at 29.29%, indicating comparable overall accuracy. However, the ACF1 values were lower for the Decomposition model than for SARIMA, suggesting that the **SARIMA model could potentially be improved** to reduce residual correlations and achieve better results. When comparing the 52-step forecasts, SARIMA significantly outperformed the Decomposition model, with a MAPE of 30.87% compared to 85.45% for Decomposition, highlighting the Decomposition model's struggles with longer-term predictions.

8. Conclusion

To conclude, the **SARIMA(1,1,1)(1,1,1)$_{52}$** model proved to be the **most accurate**, particularly for long-term predictions such as the 52-step forecast. However, it is also the most **computationally expensive** and, as it requires seasonal differencing (D=1), it results in the loss of one year's worth of data. Additionally, the **wide confidence intervals** for the SARIMA model suggest uncertainty in the forecasts which do not adapt completely to the variance observed in this dataset. Despite these drawbacks, for long-term predictions where only one fit and one prediction are made, the computational cost is less of a concern. In this context, the SARIMA model demonstrated strong forecasting accuracy, making it a valuable option for forecasting over longer periods of up to one year.

On the other hand, the *combination of a Linear Model for Trend, Fourier Series for Seasonality, and ARIMA(1,2,2) for Residuals* provided **comparable results for rolling (1-step) forecasts**, while being far less computationally expensive. Although this model was less accurate for 52-step predictions compared to the SARIMA model, it still outperformed the ARIMA model alone, offering a more efficient alternative with relatively good forecasting performance. This decomposition method **strikes a balance between accuracy and computational efficiency, making it a practical option for scenarios where computational resources are a limiting factor.**

Notably, the long-term predictions from this method converged to the mean relatively quickly, highlighting a limitation in capturing extended patterns. However, using **shorter test sets could have mitigated this issue,** potentially improving the model's performance by focusing on smaller time horizons where the residual dynamics are more effectively modeled.

Overall, the best forecasting strategy was the **expanding windows** approach, which delivered the most reliable results across both models. This method's ability to continuously update the training data and adapt to new information provided a strong advantage in terms of forecast accuracy. While other strategies, such as recursive windows and interaction, also performed well, expanding windows stood out as the most consistent and robust method for generating accurate predictions.

9. References

[1] Zdravkovic, M. (2025). Pharma Sales Data. Kaggle. https://www.kaggle.com/datasets/milanzdravkovic/pharma-sales-data

[2] Cotta, P. V. (2020). Introdução ao ARIMA e SARIMA: Modelos de Séries Temporais. Medium. https://medium.com/@paulovpcotta/introdu%C3%A7%C3%A3o-ao-arima-e-sarima-modelos-de-s%C3%A9ries-temporais-51992ae5abfd

[3] Hyndman, R. J., & Athanasopoulos, G. (2020). STL Decomposition. In Forecasting: Principles and Practice (2nd ed.). OTexts. https://otexts.com/fpp2/stl.html

[4] Tankard, J. C. (2020). Forecasting with Fourier Series. Medium. https://medium.com/@jcatankard_76170/forecasting-with-fourier-series-8196721e7a3a

[5] Chakrabarti, Arijit, & Ghosh J. K. AIC, BIC and Recent Advances in Model Selection (2011) https://doi.org/10.1016/B978-0-444-51862-0.50018-6

[6] Moodle Course Info. (2025). Moodle Course. https://moodle2425.up.pt/course/view.php?id=6699