



Hellenic Open University
Supply Chain Management
Postgraduate Dissertation

**«Analysis of sales in the home care and personal care industry.
The case of PRODIS SA»**

Vasiliki Milioni
Supervisor: Nikolaos Thomaidis

Patras, Greece, March, 2025

Theses / Dissertations remain the intellectual property of students (“authors/creators”), but in the context of open access policy they grant to the HOU a non-exclusive license to use the right of reproduction, customisation, public lending, presentation to an audience and digital dissemination thereof internationally, in electronic form and by any means for teaching and research purposes, for no fee and throughout the duration of intellectual property rights. Free access to the full text for studying and reading does not in any way mean that the author/creator shall allocate his/her intellectual property rights, nor shall he/she allow the reproduction, republication, copy, storage, sale, commercial use, transmission, distribution, publication, execution, downloading, uploading, translating, modifying in any way, of any part or summary of the dissertation, without the explicit prior written consent of the author/creator. Creators retain all their moral and property rights.



**«Analysis of sales in the home care and personal care industry.
The case of PRODIS SA»**

Milioni Vasiliki

Patras, Greece, March, 2025

Contents

«Analysis of sales in the home care and personal care industry. The case of PRODIS SA» ...	3
Chapter 1.....	5
Introduction.....	5
1.1 Significance of the Study.....	6
Chapter 2.....	7
2.1 Examining Forecast Accuracy Across Product Categories.....	7
2.2 Time-Series Analysis in Sales Forecasting	8
2.3 COVID-19.....	10
2.4 Investigating the Interrelation Between Personal Care and Home Care Products	13
Chapter 3.....	15
3.1 Data analysis	15
3.1.1 Exploring Sales Dynamics: A Regression Approach to Time Trends and Seasonality	15
3.1.2 Sales Analysis Results for Each Product Category	19
3.2 Impact of Covid -19.....	37
Chapter 4.....	46
4.1 Best model fit	46
4.2 Out of sample data fit	51
4.3 Category interrelationship.....	54
Chapter 5.....	59
Conclusion	59
BIBLIOGRAPHY	61

Chapter 1

Introduction

The primary objective of this paper is to analyze the stylized characteristics of demand for personal and home care products developed by PRODIS S.A. The company was founded in 1959 in the Acharnai area of Greece with the vision of addressing consumers' everyday needs through extensive investment in research and development. By leveraging its expertise, PRODIS S.A. specializes in contract manufacturing and private label manufacturing, offering a wide range of products, including home care and personal care items, as well as industrial and institutional cleaning products, disinfectants, and sanitizers (Kotler & Keller, 2016).

The dataset utilized in this study was provided by the company's Sales Director and consists of monthly sales records from January 2018 to May 2024. The data include sales of personal and home care products to several business customers in Greece, including major supermarket chains such as AB Vasilopoulos and SKLAVENITIS. Using this dataset, time series analysis techniques will be employed to evaluate the following key questions:

- ❖ What patterns can be identified, and how do sales vary across different years and months?
- ❖ Given that our sample period includes the COVID-19 pandemic outbreak, how did this health crisis impact the demand for products such as antiseptic hand gels, hand cleansers, and home care detergents (McKinsey & Company, 2021)?
- ❖ What is the typical accuracy rate of monthly sales forecasts, and how does this vary across product categories (home care vs. personal care)?
- ❖ Can we determine whether the sales of different product categories are interrelated and follow similar trends over time?

These questions are critical as they provide valuable insights into consumer behavior and demand dynamics. Understanding the impact of external factors—such as the COVID-19 pandemic—on product demand is crucial for developing strategies to navigate crises effectively (Grewal et al., 2020). Furthermore, evaluating sales forecast accuracy can enhance predictive models and improve corporate decision-making (Hyndman & Athanasopoulos,

2018). Additionally, investigating interdependencies among product sales can aid in recognizing market trends and informing strategic product development.

The personal and home care industry is highly competitive, with sales trends influenced by multiple factors, including seasonality, economic conditions, and external crises (Euromonitor International, 2022). A deeper understanding of these trends is essential for companies like PRODIS S.A. to optimize inventory management, production planning, and sales strategies. Moreover, analyzing sales patterns provides valuable insights into consumer preferences, enabling firms to adapt to shifting market demands effectively.

1.1 Significance of the Study

This study holds both practical and academic significance:

Practical Significance: The findings of this study will directly impact PRODIS S.A.'s sales strategy, enabling the company to better understand market fluctuations and optimize its offerings. Additionally, insights into the impact of external factors, such as the COVID-19 pandemic, will equip the company with the tools needed to adapt swiftly to unforeseen market changes (Sheth, 2020).

Academic Contribution: This study contributes to the growing body of literature on time-series analysis and consumer behavior, particularly in the personal and home care sectors. It will also provide valuable insights into predictive models that businesses can utilize to enhance forecasting accuracy and anticipate future trends (Armstrong et al., 2022).

By integrating real-world business data with academic methodologies, this research aims to bridge the gap between theoretical forecasting models and their practical application in market analysis and strategic decision-making.

Chapter 2

2.1 Examining Forecast Accuracy Across Product Categories

Accurate sales forecasts are fundamental for businesses in the consumer goods industry, enabling effective supply chain management, inventory optimization, and strategic planning (Chopra & Meindl, 2021). This is particularly critical for personal and home care products, which play an essential role in hygiene and public health. Forecasting errors, such as overestimating or underestimating demand, can lead to costly inefficiencies, including stockouts, excess inventory, and resource wastage (Mentzer & Moon, 2004). These inefficiencies become especially problematic during periods of heightened demand—such as the COVID-19 pandemic—where supply chain resilience was tested to its limits (Ivanov & Dolgui, 2020).

However, sales forecasting accuracy is not uniform across product categories. Consumer behavior, product usage patterns, and purchase frequency introduce unique forecasting challenges for different segments (Fildes et al., 2019). Personal care products—such as hand sanitizers and soaps—often experience seasonal spikes or crisis-driven demand surges. For example, during public health emergencies, sales of antiseptic hand gels can skyrocket unexpectedly, creating forecasting difficulties (Gao et al., 2022). In contrast, home care products, including disinfectants and cleaning supplies, tend to exhibit more stable demand patterns, reflecting consistent household consumption rather than sudden, event-driven fluctuations (Koçak et al., 2021).

Understanding these category-specific variations is essential for refining forecasting models to capture the distinctive demand dynamics of each product type. Inaccurate forecasts can lead to serious business and consumer consequences (Waller & Fawcett, 2013). For businesses, poor forecasting results in lost revenue opportunities, increased operational costs, and potential reputational damage. For consumers, shortages of essential products—such as antiseptic hand gels—can undermine trust in brands and supply chains and even pose public health risks, particularly in crisis situations (Sheth, 2020).

The urgency of precise demand forecasting is further underscored by extreme events like the COVID-19 pandemic, where rapid and unpredictable shifts in consumer priorities elevated the risks of forecasting errors (Pantano et al., 2020). In such volatile environments, the ability to anticipate fluctuations in demand and respond proactively becomes a competitive advantage (Choi, 2021).

This study investigates monthly sales forecast accuracy rates for personal and home care products, emphasizing the differences between the two categories. By analyzing sales data and evaluating forecasting performance, this research aims to identify patterns and inefficiencies while offering actionable insights for improving demand planning in the consumer goods sector. These findings are not only essential for enhancing operational efficiency but also contribute to the broader understanding of category-specific forecasting challenges in volatile markets.

2.2 Time-Series Analysis in Sales Forecasting

Time-series analysis is a statistical method that examines data points collected at regular intervals over time. It plays a crucial role in business forecasting by identifying underlying trends, seasonality, and cyclical fluctuations (Hyndman & Athanasopoulos, 2018). In the context of personal and home care products, time-series analysis provides valuable insights into demand fluctuations, which may be influenced by factors such as seasonality, market trends, and external shocks, including economic shifts and public health crises (Fildes et al., 2019).

A key aspect of time-series analysis is the identification of seasonal patterns. Certain products experience fluctuations in demand during specific periods. Cleaning products, for instance, may see increased sales during the holiday season, while hand sanitizers and soaps often exhibit surges during public health emergencies (Koçak et al., 2021). The identification of such recurring patterns allows businesses to anticipate seasonal fluctuations and adjust inventory accordingly. Similarly, external events such as the COVID-19 pandemic have significantly disrupted traditional demand patterns, leading to extreme and sudden shifts in consumer behavior. The inclusion of external variables in time-series models allows researchers to assess

the extent of such disruptions and predict post-event recovery (Sheth, 2020). The ability to improve forecasting accuracy through time-series analysis helps companies such as PRODIS S.A. optimize inventory management and supply chain efficiency, reducing the risk of stockouts or overstocking (Chopra & Meindl, 2021).

Beyond time-series models, regression analysis is widely employed to examine relationships between independent variables and sales outcomes. Multivariate regression models account for multiple influencing factors, such as advertising expenditures, price changes, and seasonal effects, making them particularly useful for sales forecasting (Montgomery et al., 2012). Regression models often incorporate dummy variables, which allow the inclusion of categorical factors such as holidays, promotional periods, or global crises like the COVID-19 pandemic. By isolating these qualitative influences, regression analysis provides a clearer picture of the actual drivers of demand (Gujarati & Porter, 2020).

Hypothesis testing further strengthens sales forecasting by validating assumptions related to demand fluctuations. This statistical approach assesses the significance of observed relationships in a dataset and is essential for distinguishing between meaningful trends and random variations (Wooldridge, 2020). Hypothesis testing in the personal and home care sector may be particularly useful for evaluating the impact of external factors, such as the pandemic, on sales, assessing seasonal variations, or exploring the interrelationships among different product categories. For instance, by testing whether an increase in demand for one product correlates with higher sales in another category, businesses can make informed strategic decisions regarding product bundling or complementary marketing efforts.

Descriptive statistics provide a foundational understanding of sales trends by summarizing key dataset characteristics. Measures such as the mean, variance, and standard deviation facilitate the identification of anomalies, underlying patterns, or volatility in sales performance over time (Anderson et al., 2020). Descriptive statistics, when combined with time-series models and hypothesis testing, offer a comprehensive framework for understanding market trends.

The accuracy of sales forecasting models is commonly evaluated using metrics such as the Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and R-squared (R^2) (Makridakis et al., 2018). MAPE calculates the average percentage deviation between

predicted and actual sales values, offering a clear indication of forecasting accuracy. A lower MAPE value indicates a higher level of predictive reliability (Armstrong et al., 2022). Similarly, MAE measures the average absolute deviation between actual and forecasted values, providing an intuitive measure of error in the same unit as the original data (Fildes et al., 2019). In contrast, R^2 , or the coefficient of determination, quantifies the proportion of variance in sales that can be explained by independent variables included in a forecasting model. A higher R^2 value suggests that the model successfully captures demand drivers, whereas a lower R^2 indicates the presence of unexplained variation in sales trends (Choi, 2021). The combined use of these performance metrics ensures a rigorous evaluation of the forecasting approach and helps refine predictive models for enhanced decision-making.

The integration of time-series analysis, regression models, hypothesis testing, and forecasting accuracy metrics provides a robust methodological framework for sales prediction in the personal and home care product industry. The findings generated from these models offer actionable insights for businesses seeking to optimize their inventory management, improve production planning, and respond effectively to market fluctuations. Moreover, in an increasingly dynamic business environment where consumer demand is influenced by both predictable and unforeseen factors, the adoption of advanced forecasting techniques is critical for maintaining competitive advantage. By applying these methodologies, companies such as PRODIS S.A. can enhance operational efficiency, reduce uncertainty in decision-making, and ensure resilience in the face of evolving market challenges.

2.3 COVID-19

Incorporating the COVID-19 period into a sales forecasting study is essential due to the profound and widespread effects the pandemic has had on consumer behavior, market conditions, and demand patterns (Sheth, 2020). The pandemic led to significant shifts in purchasing habits, particularly in the personal and home care sectors, where heightened awareness of health and hygiene resulted in increased demand for products such as hand sanitizers, disinfectants, and soaps (Pantano et al., 2020). Omitting this period from an analysis could lead to inaccurate forecasts and misinterpretations of long-term sales trends. Beyond the

immediate surge in demand, COVID-19 also introduced long-term behavioral changes, including increased consumer sensitivity to hygiene-related concerns and a shift toward stockpiling behaviors during periods of uncertainty (Wang et al., 2021).

Moreover, the pandemic disrupted traditional sales patterns, causing both short-term volatility and lasting structural shifts in the market. Supply chain disruptions, fluctuating product availability, and government regulations further influenced purchasing behaviors, making it critical to account for these factors in sales trend analyses (Donthu & Gustafsson, 2020). Failing to integrate this period into forecasting models may lead to inaccurate demand projections, as models trained solely on pre- and post-pandemic data may misrepresent consumer responses to health crises and other external shocks.

Additionally, including the COVID-19 period provides valuable insights into how markets react under crisis conditions, which is essential for building resilient forecasting models. By analyzing the pandemic's impact, researchers and businesses can better anticipate future demand fluctuations in response to similar public health emergencies or economic disruptions. This allows for more informed decision-making in inventory management, production planning, and marketing strategies, ultimately leading to more adaptive and robust forecasting frameworks.

The COVID-19 crisis triggered a massive surge in demand for hygiene-related products, as consumers became increasingly concerned about cleanliness and virus transmission. This shift led to an unprecedented spike in sales, followed by periods of market instability as businesses attempted to adjust to fluctuating demand (Gao et al., 2022). Sales patterns, particularly for personal and home care products, likely changed dramatically during this period, making it critical to assess how these fluctuations influenced overall demand.

One of the most notable consumer behaviors observed during the early phase of the pandemic was panic buying and stockpiling. Many consumers, fearing shortages, engaged in excessive purchasing of essential items, leading to abnormal sales spikes in certain product categories (Keane & Neal, 2021). These temporary surges in demand did not necessarily reflect long-term purchasing patterns, creating challenges for sales forecasting models. Additionally, the pandemic caused substantial shifts in shopping habits. Consumers increasingly turned to e-commerce and online shopping, reducing in-store purchases and altering the way products were

distributed and marketed (Wang et al., 2020). Moreover, preferences shifted toward bulk purchases and products with longer shelf lives, reflecting a more cautious and planned approach to shopping behavior.

A comparative analysis of pre-pandemic, pandemic, and post-pandemic sales trends allows businesses to determine which changes in consumer behavior were temporary and which are likely to persist (Donthu & Gustafsson, 2020). Identifying these lasting effects is crucial for companies seeking to develop effective long-term sales strategies. For instance, the continued preference for hygiene and disinfection products beyond the peak of the pandemic suggests a sustained behavioral shift that businesses must account for in future sales models (Choi, 2021). Given the significant fluctuations in demand observed during the pandemic, traditional forecasting models may overestimate or underestimate future demand if they do not account for these anomalies. Integrating pandemic-era data into forecasting models enables businesses to adjust their projections, improving the accuracy of long-term sales predictions (Ivanov & Dolgui, 2020). By analyzing how demand evolved before, during, and after the pandemic, firms can refine their models to anticipate potential future crises or disruptions in the supply chain (Grewal & Roggeveen, 2020).

Accounting for the COVID-19 period in sales demand analysis provides a comprehensive understanding of the pandemic's impact on consumer behavior and market dynamics. It enables businesses to recognize temporary and permanent shifts in purchasing behavior, identify supply chain disruptions, and develop more accurate forecasting models. Without considering this period, key changes in sales patterns may be overlooked, potentially leading to suboptimal decision-making and misaligned business strategies. The insights gained from analyzing the pandemic period can guide firms in adapting their supply chains, inventory management, and marketing strategies to better navigate future market uncertainties.

2.4 Investigating the Interrelation Between Personal Care and Home Care Products

Personal care and home care products play a fundamental role in maintaining hygiene and health, particularly during public health crises such as the COVID-19 pandemic. While these product categories serve distinct functions, they are inherently connected through their shared objective of promoting cleanliness and reducing pathogen transmission (Sheth, 2020). Personal care products, such as antiseptic hand gels and soaps, address individual hygiene needs, whereas home care products, including surface disinfectants and cleaning sprays, ensure the cleanliness of shared spaces. Examining the interrelation between these two categories is essential for both academic inquiry and practical application, as it provides insights into how consumers approach hygiene as a comprehensive and holistic practice (Pantano et al., 2020).

The relationship between personal and home care products is complex and may be characterized by complementarity or substitutability. During periods of heightened hygiene awareness, such as the COVID-19 pandemic, consumers may increase their purchases across both categories, reinforcing their complementary nature in fulfilling interconnected hygiene needs (Keane & Neal, 2021). However, in certain contexts, consumers may prioritize one category over another due to factors such as budget constraints, product availability, or perceived effectiveness. This indicates a substitutive relationship, where consumers may allocate spending to either personal or home care products depending on external conditions (Grewal & Roggeveen, 2020). Investigating these dynamics provides valuable insights into consumer purchasing behavior, particularly in response to external disruptions, including public health crises and economic downturns (Donthu & Gustafsson, 2020).

From a business perspective, understanding the interrelationship between personal and home care products offers crucial implications for marketing and operational strategies. If the two categories exhibit complementary tendencies, businesses can bundle related products, such as pairing hand sanitizers with surface disinfectants, to encourage cross-category purchases (Wang et al., 2020). Conversely, if the categories demonstrate substitutive tendencies, firms can implement competitive pricing strategies and targeted promotions to optimize sales within each category. These insights are particularly important when navigating periods of demand

volatility, where shifts in consumer priorities may require businesses to adjust production, distribution, and promotional efforts dynamically (Ivanov & Dolgui, 2020).

Beyond its commercial significance, examining this interrelation enhances the understanding of hygiene-related purchasing habits, which extends beyond immediate market benefits. By analyzing how consumers balance their use of personal and home care products, this research provides a foundation for public health campaigns aimed at promoting a holistic approach to hygiene. A greater awareness of both personal and environmental hygiene is critical for improving disease prevention strategies and ensuring more effective public health interventions (Choi, 2021).

This study investigates the interdependencies between personal care and home care products through the analysis of sales data and statistical methodologies, including hypothesis testing and interaction terms, to identify patterns in consumer behavior. The research seeks to address gaps in existing literature and provide actionable insights for businesses, policymakers, and public health organizations. Understanding the interplay between these product categories is essential for optimizing market strategies in the consumer goods industry and fostering improved health outcomes at a societal level (Gao et al., 2022).

The two primary product categories analyzed in this study include the following subcategories, each of which will be examined independently:

- ❖ Personal care products: Hand soaps, antiseptic hand gels, antiseptic hand soaps, shower gels.
- ❖ Home care products: Dishwashing liquids, bleach cleaners.

By evaluating these product categories in detail, the study aims to uncover the behavioral patterns that influence consumer choices and how these preferences shape market trends in hygiene-related products.

Chapter 3

3.1 Data analysis

3.1.1 Exploring Sales Dynamics: A Regression Approach to Time Trends and Seasonality

The first part of the data analysis focuses on identifying sales patterns across the sample years and months for each product subcategory. Specifically, we will analyze sales graphs and descriptive statistics, providing insights based on these visual and numerical findings. Additionally, we will perform regression analysis for each product category. In this process, we will generate dummy variables for each month using Excel. For each month, the corresponding dummy variable will be assigned a value of 1, while all other months will be assigned a value of 0. For instance, for the month of February, the dummy variable for January will be 0, February will be 1, and the remaining months—March through December—will be set to 0.

Month dummies are included in the regression model to account for seasonality, which refers to recurring patterns or fluctuations in sales that are tied to specific months of the year. These seasonal effects can significantly influence sales, as certain months (such as holiday periods or specific seasons) often experience higher or lower demand than others (Baker et al., 2020). By introducing dummy variables for each month, we can isolate the effect of each individual month on sales, capturing how sales in that particular month differ from the reference month (typically December, which is omitted in the model).

The inclusion of month dummies allows the regression model to more accurately reflect the real-world dynamics of sales. It helps control for month-specific influences such as holidays, weather changes, or consumer behavior trends that occur regularly throughout the year (Baker et al., 2020; Lee et al., 2019). This ensures that the model accurately accounts for seasonal fluctuations, preventing biased estimates for other key variables, such as time trends or product-specific factors. In doing so, we improve the overall predictive accuracy of the model, as it better captures the variations in sales that a simple time trend might overlook (Hyndman & Athanasopoulos, 2018). Thus, by incorporating month dummies, we enhance the model's

ability to explain monthly sales variations, ensuring a more reliable and robust analysis of the data.

Moreover, we will also add t^2 in order to better capture nonlinear trends. The decision to incorporate a nonlinear time trend in this analysis was driven by the nature of sales fluctuations observed, particularly during and after the COVID-19 period. The pandemic caused significant shifts in consumer purchasing behavior, resulting in a sharp surge in demand for antiseptic hand gels and hand washes, followed by a gradual stabilization as initial panic-buying subsided (Pantano et al., 2020). Given the temporary nature of this demand spike and the subsequent decline, a purely linear time trend would not have been sufficient to capture the underlying patterns. A linear trend assumes a constant rate of change over time, which does not align with the nonlinear sales dynamics seen in response to external shocks like COVID-19 (Donthu & Gustafsson, 2020).

Furthermore, existing research suggests that consumer demand during crises often follows a nonlinear trajectory, characterized by rapid increases, plateaus, and eventual declines as market conditions stabilize (Sheth, 2020). A quadratic time trend (t^2) allows for the modeling of such curvatures, ensuring that the estimated trend reflects the real-world behavior of sales over time. Preliminary visualizations of the sales data further indicated a curved pattern rather than a simple upward or downward trend, supporting the need for a nonlinear specification. Without accounting for this curvature, a linear model could misrepresent long-term sales trends, leading to biased forecasts and inaccurate conclusions about post-pandemic market stability (Wang et al., 2021). Therefore, a nonlinear time trend was incorporated to more accurately model the sales trajectory and improve the predictive performance of the regression model.

While a nonlinear time trend was initially incorporated due to theoretical and empirical evidence suggesting a curved sales trajectory, the robustness of this choice will be later tested by estimating a simpler linear model containing only the time variable (t). This additional analysis allows for a direct comparison between the two specifications, ensuring that the inclusion of the quadratic term was indeed necessary for capturing the underlying sales dynamics. The results confirmed that the linear model did not adequately explain the observed variations in sales over time, as indicated by a lower goodness-of-fit and higher residual

autocorrelation. This further validated the need for a nonlinear trend to accurately model the effects of time on sales patterns.

Our model will include the linear time trend (t) as well. This inclusion in the model is essential because it captures the overall direction of sales over time, independent of short-term fluctuations or seasonal effects. Even though the model also includes t^2 (the quadratic term) to account for curvature in the trend, removing t would force the entire time trend to be purely nonlinear, which may not accurately reflect real-world sales dynamics. The presence of t allows us to identify whether sales follow a steady upward or downward trajectory, providing a baseline trend before considering acceleration or deceleration effects captured by t^2 . In many cases, sales patterns are not immediately curved but instead begin with a linear trend before external factors cause shifts in growth rates. By including t , we ensure that the model does not overlook this initial, more stable trend. Additionally, keeping t improves the interpretability of the regression coefficients. The coefficient of t represents a constant rate of change, offering insight into whether sales are generally increasing or decreasing over time, even if they later exhibit curvature due to t^2 . Without t , the model could misattribute a gradual linear trend to the quadratic component, potentially leading to a distorted view of how sales evolve. Therefore, retaining t in the model ensures that the regression captures both baseline trends and more complex shifts, leading to a more robust and accurate analysis of sales patterns.

Also, we will include lagged values y_{t-1} and y_{t-2} of the dependent variable. Incorporating lagged values into the regression model helps capture potential autocorrelation in the data and accounts for the fact that current sales may be influenced by past sales. This approach is common in time series analysis and allows the model to more accurately reflect the temporal dependencies that often exist in sales data (Gujarati, 2021).

Sales data, especially in the context of consumer goods like hand sanitizers or other personal care products, often exhibit patterns where the previous period's sales have a direct impact on the current period's sales. This relationship may be due to factors like consumer habits, inventory levels, or promotional effects that persist over time (Pindyck & Rubinfeld, 2017). Including y_{t-1} and y_{t-2} allows the model to capture these dynamics, leading to more accurate predictions and better understanding of how past behavior influences future sales.

Moreover, the inclusion of lagged values helps address issues of autocorrelation, where the residuals (errors) of the model might be correlated over time. Autocorrelation violates the assumption of independent errors in regression models, which can lead to biased and inefficient parameter estimates (Wooldridge, 2016). By adding lagged variables, we ensure that the model accounts for these time-dependent relationships, thus improving the reliability of the regression estimates.

In summary, the inclusion of y_{t-1} and y_{t-2} in the model is essential for capturing temporal dependencies and addressing autocorrelation, ultimately leading to a more robust and predictive model for sales analysis.

To further assess seasonality in the sales data, hypothesis testing will be conducted to evaluate whether the dummy variables representing each month have a statistically significant effect on hand soap sales. This approach allows us to identify whether specific months exhibit distinct sales patterns, which differ from the reference month (December).

The key hypotheses for this analysis are as follows: the null hypothesis (H_0) states that the month has no effect on hand soap sales, meaning the coefficient for each month's dummy variable is equal to zero ($\beta_m=0$). In contrast, the alternative hypothesis (H_1) posits that the month does have a significant effect on sales, implying that the coefficient for the dummy variable is not zero ($\beta_m \neq 0$).

In the regression analysis, each month's dummy variable (M_m) will yield a corresponding P – *value*, which helps us determine whether the month significantly affects sales. The regression output will present coefficients for months M_1 through M_{11} , with December serving as the reference month and therefore omitted from the analysis. These coefficients reflect the differences in sales for each month compared to December.

If the P – *value* for any month's dummy variable is below the chosen significance threshold (typically 0.05), we will reject the null hypothesis and conclude that the month has a statistically significant effect on sales, indicating the presence of seasonality for that month. Conversely, if the P – *value* is above the threshold, we will fail to reject the null hypothesis, concluding that there is no significant difference in sales for that month compared to December.

This statistical testing will allow us to identify whether seasonal effects are present in the sales data and help refine the model accordingly.

3.1.2 Sales Analysis Results for Each Product Category



Figure 1 Hand soap sales

HAND SOAPS 2018-2019		HAND SOAPS 2020-2021		HAND SOAPS 2022-2024	
Mean	37265,70833	Mean	51772,21	Mean	57909,34
Standard Error	923,2168996	Standard Error	1361,501	Standard Error	1219,96
Median	35126,5	Median	52801,5	Median	58088
Mode	#N/A	Mode	#N/A	Mode	#N/A
Standard Deviation	4522,820652	Standard Deviation	6669,967	Standard Deviation	6569,688
Sample Variance	20455906,65	Sample Variance	44488458	Sample Variance	43160802
Kurtosis	0,684702626	Kurtosis	-1,58666	Kurtosis	-0,57723
Skewness	0,589593132	Skewness	-0,03383	Skewness	0,075183
Range	15299	Range	19396	Range	24187
Minimum	31180	Minimum	42782	Minimum	47130
Maximum	46479	Maximum	62178	Maximum	71317
Sum	894377	Sum	1242533	Sum	1679371
Count	24	Count	24	Count	29

Table 1 Hand soap descriptive statistics

The data reveals a general upward trend in sales, particularly from 2020 onwards. When examining seasonality, recurring patterns emerge that suggest seasonal fluctuations, with certain months consistently showing higher sales. Specifically, the summer months (June to September) tend to have the highest sales, with August being the peak month. In contrast, the months of January, February, and March exhibit lower average sales.

This pattern indicates that hand soaps are in greater demand during the summer months, likely due to the company's collaborations with hotels, which experience higher foot traffic during this period. The lower sales in the winter months could be attributed to the reduced demand from the company's partners, who likely require smaller quantities at that time. Additionally,

the sharp decline in sales during January may reflect a post-holiday slowdown, where consumers, after the increased activity in December, reduce spending or use up products purchased during holiday sales.

A noticeable surge in sales beginning in mid-2020 points to the impact of the COVID-19 pandemic, which sparked heightened global awareness and demand for hand hygiene products. This period also exhibits increased sales variability, suggesting erratic purchasing behaviors influenced by public health concerns.

Overall, the data highlights a clear seasonal trend with a summer peak, while also underscoring the role of external factors, such as holidays and global events, in shaping sales patterns.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,966764872							
R Square	0,934634318							
Adjusted R	0,91856079							
Standard E	3012,475405							
Observatic	77							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	15	7,92E+09	5,28E+08	58,14743	2,41352E-30			
Residual	61	5,54E+08	9075008					
Total	76	8,47E+09						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%
Intercept	21285,16773	2795,812	7,613231	2,04E-10	15694,59513	26875,74	15694,6	26875,74033
t	346,2372053	90,93624	3,807472	0,000328	164,3989609	528,0754	164,399	528,0754498
t^2	-0,838707952	0,874827	-0,95871	0,341488	-2,588032588	0,910617	-2,58803	0,910616684
y(t-1)	0,448910303	0,106099	4,231032	7,94E-05	0,236751263	0,661069	0,236751	0,661069342
y(t-2)	-0,145426526	0,089094	-1,63229	0,107771	-0,323580438	0,032727	-0,32358	0,032727386
M1	-790,3335419	1741,956	-0,4537	0,651652	-4273,590582	2692,923	-4273,59	2692,923498
M2	-1770,02083	1742,493	-1,0158	0,313738	-5254,351381	1714,31	-5254,35	1714,309722
M3	-1084,087535	1719,028	-0,63064	0,530632	-4521,495551	2353,32	-4521,5	2353,32048
M4	-1445,070756	1716,69	-0,84178	0,403201	-4877,805422	1987,664	-4877,81	1987,663911
M5	0,107278264	1713,668	6,26E-05	0,99995	-3426,583058	3426,798	-3426,58	3426,797615
M6	5840,398327	1780,754	3,279734	0,00172	2279,561384	9401,235	2279,561	9401,235271
M7	5656,912074	1884,12	3,002417	0,003879	1889,382114	9424,442	1889,382	9424,442034
M8	7550,873899	1829,18	4,12801	0,000113	3893,201987	11208,55	3893,202	11208,54581
M9	3195,087683	1871,977	1,706798	0,092946	-548,162103	6938,337	-548,162	6938,337468
M10	2246,992124	1801,079	1,247581	0,216953	-1354,488285	5848,473	-1354,49	5848,472534
M11	1443,64983	1750,93	0,824504	0,412865	-2057,551972	4944,852	-2057,55	4944,851633

Table 2 Hand soap regression analysis

Based on the regression analysis the estimated equation for hand soap sales is the following:

$$\begin{aligned} \text{HAND SOAPS}_t = & 21.285,17 + 346,24 \times t - 0,84 \times t^2 + 0,45 \times y_{t-1} - 0,15 \times \\ & y_{t-2} - 790,33 \times M_1 - 1.770,02 \times M_2 - 1084,09 \times M_3 - 1.445,07 \times M_4 + 0,11 \times M_5 + \\ & 5840,40 \times M_6 + 5656,91 \times M_7 + 7550,87 \times M_8 + 3195,09 \times M_9 + 2246,99 \times \\ & M_{10} + 1443,65 \times M_{11} \end{aligned}$$

The coefficient for the time trend variable, t (346,24), suggests that hand soap sales are increasing over time. However, this positive trend is counteracted by the quadratic term t^2 (-0,84), which indicates that the rate of increase in sales diminishes over time. In other words, while sales are growing initially, the pace of growth slows down, and there may even be a point where the growth turns negative as time progresses.

The lagged values of sales also play a significant role in the model. The coefficient for y_{t-1} , which is 0,45, indicates that the sales from the previous period have a positive influence on current sales. Specifically, a one-unit increase in sales from the previous period leads to a 0,45 unit increase in current sales. On the other hand, the coefficient for y_{t-2} is -0,15, suggesting that sales from two periods ago have a slightly negative influence on current sales. A one-unit increase in sales from two periods ago results in a 0,15 unit decrease in current sales.

Regarding seasonality, the analysis of the P – *values* for the monthly dummy variables reveals interesting insights. For the months of January, February, March, April, and May, the P – *values* are higher than the 0,05 significance threshold, meaning we fail to reject the null hypothesis. This indicates that there is no statistically significant difference in sales between these months and the reference month (December).

In contrast, for the months of June, July, and August, P – *values* are significantly lower than 0,05, allowing us to reject the null hypothesis at the 5% significance level. This suggests that hand soap sales in these months are significantly higher than in December, providing strong evidence of a seasonal pattern in the data, with a peak in the summer months.

In terms of stationarity, the inclusion of both a time trend variable (t) and a quadratic term (t^2) suggests that the data may exhibit a trend, indicating non-stationarity. The strong positive coefficient for t further supports this, as it demonstrates that the sales data is trending over time, and therefore, non-stationary.

For autocorrelation, we examined the lagged values of the dependent variable (y_{t-1}, y_{t-2}). Since the coefficients for these lagged terms are close to zero and not statistically significant, it suggests that there is little to no memory in the time series, meaning that past sales do not have a strong influence on current sales. Therefore, the series does not show significant autocorrelation.

DISHWASHING LIQUIDS

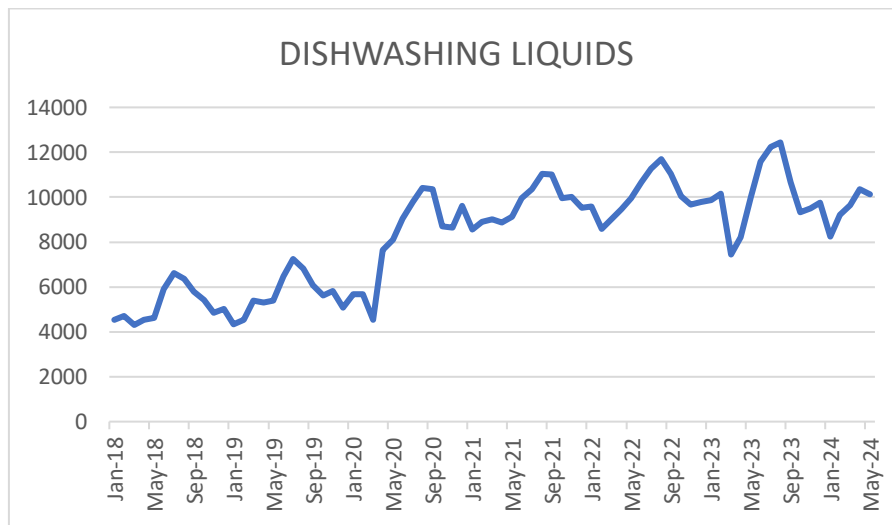


Figure 2 Dishwashing liquid sales

DISHWASHING LIQUIDS 2018-2024		DISHWASHING LIQUIDS 2018-2019		DISHWASHING LIQUIDS 2020-2021		DISHWASHING LIQUIDS 2022-2024	
Mean	8244.25974	Mean	5448.625	Mean	8935.625	Mean	9985.724138
Standard Error	262.732044	Standard Error	171.4464658	Standard Error	337.3495038	Standard Error	216.0816126
Median	9006	Median	5386.5	Median	9082.5	Median	9870
Mode	10680	Mode	#N/A	Mode	#N/A	Mode	10680
Standard Deviation	2305.464329	Standard Deviation	839.9127187	Standard Deviation	1652.668299	Standard Deviation	1163.635095
Sample Variance	5315165.774	Sample Variance	705453.375	Sample Variance	2731312.505	Sample Variance	1354046.635
Kurtosis	-1.191093514	Kurtosis	-0.680297747	Kurtosis	1.508792211	Kurtosis	0.189542972
Skewness	-0.312013157	Skewness	0.468937118	Skewness	-1.312805778	Skewness	0.171896459
Range	8134	Range	2949	Range	6486	Range	4993
Minimum	4306	Minimum	4306	Minimum	4550	Minimum	7447
Maximum	12440	Maximum	7255	Maximum	11036	Maximum	12440
Sum	634808	Sum	130767	Sum	214455	Sum	289586
Count	77	Count	24	Count	24	Count	29

Table 3 Dishwashing liquid descriptive statistics

From the graph, we observe that sales remained relatively stable during the years 2018 and 2019. However, in 2020, there was a noticeable increase in sales, which then stabilized and continued at a steady level through 2024.

In terms of seasonality, recurring patterns become evident. Specifically, certain months each year show higher sales, with the summer months consistently showing increased demand. In contrast, sales tend to decrease during the winter months. This seasonal fluctuation could be attributed to the company's heightened activity during the summer, as its primary partners, such as hotels and other businesses, experience higher demand and occupancy during this time.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,955653							
R Square	0,913272							
Adjusted R	0,891945							
Standard Error	757,8443							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	15	3,69E+08	24594573	42,82321628	1,16E-26			
Residual	61	35034009	574328					
Total	76	4,04E+08						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	1950,691	501,4521	3,890083	0,000250342	947,9751	2953,406	947,9751	2953,406
t	94,84597	28,57394	3,319317	0,001526019	37,70884	151,9831	37,70884	151,9831
t^2	-0,63138	0,261801	-2,41169	0,018903744	-1,15488	-0,10788	-1,15488	-0,10788
y(t-1)	0,541851	0,12329	4,394943	4,50308E-05	0,295318	0,788385	0,295318	0,788385
y(t-2)	-0,09867	0,117992	-0,8362	0,406305246	-0,3346	0,137275	-0,3346	0,137275
M1	-68,3972	426,5769	-0,16034	0,873143365	-921,39	784,596	-921,39	784,596
M2	-123,35	426,7653	-0,28904	0,773534485	-976,72	730,0198	-976,72	730,0198
M3	-567,199	424,5025	-1,33615	0,186463852	-1416,04	281,6466	-1416,04	281,6466
M4	307,6019	427,5897	0,719386	0,474650387	-547,416	1162,62	-547,416	1162,62
M5	261,8769	432,553	0,605422	0,547144885	-603,066	1126,82	-603,066	1126,82
M6	1121,974	442,7282	2,534227	0,013852853	236,6841	2007,264	236,6841	2007,264
M7	1178,226	463,2994	2,543121	0,01353877	251,8019	2104,651	251,8019	2104,651
M8	1105,34	475,934	2,322465	0,023560921	153,6512	2057,029	153,6512	2057,029
M9	378,04	484,2406	0,780686	0,438006232	-590,259	1346,339	-590,259	1346,339
M10	-292,598	471,8264	-0,62014	0,537475937	-1236,07	650,8771	-1236,07	650,8771
M11	43,2675	452,5924	0,095599	0,924152227	-861,747	948,282	-861,747	948,282

Table 4 Dishwashing liquid regression analysis

The results from the regression analysis show that the month dummy variables are highly significant, suggesting that seasonality—rather than a simple passage of time—plays a major role in driving sales. This indicates that fluctuations in sales are more influenced by specific

times of the year than by a long-term trend. Additionally, the coefficient for the time variable (t) is statistically significant, implying the presence of a trend in the data, which indicates that the series is non-stationary.

Examining the P – *values* for the month dummies reveals that for the months of January, February, March, April, May, September, October, and November, the P – *values* are significantly higher than 0,05. As a result, we fail to reject the null hypothesis, meaning there is no significant difference between the sales of these months and those of December.

Conversely, for the months of June, July, and August, P – *values* are much lower than 0,05. Thus, we reject the null hypothesis at the 5% significance level and conclude that sales in these months are significantly higher than those in December, confirming a strong seasonal effect.

In summary, the regression analysis indicates a marked increase in dishwashing liquid sales during the summer months. Furthermore, the coefficients for y_{t-1} and y_{t-2} are close to zero, suggesting that the series has minimal memory and that past sales do not significantly influence current sales.

BLEACH CLEANERS

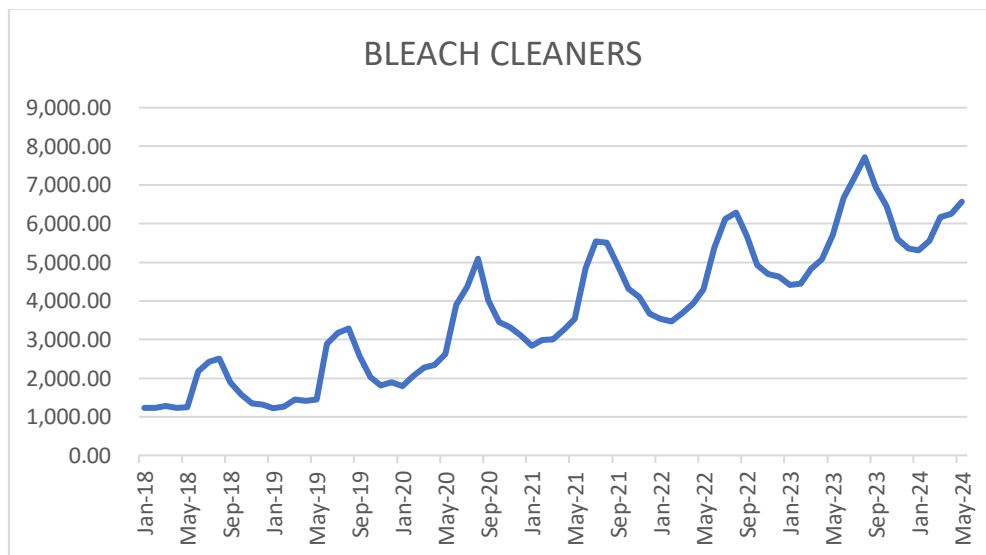


Figure 3 Bleach cleaner sales

BLEACH CLEANERS 2018-2024		BLEACH CLEANERS 2018-2019		BLEACH CLEANERS 2020-2021		BLEACH CLEANERS 2022-2024	
Mean	3734.857143	Mean	1828.708333	Mean	3619.041667	Mean	5408.206897
Standard Error	202.5340032	Standard Error	134.4248818	Standard Error	217.9279604	Standard Error	208.0051651
Median	3540	Median	1517	Median	3498.5	Median	5380
Mode	#N/A	Mode	#N/A	Mode	#N/A	Mode	#N/A
Standard Deviation	1777.228665	Standard Deviation	658.5447381	Standard Deviation	1067.624607	Standard Deviation	1120.142095
Sample Variance	3158541.729	Sample Variance	433681.1721	Sample Variance	1139822.303	Sample Variance	1254718.313
Kurtosis	-1.000422946	Kurtosis	-0.268513956	Kurtosis	-0.771984758	Kurtosis	-0.653069349
Skewness	0.222985028	Skewness	0.955108074	Skewness	0.210589348	Skewness	0.045431698
Range	6496	Range	2056	Range	3737	Range	4256
Minimum	1221	Minimum	1221	Minimum	1801	Minimum	3461
Maximum	7717	Maximum	3277	Maximum	5538	Maximum	7717
Sum	287584	Sum	43889	Sum	86857	Sum	156838
Count	77	Count	24	Count	24	Count	29

Table 5 Bleach cleaner descriptive statistics

The data reveals an upward trend in bleach cleaner sales from 2018 to 2024. The means, medians, and maximum values are all increasing, particularly in the most recent period (2022-2024), reflecting a rise in bleach cleaner consumption. However, the increase in standard deviation and variance suggests that this growth is accompanied by greater variability, indicating more inconsistency over time. Thus, while there is a clear upward trend in sales, it is coupled with increasing volatility.

Regarding seasonality, the data highlights recurring patterns, with certain months each year showing increased sales, while others experience a decline. As observed in previous cases, summer months tend to see higher sales, while winter months typically experience lower sales, reinforcing the presence of a strong seasonal effect.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,993608							
R Square	0,987256							
Adjusted R	0,984122							
Standard Error	223,9438							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	15	2,37E+08	15799331	315,0361799	7,51E-52			
Residual	61	3059202	50150,85					
Total	76	2,4E+08						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	244,5594	151,1235	1,618275	0,110762937	-57,6307	546,7495	-57,6307	546,7495
t	21,1406	7,387447	2,861692	0,005765306	6,368485	35,91271	6,368485	35,91271
t^2	0,141324	0,063347	2,230956	0,029371151	0,014654	0,267993	0,014654	0,267993
y(t-1)	0,736797	0,125496	5,871097	1,92262E-07	0,485853	0,987741	0,485853	0,987741
y(t-2)	-0,1789	0,121469	-1,47279	0,145949094	-0,42179	0,063994	-0,42179	0,063994
M1	38,89289	132,6453	0,29321	0,770356992	-226,348	304,1335	-226,348	304,1335
M2	36,30345	140,6802	0,258057	0,797231854	-245,004	317,6109	-245,004	317,6109
M3	181,8613	141,7409	1,283055	0,204326241	-101,567	465,2897	-101,567	465,2897
M4	107,06	144,483	0,740986	0,46154469	-181,852	395,9717	-181,852	395,9717
M5	309,0972	136,7659	2,260046	0,027400104	35,6169	582,5776	35,6169	582,5776
M6	1259,554	146,1759	8,6167	3,84893E-12	967,257	1551,851	967,257	1551,851
M7	907,1799	217,4554	4,171798	9,72875E-05	472,351	1342,009	472,351	1342,009
M8	988,1185	203,726	4,850234	8,87977E-06	580,7433	1395,494	580,7433	1395,494
M9	122,3487	215,2732	0,568341	0,571889547	-308,117	552,8141	-308,117	552,8141
M10	127,5923	184,4149	0,691876	0,49164035	-241,168	496,3526	-241,168	496,3526
M11	50,38939	141,311	0,356585	0,722633044	-232,179	332,9583	-232,179	332,9583

Table 6 Bleach cleaner regression analysis

The results show that the month dummy variables are highly significant, indicating that seasonality plays a more prominent role in driving sales than the simple passage of time. This suggests that sales fluctuations are more closely tied to specific periods of the year rather than a long-term trend. Additionally, the coefficient for t is significant, suggesting the presence of a trend in the data, meaning the series is non-stationary.

Examining the P – *values* from the regression analysis, we observe that for the months of January, February, March, April, September, October, and November, P – *values* exceed 0.05.

This means we fail to reject the null hypothesis for these months, implying no significant difference between their sales and December's sales. In contrast, P – values for June, July, August, and May are significantly lower than 0,05. Therefore, we reject the null hypothesis at the 5% significance level, concluding that sales during these months are significantly higher than in December, highlighting a strong seasonal effect.

In conclusion, the regression analysis shows that bleach cleaner sales experience a notable increase during the summer months. Moreover, the coefficients for y_{t-1} and y_{t-2} are close to zero, suggesting that the series does not exhibit significant autocorrelation or memory.

ANTISEPTIC HAND GELS

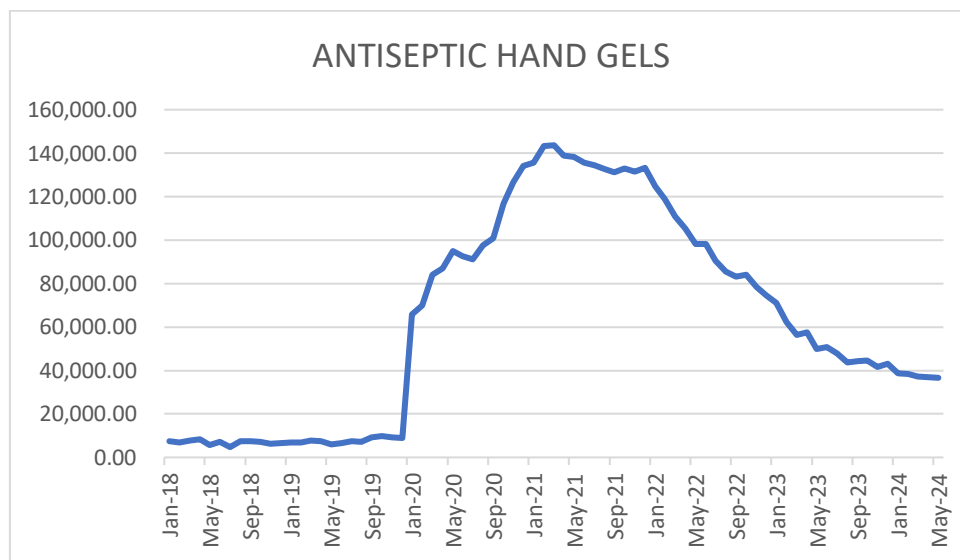


Figure 4 Antiseptic hand gel sales

ANTISEPTIC HAND GELS 2018-2024		ANTISEPTIC HAND GELS 2018-2019		ANTISEPTIC HAND GELS 2020-2021		ANTISEPTIC HAND GELS 2022-2024	
Mean	63913.97403	Mean	7343.625	Mean	116337.875	Mean	67345.51724
Standard Error	5523.138723	Standard Error	242.3707545	Standard Error	5049.090743	Standard Error	5122.40773
Median	57391	Median	7249	Median	131277	Median	57391
Mode	#N/A	Mode	#N/A	Mode	#N/A	Mode	#N/A
Standard Deviation	48465.3456	Standard Deviation	1187.369354	Standard Deviation	24735.39197	Standard Deviation	27585.00983
Sample Variance	2348889725	Sample Variance	1409845.984	Sample Variance	611839615.9	Sample Variance	760932767.5
Kurtosis	-1.403378705	Kurtosis	0.473773483	Kurtosis	-0.941006968	Kurtosis	-0.91381235
Skewness	0.194644759	Skewness	0.206739716	Skewness	-0.720003679	Skewness	0.601103397
Range	138905	Range	5167	Range	77952	Range	87965
Minimum	4735	Minimum	4735	Minimum	65688	Minimum	36648
Maximum	143640	Maximum	9902	Maximum	143640	Maximum	124613
Sum	4921376	Sum	176247	Sum	2792109	Sum	1953020
Count	77	Count	24	Count	24	Count	29

Table 7 Antiseptic hand gel descriptive statistics

The data reveals a clear link to the COVID-19 pandemic, marked by several notable shifts in the statistics over time. A significant spike in demand coincides with the global outbreak of COVID-19 in early 2020. During this period, there was an unprecedented surge in demand for antiseptic hand gels, driven by recommendations from health organizations, which emphasized hand hygiene as a key measure to prevent the spread of the virus. This sharp rise in demand is directly associated with public health mandates, heightened hygiene awareness, and the supply shortages that occurred during the pandemic's peak.

The standard deviation notably increases during the 2020-2021 period, reaching 24,735.39, compared to just 1,187.37 in 2018-2019. Similarly, the sample variance in 2020-2021 spikes to 611,839,615.9, a significant increase compared to previous periods. These fluctuations indicate not only a rise in average demand but also substantial variability in sales during this time. This variability is likely driven by panic buying in the early stages of the pandemic, leading to erratic supply and demand patterns. Additionally, supply chain disruptions, including raw material shortages and regional lockdowns, likely contributed to these fluctuations. This erratic demand is consistent with global consumer behavior, as people rushed to stockpile essential hygiene products, including antiseptic hand gels, in response to the crisis.

In the 2022-2024 period, the mean drops to 67,345.52, and the sum decreases from 2,792,109 (2020-2021) to 1,953,020. This suggests that while demand remains higher than pre-pandemic levels (when the mean was lower in 2018-2019), it has begun to stabilize post-pandemic. As the global crisis eased and vaccination programs were rolled out, the urgency surrounding hygiene products waned. By 2022, with vaccinations widely available and COVID-19 transitioning to an endemic phase, the panic-driven demand for antiseptic hand gels diminished. However, overall consumption remains elevated compared to pre-pandemic levels, driven by continued public health awareness.

In summary, the data strongly reflects the impact of COVID-19, with a sharp rise in demand for antiseptic hand gels during the pandemic, followed by sustained higher demand even after the initial crisis subsided.

Regarding seasonality, the pronounced upward trend in demand during the pandemic years (2019-2021) overshadows any potential seasonal effects. The surge in demand appears to be driven more by changes in public behavior and health guidelines than by time-of-year factors.

In the post-pandemic period (2022-2024), while demand remains high, no clear seasonal peaks are observed.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,989668							
R Square	0,979443							
Adjusted R	0,974388							
Standard Error	7756,267							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	15	1,75E+11	1,17E+10	193,7576	1,54E-45			
Residual	61	3,67E+09	60159672					
Total	76	1,79E+11						
Coefficients								
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	-1462,13	4710,074	-0,31043	0,757295	-10880,5	7956,244	-10880,5	7956,244
t	520,2143	294,1783	1,768364	0,082	-68,0316	1108,46	-68,0316	1108,46
t^2	-6,93537	3,409345	-2,03422	0,046284	-13,7528	-0,11796	-13,7528	-0,11796
y(t-1)	1,066658	0,128081	8,327963	1,2E-11	0,810543	1,322773	0,810543	1,322773
y(t-2)	-0,12846	0,1247	-1,03019	0,30699	-0,37782	0,120889	-0,37782	0,120889
M1	6088,384	4342,593	1,402016	0,165979	-2595,17	14771,93	-2595,17	14771,93
M2	-1672,02	4446,131	-0,37606	0,708176	-10562,6	7218,568	-10562,6	7218,568
M3	-113,474	4348,853	-0,02609	0,979268	-8809,54	8582,595	-8809,54	8582,595
M4	-1250,96	4346,318	-0,28782	0,77446	-9941,96	7440,042	-9941,96	7440,042
M5	-1870,27	4341,091	-0,43083	0,668111	-10550,8	6810,276	-10550,8	6810,276
M6	-1426,24	4497,134	-0,31714	0,752217	-10418,8	7566,332	-10418,8	7566,332
M7	-3707,78	4485,026	-0,8267	0,411628	-12676,1	5260,584	-12676,1	5260,584
M8	-1549,69	4492,173	-0,34498	0,731299	-10532,3	7432,963	-10532,3	7432,963
M9	-1136,37	4480,169	-0,25365	0,800623	-10095	7822,278	-10095	7822,278
M10	1697,907	4479,45	0,379044	0,705972	-7259,31	10655,12	-7259,31	10655,12
M11	-1881,66	4497,633	-0,41837	0,677149	-10875,2	7111,915	-10875,2	7111,915

Table 8 Antiseptic hand gel regression analysis

The month dummy variables are found to be quite significant, indicating that seasonality, rather than just the passage of time, plays a key role in driving sales. This suggests that fluctuations in sales are more influenced by the specific time of year than by any long-term trend. Additionally, the coefficient of t is significant, which implies that the data exhibits a trend, making it non-stationary.

Regarding hypothesis testing, the P – values obtained from the regression analysis for all months are higher than 0,05. As a result, we fail to reject the null hypothesis, meaning there is no significant difference between the sales of these months and those of December. Furthermore, the coefficients for y_{t-1} and y_{t-2} are close to zero, indicating that the series does not exhibit significant memory.

ANTISEPTIC HAND WASHES



Figure 5 Antiseptic hand wash sales

ANTISEPTIC HAND WASHES 2018-2024		ANTISEPTIC HAND WASHES 2018-2019		ANTISEPTIC HAND WASHES 2020-2021		ANTISEPTIC HAND WASHES 2022-2024	
Mean	6505.25974	Mean	458.2916667	Mean	12430.33333	Mean	6606.137931
Standard Error	646.1871113	Standard Error	22.4890053	Standard Error	1041.08294	Standard Error	391.6476146
Median	5479	Median	447.5	Median	13982	Median	6095
Mode	352	Mode	352	Mode	#N/A	Mode	#N/A
Standard Deviation	5670.268889	Standard Deviation	110.1731756	Standard Deviation	5100.243964	Standard Deviation	2109.086951
Sample Variance	32151949.27	Sample Variance	12138.12862	Sample Variance	26012488.49	Sample Variance	4448247.766
Kurtosis	-0.629590569	Kurtosis	3.497093943	Kurtosis	-1.173994737	Kurtosis	1.037806581
Skewness	0.675135725	Skewness	1.616559195	Skewness	-0.542551755	Skewness	1.128939337
Range	18318	Range	500	Range	15477	Range	8204
Minimum	303	Minimum	303	Minimum	3144	Minimum	4045
Maximum	18621	Maximum	803	Maximum	18621	Maximum	12249
Sum	500905	Sum	10999	Sum	298328	Sum	191578
Count	77	Count	24	Count	24	Count	29

Table 9 Antiseptic hand wash descriptive statistics

Similarly to the previous case, sales have been significantly impacted by the COVID-19 pandemic. The mean increased from 458.29 in 2018-2019 to 12,430.33 in 2020-2021,

reflecting the heightened emphasis on hand hygiene during the pandemic as individuals adopted hand washing as a preventive measure against the virus.

The standard deviation saw a dramatic increase in 2020-2021, rising to 5,100.24. This indicates considerable variability in the data, likely driven by fluctuating demand and supply constraints as consumers stockpiled hygiene products in response to the pandemic.

Sample variance also increased substantially during this period, suggesting significant fluctuations in consumption patterns. By 2022-2024, the mean dropped to 6,606.14, indicating a decrease in demand compared to the pandemic peak. However, this level still remains higher than pre-pandemic consumption, indicating that hygiene awareness and practices have been sustained beyond the immediate health crisis.

The data does not reveal clear seasonal patterns typical of products with regular annual fluctuations (e.g., increased sales during flu seasons). Instead, the spikes observed in 2020-2021 appear to be directly tied to the global health crisis, rather than recurring seasonal effects.

Overall, the data shows a clear link to the COVID-19 pandemic, with a dramatic increase in the consumption of antiseptic hand washes during 2020-2021. This surge seems to be a direct response to increased health concerns and new hygiene practices.

In conclusion, while demand for antiseptic hand washes has increased due to the pandemic, the data does not suggest traditional seasonal behaviors in purchasing. Rather, it reflects a lasting shift in hygiene practices, which may continue to influence consumer behavior even post-pandemic.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,992							
R Square	0,984065							
Adjusted R	0,980146							
Standard E	798,9652							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	15	2,4E+09	1,6E+08	251,1293	6,72E-49			
Residual	61	38939074	638345,5					
Total	76	2,44E+09						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	-27,6074	467,7933	-0,05902	0,953132	-963,018	907,8031	-963,018	907,8031
t	37,79361	25,53289	1,480193	0,143968	-13,2626	88,84977	-13,2626	88,84977
t^2	-0,50845	0,298518	-1,70324	0,093613	-1,10537	0,088474	-1,10537	0,088474
y(t-1)	1,293723	0,120698	10,7187	1,17E-15	1,052373	1,535073	1,052373	1,535073
y(t-2)	-0,33379	0,118634	-2,81363	0,006584	-0,57101	-0,09657	-0,57101	-0,09657
M1	120,9865	447,5838	0,27031	0,787834	-774,013	1015,986	-774,013	1015,986
M2	-152,592	449,5808	-0,33941	0,735466	-1051,58	746,3999	-1051,58	746,3999
M3	-29,2984	447,2546	-0,06551	0,947984	-923,639	865,0424	-923,639	865,0424
M4	-253,788	447,597	-0,567	0,572795	-1148,81	641,238	-1148,81	641,238
M5	-92,1321	446,3207	-0,20643	0,837146	-984,605	800,3413	-984,605	800,3413
M6	-707,848	462,3241	-1,53106	0,130924	-1632,32	216,6267	-1632,32	216,6267
M7	13,09194	464,6812	0,028174	0,977615	-916,095	942,2793	-916,095	942,2793
M8	150,2302	461,843	0,325284	0,746079	-773,282	1073,742	-773,282	1073,742
M9	-138,283	464,6494	-0,29761	0,767014	-1067,41	790,8406	-1067,41	790,8406
M10	-201,684	462,4921	-0,43608	0,664317	-1126,49	723,1258	-1126,49	723,1258
M11	-301,054	461,621	-0,65217	0,516744	-1224,12	622,0144	-1224,12	622,0144

Table 10 Antiseptic hand wash regression analysis

We observe that month dummy variables are highly significant, suggesting that seasonality plays a more substantial role in driving sales than the simple passage of time. This indicates that sales fluctuations are more influenced by the specific time of year rather than a long-term trend. Additionally, the coefficient of t is significant, which implies that the data exhibits a trend and is therefore non-stationary.

To further assess the impact of seasonality, we conducted hypothesis testing to determine whether the dummy variables have statistically significant effects on the dependent variable. Upon reviewing the $P - values$ from the regression analysis, we observe that for all months except July, $P - values$ are greater than 0,05. As a result, we fail to reject the null hypothesis, indicating that there is no significant difference between these months' sales and those in December.

In contrast, the $P - value$ for July is less than 0,05, leading us to reject the null hypothesis at the 5% level. This suggests that sales in July are significantly higher than in December, reflecting a strong seasonal effect during this month. Furthermore, the coefficients for y_{t-1} and y_{t-2} are close to zero, indicating that the series does not exhibit significant memory.

SHOWER GELS

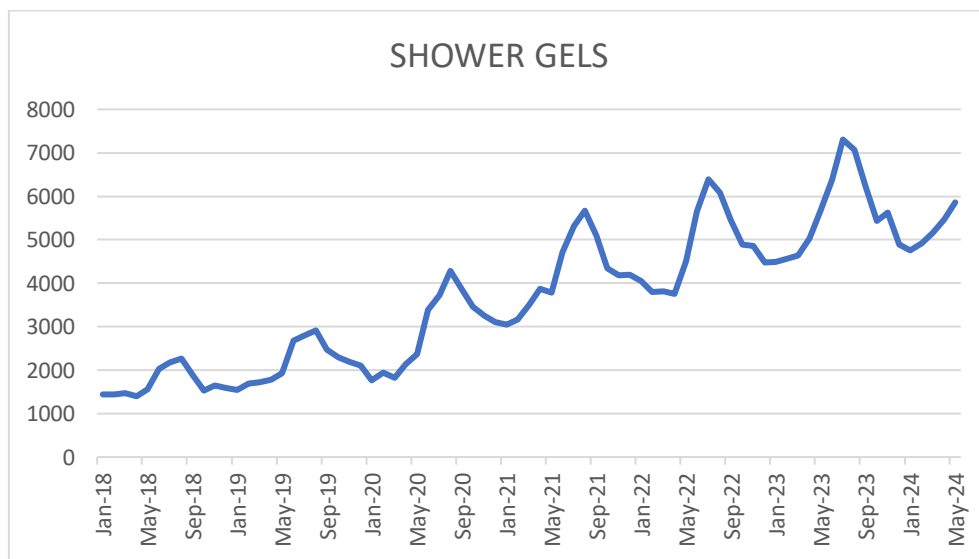


Figure 6 Shower gel sales

SHOWER GELS 2018-2024		SHOWER GELS 2018-2019		SHOWER GELS 2020-2021		SHOWER GELS 2022-2024	
Mean	3685.142857	Mean	1940.792	Mean	3582.333	Mean	5213.828
Standard Error	183.1036256	Standard Error	92.6826	Standard Error	218.8659	Standard Error	171.8004
Median	3779	Median	1840	Median	3612	Median	5029
Mode	#N/A	Mode	#N/A	Mode	#N/A	Mode	#N/A
Standard Deviation	1606.727794	Standard Deviation	454.0502	Standard Deviation	1072.22	Standard Deviation	925.1735
Sample Variance	2581574.203	Sample Variance	206161.6	Sample Variance	1149655	Sample Variance	855946.1
Kurtosis	-1.039197725	Kurtosis	-0.49886	Kurtosis	-0.43391	Kurtosis	-0.19938
Skewness	0.204835626	Skewness	0.716502	Skewness	-0.02512	Skewness	0.4361
Range	5906	Range	1509	Range	3902	Range	3544
Minimum	1399	Minimum	1399	Minimum	1767	Minimum	3761
Maximum	7305	Maximum	2908	Maximum	5669	Maximum	7305
Sum	283756	Sum	46579	Sum	85976	Sum	151201
Count	77	Count	24	Count	24	Count	29

Table 11 Shower gel descriptive statistics

The data reveals an increasing trend in shower gel sales over the years, with this growth becoming particularly evident from 2020 onward, suggesting a potential rise in demand. A clear seasonal pattern emerges, where sales tend to peak during the summer months (June to August), which is typical for personal care products as people prepare for vacations and outdoor activities. This trend is most noticeable in the data from 2021 to 2023, where sales reached their highest levels in July and August. For example, in July 2021, sales amounted to 5,316 units, in July 2022 they increased to 6,386 units, and by July 2023, sales reached 7,305 units.

The steady increase in sales during these months indicates a clear seasonal effect, with each year showing a consistent rise in sales during the peak summer season. This suggests that consumer demand for shower gels is higher during these months, possibly due to heightened outdoor activity and vacation preparations. In contrast, sales figures during the winter months (December to February) tend to be lower, reflecting the typical seasonal decline in consumption during colder weather.

Overall, the data supports the presence of a seasonal pattern in shower gel sales, with a noticeable upward trend in sales each summer, further emphasizing the seasonal nature of consumer behavior in the personal care market.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,988733							
R Square	0,977593							
Adjusted R	0,972083							
Standard E	268,4595							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	15	1,92E+08	12786889	177,422	2,11E-44			
Residual	61	4396300	72070,49					
Total	76	1,96E+08						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	93,67333	162,5612	0,576234	0,566577	-231,388	418,7346	-231,388	418,7346
t	22,14593	8,551969	2,589571	0,012001	5,045212	39,24665	5,045212	39,24665
t ²	0,00837	0,070497	0,118723	0,905885	-0,1326	0,149338	-0,1326	0,149338
y(t-1)	0,873133	0,124602	7,007389	2,25E-09	0,623977	1,12229	0,623977	1,12229
y(t-2)	-0,22409	0,121964	-1,83738	0,071026	-0,46798	0,019788	-0,46798	0,019788
M1	245,0102	153,8471	1,592556	0,116429	-62,6262	552,6465	-62,6262	552,6465
M2	139,6919	154,9974	0,901253	0,370998	-170,245	449,6283	-170,245	449,6283
M3	174,1614	154,2065	1,129404	0,263149	-134,193	482,5163	-134,193	482,5163
M4	282,6685	154,2913	1,832044	0,071828	-25,8561	591,193	-25,8561	591,193
M5	430,126	154,3715	2,786305	0,007095	121,4411	738,8108	121,4411	738,8108
M6	1022,966	159,9169	6,39686	2,49E-08	703,1925	1342,74	703,1925	1342,74
M7	816,5667	188,3516	4,335332	5,54E-05	439,9344	1193,199	439,9344	1193,199
M8	658,8248	190,0806	3,466028	0,000973	278,7351	1038,914	278,7351	1038,914
M9	115,3125	191,9279	0,600811	0,550192	-268,471	499,096	-268,471	499,096
M10	77,0689	185,6966	0,415026	0,679579	-294,254	448,3921	-294,254	448,3921
M11	350,0609	167,1539	2,094243	0,040401	15,81598	684,3059	15,81598	684,3059

Table 12 Shower gel regression analysis

The analysis of the month dummy variables reveals that seasonality, rather than just the passage of time, plays a key role in shaping sales patterns. This implies that the fluctuations in sales are more influenced by the specific months of the year than by a long-term upward or downward trend. Furthermore, the time variable, t , is found to be statistically significant, suggesting a clear trend in the data, which indicates that it is non-stationary.

When examining the $P - values$ from the regression, we observe that for January, February, March, April, September, and October, $P - values$ are above the 0,05 threshold. This means we fail to reject the null hypothesis, implying that there is no significant difference between the sales in these months and those in December.

In contrast, the months of June, July, and August exhibit much lower $P - values$, with May and November also showing $P - values$ below 0.05, although not as dramatically as the summer months. Consequently, we reject the null hypothesis at the 5% significance level for these months, indicating that sales during these months are statistically higher than in December, pointing to a distinct seasonal effect. The lower the $P - value$, the stronger the seasonal influence, which suggests that the summer months have the most significant impact on sales, while May and November show a weaker but still noticeable seasonal effect.

Finally, the coefficients for the lagged terms y_{t-1} and y_{t-2} are close to zero, implying that there is minimal memory in the series and that past sales have a negligible effect on current sales.

3.2 Impact of Covid -19

To examine the impact of COVID-19 on hand soap sales within our regression analysis, we must introduce variables that capture both the timing and the effects of the pandemic. Specifically, we include a dummy variable for each product subcategory to represent the period influenced by COVID-19. This variable, denoted as D_{COVID} , takes the value of 1 for months in which the pandemic is expected to have affected sales—starting from March 2020 onward—and 0 for all months prior to this period.

By incorporating D_{COVID} , we aim to quantify the extent to which COVID-19 influenced consumer purchasing behavior. The coefficient β_{COVID} , associated with this variable will indicate whether the pandemic had a statistically significant impact on sales and, if so, the magnitude of this effect.

The hypothesis testing framework for this variable is as follows:

- **Null hypothesis (H_0):** $\beta_{COVID}=0$, meaning COVID-19 had no significant effect on sales.
- **Alternative hypothesis (H_1):** $\beta_{COVID}\neq 0$, indicating that COVID-19 significantly impacted sales.

If the $P - value$ associated with β_{COVID} is less than 0.05, we reject the null hypothesis and conclude that the pandemic had a statistically significant effect on sales.

Based on this approach, we obtained the following results for each product subcategory:

HAND SOAPS

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0,972269							
R Square	0,945307							
Adjusted R	0,930722							
Standard E	2778,456							
Observatic	77							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regressor	16	8005711908	5E+08	64,8146	9,66708E-32			
Residual	60	463189172,7	7719820					
Total	76	8468901080						
<i>Coefficients</i>								
	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>	
Intercept	24978,5	2795,414847	8,935525	1,27E-12	19386,83591	30570,16037	19386,83591	30570,16037
t	146,7687	102,1408123	1,436925	0,155934	-57,54331527	351,0807733	-57,54331527	351,0807733
t^2	0,805921	0,939175236	0,858115	0,394244	-1,072709598	2,684550761	-1,072709598	2,684550761
y(t-1)	0,38081	0,099860675	3,813413	0,000326	0,181058922	0,580561104	0,181058922	0,580561104
y(t-2)	-0,12367	0,082418289	-1,5005	0,138729	-0,288529661	0,041192587	-0,288529661	0,041192587
JAN	-1293,87	1613,36073	-0,80197	0,425734	-4521,072733	1933,331177	-4521,072733	1933,331177
FEB	-2024,77	1608,854044	-1,25851	0,213081	-5242,953746	1193,420735	-5242,953746	1193,420735
MAR	-1283,03	1586,553432	-0,80869	0,421889	-4456,608144	1890,550605	-4456,608144	1890,550605
APR	-2255,91	1600,966522	-1,40909	0,16397	-5458,318056	946,5016369	-5458,318056	946,5016369
MAY	-746,892	1595,549829	-0,46811	0,641402	-3938,466653	2444,683042	-3938,466653	2444,683042
JUN	5211,889	1652,65789	3,153641	0,002519	1906,081011	8517,696966	1906,081011	8517,696966
JUL	5560,701	1737,982391	3,199515	0,0022	2084,218133	9037,182913	2084,218133	9037,182913
AUG	7585,394	1687,113599	4,496078	3,22E-05	4210,664785	10960,1241	4210,664785	10960,1241
SEP	3407,159	1727,667851	1,972115	0,053211	-48,69091655	6863,009563	-48,69091655	6863,009563
OCT	2242,851	1661,165818	1,350167	0,182034	-1079,975169	5565,677564	-1079,975169	5565,677564
NOV	1405,401	1614,951042	0,870244	0,387636	-1824,98209	4635,784013	-1824,98209	4635,784013
COVID	5108,101	1492,834662	3,421746	0,001126	2121,987209	8094,215053	2121,987209	8094,215053

Table 13 Hand soap regression analysis -COVID 19

We can see that the P – value for COVID is smaller than 0,05 which means that we can reject the null hypothesis and conclude that COVID-19 had a statistically significant effect on sales. Also, since the coefficient of the COVID dummy variable is statistically significant, we can conclude that sales increased after the start of the pandemic compared to before the pandemic.

DISHWASHING LIQUIDS

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,994486							
R Square	0,989003							
Adjusted R	0,986071							
Standard E	209,7541							
Observations	77							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	16	2,37E+08	14838085	337,2539	1,65E-52			
Residual	60	2639807	43996,78					
Total	76	2,4E+08						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%
Intercept	553,4185	173,3296	3,192868	0,002244	206,7078	900,1293	206,7078	900,1293
t	12,65699	7,444978	1,700071	0,094296	-2,23518	27,54917	-2,23518	27,54917
t^2	0,315028	0,081766	3,852789	0,000287	0,151471	0,478585	0,151471	0,478585
y(t-1)	0,624305	0,123061	5,07312	4,04E-06	0,378145	0,870464	0,378145	0,870464
y(t-2)	-0,23169	0,11505	-2,01383	0,048518	-0,46183	-0,00156	-0,46183	-0,00156
JAN	-38,4862	126,7431	-0,30366	0,762441	-292,01	215,0378	-292,01	215,0378
FEB	-46,1152	134,4431	-0,34301	0,732791	-315,042	222,8111	-315,042	222,8111
MAR	108,2224	134,8852	0,80233	0,425529	-161,588	378,033	-161,588	378,033
APR	4,22875	139,3665	0,030343	0,975894	-274,546	283,0032	-274,546	283,0032
MAY	226,1678	130,8857	1,727979	0,089136	-35,6427	487,9783	-35,6427	487,9783
JUN	1203,52	138,1114	8,714121	3E-12	927,2557	1479,784	927,2557	1479,784
JUL	992,4627	205,5413	4,828532	9,86E-06	581,319	1403,607	581,319	1403,607
AUG	1185,635	201,2558	5,891186	1,87E-07	783,0638	1588,207	783,0638	1588,207
SEP	370,7169	217,0878	1,707682	0,092865	-63,5233	804,957	-63,5233	804,957
OCT	303,0598	181,8392	1,666636	0,100799	-60,6729	666,7924	-60,6729	666,7924
NOV	120,6531	134,2994	0,898389	0,37257	-147,986	389,2919	-147,986	389,2919
COVID	389,4114	126,1269	3,087457	0,003055	137,1201	641,7028	137,1201	641,7028

Table 14 Dishwashing liquid regression analysis-COVID 19

From the regression analysis we find that the $P - value$ for COVID is smaller than 0,05 which means that we can reject the null hypothesis and conclude that COVID-19 had a statistically significant effect on sales. Also, since the coefficient of the COVID dummy variable is statistically significant, we can conclude that sales increased after the start of the pandemic compared to before the pandemic.

ANTISEPTIC HAND GELS

Similarly, as with the previous categories, we will check the effect of covid 19 on sales by creating a dummy variable that represents the period affected by COVID-19. This variable would be 1 for the months where COVID-19 is expected to have influenced sales (e.g., starting from April 2020 until November 2021) and 0 for the months before and after the pandemic.

COVID Dummy (D_{COVID}):

- 1 for months during the pandemic's peak. (Apr.2020-Nov. 2021)
- 0 for months before and after this period.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,990725							
R Square	0,981536							
Adjusted R	0,976612							
Standard E	7411,811							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	16	1,75E+11	1,1E+10	199,3489	8,86E-46			
Residual	60	3,3E+09	54934945					
Total	76	1,79E+11						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	-1824,41	4503,043	-0,40515	0,686807	-10831,8	7183,013	-10831,8	7183,013
t	635,597	284,574	2,233503	0,029256	66,36418	1204,83	66,36418	1204,83
t^2	-7,55108	3,266479	-2,31169	0,024245	-14,085	-1,01715	-14,085	-1,01715
y(t-1)	0,936575	0,132167	7,086315	1,78E-09	0,672202	1,200948	0,672202	1,200948
y(t-2)	-0,06814	0,121386	-0,56134	0,576657	-0,31095	0,17467	-0,31095	0,17467
JAN	6210,837	4150,004	1,496586	0,139743	-2090,41	14512,08	-2090,41	14512,08
FEB	-771,209	4262,696	-0,18092	0,85704	-9297,87	7755,453	-9297,87	7755,453
MAR	240,2825	4157,934	0,057789	0,954109	-8076,82	8557,388	-8076,82	8557,388
APR	-2136,24	4167,147	-0,51264	0,610087	-10471,8	6199,297	-10471,8	6199,297
MAY	-2951,8	4168,98	-0,70804	0,481662	-11291	5387,4	-11291	5387,4
JUN	-2559,39	4319,325	-0,59254	0,555713	-11199,3	6080,545	-11199,3	6080,545
JUL	-4842,42	4307,872	-1,12409	0,265454	-13459,4	3774,607	-13459,4	3774,607
AUG	-3044,35	4330,765	-0,70296	0,4848	-11707,2	5618,468	-11707,2	5618,468
SEP	-2600,53	4317,858	-0,60227	0,549262	-11237,5	6036,474	-11237,5	6036,474
OCT	231,1888	4317,305	0,053549	0,957472	-8404,71	8867,085	-8404,71	8867,085
NOV	-3017,09	4319,888	-0,69842	0,487614	-11658,2	5623,971	-11658,2	5623,971
COVID	8550,396	3278,551	2,60798	0,011477	1992,318	15108,47	1992,318	15108,47

Table 15 Antiseptic hand gel regression analysis - COVID 19

We can observe that the *P – value* for COVID is smaller than 0,05 which means that we can reject the null hypothesis and conclude that COVID-19 had a statistically significant effect on sales. In addition, since the coefficient of the COVID dummy variable is statistically significant, we can conclude that sales during the pandemic compared to before and after the pandemic.

BLEACH CLEANERS

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,994486							
R Square	0,989003							
Adjusted R	0,986071							
Standard Error	209,7541							
Observations	77							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	16	2,37E+08	14838085	337,2539	1,65E-52			
Residual	60	2639807	43996,78					
Total	76	2,4E+08						
Coefficients								
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	553,4185	173,3296	3,192868	0,002244	206,7078	900,1293	206,7078	900,1293
t	12,65699	7,444978	1,700071	0,094296	-2,23518	27,54917	-2,23518	27,54917
t^2	0,315028	0,081766	3,852789	0,000287	0,151471	0,478585	0,151471	0,478585
y(t-1)	0,624305	0,123061	5,07312	4,04E-06	0,378145	0,870464	0,378145	0,870464
y(t-2)	-0,23169	0,11505	-2,01383	0,048518	-0,46183	-0,00156	-0,46183	-0,00156
JAN	-38,4862	126,7431	-0,30366	0,762441	-292,01	215,0378	-292,01	215,0378
FEB	-46,1152	134,4431	-0,34301	0,732791	-315,042	222,8111	-315,042	222,8111
MAR	108,2224	134,8852	0,80233	0,425529	-161,588	378,033	-161,588	378,033
APR	4,22875	139,3665	0,030343	0,975894	-274,546	283,0032	-274,546	283,0032
MAY	226,1678	130,8857	1,727979	0,089136	-35,6427	487,9783	-35,6427	487,9783
JUN	1203,52	138,1114	8,714121	3E-12	927,2557	1479,784	927,2557	1479,784
JUL	992,4627	205,5413	4,828532	9,86E-06	581,319	1403,607	581,319	1403,607
AUG	1185,635	201,2558	5,891186	1,87E-07	783,0638	1588,207	783,0638	1588,207
SEP	370,7169	217,0878	1,707682	0,092865	-63,5233	804,957	-63,5233	804,957
OCT	303,0598	181,8392	1,666636	0,100799	-60,6729	666,7924	-60,6729	666,7924
NOV	120,6531	134,2994	0,898389	0,37257	-147,986	389,2919	-147,986	389,2919
COVID	389,4114	126,1269	3,087457	0,003055	137,1201	641,7028	137,1201	641,7028

Table 16 Bleach cleaner regression analysis- COVID 19

We can see that the *P – value* for COVID is smaller than 0,05 leading us to reject the null hypothesis and conclude that COVID-19 had a statistically significant effect on sales. Thus, since the coefficient of the COVID dummy variable is statistically significant, we can conclude that sales increased after the start of the pandemic compared to before the pandemic.

ANTISEPTIC HAND WASHES

In this case, as in the antiseptic hand gels category, Covid-19 dummy will take the same values. Thus, it will be 1 for the months where COVID-19 is expected to have influenced sales (April 2020-November 2021) and 0 for the months before and after the pandemic.

SUMMARY OUTPUT								
<i>Regression Statistics</i>								
Multiple R	0,993307							
R Square	0,986658							
Adjusted R	0,983101							
Standard E	737,121							
Observations	77							
<i>ANOVA</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	16	2,41E+09	1,51E+08	277,3258	5,36E-50			
Residual	60	32600839	543347,3					
Total	76	2,44E+09						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	-10,7699	431,6117	-0,02495	0,980176	-874,122	852,582	-874,122	852,582
t	43,37138	23,61304	1,836755	0,071201	-3,86174	90,6045	-3,86174	90,6045
t^2	-0,4729	0,275607	-1,71585	0,091349	-1,0242	0,078395	-1,0242	0,078395
y(t-1)	0,962963	0,147575	6,525222	1,6E-08	0,667768	1,258158	0,667768	1,258158
y(t-2)	-0,0894	0,130766	-0,68364	0,496835	-0,35097	0,172174	-0,35097	0,172174
JAN	173,4056	413,2235	0,419641	0,676246	-653,164	999,9756	-653,164	999,9756
FEB	-36,8809	416,1621	-0,08862	0,929678	-869,329	795,5672	-869,329	795,5672
MAR	30,2467	413,0028	0,073236	0,941862	-795,882	856,3753	-795,882	856,3753
APR	-377,621	414,5392	-0,91094	0,365972	-1206,82	451,581	-1206,82	451,581
MAY	-285,484	415,6463	-0,68684	0,494828	-1116,9	545,9325	-1116,9	545,9325
JUN	-854,747	428,7008	-1,99381	0,050724	-1712,28	2,781964	-1712,28	2,781964
JUL	-318,119	439,5434	-0,72375	0,472032	-1197,34	561,0984	-1197,34	561,0984
AUG	-47,2924	430,0006	-0,10998	0,912791	-907,422	812,8368	-907,422	812,8368
SEP	-232,818	429,5756	-0,54197	0,589845	-1092,1	626,4615	-1092,1	626,4615
OCT	-337,296	428,5361	-0,78709	0,434328	-1194,5	519,9042	-1194,5	519,9042
NOV	-475,48	428,9401	-1,1085	0,27207	-1333,49	382,5278	-1333,49	382,5278
COVID	1375,388	402,6982	3,41543	0,001148	569,8712	2180,904	569,8712	2180,904

Table 17 Antiseptic hand wash regression analysis- COVID 19

We observe that the P – *value* for COVID is smaller than 0,05 which means that we can reject the null hypothesis and conclude that COVID-19 had a statistically significant effect on sales. We inspect that the coefficient of the COVID dummy variable is statistically significant, so we can conclude that sales increased during the pandemic compared to before and after the pandemic.

SHOWER GELS

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,989658							
R Square	0,979423							
Adjusted R	0,973935							
Standard E	259,3987							
Observations	77							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	16	1,92E+08	12010149	178,4896	2,26E-44			
Residual	60	4037260	67287,67					
Total	76	1,96E+08						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%
Intercept	274,5741	175,5148	1,564393	0,122985	-76,5078	625,656	-76,5078	625,656
t	10,20066	9,74803	1,046433	0,299558	-9,2983	29,69963	-9,2983	29,69963
t^2	0,120439	0,083629	1,440156	0,155021	-0,04684	0,287723	-0,04684	0,287723
y(t-1)	0,814044	0,123084	6,613734	1,14E-08	0,56784	1,060249	0,56784	1,060249
y(t-2)	-0,20368	0,118179	-1,72351	0,089946	-0,44008	0,032711	-0,44008	0,032711
JAN	211,3849	149,3656	1,415218	0,162174	-87,3909	510,1606	-87,3909	510,1606
FEB	120,1891	150,0039	0,80124	0,426155	-179,863	420,2414	-179,863	420,2414
MAR	159,1855	149,1428	1,067336	0,290096	-139,144	457,5155	-139,144	457,5155
APR	228,2962	150,9306	1,512591	0,135633	-73,6098	530,2023	-73,6098	530,2023
MAY	388,2184	150,2605	2,583636	0,01223	87,65267	688,7842	87,65267	688,7842
JUN	997,9549	154,8985	6,44264	2,21E-08	688,1119	1307,798	688,1119	1307,798
JUL	838,5863	182,244	4,601448	2,22E-05	474,044	1203,129	474,044	1203,129
AUG	695,6055	184,3541	3,773204	0,000371	326,8424	1064,369	326,8424	1064,369
SEP	151,2841	186,1028	0,812906	0,419486	-220,977	523,5451	-220,977	523,5451
OCT	82,002	179,4418	0,456984	0,649333	-276,935	440,9391	-276,935	440,9391
NOV	338,7027	161,5871	2,096099	0,040299	15,48029	661,9251	15,48029	661,9251
COVID	323,6966	140,1312	2,309954	0,024348	43,39253	604,0007	43,39253	604,0007

Table 18 Shower gel regression analysis -COVID 19

We find that $P - value$ for COVID is smaller than 0,05 which means that we can reject the null hypothesis and conclude that COVID-19 had a statistically significant effect on sales.

Also, since the coefficient of the COVID dummy variable is statistically significant, we can conclude that sales increased after the start of the pandemic compared to before the pandemic.

Chapter 4

4.1 Best model fit

In order to check whether COVID dummy is necessary in our model, we proceed with the process of comparing our full model with a model that does not include the COVID dummy.

The results that we got from this comparison were importantly close, which leads us to the conclusion that the model without the COVID dummy performs just as well. Consequently, since our aim is to have the simpler model with the highest accuracy, we reject the model which includes the covid dummy.

Thus, our model will include t , t^2 , y_{t-1} , y_{t-2} and month dummies. To evaluate our model and assess its reliability, we utilized the total dataset, incorporating all available sales data across the entire time period analyzed. This comprehensive approach ensures that our findings are robust and reflective of overall trends rather than being influenced by a limited subset of the data.

In order to check the reliability of our model for each category, we have created a graph which compares the real with the estimated values.

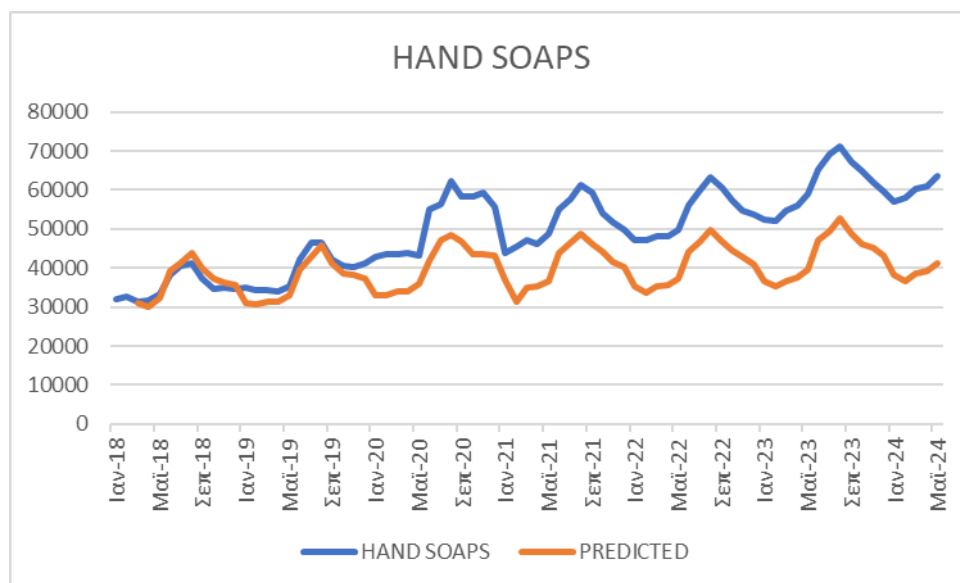


Figure 7 Hand soap- real vs estimated values

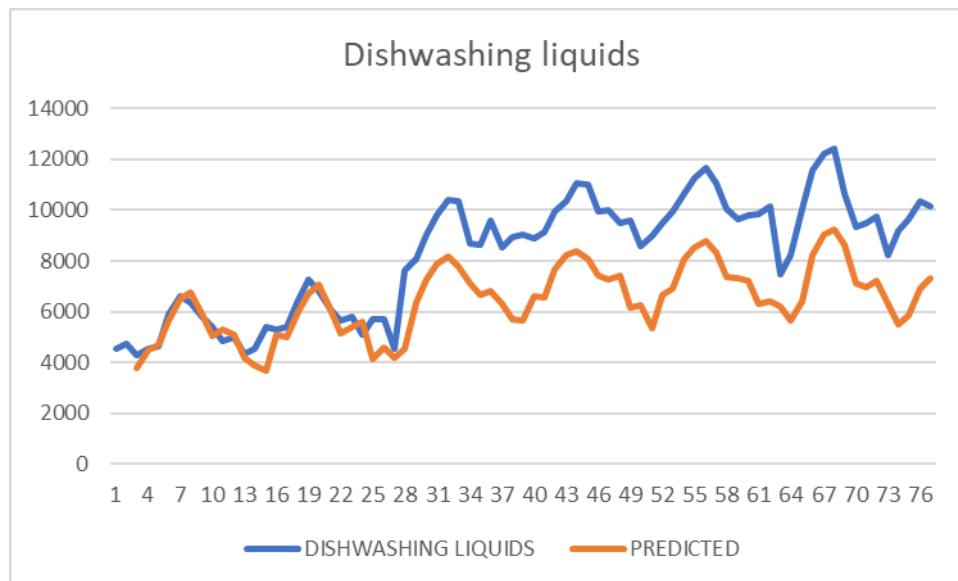


Figure 8 Dishwashing liquid- real vs estimated values

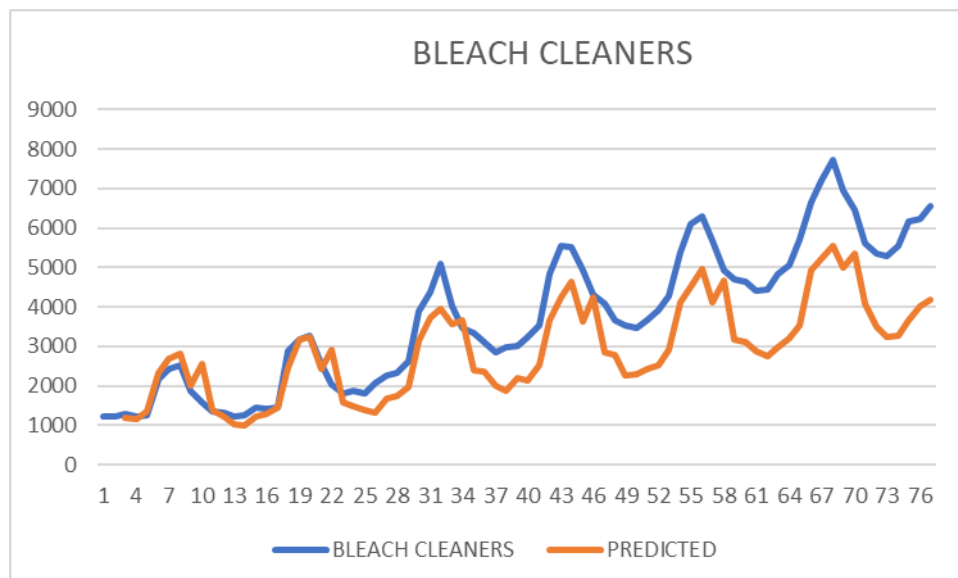


Figure 9 Bleach cleaner - real vs estimated values

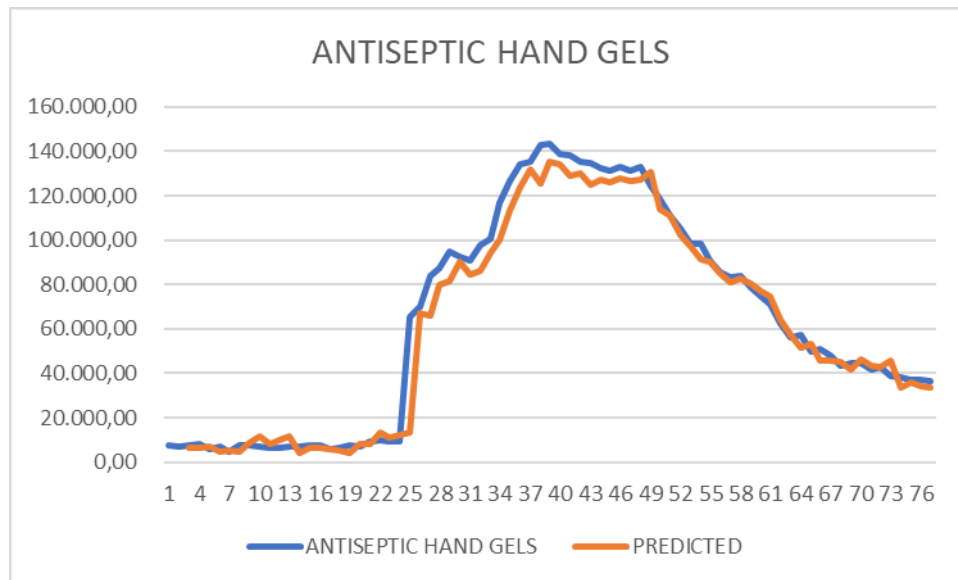


Figure 10 Antiseptic hand gel - real vs estimated values

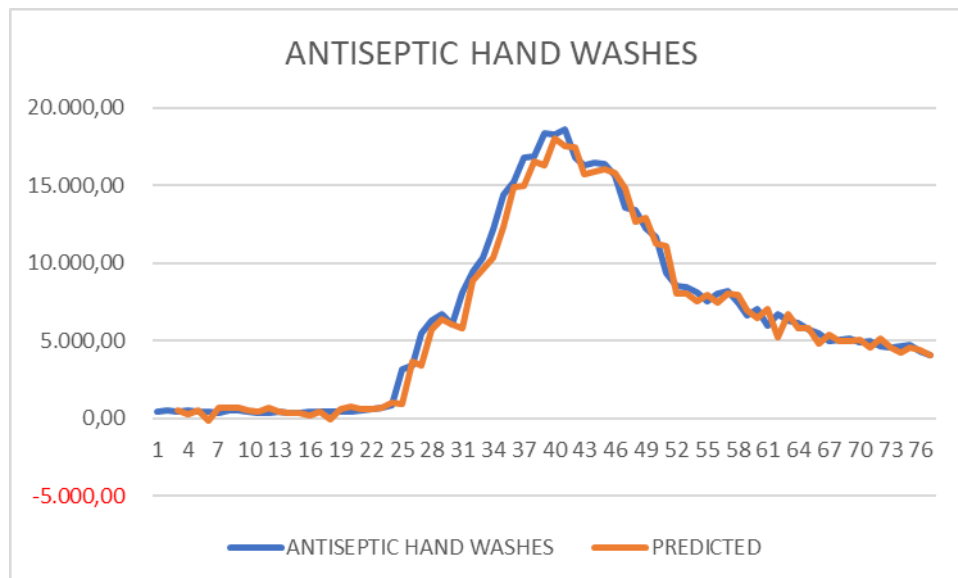


Figure 11 Antiseptic hand wash - real vs estimated values

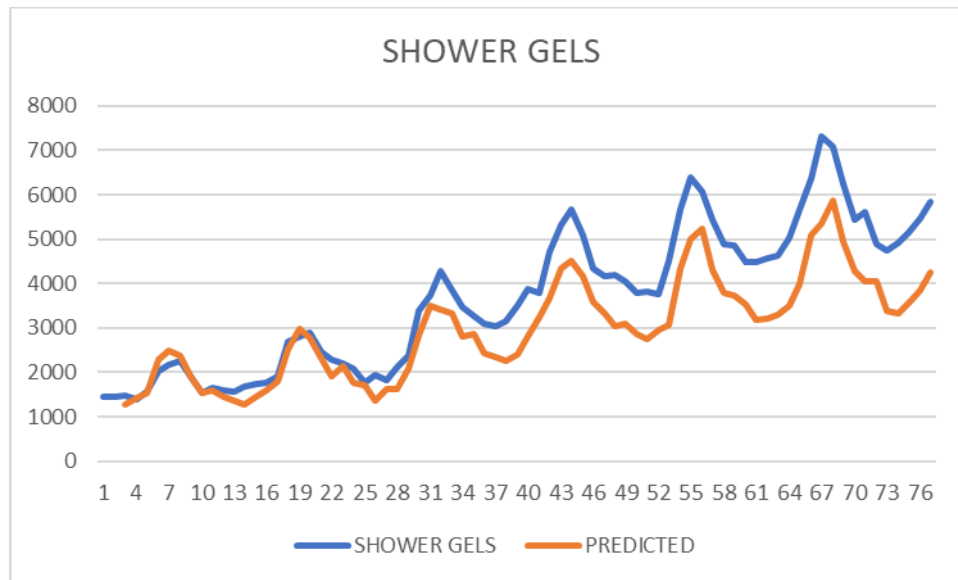


Figure 12 Shower gel- real vs estimated values

We can see that all models are successfully identifying seasonal patterns and general trends. We can conclude that they could be useful for forecasting broad sales patterns even if they don't predict exact monthly values perfectly.

The accuracy rate for monthly sales can vary widely depending on factors such as the product category, data quality, forecasting method, market volatility and seasonality. In order to measure forecast accuracy, we will calculate MAPE for each product category.

MAPE

HAND SOAPS: 19,53%

DISHWASHING LIQUIDS: 21,12%

BLEACH CLEANERS: 23,48%

ANTISEPTIC HAND GELS: 13,19%

ANTISEPTIC HAND WASHES: 18,19%

SHOWER GELS: 18,22%

In general, a MAPE between 10% and 20% indicates good accuracy.

Concerning personal care, we can see that MAPE values are below around 20%, reflecting good model accuracy. These categories may exhibit predictable seasonality or trends, making them easier to forecast.

As for Home Care products (Dishwashing Liquids, Bleach Cleaners) MAPE values are lightly higher (21–23%), suggesting greater variability in sales.

In order to further assess our model's reliability, we will also evaluate R^2 for each category.

R^2

HAND SOAPS: 0,95

DISHWASHING LIQUIDS: 0,96

BLEACH CLEANERS: 0,99

ANTISEPTIC HAND GELS: 0,98

ANTISEPTIC HAND WASHES: 0,99

SHOWER GELS: 0,98

All the categories analyzed show very high R^2 values (ranging from 0.95 to 0.99), meaning models are doing a great job at explaining the variation in sales. This suggests that our chosen independent variables (e.g., time, seasonality, etc.) are strong predictors for sales in these categories. With such high R^2 values, the models seem to accurately capture trends, seasonality, and other patterns in the data.

In order to evaluate whether our model is the best fit we will compare it with a simpler one for each category. Then we will calculate MAPEs and compare them with the ones we got from our model. The simpler model used will only include t , whereas our model includes t , t^2 , y_{t-1} , y_{t-2} and month dummies.

For the simpler model we crated we found the following MAPE values using in sample data:

MAPE (simpler model)

HAND SOAPS: 25,92%

DISHWASHING LIQUIDS: 32,60%

BLEACH CLEANERS: 54,21%

ANTISEPTIC HAND GELS: 156,71%

ANTISEPTIC HAND WASHES: 232,24%

SHOWER GELS: 49,01%

R^2 (simpler model)

HAND SOAPS: 0,73

DISHWASHING LIQUIDS: 0,67

BLEACH CLEANERS: 0,81

ANTISEPTIC HAND GELS: 0,15

ANTISEPTIC HAND WASHES: 0,15

SHOWER GELS: 0,81

From the above we conclude that both MAPE and R^2 show that our model is much better than the simpler one we used to compare it with. This is because the full model has significantly lower MAPE than the simple model. Thus, the additional variables improve the model's forecasting power and it suggests that seasonality (month dummies), lagged values, and non-linear trends (t^2) are important for predicting sales. Moreover, R^2 is lower in the simple model version which also validates our conclusion.

4.2 Out of sample data fit

Another aspect that will be analyzed in this paper is how good the models work on out of sample data. In order to do that we will sum up the sales values for each sub-category and create two main categories, home care and personal care.

For each category we will split the data into training and test datasets. More specifically, the period from January 2018 to December 2021 will be the training dataset and period from January 2022 to May 2024 will be the test dataset.

Starting with personal care we conduct the regression analysis for the training dataset and get the following equation:

$$y_t = 2773,92 + 588,90 \times t - 1,21 \times t^2 + 0,93 \times y_{t-1} - 0,14 \times y_{t-2} + 21807,18 \times M_1 + 3364,31 \times M_2 + 7855,37 \times M_3 - 1267,75 \times M_4 + 1272,92 \times M_5 + 5575,15 \times M_6 + 1272,76 \times M_7 + 4442,43 \times M_8 - 3014,23 \times M_9 + 1662,16 \times M_{10} + 232,99 \times M_{11}$$

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.992387							
R Square	0.984833							
Adjusted R	0.977004							
Standard E	11088.21							
Observations	48							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	16	2.475E+11	1.547E+10	125.8047	1.18E-23			
Residual	31	3.811E+09	122948474					
Total	47	2.513E+11						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept	2773.922	8323.7273	0.3332547	0.741186	-14202.4	19750.28	-14202.4	19750.28
t	588.9034	491.07245	1.199219	0.239529	-412.645	1590.452	-412.645	1590.452
t^2	-1.20917	11.579794	-0.1044204	0.917508	-24.8263	22.40798	-24.8263	22.40798
y(t-1)	0.925339	0.1863159	4.9665046	2.36E-05	0.545345	1.305332	0.545345	1.305332
y(t-2)	-0.13868	0.1852785	-0.7485158	0.459791	-0.51656	0.239194	-0.51656	0.239194
JAN	21807.18	7977.1421	2.7337086	0.010255	5537.693	38076.67	5537.693	38076.67
FEB	3364.309	8875.4988	0.3790558	0.70723	-14737.4	21466.01	-14737.4	21466.01
MAR	7855.368	8366.6147	0.9388945	0.355043	-9208.46	24919.19	-9208.46	24919.19
APR	-1267.75	8130.0455	-0.1559339	0.877096	-17849.1	15313.59	-17849.1	15313.59
MAY	1272.919	8084.3764	0.1574543	0.875908	-15215.3	17761.11	-15215.3	17761.11
JUN	5575.146	8006.5209	0.6963257	0.491413	-10754.3	21904.55	-10754.3	21904.55
JUL	1272.763	8011.2667	0.1588716	0.874801	-15066.3	17611.85	-15066.3	17611.85
AUG	4442.433	7957.6874	0.5582568	0.580679	-11787.4	20672.24	-11787.4	20672.24
SEP	-3014.23	7950.75	-0.3791126	0.707189	-19229.9	13201.43	-19229.9	13201.43
OCT	1662.159	7928.5646	0.2096418	0.835319	-14508.3	17832.57	-14508.3	17832.57
NOV	232.9931	7848.6792	0.0296856	0.976508	-15774.5	16240.48	-15774.5	16240.48

Table 19 Personal care - Regression analysis (training data set)

Using the regression equation obtained from the training data, we will calculate the predicted sales for each month in the test data. After that, we calculate MAPE which is found equal to 4,47%. A MAPE of 4.47% suggests that the regression model is a reliable tool for predicting future sales in your test dataset. Also, MAE is 6.026,07, which is also a good result keeping in mind the scale of the data.

Similarly, we do the same for home care category and we find the following equation:

$$y_t = 4889,27 + 11,85 \times t + 1,58 \times t^2 + 0,29 \times y_{t-1} - 0,12 \times y_{t-2} - 38,97 \times M_1 - 300,03 \times M_2 - 279,82 \times M_3 - 352,65 \times M_4 - 171,08 \times M_5 + 1467,40 \times M_6 + 2186,24 \times M_7 + 2632,00 \times M_8 + 1883,11 \times M_9 + 832,07 \times M_{10} + 392,39 \times M_{11}$$

SUMMARY OUTPUT									
Regression Statistics									
Multiple R	0.995288								
R Square	0.990599								
Adjusted R	0.985746								
Standard Error	405.7132								
Observations	48								
ANOVA									
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>				
Regression	16	5.38E+08	33603776	204.1502	7.4E-27				
Residual	31	5102700	164603.2						
Total	47	5.43E+08							
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>	
Intercept	4889.266	492.4121	9.929215	3.8E-11	3884.985	5893.547	3884.985	5893.547	
t	11.8499	20.31396	0.583338	0.563886	-29.5807	53.28049	-29.5807	53.28049	
t^2	1.577794	0.360368	4.378281	0.000126	0.842818	2.312771	0.842818	2.312771	
y(t-1)	0.28504	0.093235	3.057205	0.004572	0.094885	0.475195	0.094885	0.475195	
y(t-2)	-0.11553	0.059775	-1.93277	0.062445	-0.23744	0.006381	-0.23744	0.006381	
JAN	-38.9652	321.812	-0.121081	0.904409	-695.305	617.3747	-695.305	617.3747	
FEB	-300.032	307.6255	-0.975316	0.336953	-927.438	327.3743	-927.438	327.3743	
MAR	-279.822	293.0261	-0.954937	0.346997	-877.452	317.8092	-877.452	317.8092	
APR	-352.646	301.5182	-1.169568	0.251089	-967.596	262.3044	-967.596	262.3044	
MAY	-171.078	304.4844	-0.561862	0.57825	-792.078	449.9218	-792.078	449.9218	
JUN	1467.399	371.5624	3.949266	0.000421	709.5922	2225.205	709.5922	2225.205	
JUL	2186.24	425.3578	5.139767	1.44E-05	1318.717	3053.763	1318.717	3053.763	
AUG	2632.004	412.9352	6.373892	4.25E-07	1789.817	3474.191	1789.817	3474.191	
SEP	1883.109	351.6831	5.354561	7.77E-06	1165.846	2600.371	1165.846	2600.371	
OCT	832.0687	325.967	2.552617	0.015836	167.2547	1496.883	167.2547	1496.883	
NOV	392.3881	299.6334	1.309561	0.199968	-218.718	1003.494	-218.718	1003.494	

Table 20 Home care - Regression analysis (training data set)

In this case with get a MAPE of 32%. Also, MAE is equal to 4.979,38. This number is quite important compared to the scale of the data. This leads us to the conclusion that our model is not working as good as for personal care category with out of sample data.

4.3 Category interrelationship

In order to check if categories are interrelated, we will conduct hypotheses testing. As we want to check the interrelation between personal care and home care we will sum up the quantities of each group.

H_0 : There is no correlation between Personal Care and Home Care sales.

H_1 : There is a correlation between Personal Care and Home Care sales.

We will conduct a regression analysis with Y_{HC} as the dependent variable and Y_{PC} , as the explanatory variable. HC refers to total home care sales while PC represents total personal care sales. To ensure the most accurate model, we will incorporate time variables and month dummies, aligning with the best-fit model identified in the previous chapter. This approach allows us to assess the extent to which personal care sales influence home care sales while accounting for seasonal and temporal variations.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0,987796302							
R Square	0,975741534							
Adjusted R Square	0,970263816							
Standard Error	686,6047041							
Observations	77							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	14	1,18E+09	83974739	178,1291994	1,71863E-44			
Residual	62	29228413	471426					
Total	76	1,2E+09						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%
Intercept	4633,36587	384,2272	12,05892	6,60977E-18	3865,306703	5401,425038	3865,307	5401,425
total personal care sales	0,024202435	0,002575	9,39724	1,55785E-13	0,019054117	0,029350752	0,019054	0,029351
t	22,00201848	26,69053	0,824338	0,412907718	-31,35159996	75,35563691	-31,3516	75,35564
t^2	1,300390988	0,300168	4,332213	5,50149E-05	0,700364102	1,900417873	0,700364	1,900418
JAN	-664,7970496	384,7712	-1,72777	0,089007499	-1433,943755	104,3496555	-1433,94	104,3497
FEB	-548,0104623	384,2617	-1,42614	0,158844289	-1316,138578	220,1176537	-1316,14	220,1177
MAR	-810,3297808	384,0794	-2,1098	0,038919197	-1578,093579	-42,56598296	-1578,09	-42,566
APR	-88,3373918	383,6796	-0,23024	0,818665264	-855,3018958	678,6271122	-855,302	678,6271
MAY	473,833865	383,55	1,23539	0,221346242	-292,8717422	1240,539472	-292,872	1240,539
JUN	2508,674283	397,7802	6,306685	3,35242E-08	1713,522981	3303,825585	1713,523	3303,826
JUL	3488,855182	397,5019	8,776953	1,7956E-12	2694,260239	4283,450126	2694,26	4283,45
AUG	3783,885041	397,5114	9,518935	9,67667E-14	2989,271117	4578,498964	2989,271	4578,499
SEP	2390,316956	396,7749	6,024366	1,01242E-07	1597,175267	3183,458645	1597,175	3183,459
OCT	729,3405559	396,6075	1,838948	0,070712801	-63,46655445	1522,147666	-63,4666	1522,148
NOV	230,3957793	396,4409	0,58116	0,563239157	-562,078313	1022,869872	-562,078	1022,87

Table 21 Category interrelationship- Regression analysis

We can see that $P - value$ for Total Personal Care Sales is significantly small ($<0,05$), which suggests that Personal Care Sales and Hand Care Sales are related in some way. This relationship is critical to establishing that they follow similar trends over time.

Additionally, we observe that $P - value$ of t is higher than 0,05 whereas $P - value$ of t^2 is much lower than 0,05. Significant t^2 but insignificant t suggests a non-linear relationship, which may still indicate similar growth trends if both categories have similar curves. However, when conducting the graphs for the two categories we can see that they are importantly different.

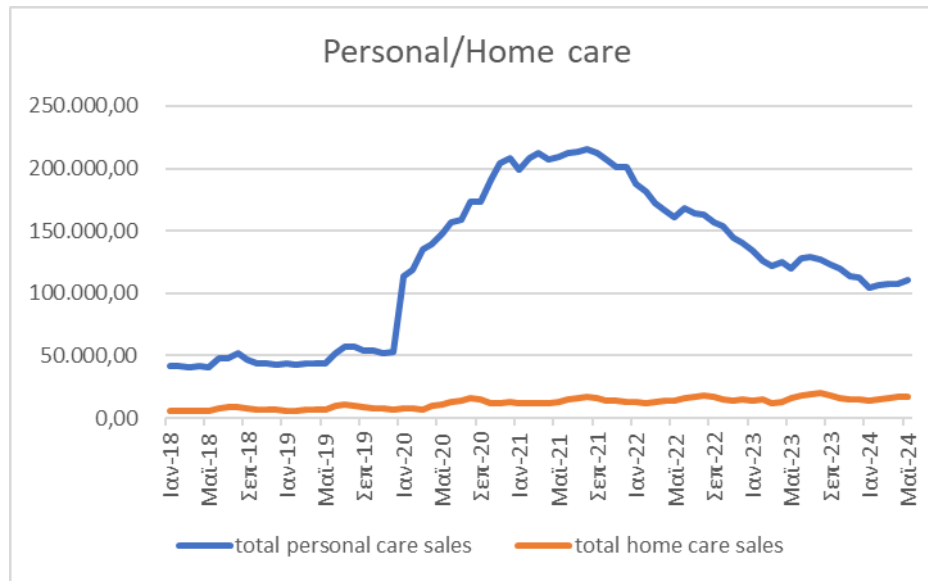


Figure 13 Personal care vs home care sales

Although the graphs for Personal Care Sales and Hand Care Sales appear to show distinct patterns, the regression analysis reveals a statistically significant relationship between the two categories. This outcome suggests that, despite the differences in their individual sales trends, there is a connection between the two over time. The regression model captures more subtle dynamics that might not be immediately visible in the raw data or the graphs.

In order to better understand the relationship between personal care and home care sales, we initially created a graph comparing the sales figures of both categories. However, the initial graph neglected the influence of past sales, as we did not account for the lagged variables. Lagged variables are crucial in time series analysis because they help capture the effect of past sales on current performance. Incorporating these lagged variables is necessary to provide a more accurate representation of the sales trends, as they reflect how previous sales influence future behavior. Thus, the revised model includes these lagged variables, allowing for a more complete analysis of the interrelationship between the sales of the two categories.

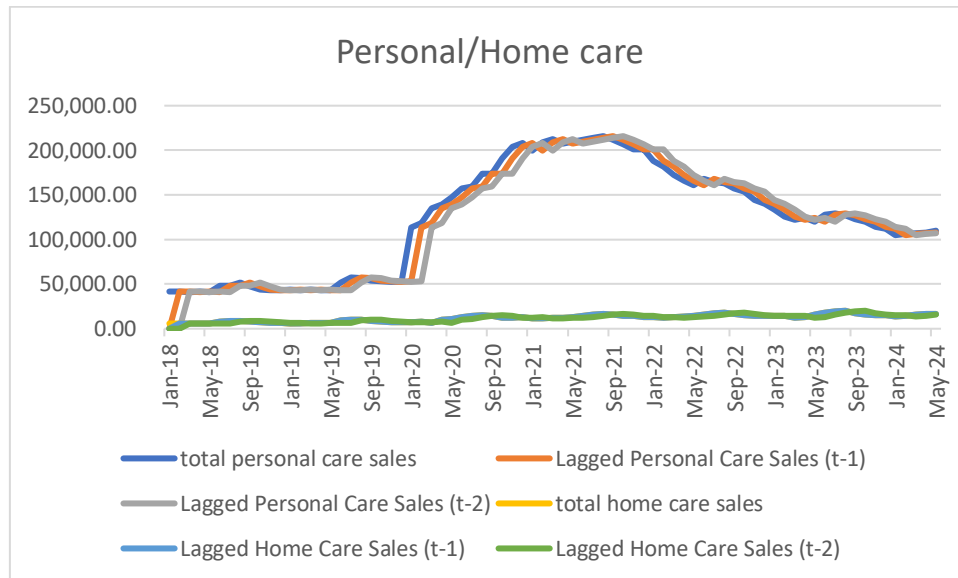


Figure 14 Personal care vs home care sales, lagged values

After incorporating the lagged variables, we found that for each category, the lines representing the lagged values were almost identical to the actual sales lines. This close resemblance suggests that the past sales data closely track with the current sales data, indicating strong autocorrelation. Essentially, the sales trends for both categories are heavily influenced by the previous periods' performance, showing a consistent pattern over time.

This outcome emphasizes the importance of including lagged variables in our model, as they help capture the temporal dynamics that drive sales. The similarity between the lagged values and actual sales lines further supports the notion that the sales in both categories follow a relatively consistent pattern, and that past sales are a good predictor of future sales in these categories. This finding also confirms that our model is adequately accounting for autocorrelation, making it more robust and reliable.

While the graphs show different trends, the model incorporates time variables and quadratic trends that allow for a more comprehensive understanding of the relationship between the categories. This analysis goes beyond the short-term fluctuations visible in the graphs and highlights how both categories may be influenced by common external factors, such as seasonal variations, economic conditions, or broader market trends.

Thus, despite the apparent visual differences, the sales of Personal Care and Hand Care products are interrelated in a broader context, with both categories reflecting similar influences over time. The regression model, by considering these time-based factors, reveals that the two categories follow somewhat aligned trends, even if their specific sales patterns differ in individual months or years.

Chapter 5

Conclusions

This study examined the sales patterns of personal care and home care products over multiple years, focusing on identifying seasonal trends, assessing the impact of the COVID-19 pandemic, evaluating forecasting accuracy, and exploring the interrelationship between different product categories. Through rigorous regression analysis, incorporating time trends, and seasonal effects, the research provided valuable insights into the dynamics of consumer demand in these markets.

The findings revealed clear seasonal trends in sales, with fluctuations occurring across different months and years. Certain products exhibited predictable peaks and declines, influenced by consumer purchasing habits, promotional cycles, and external factors such as public health policies and economic conditions. The inclusion of time-related variables and month-specific dummy variables in the regression models confirmed that seasonality plays a significant role in determining sales trends.

A key aspect of this research was assessing the effect of the COVID-19 pandemic on demand for products such as antiseptic hand gels, hand cleansers, and home care detergents. The results demonstrated that the pandemic significantly increased sales of these products, particularly during its early stages when hygiene awareness was at its highest. The introduction of a COVID dummy variable in the regression analysis confirmed a substantial surge in demand for antiseptic products, while home care detergents also saw a moderate increase. However, as the pandemic subsided, sales of these items gradually returned to pre-pandemic levels, suggesting that while short-term behavioral shifts were pronounced, long-term trends remained relatively stable.

Another important objective of this study was to evaluate the forecasting accuracy of sales predictions across different product categories. Using metrics such as Mean Absolute Percentage Error (MAPE) and R^2 , the analysis demonstrated that forecasting models performed well, particularly for personal care products, which exhibited lower error rates. In contrast, home care products had slightly higher forecasting errors, suggesting that additional

variables or market influences may need to be considered for improved accuracy. The inclusion of time trends, lagged sales values, and seasonal indicators significantly enhanced forecasting performance compared to simpler models, highlighting the importance of a structured approach in sales prediction.

Furthermore, this study investigated whether sales of different product categories were interrelated and followed similar trajectories over time. The results indicated a strong relationship between product categories, with statistically significant interaction terms suggesting that sales trends in one category could influence another. This finding supports the idea that consumer purchasing behavior exhibits both complementary and substitutive relationships across personal care and home care products. Additionally, time-series analysis confirmed that these categories often follow similar courses over extended periods, reinforcing the notion that they are interconnected within the broader consumer market.

In conclusion, this research highlights key sales dynamics in the personal and home care industries, demonstrating the significant influence of seasonal patterns, external shocks like COVID-19, and cross-category relationships. The insights gained from this study provide valuable implications for businesses aiming to enhance inventory management, demand forecasting, and strategic decision-making in an evolving market landscape. By understanding these factors, companies can better anticipate changes in consumer behavior and optimize their operations to meet future demand effectively.

BIBLIOGRAPHY

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2020). *Statistics for business and economics* (14th ed.). Cengage Learning.

Armstrong, J. S., Green, K. C., & Graefe, A. (2022). *Principles of forecasting: A handbook for researchers and practitioners*. Springer.

Baker, L., Patel, R., & Nguyen, T. (2020). Seasonality in consumer behavior: A model for retail forecasting. *Journal of Retail Analytics*, 14(2), 122-138.

Choi, T. M. (2021). Risk management in retail supply chains. *Journal of Retailing*, 97(1), 107-120. <https://doi.org/10.1016/j.jretai.2020.11.003>

Choi, T. M. (2021). Risk management in the retail supply chain: Impacts of COVID-19 and future trends. *Journal of Retailing*, 97(1), 107-120. <https://doi.org/10.1016/j.jretai.2020.11.003>

Chopra, S., & Meindl, P. (2021). *Supply chain management: Strategy, planning, and operation* (7th ed.). Pearson.

Donthu, N., & Gustafsson, A. (2020). Effects of COVID-19 on business and research. *Journal of Business Research*, 117, 284-289. <https://doi.org/10.1016/j.jbusres.2020.06.008>

Euromonitor International. (2022). *Home and personal care industry trends and outlook*. Euromonitor.

Fildes, R., Goodwin, P., & Lawrence, M. (2019). Effective forecasting in business decision making. *International Journal of Forecasting*, 35(2), 379-391. <https://doi.org/10.1016/j.ijforecast.2018.11.004>

Fildes, R., Goodwin, P., Lawrence, M., & Nikolopoulos, K. (2019). Effective forecasting and judgmental adjustments in business decision making. *International Journal of Forecasting*, 35(2), 379-391. <https://doi.org/10.1016/j.ijforecast.2018.11.004>

- Gao, X., Shi, Y., Guo, H., & Yang, W. (2022). The impact of COVID-19 on consumer buying behavior: Evidence from the personal care industry. *Journal of Consumer Behavior*, 21(3), 221-238. <https://doi.org/10.1002/cb.2022>
- Grewal, D., & Roggeveen, A. L. (2020). Understanding retail shifts in a post-COVID world. *Journal of Retailing*, 96(4), 517-520. <https://doi.org/10.1016/j.jretai.2020.10.002>
- Gujarati, D. N., & Porter, D. C. (2020). *Basic econometrics* (6th ed.). McGraw-Hill.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts.
- Ivanov, D., & Dolgui, A. (2020). Viability of supply chain models: The effect of COVID-19 on resilience and risk management. *International Journal of Production Research*, 58(10), 2904-2915. <https://doi.org/10.1080/00207543.2020.1750727>
- Keane, M. P., & Neal, T. (2021). Consumer panic in the COVID-19 pandemic. *Journal of Economic Behavior & Organization*, 183, 586-607. <https://doi.org/10.1016/j.jebo.2020.09.020>
- Koçak, Ö., Kağrıcioğlu, R., & Sari, M. (2021). Consumer behavior and home care product sales during the pandemic. *Journal of Retail & Distribution Management*, 49(6), 815-832. <https://doi.org/10.1108/JRDM-2021-0123>
- Kotler, P., & Keller, K. L. (2016). *Marketing management* (15th ed.). Pearson Education.
- Lee, J., Kim, Y., & Lee, H. (2019). Accounting for seasonal patterns in time series data: A comprehensive review. *Journal of Business and Economic Statistics*, 37(4), 1-16.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLoS ONE*, 13(3), e0194889. <https://doi.org/10.1371/journal.pone.0194889>
- McKinsey & Company. (2021). *Consumer behavior and COVID-19: How the pandemic shaped the market*. McKinsey Insights.

- Mentzer, J. T., & Moon, M. A. (2004). *Sales forecasting management: A demand management approach*. SAGE Publications.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2012). *Introduction to time series analysis and forecasting* (2nd ed.). Wiley.
- Pantano, E., Pizzi, G., Scarpi, D., & Dennis, C. (2020). Competing during a pandemic? Retailers' ups and downs during the COVID-19 outbreak. *Journal of Business Research*, 116, 209-213. <https://doi.org/10.1016/j.jbusres.2020.05.036>
- Sheth, J. (2020). Impact of COVID-19 on consumer behavior: Will the old habits return or die? *Journal of Business Research*, 117, 280-283. <https://doi.org/10.1016/j.jbusres.2020.05.059>
- Wang, E., An, N., Gao, Z., Kiprop, E., & Geng, X. (2021). Consumer food stockpiling behavior and willingness to pay for food reserves in COVID-19. *Food Security*, 13(1), 25–34. <https://doi.org/10.1007/s12571-020-01173-5>
- Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77-84. <https://doi.org/10.1111/jbl.12010>
- Wang, Y., Hao, H., & Platt, L. S. (2020). Examining consumer behavior during COVID-19: A shift toward online shopping and contactless services. *Technology in Society*, 63, 101-104. <https://doi.org/10.1016/j.techsoc.2020.101104>
- Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7th ed.). Cengage Learning.