School of Social Sciences

Supply Chain Management

Postgraduate Dissertation

ANALYSING SALES PATTERNS AND TRENDS IN THE FOOD AND HOUSEHOLD WHOLESALE SECTOR

The Case of ACHTIDA S.A.

Dialekti P. Kappa

Supervisor: Nikolaos S. Thomaidis

Patras, Greece, June 2023

ANALYSING SALES PATTERNS AND TRENDS IN THE FOOD AND HOUSEHOLD WHOLESALE SECTOR

The Case of ACHTIDA S.A.

Dialekti P. Kappa

Supervising Committee

Supervisor:	Co-Supervisor:	Co-Supervisor:
Nikolaos Thomaidis	Ioannis Giannikos	Michael Geranios
Tutor	Tutor	Tutor

Patras, Greece, June 2023

Special thanks
to Sokratis Kouvrakis
for his time and his trust

Abstract

The aim of the proposed dissertation is to analyze sales evolution, as well as explore the distinct characteristics of several product categories classified under the food and household umbrella. The analysis is based on the examination of ACHTIDA S.A. sales, a company active for over thirty years in the wholesaling of food and snack products, as well as of cosmetics and household commodities.

Through the use of several statistical tools the study aims to address the following issues:

- Explore the statistical properties of each product category's sales
- Detect the possible correlations or differences among each category's sales evolution over time
- Examine the degree in which seasonality or other external factors determine the sales volume of each category
- Explore the COVID-19 outbreak effect on sales, as well as the reaction of retailers and final customers to major incidents throughout the pandemic era, such as WHO announcements, lockdown imposition, EU's vaccination approval, or else.

The study is based on timeseries examination and regression analysis, taking into account the direct relationship among sales and operational aspects such as inventory handling, replenishing frequency, and procurement practices in general. The central objective of the study is sales interpretation, with respect to patterns and trends, as well as the effects of forces external to the company's' control, that can provide the basis to identify major sales determinants, take advantage of downstream demand information, minimize lost sales by avoiding out-of-stock conditions and adjust effectively to supply chain disruptions, as well as to create a foundation of knowledge, sufficient to formulate the basic guidelines for adequate and accurate forecasting models.

Keywords

Wholesale, Pattern Recognition, Timeseries, Consumer Products, Regression, Seasonality, COVID-19 effect.

Ανάλυση Προτύπων και Τάσεων Πωλήσεων στον Κλάδο Χονδρικού Εμπορίου Τροφίμων και Ειδών Οικιακής Χρήσης

Η περίπτωση της ΑΧΤΙΔΑ ΑΕΒΕ

Διαλεκτή Π. Κάππα

Περίληψη

Ο στόχος της παρούσας μεταπτυχιακής εργασίας είναι η ανάλυση της εξέλιξης των πωλήσεων, καθώς και η διερεύνηση εξειδικευμένων χαρακτηριστικών διαφόρων προϊοντικών κατηγοριών που ταξινομούνται ως τρόφιμα και είδη οικιακής χρήσεως. Η ανάλυση βασίζεται στην εξέταση των πωλήσεων της ΑΧΤΙΔΑ Α.Ε.Β.Ε., μιας επιχείρησης που ενεργεί για πάνω από τριάντα έτη στο χονδρικό εμπόριο τροφίμων και σνακ, καλλυντικών και προϊόντων οικιακής χρήσης.

Χρησιμοποιώντας πλήθος στατιστικών εργαλείων, η εργασία στοχεύει στην κάλυψη των ακόλουθων ζητημάτων:

- Διερεύνηση των στατιστικών ιδιοτήτων κάθε προϊοντικής κατηγορίας
- Ανάδειξη πιθανών συσχετισμών ή διαφοροποιήσεων μεταξύ των προϊοντικών κατηγοριών όσο αφορά στην διαχρονική τους εξέλιξη
- Εξέταση του βαθμού κατά τον οποίο η εποχικότητα ή άλλοι εξωγενείς παράγοντες καθορίζουν την εξέλιξη των πωλήσεων
- Διερεύνηση της επίδρασης της πανδημίας στις πωλήσεις, καθώς και των αντιδράσεων λιανεμπόρων και τελικών καταναλωτών σε σημαντικά γεγονότα της εν λόγω περιόδου, όπως οι ανακοινώσεις του Παγκοσμίου Οργανισμού Υγείας, η επιβολή περιοριστικών μέτρων, η έγκριση εμβολίων από την Ευρωπαϊκή Ένωση, ή άλλων.

Η μελέτη βασίζεται στην ανάλυση χρονοσειρών και την ανάλυση παλινδρόμησης, λαμβάνοντας υπόψη την άμεση συσχέτιση των πωλήσεων με επιχειρησιακούς παράγοντες,

όπως η διαχείριση αποθεμάτων, η συχνότητα αναπλήρωσής τους και οι γενικότερες πρακτικές προμηθειών. Κύριο στόχο αποτελεί η ερμηνεία της εξέλιξης των πωλήσεων, με αναφορά σε πρότυπα και τάσεις, καθώς και σε επιδράσεις εξωγενείς ως προς τον επιχειρησιακό έλεγχο, που θα μπορέσουν να δημιουργήσουν την βάση ανάδειξης κύριων καθοριστικών παραγόντων για τις πωλήσεις, οδηγώντας στην εκμετάλλευση των δεδομένων ζήτησης, στον περιορισμό του κόστους που προκύπτει από απώλεια πωλήσεων και στην προσαρμογή σε καταστάσεις διάρρηξης της εφοδιαστικής αλυσίδας. Επίσης, στην δημιουργία γνωστικού υποβάθρου, ικανού να αποτελέσει την βάση σχεδιασμού κατάλληλων, χαρακτιριζόμενων από ακρίβεια, μοντέλων πρόβλεψης.

Λέξεις – Κλειδιά

Χονδρικό Εμπόριο, Αναγνώριση Προτύπων, Χρονοσειρές, Καταναλωτικά Προϊόντα, Παλινδρόμηση, Εποχικότητα, COVID-19.

Table of Contents

List of Figures

List of Tables

List of Abbreviations & Acronyms

EU	European Union
WHO	World Health Organization
ELSTAT	Hellenic Statistical Authority

1 INTRODUCTION

The aim of the proposed dissertation is to analyze sales evolution, as well as explore the distinct characteristics of several product categories classified under the food and household umbrella. Subject selection was triggered by the authors' professional involvement in wholesaling and retailing of the respective product ranges, which was further stimulated by the academic background that supply chain studies cultivated.

The analysis is based on the examination of ACHTIDA S.A. sales, a company active for over thirty years in the wholesaling of food and snack products, as well as of cosmetics and household commodities. The company was founded in 1988 in Moudania, Chalkidiki from Sokratis Kouvrakis and its activity was built around consumer product trading as well as logistics services. Additionally, business cycle incorporates a fraction of manufacturing activities relevant with coffee processing and packaging. Its distribution range expands in three prefectures of Northern Greece, distinct for their differentiated market attributes. Achtida's customer profile varies from supermarkets, local retail shops and hospitality units in Thessaloniki, Chalkidiki and Pieria. Business activity is exclusively coordinated from the central warehouse in Moudania, from which either the salesforce, or the distribution fleet (vans and trucks) serves wholesale customers in multiple locations.

ACHTIDA is the official distributor for several product lines, commercialized in the Greek market by outstanding, global companies such as Essity, Colgate Palmolive, Johnson & Johnson, Pepsico, Sarandis, Jacobs – Dowe Egberts, Mondelez and many others. Product range incorporates food, snacks, detergents, personal hygiene products, household commodities as well as cosmetics.

The study is based on timeseries examination and regression analysis, taking into account the direct relationship among sales and operational aspects such as inventory handling, replenishing frequency, and procurement practices in general. The central objective of the study is sales interpretation, with respect to patterns and trends, as well as the effects of forces external to the company's' control, that can provide the basis to

- Identify major sales determinants and exploit them in order to specify best practices concerning inventory levels, capacity and replenishing schemes
- Take advantage of downstream demand information (retailers – consumers), utilize it to enhance operational practice and diffuse it to upstream supply chain members, aiming to improve agility, adaptability and efficiency.
- Minimize lost sales by avoiding out-of-stock conditions and adjust effectively to supply chain disruptions

At last, but not least, although the dissertation does not aim to provide a forecasting model, it is anticipated to create a foundation of knowledge, sufficient to formulate the basic guidelines for adequate and accurate forecasting models.

1.1 Literature Review

The majority of Achtida's customers are retailers of various sizes, from mini markets to hypermarkets. Throughout literature, analyzing and foreseeing the demand of retail industry is said to lead to several benefits in terms of supply chain operations and control. Indicatively, Bottani, et al., suggest taking advantage of retailers' demand forecasting, which combined with pricing policies could lead to stockout reduction. Results of the study demonstrated the ability to limit stockout losses more than 50 percent. Demand prediction can also lead to various benefits for the wholesaler not only in direct economic turns relevant with lost sales, but also in terms of supply chain relationships improvement as well as customer (retailer) satisfaction (Bottani, et al., 2019). As observed in other papers, forecasting the demand of retail industry is a crucial determinant of agility throughout the supply chain. Information acquiring and data processing capabilities, combined in accurate forecasting practices integrated into procurement operations can enhance echelon versatility and the overall supply chain adaptability (Bai, 2022).

According to studies applied to other sectors, such as fuel industry, timeseries analysis and regression applications can assist procurement improvement (Salaverry & White III, 2009), while other researchers argue that different products require different forecasting techniques in order to achieve competitiveness and profitability (Stanton

& Baglione, 2021). Particularly, supply chain members should involve in collaborative planning processes by sharing information and combining supply and demand signals upstream and downstream. Regression applications were also made in studies covering sales forecasting in food retail industry (Sivanandam Arunraj & Ahrens, 2015).

In manufacturing contexts, the value of downstream demand visibility is analyzed by Lehtonen, Småros, & Holmström, highlighting the need for harmonization among production, purchasing and consumer demand. In particular, companies are urged to invest in acquiring access to downstream demand information, in order to enhance operational competence (Lehtonen, Småros, & Holmström, 2005). Wholesaling of consumer products, as being a level above retail stores, can admittedly be supported by forecasting applications especially in terms of deploying pricing policies and inventory strategies (Wang, 2022).

According to Hançerlioğulları, Sen, & Aktunç (2016), demand volatility is directly related to inventory turnover performance of retailing companies. Specifically, demand uncertainty is highlighted as a major variability generator in any supply chain, not to mention retail, in which supply fluctuations can easily result to out-of-stock instances which in turn result to unsatisfied customers (Hançerlioğulları, Sen, & Aktunç, 2016). Sweeney, Riley, & Duan, (2022) argue that product variety as well as demand volatility affect inventory and replenishing rates, suggesting that the higher the product variety in grocery retail environments, the better the inventory performance (Bai, 2022). Additionally, stockouts seem to impose direct effects on demand, mainly among substitute products (Kim, Kim, & Lu, 2019).

As ACHTIDA commercializes an appreciable volume of seasonal commodities, literature review also stressed on the diversity of practices in selling seasonal products, when it comes to product assortment (types, number of items and stockkeeping levels) and forecasting choices (Sodero, Namin, Gauri, & Bhaskaran, 2021). Another factor that has to be mentioned and affects Achtida's sales directly is the effect that promotional activities impose on demand formation. Association exists among timeseries attributes and demand variability, with the latter having major impacts on sales forecasting (Abolghasemi, Beh, Tarr, & Gerlach, 2022).

As far as supply chain disruptions are concerned, Stanton & Baglione, (2021) outline that COVID-19 pandemic effect on supply chains was a sheer proof of how intense demand variability impacts can be, as it altered not only the mix of products purchased but also the point of sales itself. Grocery shops, supermarkets and hypermarkets faced differentiated conditions and were forced to modify procurement priorities, inventory levels and pricing strategies. Additionally, online consumer product sales were amplified, further diversifying the commercial framework of consumer commodities (Stanton & Baglione, 2021).

1.2 Methodology & Limitations

Through the use of several statistical tools the study aims to address, indicatively but not exhaustively, the following issues:

- Explore the statistical properties of each product category's sales
- Detect the possible correlations or differences among each category's sales evolution over time
- Examine the degree in which seasonality or other external factors determine the sales volume of each category
- Explore the COVID-19 outbreak effect on sales, as well as the reaction of retailers and final customers to major incidents throughout the pandemic era, such as WHO announcements, lockdown imposition, EU's vaccination approval, or else.

The analysis expands in four product groups:

1. Cosmetics
2. Commodities with reference to personal hygiene and cleanness
3. Household commodities, detergents and paper products
4. Snacks,

all of them further analyzed in distinct subgroups.

During the design phase, two more product groups were analyzed in order to be incorporated in the study, food products and beverages. However, due to methodological issues such as timeseries discontinuity and fragmentary observations concerning specific but significant product lines, their induction was abolished. Product group formation was decided after scrutinizing firstly their fixed characteristics (use, usefulness) and secondly, after determining, together with company's administration, combination schemes that would render greater commercial significance to groups and subgroups. Since the primary objective of the analysis is to provide ACHTIDA with tools that would assist sales evolution interpretation, the composition of product categories becomes a prominent determinant towards the effectiveness of the study.

The study can be categorized as a quantitative application on business data. The data employed were generated from primary sources, namely Achtida's net sales reports on value per item, on a monthly frequency. The timeseries frame expands from January 2012 up to December 2022, including twelve reports per year. Microsoft Excel 2016 was utilized for the group formation procedure, which, through filter use generated aggregations that led to product, subgroup and ultimately aggregated monthly group sales. The terms on which filter application was executed were based on the company's coding scheme, which allocates individual items in specific coding categories, from bottom (closer to SKU level) to top: statistical category, subgroup, group, supplier. The major determinant of the analysis group formation was Achtida's subgroups, accompanied with additional statistical category filters were needed.

Clarifications should be provided concerning the sales values that constitute group and subgroup aggregated sales. For the vast majority of items that ACHTIDA S.A. trades, distinct return and refund policies exist. In the majority of sales categories, exempt from instances that are clearly mentioned, the value of expired or perished items is subtracted from net sales, in order to fully comply with actual revenue per item. On the other hand, refunds not relevant to item condition or expiration date, as well as refunds that cannot be directly connected to item units, such as discount coupons or monetary provisions towards commercial objectives (e.g., promotional activities) are not incorporated to the presented sales values. Appendix I provides a full list of monetary values not included to the study.

Accordingly, the study does not incorporate sales data that were fragmented or of a partial presence throughout the analyzed period, namely 2012 – 2022. That choice was made upon the aim to obtain consistent timeseries, discharged of excerpts that could distort sales evolution illustration as well as hinder regression analysis. For instance, items relevant with sales agreements or distribution contracts active only during parts of the period under study and not through its full length were excluded. Furthermore, in specific instances subgroup coding necessitated item or product line exclusion, as there were cases of subgroup codes that referred to different items during the timeseries frame.

The analysis for every product group is structured in four distinct parts. At first, an introduction of the product group and a presentation of its components, namely subgroups. Data relevant with brands incorporated in each category as well as the producers or distributors of them are displayed, either to assist the curious reader in comprehending product mix, or provide the company with hints that can be utilized towards commercial agreement assessment. At the second part, exploration of data is accomplished through visual observation of graphs and employment of descriptive statistics, central tendency and variation measures. Seasonality is examined to a certain extent, as already in the early stages of data processing its implications became distinguishable.

The third part concentrates on regression analysis, employing variables that either became evident during the second stage analysis, or were bound to be examined as part of the initial research questions. Indeed, in most of the cases seasonal factors, forces generated by COVID-19 pandemic as well as other, macroeconomic variables are examined, leading to differentiated results. Finally, at the last part of the research structure, concluding remarks are presented, summarizing the main features obtained in every chapter.

1.3 Findings

Study findings are notably diversified among product groups, as either seasonal, pandemic, or macroeconomic variables seem to impose different effects on each groups' sales. Yet, there are cases in which regression analysis highlighted significant

seasonal forces, affecting sales evolution in a considerable degree. Covid-19 effect was proved to have also differentiated impacts among product categories, as its repercussions appear to be extremely dependent on product characteristics.

2 GROUP 1 – COSMETICS

2.1 Group Introduction

The first group of products contains commodities relevant to face, body and hair cosmetic treatment. Four distinct product categories (subgroups) compose an aggregated revenue that moves along a range of 213.169€, with minimum value of sales 2.847€ and maximum value 216.017€ during the eleven-year period under study.

2.1.1. FRAGRANCES & GIFT SETS (101)

The first subgroup consists of eau de toilette (EDT) fragrances, eau de parfum (EDP) as well as associated multipacks containing either EDT or EDP together with deodorant, shower gel, body lotion, after shave or other products. The multipacks are designed as gift sets and the major brand names that constitute the category are BU, STR8, TESORI, Antonio Banderas, C-THRU, Elode, Denim, Adidas, Playboy, all of them distributed in the Greek market by Sarandis S.A.

The average contribution of 101 subgroup to the formation of Group 1 is 9.96% over the eleven-year period.

2.1.2. FACE & BODY COSMETICS (102)

The second subgroup is formatted from the aggregation of sales of commodities suitable for face cleaning (face wipes, de-makeup lotions, scrubs etc.) face and body hydration (moisturizers, body oils, face creams) as well as hair removal products. The main distributors are Johnson & Johnson, marketing products under the same brand name, as well as Sarandis S.A. with brands such as Bioten, Strep, Le Petite Marsellais. Subgroup 102 contributed by 14% on average to Group 1 sales.

2.1.3. SUN PROTECTION (103)

The sun protection category consists of sunscreen and tanning products, milk, oil, after sun and protective face creams. The brands that compose the subgroup are Carroten and Piz Buin, distributed by Sarandis S.A. and Johnson & Johnson respectively. For that particular product category, two facts should be underlined:

- Firstly, the distinct trading practices of sunscreen products. At the beginning of summer season, fully composed stands with specific range of products are placed to retail points (super markets, mini markets etc). At the end of the season, retailers return the unsold items to the distributor, being fully compensated on the respective value. Given that fact, it should be noted that this study analyzes only the sales value of sun protection, by completely excluding return value, as the value of the sold items reflects more appropriately the trends present in every season. Additionally, negative values of sales through the autumn season would result to a misleading timeseries.
- Secondly, through the period 2012 - 2022, Achtida S.A. has periodically commercialized other brands such as Hawaian Tropic (from 2014 to 2016), Coppertone and Kolastyna. However, because of the fragmented presence of those brands throughout the period under study, the respective items were excluded from the timeseries, aiming to provide a solid reflection of subgroup evolution by keeping out occasional sales. Moreover, the purposefulness of the study would become questionable for brands no longer traded by Achtida S.A.

Sun protection products account for nearly 75% on average of Group 1 total sales.

2.1.4. HAIR PRODUCTS (104)

Hair products subgroup occupies only 1% on average of period sales, consisting by hairsprays and dry shampoos. Again, as in 103 subgroups' case, a vast number of products has been excluded from the study (Wella and Wella New Wave hair products, Koleston hair dyes) due to their fragmented presence through the eleven-year period.

2.2 Data Exploration

In the next page, Figure 2-1 depicts the aggregated Group 1 sales evolution over time, along with distinct subgroup representation.

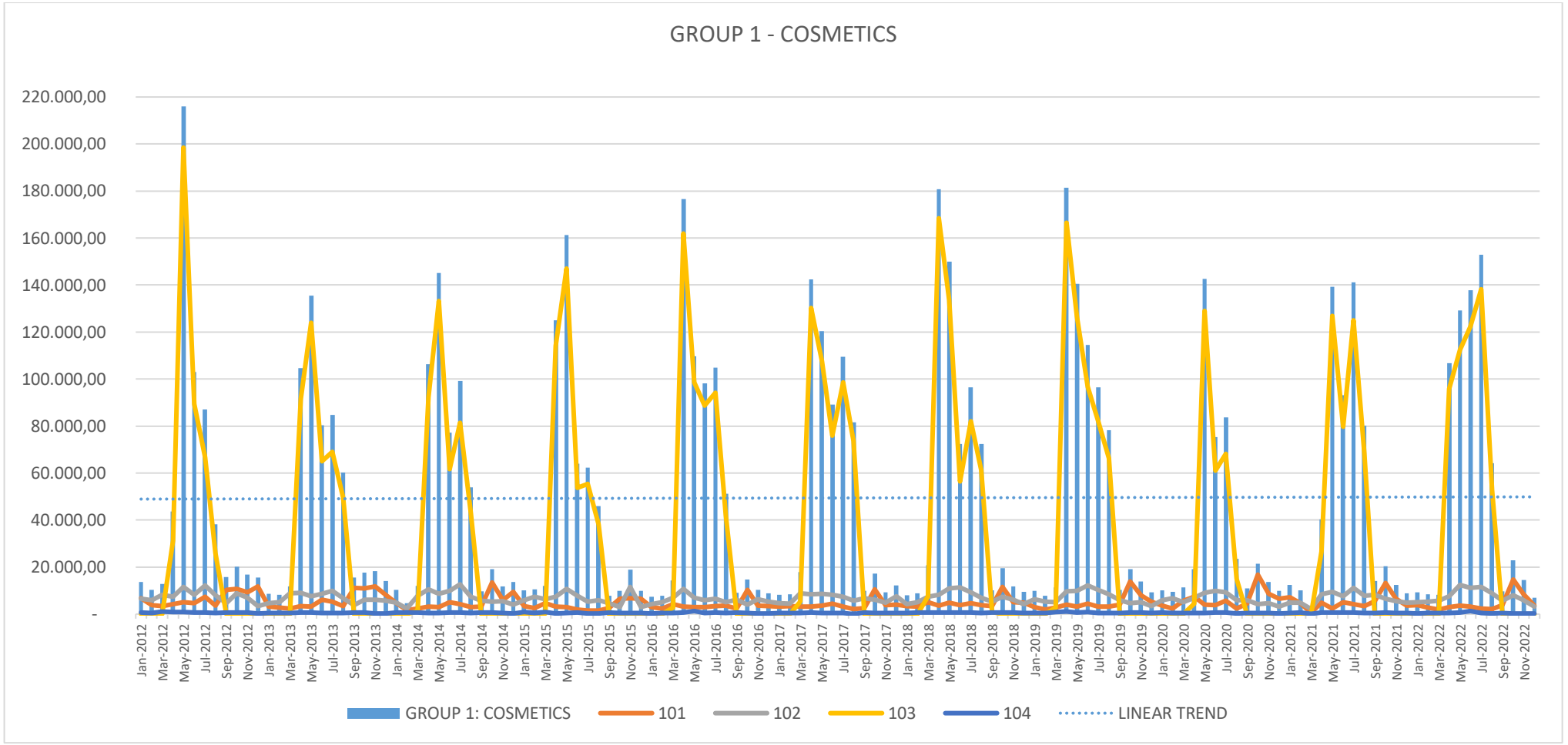

Figure 2-1: Group 1 sales evolution (2012-2022)

2.2.1. VISUAL OBSERVATIONS & DESCRIPTIVE FEATURES

Figure 2-1 visual examination leads to several presumptions. First of all, the dominance of 103 (Sun Protection) subgroup becomes apparent, as its evolution nearly coincides with the aggregated sales of Group 1. As presented in Table 2-1, 103 subgroup sales extend to a range of 198.477€ whereas Group 1 range is marginally higher, in 213.169€. The minimum sales value of Sun Protection is zero, a fact expectable during the winter season, with the higher being 198.479€, accomplished on May 2012. By examining the subgroup while taking into account off-season absence of sales, the median is also zero, as half of the monthly observations contain no sales value. Under the same logic, the mean value of sales is 37.034€. However, by considering only the observations with positive sales value, 103 subgroup median escalates to 73.671€ and the mean on 75.208€. In the first case (on and off-season sales) the standard deviation from the mean is 51.225€ ant in the second case (calculated only on positive sales) 49.549€.

The linear trend of the aggregated group sales is nearly stable from 2012 to 2022, however incorporating an explicitly distinct seasonal pattern. Indeed, owing mainly to subgroup 103 contribution, sales spike at the beginning of summer season in most of the cases, keep a high volume throughout the summer and finally decline at the beginning of autumn.

	GROUP 1	**101**	**102**	**103**	**104**
MINIMUM	2.847,04	592,82	2.131,79	-	122,43
MAXIMUM	216.016,53	16.786,12	12.666,57	198.476,79	1.301,85
RANGE	213.169,49	16.193,30	10.534,78	198.476,79	1.179,42
MEAN	49.408,89	4.921,10	6.918,35	***37.034,25 / 75.208,01***	535,20
MEDIAN	17.780,12	3.724,08	6.498,97	***- / 73.670,55***	529,16
STANDARD DEVIATION	52.164,69	3.059,52	2.377,12	51.224,70	195,61

Table 2-1: Group 1 - Basic Descriptive Features

As far as the rest of subgroups are concerned, 101 (Fragrances & Gift Sets) performs a marginally downward trend as well, which incorporates seasonal spikes before Christmas season, mostly on October or November every year. Maximum sales value is at 16.786€ in October 2020, while minimum value is located on March 2021. Subgroup 102 (Face & Body Cosmetics) follows a stable to marginally declining trend.

Seasonal effects are present as well, with peaks usually around July, fact attributed to the vast number of tourists visiting Chalkidiki during that period (Kouvrakis, 2023). Subgroup sales range from 2.132€ to 12.667€ with mean and median on 6.918€ and 6.499 respectively. The timeseries behavior resembles schematically that of a stationary timeseries (Thomaidis N. S., Principles of Quantitative Forecasting, 2021), with a mean on 6.918€. Finally, subgroup 104 in its minor presence, acquires a downward trend, performing in a narrow range of sales (1.179€) around a mean of 535€.

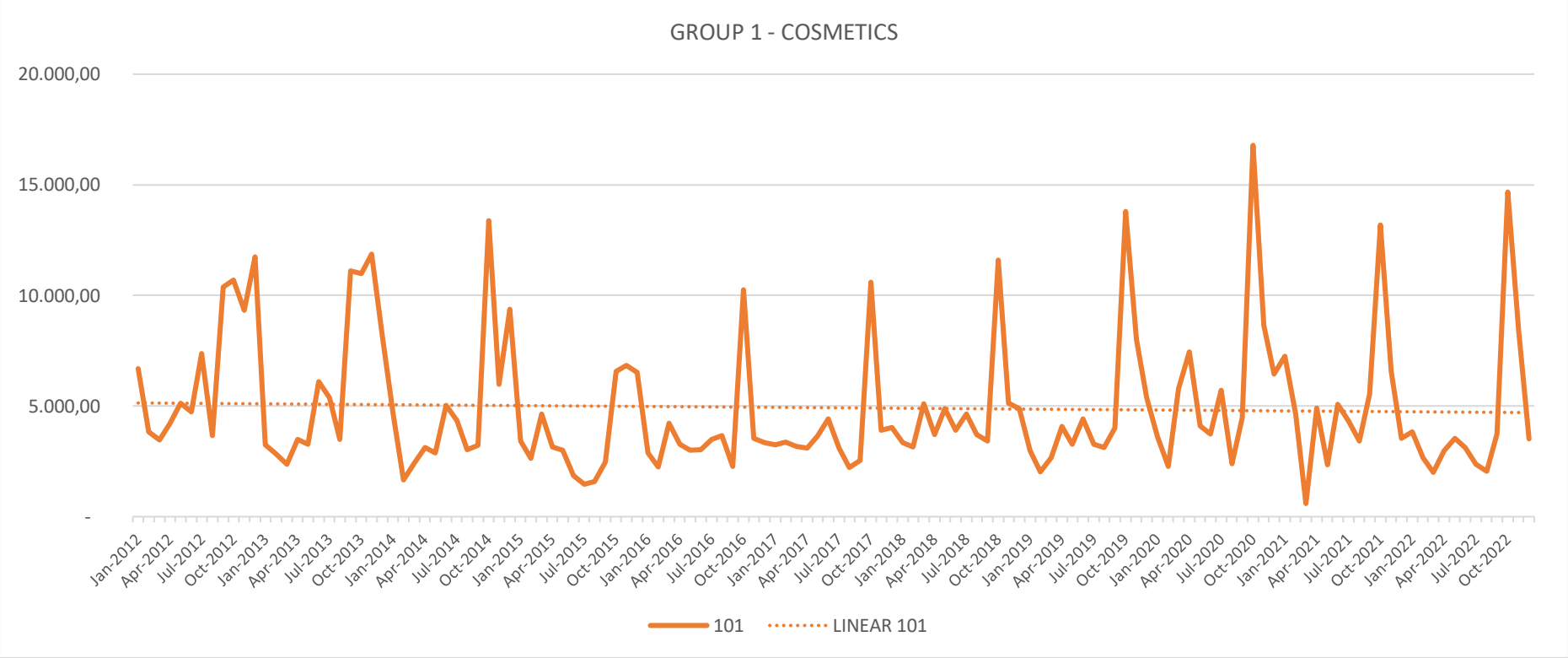

Figure 2-2: Subgroup 101 sales evolution (2012-2022)

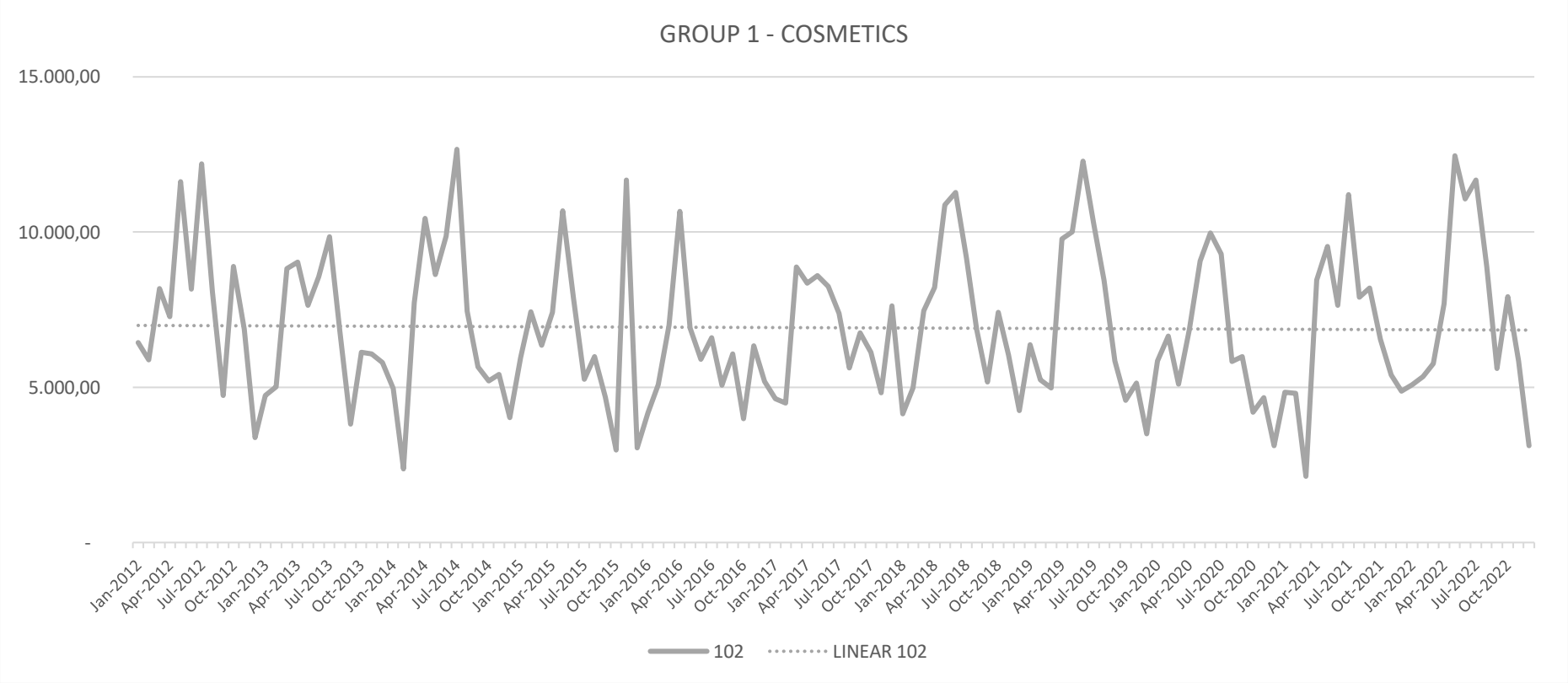

Figure 2-3: Subgroup 102 sales evolution (2012-2022)

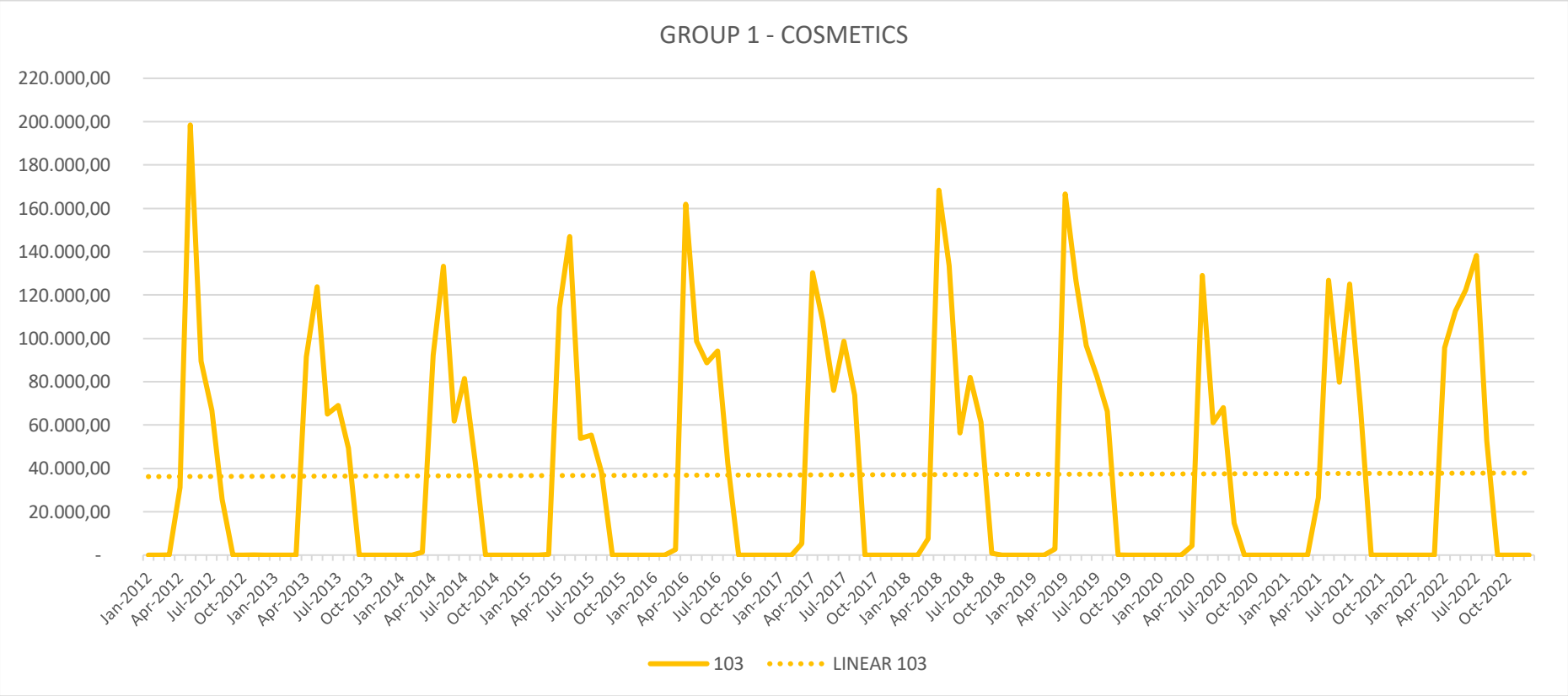

Figure 2-4: Subgroup 103 sales evolution (2012-2022)

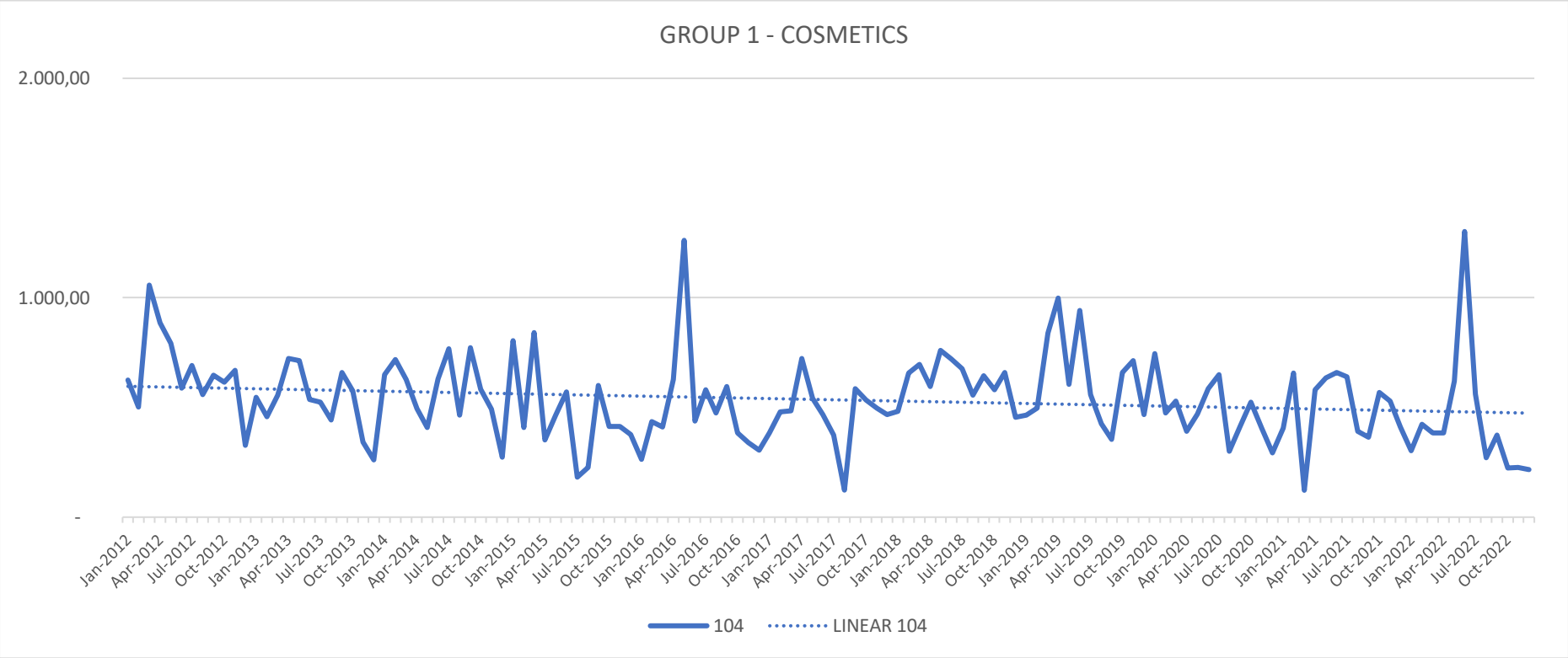

Figure 2-5: Subgroup 104 sales evolution (2012-2022)

Exploring the possible relationship among subgroup sales evolution, as mentioned in the relevant literature, covariance is a means of assessing whether there is any association among the variations of two variables. Positive covariance indicates that the paired values of the two variables increase together, and on the other hand, a negative covariance reveals that when the first variable increases the second decreases and vise versa. However, covariance does not interpret the magnitude, or the strength of the relationship among the two variables (King & Eckersley, 2019) (Albright, Winston, & Zappe, 2011, p. 108). Correlation, on the contrary, which is less sensitive to measurement scale (Albright, Winston, & Zappe, 2011, p. 108) does indicate the strength of the linear relationship.

Since it is preferable to summarize correlations than provide the reader with an enormous number of scatterplots (Albright, Winston, & Zappe, 2011, p. 109), Table 2-2 presents the relationships among subgroup variances.

	101 vs 102	**101 vs 103**	**101 vs 104**
CORRELATION	-0,166	-0,294	-0,036
		102 vs 103	**102 vs 104**
CORRELATION		0,700	0,387
			103 vs 104
CORRELATION			0,300

Table 2-2: Correlation of variance among Group 1 subgroups

The closest the correlation coefficient to 1 or -1, the strongest the relationship among the variable's variations (Albright, Winston, & Zappe, 2011, p. 109), so the only pair of subgroups that present absolute correlation greater than 0.5 are 102 and 103 (0.70) a circumstance anticipated due to their concurrent summer peak sales and winter lows.

2.2.2. FREQUENCY

Group 1 distribution is extremely asymmetrical, as skewness is higher than 1 (1.12) positively skewed and platykurtic, with kurtosis 0.12 (Giannikos, Descriptive Statistics, 2020). The Group's histogram (Figure 2-6) reveals this significant abnormality, as the

vast majority of observations fall below the sales mean, while the median is also located to the first bin of observations, containing sales from 2.847,04€ up to 38.375€.

The interquartile range demonstrates as well the distributions' skewness, as the distance up to the third quartile (75%) is composed with sales values up to 86.460,71€ (Figure 2-7), and the median (17.780€) is found extremely close to the 25th percentile (10.175.85€) (Zaiontz, Retrieved on May 10th, 2023).

	GROUP 1	**101**	**102**	**103**	**104**
MEAN	**49.408,89**	4.921,10	6.918,35	***37.034,25 / 75.208,01***	535,20
MEDIAN	**17.780,12**	3.724,08	6.498,97	***- / 73.670,55***	529,16
SKEWNESS	**1,12**	1,69	0,47	1,12	0,90
KURTOSIS	**0,12**	2,54	-0,38	0,10	2,37

Table 2-3: Group 1 - Frequency Features

As far as distinct subgroups are concerned, 103 is found to be in absolute accordance with Group 1 aggregated frequency features in terms of skewness and kurtosis, an anticipated fact due to its great participation to the aggregated group data.

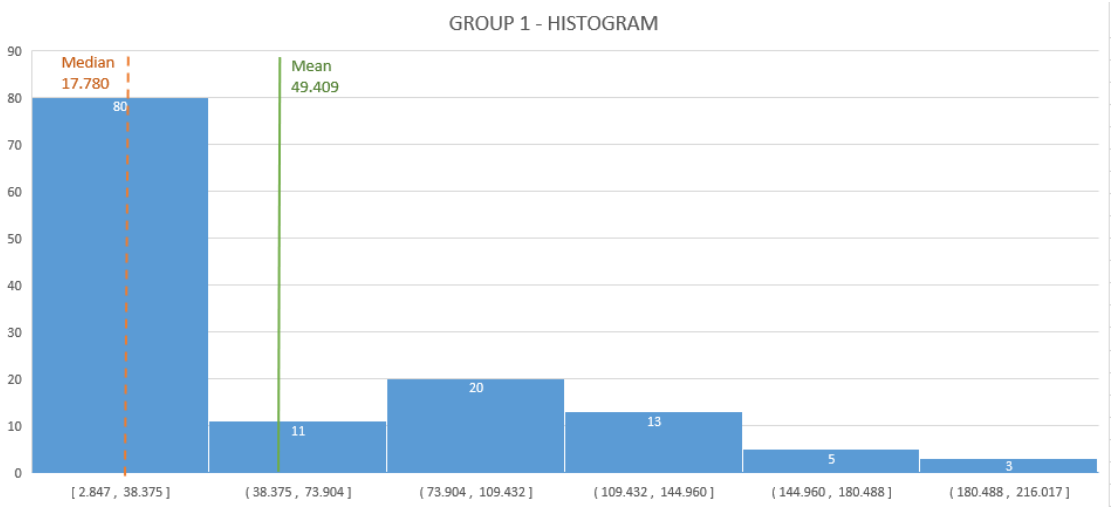

Figure 2-6: Histogram – Group 1

Subgroups 101 and 104 demonstrate kurtosis under but close to 3 units, characterized platykurtic to mesokurtic (Giannikos, Descriptive Statistics, 2020). 102 and 104 subgroups are moderately skewed to the left (positively), whereas subgroup 101 is highly skewed to the same direction (Giannikos, Descriptive Statistics, 2020).

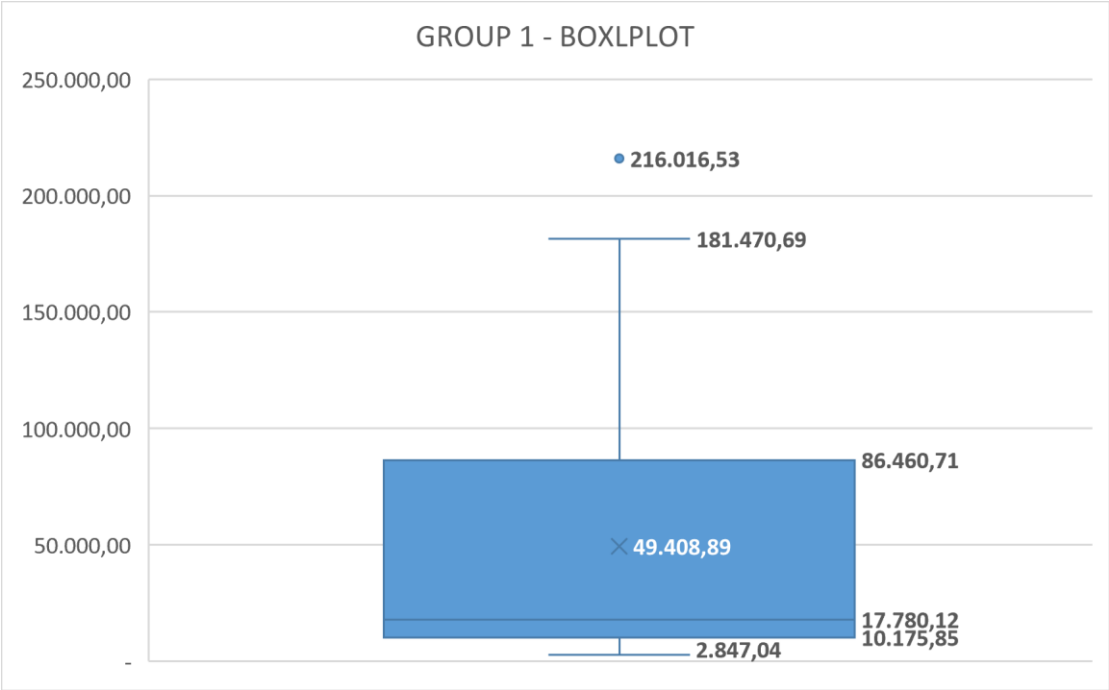

Figure 2-7: Boxplot - Group 1

2.2.3. SEASONALITY

Figure 2-8 compares dispersion of observations in a monthly level while depicting a seasonal effect present in the timeseries through a monthly boxplot. The extreme seasonal effect was firstly demonstrated in Figure 2-1 and it results in a sales spike in spring, mostly in April for the years up to 2019. From 2020 the spike lagged one month, probably due to the retailer's pessimism on the expected tourist arrivals due to the pandemic (Kouvrakis, 2023).

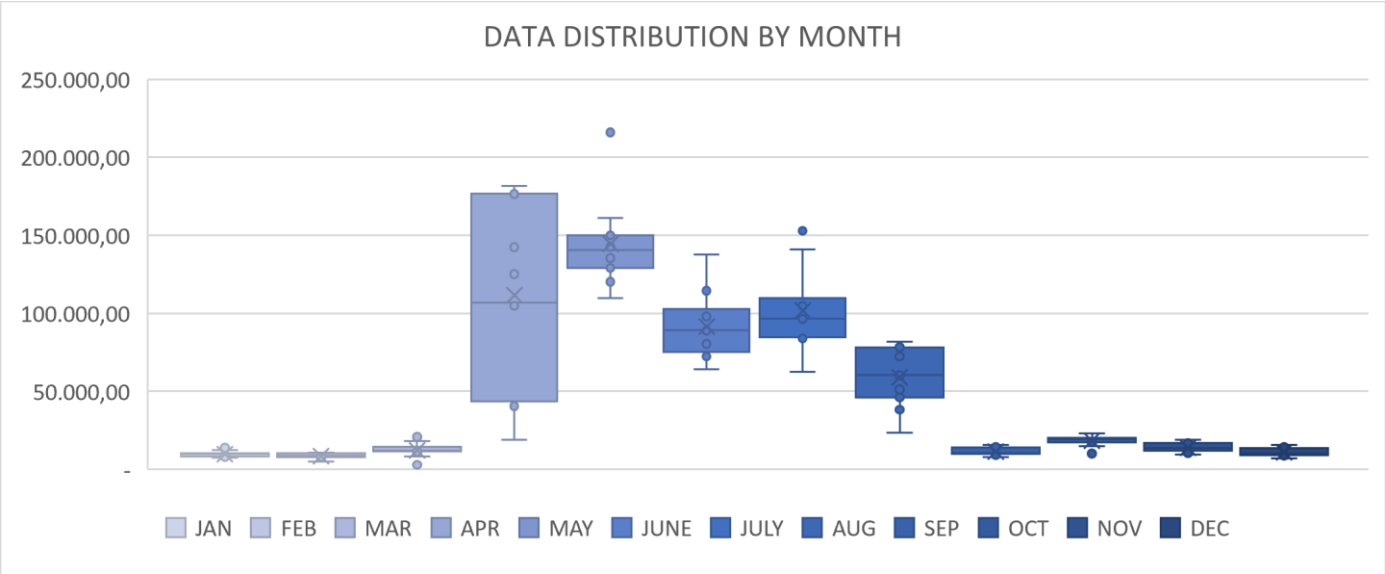

Figure 2-8: Monthly Boxplot – Group 1

April, in particular, performs the wider standard deviation among twelve months (57.778,20€) and the second larger average sales (111.558,87€) after May, with mean sales on 144.499,53€. The observed facts come in absolute compliance with the distinct seasonal effect that subgroup 103 market practice imposes. As exhibited in Figure 2-9, April also acquires the wider interquartile range of sales among the rest months of the year, furtherly establishing monthly sales variance.

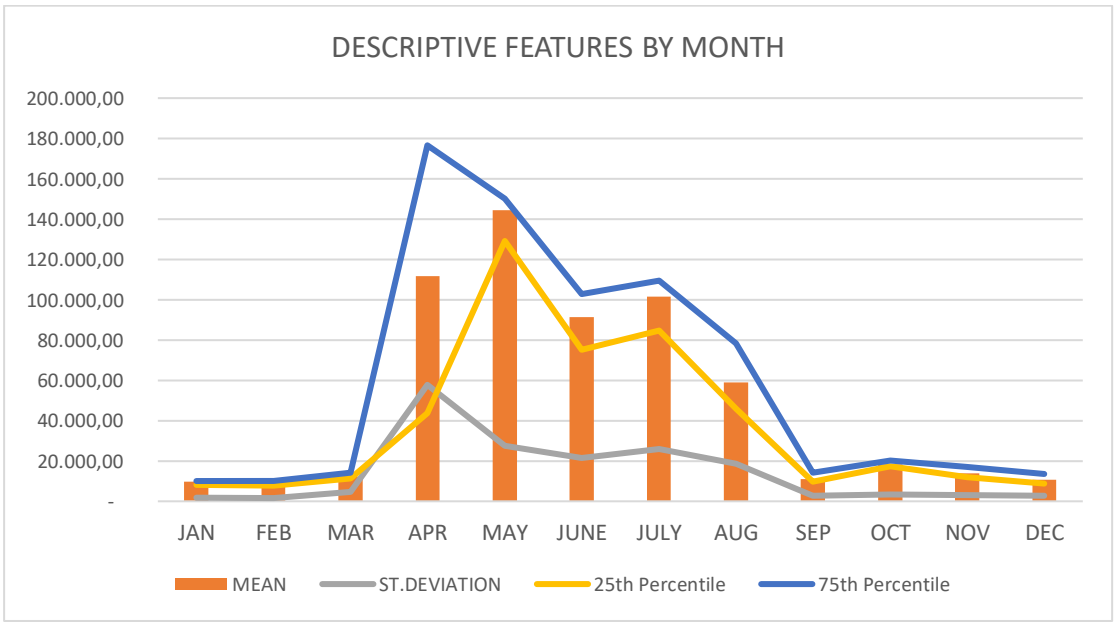

Figure 2-9: Descriptive Features by Month

The effect that seasonality imposes on Group 1 sales is better appreciated through regression analysis further on.

2.3. Regression

2.3.1. SEASONAL EFFECTS

Since seasonality was brought up as a major driver of sales evolution, the first step of regression analysis will aim to record its effect through the incorporation of dummy variables for each month of the year (Thomaidis N. S., Special Forecasting Topics, 2021). Regression results are presented in Table 2-4, at a confidence level 95%.

Insignificant factors with P-value above 0.05 (Thomaidis N. S., Special Forecasting Topics, 2021), Linear Trend, February, March, September and December, are removed, so as to obtain the results presented in Table 2-5.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	9,215	0,113	81,289	0,000	8,991	9,440
Linear Trend	-0,001	0,001	-0,843	0,401	-0,002	0,001
FEB	-0,137	0,145	-0,939	0,350	-0,425	0,151
MAR	0,149	0,145	1,027	0,307	-0,139	0,437
APR	2,259	0,145	15,534	0,000	1,971	2,547
MAY	2,694	0,145	18,522	0,000	2,406	2,982
JUNE	2,227	0,145	15,311	0,000	1,939	2,515
JYLY	2,330	0,146	16,016	0,000	2,042	2,618
AYG	1,759	0,146	12,085	0,000	1,471	2,047
SEP	0,124	0,146	0,852	0,396	-0,164	0,412
OCT	0,629	0,146	4,319	0,000	0,341	0,917
NOV	0,341	0,146	2,341	0,021	0,053	0,629
DEC	0,091	0,146	0,627	0,532	-0,197	0,380

Table 2-4: Regression results on Seasonality – 1st step

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	9,218	0,046	199,874	0,000	9,126	9,309
APR	2,215	0,113	19,604	0,000	1,991	2,438
MAY	2,649	0,113	23,449	0,000	2,425	2,872
JUNE	2,181	0,113	19,310	0,000	1,958	2,405
JYLY	2,284	0,113	20,215	0,000	2,060	2,507
AYG	1,711	0,113	15,150	0,000	1,488	1,935
OCT	0,580	0,113	5,136	0,000	0,357	0,804
NOV	0,292	0,113	2,581	0,011	0,068	0,515

Table 2-5: Regression results on Seasonality - 2nd step

Regression statistics at the second step of seasonality examination provide us with an 0,911 R Square, meaning that seasonal effects can explain approximately 91% of the variability in sales evolution (Albright, Winston, & Zappe, 2011, p. 557). The dominance of seasonal effects is incorporated in the regression line produced by Equation 1.

(1) $$log\left(\widehat{S_t}\right) = 9.218 + 2.215 * APR + 2.649 * MAY + 2.181 * JUN + 2.284 * JUL + 1.711 * AUG + 0.580 * OCT + 0.292 * NOV$$

Given the high value of R Square, as well as the adjusted R Square for Equation 1, a Goodness-of–Fit graph (Thomaidis N. S., Special Forecasting Topics, 2021) is presented in Figure 2-10: Regression Performance on Equation 1, in order to better appreciate the models' performance as an explanatory means on sales evolution. Indeed, estimated sales line follows actual sales line pattern, producing residuals with Mean Absolute Error (MAE) 11.059,02€ (Thomaidis N. S., Simple Forecasting Techniques, 2021) and Mean Absolute Percentage Error (MAPE) 25.73% (Thomaidis N. S., Simple Forecasting Techniques, 2021).

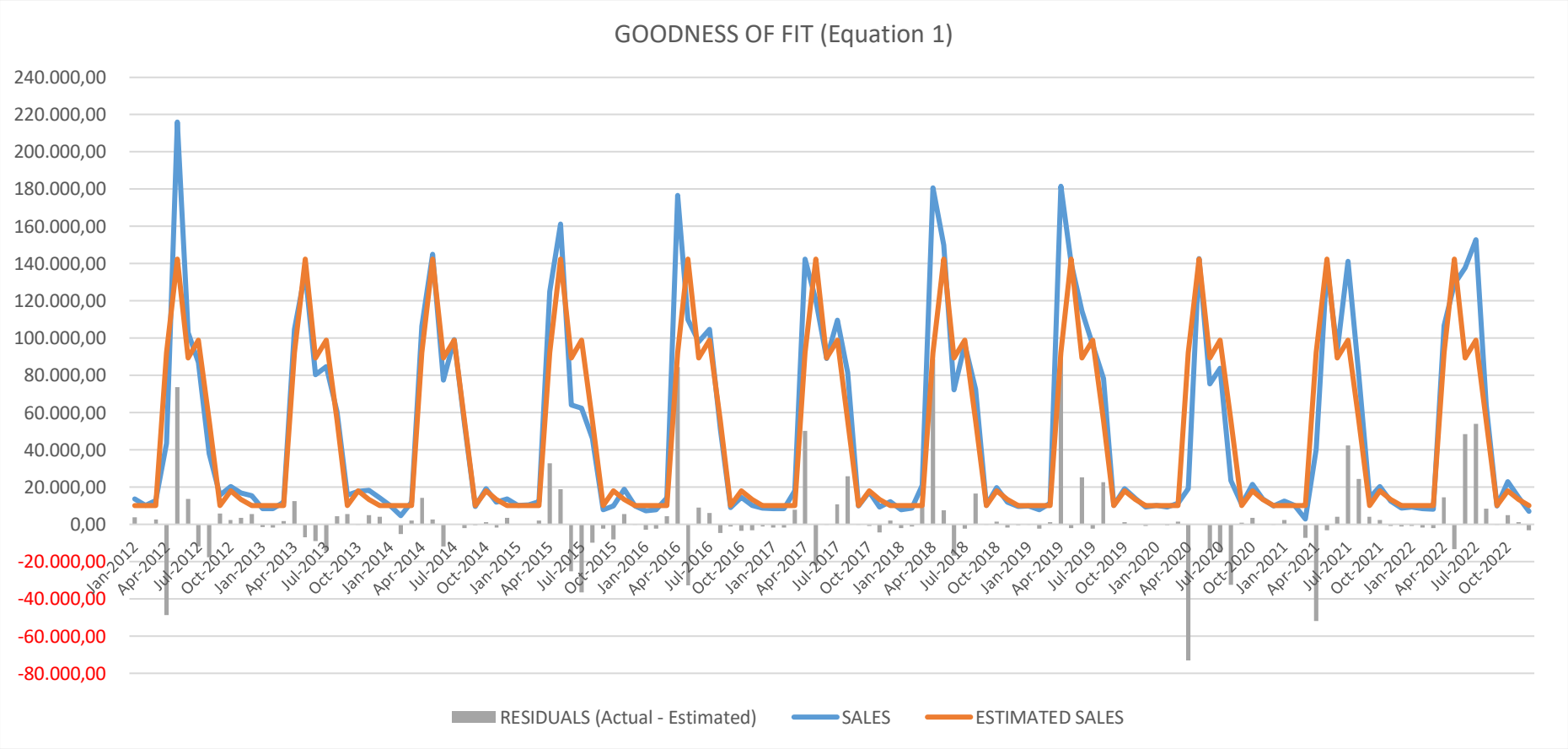

Figure 2-10: Regression Performance on Equation 1

2.3.2. PANDEMIC EFFECT

The effect that COVID-19 pandemic has imposed on sales will be examined through several variables:

- WHO announcements. On 30/1/2020, after observing the prevailing number of coronavirus cases in China from December 2019 and the rapid spread to other countries across the globe, the World Health Organization declared a Public Health Emergency of International Concern (PHEIC). Thereupon, on 11/3/2020 WHO characterized the COVID-19 outbreak as a pandemic (World Health Organization, Retrieved on May 5th, 2023). The characterization was substantially raised on May 5th 2023, as WHO declared that present situation does no longer fit the PHEIC definition (World Health Organization, Retrieved on May 12th 2023). Thus, in the present study frame, pandemic is considered active from March 2020 up to December 2022, at the end of the timeseries.
- 1st lockdown imposition. In Greece, first restrictive measures against the spread of COVID-19 were imposed on March 2020, initially with schools and universities that ceased works (Sotiriou, 2021) and then with additional measures that restricted economic activity and citizen movements (Ministry of Climate Change & Civil Protection, 2020). Those measures were progressively raised during May 2020 (Sotiriou, 2021), thus 1st lockdown dummy variable remains active from March 2020 to June 2020.
- 2nd lockdown imposition. On November 2020 restrictions were imposed again, with respect to economic activity, citizen movements and the operation of educational institutes (Ministry of Climate Change & Civil Protection, 2020). Although being of a more moderate character compared to the first lockdown, the restrictions deescalated progressively from May 2021. The 2nd lockdown dummy factor is active from November 2020 up to April 2021.
- COVID-19 vaccination approval from ECDC and the launch of vaccination program in Greece. The first approval of COVID-19 vaccine was made on 21/12/2020 from the European Centre for Disease Prevention and Control and referred to Comirnaty, a vaccine developed by BioNTech and Pfizer (European Centre for Disease Prevention and Control, 2020). At the same time, on

27/12/2020 Greek citizens are being vaccinated for the first time, namely nurse Kampisiouli E., Prime Minister Kiriakos Mitsotakis and the Greek President Katerina Sakellaropoulou (Sotiriou, 2021). Since EU's vaccine approval and the launch of vaccination program in Greece almost coincide, a unique dummy (VAC) is introduced to the analysis, active from January 2021 up to the end of the timeseries.

Proceeding with data exploration, WHO dummy variable is added in Equation 1, producing the results demonstrated in Table 2-6. WHO independent variable acquires a not significant coefficient at a 95% confidence level, but marginally significant at a 90% confidence level.

However, given that regression analysis proceeds on an α=0.05, WHO will not be incorporated in the regression equation. The next independent variable tested is the first lockdown and the respective results are summarized in Table 2-7. Regression analysis provided with a significant (p=0.011) negative coefficient (-0.518), resembling the declining effect that the first bundle of restrictive measures had on sales.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	9,244	0,049	190,548	0,000	9,148	9,340
APR	2,219	0,112	19,771	0,000	1,996	2,441
MAY	2,653	0,112	23,641	0,000	2,431	2,875
JUNE	2,185	0,112	19,475	0,000	1,963	2,408
JYLY	2,288	0,112	20,386	0,000	2,066	2,510
AYG	1,715	0,112	15,287	0,000	1,493	1,938
OCT	0,584	0,112	5,207	0,000	0,362	0,806
NOV	0,296	0,112	2,635	0,010	0,074	0,518
WHO	-0,112	0,068	-1,649	0,102	-0,246	0,022

Table 2-6: Regression results on Seasonality & Pandemic (WHO variable)

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	9,227	0,045	203,979	0,000	9,138	9,317
APR	2,252	0,111	20,218	0,000	2,032	2,473
MAY	2,687	0,111	24,117	0,000	2,466	2,907
JUNE	2,172	0,111	19,654	0,000	1,953	2,391
JYLY	2,274	0,111	20,580	0,000	2,055	2,493
AYG	1,702	0,111	15,401	0,000	1,483	1,921
OCT	0,571	0,111	5,165	0,000	0,352	0,790
NOV	0,282	0,111	2,554	0,012	0,063	0,501
1st LOCKDOWN	-0,518	0,200	-2,592	0,011	-0,914	-0,122

Table 2-7: Regression results on Seasonality & 1st Lockdown

Further on, the effects of second lockdown are examined, as demonstrated in Table 2-8. Undeniably, second lockdown imposes negative repercussions on sales, acquiring a significant (p=0.012) negative coefficient (-0.356) as well.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	9,253	0,045	203,676	0,000	9,163	9,343
APR	2,260	0,109	20,731	0,000	2,044	2,476
MAY	2,662	0,109	24,335	0,000	2,446	2,879
JUNE	2,146	0,109	19,759	0,000	1,931	2,361
JYLY	2,248	0,109	20,700	0,000	2,033	2,463
AYG	1,676	0,109	15,431	0,000	1,461	1,891
OCT	0,545	0,109	5,015	0,000	0,330	0,760
NOV	0,288	0,108	2,666	0,009	0,074	0,502
1st LOCKDOWN	-0,539	0,196	-2,753	0,007	-0,926	-0,151
2nd LOCKDOWN	-0,356	0,139	-2,553	0,012	-0,631	-0,080

Table 2-8: Regression results on Seasonality, 1st Lockdown & 2nd Lockdown

Subsequently, the last pandemic-related variable is tested, the one relevant with vaccination program effects (VAC). As shown in Table 2-9, although VAC acquired a

reasonable, positive coefficient, its effect is extremely questionable with p=0.811, thus not introduced to the regression equation.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	9,251	0,047	196,814	0,000	9,158	9,344
APR	2,260	0,109	20,650	0,000	2,043	2,477
MAY	2,661	0,110	24,218	0,000	2,444	2,879
JUNE	2,145	0,109	19,670	0,000	1,929	2,361
JYLY	2,247	0,109	20,608	0,000	2,031	2,463
AYG	1,675	0,109	15,360	0,000	1,459	1,891
OCT	0,544	0,109	4,988	0,000	0,328	0,760
NOV	0,288	0,109	2,657	0,009	0,074	0,503
1st LOCKDOWN	-0,536	0,197	-2,721	0,007	-0,925	-0,146
2nd LOCKDOWN	-0,365	0,146	-2,509	0,013	-0,653	-0,077
VAC	0,018	0,077	0,239	0,811	-0,135	0,172

Table 2-9: Regression results on Seasonality, 1st Lockdown, 2nd Lockdown & VAC

Consequently, taking into account the data of the previous step (Table 2-8), as well as the satisfactory regression statistics (R^2= 0.919 and adjusted R^2 = 0.914) the equation that finalizes the regression analysis is Equation 2.

(2) $$log\left(\widehat{S_t}\right) = 9.253 + 2.260 * APR + 2.662 * MAY + 2.146 * JUN + 2.248 * JUL + 1.676 * AUG + 0.545 * OCT + 0.288 * NOV - 0.539 * 1st\ LOCKDOWN - 0.356 * 2nd\ LOCKDOWN$$

Figure 2-11 depicts Equation 2 performance by comparing actual to estimated sales (Thomaidis N. S., Special Forecasting Topics, 2021), while at the same time presents residuals. At that case, Mean Absolute Error (MAE) declines from 11.059,02€ in Equation 1: Expected sales based on seasonality, to 10.939,42€, while Mean Absolute Percentage Error (MAPE) to 25.11% from 25.73% with Equation 1 (Thomaidis N. S., Simple Forecasting Techniques, 2021).

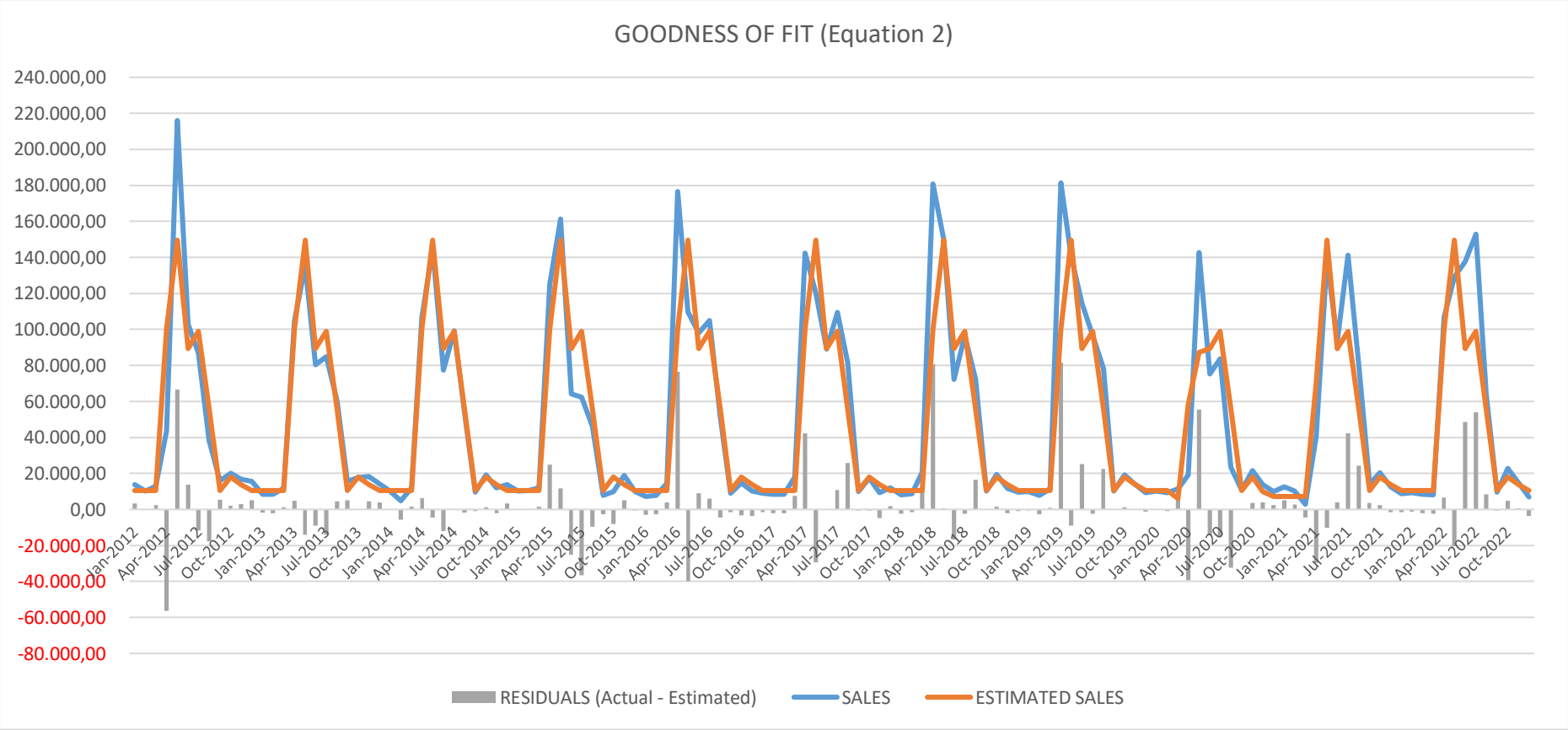

Figure 2-11:Regression performance on Equation 2

2.4. Concluding Remarks

Taking into account that the main objective of this study is to identify and analyze specific patterns and trends that may shape Achtida's sales evolution over time, the main reflections resulting from the analysis of the first Group of products are:

- In aggregate, Group 1 demonstrates a stable behavior, by not incorporating a distinct trend, either upwards or downwards. The timeseries itself performs in a strict cyclical pattern, mainly triggered by Subgroup 103 which constitutes approximately 75% of Group 1 average total sales. The seasonal norm is expressed as an increase in sales on spring months, mainly in April or May of each year, on which peak is accomplished with mean monthly sales 111.559€ and 144.500€ respectively. Subsequently, after keeping a high volume throughout summer, sales decline rapidly on August, proceeding to September on a 11.146€ mean level of monthly sales.
- Apart from subgroup 103 (Sun Protection) that reflects perfectly the aggregate Group 1 behavior, subgroup 102 (Face & Body Cosmetics) seems to acquire a correlated evolution, as its observed pattern coincides to a certain extend with subgroup 103, with summer peaks and winter lows. Subgroup 101 (Fragrances & Gift Sets) on the other hand, reflecting its reference to the local (and not tourist) markets (Kouvrakis, 2023), displays a contrasting behavior, with sales increasing during winter. At last, subgroup 104 (Hair Products) which constitutes a minor portion of Group 1 sales can't be directly related or opposed to other subgroup patterns.
- The data distribution results in a histogram highly skewed to the left, as the majority of observations falls below the distribution mean. Translated in simple words, most of the monthly sales observations are smaller than the average monthly sales, highlighting the sheer effect of increased summer sales imposed to the timeseries.
- Regression analysis underscored the importance of seasonality and its implications towards sales formation, as in a 95% significance level seasonality is proved to be responsible for approximately 91% of variation.

- Apart from seasonal effects, the COVID-19 pandemic era as a continuous, lasting condition did not demonstrate significant impacts on sales, although the 1st and the 2nd Lockdown did. Veritably, through regression applications the first bunch of restrictions applied from March 2020 up to May 2020, as well as the second bunch imposed from November 2020 through to April 2021 had a significant declining effect on sales.

3 GROUP 2 – PERSONAL HYGIENE & CLEANNESS

3.1. Group Introduction

The second group of products consists of seven discreet subgroups. All of them fall under the category of personal hygiene and cleanness, as their commercial characteristics as well as their usefulness for the final consumer are similar.

For the eleven-year period under study, the monthly observations representing the sales of Group 2 move around a mean of 109.936.54€, with average deviation by 31.188.82€. The minimum value of sales throughout the eleven-year period is 38.422.55€, whereas the maximum is 176.535,65€.

3.1.1. HAIR AND BODY CLEANNESS (201)

The first subgroup consists of products used for hair and body cleanness, such as shampoo, shower gel, foam bath and hair conditioner. The major brand names under those categories are Wash & Go, Orzene, Noxzema, Palmolive, Sanex, Vidal, Le Petit Marseillais, Str8, Papoutsanis, Johnson Baby and many other less recognizable brands. The majority of those brands are produced or distributed in the Greek market from companies like Johnson & Johnson, Sarandis, Elgeka and Colgate Palmolive.

The average contribution of 201 subgroup to the formation of Group 2 total is 20.22%.

3.1.2. DEODORANTS (202)

The second subgroup is formatted from the aggregation of sales of roll-on and spray deodorants, manly from the brands Noxzema, Str8, Bioten and Sanex. In accordance with 201 subgroup, major distributors for 202 are Sarandis and Colgate Palmolive. Throughout the period under study, subgroup 202 accounted for 9.47% of Group 2 total sales.

3.1.3. MOUTH HYGIENE (203)

The majority of the components that form subgroup 203 is traded by Johnson & Johnson through Listerine brand, as well as by Colgate Palmolive through a wide range of toothpastes, toothbrushes and other complementary commodities such as mouth

wash, dental floss etc. Mouth hygiene contributed by 8.38% on period average to Group 2 sales.

3.1.4. ADULT DIAPERS & SANITARY NAPKINS (204)

Subgroup 204 is composed by items merchandised in the Greek market by Essity Hellas, a distributor of an extreme range of personal hygiene products with brands like Tena and Libresse. Likewise, Johnson & Johnson accounts for a remarkable part of subgroups' total, with O.B. tampons and Carefree sanitary napkins. The great dynamics and potential of 204 subgroup is reflected to its contribution to Group 2 sales, that on average accounted for 43.32%.

3.1.5. BABY CARE & DIAPERS (205)

Baby Care is a subgroup with characteristics similar to 204, since it is composed by commodities that reserve basic, or even urgent needs of hygiene and personal cleanness. Again, the majority of products are marketed by Johnson & Johnson and Essity Hellas. Indicatively but not exhaustively those products are Johnson Baby shampoos, baby powder, hair conditioners, foam baths, and Libero diapers.

205 subgroup accounted on average for 9.53% of periods' sales.

3.1.6. SHAVING, COTTON, PHARMACY (206)

Through the first steps of the analysis, Shaving, Cotton and Pharmacy, had been treated as three distinct subgroups. However, as data processing progressed, it became noticeable that those distinct subgroups each accounted for minor portions of total Group 2 sales, particularly under 3% each. Accepting that such a small contribution would not provide the study with any sufficient hints either of commercial value or of statistical importance, the three subgroups were aggregated to form 206 subgroup, that accounts on average for 5.82% of the eleven-year period average sales.

Indicatively, the major brands that compose 206 subgroup are Noxzema, Str8, Denim and Prosar (Shaving), Johnson Baby and Demak-Up (Cotton), as well as Depon (Pharmacy).

3.1.7. HANDWASHING & PERSONAL PROTECTION (207)

Subgroup 207, composed of liquid soap, soap bars, face masks and medical gloves had again been treated autonomously, as a distinct group. That choice was triggered by the

appeal those product families had through the COVID-19 pandemic outbreak and the sales spike on the first months of the pandemic, as remarked by wholesalers and retailers. However, the observed minor performance in terms of sales percentage through the period 2011-2022 (3.26% on average) led to its integration to Group 2.

Brands that compose the subgroup are Le Petit Marseillais, Palmolive, Vidal, Noxzema, Johnson & Johnson, Cleopatra, Natura liquid and bar shoaps, as well as other non-recognizable brands of face masks and medical gloves.

3.2. Data Exploration

Figure 3-1 presents Group 2 sales' evolution over time, together with individual subgroup representation.

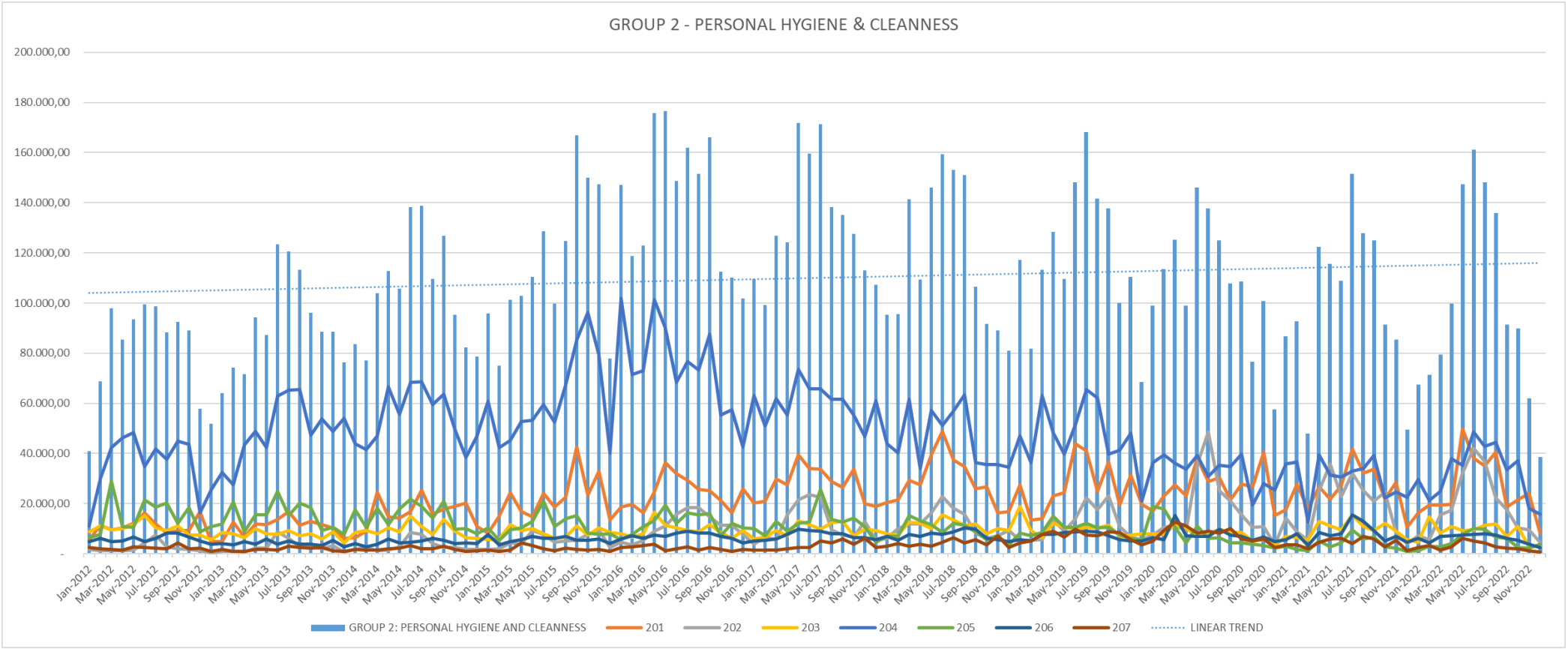

Figure 3-1: Group 2 sales evolution (2012-2022)

3.2.1. VISUAL OBSERVATIONS & DESCRIPTIVE FEATURES

At the first level of data exploration, an observation of Figure 3-1 leads to several primary claims.

Firstly, the aggregated Group sales reveal an upward trend throughout the eleven-year period. Hints of seasonal effects are demonstrated, as in every year a spike in sales volume is detected through the summer months, noticeably from May to September. Up to 2017 that seasonal pattern incorporates a distinct upward trend, performing higher volumes on summer, however from 2018 and on the pattern is differentiated as at peak season sales keep volume over 145.000€ while in low season sales volume becomes weaker. For instance, in March 2021, December 2021 and December 2022 the aggregated Group sales adjust to a volume around 40.000€, a level previously kept only in the winter of 2012 and 2013. Consequently, the positive linear trend of the eleven-year period forced by strong sales volume performed through summer months is moderated after being negatively affected by the winter decline.

Throughout the period under study, Group 2 sales volume floats within a range of 138.113,10€, with a minimum value of 38.422 (December 2022) and a maximum of 176.535.65€ (May 2016). As far as central tendency is concerned (Fraser, 2016, p. 10), the mean value of sales is 109.936,54€, while the median, the point that splits observation set in two halves (Giannikos, Descriptive Statistics, 2020) is very close, at 108.220,04€.

	GROUP 2	**201**	**202**	**203**	**204**	**205**	**206**	**207**
MINIMUM	**38.422,55**	4.803,12	525,51	2.793,38	11.077,83	936,66	2.524,58	576,04
MAXIMUM	**176.535,65**	49.780,93	48.527,39	19.021,34	101.836,35	29.171,31	15.562,94	12.774,54
RANGE	**138.113,10**	44.977,81	48.001,88	16.227,96	90.758,52	28.234,65	13.038,36	12.198,50
MEAN	**109.936,54**	22.233,98	10.414,89	9.210,41	47.620,64	10.472,86	6.394,96	3.588,79
MEDIAN	**108.220,04**	21.137,18	7.628,38	9.196,65	44.714,83	10.187,34	6.164,02	2.618,86
STANDARD DEVIATION	**31.188,82**	9.968,38	9.383,21	2.757,21	17.944,72	5.703,34	2.128,49	2.495,26

Table 3-1: Group 2 - Basic Descriptive Features

The standard deviation of observations, that equals the square root of variance and represents the average deviation from the mean, fosters dispersion assessment (Fraser, 2016, p. 10) and is located at 31.188.82€.

As far as individual subgroups are concerned, 201 (Hair & Body Cleanness) performs an upward trend up to 2015, while after that time point stabilizes and moves alongside with Group 2 pattern, performing seasonal spikes and lows. Mean and median resemble in that case as well, the former being at 22.233,98€ and the latter at 21.137,18€. Maximum sales value is at 49.780,93€ in May 2022, while minimum value is located at the beginning of the timeseries, on December 2012. The observed standard deviation from the mean is 9.968,38€.

Subgroup 202 (Deodorants) performance is nearly stable through the first four years, while in July 2016 reaches a peak of 18.452 € in sales volume. After that time point sales demonstrate a clear seasonal pattern with upward trend, which becomes even stronger between 2020 – 2022. Indicatively, after 2020 the minimum sales volume is 3.622€ and the maximum volume 48.527€, contrary to 525€€ and 23.687€ respectively from 2012 to 2019. Subgroup 202 average percentage change throughout the period under study is 19.06%, the largest among all subgroups.

Subgroup 203 (Mouth Hygiene) together with 206 (Shaving, Cotton, Pharmacy), demonstrate the slightest fluctuation of sales among all subgroups, with 203 performing a nearly stable evolution over time (+6.19% average change over eleven years) and 206 +4.29% respectively. Indicatively, the standard deviation of those components is 2.757,21 for 203 and only 2.128,49 for 206, a magnitude relevant with their narrow range.

Subgroup 204 (Adult Diapers & Sanitary Napkins) noticeably occupies the majority of sales volume up to 2020, while, however, its participation to Group 2 total sales is moving downwards from 2016 and on. The average contribution to Group 2 formation is 43.32%. Monthly, 204 sales seem to adjust to Group 2 sales, having seasonal behavior. Mean sales are located around 47 thousand euros, with a standard deviation of 17.944,72 €.

Subgroup 205 (Baby Care & Diapers) seems to have a stable evolution up to 2019 moving from a minimum of 3.556€ to a maximum of 29.171€, while from 2020 and on

loses robustness acquiring a minimum volume of 936€ and a maximum one of 18.838€. The downward trend is clearly presented schematically in Figure 3-6. Finally, subgroup 207 which is highly related with COVID-19 outbreak hits its peak volume on March 2020, a time point related with the pandemic outbreak as well as the first lockdown imposition in Greece. At that time sales reached 12.774.,54€, a level never again accomplished. Noticeably, after March 2020 subgroup 207 kept only its seasonal pattern while completely eliminating upward trend.

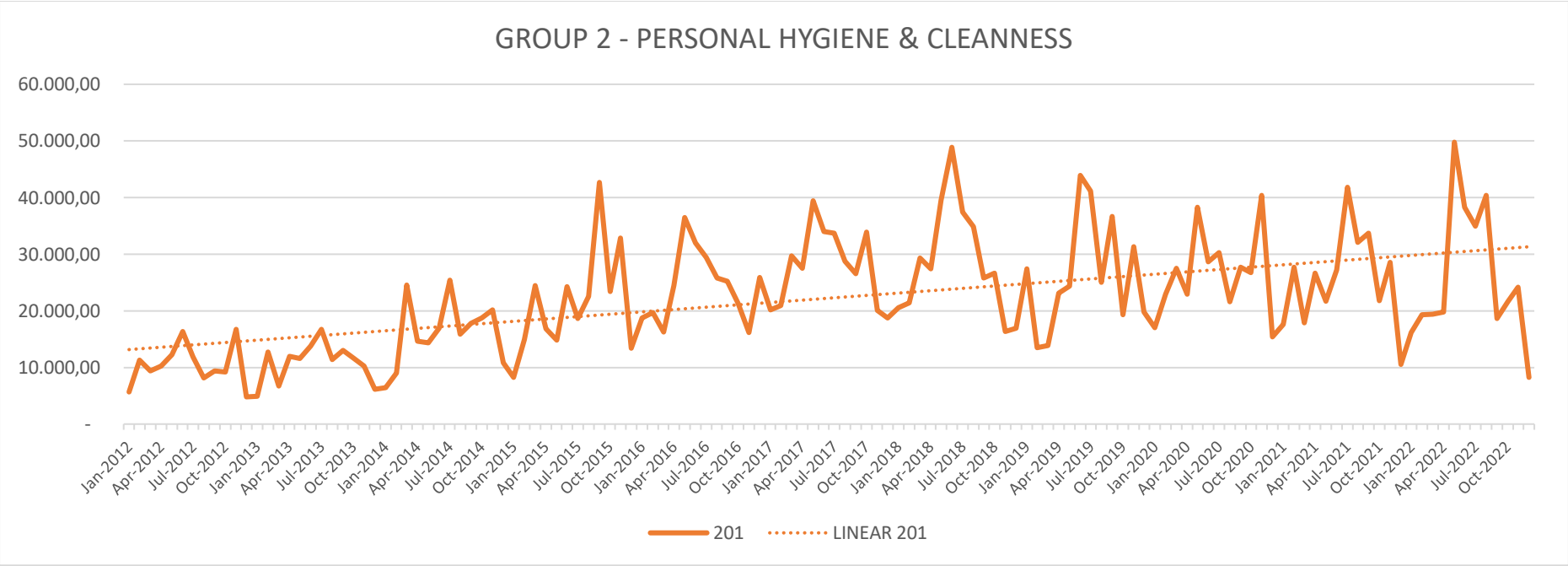

Figure 3-2: Subgroup 201 sales evolution (2012-2022)

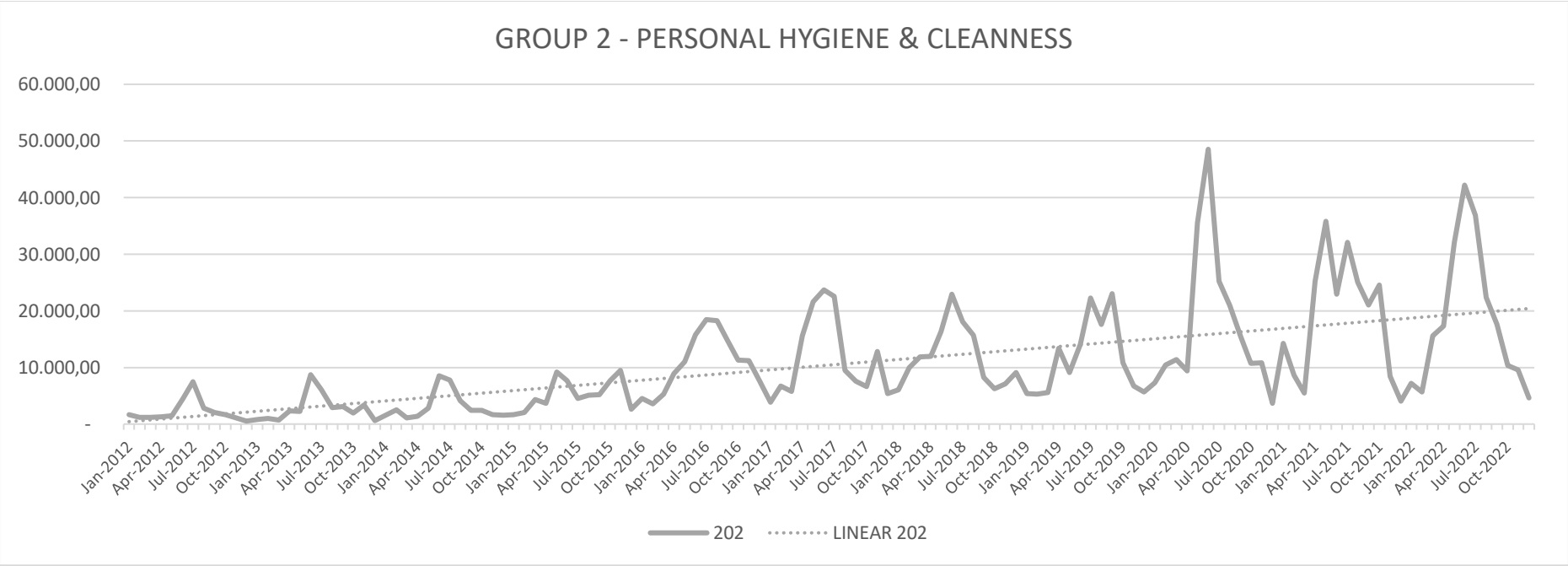

Figure 3-3: Subgroup 202 sales evolution (2012-2022)

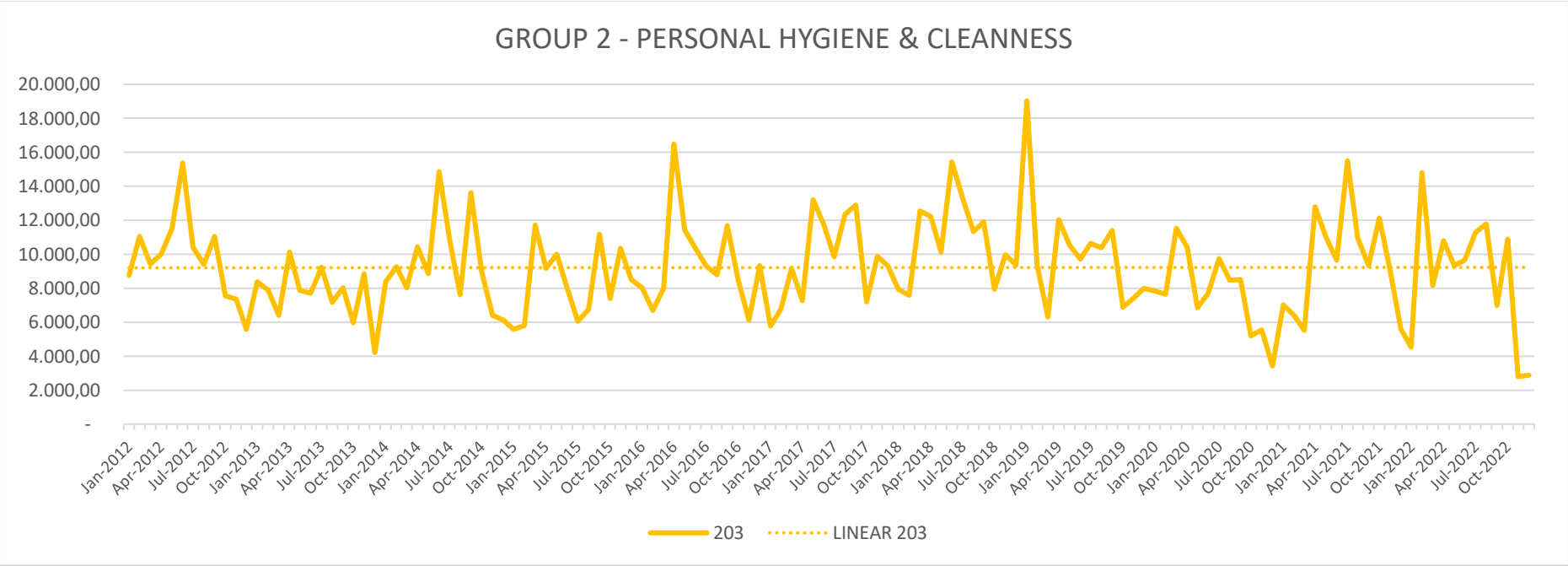

Figure 3-4: Subgroup 203 sales evolution (2012-2022)

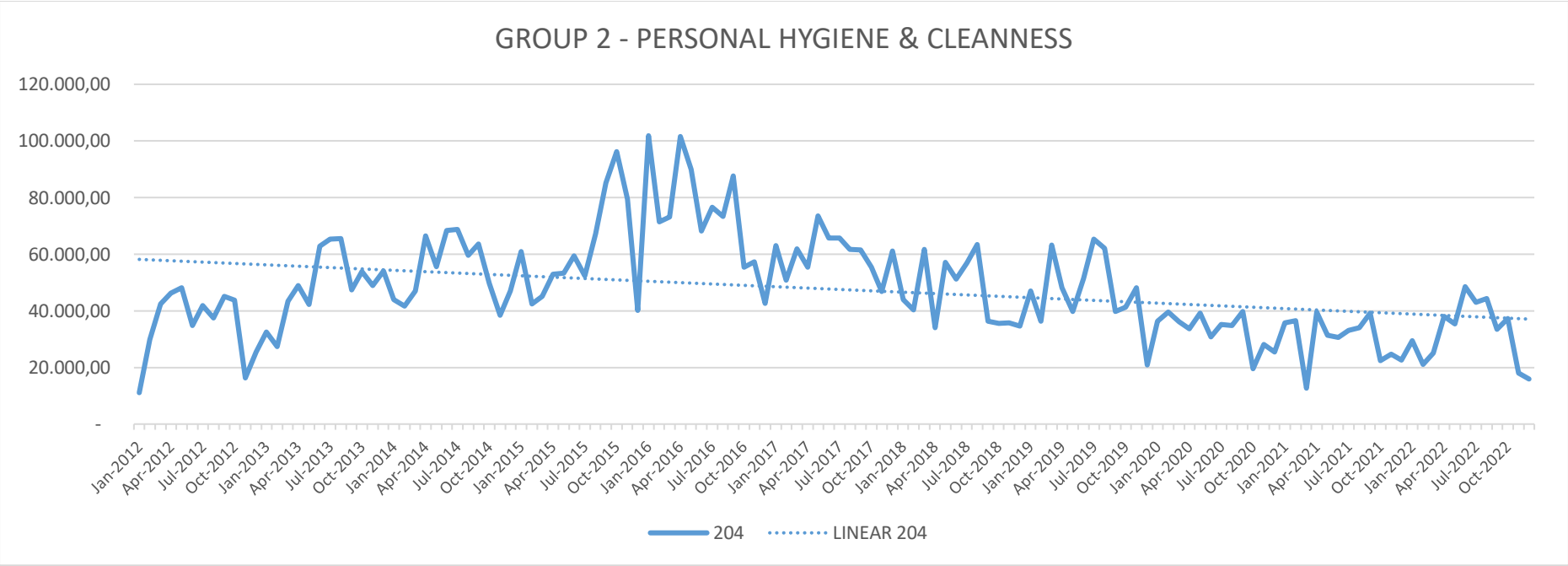

Figure 3-5: Subgroup 204 sales evolution (2012-2022)

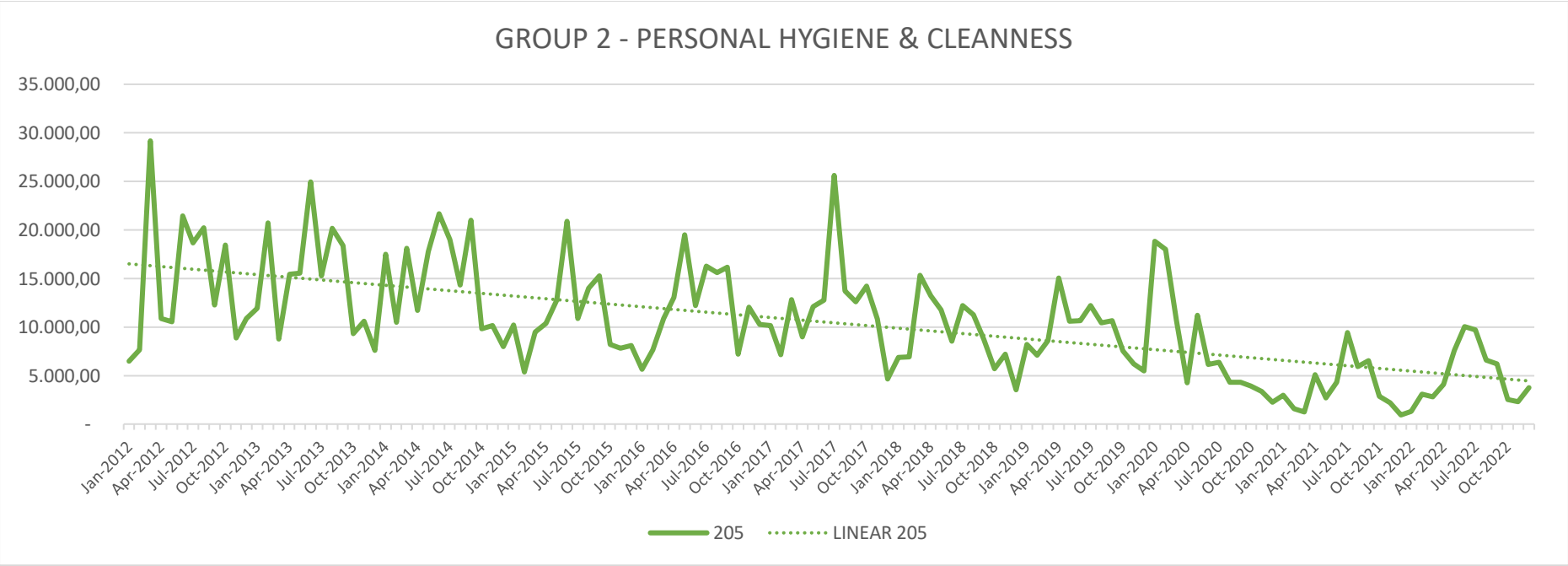

Figure 3-6: Subgroup 205 sales evolution (2012-2022)

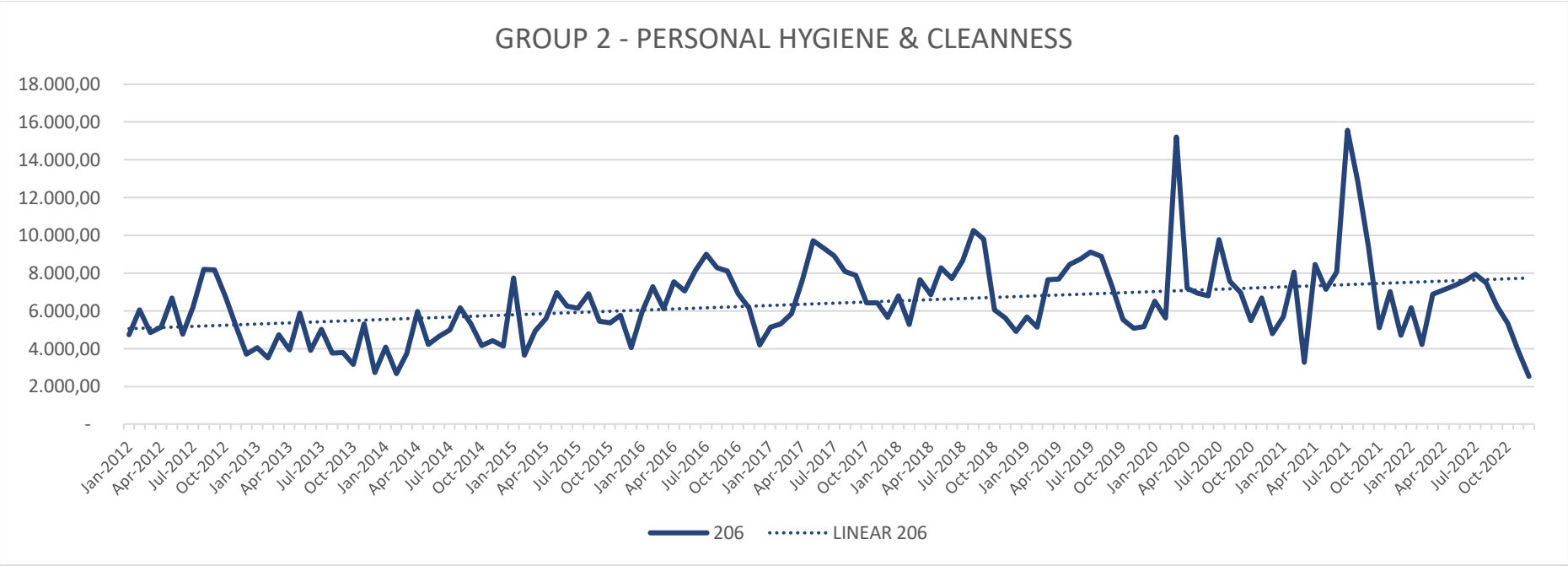

Figure 3-7: Subgroup 206 sales evolution (2012-2022)

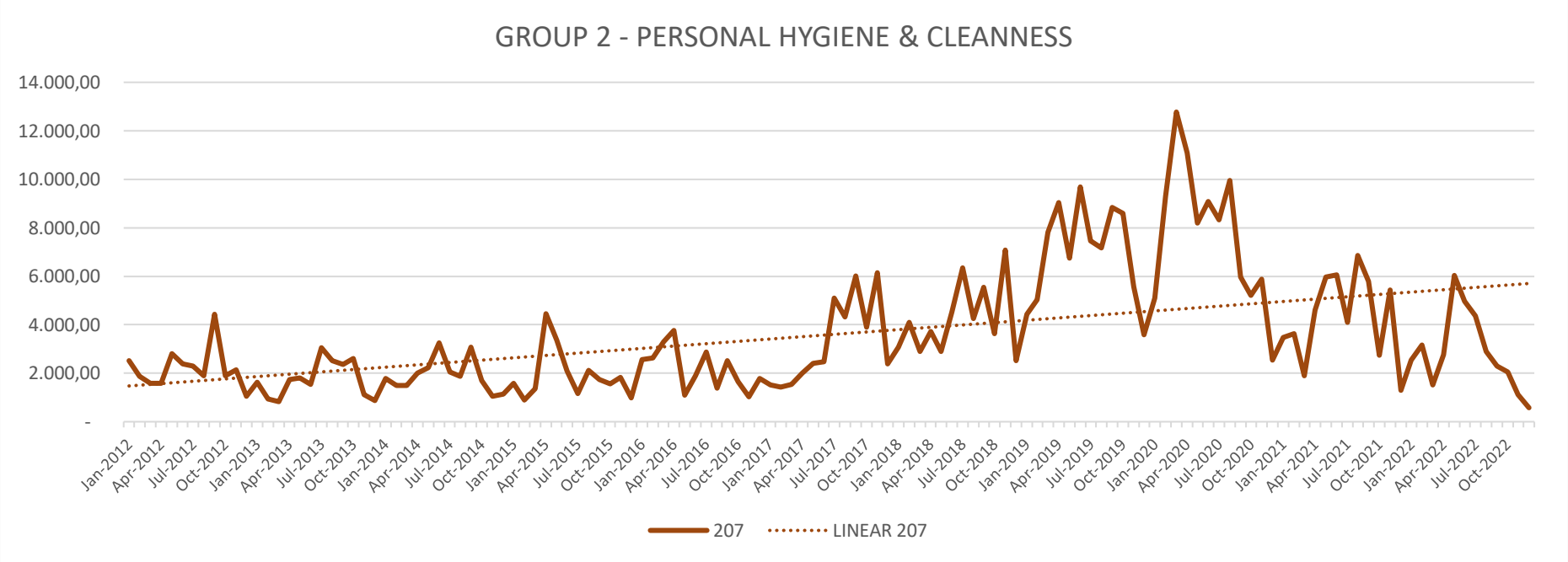

Figure 3-8: Subgroup 207 sales evolution (2012-2022)

Concerning the presence of relationship among subgroup evolution, correlations in Table 3-2 present the extend of concurrence among subgroup variances.

201 vs 202	**201 vs 203**	**201 vs 204**	**201 vs 205**	**201 vs 206**	**201 vs 207**
0,678	0,376	0,208	-0,087	0,581	0,426
	202 vs 203	**202 vs 204**	**202 vs 205**	**202 vs 206**	**202 vs 207**
	0,273	-0,045	-0,185	0,566	0,493
		203 vs 204	**203 vs 205**	**203 vs 206**	**203 vs 207**
		0,238	0,243	0,419	0,230
			204 vs 205	**204 vs 206**	**204 vs 207**
			0,450	0,150	-0,153
				205 vs 206	**205 vs 207**
				-0,024	-0,152
					206 vs 207
					0,533

Table 3-2: Correlation of variance among Group 2 subgroups

The main points captured in Table 3-2 data are in absolute accordance with the trends observed in Figure 3-2 through to Figure 3-8. Taking into account that the closest the correlation coefficient to 1 or -1, the strongest the relationship among the variable's variations (Albright, Winston, & Zappe, 2011, p. 109), the only pairs that present absolute correlation bigger than 0.5 are 201 & 202 (0.68) and 201 & 206 (0.58), 202 & 206 (0.57) and 206 & 207 (0.53).

In all those cases correlation is marginally over 0.5, displaying a relationship which is better depicted in the scatterplots in Figure 3-9 through to Figure 3-12. Indeed, the presence of the relationship becomes evident, although its strength is moderate in all cases.

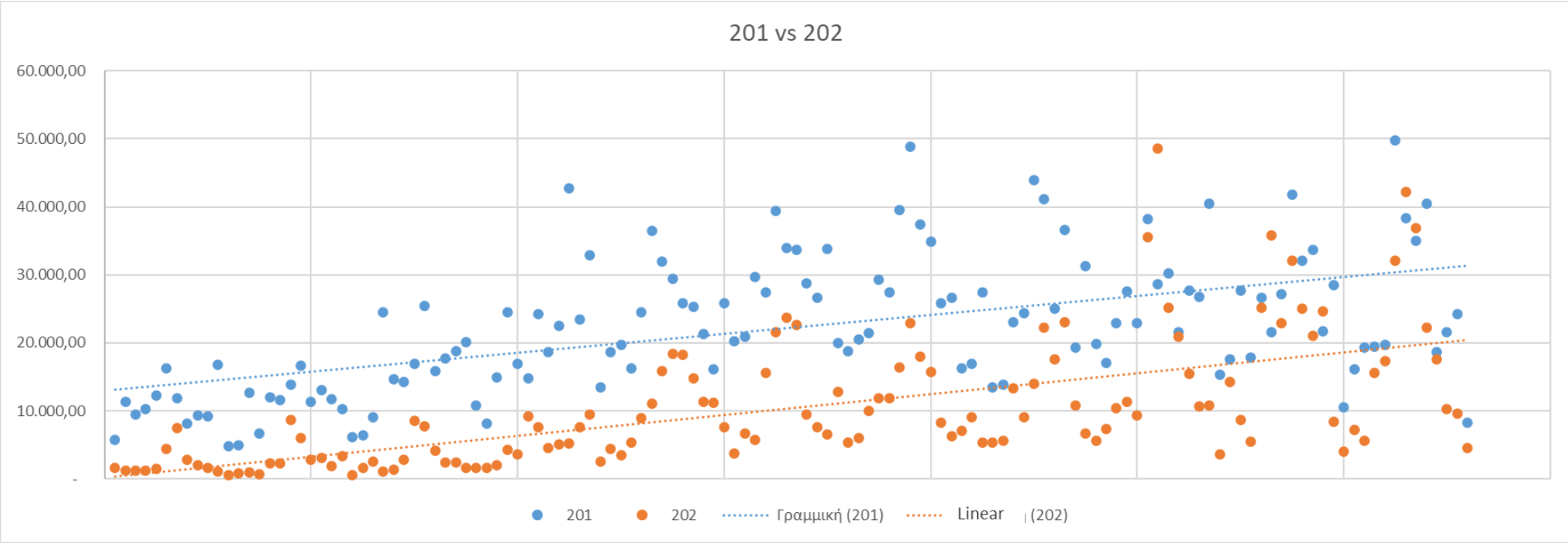

Figure 3-9: Correlation among 201 & 202

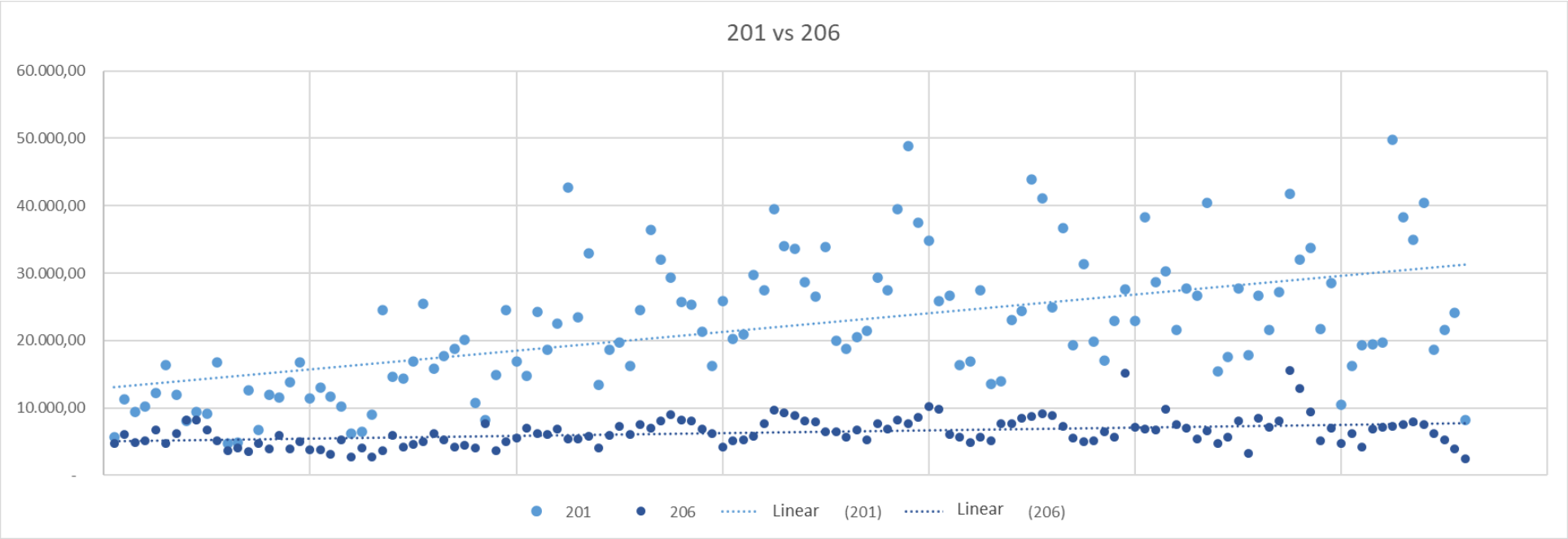

Figure 3-10: Correlation among 201 & 206

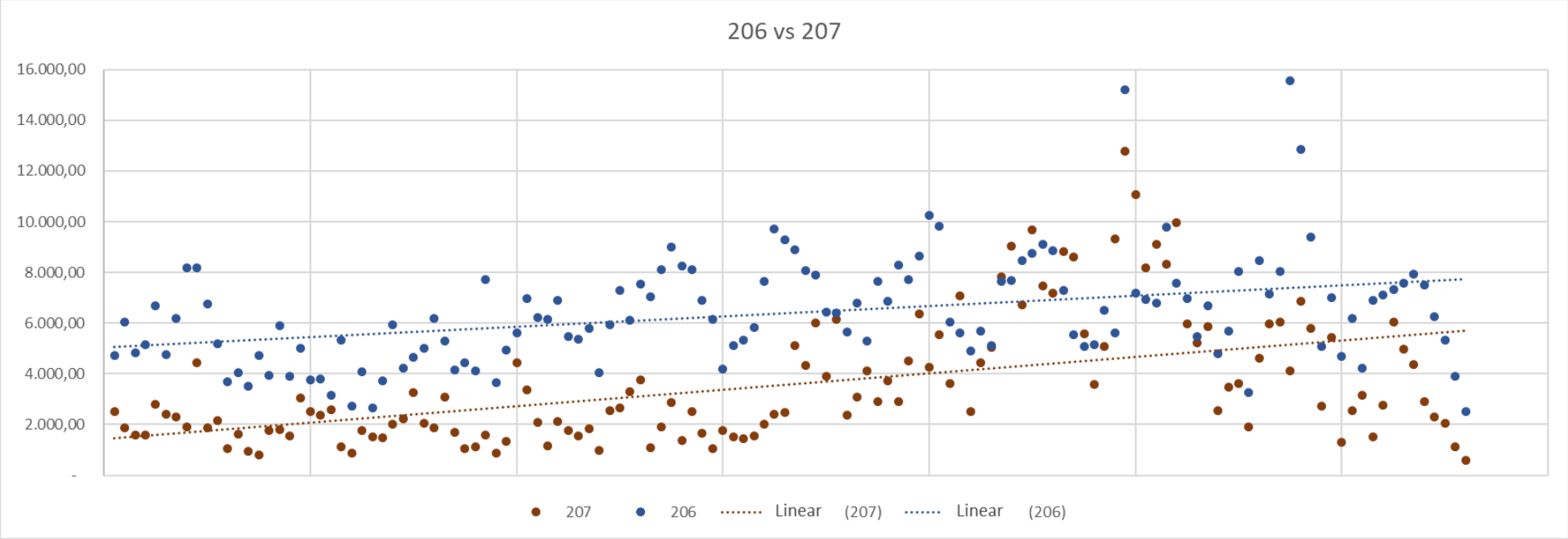

Figure 3-11: Correlation among 206 & 207

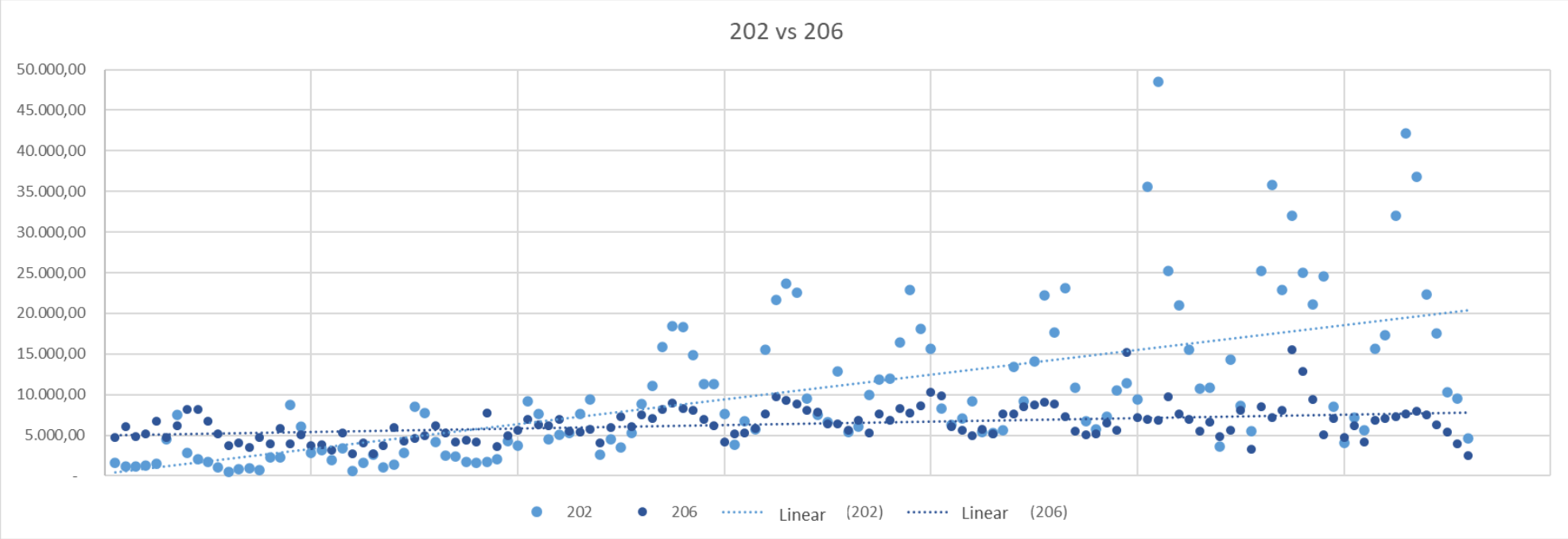

Figure 3-12: Correlation among 202 & 206

3.2.2. FREQUENCY

Group 2 data distribution is fairly symmetrical, as skewness is slightly positive (0.11) and platykurtic, with kurtosis -0.49 (Giannikos, Descriptive Statistics, 2020). As confirmed by Figure 3-13 and the relevant boxplot (Figure 3-14) a marginal majority of observations is located below the mean. As far as interquartile range is concerned, the distance between the 25th percentile and 75th percentile, it is found between 88.605€ and 133.487€ (Zaiontz, Retrieved on May 10th, 2023).

	GROUP 2	**201**	**202**	**203**	**204**	**205**	**206**	**207**
MEAN	**109.936,54**	22.233,98	10.414,89	9.210,41	47.620,64	10.472,86	6.394,96	3.588,79
MEDIAN	**108.220,04**	21.137,18	7.628,38	9.196,65	44.714,83	10.187,34	6.164,02	2.618,86
SKEWNESS	**0,11**	0,50	1,56	0,50	0,61	0,66	1,27	1,31
KURTOSIS	**-0,49**	-0,23	2,59	0,90	0,52	0,23	3,76	1,28

Table 3-3: Group 2 - Frequency Feature

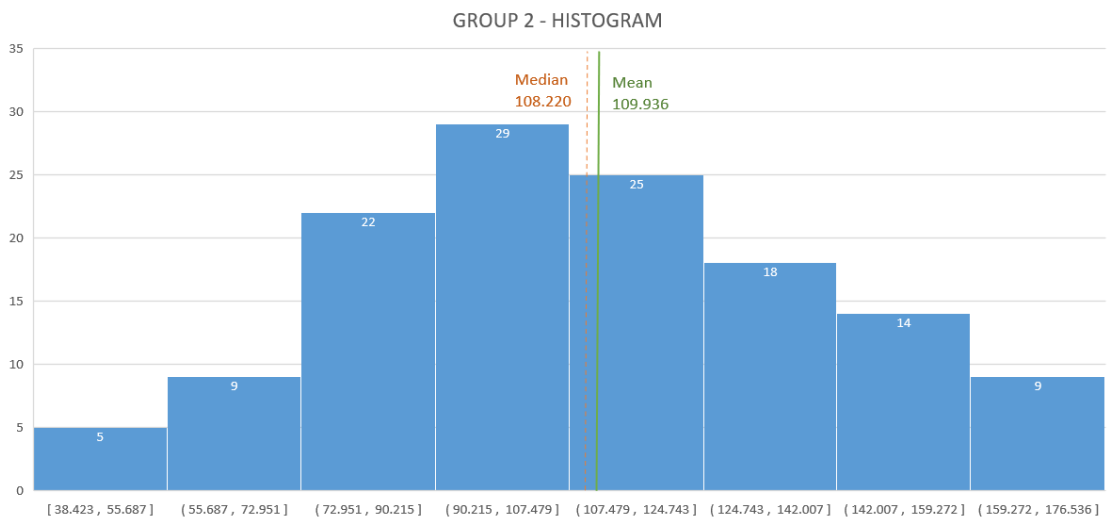

Figure 3-13: Histogram – Group 2

For individual groups, in accordance with features displayed in part 2.2.1., it is worth noting that against the fairly symmetrical distribution of Group 2 aggregated data, subgroups 202, 206 and 207 display distributions highly skewed (>1) to the left, meaning that the majority of observations fall below the distribution mean (Giannikos, Descriptive Statistics, 2020), furtherly denoting the effect of higher volume of sales in shorter time periods (Albright, Winston, & Zappe, 2011, p. 42).

Subgroups 201, 203, 204 and 205 (0.5 – 1) display moderate positive skewness (Giannikos, Descriptive Statistics, 2020).

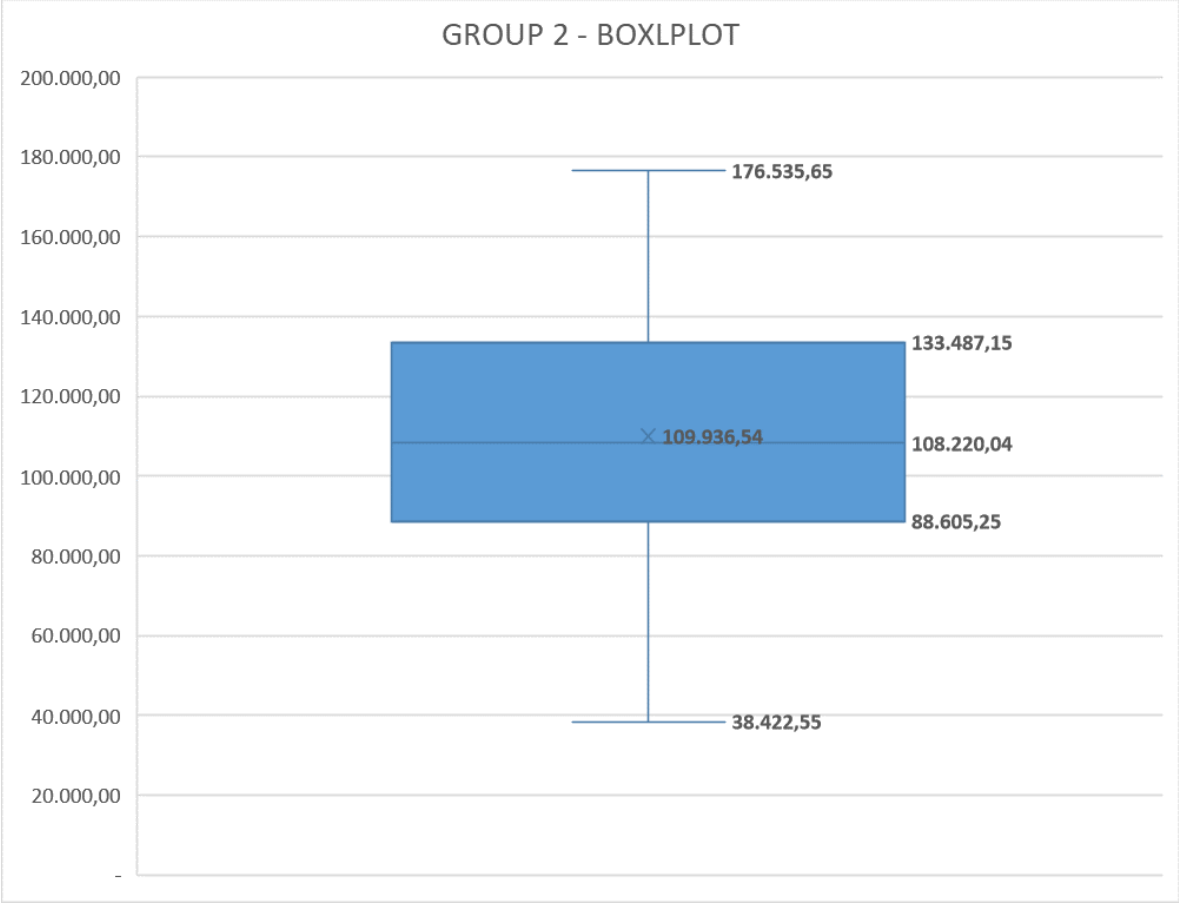

Figure 3-14: Boxplot - Group 2

3.2.3. SEASONALITY

Indications from Figure 3-1, as well as all other subgroup depiction analysis led to further exploration of cyclicality patterns in Group 2 sales evolution, and particularly

in seasonality effects. Figure 3-15 demonstrates the respective seasonal effect, through a boxplot that compares dispersion of observations in a monthly level.

It becomes clear that Group 2 sales do incorporate a seasonal trend, extremely evident from May to September each year. Indicatively, the interquartile range of May sales is located well above February's interquartile range as well as the mean value of sales from May to September stay well above the interquartile range of all the other months.

The contribution of seasonality to sales evolution will be thoroughly examined through regression analysis further on.

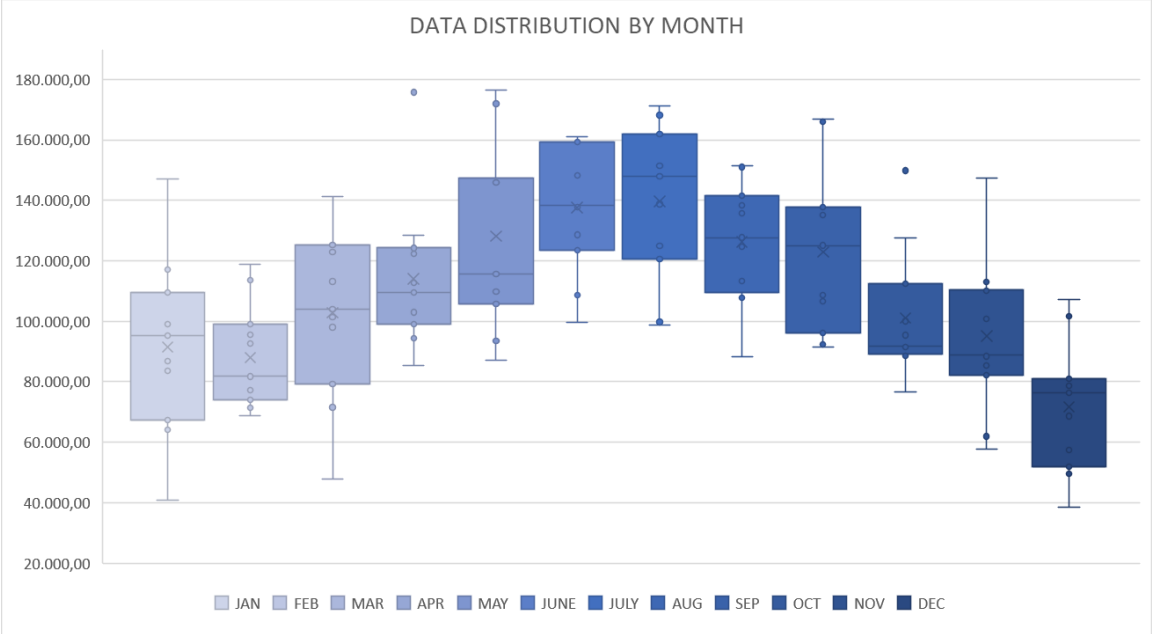

Figure 3-15: Monthly Boxplot - Group2

3.3. Regression

3.3.1. SEASONAL EFFECTS

At the aim of building a regression model that depicts the statistical properties and the patterns who contribute to sales formation, seasonality is the first factor incorporated in

the regression equation, since signs of its major impact are already provident from the conducted analysis. Through the use of dummy variables for each month of the year, regression results are presented in Table 3-4, at a confidence level 95%.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,327	0,080	142,285	0,000	11,170	11,485
LINEAR TREND	0,001	0,001	1,433	0,154	0,000	0,002
FEB	-0,008	0,102	-0,074	0,941	-0,210	0,195
MAR	0,124	0,102	1,217	0,226	-0,078	0,326
APR	0,249	0,102	2,433	0,016	0,046	0,451
MAY	0,357	0,102	3,491	0,001	0,154	0,559
JUNE	0,442	0,102	4,323	0,000	0,239	0,644
JYLY	0,451	0,102	4,413	0,000	0,249	0,653
AYG	0,354	0,102	3,461	0,001	0,151	0,556
SEP	0,317	0,102	3,100	0,002	0,114	0,519
OCT	0,125	0,102	1,218	0,226	-0,078	0,327
NOV	0,047	0,102	0,463	0,644	-0,155	0,250
DEC	-0,246	0,102	-2,408	0,018	-0,449	-0,044

Table 3-4: Regression results on Seasonality – 1st step

Insignificant factors with P-value above 0.05 (Linear Trend, February, March, October and November) are removed, so as to obtain the results presented in Table 3-5.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,436	0,032	353,534	0,000	11,372	11,500
APR	0,190	0,079	2,394	0,018	0,033	0,347
MAY	0,299	0,079	3,768	0,000	0,142	0,455
JUNE	0,384	0,079	4,851	0,000	0,228	0,541
JYLY	0,394	0,079	4,977	0,000	0,238	0,551
AYG	0,298	0,079	3,761	0,000	0,141	0,455
SEP	0,262	0,079	3,306	0,001	0,105	0,419
DEC	-0,299	0,079	-3,772	0,000	-0,456	-0,142

Table 3-5: Regression results on Seasonality - 2nd step

Regression statistics at the second step of seasonality exploration provide us with an 0.426 R Square, meaning that seasonal effects can explain approximately 43% of the variability in sales evolution (Thomaidis N. S., Special Forecasting Topics, 2021). The regression equation is

(3) $log\left(\widehat{S_t}\right) = 11.436 + 0.190 * APR + 0.299 * MAY + 0.384 * JUN + 0.394 * JUL + 0.298 * AUG + 0.262 * SEP - 0.299 * DEC$

3.3.2. PANDEMIC EFFECT

Proceeding with data exploration, WHO dummy variable is added in **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**, producing the following results:

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,466	0,033	342,681	0,000	11,399	11,532
APR	0,194	0,077	2,509	0,013	0,041	0,347
MAY	0,303	0,077	3,916	0,000	0,150	0,456
JUNE	0,389	0,077	5,025	0,000	0,236	0,542
JYLY	0,399	0,077	5,154	0,000	0,246	0,552
AYG	0,303	0,077	3,909	0,000	0,149	0,456
SEP	0,266	0,077	3,443	0,001	0,113	0,420
DEC	-0,294	0,077	-3,804	0,000	-0,448	-0,141
WHO	-0,124	0,047	-2,654	0,009	-0,216	-0,031

Table 3-6: Regression results on Seasonality & Pandemic (WHO variable)

Since all P-values are lower than α=95%, Equation 3 is transformed to:

(4) $log\left(\widehat{S_t}\right) = 11.466 + 0.194 * APR + 0.303 * MAY + 0.389 * JUN + 0.399 * JUL + 0.303 * AUG + 0.266 * SEP - 0.294 * DEC - 0.124 * WHO$

It is more than apparent that the declaration of the WHO with respect to COVID-19 being characterized as a pandemic, resulted in sales decline. Proceeding with other pandemic incidents able to constitute regression variables, the next regression attempt incorporates 1st lockdown imposition (Table 3-7).

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*

Constant Term	11,466	0,033	344,953	0,000	11,400	11,532
APR	0,178	0,078	2,294	0,023	0,024	0,331
MAY	0,287	0,078	3,699	0,000	0,133	0,440
JUNE	0,394	0,077	5,119	0,000	0,242	0,546
JYLY	0,404	0,077	5,249	0,000	0,252	0,556
AYG	0,308	0,077	3,997	0,000	0,155	0,460
SEP	0,271	0,077	3,528	0,001	0,119	0,424
DEC	-0,289	0,077	-3,761	0,000	-0,442	-0,137
WHO	-0,144	0,048	-3,006	0,003	-0,240	-0,049
1st LOCKDOWN	0,234	0,144	1,625	0,107	-0,051	0,520

Table 3-7: Regression results on Seasonality, Pandemic (WHO) & 1st Lockdown

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,467	0,034	339,846	0,000	11,400	11,534
APR	0,195	0,078	2,505	0,014	0,041	0,348
MAY	0,300	0,078	3,841	0,000	0,145	0,455
JUNE	0,386	0,078	4,940	0,000	0,231	0,541
JYLY	0,396	0,078	5,067	0,000	0,241	0,551
AYG	0,300	0,078	3,834	0,000	0,145	0,454
SEP	0,264	0,078	3,373	0,001	0,109	0,418
DEC	-0,294	0,078	-3,784	0,000	-0,448	-0,140
WHO	-0,117	0,051	-2,315	0,022	-0,218	-0,017
2nd LOCKDOWN	-0,036	0,108	-0,336	0,737	-0,251	0,178

Table 3-8: Regression results on Seasonality, Pandemic (WHO) & 2nd Lockdown

Apart from the p-value of 1st lockdown variable that could be marginally insignificant on a 90% level of confidence, it is worth noting the positive sign of its coefficient, being in sheer contrast with the observed negative pandemic effects. In any case the 1st lockdown effect will not be incorporated to the regression equation as the chosen level of confidence is 95% and the coefficients p-value is over 5%.

Reverting to the basic regression that incorporates only seasonal and WHO variables, the next regression attempt integrates the 2nd lockdown variable. 2nd lockdown effect is also questionable, since the lower value of the respective coefficient is lower than zero, the upper over zero and its p-value 0.737. Consequently, 2nd lockdown will not be incorporated in the regression as well.

Finalizing the examination of pandemic effects, the last attempt deals with the possible implications from vaccination approval and the launch of vaccination program in Greece. VAC variable is introduced in the regression ending up in the results in Table 3-9.

VAC coefficient is proved to be insignificant (p-value > α), while also affecting WHO coefficient by declining its significance. Since the pandemic effects analysis brought the most reliable results on the first level of exploration (Equation 4: Seasonality and Pandemic (WHO variable)) that incorporated WHO variable, the examination will proceed on that basis.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,467	0,033	342,519	0,000	11,401	11,533
APR	0,192	0,077	2,479	0,015	0,039	0,345
MAY	0,301	0,077	3,885	0,000	0,147	0,454
JUNE	0,387	0,077	4,994	0,000	0,233	0,540
JYLY	0,397	0,077	5,123	0,000	0,243	0,550
AYG	0,300	0,077	3,878	0,000	0,147	0,453
SEP	0,264	0,077	3,412	0,001	0,111	0,417
DEC	-0,297	0,077	-3,833	0,000	-0,450	-0,143
WHO	-0,060	0,078	-0,772	0,442	-0,214	0,094
VAC	-0,090	0,088	-1,020	0,310	-0,265	0,085

Table 3-9: Regression results on Seasonality, Pandemic (WHO) & Vaccination Program (VAC)

3.3.3. MACROECONOMIC EFFECTS

The regression results on Equation 4 produced an R square value 0.457. Since R square provides us with knowledge on the percentage of variation of sales that can be explained

by the regression line and the proposed explanatory variables (Albright, Winston, & Zappe, 2011, p. 558), further examination is necessary in order to discover other possible explanatory factors.

Firstly, the variable of Gross Domestic Product (GDP) is introduced, as an indicator of economic activity. It is worth noting that Hellenic Statistical Authority (ELSTAT) publicizes GDP data only in yearly and trimester (quarterly accounts) frequency on country level up to December 2022, as well as yearly GDP data on municipal level up to December 2022 (Hellenic Statistical Authority, Retrieved on April 29, 2023). Although the writers' aim was to explore GDP effect on a monthly municipal basis by aggregating Thessaloniki's, Chalkidiki's and Pieria's data (prefectures that Achtida is commercially active), the data finally used were quarterly GDP in current prices (Hellenic Statistical Authority, Retrieved on April 30, 2023). In order to match the observation frequency of actual sales (monthly) and constitute a comparable variable, the publicized GDP value for each trimester was copied to all three months belonging to it.

The same practice was applied during the examination of the tourism turnover index, whose evolution has been again recorded in a quarterly basis by ELSTAT (Hellenic Statistical Authority, Retrieved on May 1st). The produced regression results are demonstrated in Table 3-10 (GDP addition) and Table 3-11 (Tourism turnover index addition).

In the first case, by exploring GDP's contribution to sales evolution, the addition of GDP variable automatically constitutes WHO variable marginally insignificant. That fact, by using common sense, can be attributed to the parallel effects that WHO and GDP should have on sales. Indeed, pandemic initiation influences sales negatively, while GDP, whose value declines during the pandemic, affects sales negatively as well. Taking also into account that GDP's coefficient is almost zero, its addition to the regression equation will not provide the study with more explanatory evidence. Consequently, GDP variable is discarded over WHO variable.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	12,481	0,348	35,892	0,000	11,792	13,169
APR	0,183	0,075	2,439	0,016	0,035	0,332
MAY	0,292	0,075	3,887	0,000	0,143	0,441
JUNE	0,378	0,075	5,028	0,000	0,229	0,527
JYLY	0,393	0,075	5,234	0,000	0,245	0,542
AYG	0,297	0,075	3,952	0,000	0,148	0,446
SEP	0,261	0,075	3,472	0,001	0,112	0,410
DEC	-0,293	0,075	-3,898	0,000	-0,441	-0,144
WHO	-0,090	0,047	-1,934	0,055	-0,183	0,002
GDP (Million euros)	0,000	0,000	-2,931	0,004	0,000	0,000

Table 3-10: Regression results on Seasonality, Pandemic (WHO) & GDP

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,364	0,050	225,601	0,000	11,265	11,464
APR	0,125	0,080	1,557	0,122	-0,034	0,283
MAY	0,233	0,080	2,917	0,004	0,075	0,392
JUNE	0,319	0,080	3,990	0,000	0,161	0,478
JYLY	0,180	0,112	1,609	0,110	-0,042	0,402
AYG	0,084	0,112	0,749	0,455	-0,138	0,306
SEP	0,048	0,112	0,427	0,670	-0,174	0,270
DEC	-0,317	0,076	-4,163	0,000	-0,467	-0,166
WHO	-0,096	0,047	-2,044	0,043	-0,188	-0,003
TOURISM TURNOVER	0,002	0,001	2,642	0,009	0,000	0,003

Table 3-11: Regression results on Seasonality, Pandemic (WHO) & Tourism Turnover Index

On the other hand, the addition of Tourism Turnover variable has again similar effects, affecting the significance of the previously proposed monthly variables. As observed, Tourism Index introduction constitutes APR, JULY, AUG, SEP variable coefficients insignificant, providing no explanatory value by recreating the effect of identical forces, as Tourism Turnover definitely reflects the seasonal effect. In that way, Tourism

Turnover variable is rejected over the seasonal effects that monthly dummy variables have already illustrated.

3.3.4. SALES MOMENTUM

Proceeding with the examination of factors that can explain sales evolution and after the exploration of pandemic and macroeconomic effects, the influence of sales impetus is examined, using as base model the one built in Equation 4.

$$log\left(\widehat{S_t}\right) = 11.466 + 0.194 * APR + 0.303 * MAY + 0.389 * JUN + 0.399 * JUL + 0.303 * AUG + 0.266 * SEP - 0.294 * DEC - 0.124 * WHO$$

At that level, the regression model produced an R square at 45.7% and an adjusted R square at 0.421. Incorporating sales of the previous time period (previous month) as a lagged variable, is an attempt to enhance the explanatory value of the model and bring to Achtida's administration an alternative way to interpret sales evolution.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	6,402	0,787	8,136	0,000	4,844	7,960
APR	0,147	0,064	2,310	0,023	0,021	0,274
MAY	0,201	0,065	3,088	0,003	0,072	0,330
JUNE	0,238	0,067	3,559	0,001	0,106	0,371
JYLY	0,210	0,069	3,045	0,003	0,074	0,347
AYG	0,109	0,069	1,579	0,117	-0,028	0,247
SEP	0,116	0,067	1,734	0,085	-0,016	0,249
DEC	-0,310	0,064	-4,871	0,000	-0,436	-0,184
WHO	-0,092	0,039	-2,360	0,020	-0,169	-0,015
LAGGED LN SALES	0,444	0,069	6,461	0,000	0,308	0,579

Table 3-12: Regression results on Seasonality, Pandemic (WHO) & Lagged Sales - 1st Step

At the first step of exploration, as depicted in Table 3-12, Lagged Sales acquire a significant and dominant coefficient, demonstrating the repercussion of previous level of sales on next period. The variables that became insignificant are excluded from the model by proceeding to the next step. Table 3-13 illustrates the results, depicting the need to perform another regression round by excluding APR from the calculated variables.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	5,709	0,716	7,978	0,000	4,293	7,126
APR	0,117	0,063	1,862	0,065	-0,007	0,241
MAY	0,162	0,063	2,579	0,011	0,038	0,287
JUNE	0,193	0,064	3,021	0,003	0,067	0,319
JYLY	0,159	0,065	2,449	0,016	0,031	0,288
DEC	-0,336	0,063	-5,341	0,000	-0,461	-0,212
WHO	-0,084	0,039	-2,156	0,033	-0,161	-0,007
LAGGED LN SALES	0,506	0,062	8,181	0,000	0,384	0,629

Table 3-13: Regression results on Seasonality, Pandemic (WHO) & Lagged Sales - 2nd Step

The final results are presented in Table 3-14 and are used to constitute Equation 5.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	5,765	0,722	7,984	0,000	4,336	7,194
MAY	0,148	0,063	2,345	0,021	0,023	0,273
JUNE	0,179	0,064	2,793	0,006	0,052	0,306
JYLY	0,146	0,065	2,230	0,028	0,016	0,275
DEC	-0,351	0,063	-5,573	0,000	-0,476	-0,227
WHO	-0,083	0,039	-2,117	0,036	-0,162	-0,005
LAGGED LN SALES	0,503	0,062	8,047	0,000	0,379	0,626

Table 3-14: Regression results on Seasonality, Pandemic (WHO) & Lagged Sales - Final step

(5) $log\left(\widehat{S_t}\right) = 5{,}765 + 0{,}148 * MAY + 0.179 * JUN + 0{,}146 * JUL - 0.351 * DEC - 0.083 * WHO + 0{,}503 * LAGGED\ SALES$

3.3.4. LINEAR TREND

In section 2.3.1. SEASONAL EFFECTS the regression application led to the exclusion of linear trend from Equation 3, and subsequently from Equation 4 that incorporated both seasonal and pandemic variables.

In order to further appreciate the regression performance, a Goodness – of – Fit graph (Thomaidis N. S., Special Forecasting Topics, 2021) was created (Figure 3-16), which demonstrated a clear trend of sales overestimation up to August 2015, which then reversed to underestimation up to the end of the timeseries. Therefore, Linear Trend was re-inducted to the regression, acquiring significant coefficient that led to Equation 6:

$$(6)\ log\left(\widehat{S_t}\right) = 11.243 + 0.004 * t + 0.211 * APR + 0.316 * MAY + 0.397 * JUN + 0.402 * JUL + 0.302 * AUG + 0.261 * SEP - 0.313 * DEC - 0.419 * WHO$$

Following the application of Equation 6, Goodness – of – Fit graph (Figure 3-17) improved as residuals discarded the absolute pattern of overestimation and underestimation, providing a less repetitive image. Indicatively, it is worth noting that on Figure 3-16 depiction the Mean Absolute Error (MAE) (Thomaidis N. S., Simple Forecasting Techniques, 2021) is 18.750,60 and the Mean Absolute Percentage Error (MAPE) (Thomaidis N. S., Simple Forecasting Techniques, 2021) 18.46%. In the latter case of Linear Trend addition in Figure 3-17, MAE is 15.713,70 and MAPE 15.38%. Moreover, R square was boosted from 0.457 to 0.587 and adjusted R square from 0.421 to 0.557.

As data exploration proceeded, the re-induction of Linear Trend was considered appropriate, in order to check if it would provide extra explanatory value to Equation 5: Seasonality, Pandemic & Lagged Sales as well. Indeed, the R square 0.586 and adjusted R square 0. 566 from Equation 5 were improved to 0.611 and 0.589 respectively in Equation 7.

$$(7)\ log\left(\widehat{S_t}\right) = 6.523 + 0.002 * t + 0{,}161 * MAY + 0.198 * JUN + 0{,}170 * JUL - 0.367 * DEC - 0.225 * WHO + 0{,}428 * LAGGED\ SALES$$

As far as residual behavior is concerned, a 15.996,20 MAE and a 15.38% MAPE improved to 15.045,05 and 14.54% respectively. The results are depicted on the Goodness – of – Fit graphs in Figure 3-18 and Figure 3-19.

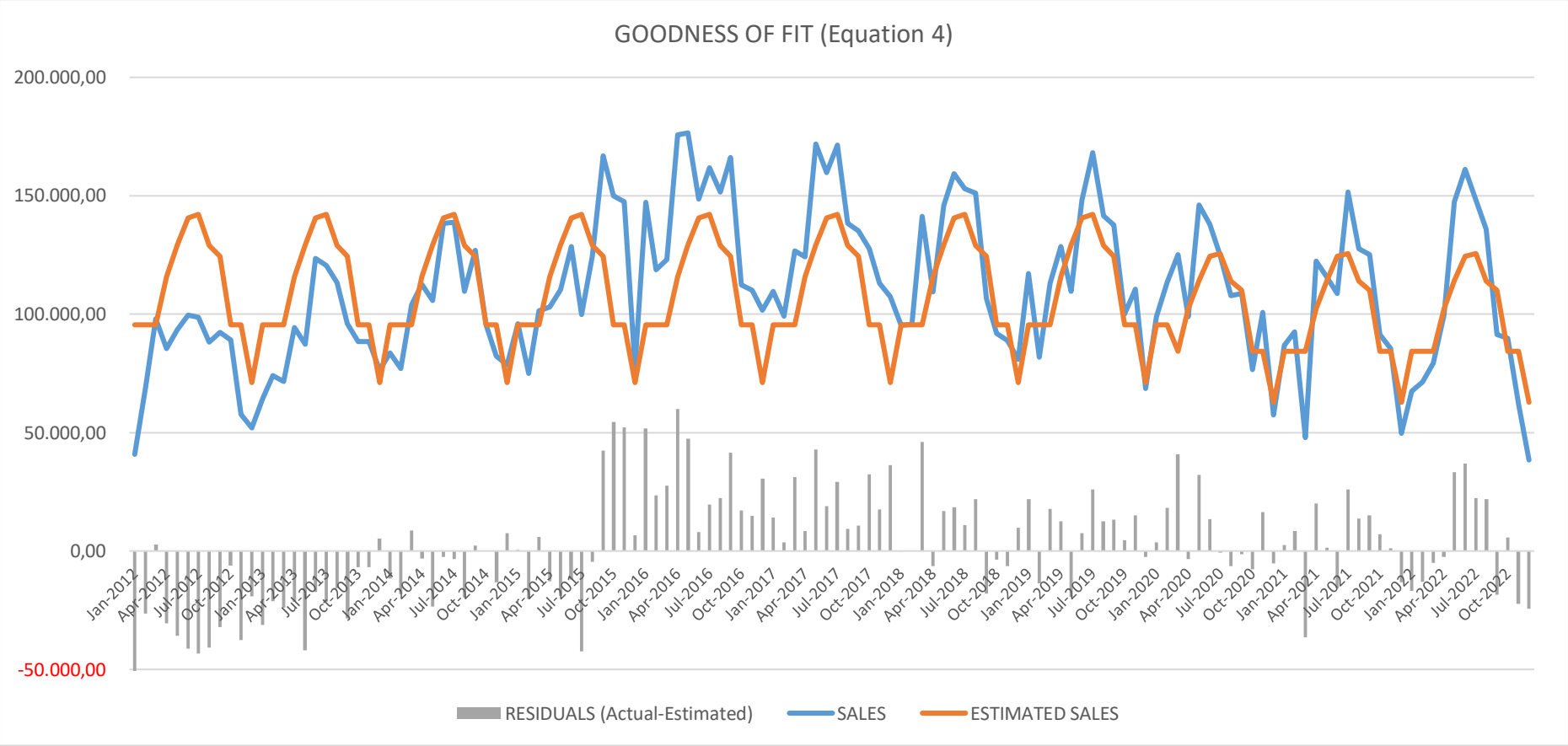

Figure 3-16: Regression performance on Equation 4

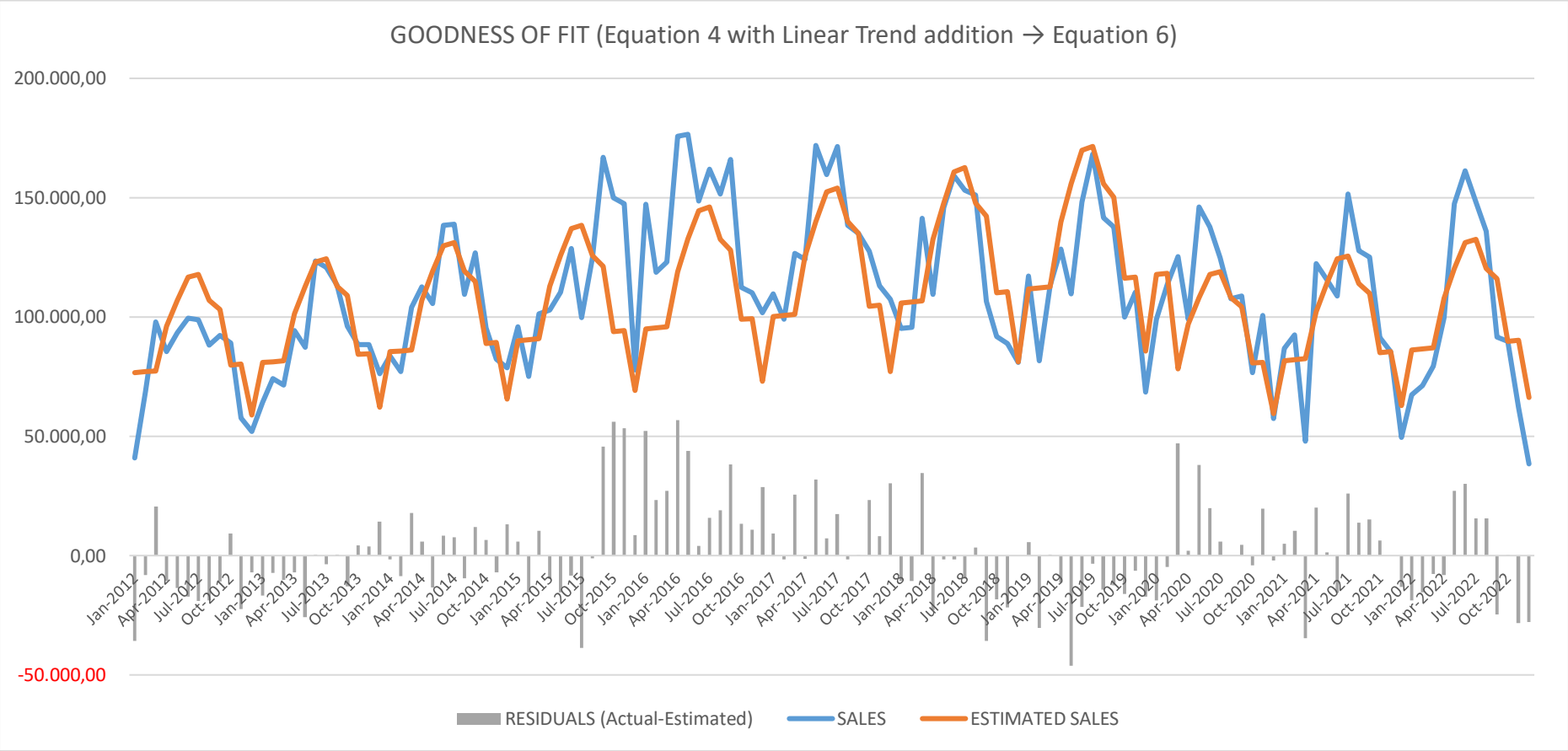

Figure 3-17: Regression performance on Equation 6

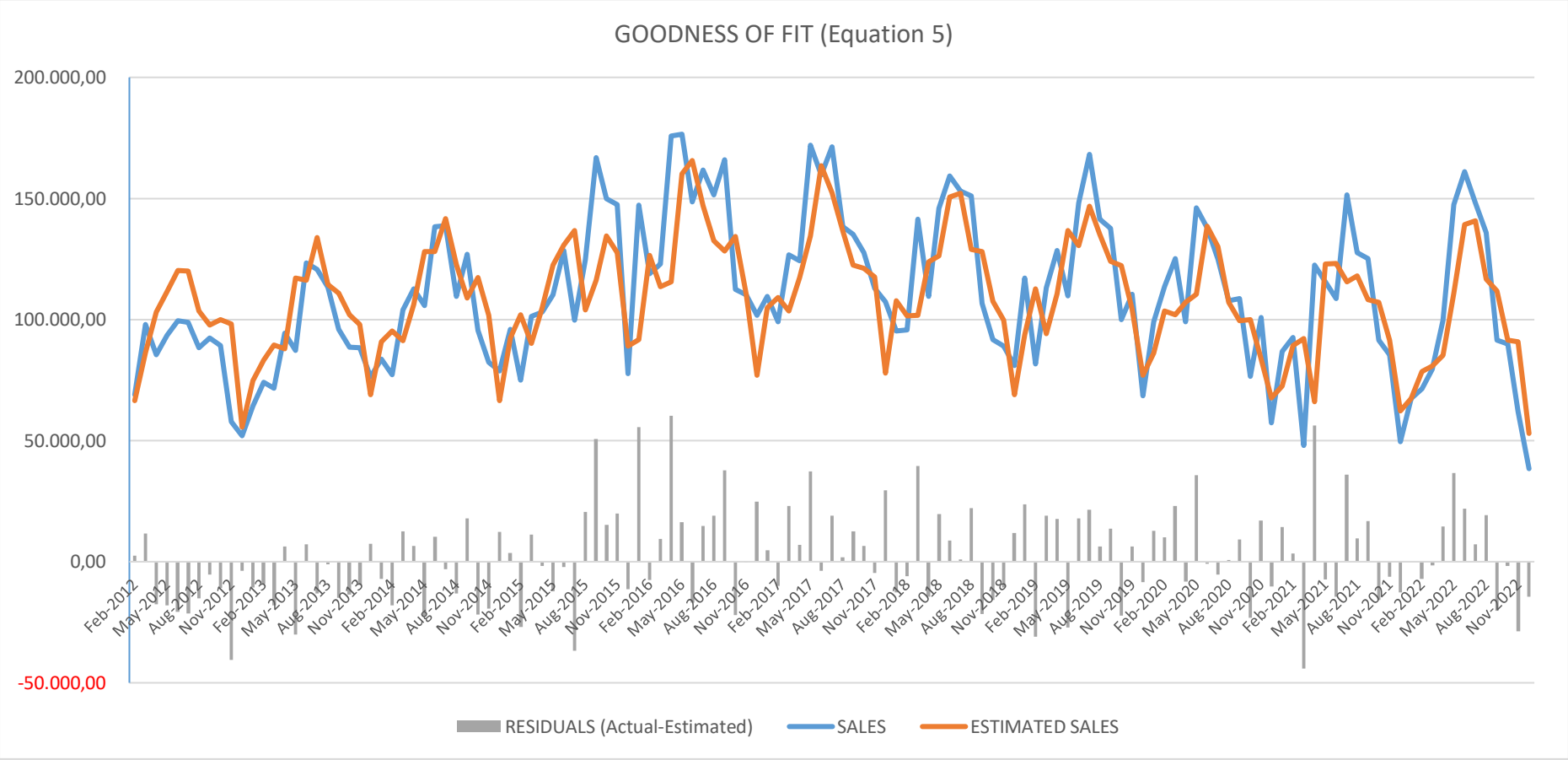

Figure 3-18: Regression performance on Equation 5

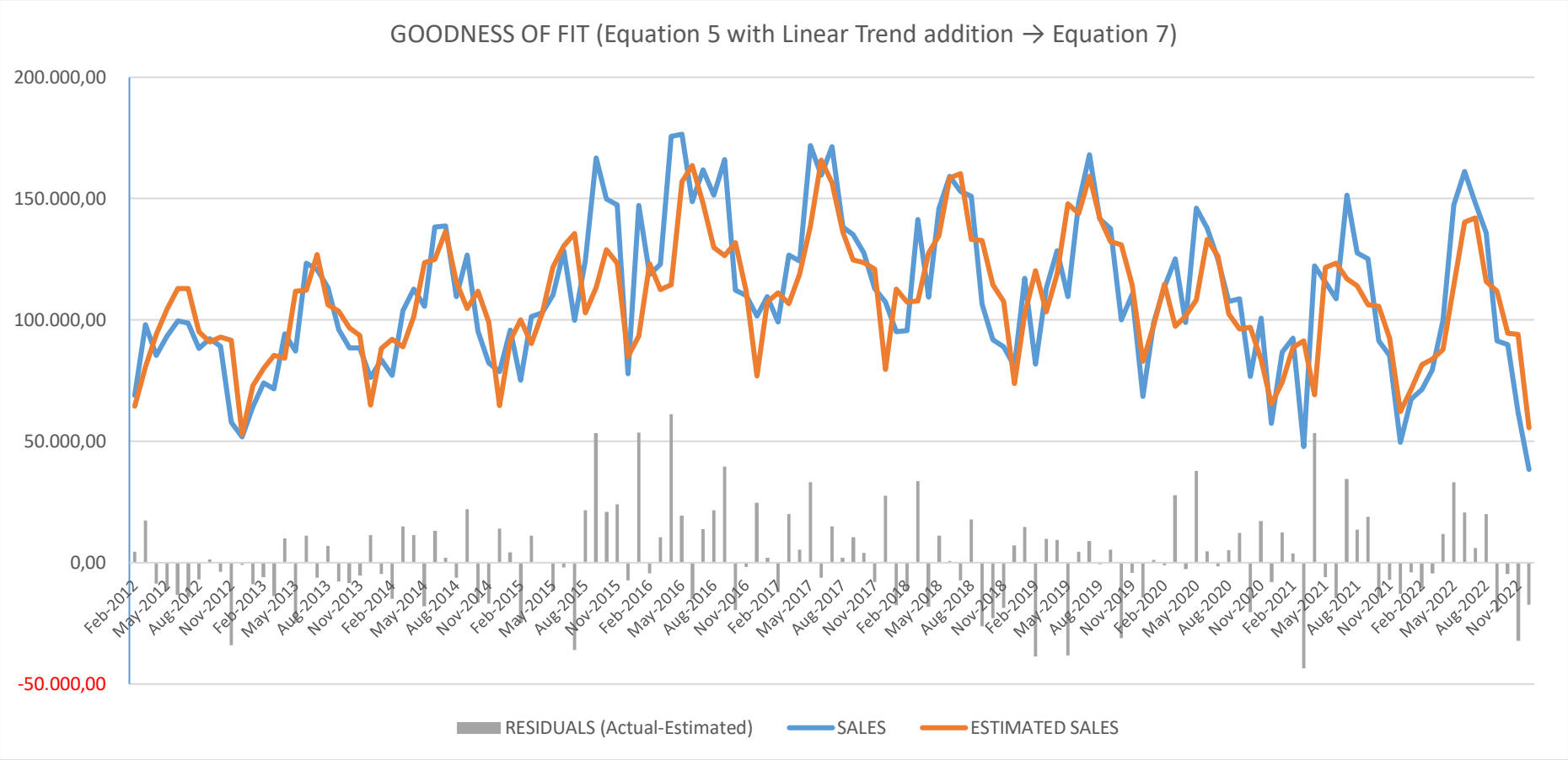

Figure 3-19: Regression performance on Equation 7

3.4. Concluding Remarks

The purpose of the present paper is primarily to explore and identify sales patterns and trends that characterize Achtida's sales over the past eleven complete years, from 2012 up to 2022. By examining the second Group of products, under the umbrella of Personal Hygiene and Cleanness, several observations arise.

The aggregated Group behavior is determined by an upward trend, extremely evident after 2015. Among the distinct subgroups the major contributors to Group 2 formation were Adult Diapers & Sanitary Napkins (204) and Hair and Body Cleanness (201), with average eleven – year contribution 43.32% and 20.22% respectively.

The timeseries revealed several differences among the subgroups' evolution. An upward trend was evident for 201, 202 (Deodorants), and 206 (Shaving, Cotton and Pharmacy), whereas subgroup 207 demonstrated also an increasing trend, attributed mainly to pandemic effects. On the other hand, subgroup 205 (Baby Care & Diapers) as well as 204 (Adult Diapers % Sanitary Napkins) presented a reverse evolution, with sales decline over time. Subgroup 203 (Mouth Hygiene) demonstrated minor deviations around a mean of 9.210€ monthly.

For the Group as a whole, a strong seasonal pattern (Figure 3-15) became evident, constituting April, June, July, August and September the months whose mean sales never fell under 110.000€. Seasonality was furtherly established during the regression rounds. As far as pandemic repercussions are concerned, neither lockdown impositions, nor the vaccination program initiation were proved to inflict direct effects on sales. Yet, the pandemic era itself, embodied in a dummy variable active from March 2020 when COVID-19 has been characterized as a pandemic by the WHO (World Health Organization, Retrieved on May 12th 2023), was proved to have significant, negative effects on sales. Either the seasonal, or the pandemic outbreak effect are becoming clearly evident through the graphs that decompose the timeseries (Thomaidis N. S., Special Forecasting Topics, 2021) based on Equation 6: Seasonality, Pandemic & Linear Trend and Equation 7: Seasonality, Pandemic, Lagged Sales & Linear Trend.

Aiming to explain sales evolution as deep as possible, it is worth to underline that, the better R square value achieved was 61.1% for Equation 7: Seasonality, Pandemic, Lagged Sales & Linear Trend and after that, 58.7% for Equation 6: Seasonality,

Pandemic & Linear Trend. Those equations remain both present in the study, although the latter essentially abolishes the former. The reason is that Equation 6, taking in account mainly factors external to the company's control, provides a tool for better assessment of external, independent forces that affected sales over time. On the other hand, Equation 7 which incorporates previous month's sales includes a feature that can be used for short term predictions (next month's sales) and can provide a basis for building a forecasting model.

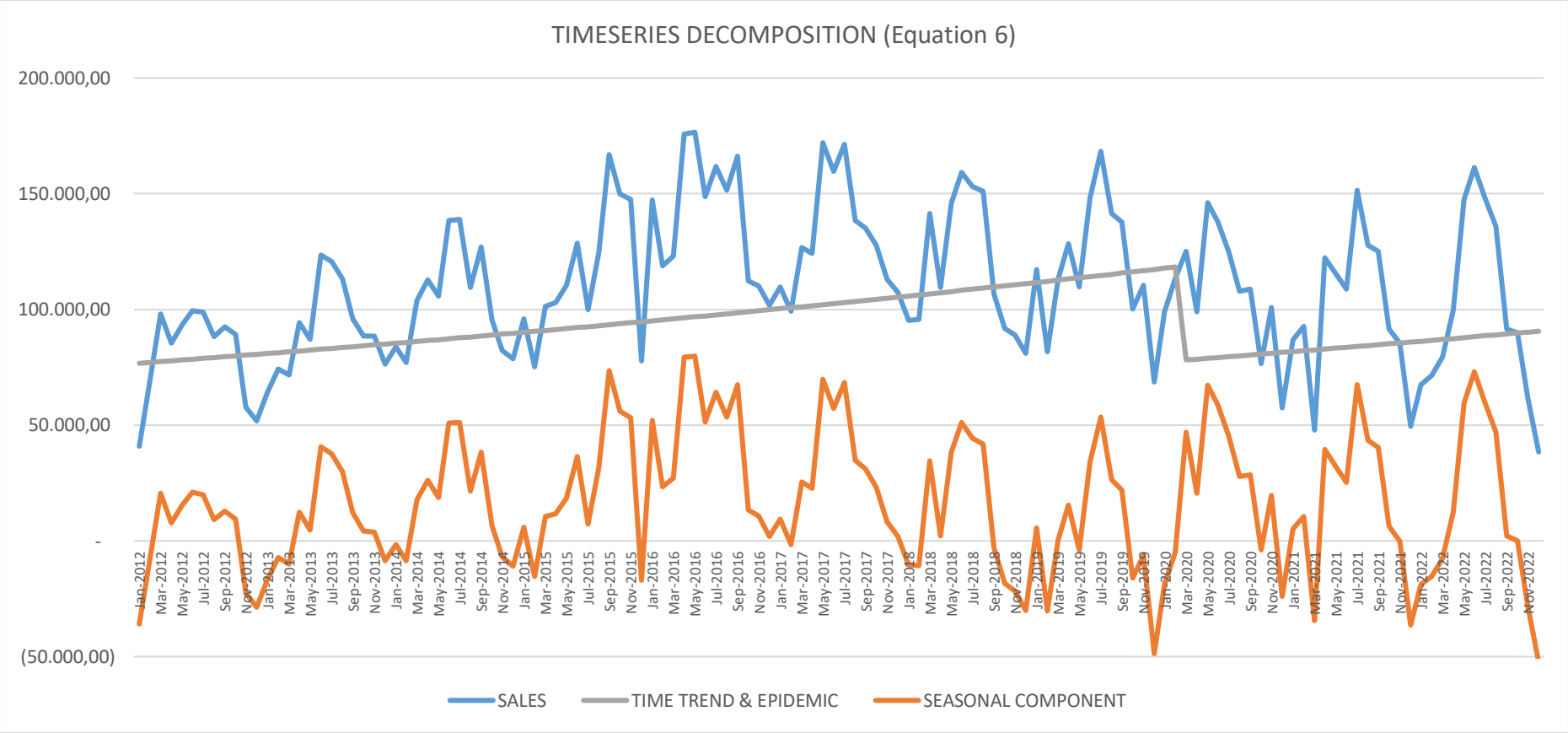

Figure 3-20: Timeseries Decomposition on Equation 6

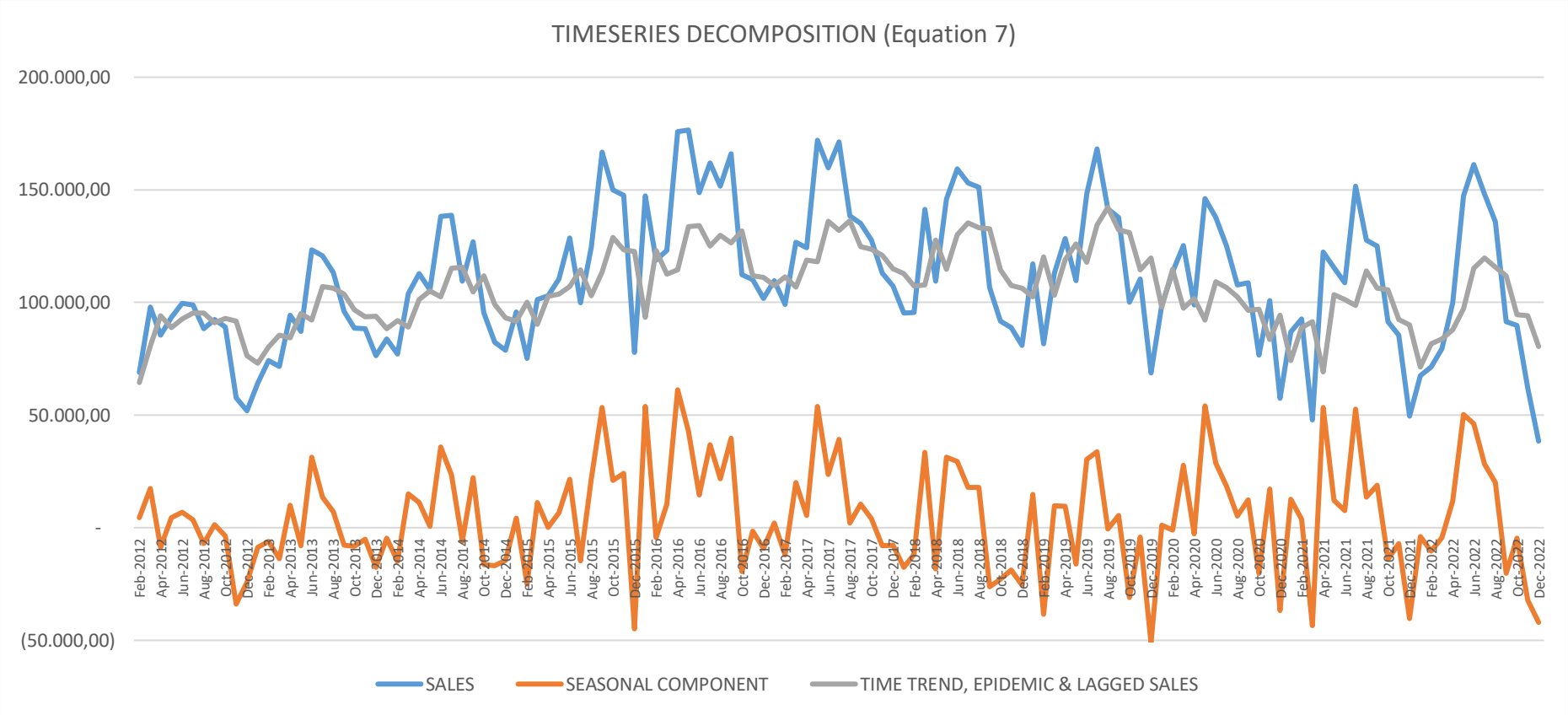

Figure 3-21: Timeseries Decomposition on Equation 7

4 GROUP – 3 HOUSEHOLD, DETERGENTS & PAPER PRODUCTS

4.1. Group Introduction

The third group of commodities incorporates several product categories pertinent to housekeeping activities. The products under that group, summarized under six distinct subgroups, all reply to the needs of household care, either through durable and repetitive, or by consumable goods. Group 3 aggregate sales value ranges from 40.647€ to 203.421 throughout the years under examination.

4.1.1. HOUSEHOLD EQUIPEMENT & CONSUMABLES (301)

The first subgroup consists of consumable goods such as metal food and baking containers, baking paper, aluminum foil, food-wrapping films, food bags, ice packs, garbage bags, as well as durable equipment like plastic food containers, sponges, mops, buckets, multipurpose gloves and other supplies with reference to housekeeping and cleaning. All of the above-mentioned commodities are marketed under the Sanitas brand name, commercialized in the Greek marketplace by Sarandis S.A. 301 subgroup accounts on average for 32.9% of Group 3 sales over the period 2012 – 2022, the greatest share among all subgroups.

4.1.2. INSECTICIDES & MOSQUITO REPELLENTS (302)

302 subgroup includes insecticides such as Teza and Pyrox sprays, foams and gels against crawling and flying insects, as well as Teza and Pyrox repellents (tablets, liquids, candles, spirals) against mosquitos and flies. A small portion of 302 sales is generated from body mosquito repellents, under the brand Sam Loves Betty.

Sarandis S.A. is the distributor of 302 products as well, which constitute on average 13.3% of the eleven-year Group sales.

4.1.3. DISWASHING (303)

Golgate Palmolive and Sarandis fill this subgroup with handwashing dish detergents such as Ajax Excel, Palmolive and Ava. 303 contribution on Group 3 sales formation was on average 11.9%

4.1.4. LAUNDRY (304)

The subgroup consists mainly of Soupline fabric softeners, Soflan detergent for delicate clothes (Colgate Palmolive), as well as Age Gentile washing boosters (Elgeka). Its share to Group 3 sales was 11.3% on average.

4.1.5. HOUSE & WC DETERGENTS (305)

Subgroup 305 is the second largest of Group 3, with average 22.9% contribution to sales and an extremely wide range of products: Indicatively, Fabuloso general cleaning liquid, Ajax glass, WC, grease or multipurpose cleaning liquids, Ajax Ultra, Ajax Kloron, Ajax FDF (Colgate Palmolive). Under Sarandis handling: Tuboflo clog removers, Afroso gels, blocks and WC liquids.

4.1.6. PAPER PRODUCTS (306)

The last subgroup encompasses paper consumable products (tissues, napkins, towel and toilet paper) handled by Essity Hellas, mainly under Zewa brand. Its contribution to aggregated Group 3 sales is the minor of all, 7.7%.

4.2. Data Exploration

In order to depict Group 3 timeseries characteristics as well as assist pattern and trend recognition, Figure 4-1 summarizes aggregate as well as distinct subgroup sales.

4.2.1. VISUAL OBSERVATIONS & DESCRIPTIVE FEATURES

The aggregate group sales demonstrate a considerable cyclical pattern, characterized by summer peaks and winter lows. Sales range from 40.647€ through to 203.421€, with a median (116.174€) very close to the observed mean monthly sales, on 119.442€. The standard deviation of sales is 38.146,75€. The intensity of seasonal effect seems to be lighter during the first half of the timeseries, contrary to the second half within which sales seem to acquire greater variations. Indicatively, the first 66 monthly observations are located in a range of 135.212€ with a 36.280 standard deviation, whereas sales from

July 2017 and on are placed in a 162.774€ range, with the respective deviation being 38.677€.

The timeseries demonstrates an upward trend, with average monthly change on 5.8% for the eleven-year period. Up to July 2017 the greatest level of sales accomplished was 184.120€, while for the latter part of it was 203.421€, on May 2022.

	GROUP 3	**301**	**302**	**303**	**304**	**305**	**306**
MINIMUM	**40.647,32**	14.591,23	8,96	1.819,45	2.489,04	8.446,68	633,54
MAXIMUM	**203.421,00**	110.580,43	81.530,37	52.420,69	31.126,91	53.936,04	33.057,07
RANGE	**162.773,68**	95.989,20	81.539,33	50.601,24	28.637,87	45.489,36	32.423,53
MEAN	**119.441,86**	39.348,62	15.867,79	14.171,32	13.513,94	27.365,13	9.175,05
MEDIAN	**116.174,43**	37.711,85	3.492,17	12.607,50	13.039,58	27.019,00	6.770,29
STANDARD DEVIATION	**38.146,75**	15.313,64	20.287,18	9.750,92	5.494,38	7.236,85	6.402,92

Table 4-1: Group 3 - Basic Descriptive Features

Proceeding to individual subgroup examination, Figure 4-2 up to Figure 4-7 present individual product group sales and their evolution over time.

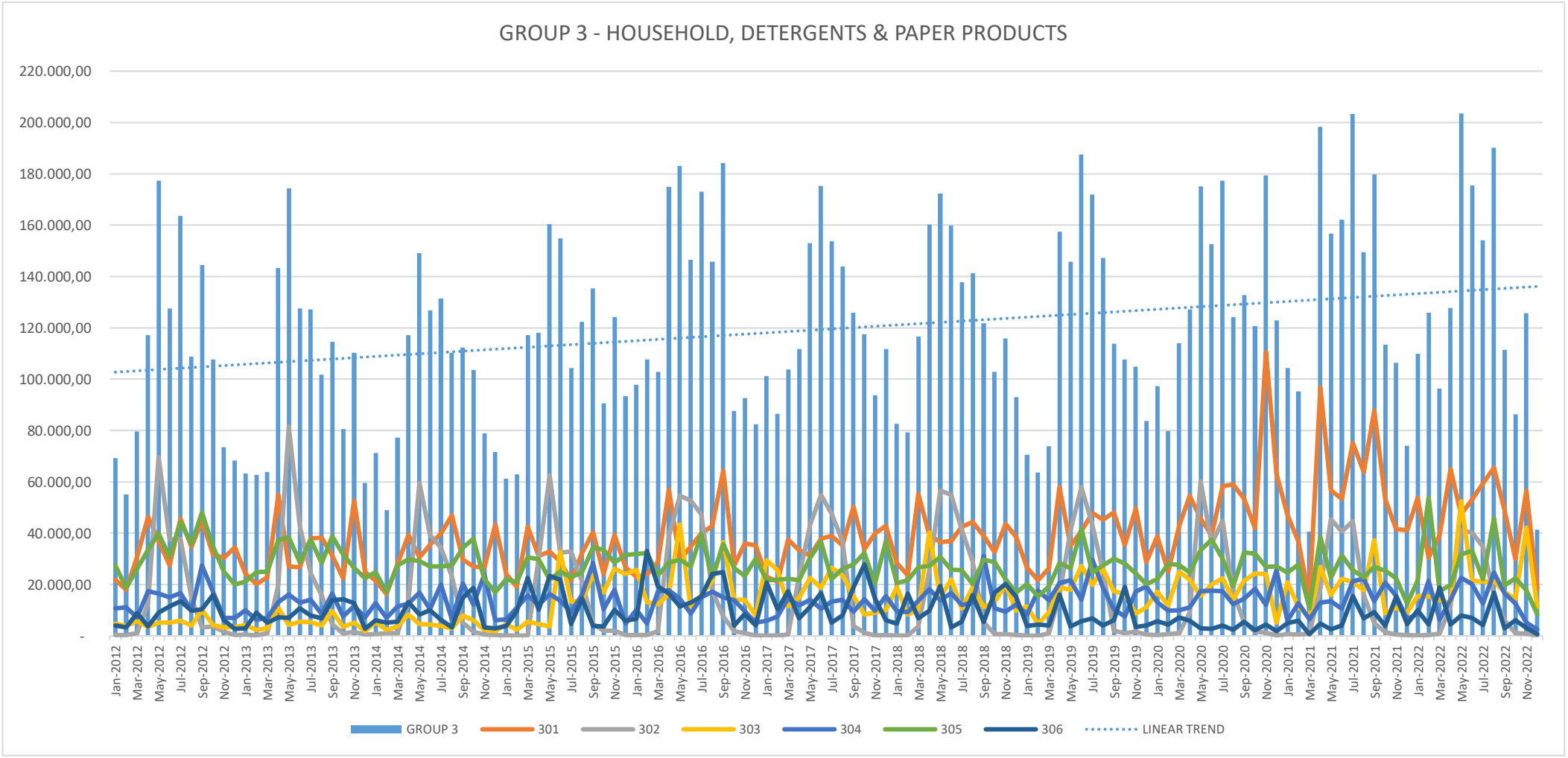

Figure 4-1: Group 3 sales evolution (2012-2022)

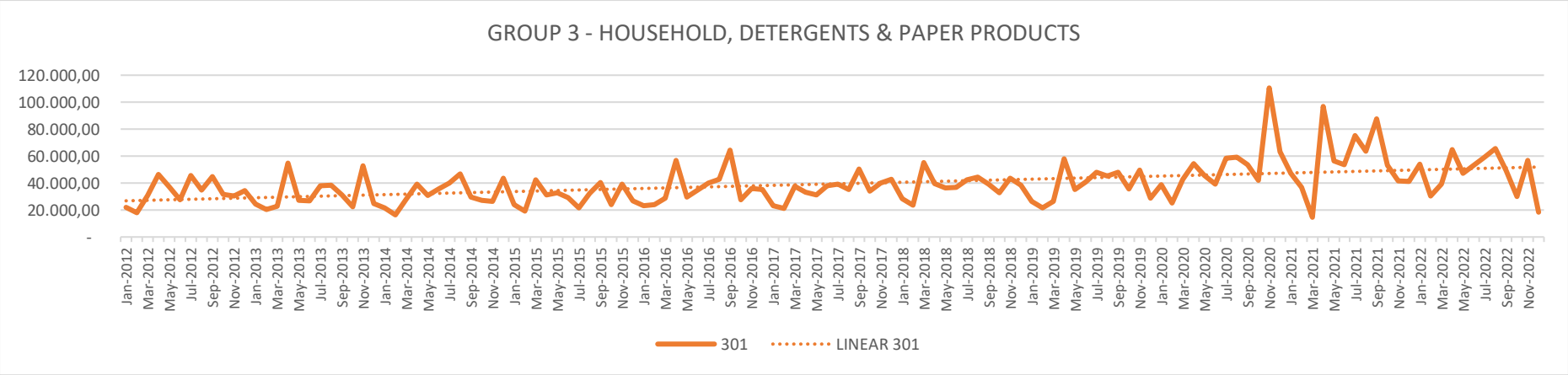

Figure 4-2: Subgroup 301 sales evolution (2012-2022)

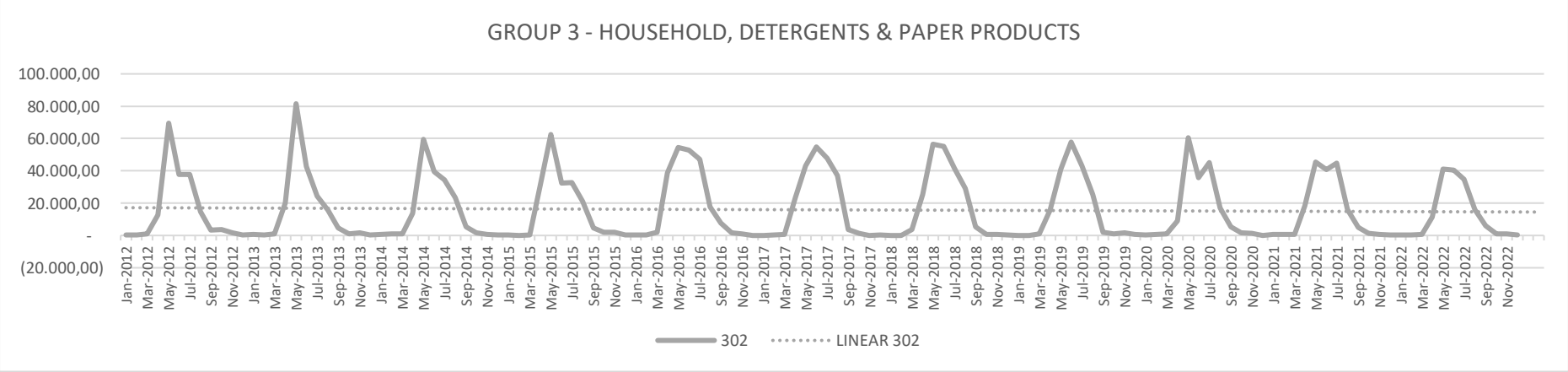

Figure 4-3: Subgroup 302 sales evolution (2012-2022)

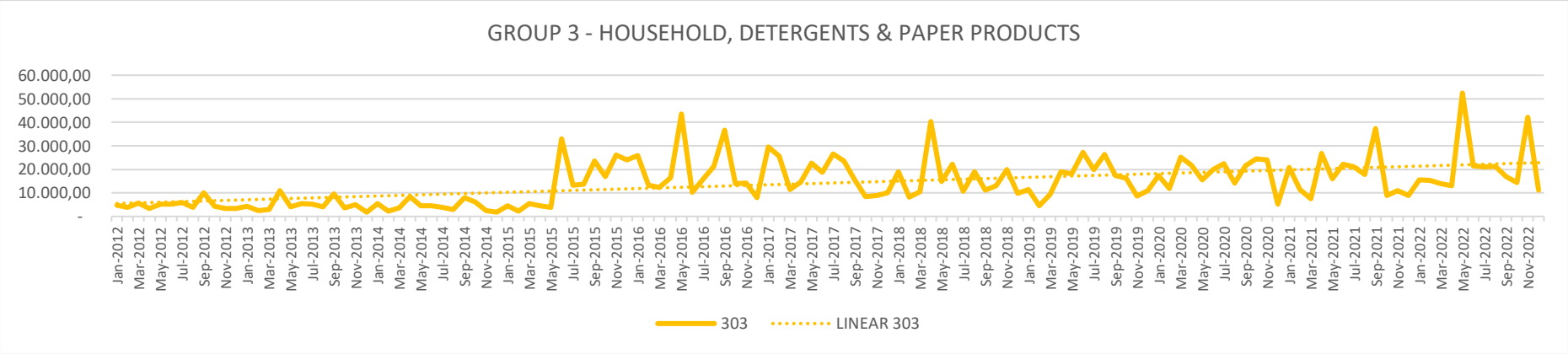

Figure 4-4: Subgroup 303 sales evolution (2012-2022)

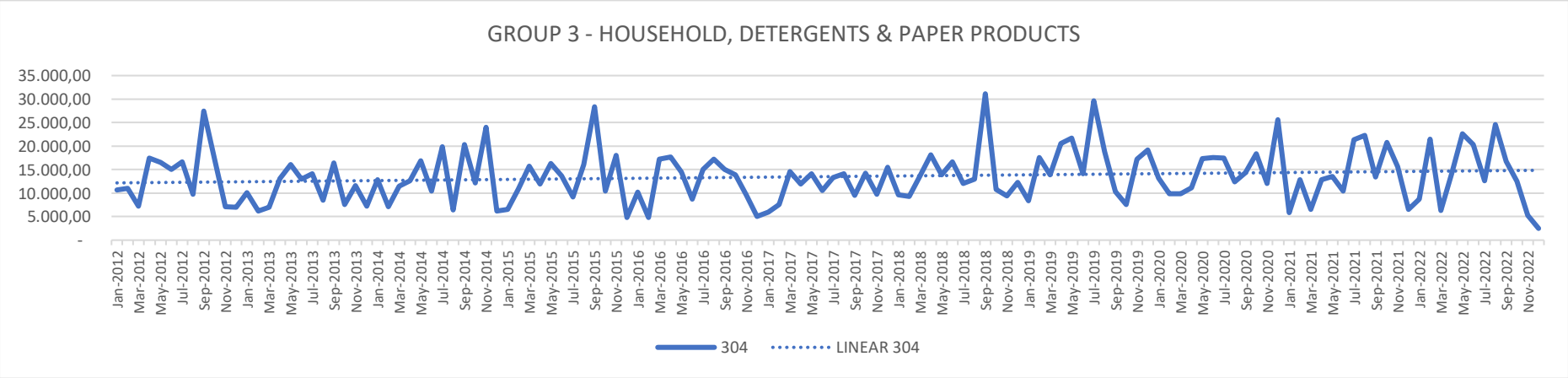

Figure 4-5: Subgroup 304 sales evolution (2012-2022)

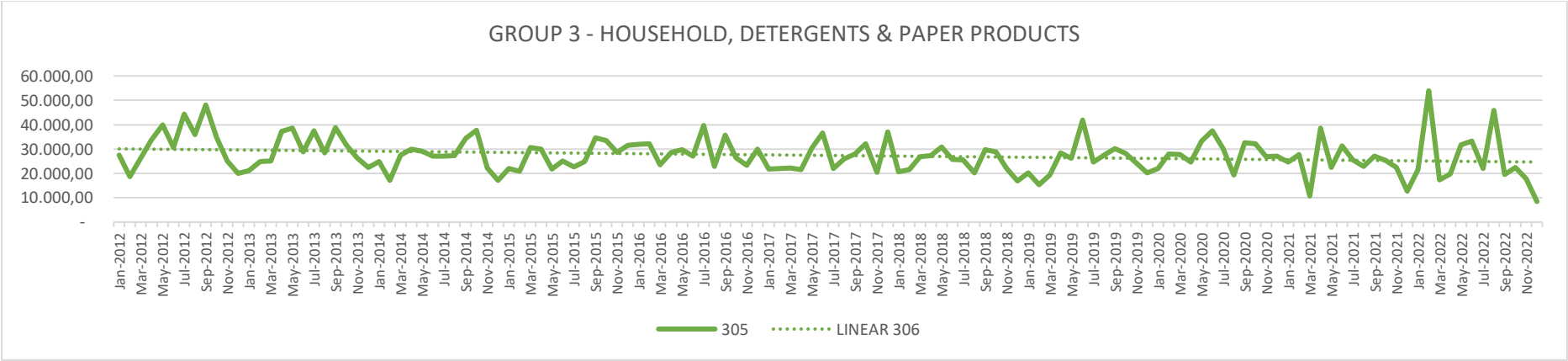

Figure 4-6: Subgroup 305 sales evolution (2012-2022)

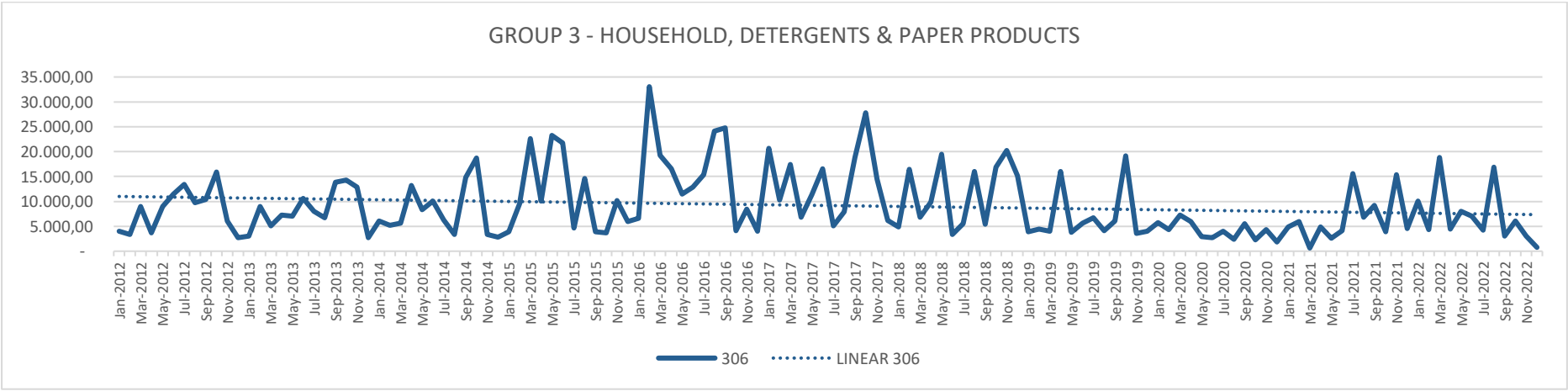

Figure 4-7: Subgroup 306 sales evolution (2012-2022)

Subgroup 301 (Household Equipment & Consumables) demonstrates an upward trend, having an average monthly change +10.6%. The highest value of sales is reached on November 2020 (110.580,43€). Subgroups' monthly sales fluctuate around a 39.348,62€ mean, within a 95.989,20€ range. Although sales spike in most of the cases on April, the seasonal effect demonstrated in Figure 4-1 aggregate sales evolution is not distinguishable.

On the contrary, subgroup 302 is by all means subject to seasonal forces, as presented in Figure 4-3. Indeed, 302 products demonstrate the higher standard deviation among all subgroups as compared to its sales range, not to mention that the vast majority of observations falls below mean sales (15.867,79€). Those facts underline subgroups clear dependence on tourist season and warm weather, a circumstance anticipated for insecticides and mosquito repellents. The overall trend heads downwards, as peak season sales diminish progressively. Higher sales are performed on May every year, except from two cases that occurred on June (2017 and 2019), mostly due to May weather conditions (Kouvrakis, 2023). Apart from that, 2020 is the first summer season in the time series with two distinguishable peaks, attributed to the low anticipation of retailers concerning proximate sales. Indeed, in 2020 and 2021 demand seemed to exceed early season supplies, with high volume orders being placed again in July (Kouvrakis, 2023).

The Dishwashing Detergents subgroup (303) performs a clearly upward trend, mainly triggered by the incorporation of AVA liquid detergents to Sarandis product mix. Up to that point (May 2015) mean 303 sales are 4.716.17€, whereas from June 2015 up December 2022 mean sales escalate to 18.431.34€. Throughout the timeseries, mean sales are located to 14.171€, with standard deviation at 9.750,92. Neither seasonal patterns, nor pandemic effects are evident.

Subgroup 304 (Laundry) demonstrates a moderately upward trend, having its mean sales around 13.514€. The standard deviation of observations is 5.494,38, within a relatively narrow range (28.637,87), the smallest among all subgroups. Subgroup 305 (House & WC Detergents) acquires a diminishing overall trend, having its mean sales at 27.365€. However, the range of observations becomes broader after 2020, encompassing the higher sales volume achieved, on 53.936€ in February 2022. At last, but not least, Paper Products (306) acquire a downward trend as well. Its sales range

from 633€ up to 33.057€ in February 2016, a time point nearly in the middle of the timeseries. Seasonal patterns and pandemic effects are not detected at the last three subgroups. Exploring the possible relationship among subgroup sales evolution, Table 4-2 demonstrates that none of the pairs displays a relationship worth noting.

Στήλη1	**301 vs 302**	**301 vs 303**	**301 vs 304**	**301 vs 305**	**301 vs 306**
CORRELATION	0,121	0,391	0,294	0,172	-0,034
		302 vs 303	**302 vs 304**	**302 vs 305**	**302 vs 306**
CORRELATION		0,187	0,275	0,288	0,043
			303 vs 304	**303 vs 305**	**303 vs 306**
CORRELATION			0,159	0,097	0,071
				304 vs 305	**304 vs 306**
CORRELATION				0,389	-0,009
					305 vs 306
CORRELATION					0,180

Table 4-2: Correlation of variance among Group 3 subgroups

4.2.2. FREQUENCY

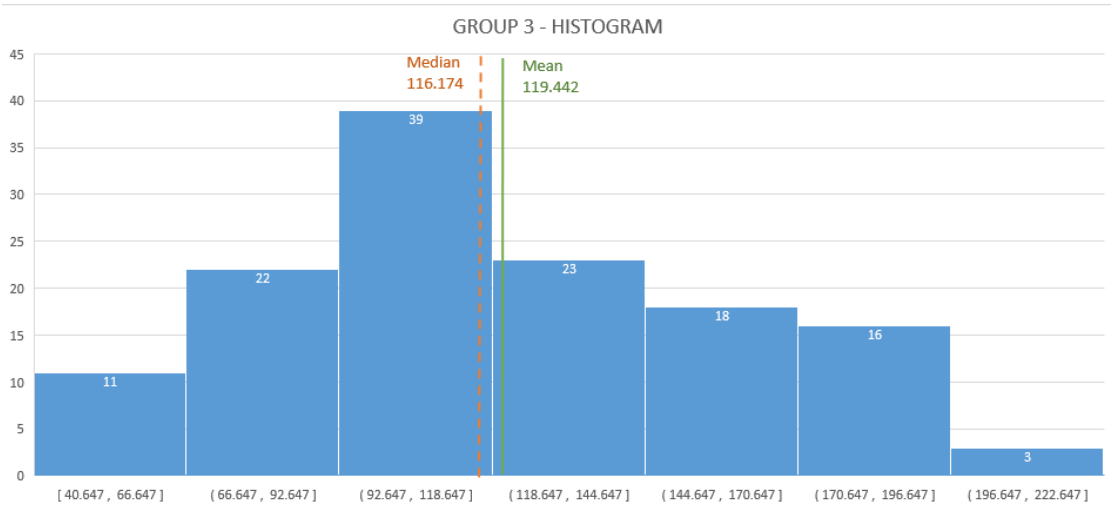

Figure 4-8: Histogram – Group 3

Group 3 observation distribution is quite symmetrical, with a slight positive skewness at 0.18. Approximately 55% of data points fall below the mean while kurtosis is -0.68, constituting the distribution platykurtic (Giannikos, Descriptive Statistics, 2020).

Among discrete subgroups data distribution is differentiated only for 301, being leptokurtic (4.19). In terms of skewness 304 and 305 are moderately skewed, whereas all other subgroups, having been more affected by outlier effects (high sales volume), are highly skewed to the right (Albright, Winston, & Zappe, 2011, p. 42).

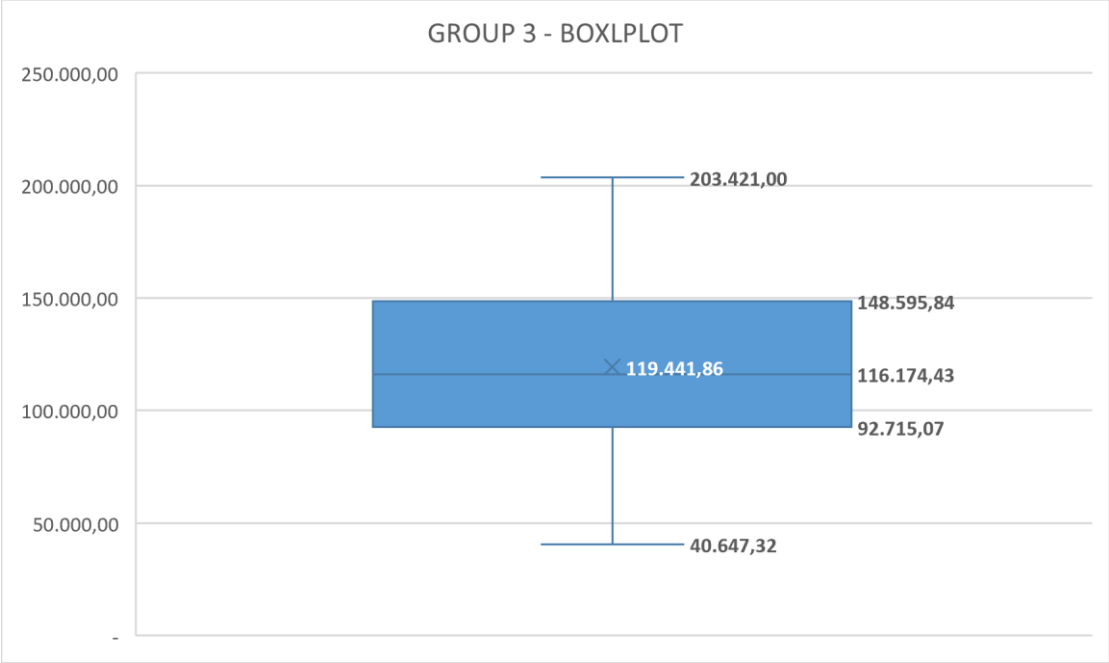

Figure 4-9: Boxplot – Group 3

	GROUP 3	301	302	303	304	305	306
MEAN	**119.441,86**	39.348,62	15.867,79	14.171,32	13.513,94	27.365,13	9.175,05
MEDIAN	**116.174,43**	37.711,85	3.492,17	12.607,50	13.039,58	27.019,00	6.770,29
SKEWNESS	**0,18**	1,52	1,12	1,14	0,70	0,62	1,15
KURTOSIS	**-0,68**	4,19	0,10	1,67	0,62	1,27	0,91

Table 4-3: Frequency Features - Group 3

4.2.3. SEASONALITY

Up to that point, signals of seasonal behavior became evident from the figures depicting Group 3 sales evolution (2012-2022). In order to better evaluate the magnitude of seasonal forces, Figure 4-9 summarizes each months' data distribution.

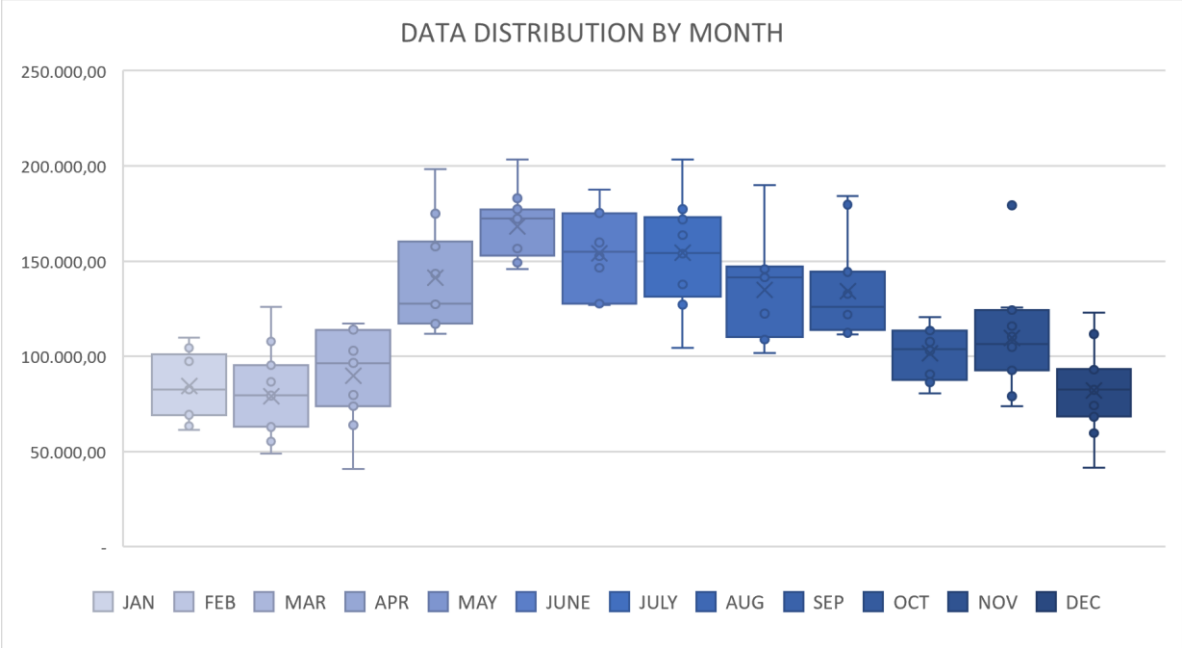

Figure 4-10: Monthly Boxplot – Group 3

Indeed, from April up to September mean sales remain steadily over 130.000€. Additionally, for that particular time frame, the 25th percentile of observations remains almost over than the 75th percentile of sales of the rest of months, with November being the only exemption, on which outliers are also demonstrated.

4.3. Regression

4.3.1. SEASONAL EFFECTS

Seasonality has been an evident factor affecting sales formation up to now and Group 3 has been proved to be influenced as well. The extent to which seasonal forces shape sales volume will be explored at initial regression rounds. Apart from the linear trend,

monthly dummy variables are inserted to regression equation, producing the results in Table 4-4. Since insignificant variables emerge (FEB, MAR, DEC), Table 4-5 summarizes the final results of seasonality analysis.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,195	0,067	166,362	0,000	11,062	11,329
Linear Trend	0,002	0,000	4,473	0,000	0,001	0,003
FEB	-0,087	0,086	-1,010	0,315	-0,258	0,084
MAR	0,035	0,086	0,406	0,686	-0,136	0,206
APR	0,513	0,086	5,939	0,000	0,342	0,684
MAY	0,698	0,086	8,083	0,000	0,527	0,869
JUNE	0,605	0,086	7,006	0,000	0,434	0,776
JYLY	0,597	0,086	6,911	0,000	0,426	0,768
AYG	0,461	0,086	5,340	0,000	0,290	0,632
SEP	0,453	0,086	5,241	0,000	0,282	0,624
OCT	0,180	0,086	2,087	0,039	0,009	0,351
NOV	0,234	0,086	2,702	0,008	0,062	0,405
DEC	-0,070	0,086	-0,806	0,422	-0,241	0,102

Table 4-4: Regression results on Seasonality – 1st step

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,166	0,043	261,751	0,000	11,082	11,251
Linear Trend	0,002	0,000	4,444	0,000	0,001	0,003
APR	0,543	0,068	7,970	0,000	0,408	0,678
MAY	0,728	0,068	10,687	0,000	0,594	0,863
JUNE	0,636	0,068	9,324	0,000	0,501	0,770
JYLY	0,627	0,068	9,204	0,000	0,492	0,762
AYG	0,492	0,068	7,214	0,000	0,357	0,627
SEP	0,483	0,068	7,089	0,000	0,348	0,618
OCT	0,211	0,068	3,093	0,002	0,076	0,346
NOV	0,264	0,068	3,872	0,000	0,129	0,399

Table 4-5: Regression results on Seasonality - 2nd step

At that point, regression statistics generate an R Square at 0.681, implying that the model incorporating time trend and seasonality alone is responsible for approximately 68% of sales variation (Albright, Winston, & Zappe, 2011, p. 557). Equation 8 portrays the respective coefficients.

$$(8)\ log\left(\widehat{S_t}\right) = 11.166 + 0.002 + 0.543 * APR + 0.728 * MAY + 0.636 * JUN + 0.627 * JUL + 0.492 * AUG + 0.483 * SEP + 0.211 * OCT + 0.264 * NOV$$

4.3.2. PANDEMIC EFFECT

The pandemic effect on sales is tested for Group 3 with several variables, as in aforementioned product groups. The regression results are presented on Table 4-6 up to Table 4-9, in which, in a 95% significance level none of the pandemic variables was proved to have a notable result in sales formation. However, although VAC variable could be incorporated to the model in a 90% significance level (since 0.05 < p-value < 0.10) its negative sign demonstrates effects against the anticipated sales rise.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,155	0,048	234,695	0,000	11,061	11,249
Linear Trend	0,002	0,001	3,299	0,001	0,001	0,004
APR	0,545	0,068	7,963	0,000	0,409	0,680
MAY	0,730	0,068	10,669	0,000	0,594	0,865
JUNE	0,637	0,068	9,309	0,000	0,501	0,772
JYLY	0,628	0,068	9,187	0,000	0,493	0,764
AYG	0,492	0,068	7,200	0,000	0,357	0,628
SEP	0,484	0,068	7,072	0,000	0,348	0,619
OCT	0,211	0,068	3,082	0,003	0,075	0,346
NOV	0,264	0,068	3,855	0,000	0,128	0,399
WHO	-0,033	0,062	-0,540	0,591	-0,156	0,089

Table 4-6: Group 3 - Regression results on Seasonality & Pandemic (WHO)

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,167	0,043	260,564	0,000	11,082	11,251
Linear Trend	0,002	0,000	4,344	0,000	0,001	0,003
APR	0,541	0,069	7,850	0,000	0,405	0,678
MAY	0,726	0,069	10,538	0,000	0,590	0,863
JUNE	0,636	0,068	9,290	0,000	0,501	0,772
JYLY	0,628	0,068	9,171	0,000	0,493	0,764
AYG	0,493	0,069	7,190	0,000	0,357	0,628
SEP	0,484	0,069	7,066	0,000	0,349	0,620
OCT	0,212	0,069	3,089	0,002	0,076	0,347
NOV	0,265	0,069	3,865	0,000	0,129	0,401
1st LOCKDOWN	0,030	0,123	0,248	0,805	-0,213	0,273

Table 4-7: Group 3 - Regression results on Seasonality & 1st Lockdown

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,166	0,043	261,917	0,000	11,082	11,250
Linear Trend	0,002	0,000	4,017	0,000	0,001	0,003
APR	0,543	0,068	7,974	0,000	0,408	0,678
MAY	0,737	0,069	10,746	0,000	0,601	0,873
JUNE	0,644	0,069	9,392	0,000	0,509	0,780
JYLY	0,637	0,069	9,273	0,000	0,501	0,772
AYG	0,501	0,069	7,297	0,000	0,365	0,637
SEP	0,493	0,069	7,173	0,000	0,357	0,629
OCT	0,220	0,069	3,207	0,002	0,084	0,356
NOV	0,265	0,068	3,887	0,000	0,130	0,400
2nd LOCKDOWN	0,096	0,090	1,077	0,284	-0,081	0,274

Table 4-8: Group 3 - Regression results on Seasonality & 2nd Lockdown

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,140	0,045	246,876	0,000	11,051	11,229
Linear Trend	0,003	0,001	4,457	0,000	0,002	0,004
APR	0,544	0,068	8,035	0,000	0,410	0,677
MAY	0,728	0,068	10,763	0,000	0,594	0,862
JUNE	0,634	0,068	9,379	0,000	0,501	0,768
JYLY	0,626	0,068	9,247	0,000	0,492	0,760
AYG	0,489	0,068	7,231	0,000	0,355	0,623
SEP	0,480	0,068	7,094	0,000	0,346	0,614
OCT	0,207	0,068	3,057	0,003	0,073	0,341
NOV	0,260	0,068	3,832	0,000	0,125	0,394
VAC	-0,103	0,061	-1,689	0,094	-0,224	0,018

Table 4-9: Group 3 - Regression results on Seasonality & VAC

4.3.3. MACROECONOMIC EFFECTS

Proceeding with microeconomic factors that might have affected sales evolution, GDP and LAGGED GDP variables are tested in Table 4-10 and Table 4-11 respectively.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,670	0,303	38,574	0,000	11,071	12,269
Linear Trend	0,002	0,000	4,743	0,000	0,001	0,003
APR	0,539	0,068	7,959	0,000	0,405	0,673
MAY	0,724	0,068	10,692	0,000	0,590	0,858
JUNE	0,631	0,068	9,316	0,000	0,497	0,765
JYLY	0,625	0,068	9,237	0,000	0,491	0,759
AYG	0,489	0,068	7,229	0,000	0,355	0,623
SEP	0,481	0,068	7,101	0,000	0,347	0,615
OCT	0,212	0,068	3,127	0,002	0,078	0,346
NOV	0,265	0,068	3,910	0,000	0,131	0,399
GDP (Million euros)	0,000	0,000	-1,683	0,095	0,000	0,000

Table 4-10: Group 3 - Regression results on Seasonality & GPD

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,781	0,311	37,900	0,000	11,166	12,397
Linear Trend	0,002	0,000	4,721	0,000	0,001	0,003
APR	0,544	0,068	8,027	0,000	0,410	0,678
MAY	0,724	0,068	10,686	0,000	0,590	0,859
JUNE	0,631	0,068	9,313	0,000	0,497	0,766
JYLY	0,623	0,068	9,190	0,000	0,489	0,757
AYG	0,490	0,068	7,236	0,000	0,356	0,625
SEP	0,482	0,068	7,109	0,000	0,348	0,616
OCT	0,209	0,068	3,085	0,003	0,075	0,343
NOV	0,267	0,068	3,932	0,000	0,132	0,401
LAGGED GDP	0,000	0,000	-1,995	0,048	0,000	0,000

Table 4-11: Group 3 - Regression results on Seasonality & Lagged GPD

Lagged GDP is tested based on the assumption that previous period economic conditions could have affected current period sales. It should be noted again that GDP data are only available on quarterly accounts (Hellenic Statistical Authority, Retrieved on April 30, 2023), so in order to match the observation frequency of actual sales (monthly) and constitute a comparable variable, the publicized GDP value for each trimester was copied to all three months belonging to it. According to Table 4-10, GDP does not acquire a significant coefficient in a 95% significance level. On the other hand, Table 4-11 demonstrates a Lagged GDP coefficient that could be part of the regression equation in a 95% significance level. However, the coefficient itself (-0,0000139) provides no explanatory value to the model. Consequently, none of the GDP variables is used.

Next attempts on sales evolution interpretation concentrate on tourism aspects, especially tourism turnover index and overnight stops. As presented in Table 4-12, tourism turnover on country level is not a significant sales driver for Group 3 products, as its p-value falls far beyond the chosen confidence level.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,171	0,050	224,077	0,000	11,072	11,270
Linear Trend	0,002	0,000	4,410	0,000	0,001	0,003
APR	0,548	0,073	7,534	0,000	0,404	0,692
MAY	0,733	0,073	10,085	0,000	0,589	0,877
JUNE	0,640	0,073	8,809	0,000	0,496	0,784
JYLY	0,642	0,101	6,364	0,000	0,442	0,841
AYG	0,506	0,101	5,021	0,000	0,306	0,705
SEP	0,498	0,101	4,940	0,000	0,298	0,697
OCT	0,213	0,069	3,081	0,003	0,076	0,349
NOV	0,266	0,069	3,852	0,000	0,129	0,402
TOURISM TURNOVER	0,000	0,001	-0,191	0,849	-0,001	0,001

Table 4-12: Group 3 - Regression results on Seasonality & TTI

Overnight stops constitute a monthly numerical variable that has been constructed based on data ELSTAT provided upon request, on prefectural level for Central Macedonia area (Hellenic Statistical Authority, 2023).

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,140	0,044	250,452	0,000	11,051	11,228
Linear Trend	0,002	0,000	4,995	0,000	0,002	0,003
APR	0,554	0,067	8,279	0,000	0,421	0,686
MAY	0,698	0,073	9,497	0,000	0,552	0,843
JUNE	0,588	0,097	6,045	0,000	0,395	0,781
JYLY	0,574	0,123	4,663	0,000	0,330	0,818
AYG	0,399	0,130	3,064	0,003	0,141	0,657
SEP	0,468	0,095	4,933	0,000	0,280	0,656
OCT	0,218	0,070	3,103	0,002	0,079	0,357
NOV	0,251	0,066	3,777	0,000	0,119	0,383
OVERNIGHT	0,000	0,000	0,555	0,580	0,000	0,000

Table 4-13: Group 3 - Regression results on Seasonality & Overnights

The data used are the aggregate overnight stays for Chalkidiki, Thessaloniki and Pieria prefectures, locations that Achtida S.A. is commercially active. However, since data for 2022 were not provided, the regression was performed incorporating sales up to December 2021. The results were again worthless, since Overnight p-value was 0.580.

4.3.4. SALES MOMENTUM

The last attempt to enhance the explanatory value of the model is accomplished through the examination of previous months' sales effect on current sales. Apparently, regression results demonstrate a not significant p-value.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,310	1,042	10,854	0,000	9,247	13,373
Linear Trend	0,002	0,001	3,971	0,000	0,001	0,003
APR	0,543	0,069	7,877	0,000	0,406	0,679
MAY	0,734	0,082	8,978	0,000	0,572	0,896
JUNE	0,643	0,092	6,983	0,000	0,461	0,826
JYLY	0,634	0,087	7,320	0,000	0,463	0,806
AYG	0,498	0,086	5,785	0,000	0,328	0,669
SEP	0,488	0,079	6,164	0,000	0,331	0,645
OCT	0,216	0,079	2,736	0,007	0,060	0,372
NOV	0,265	0,070	3,786	0,000	0,127	0,404
LAGGED SALES	-0,013	0,093	-0,137	0,891	-0,197	0,171

Table 4-14: Group 3 - Regression results on Seasonality & Lagged Sales

Up to that point, the level of regression analysis that produced significant results for Group 3 is that of Equation 8, which demonstrates **Σφάλμα! Το αρχείο προέλευσης της αναφοράς δεν βρέθηκε.**. The model's regression statistics establish an R Square 0.681 and an adjusted R Square 0.658.

Equation 8 performance is depicted in Figure 4-11, in which residuals do not seem to adopt a specific or repetitive pattern. Equation application generates a Mean Absolute

Error (MAE) (Thomaidis N. , 2021) is 16.362,33 and a Mean Absolute Percentage Error (MAPE) (Thomaidis N. , 2021) 15.26%.

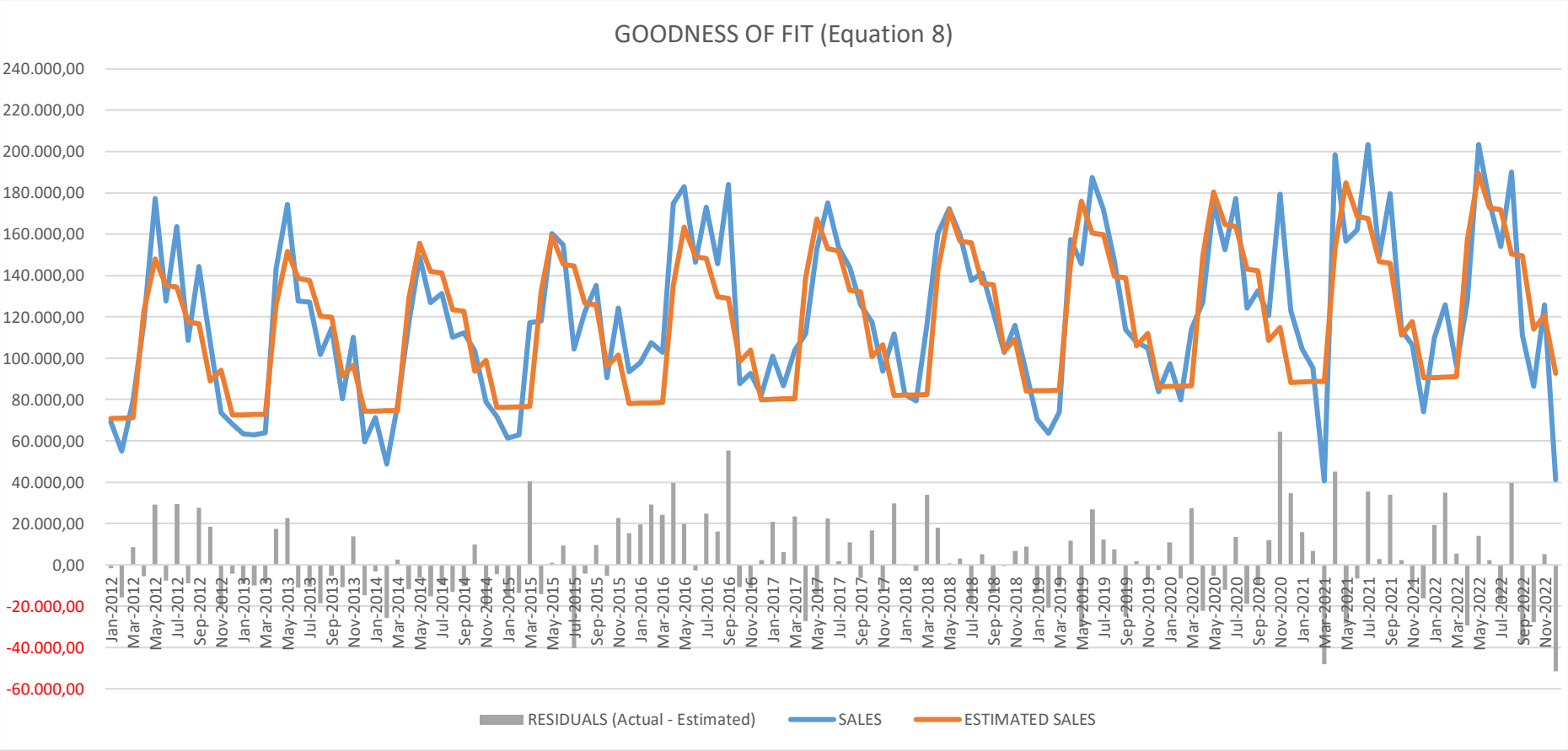

Figure 4-11: Regression performance on Equation 8

4.4. Concluding Remarks

The exploration of Group 3 statistical properties yielded a variety of remarks, most of them quite contrasting with product Groups previously analyzed. In particular:

- Group 3 in aggregate demonstrates a clear seasonal pattern, realized through summer season peaks (April to September) and winter lows. The pattern is evident throughout the timeseries, incorporating at the same time a distinct upward trend. Indicatively, 2012 annual sales generated 1.292.099€ revenue, whereas 2022 sales 1.545.514€ respectively.
- The seasonal pattern is not reproduced by all Group 3 components, as several Subgroups demonstrate a differentiated behavior that can not be standardized. Specifically, only Subgroup 302 (Insecticides & Mosquito Repellents) acquires clear cyclical behavior, affecting the overall Group 3 seasonal picture.
- As far as trend is concerned, Subgroups 301 (Household Equipment & Consumables), 303 (Dishwashing) and 304 (Laundry) perform as drivers of Groups' overall upward trend, contributing to positive trend formation. Their aggregated average share on Group 3 sales is 56,1%, a fact which constitutes them capable of formatting the overall tendency. On the other hand, Subgroups 302, 305 and 306, all demonstrate diminishing trends.
- The data distribution generates a fairly symmetrical Group 3 histogram, slightly skewed to the right, highlighting the effect of extreme, high value sales. The subgroup under the greatest skewness is that of Household Equipment & Consumables (301), whose vast majority of observations is placed to the left side of the distribution.
- Correlation of variance among subgroups is not evident.
- Regression analysis confirmed that seasonality and time trend are the sole drivers of sales formation. Neither pandemic variables nor macroeconomic factors seemed to impose significant alterations on sales evolution.
- The time trend present in Equation 8 can be considered substitute to the effects generated by the incorporation of AVA liquid dish detergents in 303 product mix in 2015. Specifically, by adding AVA dummy variable, active from May

2015 up to the end of the timeseries, linear trend becomes insignificant, as AVA acquires a significant coefficient (0.183 with 0.004 p-value) that reproduces Equation 8 performance. By exempting linear trend and adding AVA to the regression equation, residual behavior improves (MAE 15.591 versus 16.362 And MAPE 14.32% against 15.26% in Equation 8). The new model performance, based on Equation 9 is presented in Figure 4-12.

(9) $$log\left(\widehat{S_t}\right) = 11.165 + 0.547 * APR + 0.716 * MAY + 0.625 * JUN + 0.619 * JUL + 0.485 * AUG + 0.479 * SEP + 0.208 * OCT + 0.264 * NOV + 0.203 * AVA$$

Although Equation 9 model is strictly based on single product influence and can be used as grounds for individual brand – effect examination, provides appreciable explanatory value to the model.

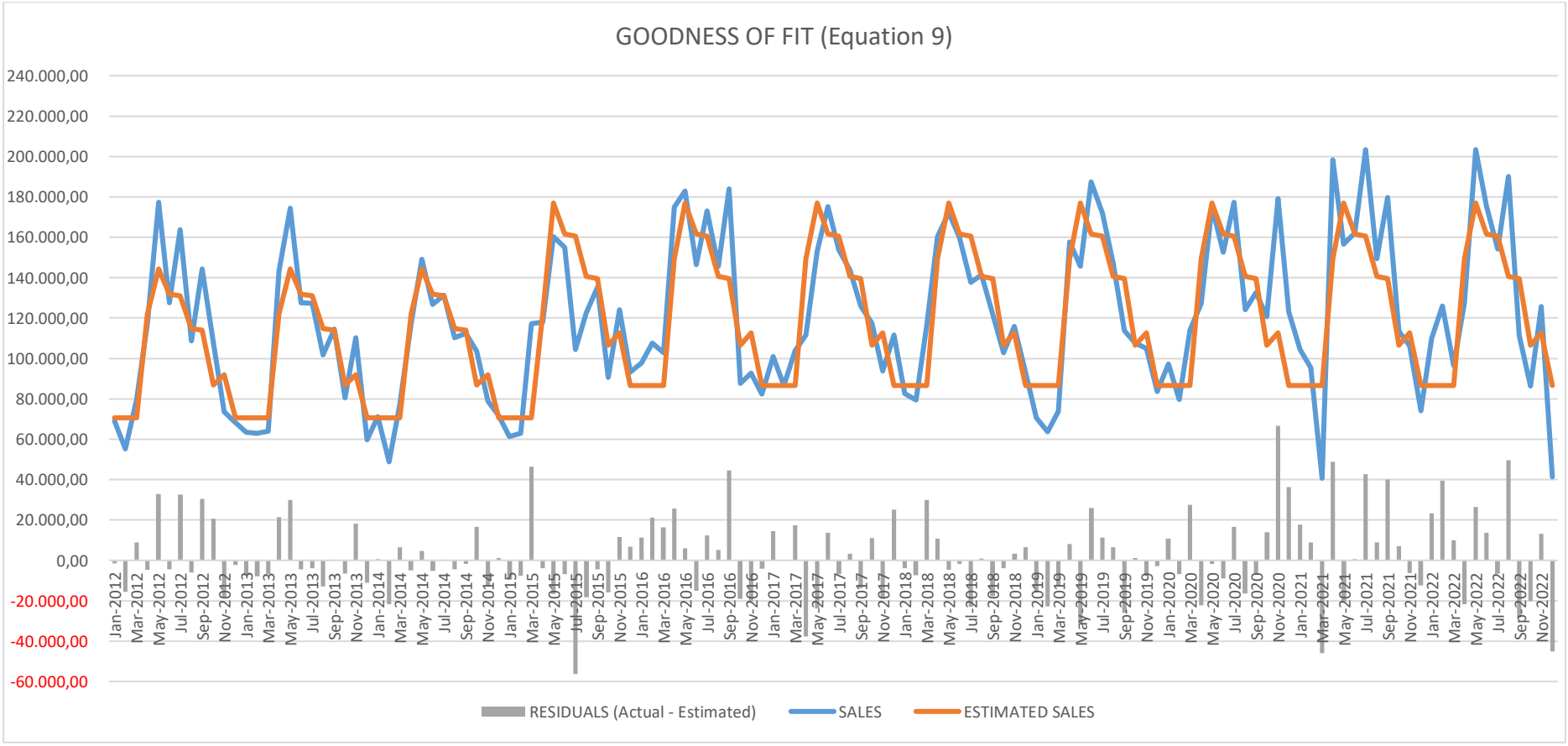

Figure 4-12: Regression performance on Equation 9

5 GROUP 4 - SNACKS

5.1. Group Introduction

The fourth group of products is composed of five individual subgroups, among which only two hold the widest sales share. Snacks are a commercial category with great dynamics for Achtida, as its share on annual revenue is admittedly of considerable size (Kouvrakis, 2023). For the period under study, sales range from 94.485€ monthly, up to 428.520€. Mean sales are located to 220.487€.

It is worth noting that among all product categories examined in the present study frame, Snacks were the group with the greatest variations in product mix terms. In order to rationalize the analysis and provide meaningful results, brands that have not been commercialized by Achtida throughout the full length of the timeseries, were excluded from the study. The most outstanding of them were:

- Elbisco products: Snacks under the brand name Allatini (biscuits, crackers), as Achtida dismissed contract of commercial representation with Elbisco in 2020.
- Chipita Croisants: Molto and 7Days croissants are not included in the timeseries, as collaboration with Achtida begun in 2022, after the Chipita – Mondelez merge.
- Chocolates commercialized from Mars Hellas: fragmentary presence in Achtidas' product mix, as representation agreement was terminated approximately at the beginning of the timeseries.
- Other brands: Healthy Habbits nutritious snacks, McVites biscuits, chewing gums partially promoted by Pepsico (Tasty), Canderel chocolates and others.

Consequently, the observations presented below are in respect with a) current product mix and b) brands that had permanent presence all through the timeseries.

5.1.1. CHEWING GUMS & CANDIES (401)

The first subgroup incorporates products such as chewing gums, sugar-based candies and lollipops. For items with such a small unit value, the average share to groups on

11.6%, denotes the vast amount of units sold monthly. The most dominant brands are Trident, Dentyne, Babaloo and Elma chewing gums, Halls, Ibis, O'Mamy, Viap, as well as no-name, professional packaging products sold to mini-markets and patisseries.

5.1.2. CHOCOLATE (402)

402 subgroup consists of a vast variety of brands, namely Lacta, Pavlidis, Kinder, Milka, Toblerone. Additionally, it has to be mentioned that although Achtidas' representation contract as the official distributor of Afoi Kotsiopouli (Ion products: Ion chocolates, Sokofreta, Derby, Break and other brands) ended some years ago (Kouvrakis, 2023), the company continued to market those brands in an equivalent volume. Therefore, Ion products remained in the timeseries and represent a considerable percentage (indicatively 9.2% for July 2022) of the 402 participation (46.7%) to average Group 4 sales formation.

5.1.3. NATURAL & TRADITIONAL SNACKS (403)

403 subgroup presents the aggregate sales of sesame bars (pasteli), halvah, loukoumi, sweet preserved fruits in jar, as well as dry nuts. Its participation to Group 4 sales is only 1.9% in average, but due to the differentiated characteristics of items incorporated to it, it couldn't be aggregated with other snack categories. The leading brands participating in 403 are Pasteli Jannis, Papageorgiou, Tasty Naturals.

5.1.4. DOUGH BASED SNACKS (404)

Dough based snacks refer to sweet biscuits and crackers, salted crackers and croissants.

Dominant brand names are Trendy and Folie croissants, Oreo biscuits and Tuc crackers. With the exclusion of Allatini brand as mentioned above, subgroup 404 is responsible for 7.2% of Group 4 average eleven-year sales.

5.1.5. CHIPS & EXTRUDED SNACKS (405)

405 subgroup incorporates the wide variety of snacks produced and commercialized in Greece by (former Tasty Foods) Pepsico Hellas. Ruffles and Lays are the leading chips brands and, as far as other extruded snacks are concerned, Cheetos (in a variety of flavors), Dracoulinia, Foudounia, Pizzinia, Pacotinia, Poppers, as well as Doritos nachos.

The contribution of Chips and Extruded Snacks in Group 4 average sales formation is 32.6%.

5.2. Data Exploration

Figure 5-1 depicts Group 4 sales' evolution during the period 2012 – 2022, accompanied by distinct subgroup presentation.

5.2.1. VISUAL OBSERVATIONS

Group 4 total sales display a distinct cyclical pattern across the period under study. Especially, there seem to be a sheer seasonal effect characterized by a progressive sales increase from winter up to spring months, which finally ends up in a great spike in July in most of the cases. An explicit upward trend is apparent, mainly evident on the second half of the timeseries. Indeed, while mean sales are located on 182.052€, with a standard deviation 42.822€ for the first 66 monthly observations, the picture is quite different for the second part of the timeseries, with mean sales 258.922€ and standard deviation 77.642. Those facts denote, apart from the upward trend, the widening variation of sales as well.

Mean sales throughout the period are 220.487€ and standard deviation 73.413€. Total sales range monthly from 94.485,24€ to 428.520,10€, while median sales value is 210.161,08€. The highest sales value was accomplished off-season, in November 2020, mainly triggered by a huge spike in chocolate (subgroup 402) sales.

Concerning the pandemic era, the only differences visually detected are

a) the restrained sales on 2020 summer months, as July provided only 21% sales increase over June, compared to an average 45% on all other years.

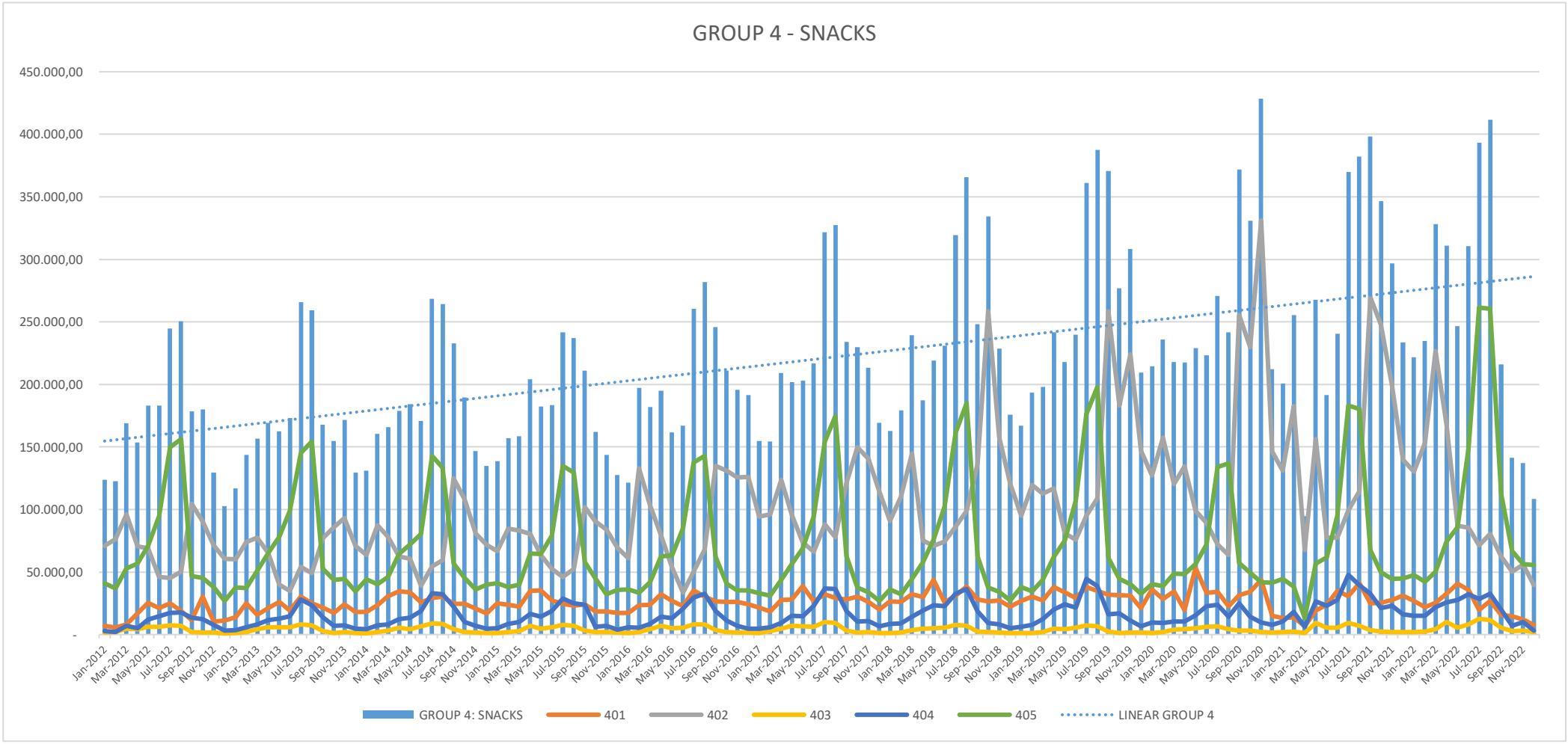

Figure 5-1: Group 4 sales evolution (2012-2022)

b) the delayed 2020 peak season, as highest sales were located from September up to November

c) the differentiation in annual pick sales, which were accomplished on November

d) the distortion of seasonal pattern from 2020 and on, as greater variations occurred from month to month.

	GROUP 4	**401**	**402**	**403**	**404**	**405**
MINIMUM	**94.485,24**	5.601,16	33.519,23	953,44	2.077,51	13.121,85
MAXIMUM	**428.520,10**	53.232,08	331.074,60	12.604,29	47.761,62	261.497,41
RANGE	**334.034,86**	47.630,92	297.555,37	11.650,85	45.684,11	248.375,56
MEAN	**220.487,15**	25.666,05	102.881,94	4.135,85	15.900,60	71.902,71
MEDIAN	**210.161,08**	25.969,61	86.671,41	3.406,89	14.072,18	54.310,29
STANDARD DEVIATION	**73.413,37**	8.204,22	54.096,69	2.665,05	9.936,20	48.339,85

Table 5-1: Group 4 - Basic Descriptive Features

Figure 5-2 up to Figure 5-6 demonstrate graphically subgroup evolution. 401 (Chewing Gums & Candies) is characterized from a moderately upward trend, within a range of 47.630,92€. The observations vary by a standard deviation 8.204,22€, the smallest variation against all subgroups, when compared to their range. The greater sales value is accomplished on May 2020, and the lower on February 2012. Signs of seasonal forces are evident mostly in the years up to 2018, as lowest sales are located around December or January, while highest sales happen mostly on spring and summer months. After 2018 – 2019 the pattern is quite distorted, keeping however hints of seasonality.

Subgroup 402 (Chocolate) demonstrates a profound upward trend, having the widest sales range (297.555,37€) among all subgroups, but the second smallest deviation compared to its range. Mean sales are located on 102.881,94€, the highest again among subgroups. Chocolates' seasonal spikes are quite differentiated, as they occur mostly in autumn, a fact expected for that specific product group (Kouvrakis, 2023). From 2018 and then again on the beginning of the pandemic era the variation of observation is amplified, keeping, however the distinct "two times in a year" peaks (September / October and March / April).

Natural and traditional snacks (403) occupy a small share of Group 4 sales, approximately 2% on the eleven-year period, while at the same time demonstrate an approximately stable trend. However, it is worth noting that 403 is a subgroup with distinguishable seasonal pattern over time, demonstrating at most of the cases two distinct peaks, the first in April and the second on July or August every year. Sales range from 953,44€ monthly to 12.604,29€, a level accomplished for the first time in July 2022, with mean sales located on 4.136€.

Subgroup 404 (Dough based snacks) follows a clear seasonal pattern as well, exhibiting a raising tendency strictly up to 2019. On 2020, due to pandemic disruptions (Kouvrakis, 2023) peak sales take place on September, conveying only 24.804€ of revenue, being 44% lower than previous year peak sales, on July (44.354€). From then on, sales recover next year, accomplishing 47.761€ on July 2021. Mean sales are placed on 15.900,60€, with a standard deviation 9.936,20€.

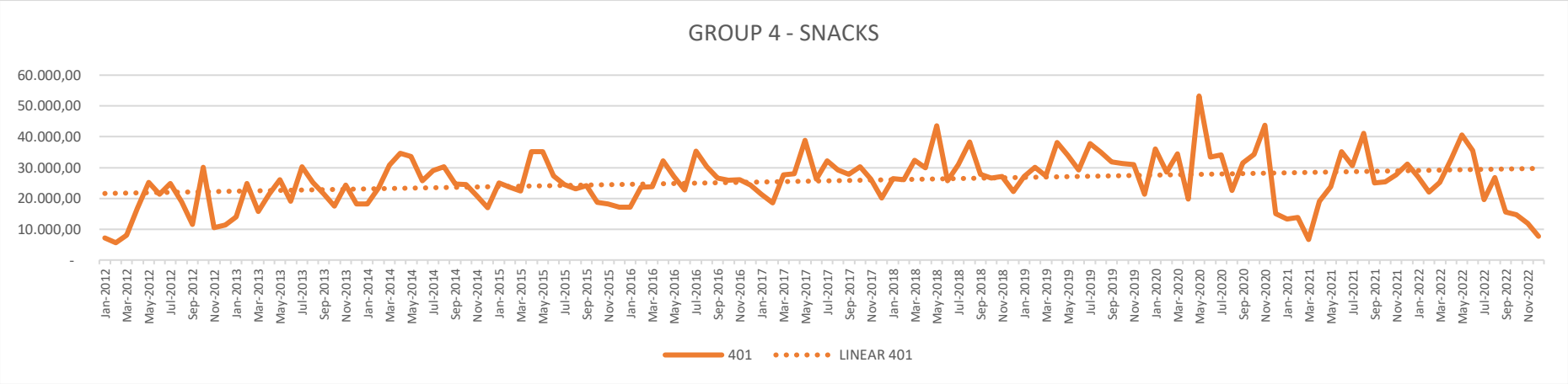

Figure 5-2:Subgroup 401 sales evolution (2012-2022)

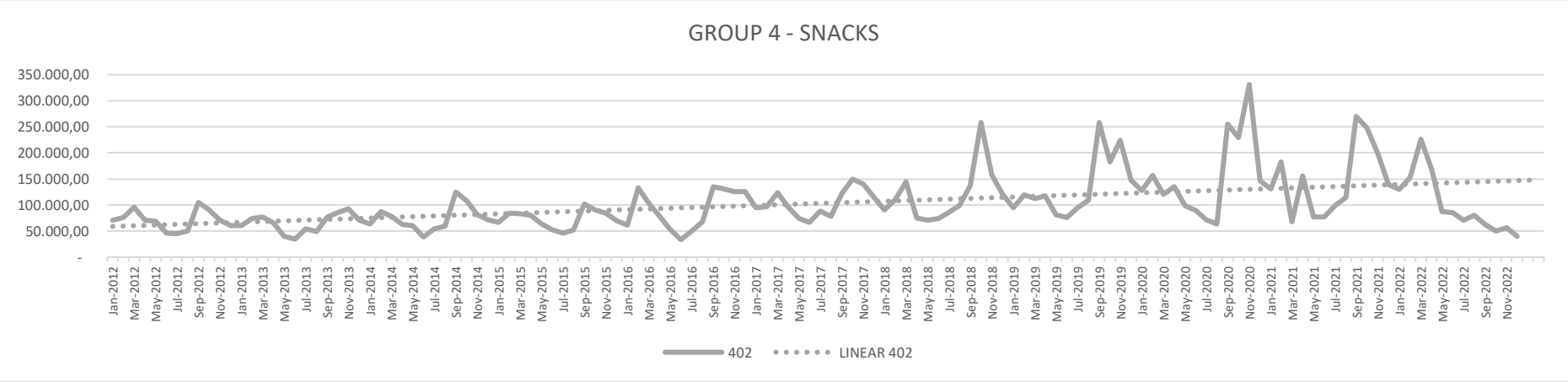

Figure 5-3: Subgroup 402 sales evolution (2012-2022)

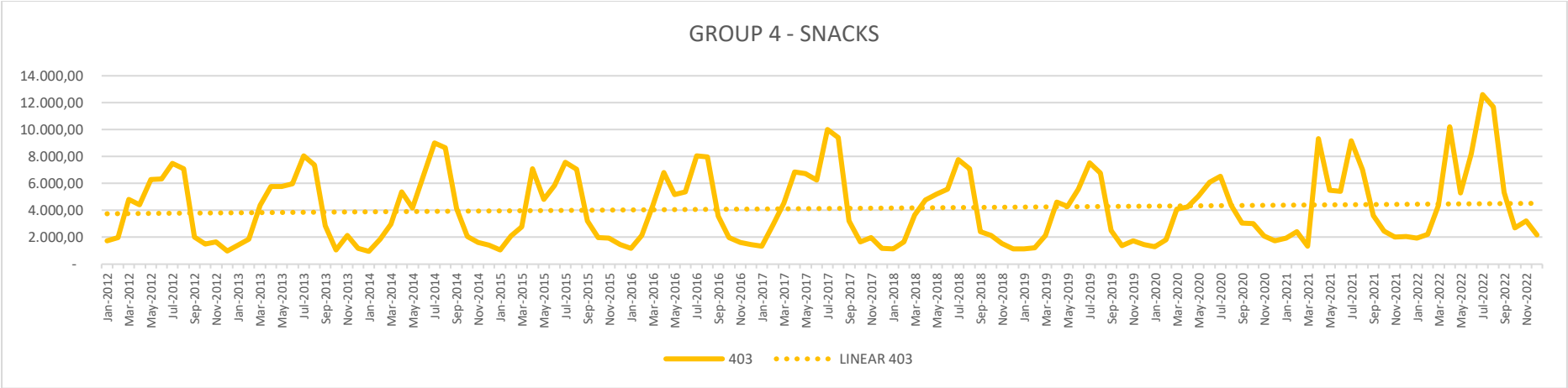

Figure 5-4: Subgroup 403 sales evolution (2012-2022)

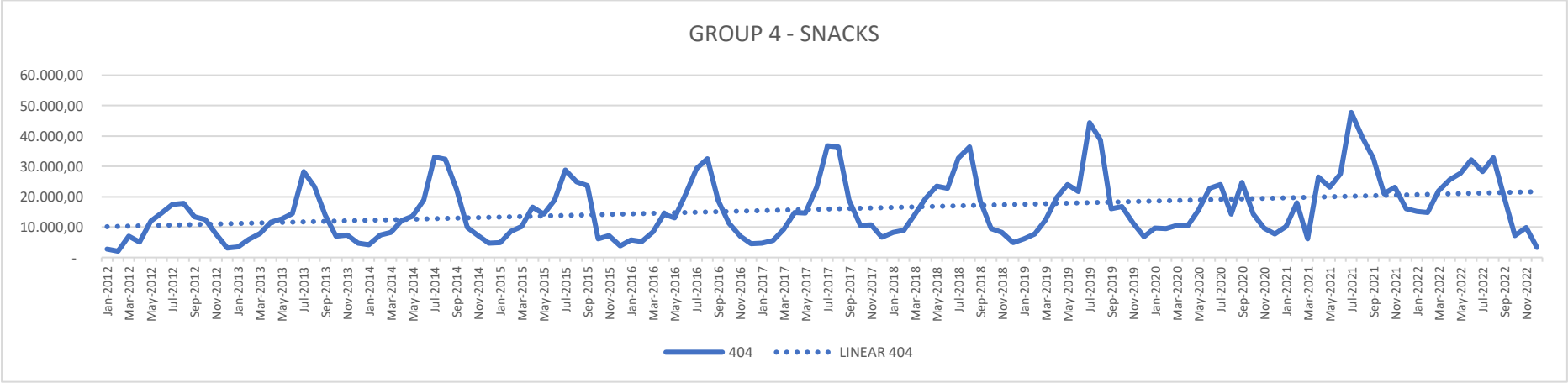

Figure 5-5: Subgroup 404 sales evolution (2012-2022)

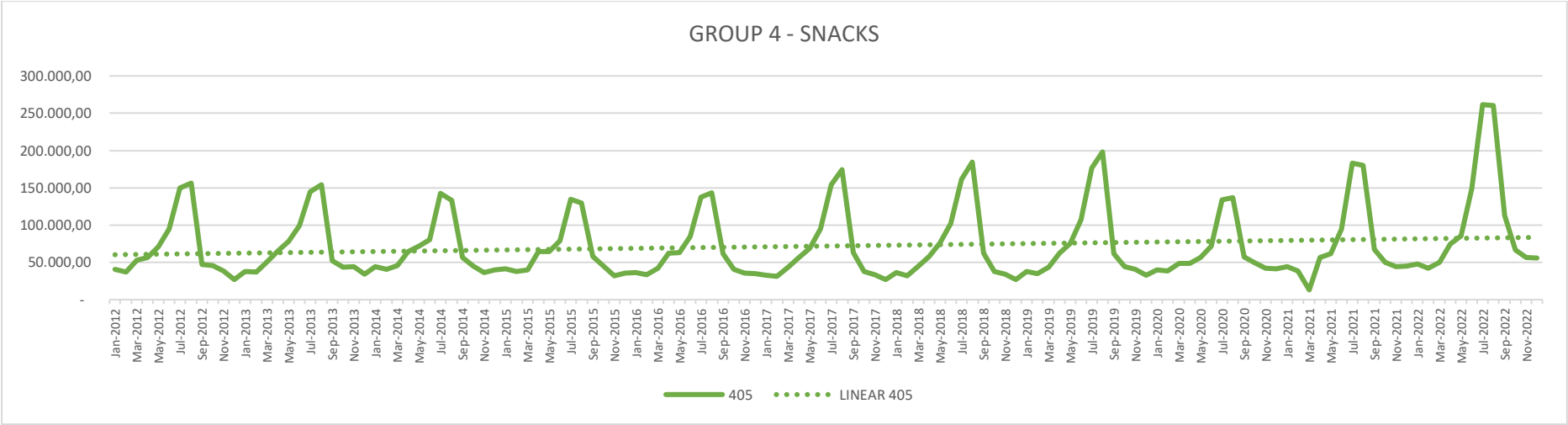

Figure 5-6: Subgroup 405 sales evolution (2012-2022)

Finally, subgroup 405 (Chips & Extruded Snacks), being the second most robust after chocolate snacks, forms 32.6% of Group 4 average eleven-year sales, demonstrating at the same time an explicitly seasonal pattern and an upward trend. The increasing tendency is distinguishable from 2017 and on, yet being restrained during the first summer under pandemic restrictions. Indeed, in July 2020 sales attain 136.707€, a level comparable with July 2015 sales, on 134.755€. Lowest sales value is located on March 2021, on 13.121,85€. On the other hand, highest sales value was accomplished on July 2022, at 261.497,41 €. Mean sales of the eleven-year period are 71.903€, with 48.339,85€ standard deviation.

Concerning the existence of linear relationship among subgroups (Giannikos, Forecasting with Regression, 2020), Table 5-2 summarizes correlation data, in order to assess its strength.

	401 vs 402	**401 vs 403**	**401 vs 404**	**401 vs 405**
CORRELATION	0,224	0,347	0,509	0,295
		402 vs 403	**402 vs 404**	**402 vs 405**
CORRELATION		-0,301	0,009	-0,293
			403 vs 404	**403 vs 405**
CORRELATION			0,776	0,843
				404 vs 405
CORRELATION				0,804

Table 5-2: Correlation of variance among Group 4 subgroups

The pairs that demonstrate correlation closer to the unit are subgroups 403 and 405 as well as 404 and 405. The relevant scatterplots are presented below.

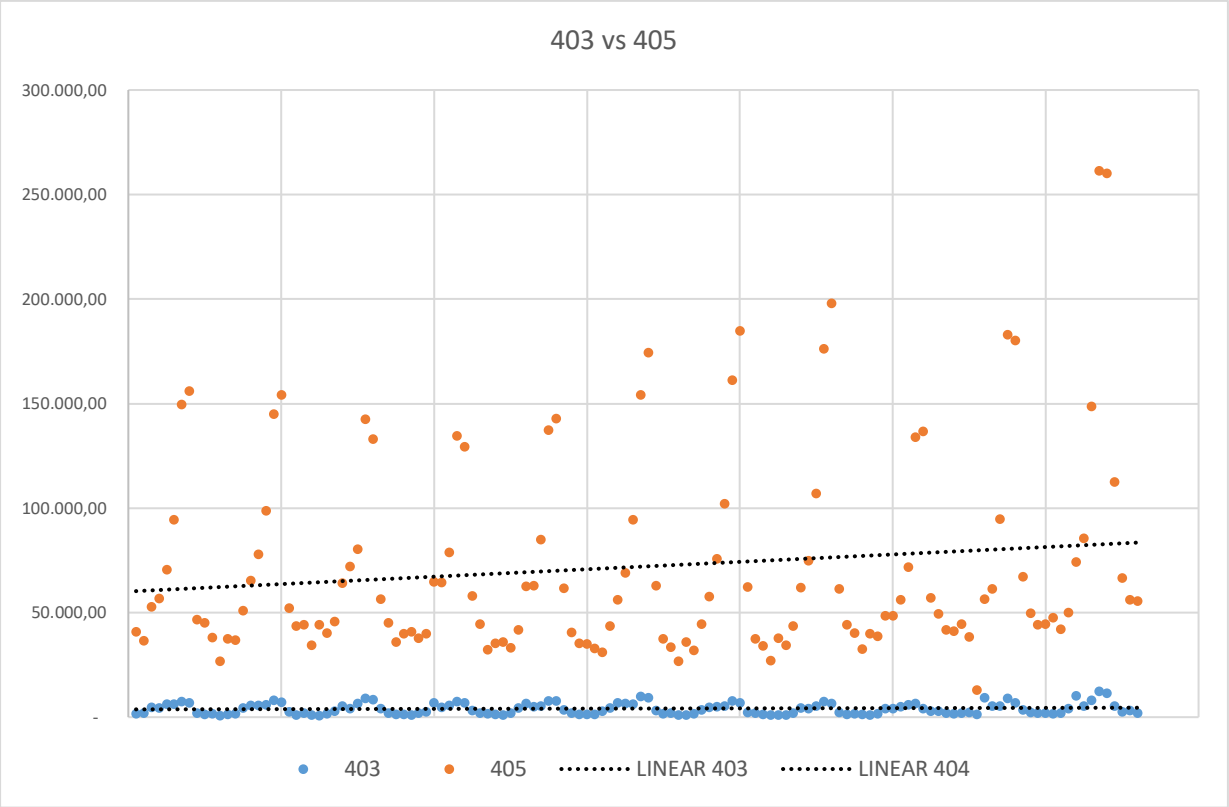

Figure 5-7: Correlation among 403 & 405

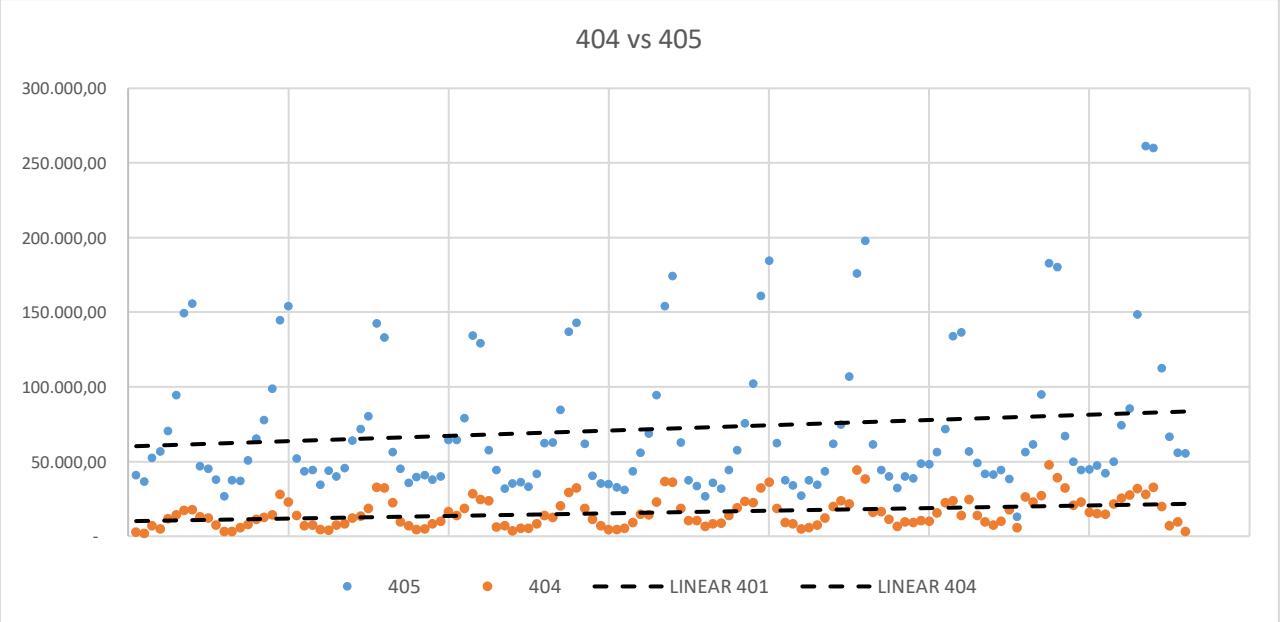

Figure 5-8: Correlation among 404 & 405

5.2.2. FREQUENCY

Group 4 data distribution is moderately skewed to the left, having a positive skew (Giannikos, Descriptive Statistics, 2020) that denotes the existence of extreme, high values of sales among the observations. The majority of observations is found below the distribution mean (220.487,15), while median is located on 210.161,08€.

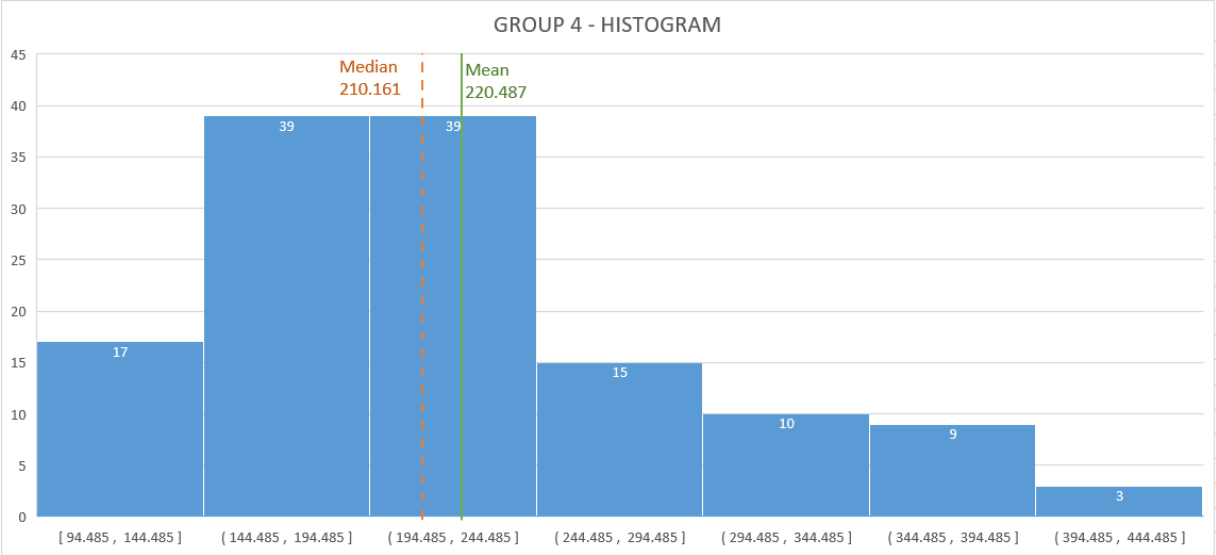

Figure 5-9: Histogram - Group 4

Individual subgroups are partially differentiated from Group 4 as a whole. Particularly, as far as kurtosis is concerned, group aggregation as well as the majority of subgroups display values lower than unit, constituting their distributions platykurtic. However, subgroup 405 of Chips and Extruded Snacks acquires a 2.84 kurtosis, introducing a nearly mesokurtic shape and 402 (Chocolate) a 3.31, establishing the first case of leptokurtic distribution (Giannikos, Descriptive Statistics, 2020). In the latter case, tail thickness indicates the presence of many outliers (Lord, Qin, & Geedipally, 2021).

In terms of skewness, subgroups 405 and 402 demonstrate again the higher values (1.74 and 1.71 respectively) highlighting again the existence of extreme values. Subgroups 403 and 404 are moderately skewed, while subgroup 401 demonstrates an approximately symmetrical distribution, with skewness 0.0014.

	GROUP 4	**401**	**402**	**403**	**404**	**405**
MEAN	**220.487,15**	25.666,05	102.881,94	4.135,85	15.900,60	71.902,71
MEDIAN	**210.161,08**	25.969,61	86.671,41	3.406,89	14.072,18	54.310,29
SKEWNESS	**0,86**	0,00	1,71	0,80	0,90	1,74
KURTOSIS	**0,25**	0,56	3,31	-0,08	0,25	2,84

Table 5-3: Group 4 - Frequency Features

The boxplot in Figure 5-10 reproduces the effects demonstrated on Group 4 Histogram.

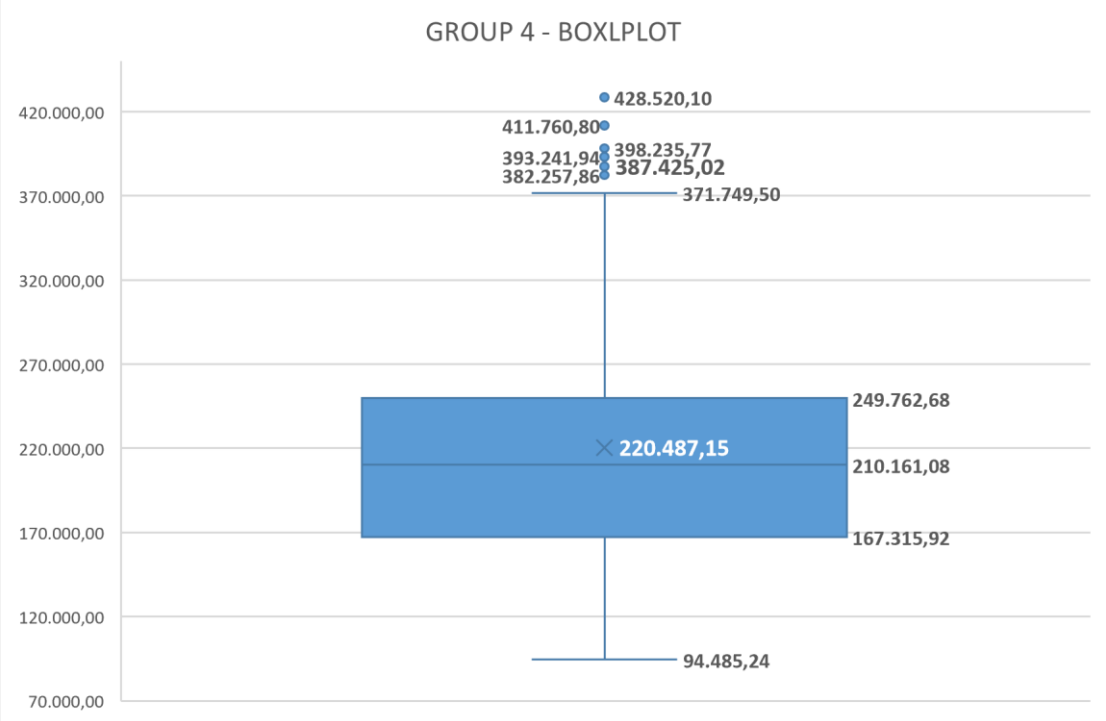

Figure 5-10: Boxplot - Group 4

5.2.3. SEASONALITY

Seasonal implications were made clear from the graphic representation of Group 4 sales evolution in Figure 5-1. In order to better asses the phenomenon, Figure 5-11 presents monthly data distribution.

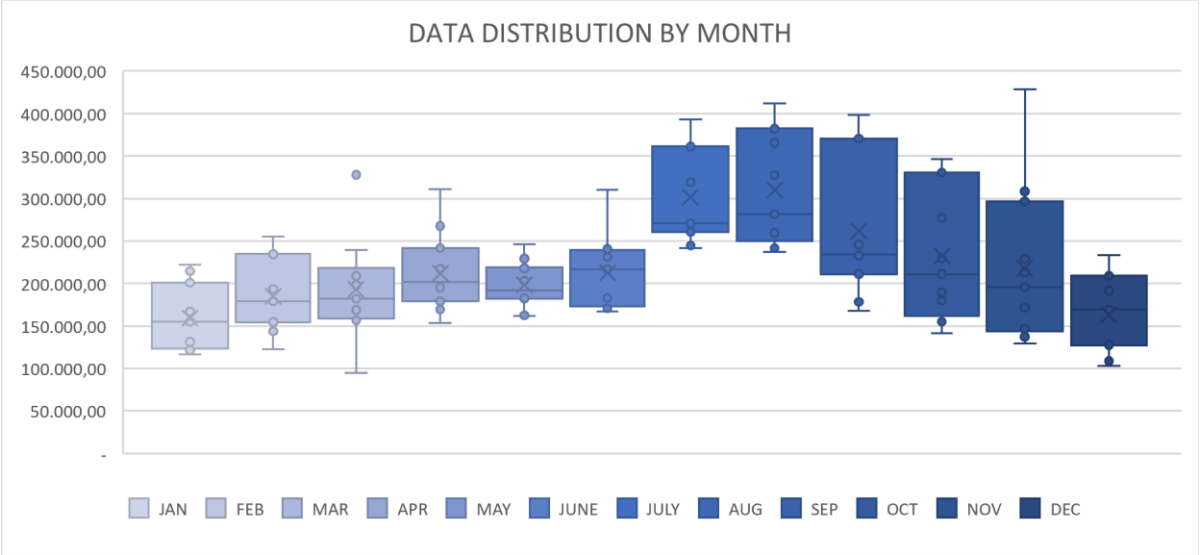

Figure 5-11: Monthly Boxplot - Group 4

The seasonal effect becomes clearly evident since in July, the light upward trend detected up to June creates a clear spike which elevates mean July sales to 301.452€. Subsequently, mean sales per month remain steadily over 218.000€ up to November.

The most significant alteration noticed in Figure 5-11 compared to other Groups' monthly distributions is the endurance of sales distribution to higher values up to October and November. That observation is mainly stimulated by subgroup 402 (Chocolate) autumn sales and its differentiated seasonal cycle.

5.3. Regression

5.3.1. SEASONAL EFFECTS

Seasonality is the first factor incorporated in the regression equation, since signs of its effects are already evident from the aforementioned analysis. With the incorporation of dummy variables for each month of the year, regression results are presented in Table 5-4, at a confidence level 95%. Since insignificant coefficients exist, regression is applied again after excluding them (Table 5-5). At that point, regression statistics

generate an R Square value at 0.634 and an Adjusted R Square at 0.606. The results are integrated in Equation 10.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,698	0,066	177,681	0,000	11,567	11,828
Linear Trend	0,004	0,000	9,258	0,000	0,003	0,005
FEB	0,145	0,084	1,716	0,089	-0,022	0,312
MAR	0,163	0,084	1,935	0,055	-0,004	0,331
APR	0,276	0,084	3,264	0,001	0,108	0,443
MAY	0,218	0,084	2,581	0,011	0,051	0,385
JUNE	0,275	0,084	3,253	0,001	0,108	0,442
JYLY	0,623	0,084	7,372	0,000	0,456	0,790
AYG	0,640	0,085	7,571	0,000	0,473	0,807
SEP	0,445	0,085	5,268	0,000	0,278	0,613
OCT	0,315	0,085	3,727	0,000	0,148	0,483
NOV	0,224	0,085	2,653	0,009	0,057	0,392
DEC	-0,034	0,085	-0,400	0,690	-0,201	0,134

Table 5-4: Group 4 - Regression results on Seasonality - 1st step

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,769	0,043	275,596	0,000	11,685	11,854
Linear Trend	0,004	0,000	8,973	0,000	0,003	0,005
APR	0,207	0,068	3,034	0,003	0,072	0,342
MAY	0,149	0,068	2,190	0,030	0,014	0,284
JUNE	0,206	0,068	3,023	0,003	0,071	0,341
JYLY	0,554	0,068	8,124	0,000	0,419	0,689
AYG	0,571	0,068	8,372	0,000	0,436	0,707
SEP	0,377	0,068	5,522	0,000	0,242	0,512
OCT	0,247	0,068	3,615	0,000	0,112	0,382
NOV	0,156	0,068	2,285	0,024	0,021	0,291

Table 5-5: Group 4 - Regression results on Seasonality - 2nd step

(10) $$log\left(\widehat{S_t}\right) = 11.769 + 0.004 * t + 0.207 * APR + 0.149 * MAY + 0.206 * JUN + 0.554 * JUL + 0.571 * AUG + 0.377 * SEP + 0.247 * OCT + 0.156 * NOV$$

5.3.2. PANDEMIC EFFECT

Proceeding with pandemic effect exploration, WHO, 1st LOCKDOWN, 2ND LOCKDOWN and VAC variables are tested as follows. Table 5-6, Table 5-7 and Table 5-8, all generate insignificant results concerning the pandemic variables.

On the other hand, as presented in Table 5-9, VAC acquires a p-value that can be accepted in a 95% percent significance level. Nevertheless, the coefficient itself yields a value (-0.139) that is unable to be interpreted sufficiently. Particularly, VAC variable, referring to the initiation of vaccination program against COVID-19 is supposed to generate beneficial implications towards sales evolution, through restriction decline and other positive suggestions on consumer behavior. This opposition leads to the exclusion of VAC variable as well, as the contrasting effect is believed to be an outcome of random or incidental factors.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,736	0,047	248,993	0,000	11,643	11,830
Linear Trend	0,005	0,001	7,112	0,000	0,004	0,006
APR	0,212	0,068	3,124	0,002	0,078	0,346
MAY	0,153	0,068	2,263	0,025	0,019	0,288
JUNE	0,209	0,068	3,089	0,002	0,075	0,344
JYLY	0,557	0,068	8,210	0,000	0,422	0,691
AYG	0,573	0,068	8,448	0,000	0,439	0,707
SEP	0,378	0,068	5,567	0,000	0,243	0,512
OCT	0,247	0,068	3,635	0,000	0,112	0,381
NOV	0,155	0,068	2,284	0,024	0,021	0,289
WHO	-0,099	0,061	-1,611	0,110	-0,221	0,023

Table 5-6: Group 4 - Regression results on Seasonality & Pandemic (WHO)

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,769	0,043	274,269	0,000	11,684	11,854
Linear Trend	0,004	0,000	8,835	0,000	0,003	0,005
APR	0,207	0,069	2,995	0,003	0,070	0,343
MAY	0,149	0,069	2,160	0,033	0,012	0,286
JUNE	0,206	0,069	3,010	0,003	0,071	0,342
JYLY	0,554	0,069	8,085	0,000	0,419	0,690
AYG	0,572	0,069	8,332	0,000	0,436	0,707
SEP	0,377	0,069	5,496	0,000	0,241	0,513
OCT	0,247	0,069	3,598	0,000	0,111	0,383
NOV	0,156	0,069	2,275	0,025	0,020	0,292
1st LOCKDOWN	0,004	0,123	0,036	0,971	-0,239	0,248

Table 5-7: Group 4 - Regression results on Seasonality & 1st Lockdown

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,769	0,043	274,563	0,000	11,684	11,854
Linear Trend	0,004	0,000	8,557	0,000	0,003	0,005
APR	0,207	0,068	3,023	0,003	0,071	0,343
MAY	0,152	0,069	2,202	0,030	0,015	0,288
JUNE	0,209	0,069	3,026	0,003	0,072	0,345
JYLY	0,557	0,069	8,069	0,000	0,420	0,694
AYG	0,574	0,069	8,313	0,000	0,437	0,711
SEP	0,380	0,069	5,495	0,000	0,243	0,516
OCT	0,249	0,069	3,610	0,000	0,113	0,386
NOV	0,156	0,069	2,280	0,024	0,021	0,292
2nd LOCKDOWN	0,027	0,090	0,300	0,765	-0,151	0,205

Table 5-8: Group 4 - Regression results on Seasonality & 2nd Lockdown

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,734	0,045	262,325	0,000	11,645	11,822
Linear Trend	0,005	0,001	8,326	0,000	0,004	0,006
APR	0,208	0,067	3,095	0,002	0,075	0,340
MAY	0,149	0,067	2,221	0,028	0,016	0,282
JUNE	0,205	0,067	3,055	0,003	0,072	0,338
JYLY	0,552	0,067	8,230	0,000	0,419	0,685
AYG	0,568	0,067	8,468	0,000	0,435	0,701
SEP	0,373	0,067	5,553	0,000	0,240	0,506
OCT	0,242	0,067	3,598	0,000	0,109	0,375
NOV	0,150	0,067	2,232	0,027	0,017	0,283
VAC	-0,139	0,060	-2,304	0,023	-0,259	-0,020

Table 5-9: Group 4 - Regression results on Seasonality & VAC

5.3.3. MACROECONOMIC EFFECTS

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	12,220	0,304	40,253	0,000	11,619	12,821
Linear Trend	0,004	0,000	9,123	0,000	0,003	0,005
APR	0,203	0,068	2,990	0,003	0,069	0,338
MAY	0,145	0,068	2,139	0,034	0,011	0,280
JUNE	0,202	0,068	2,974	0,004	0,068	0,337
JYLY	0,552	0,068	8,134	0,000	0,418	0,687
AYG	0,569	0,068	8,380	0,000	0,435	0,704
SEP	0,375	0,068	5,514	0,000	0,240	0,509
OCT	0,248	0,068	3,644	0,000	0,113	0,382
NOV	0,157	0,068	2,305	0,023	0,022	0,291
GDP (Million euros)	0,000	0,000	-1,500	0,136	0,000	0,000

Table 5-10: Group 4 - Regression results on Seasonality & GPD

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	12,261	0,313	39,172	0,000	11,641	12,881
Linear Trend	0,004	0,000	8,997	0,000	0,003	0,005
APR	0,207	0,068	3,033	0,003	0,072	0,342
MAY	0,145	0,068	2,131	0,035	0,010	0,281
JUNE	0,202	0,068	2,963	0,004	0,067	0,337
JYLY	0,550	0,068	8,060	0,000	0,415	0,685
AYG	0,570	0,068	8,347	0,000	0,435	0,705
SEP	0,375	0,068	5,495	0,000	0,240	0,510
OCT	0,245	0,068	3,585	0,000	0,110	0,380
NOV	0,157	0,068	2,305	0,023	0,022	0,293
LAGGED GDP (1 month, in Million euros)	0,000	0,000	-1,579	0,117	0,000	0,000

Table 5-11: Group 4 - Regression results on Seasonality & Lagged GPD

As in previously analyzed product Groups, GDP and Lagged GDP are variables tested on their ability to explain sales variation. According to Table 5-10 and Table 5-11, this was not confirmed. Subsequently, Tourism Turnover Index was proved insignificant as well.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,763	0,050	235,723	0,000	11,664	11,862
Linear Trend	0,004	0,000	8,815	0,000	0,003	0,005
APR	0,201	0,073	2,758	0,007	0,057	0,345
MAY	0,143	0,073	1,967	0,051	-0,001	0,287
JUNE	0,200	0,073	2,750	0,007	0,056	0,344
JYLY	0,536	0,101	5,307	0,000	0,336	0,735
AYG	0,553	0,101	5,478	0,000	0,353	0,752
SEP	0,358	0,101	3,552	0,001	0,159	0,558
OCT	0,245	0,069	3,543	0,001	0,108	0,381
NOV	0,154	0,069	2,229	0,028	0,017	0,291
TOURISM TURNOVER	0,000	0,001	0,254	0,800	-0,001	0,001

Table 5-12: Group 4 - Regression results on Seasonality & TTI

Proceeding to Overnight stops examination, results were quite different as the variable acquired a significant coefficient, receiving however a coefficient extremely close to zero. Nevertheless, the analysis progressed to the exclusion of insignificant factors on Table 5-14.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,691	0,038	308,540	0,000	11,616	11,766
Linear Trend	0,005	0,000	11,885	0,000	0,004	0,006
APR	0,168	0,057	2,947	0,004	0,055	0,281
MAY	0,072	0,063	1,152	0,252	-0,052	0,196
JUNE	0,017	0,083	0,208	0,836	-0,147	0,182
JYLY	0,299	0,105	2,851	0,005	0,091	0,507
AYG	0,293	0,111	2,640	0,009	0,073	0,513
SEP	0,249	0,081	3,077	0,003	0,088	0,409
OCT	0,254	0,060	4,237	0,000	0,135	0,372
NOV	0,208	0,057	3,681	0,000	0,096	0,321
OVERNIGHT	0,00000017	0,000	2,785	0,006	0,000	0,000

Table 5-13: Group 4 - Regression results on Seasonality & Overnights, 1st Step

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	11,698	0,037	314,058	0,000	11,624	11,772
Linear Trend	0,005	0,000	12,006	0,000	0,004	0,006
APR	0,155	0,055	2,837	0,005	0,047	0,264
JYLY	0,262	0,074	3,568	0,001	0,117	0,408
AYG	0,254	0,077	3,309	0,001	0,102	0,406
SEP	0,221	0,062	3,582	0,001	0,099	0,343
OCT	0,237	0,055	4,343	0,000	0,129	0,345
NOV	0,198	0,055	3,571	0,001	0,088	0,307
OVERNIGHT	0,00000018	0,000	4,602	0,000	0,000	0,000

Table 5-14: Group 4 - Regression results on Seasonality & Overnights, 2nd Step

Although the coefficient remained close to zero, the model produced improved regression statistics: R Square: 0.753 and Adjusted R Square 0.735. A matter that

should not be underestimated is that overnight stop data are available only up to December 2021, so the model ultimately excludes one complete year of sales observations. However, its performance will be furtherly assessed through Goodness – of – Fit graphs as well as residual behavior.

$$(11)\ log\left(\widehat{S_t}\right) = 11.698 + 0.005 * t + 0.155 * APR + 0.262 * JUL + 0.254 * AUG + 0.221 * SEP + 0.237 * OCT + 0.198 * NOV + 0{,}00000018 * OVERNIGHS$$

5.3.4. SALES MOMENTUM

Based on Equation 11, the analysis proceeded to the last regression round that incorporated Lagged Sales.

	Coefficients	*Standard error*	*t*	*P value*	*Lower 95%*	*Upper 95%*
Constant Term	10,688	1,052	10,164	0,000	8,603	12,772
Linear Trend	0,005	0,001	7,356	0,000	0,003	0,006
APR	0,155	0,055	2,820	0,006	0,046	0,264
JYLY	0,269	0,074	3,629	0,000	0,122	0,417
AYG	0,232	0,081	2,879	0,005	0,072	0,392
SEP	0,189	0,070	2,695	0,008	0,050	0,328
OCT	0,210	0,062	3,400	0,001	0,088	0,332
NOV	0,175	0,060	2,895	0,005	0,055	0,295
OVERNIGHT	0,000000171	0,000	4,013	0,000	0,000	0,000
LAGGED LN Sales	0,086	0,090	0,962	0,338	-0,091	0,264

Table 5-15: Group 4 - Regression results on Seasonality, Overnights & Lagged Sales

According to Table 5-15, Lagged Sales is not a significant variable towards sales formation in a 95% confidence level.

Consequently, the models presented in Equation 10: Expected sales based on time trend and seasonality and in Equation 11: Expected sales based on time trend, seasonality & Overnights are the ones applied to the following figures.

Equation 10 concluded in a Mean Absolute Error (MAE) on 28.775,25€ and a Mean Absolute Percentage Error (MAPE) on 13.93%. Equation 11, regardless of the fragmented information that was built on, produces a MAE 22.946,28 and a MAPE 11.25%.

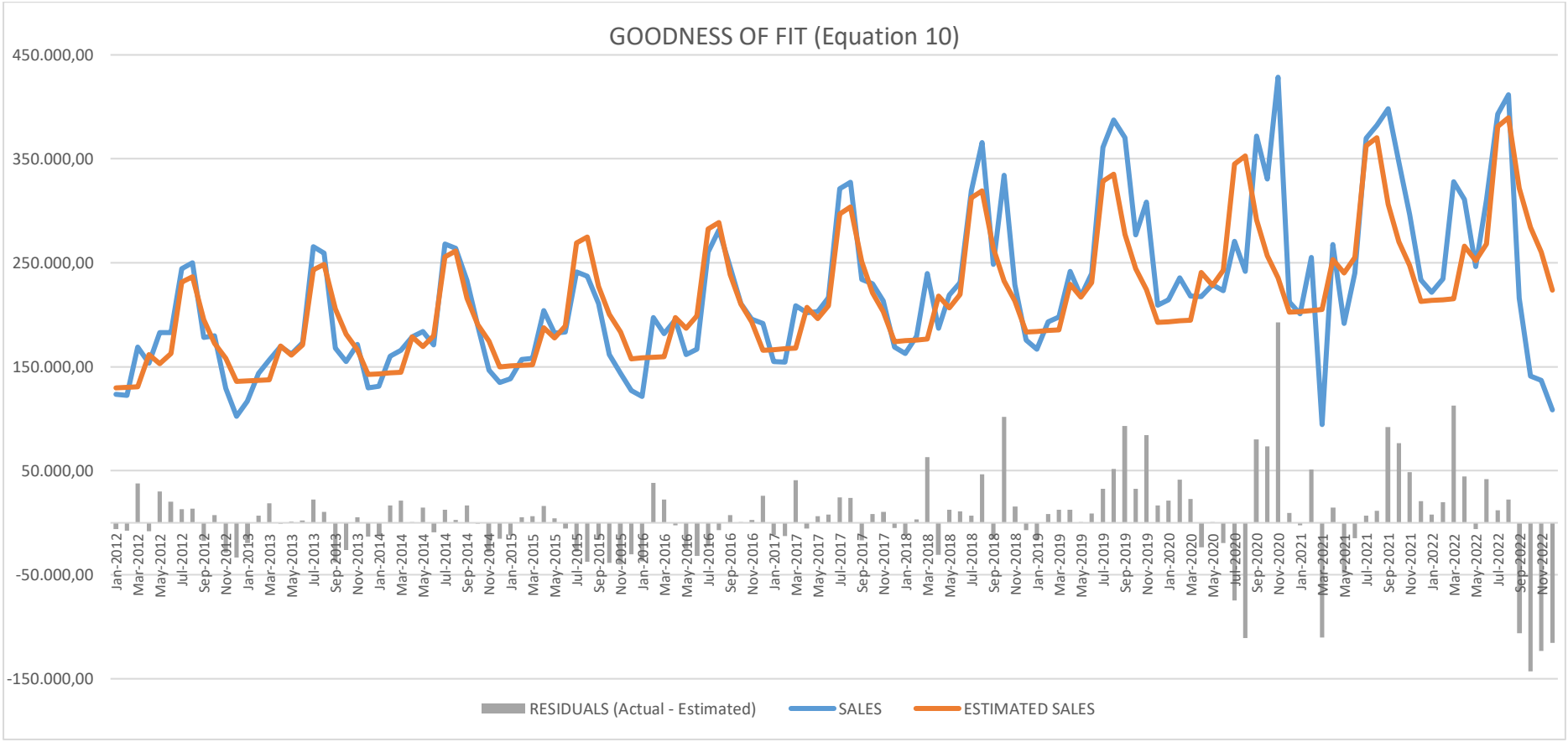

Figure 5-12: Regression performance on Equation 10

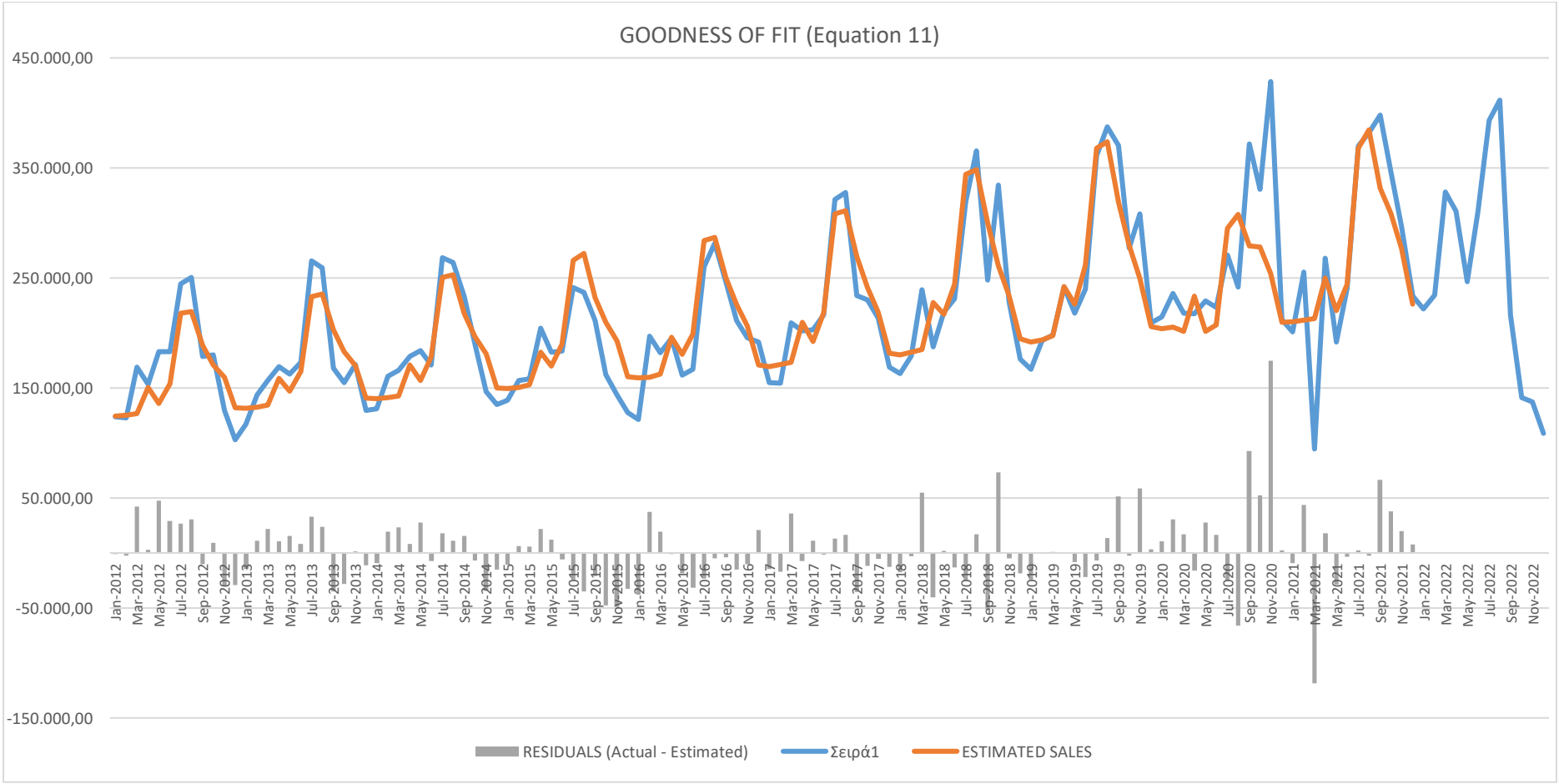

Figure 5-13: Regression performance on Equation 11

5.4. Concluding Remarks

The conclusions of the conducted analysis over the fourth product Group concentrate to the following bullet points.

- Group 4 total revenue is dominated by a seasonal pattern peaking up during summer months while exhibiting lower sales on winter. The pattern is replicated throughout the timeseries, displaying a corrugated effect. The wavy effect is distorted only in 2020 up to 2022, due to disruptions forced by the pandemic.
- The overall trend is heading upwards, providing greater variation to the aforementioned wavy effect, mainly through higher summer peaks. Indicatively, after 2016 a new sales record is accomplished year after year, with sole exception 2021, whose maximum sales were a little bit lower than 2020.
- Concerning individual subgroups, all of them demonstrate seasonal patterns, mainly between summer and winter season. Chocolate (402 subgroup) is the most differentiated among them, having its peak season in autumn, owing to product characteristics (Kouvrakis, 2023).
- Subgroups' trend also corresponds to aggregate group trend, exempt from 403 (Natural & Traditional Snacks) who exhibits a more stable performance. Interestingly however, it should by underlined that 403 is a subgroup in which several product line introductions were made during the last years (rice chips, corn cakes and others), that were not presented in the study due to timeseries inconsistency. So, in actual fact, the subgroup is enhanced, producing at the moment higher revenue than the value presented in the study.
- Correlations among subgroup sales variation were evident, yet not considerable. Only subgroup pairs 403 & 405, 404 & 405 and to a smaller extent 403 & 404 could be linearly related.
- Contrary to the alterations present in the graphical depictions of aggregated, as well as individual subgroup sales, COVID-19 pandemic was not a significant sales determinant.

- Seasonality was proved to be the dominant sales driver, either by its pure, time-related substance, or by its association with tourism arrivals in the areas were Achtida is active.
- Apart from time trend and seasonal effects (Equation 10), a regression model that held Overnights as a significant sales driver (Equation 11) will assist company's administration to corelate revenue evolution to the number of tourists in areas of interest.

6 CONCLUSION

6.1. Summary and Insights

The main reflections obtained of the analysis are summarized here, in order to provide the company with useful insights concerning the distinct characteristics of each product category, the factors that affect sales evolution, the possible correlations among each groups' evolution, as well as sales sensitivity to external factors and adverse effects stemming from COVID-19 pandemic.

To begin with, the Cosmetics product Group demonstrates a stable trend by incorporating at the same time a sheer seasonal pattern. The seasonal influence is directly related with Subgroup 103 (Sun protection) performance, which constitutes a product team under great variability, due to its use. Other subgroups also preserve the seasonal pattern, mostly Face & Body cosmetics which in essence duplicates 103 behavior. Fragrances & Gift sets display a reverse seasonal line, as sales spike during winter. For Group 1 as a whole, data distribution results in a histogram highly skewed to the left, as the majority of observations falls below the distribution mean. That is because the majority of observations contain monthly sales smaller than the average monthly sales, highlighting the distinguishable effect of increased summer sales imposed to the timeseries

Regression analysis on Group 1 verified the sheer impact of seasonal variables towards sales formation, as in a 95% significance level seasonality is proved to be responsible for approximately 91% of sales variation. Apart from that, COVID-19 pandemic era did not demonstrate significant impacts on sales, although the 1st and the 2nd Lockdown did. Indeed, restrictions applied from March 2020 up to May 2020, as well as from November 2020 through to April 2021 had a significant declining effect on sales value.

The aggregated Personal Hygiene and Cleanness Group performance is determined by an upward trend, extremely evident after 2015. Among the distinct subgroups the major contributors to Group 2 formation were Adult Diapers & Sanitary Napkins (204) and Hair and Body Cleanness (201). For the Group as a whole, a distinct seasonal pattern became apparent, constituting April, June, July, August and September the months whose mean sales kept level over 110.000€. Seasonality was further established during the regression rounds. As far as pandemic repercussions are concerned, neither lockdown impositions, nor

the vaccination program initiation were proved to inflict obstacles on sales evolution. Yet, the pandemic era itself was proved to have significant, negative effects on sales.

The timeseries revealed several differences among the subgroups' evolution as upward trend was evident for Hair & Body Cleanness products, Deodorants, and Shaving, Cotton and Pharmacy. On the other hand, Baby Care & Diapers as well as Adult Diapers & Sanitary Napkins presented a reverse evolution, with sales declining over time. Hygiene & Personal Protection subgroup demonstrated also an increasing trend, attributed mainly to pandemic effects.

The best regression performance was accomplished by two models. The best R square value achieved was 61.1% for Equation 7: Seasonality, Pandemic, Lagged Sales & Linear Trend and after that, 58.7% for Equation 6: Seasonality, Pandemic & Linear Trend. Those equations remained both present in the study, as in Equation 6, taking in account mainly factors external to the company's control, provides a tool for better assessment of external, independent forces that affect sales. On the other hand, Equation 7 which incorporates previous month's sales includes a feature that can be used for short term predictions (next month's sales) and can provide a basis for building a forecasting model.

Household, Detergents & Paper Products Group in aggregate demonstrates a clear seasonal pattern, realized through summer season peaks (April to September) and winter lows. The pattern is evident throughout the timeseries, incorporating at the same time a distinct upward trend. The seasonal pattern is not reproduced by all subgroups, as some of them demonstrate a differentiated behavior that cannot be standardized. Specifically, only Insecticides & Mosquito Repellents subgroup acquires clear cyclical behavior, affecting the overall Group 3 seasonal picture. Regarding subgroup trends, Household Equipment & Consumables, Dishwashing and Laundry perform as drivers of Groups' overall upward trend, contributing to positive sales evolution. On the other hand, Insecticides, House & WC Detergents as well as Paper Products, all demonstrate diminishing trends.

Group 3 data generates a fairly symmetrical distribution, slightly skewed to the right. The subgroup under the greatest skewness is that of Household Equipment & Consumables, whose vast majority of observations is placed to the left side of the distribution.

In contrast with Groups 1 and 2, the exploration of Group 3 statistical properties yielded differentiated remarks. Particularly, regression analysis confirmed that seasonality and time

trend are the sole drivers of sales formation. Neither pandemic variables nor macroeconomic factors seemed to impose significant alterations on sales evolution.

Snacks product Group is dominated by a seasonal pattern, as sales peak up during summer months, while exhibiting lower sales on winter. The pattern is replicated throughout the timeseries, although being quite distorted only in 2020 up to 2022, due to pandemic disruptions. The overall trend is heading upwards, as mainly after 2016 a new sales record is accomplished year after year, with sole exception 2021, whose maximum sales were a little bit lower than 2020.

Subgroups demonstrate seasonal patterns, mainly between summer and winter season. Chocolate is the most differentiated among them, having its peak season in autumn. Subgroup trend also corresponds to aggregate trend, exempt from Natural & Traditional Snacks who exhibit a more stable performance. Correlations among subgroup sales variation were evident, namely among 403 & 405, 404 & 405 and to a smaller extent 403 & 404, yet not considerable.

Regression results highlighted seasonality as the dominant sales driver, either by its pure, time-related substance, or by its association with tourism arrivals in the areas Achtida is active. On the contrary, COVID-19 related variables were not significant sales determinants.

As a general remark on data distribution symmetry, the vast majority of product Groups generated distorted results with positive skewness, indicating the great effect of extreme, high values. The frequent presence of outliers in terms of high sales values further emphasizes on seasonality effect significance. In all Group cases, seasonality proved to be the most outstanding sales driver.

COVID-19 pandemic had ambiguous implications on sales evolution. Particularly, Cosmetics' sales were affected only by restriction impositions (Lockdowns), while Personal Hygiene & Cleanness was subject to the general implications of the pandemic era. Household & Detergents were not affected at all, as well as Snacks, although visual observations led to anticipating different results.

Cosmetics' dependence on seasonal forces did not leave room for further examination of sales drivers. In contrast, Personal Hygiene & Cleanness displayed a clear dependence on previous periods sales, implying that a sales momentum is active. As mentioned above,

Household & Detergents are dominated only by seasonality and time trend, whereas Snacks are clearly affected by tourist traffic.

6.2. Field for future research

Based on those general remarks that confirm the differences among product categories, Achtida is urged to prioritize product groups or even product lines in a more detailed level, in order to perform thorough analysis directed to items with wider profit margins or those of suppliers that provide better commercial agreements. Sales exploration, regression analysis or even forecasting model deployment, after taking into account the broad reflections obtained here, can be addressed to items having greater commercial value for the company.

As mentioned in previous chapters, the capability of brands to be successful is highly dependent on the degree which supply successfully meets demand (Stanton & Baglione, 2021). Taking in account that forecasting is bound to be imprecise in a certain level, ensuring the use of proper forecasting techniques improves the probability of achieving more accurate forecasts and consequently affects operational efficiency. Since accuracy is subject to the chosen forecasting methods (Stanton & Baglione, 2021), studies highlight the need for deploying diverse forecasting applications and even combining forecasting schemes in order to obtain more versatile and accurate models (Aye, Balcilar, Gupta, & Majumdar, 2015).

The use of diverse forecasting techniques is thoroughly analyzed across literature. The utilization of neural networks and machine learning has been studied in various sectors (Agha & Alnahhal, 2012) (Sivanandam Arunraj & Ahrens, 2015) (Barboza, Kimura, & Altman, 2017) (de Campos Souza, 2020) (Afzal, Ziapour, Shokri, Shakibi, & Sobhani, 2023) (Liao, He, Wu, Wu, & Bausys, 2023). In any case, the forecasting model specifications should be properly defined, in order to avoid ignoring significant timeseries features (Thomaidis & Dounias, 2012). Moreover, a neural network regression model should be built on methodical and detailed specifications, in order to incorporate as much as possible timeseries drivers and ultimately minimize errors (Thomaidis & Dounias, 2011). However, the employment of approaches incorporating neural networks or machine learning should be executed with caution, as the phenomenon of overfitting and ultimately diminishing the interpretive robustness of a model is always present.

Barring the evident need for proper model selection in order to better fit the respective data sets, studies suggest treating product lines with high promotional influence separately, using hybrid timeseries analysis and forecasting models (Abolghasemi, Beh, Tarr, & Gerlach, 2022). Since promotions are a routine embedded in consumer product trading practice (Kouvrakis, 2023) its effect on sales evolution must not only be identified but also taken into account throughout timeseries analysis or even during the design of a forecasting model. Thus, the company should reengineer its coding scheme, in order to incorporate the respective promotional data and appreciate their effect on sales evolution.

In brief, Achtida's administration can benefit from this study in terms of setting and prioritizing forecasting goals through the identification of general, descriptive characteristics of each product group. Furthermore, combining the results as well as the suggestions made on model selection with internal information, such as profit margins or even other, commercial benefits stemming from supplier agreements, Achtida will be able to emphasize in product lines which provide greater financial advantages through promotional or other kind of control. At last, but not least, the company can benefit in terms of aiding the determination of inventory holding practices, procurement and replenishing frequency designation, via sales driver identification and ultimately achieve the best possible correspondence of demand to supply.

References

Abolghasemi, M., Beh, E., Tarr, G., & Gerlach, R. (2022). Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion. *Computers & Industrial Engineering, 142*, p. 106380. doi:https://doi.org/10.1016/j.cie.2020.106380.

Afzal, S., Ziapour, B. M., Shokri, A., Shakibi, H., & Sobhani, B. (2023, November). Building energy consumption prediction using multilayer perceptron neural network-assisted models; comparison of different optimization algorithms,. *Energy, 282*. doi:https://doi.org/10.1016/j.energy.2023.128446.

Agha, S. R., & Alnahhal, M. J. (2012, November). Neural network and multiple linear regression to predict school children dimensions for ergonomic school furniture design. *Applied Ergonomics, 43*(6), pp. 979-984. doi:https://doi.org/10.1016/j.apergo.2012.01.007.

Albright, C. S., Winston, W. L., & Zappe, C. J. (2011). *Data Analysis and Decision Making.* Ohio USA: South-Western, Cengage Learning.

Aye, G. C., Balcilar, M., Gupta, R., & Majumdar, A. (2015). Forecasting aggregate retail sales: The case of South Africa. *International Journal of Production Economics, 160*, pp. 66-79. doi:https://doi.org/10.1016/j.ijpe.2014.09.033.

Bai, B. (2022). Acquiring supply chain agility through information technology capability: the role of demand forecasting in retail industry. *Kybernetes, ahead-of-print*. doi:https://doi.org/10.1108/K-09-2021-0853

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications, 83*, pp. 405-417. doi:https://doi.org/10.1016/j.eswa.2017.04.006.

Bottani, E., Centobelli, P., Gallo, M., Kaviani, M., Jain, V., & Murino, T. (2019). Modelling wholesale distribution operations: an artificial intelligence framework. *Industrial Management & Data Systems, 119*(4), pp. 698-718. doi:https://doi.org/10.1108/IMDS-04-2018-0164

de Campos Souza, P. (2020). Fuzzy neural networks and neuro-fuzzy networks: A review the main techniques and applications used in the literature. *Applied Soft Computing, 92*. doi:https://doi.org/10.1016/j.asoc.2020.106275.

European Centre for Disease Prevention and Control. (2020, December 21). *First COVID-19 vaccine authorised for use in the European Union*. Retrieved from Media Centre: https://www.ecdc.europa.eu/en/news-events/first-covid-19-vaccine-authorised-use-european-union

Fraser, C. (2016). *Business Statistics for Competitive Advantage with Excel 2016.* Switzerland: Springer.

Giannikos, I. (2020, October). Descriptive Statistics. *Presentation at the First On-Line Tutor Student Session of the Quantitative Methods for Supply Chain Management Course of the Supply Chain Management postgraduate program of the Hellenic Open University*.

Giannikos, I. (2020, November 2). Forecasting with Regression. *Presentation at the Second On-Line Tutor Student Session of the Quantitative Methods for Supply*

Chain Management Course of the Supply Chain Management postgraduate program of the Hellenic Open University.

Hançerlioğulları, G., Sen, A., & Aktunç, E. (2016). Demand uncertainty and inventory turnover performance: An empirical analysis of the US retail industry. *International Journal of Physical Distribution & Logistics Management,*(46, No 6/7), pp. 681-708. doi:https://doi.org/10.1108/IJPDLM-12-2014-0303

Hellenic Statistical Authority. (2023, May 29). Arrivals & Overnight Stays in Hotels and Campings per Region, per Month, 2012-2021, months 1-12.

Hellenic Statistical Authority. (Retrieved on April 29, 2023). *Economy, Indicies*. Retrieved from Hellenic Statistical Authority: https://www.statistics.gr/en/statistics/eco

Hellenic Statistical Authority. (Retrieved on April 30, 2023). *Gross domestic product / 4th Quarter 2022 / Timeseries.* Retrieved from Hellenic Statistical Authority: https://www.statistics.gr/en/statistics/-/publication/SEL84/-

Hellenic Statistical Authority. (Retrieved on May 1st). *Tourism (Turnover index) / 4th Quarter 2022 / Timeseries.* Retrieved from Hellenic Statistical Authority: https://www.statistics.gr/el/statistics/-/publication/DKT51/-

Kim, S., Kim, H., & Lu, J.-C. (2019). A practical approach to measuring the impacts of stockouts on demand. *Journal of Business & Industrial Marketing, 34*(4), pp. 891-901. doi:https://doi.org/10.1108/JBIM-04-2018-0126

King, A. P., & Eckersley, R. J. (2019). Chapter 2 - Descriptive Statistics II: Bivariate and Multivariate Statistics. In A. P. King, & R. J. Eckersley, *Statistics for Biomedical Engineers and Scientists* (pp. 23-56). Academic Press. doi:https://doi.org/10.1016/B978-0-08-102939-8.00011-6

Kouvrakis, S. E. (2023, May 15). CEO, Achtida S.A. (D. P. Kappa, Interviewer)

Lehtonen, J., Småros, J., & Holmström, J. (2005). The effect of demand visibility in product introductions. *International Journal of Physical Distribution & Logistics Management, 35*(2), pp. 101-115. doi:https://doi.org/10.1108/09600030510590291

Liao, H., He, Y., Wu, X., Wu, Z., & Bausys, R. (2023). Reimagining multi-criterion decision making by data-driven methods based on machine learning: A literature review. *Information Fusion*. doi:https://doi.org/10.1016/j.inffus.2023.101970.

Lord, D., Qin, X., & Geedipally, S. R. (2021). Exploratory analysis of safety data. In *Highway Safety Analytics and Modeling* (pp. 135-177). Elsevier. doi:https://doi.org/10.1016/B978-0-12-816818-9.00015-9

Ministry of Climate Change & Civil Protection. (2020, March 13). *Additional measures for Coronavirus spread restriction*. Retrieved from Press Releases: https://civilprotection.gov.gr/deltia-tupou/epiprostheta-metra-gia-ton-periorismo-tis-exaplosis-toy-koronoioy

Ministry of Climate Change & Civil Protection. (2020, November 5). *Restrictive Measures Specialization*. Retrieved from Press Releases: https://civilprotection.gov.gr/deltia-tupou/exeidikeysi-neon-metron-gia-tin-antimetopisi-tis-pandimias-apo-ton-yfypoyrgo-politikis

Salaverry, J., & White III, E. (2009). Improving procurement through regression analysis: A case study of predicting argentine jet fuel prices. *Journal of Public Procurement, 9*(1), pp. 1-16. doi:https://doi.org/10.1108/JOPP-09-01-2009-B001

Sivanandam Arunraj, N., & Ahrens, D. (2015). A hybrid seasonal autoregressive integrated moving average and quantile regression for daily food sales forecasting,. *International Journal of Production Economics, 170*(Part A), pp. 321-335. doi:https://doi.org/10.1016/j.ijpe.2015.09.039.

Sodero, A. C., Namin, A., Gauri, D. K., & Bhaskaran, S. R. (2021). The strategic drivers of drop-shipping and retail store sales for seasonal products. *Journal of Retailing, 97*(4), pp. 561-581. doi:https://doi.org/10.1016/j.jretai.2021.09.001.

Sotiriou, A. (2021, February 26). *Το ημερολόγιο της πανδημίας στην Ελλάδα*. Retrieved from Money Review: https://www.moneyreview.gr/society/19969/to-imerologio-tis-pandimias-stin-ellada/

Stanton, J., & Baglione, S. (2021). Competing Forecasting Techniques Produces Different Optimal Solutions Based on the Product. In K. Lawrence, & R. Klimberg (Ed.), *Advances in Business and Management Forecasting* (Vol. 14, pp. 101-115). Emerald Publishing Limited. doi:https://doi.org/10.1108/S1477-407020210000014008

Thomaidis, N. (2021, November 21). Simple Forecasting Techniques. *Lecture Notes and Excel Files during the 2nd Tutor - Student Session on the Forecasting Module of the Postgraduate Program Supply Chain Management of the Hellenic Open University*.

Thomaidis, N. S. (2021, October 17). Principles of Quantitative Forecasting. *Lecture Notes on the 2nd Tutor - Student Session on the Forecasting Module of the Postgraduate Program Supply Chain Management of the Hellenic Open University.*

Thomaidis, N. S. (2021, December 19). Special Forecasting Topics. *Excel Files on the 3rd Tutor - Student Session on the Forecasting Module of the Postgraduate Program Supply Chain Management of the Hellenic Open University*.

Thomaidis, N. S., & Dounias, G. D. (2011, May). On detecting the optimal structure of a neural network under strong statistical features in errors. *Journal of Time Series Analysis, 32*(3), pp. 204-222. doi:https://doi.org/10.1111/j.1467-9892.2010.00693.x

Thomaidis, N. S., & Dounias, G. D. (2012, March). A comparison of statistical tests for the adequacy of a neural network regression model. *Quantitative Finance, 12*(3), pp. 437-449. doi:https://doi.org/10.1080/14697680903426573

Wang, C.-H. (2022). Considering economic indicators and dynamic channel interactions to conduct sales forecasting for retail sectors. *Computers & Industrial Engineering, 165*. doi:https://doi.org/10.1016/j.cie.2022.107965.

World Health Organization. (Retrieved on May 12th 2023). *Statement on the fifteenth meeting of the IHR (2005) Emergency Committee on the COVID-19 pandemic*. Retrieved from News: https://www.who.int/news/item/05-05-2023-statement-on-the-fifteenth-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-coronavirus-disease-(covid-19)-pandemic?adgroupsurvey={adgroupsurvey}&gclid=Cj0KCQjwpPKiBhDvARIsAC

World Health Organization. (Retrieved on May 5th, 2023). *Overview*. Retrieved from Coronavirus disease (COVID-19) pandemic: https://www.who.int/europe/emergencies/situations/covid-19

Zaiontz, C. (Retrieved on May 10th, 2023). *Creating Boxplots in Excel*. Retrieved from Real Statistics Using Excel: https://real-statistics.com/descriptive-statistics/box-plots/

Appendix A: Excluded sales values

Based on Group Coding Scheme

901 Shelves / Stands

902 Pallets

980 Discounts

Based on Subgroup Coding Scheme

011 Mondelez Stands

015 J&J Coupons

017 J&J Stands

069 TESTERS

071 Sarandis Coupons

072 Various Stands

159 SCA Coupons

279 Colgate Palmolive Coupons

356 JDE Coupons

900 Taxes

980 Discounts

Author's Statement:

I hereby expressly declare that, according to the article 8 of Law 1559/1986, this dissertation is solely the product of my personal work, does not infringe any intellectual property, personality and personal data rights of third parties, does not contain works/contributions from third parties for which the permission of the authors/beneficiaries is required, is not the product of partial or total plagiarism, and that the sources used are limited to the literature references alone and meet the rules of scientific citations.