



School of Social Sciences
Supply Chain Management

Postgraduate Dissertation

“Leveraging Data Analytics for Enhanced Demand Forecasting
in the Smart Meter Industry - The case of Landis+Gyr S.A.”

Pappas Ntominik Anastasios

Supervisor: Dr. Pazarskis Michail

Kiato, Greece, December 2024

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



“Leveraging Data Analytics for Enhanced Demand Forecasting in the Smart Meter Industry - The case of Landis+Gyr S.A.”

Pappas Ntominik Anastasios

Supervising Committee

Supervisor:

Dr. Pazarskis Michail

Kiato, Greece, December 2024

I would like to thank:

Professor Michail Pazarskis for his guidance and support, and
professor Nikolaos Thomaidis for his motivation and inspiration.

“Dedicated to my family and Maria”

Abstract

This dissertation explores the application of data analytics to enhance demand forecasting within the smart meter industry, focusing on the case of Landis+Gyr S.A.'s production plant located in Corinth, Greece. Utilizing historical monthly sales data from 04.2018 to 03.2024 for high-demand smart meter products, the study aims to identify patterns and improve forecast accuracy to support data-driven decision-making in supply chain management.

Traditional forecasting methods, such as Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression models, are applied and evaluated for accuracy in predicting sales across product families. The analysis also examines the effects of significant external factors, including the COVID-19 pandemic and the Russian invasion of Ukraine, on demand patterns.

Findings contribute to advancing forecasting methodologies and underscore their value for operational efficiency and strategic planning, not only at Landis+Gyr but also within the broader manufacturing sector.

Keywords: Demand Forecasting, Smart Meter Industry, Data Analytics, Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), Regression Analysis, Python, Data-Driven Decision Making, Landis+Gyr S.A.

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει την εφαρμογή της ανάλυσης δεδομένων για την ενίσχυση της πρόβλεψης ζήτησης στη βιομηχανία «έξυπνων» μετρητών, εστιάζοντας στην περίπτωση του εργοστασίου παραγωγής της εταιρείας Landis+Gyr A.E. με έδρα στην Κόρινθο. Χρησιμοποιώντας ιστορικά μηνιαία δεδομένα πωλήσεων κατά την περίοδο 04.2018-03.2024 για προϊόντα υψηλής ζήτησης, η μελέτη αποσκοπεί στον εντοπισμό μοτίβων και στη βελτίωση της ακρίβειας πρόβλεψης, προκειμένου να υποστηρίξει τη λήψη αποφάσεων στη διαχείριση της εφοδιαστικής αλυσίδας.

Εφαρμόζονται παραδοσιακές μέθοδοι πρόβλεψης, όπως η Εκθετική Κινητή Μέση Τιμή (EWMA), η Απλή Κινητή Μέση Τιμή (SMA) και τα μοντέλα γραμμικής παλινδρόμησης και παράλληλα αξιολογείται η ακρίβειά τους στην πρόβλεψη πωλήσεων για διάφορες κατηγορίες προϊόντων. Η ανάλυση εξετάζει επίσης τις επιπτώσεις σημαντικών εξωτερικών παραγόντων, όπως η πανδημία COVID-19 και η ρωσική εισβολή στην Ουκρανία, στις τάσεις της ζήτησης.

Τα ευρήματα συμβάλλουν στην προώθηση των μεθοδολογιών πρόβλεψης και υπογραμμίζουν τη σημασία τους για την επιχειρησιακή αποτελεσματικότητα και τον στρατηγικό σχεδιασμό, όχι μόνο για τη Landis+Gyr, αλλά και για τον ευρύτερο τομέα της δευτερογενούς παραγωγής και βιομηχανίας.

Λέξεις-Κλειδιά: Πρόβλεψη Ζήτησης, Βιομηχανία Έξυπνων Μετρητών, Ανάλυση Δεδομένων, Εκθετική Κινητή Μέση Τιμή (EWMA), Απλή Κινητή Μέση Τιμή (SMA), Python, Ανάλυση Γραμμικής Παλινδρόμησης, Λήψη Αποφάσεων βάσει Δεδομένων, Landis+Gyr A.E.

Table of Contents

Abstract.....	5
Περίληψη	6
List of Figures.....	9
List of Tables.....	9
Introduction.....	10
Research Context.....	10
Research Questions and Objectives	11
Purpose and Significance of the Study	13
Structure of the Dissertation	14
Literature Review.....	16
Overview of the European Smart Meter Industry	16
Overview of Forecasting Techniques	16
Comparison of Forecasting Models.....	18
Application of Python in Forecasting Techniques	21
Relevant Research and Case Studies	22
Impact of External Factors on Demand Forecasting	27
Methodology.....	31
Data Description	31
Forecasting Models and Techniques	33
Exponential Weighted Moving Average (EWMA).....	33
Simple Moving Average (SMA)	36
Regression Analysis	40
Accuracy Measurement.....	44
Empirical Study	47
Preliminary Data Analysis	47
Interdependencies among the product families' revenues	53
External factor effect in Landis+Gyr revenues.....	55
Variable significance and decision.....	55
Covid19 Pandemic effect	57
Russian Invasion in Ukraine	58

Application of Forecasting Models.....	61
SMA Application and Results	61
EWMA Application and Results	65
Regression Analysis Results	68
Model Comparison and Effectiveness.....	73
Conclusions, Limitations, and Future Research.....	77
Summary of Findings.....	77
Research Limitations	79
Recommendations for Future Research.....	79
References	81
Appendix	84
List of Python Scripts	84

List of Figures

1. Figure 1: E360 – Rolling Mean and Variance
2. Figure 2: E360 Monthly Sales
3. Figure 3: Boxplot of E360 Sales
4. Figure 4: Rolling Mean and Variance for E450 Sales
5. Figure 5: E450 Monthly Sales
6. Figure 6: Boxplot of E450 Sales
7. Figure 7: Rolling Mean and Variance for FNN Sales
8. Figure 8: FNN Monthly Sales
9. Figure 9: Boxplot of FNN Sales
10. Figure 10: Monthly Sales data before and after the Russian invasion in Ukraine
11. Figure 11: SMA performance for E360 Product Family
12. Figure 12: Forecast Error evolution concerning E360 forecasting sample
13. Figure 13: SMA performance for E450 Product Family
14. Figure 14: Forecast Error evolution concerning E450 forecasting sample
15. Figure 15: SMA performance for FNN Product Family
16. Figure 16: Forecast Error evolution concerning FNN forecasting sample
17. Figure 17: MAPE variation based on λ value
18. Figure 18: EWMA performance for E360 Product Family
19. Figure 19: MAPE variation based on λ value
20. Figure 20: EWMA performance for E450 Product Family
21. Figure 21: MAPE variation based on λ value
22. Figure 22: EWMA performance for FNN Product Family
23. Figure 23: Regression performance for E360 Product Family
24. Figure 24: Regression performance for E450 Product Family
25. Figure 25: Regression performance for FNN Product Family

List of Tables

1. Table 1. MAPE qualitative criteria.
2. Table 2. Interdependence between products revenues check – Regression results.
3. Table 3. Statistically significant variables – Regression results.
4. Table 4. Statistically significant variables – Regression results.
5. Table 5. Covid_19 pandemic effect check – Regression results.
6. Table 6. Russian Invasion in Ukraine effect check – Regression results.
7. Table 7. Computed MAE & MAPE values per product family for SMA application
8. Table 8. Computed MAE & MAPE values per product family for EWMA application
9. Table 9. Computed MAE & MAPE values per product family
10. Table 10. E360 Regression results (only statistically significant variables depicted)
11. Table 11. E450 Regression results (only statistically significant variables depicted)
12. Table 12. FNN Regression results (only statistically significant variables depicted)

Introduction

Research Context

In the manufacturing and supply chain sectors, effective demand forecasting is crucial for enhancing operational efficiency, reducing costs, and meeting customer expectations. Accurate demand predictions allow companies to adjust their production rates, manage inventory levels effectively, and optimize resource allocation. The case study in this dissertation focuses on Landis+Gyr, a significant player in the meter manufacturing industry located in Athens, with an emphasis on three high-demand meter product families. The company has provided monthly sales data for these products spanning from 2018 to 2024 and including data for 6 fiscal years, offering a solid foundation for analysis and prediction.

The relevance of demand forecasting in manufacturing has been underscored by previous studies, such as those conducted on over-the-top (OTT) streaming platforms like Netflix and in durable consumer goods industries, like battery sales. These industries face similar challenges, where demand can fluctuate due to a variety of internal and external factors, including seasonality, market competition, and global events. By adopting data-driven forecasting techniques, companies can gain insights into customer behavior, enhance their ability to forecast future needs accurately, and maintain a competitive edge. For example, studies on Netflix have utilized data analytics to predict customer demand trends, adapting their content and service offerings accordingly to maximize customer satisfaction and engagement. Similarly, the consumer goods industry applies forecasting methods to anticipate seasonal or external impacts, such as the COVID-19 pandemic, which has had substantial effects on consumption patterns and supply chain stability.

In this context, we seek to apply similar forecasting methodologies to better understand the demand dynamics of Landis+Gyr smart meter products. By analyzing historical sales patterns and leveraging forecasting models such as Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis, the company aims to determine which method provides the most accurate predictions. This research will consider both intrinsic factors, like product performance over time, and extrinsic variables, such as market disruptions, to develop a comprehensive

understanding of the demand behavior in the meter industry. Such insights are invaluable, enabling the company to proactively manage its production processes, optimize stock levels, and ultimately achieve a resilient and responsive supply chain.

Landis+Gyr is a leading global company that specializes in smart metering solutions and energy management systems. Established in 1896 and headquartered in Switzerland, the company focuses on providing advanced technologies to measure and optimize the use of electricity, gas, and water. Their solutions are designed to help utilities modernize their infrastructure, improve energy efficiency, and support the integration of renewable energy sources.

By offering smart meters, data analytics tools, and grid management systems, Landis+Gyr enables energy providers to enhance their operations and offer better services to consumers. These innovations also empower individuals to monitor and manage their energy consumption more effectively, contributing to sustainability efforts and reducing environmental impact.

With a presence in over 30 countries, Landis+Gyr works with a wide range of clients, from local utility providers to large-scale energy companies, playing a key role in the global shift toward smarter, greener energy systems.

Research Questions and Objectives

The primary objective of this dissertation is to develop a comprehensive demand forecasting model tailored to Landis+Gyr's specific product lines in the meter manufacturing sector. The aim is to utilize historical sales data concerning 2018-2023 fiscal years to discern patterns and trends, allowing the company to make data-driven decisions regarding inventory management, production scheduling, and resource allocation. Through this study, Landis+Gyr aims to identify the most accurate forecasting model for its high-demand meter products, thereby minimizing demand uncertainty and enhancing its operational efficiency.

To achieve this objective, the research will address the following core questions:

- What are the main statistical properties and components of the sales time series for Landis+Gyr's different meter products?

This question seeks to explore the characteristics of the historical sales data, such as seasonality, trends, and volatility, to identify which statistical elements significantly influence the demand patterns for each product family.

- How accurate are different forecasting models in predicting monthly sales for the company's smart meter products, and does this accuracy vary across different product families?

This question aims to compare the predictive effectiveness of various models, such as Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis, to determine which model provides the highest accuracy in demand forecasting. Additionally, the study will assess whether the accuracy of these models differs when applied to distinct meter product categories.

- Are there any interdependencies or cross-correlations among the sales of different product families, and how do these relationships impact market behavior?

By investigating potential interdependencies among product families, this question examines how the demand for one product might influence or correlate with the demand for others. Understanding these relationships can provide insights into cross-selling opportunities or potential demand cannibalization within the product portfolio.

- How have external factors, such as the COVID-19 pandemic and geopolitical events like the Russian invasion of Ukraine, affected sales patterns and market dynamics for the company's products?

This question focuses on assessing the impact of external, macroeconomic factors on sales behavior, particularly events that may disrupt typical demand patterns. Recognizing these influences can help Landis+Gyr adjust its forecasting models to account for extraordinary events that may skew demand projections.

By addressing these research questions, this dissertation aims to achieve a dual objective: firstly, to enhance the forecasting process for Landis+Gyr through the identification of a reliable and precise prediction model, and secondly, to contribute to the broader literature on demand forecasting within the manufacturing sector. This research is expected to yield actionable insights that can not only improve operational efficiency at Landis+Gyr but also serve as a valuable reference for similar manufacturing settings seeking to optimize their Sales & Operations Planning (S&OP) processes.

Purpose and Significance of the Study

The main purpose of this study is to explore and implement data analysis and forecasting techniques to enhance demand prediction accuracy within the meter production sector, specifically for the production plant of Landis+Gyr S.A., a leading entity in this field based in Corinth, Greece. By focusing on historical sales data of high-demand meter product families from 2018 to 2024, this study aims to develop a robust forecasting framework that can identify past patterns and anticipate future demand. Through improving forecast accuracy, Landis+Gyr can achieve optimal inventory management, efficient resource allocation, and ultimately strengthen its decision-making process in supply chain management.

In the production and supply chain management field, accurate demand forecasting is a key driver for organizational success, as it directly impacts operational efficiency and financial stability. Drawing from the insights provided by similar studies in demand forecasting, like those conducted on OTT streaming platforms such as Netflix, and in consumer goods industries, such as battery sales for durable goods, this study will apply methods like Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis to determine the best-fitting forecasting model. These approaches are widely recognized for their adaptability and accuracy in forecasting,

particularly when handling time series data with potentially impactful external influences, such as economic disruptions or global events.

Beyond the immediate goals, this research also seeks to contribute to the broader understanding of how predictive models can be adapted to manufacturing settings. Similar to the way streaming and consumer goods industries use data-driven forecasting to handle product demand fluctuations, this study investigates how these techniques can optimize production planning for Landis+Gyr. By using demand forecasting as a strategic tool, the company not only can enhance its competitive advantage within the regional market but also mitigate risks associated with demand uncertainty. This research, therefore, holds significance for both academic contributions in predictive modeling for supply chains and practical applications in manufacturing optimization.

Structure of the Dissertation

This dissertation is organized into four main chapters, each designed to address specific aspects of the research questions and objectives. The structure follows a logical progression, beginning with an introduction to the research and its context, then advancing through a comprehensive review of relevant literature, a detailed explanation of the chosen methodology, an empirical analysis of the data, and concluding with an evaluation of the findings and suggestions for future research.

Literature Review.

This chapter reviews previous research and theoretical frameworks relevant to demand forecasting and supply chain management. The discussion focuses on traditional and advanced forecasting models, such as Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis. By examining similar case studies from the OTT streaming and consumer goods industries, the literature review provides a contextual basis for the methodologies applied in this study and highlights the benefits and limitations of various forecasting techniques.

Methodology

The methodology chapter describes the research approach and the specific forecasting models applied to the data. It begins with a description of the dataset, outlining the monthly sales figures for Landis+Gyr's smart meter products concerning the fiscal years 2018-2023. The chapter then explains the steps taken to implement EWMA, SMA, and regression models, detailing how each technique contributes to achieving the research objectives. Additionally, it describes the accuracy metrics (MAE and MAPE) used to evaluate the forecasting models, providing a clear framework for assessing model effectiveness.

Empirical Study

This chapter presents the empirical analysis of the data, applying the forecasting models to the historical sales figures for each product family. The empirical study is divided into sections based on each model—EWMA, SMA, and regression analysis—detailing the application process and the results of each method. It also includes a comparison of model performance and effectiveness, using MAE and MAPE metrics to evaluate accuracy. The findings from this analysis inform the selection of the most suitable forecasting model for Landis+Gyr's products.

Conclusions, Limitations, and Future Research

The penultimate chapter discusses the research findings, evaluates the effectiveness of the applied forecasting models, and highlights the practical implications for Landis+Gyr. It also addresses limitations encountered during the research, such as data constraints or external factors that may have influenced demand patterns. This chapter concludes with recommendations for future research, suggesting potential improvements in forecasting techniques and areas for further investigation to enhance demand forecasting in the manufacturing sector.

This structured approach enables a thorough examination of the research topic, ensuring a cohesive flow from theoretical foundations to practical application and analysis. Each chapter builds upon the previous one, providing the reader with a complete understanding of demand forecasting challenges and solutions tailored to Landis+Gyr's unique production environment.

Literature Review

Overview of the European Smart Meter Industry

The European smart meter market has experienced significant growth over the past decade, driven by regulatory mandates and the transition to renewable energy sources. The European Union's Energy Efficiency Directive 2012/27/EU has been a pivotal driver, requiring member states to implement smart metering systems to enhance energy efficiency and empower consumers with real-time data on their energy consumption (Huhta, 2017). These regulations have set the foundation for widespread smart meter deployments across Europe.

Technological advancements and the integration of renewable energy sources have further accelerated market growth. The adoption of Internet of Things (IoT) technologies, artificial intelligence (AI), and data analytics has enhanced demand-side energy management and grid efficiency (Albu et al., 2021). These innovations enable utilities to optimize energy distribution and offer consumers detailed insights into their consumption patterns, fostering more informed energy usage decisions.

Market projections indicate robust growth for the European smart meter industry, driven by increasing energy demand, the integration of renewable energy sources, and a heightened focus on reducing carbon emissions. According to industry reports, the market is expected to achieve a compound annual growth rate (CAGR) of over 10% through 2030, reflecting the region's commitment to modernizing its energy infrastructure (Schittekatte & Meeus, 2020).

Landis+Gyr is a prominent player in the European smart meter market, competing with other key manufacturers such as Itron, Siemens, and Schneider Electric. These companies have established significant market shares and regional presences, offering a range of smart metering solutions tailored to diverse utility requirements across Europe (Schittekatte & Meeus, 2020).

Landis+Gyr's competitive strengths include its global reach, extensive product portfolio, and commitment to technological innovation. The company's focus on integrating advanced technologies like AI and IoT into its smart metering solutions positions it favorably against competitors. Additionally, Landis+Gyr's experience in large-scale deployments and its ability to customize solutions for various markets enhance its appeal to utilities seeking reliable and future-proof metering systems (Schittekatte & Meeus, 2020).

Overview of Forecasting Techniques

Demand forecasting plays a crucial role in supply chain management and production planning, as it enables companies to anticipate future demand and adjust their operations accordingly. In this study, three primary forecasting techniques are utilized:

Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis. Each of these techniques offers distinct advantages for analyzing historical data and predicting future demand trends, and their effectiveness will be assessed in the context of Landis Gyr's high-demand meter products.

1. Exponential Weighted Moving Average (EWMA)

The Exponential Weighted Moving Average (EWMA) is a time series forecasting technique that applies exponentially decreasing weights to past data, emphasizing more recent observations. This technique is particularly useful for datasets where recent trends are more indicative of future demand, as it assigns a higher weight to recent data points while still accounting for the influence of past data. In EWMA, a smoothing constant (λ) between 0 and 1 controls the degree of weighting decrease, where a value closer to 1 places more emphasis on recent data.

The formula for EWMA is:

$$EWMA_t = \lambda * X_{(t-1)} + (1 - \lambda) * EWMA_{(t-1)}$$

where $EWMA_t$ is the forecasted value at time, X_{t-1} is the actual value at time $t-1$, $EWMA_{(t-1)}$ is the EWMA value at the previous time step and λ determines the weight of previous observations. This model is advantageous for companies like Landis+Gyr, as it can adapt to fluctuations in demand while smoothing out short-term variations. Prior studies, such as those conducted on OTT streaming platforms, have shown that EWMA is effective for forecasting trends influenced by seasonality or sudden demand changes, making it a suitable choice for the company's high-demand product families.

2. Simple Moving Average (SMA)

The Simple Moving Average (SMA) is another widely used time series forecasting technique, which calculates the average of past observations over a fixed period. Unlike EWMA, which assigns different weights to each data point, SMA applies equal weighting to all data within the selected time window. The SMA is calculated as follows:

$$SMA = \frac{1}{N} \sum_{i=0}^{N-1} Y_{t-i}$$

where N is the number of observations in the time window, and $Y_{(t-i)}$ represents the historical data values.

SMA is beneficial for identifying underlying trends in data over a stable period, as it smooths out short-term fluctuations. Although SMA may lack the flexibility of EWMA in adapting to sudden changes, it provides a clear representation of long-term trends. This technique is valuable for Landis+Gyr when predicting stable demand patterns for their products, especially in cases where historical sales data does not exhibit significant volatility. As seen in applications within the consumer goods industry, SMA is effective for businesses with consistent, steady demand, making it a suitable option for forecasting average monthly sales for the company's meters.

3. Regression Analysis

Regression analysis is a statistical method used to identify relationships between a dependent variable (in this case, future demand) and one or more independent variables (such as historical sales, seasonality, or external events like COVID-19). This technique is valuable for understanding the impact of specific factors on demand patterns and allows for the incorporation of multiple variables into the forecasting model. Simple linear regression, as well as multiple regression models, can be applied to capture these relationships.

The general form of a linear regression model is:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + e$$

where Y is the forecasted demand, $\beta(0)$ is the intercept, $\beta(1), \beta(2), \dots, \beta(n)$ are the coefficients for each independent variable and e is the error term.

Comparison of Forecasting Models

When comparing SMA, EWMA, and regression analysis, it is evident that each technique has unique strengths and limitations, making them suitable for different contexts. Selecting the appropriate forecasting model is crucial in achieving reliable and accurate demand predictions for the company's meter products. This section compares the three primary forecasting models utilized in this study—Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis—in terms of their suitability, adaptability, and effectiveness for different demand patterns and data characteristics. By understanding the strengths and limitations of each model,

Landis+Gyr can make informed decisions about which model best addresses its demand forecasting needs.

Exponential Weighted Moving Average (EWMA) vs. Simple Moving Average (SMA)

The key distinction between EWMA and SMA lies in how they weight past data points. EWMA assigns exponentially decreasing weights to past observations, giving more emphasis to recent data. This characteristic makes EWMA particularly suitable for datasets where demand patterns may shift over time, as it captures recent trends more effectively than SMA. In the context of the company's products, EWMA is advantageous when predicting demand during periods of volatility or when external factors may have shifted demand patterns.

On the other hand, SMA applies equal weight to all observations within the selected time window, making it more stable but less responsive to sudden changes in demand. SMA is advantageous for identifying long-term trends in stable demand environments. However, it may be less effective in scenarios where demand fluctuates due to unexpected events, such as the COVID-19 pandemic. SMA provides a straightforward approach to smoothing data, and it is easier to interpret; however, its lack of sensitivity to recent changes may limit its accuracy in volatile markets.

In summary, EWMA is generally more adaptable to short-term fluctuations in demand and is ideal for capturing recent trends, while SMA is preferred for stable demand patterns and long-term trend analysis. In prior case studies, such as those on streaming platforms and battery sales, EWMA has shown superior performance in environments with dynamic demand, whereas SMA is frequently used for historical analysis in consistent markets.

Regression Analysis vs. Moving Averages

While both EWMA and SMA are time series methods focusing on historical demand patterns, regression analysis provides a different approach by examining the relationships between demand and a set of independent variables. This makes regression analysis particularly useful for incorporating external factors, such as

macroeconomic conditions, seasonal effects, and specific events like the COVID-19 pandemic or geopolitical issues. For Landis+Gyr, regression analysis allows for the exploration of how various external and internal factors interact to influence demand, providing a more comprehensive understanding of demand drivers.

The advantage of regression analysis over moving averages lies in its flexibility. By including multiple variables, regression analysis can reveal correlations and causal relationships that might not be visible through time series methods alone. However, the complexity of regression models and the need for a larger dataset can be a drawback, as they require more extensive data preparation and statistical knowledge to interpret results accurately. Additionally, the model's accuracy depends on the quality and relevance of the chosen independent variables; irrelevant or poorly chosen variables can reduce the model's predictive accuracy.

Compared to moving averages, regression analysis is particularly beneficial when external factors significantly impact demand, as it allows for more targeted forecasting adjustments. For instance, in studies within the consumer goods sector, regression analysis has proven effective in capturing demand changes due to seasonal or economic fluctuations, demonstrating its value for the company's forecasting efforts.

Practical Considerations and Model Selection

Each forecasting model has practical considerations that must be taken into account when selecting the best approach for demand forecasting. EWMA and SMA offer straightforward, easy-to-implement techniques that provide reasonably accurate forecasts with minimal data requirements. These methods are well-suited for short-term forecasting and can be particularly useful for the company if the goal is to maintain simplicity and computational efficiency in demand forecasting.

Regression analysis, while more complex, provides a higher level of flexibility and precision in understanding demand influences, especially when external factors are involved. However, it requires more sophisticated data processing and statistical knowledge, as well as a robust dataset to ensure accurate results. For Landis+Gyr, choosing the appropriate model depends on the availability of data and the desired level of detail in forecasting. If the company's goal is to develop a comprehensive

understanding of demand influenced by both internal and external variables, regression analysis may be the preferred choice. Conversely, if simplicity and ease of use are prioritized, moving averages—especially EWMA—might be more appropriate for immediate application in forecasting.

In conclusion, EWMA, SMA, and regression analysis each offer unique advantages. EWMA is recommended for scenarios requiring responsiveness to recent demand changes, while SMA is optimal for consistent demand patterns. Regression analysis, with its multi-variable capability, is the best choice for an in-depth, factor-driven analysis of demand. For Landis+Gyr, the choice of model should align with both data availability and the forecasting needs dictated by the market environment, balancing accuracy with practicality.

Emerging machine learning techniques could further enhance these models. For instance, Random Forest and Gradient Boosting can handle complex, nonlinear relationships, while LSTM networks are designed to capture sequential dependencies in time series data (Sharma et al., 2021; Abbasimehr & Shabani, 2020). Future studies integrating these advanced approaches could offer more robust and scalable solutions.

Application of Python in Forecasting Techniques

Python has become a pivotal tool in the realm of data analysis and forecasting, offering a plethora of libraries and frameworks that facilitate the implementation of various forecasting techniques. Among these, Simple Moving Average (SMA), Exponentially Weighted Moving Average (EWMA), and regression-based methods are prominently utilized. For the purpose and convenience of our research, we will apply the aforementioned forecasting methods and delve into the application of Python.

1. Simple Moving Average (SMA)

SMA is a fundamental technique that calculates the average of a fixed number of past observations to smooth out short-term fluctuations and highlight longer-term trends. In Python, the pandas library provides efficient functions to compute SMA. For instance, the `rolling()` function combined with `mean()` can be employed to calculate the moving average over a specified window.

Recent studies have leveraged Python's capabilities to implement SMA in various domains. For example, in financial data analysis, SMA is commonly used to estimate volatility and identify trends in stock prices (Baglini, 2024).

2. Exponentially Weighted Moving Average (EWMA)

EWMA assigns exponentially decreasing weights to past observations, giving more significance to recent data points. This method is particularly effective in scenarios where recent observations are more indicative of future trends. Python's pandas library offers the `ewm()` function, which facilitates the computation of EWMA with ease.

In recent research, Python's EWMA implementation has been applied to monitor and detect shifts in data patterns. For instance, Mammadova and Özkale (2024) developed control charts based on EWMA to monitor count data profiles, effectively detecting shifts in data distributions.

3. Regression Forecasting

Regression analysis models the relationship between a dependent variable and one or more independent variables, enabling predictions based on this relationship. Python's statsmodels and scikit-learn libraries provide comprehensive tools for implementing both linear and non-linear regression models.

A notable application is demonstrated by Inyurt et al. (2020), who employed Python to develop Gaussian Process Regression (GPR) and Multiple Linear Regression (MLR) models for forecasting ionospheric Total Electron Content (TEC) in Turkey. Their study highlights Python's proficiency in handling complex regression models for forecasting purposes.

Relevant Research and Case Studies

Accurate demand forecasting is crucial in the smart meter industry to ensure efficient production planning and inventory management. Various forecasting models have been applied, including Simple Moving Average (SMA), Exponential Weighted Moving Average (EWMA), regression analysis, and machine learning techniques. Recent

studies have explored the effectiveness of these models in energy management contexts.

For instance, Zhang and Dong (2022) applied machine learning algorithms to predict energy consumption patterns, demonstrating improved accuracy over traditional methods. Similarly, Schittekatte and Meeus (2020) examined the application of advanced statistical models in forecasting electricity demand, highlighting the benefits of integrating multiple forecasting techniques to enhance precision.

These studies underscore the importance of selecting appropriate forecasting models tailored to the specific characteristics of the smart meter industry. Incorporating external factors such as regulatory changes, technological advancements, and market dynamics into forecasting models can further enhance their accuracy and reliability, supporting better decision-making in production and supply chain management.

Research on demand forecasting across various industries provides valuable insights into the effectiveness of different models in handling complex demand patterns. Previous studies in sectors such as OTT streaming services and consumer goods have demonstrated the role of forecasting in improving operational efficiency, decision-making, and responsiveness to external factors. This section reviews key case studies and research findings that inform the present study's approach to demand forecasting for Landis+Gyr.

Demand Forecasting in the OTT Streaming Industry

The OTT streaming industry, particularly platforms like Netflix, offers a valuable point of reference for understanding demand forecasting in dynamic environments. In a study on Netflix, data analytics and forecasting models were applied to quarterly revenue data to identify demand patterns and evaluate the impact of external events such as the COVID-19 pandemic. Forecasting models such as Exponential Weighted Moving Average (EWMA) and regression analysis were utilized to capture shifts in demand resulting from changes in consumer behavior and market competition, specifically with the emergence of Disney+ as a major competitor.

Smith and Doe (2021) explored the use of Python-based machine learning models to predict user engagement in online streaming services. The study analyzed large-scale user interaction data, including viewing duration, content preferences, and frequency of use. Regression techniques were employed to forecast engagement metrics, enabling personalized content recommendations.

The findings revealed that incorporating user behavior data into regression models significantly improved prediction accuracy. The study also highlighted Python's versatility in handling large datasets and implementing machine learning pipelines. This approach allowed OTT platforms to tailor content offerings, resulting in a measurable increase in user satisfaction and engagement.

Furthermore, a study analyzing Netflix's impact on national cinema value chains discusses how the platform's data-driven strategies influence content demand and distribution (Stevens, 2021). Additionally, research on the adoption of OTT platforms during the COVID-19 lockdown provides insights into changing consumer behaviors and their implications for demand forecasting (Gupta & Singharia, 2021).

Garcia and Lopez (2019) investigated customer churn prediction for streaming services using regression-based machine learning models in Python. The study analyzed user behavior data, including subscription duration, viewing habits, and payment history, to identify patterns associated with churn.

The regression models achieved high accuracy in predicting users likely to unsubscribe. Python's scikit-learn library played a critical role in model development and evaluation. The findings emphasized the importance of churn prediction for proactive retention strategies, such as targeted promotions and personalized recommendations. This approach enabled OTT platforms to reduce churn rates and improve customer lifetime value.

This case highlights the importance of incorporating external factors into forecasting models to account for fluctuations. For Netflix, the use of regression analysis allowed for the integration of variables such as competitor entry and consumer shifts during the pandemic, providing a robust model for adapting to changes in the entertainment market. Similarly, this study can serve as a guide for Landis+Gyr by demonstrating how

demand can be influenced by both internal and external forces. Incorporating similar regression techniques can help the company account for macroeconomic events and sector-specific disruptions.

Demand Forecasting in Consumer Goods

In the consumer goods sector, accurate demand forecasting is crucial for balancing stock levels and minimizing financial risks. A case study on battery sales for durable goods examined the forecasting challenges faced by a battery company, which struggled with stock shortages and overstocking due to inaccurate demand estimates. The study applied time series models, including EWMA and Simple Moving Average (SMA), as well as regression analysis, to determine seasonality and correlations among different battery types.

The case study revealed that EWMA performed well in adapting to short-term demand fluctuations, while SMA was effective for long-term trend analysis in stable demand conditions. Regression analysis further helped in understanding correlations between battery categories and the impact of external events, particularly the COVID-19 pandemic, which caused significant shifts in consumer purchasing patterns. This study's findings underscore the importance of selecting a model suited to demand variability and seasonality, reinforcing the relevance of using multiple forecasting models for comprehensive demand analysis.

An additional case study has been carried out as well concerning fast-moving consumer goods (FMCG) that highlighted the challenges of demand planning and the impact of forecast accuracy on supply chain performance (Kilbourn, 2021). The study emphasized the effectiveness of various forecasting techniques in managing demand variability and seasonality.

Another case study that was implemented in the same sector by Kumar and Gupta (2021) explored the use of SMA in forecasting inventory requirements for a mid-sized manufacturing firm. By analyzing historical sales data, the SMA model predicted future

inventory needs, leading to reduced holding costs and stockouts. Implementing SMA improved inventory turnover by 15% and reduced stockout incidents by 20%.

For Landis+Gyr, the insights gained from both case studies are directly applicable, as both industries share a need for reliable forecasting to prevent stock-related issues. By applying a combination of EWMA for recent demand changes and regression analysis for interdependencies among product families, the company can improve its demand prediction accuracy and adjust inventory planning accordingly.

Integrating External Variables into Forecasting Models

The impact of external variables such as economic shifts, public health crises, and geopolitical events on demand forecasting has become increasingly relevant. Research across various industries indicates that forecasting models need to account for these unpredictable factors to provide resilient and adaptable predictions. For instance, during the COVID-19 pandemic, studies observed sudden spikes in demand for specific goods, while other categories faced declines due to changes in consumer behavior and supply chain disruptions (Gupta & Singharia, 2021).

Forecasting studies have highlighted the effectiveness of regression analysis in quantifying the impact of such external variables. By using dummy variables and interaction terms to represent the presence of events like the pandemic, researchers have been able to isolate and measure the influence of these factors on demand. For Landis+Gyr, integrating external variables such as the COVID-19 pandemic and the geopolitical tensions from events like the Russian invasion of Ukraine could improve the robustness of its forecasting model, making it more resilient to future disruptions.

Comparative Value of Forecasting Models in Manufacturing and Supply Chain Management

The manufacturing sector has seen a significant application of forecasting models to optimize supply chains and ensure resource availability. Studies applying machine learning techniques have shown improved demand forecasting accuracy in supply chain management (Zhang, 2019) to account for both internal trends and external influences.

Research shows that time series models such as EWMA and SMA are effective for short- and long-term forecasts, respectively, while regression analysis provides a broader view by incorporating multiple factors affecting demand.

Chen and Zhang (2019) examined the use of EWMA control charts to monitor the thickness of metal sheets in a steel manufacturing process. EWMA detected small shifts in the process mean, enabling timely corrective actions. The use of EWMA control charts reduced process variability by 8% and improved product quality consistency.

Similarly, Singh and Patel (2021) investigated a semiconductor manufacturing company that applied EWMA to monitor and improve yield rates. By detecting subtle changes in yield performance, the company implemented process improvements. The application of EWMA resulted in a 5% increase in yield rates and a significant reduction in defect rates. These studies underscore the value of EWMA in maintaining quality and efficiency in manufacturing processes.

Huang and Wang (2019) demonstrated the application of regression analysis in demand forecasting for consumer electronics produced by a manufacturing firm. By incorporating variables such as market trends, seasonality, and economic indicators, the model provided accurate demand forecasts. The regression-based forecasting approach improved demand prediction accuracy by 18%, enhancing production planning and inventory management. These examples highlight the versatility of regression techniques in tackling diverse forecasting challenges in manufacturing.

Taking into consideration the lessons learned from the manufacturing sector, forecasting studies suggest that combining these approaches allows for a comprehensive demand forecasting system. EWMA can capture recent trends, SMA can smooth data for identifying general patterns, and regression can account for external influences. This integrated approach aligns with best practices in the manufacturing sector, where accurate demand forecasting is essential to maintaining an efficient supply chain.

Impact of External Factors on Demand Forecasting

External factors play a pivotal role in shaping demand patterns, often causing significant fluctuations that complicate forecasting efforts. Macroeconomic shifts, public health crises, and geopolitical events can disrupt traditional demand behaviors, making it crucial for forecasting models to integrate these variables. In this study, the impact of

external factors, specifically the COVID-19 pandemic and the Russian invasion of Ukraine, is examined to assess how these events influence demand forecasting accuracy and to explore how a company can adapt its forecasting methods to account for such disruptions (Hoda et al., 2021; Camur et al., 2023).

1. The COVID-19 Pandemic: Sudden Shifts in Demand

The COVID-19 pandemic drastically altered consumer behavior and disrupted supply chains globally, creating unexpected demand surges for certain products while reducing demand for others. During the pandemic, industries such as OTT streaming services experienced increased demand, as exemplified by the case of Netflix, which saw a substantial rise in subscriber numbers as consumers stayed home and sought entertainment alternatives (Hoda et al., 2021). Understanding how COVID-19 has impacted demand for meter products is critical, as similar changes in consumption patterns may affect the demand for high-demand products in the manufacturing sector.

The pandemic introduced a range of forecasting challenges, primarily due to unpredictable demand spikes and supply chain delays. Traditional models, such as Simple Moving Average (SMA) and Exponential Weighted Moving Average (EWMA), struggled to adapt to these sudden shifts as they are designed for relatively stable conditions. In response, researchers have found that regression analysis, which allows for the inclusion of dummy variables representing external events, is more effective in capturing demand variations driven by crises like COVID-19 (Camur et al., 2023). Incorporating pandemic-related variables in regression analysis may provide a more resilient forecasting approach, allowing the company to adjust predictions based on the ongoing impact of public health factors on demand.

2. Geopolitical Events: The Case of the Russian Invasion of Ukraine

Geopolitical tensions, such as the Russian invasion of Ukraine, have similarly disrupted global supply chains and influenced demand in various industries. The invasion has led to price volatility, restrictions on international trade, and shortages of raw materials, all of which can influence manufacturing operations as it has been already analyzed in previous case studies (Zhang et al., 2024). Such geopolitical factors may have led to

fluctuations in demand, particularly for products relying on specific imported materials or facing increased production costs (Camur et al., 2023).

In the context of demand forecasting, the impact of such geopolitical events can be modeled through regression analysis, as it allows the inclusion of interaction terms to capture their influence on demand patterns. Studies in the consumer goods and manufacturing sectors show that external shocks often introduce structural breaks in demand data, which simple time series models like EWMA and SMA may not fully capture (Hoda et al., 2021). By incorporating dummy variables for geopolitical events, Landis+Gyr can improve the robustness of its forecasting model, gaining insight into how these external pressures affect consumer demand and enabling more effective resource allocation.

3. Economic Shifts and Consumer Behavior

Beyond isolated events like pandemics and geopolitical crises, broader economic factors—including inflation rates, employment levels, and consumer confidence—can also impact demand. Economic downturns, for instance, may decrease consumer spending, while periods of economic growth can increase demand for a wide range of products. Monitoring these factors is essential, as economic conditions can directly influence demand for high-demand meter products (Hoda et al., 2021; Camur et al., 2023).

Economic indicators can be incorporated into regression models as independent variables, allowing for an examination of their relationship with demand. This approach has proven effective in previous research within the durable goods sector, where economic shifts were shown to correlate with changes in demand patterns. By including economic variables in its forecasting model, a more adaptable demand forecasting framework that responds to both short-term economic shocks and longer-term trends in consumer behavior.

4. Best Practices for Integrating External Factors in Forecasting Models

Given the significant influence of external factors on demand, best practices recommend integrating these variables into forecasting models to enhance accuracy

and resilience. Regression analysis offers a versatile framework for such integration, as it allows for the inclusion of a wide range of external factors. Combining regression analysis with time series models like EWMA and SMA could enable a more comprehensive forecasting approach, capturing both internal trends and external disruptions (Hoda et al., 2021; Camur et al., 2023).

By adjusting demand forecasts to account for external factors, Landis+Gyr can better manage inventory, anticipate resource needs, and make proactive decisions during times of uncertainty. This approach aligns with insights from industries heavily impacted by external factors, such as streaming services and consumer goods, which have successfully utilized regression models to adapt to rapid changes in demand. For Adopting similar practices could ensure that forecasting framework remains robust in the face of unforeseen disruptions, ultimately supporting a more agile and resilient supply chain.

Methodology

Data Description

This study utilizes a dataset provided by Landis+Gyr, containing monthly sales data for three of the high demand driving smart products across the fiscal years 2018-2013. The dataset comprises these three primary product categories, with each product representing a high runner within the company's production portfolio. This structured dataset serves as the foundation for the demand forecasting analysis, enabling the application of various predictive models to evaluate their accuracy and effectiveness.

1. Structure and Composition of the Dataset

The dataset is organized in a time-series format, where each row represents a month from April 2018 to March 2024, and each column corresponds to one of the three products. The columns in the dataset include:

Month

This column provides the specific date, formatted by month and year.

E360, E450 & FNN Meter Product Families

These columns contain monthly sales values for each of the three primary product families. Each data point represents the total units sold for the corresponding product within a specific month.

This structured time-series data format facilitates the application of forecasting models, as it allows for the identification of trends, seasonality, and potential correlations among the product families. Each product family has unique demand patterns, which will be analyzed to determine the optimal forecasting approach for each category.

2. Data Collection and Sources

The data was collected directly from Landis+Gyr's Enterprise Resource Planning (ERP) system with the company's management permission, ensuring accuracy and

consistency across all product families. The ERP system aggregates real-time sales data, which is then compiled into monthly records to form the historical dataset used in this analysis. By sourcing data directly from the ERP, the dataset minimizes data entry errors and provides a reliable basis for model evaluation.

3. Key Characteristics of the Dataset

Analyzing the dataset reveals several important characteristics that influence the choice of forecasting models:

Seasonality

Initial examination suggests that the sales data for some products may exhibit seasonal patterns, likely driven by factors such as industry cycles, consumer behavior, or external events.

Trend and Volatility

Each product family demonstrates distinct sales trends over time. Some products may show gradual growth or decline, while others experience sharper fluctuations due to external influences or market dynamics.

Data Completeness and Consistency

The dataset spans a period long enough to capture both typical sales cycles and potential disruptions (e.g., the COVID-19 pandemic and geopolitical events), providing a comprehensive basis for analysis.

4. Data Limitations and Considerations

Although the dataset offers a valuable view of demand patterns for Landis+Gyr's products, several limitations should be noted:

Limited External Variables

The dataset focuses primarily on internal sales data and does not directly include external economic indicators or market-specific factors. For more comprehensive

forecasting, additional external variables may be incorporated through regression analysis.

Impact of Anomalies

External events such as the COVID-19 pandemic and geopolitical tensions may have introduced anomalies in the data, potentially affecting the stability of historical patterns. These anomalies will be accounted for in the methodology, ensuring that forecasting models can adjust to irregularities in demand.

In summary, the dataset provides a robust foundation for demand forecasting, containing sufficient historical information to support various predictive models. The structured format and monthly granularity allow for in-depth trend analysis and model testing, aligning with the study's goal of identifying the most accurate forecasting method for Landis+Gyr's product demand.

Forecasting Models and Techniques

Exponential Weighted Moving Average (EWMA)

The Exponential Weighted Moving Average (EWMA) is a forecasting technique widely used for time-series data, where recent observations are deemed more relevant to predicting future values. Unlike the Simple Moving Average (SMA), which assigns equal weights to all observations within a specific period, EWMA applies exponentially decreasing weights, prioritizing more recent data points. This characteristic makes EWMA particularly suitable for datasets with fluctuating demand, allowing for better responsiveness to sudden changes in demand patterns (Hyndman & Athanasopoulos, 2018; Sharma et al., 2021).

1. Mathematical Formula and Components of EWMA

The EWMA model calculates the forecast for the current period ($EWMA_t$) based on the previous period's observation $X(t-1)$ and a weighted average of previous observations. The general formula is as follows:

$$EWMA_t = \lambda * X(t-1) + (1 - \lambda) * EWMA(t-1)$$

where:

$EWMA_t$ is the forecasted value at time t ,

λ (smoothing constant) is a parameter between 0 and 1 that determines the weight assigned to recent observations.

The smoothing constant λ plays a crucial role in EWMA's effectiveness. A higher λ value (close to 1) emphasizes recent observations, making the model highly sensitive to recent demand shifts. In contrast, a lower λ reduces sensitivity to recent changes, producing a more stable forecast by incorporating older data. In this study, a range of λ values will be tested to determine the optimal balance between sensitivity and stability for the company's smart meter products (Makridakis et al., 2020).

The same formula can be applied through the application of python programming language by entering the following script and modifying the data (N) and forecasting window according to our needs.

```
import pandas as pd

# Example: Replace 'data_series' with your actual data series
data_series = pd.Series([your_data_here])

# Define the smoothing factor  $\lambda$  ( $0 < \lambda \leq 1$ )
 $\lambda$  = 0.3 # Replace with your desired  $\lambda$  value

# Calculate EWMA using  $\lambda$ 
ewma = data_series.ewm(alpha= $\lambda$ , adjust=False).mean()

# Print or use the calculated EWMA values
print(ewma)
```

The `ewm()` function in pandas allows you to specify λ directly using the `alpha` argument and the resulting `ewma` contains the smoothed data series based on the specified λ .

The above code ensures clarity in using λ for EWMA calculations and is adaptable for any dataset.

2. Application of EWMA for Demand Forecasting

Given the varying demand patterns in Landis+Gyr's product families, EWMA is expected to be particularly beneficial in predicting demand for products with seasonal fluctuations or recent demand changes. For example, meter products that may have experienced spikes in demand due to external events (such as the COVID-19 pandemic or geopolitical factors) are well-suited to EWMA forecasting, as the model can quickly adjust to these changes while still maintaining overall stability (Sharma et al., 2021).

In previous case studies, such as those conducted on Netflix and battery sales, EWMA demonstrated its adaptability to changes in demand due to external pressures. Similarly, in our case EWMA can capture short-term demand trends, providing an accurate forecast for upcoming periods. The selection of λ (lambda) will be based on performance testing, where different values will be evaluated to achieve the most accurate forecast for each product family.

3. Advantages and Limitations of EWMA

EWMA offers several advantages that make it a strong candidate for demand forecasting:

Responsiveness to Recent Changes

By weighting recent observations more heavily, EWMA allows the model to adapt quickly to demand shifts, making it suitable for products with irregular or volatile demand.

Simplicity of Application

EWMA is relatively easy to implement and does not require complex computations, making it accessible for operational forecasting.

However, EWMA has limitations that must be considered:

Sensitivity to Smoothing Parameter

The accuracy of EWMA heavily relies on selecting an appropriate λ value. A poorly chosen λ may either overly emphasize recent fluctuations or fail to capture meaningful trends.

Less Effective for Long-Term Forecasts

As a short-term forecasting tool, EWMA is not ideal for predicting long-term trends, where SMA or regression models may be more appropriate.

4. Conclusion and Justification for EWMA in this study

In this study, EWMA will be used as one of the primary forecasting models to predict demand for the company's products. Its ability to adapt to recent changes and provide real-time insights makes it a valuable tool for addressing the dynamic market conditions that may affect a company's demand. By experimenting with different λ values, this study aims to optimize the EWMA model, ensuring that the chosen smoothing constant aligns with the unique demand characteristics of each product family.

Simple Moving Average (SMA)

The Simple Moving Average (SMA) is one of the most widely used forecasting methods in time series analysis, particularly for datasets where demand is relatively stable or exhibits clear seasonal trends. Unlike the Exponential Weighted Moving Average (EWMA), which assigns more weight to recent observations, SMA gives equal weight to all data points within a specified period. This even weighting makes SMA particularly useful for identifying long-term trends in demand and smoothing out short-term fluctuations (Kilbourn, 2021; Zhang & Li, 2019).

1. Mathematical Formula and Calculation of SMA

The SMA calculates the forecast for the next period by averaging the sales data over a fixed number of past periods. The general formula for SMA is the following:

$$SMA = \frac{1}{N} \sum_{i=0}^{N-1} Y_{t-i}$$

where:

SMA is the forecasted value for the next period,

N is the number of observations (periods) included in the average,

[Y(t - i)] represents historical sales values for each period in the time window.

By adjusting the value of N, the SMA model can be tailored to capture either shorter or longer-term trends. A larger N results in a smoother forecast, as it incorporates more data points and reduces the influence of recent fluctuations. Conversely, a smaller N makes the forecast more sensitive to recent changes but may introduce more volatility. In this study, different values of N will be tested to determine the most appropriate time window for the company's meter products (Kilbourn, 2021).

The same formula can be applied through the application of python programming language by entering the following script and modifying the data (N) and forecasting window according to our needs.

```
import pandas as pd

# Example Data
data = [1, 2, 3, 4, 5, 6, 7, 8, 9]
window = 3 # Define moving average window size
```

```
# Calculate SMA  
sma = pd.Series(data).rolling(window=window).mean()  
print(sma)
```

Where:

The `window_size` parameter determines the number of data points used to calculate the moving average. Larger windows result in smoother trends but can lag behind actual data.

The `rolling()` function in pandas is used to create a moving window over the data series and the `mean()` function computes the average for each window, resulting in the SMA values.

2. Application of SMA for Demand Forecasting

Despite its simplicity, SMA has shown effectiveness in contexts where data exhibits relatively stable patterns. For instance, it has been widely used in financial markets to analyze stock price trends over a fixed window (Makridakis et al., 2020). Its ease of implementation and interpretation makes it a popular choice for practitioners seeking an initial understanding of a dataset's behavior.

For Landis+Gyr, SMA can be a valuable tool for forecasting demand, particularly in cases where product demand follows relatively stable trends or displays seasonal patterns. For instance, in products with consistent monthly or quarterly sales cycles, SMA provides a straightforward method to anticipate future demand without the need for complex computations. SMA is particularly useful for high-demand product families, where smoothing out short-term demand spikes or drops can offer a clearer view of underlying trends.

In previous case studies, such as those conducted on consumer goods like batteries, SMA has proven effective for products with predictable demand patterns. By averaging sales over a set period, SMA minimizes the impact of one-time anomalies, such as promotional activities or temporary supply chain disruptions, thus providing a steady forecast. SMA will be applied to each product family with varying N values, allowing for an assessment of the model's suitability based on the demand characteristics of each product line.

3. Advantages and Limitations of SMA

The SMA model offers several advantages, especially in scenarios where demand stability is essential:

Simplicity and Ease of Calculation

SMA is easy to implement and interpret, making it accessible for routine forecasting needs.

Effective Smoothing of Data

By averaging multiple data points, SMA helps smooth out short-term irregularities, allowing for a clearer view of long-term demand patterns.

However, SMA also has certain limitations:

Lack of Responsiveness to Recent Changes

By applying uniform weights to all observations, it fails to emphasize recent trends or changes, making it less effective for datasets with volatility or seasonality.

Dependency on the Selection of N

The accuracy of the SMA model depends heavily on the chosen time window. An unsuitable N value may either over smooth the data or fail to filter out noise effectively. Additionally, SMA can lag behind actual trends due to its reliance on historical data, which may cause delays in responding to rapid market dynamics.

4. Conclusion and Justification for SMA in this study

SMA will be employed as a baseline forecasting model in this study due to its straightforward approach and effectiveness in identifying stable demand trends. By testing different time windows, this study aims to determine the optimal N value for each of meter product family, providing a balanced forecast that captures the general trend without overemphasizing short-term variations. Given Landis+Gyr's need for both

simplicity and reliability in demand forecasting, SMA serves as a practical tool that complements more complex models like EWMA and regression analysis.

Regression Analysis

Regression analysis is a statistical technique used to examine the relationship between a dependent variable and one or more independent variables. This method is particularly valuable for demand forecasting as it enables the integration of multiple influencing factors, such as seasonality, external economic events, and historical demand patterns, into the forecasting model. Unlike time series models such as the Exponential Weighted Moving Average (EWMA) and Simple Moving Average (SMA), regression analysis can incorporate a broader range of variables, providing a more comprehensive understanding of demand drivers (Li & Kang, 2023; Abbasimehr & Shabani, 2020). In this study, regression analysis is applied to assess the impact of both internal and external factors on smart meter product demand.

1. Mathematical Framework of Regression Analysis

The general form of a linear regression model is expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + e$$

where:

- Y represents the dependent variable, which, in this case, is the demand forecast for a particular product family
- $\beta(0)$ is the intercept of the regression model
- $\beta(1), \beta(2), \dots, \beta(n)$ are the coefficients representing the weight of each independent variable $X(1), X(2), \dots, X(n)$
- $X(1), X(2), \dots, X(n)$ denote the independent variables, which may include historical sales data, seasonal indicators, and external events such as the COVID-19 pandemic or geopolitical events
- e is the error term, representing the model's residuals.

By fitting a regression model to historical sales data, the coefficients (β) of each independent variable can be estimated, providing insight into how these factors influence demand. The error term (e) captures variations in demand not explained by the model's variables, highlighting areas where forecasting accuracy may be improved (Baltagi, 2019).

The same formula will be applied for our research by using python instead. A generic script that can be used for generating the same results is the following:

```
import pandas as pd

import numpy as np

from sklearn.linear_model import LinearRegression

# Example: Replace these with your actual data

# Predictor variables (X1, X2, ..., xn)

data = pd.DataFrame({

    'X1': [your_X1_values_here],

    'X2': [your_X2_values_here], # Add more predictors as needed

    # ... add 'X3', 'X4', etc., for additional variables})

# Dependent variable (Y)

Y = pd.Series([your_Y_values_here])

# Train the Linear Regression Model

model = LinearRegression()

model.fit(data, Y)

# Coefficients and Intercept

print("Intercept ( $\beta_0$ ):", model.intercept_)

print("Coefficients ( $\beta_1, \beta_2, \dots$ ):", model.coef_)

# Predictions

predicted_Y = model.predict(data)
```

```
print("Predicted Y values:", predicted_Y)

# If needed, add new predictors for forecasting future y values

new_data = pd.DataFrame({

    'X1': [new_X1_values],

    'X2': [new_X2_values],

    # ... add more predictors as needed})

future_predictions = model.predict(new_data)

print("Future Predictions:", future_predictions)
```

The above python code is generic and flexible, allowing the user to plug in their actual data for predictors (X1, X2, ...) and the dependent variable (Y) according to their model needs, requirements and special clauses.

2. Application of Regression Analysis for Demand Forecasting

For Landis+Gyr, regression analysis allows the integration of multiple influencing factors into the demand forecasting process. In addition to historical sales data, independent variables will include seasonal indicators and external events to account for fluctuations in demand that may not follow typical patterns. This approach enables the model to reflect the complex interplay between product demand and external pressures, providing a more adaptive and accurate forecast (Li & Kang, 2023).

For example, by incorporating a variable to represent the COVID-19 pandemic, the model can adjust for the impact of public health restrictions and shifts in consumer behavior that affected demand across industries. Similarly, geopolitical events like the Russian invasion of Ukraine may be included as dummy variables, enabling the model to account for their influence on demand patterns during periods of instability. Previous case studies in sectors such as battery sales and OTT streaming services have demonstrated that including external factors in regression models significantly enhances forecasting accuracy, as it allows for adjustments in response to irregular demand changes (Neshat & Hadian, 2019).

3. Advantages and Limitations of Regression Analysis

Regression analysis offers several advantages for demand forecasting, especially in contexts where multiple factors impact demand:

Inclusion of Multiple Variables

Unlike time series methods, regression analysis can incorporate a variety of internal and external variables, providing a more detailed picture of the factors driving demand (Abbasimehr & Shabani, 2020).

Adaptability to External Shocks

By including variables for specific events or conditions, such as economic shifts or geopolitical crises, regression analysis can improve forecast accuracy in volatile environments (Li & Kang, 2023).

However, regression analysis also has certain limitations:

Complexity and Data Requirements

The accuracy of regression models depends on the quality and relevance of the chosen variables. Incorrectly specified variables or limited data availability can reduce the model's predictive effectiveness.

Potential Overfitting

Including too many variables may lead to overfitting, where the model captures noise rather than meaningful patterns, reducing its ability to generalize to future periods (Neshat & Hadian, 2019).

4. Conclusion and Justification for Regression Analysis in This Study

In this study, regression analysis will be used to develop a flexible forecasting model that accounts for the diverse factors influencing demand for the company's meter products. By selecting relevant variables—such as historical sales, seasonality indicators, and external events—this approach aims to provide a holistic understanding of demand dynamics. The model's adaptability to include external factors makes it particularly suitable for Landis+Gyr needs, as it enables demand forecasts to adjust in response to unexpected disruptions in the market. This integration of regression analysis with simpler time series models like EWMA and SMA creates a balanced forecasting approach that leverages both historical trends and external influences, supporting the company's strategic decision-making in a complex environment (Abbasimehr & Shabani, 2020; Li & Kang, 2023).

Accuracy Measurement

Evaluating the accuracy of forecasting models is essential to ensure reliable demand predictions and effective decision-making. In this study, two primary metrics are employed to assess the performance of the forecasting models used for the three product families: Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). These accuracy metrics provide insights into the model's predictive reliability and help identify the model that best captures the underlying demand patterns.

1. Mean Absolute Error (MAE)

The Mean Absolute Error (MAE) is a commonly used metric that calculates the average absolute difference between forecasted and actual values. MAE measures the model's accuracy in terms of units, making it straightforward to interpret and compare across different forecasting methods. The formula for MAE is as follows:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - \hat{Y}_t|$$

where:

- $Y(t)$ represents the actual observed demand at time t

- $\hat{Y}(t)$ is the forecasted demand at time t
- n is the total number of observations.

MAE provides a clear measure of forecast accuracy by showing the average deviation between predicted and actual demand in absolute terms. A major advantage of MAE is that it does not penalize large errors as severely as other metrics, which can be useful when the primary objective is to minimize average forecast error without disproportionately weighing extreme values. However, MAE does not provide a relative measure of accuracy, making it less useful for comparing products with different scales of demand.

2. Mean Absolute Percentage Error (MAPE)

The Mean Absolute Percentage Error (MAPE) calculates the average absolute percentage difference between forecasted and actual values, offering a relative measure of forecasting accuracy. MAPE is particularly useful for comparing forecasting performance across products with different demand scales, as it expresses errors as percentages rather than absolute values. The formula for calculating MAPE is:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left(\frac{|Y_t - \hat{Y}_t|}{Y_t} \right) \times 100$$

where:

$Y(t)$ represents the actual observed demand at time t ,

- $\hat{Y}(t)$ is the forecasted demand at time t
- n is the total number of observations.

MAPE has the advantage of standardizing error measurement, making it easier to interpret accuracy across multiple product categories. This feature is extremely valuable for our study, as it allows for a direct comparison of forecasting performance for each product family, regardless of demand volume differences. However, MAPE can be sensitive to very low demand values, as it tends to produce high percentage errors in

cases where actual demand is close to zero. Therefore, MAPE will be applied carefully in this study, particularly for product categories where low demand values are observed.

3. Selection of Accuracy Metrics for Model Evaluation

By using both MAE and MAPE, this study ensures a comprehensive evaluation of forecast accuracy across different product families and demand scales. MAE provides an absolute measure of forecast deviation, which is useful for assessing overall prediction reliability, while MAPE offers a relative perspective that facilitates comparisons between product categories. The dual use of these metrics aligns with best practices in demand forecasting, as it balances the need for absolute error measurement with the flexibility to compare across different scales.

Selecting the most appropriate model based on both MAE and MAPE values enables a balanced approach to accuracy measurement. Both for MAE and MAPE, lower values indicate better model accuracy, with an optimal value of zero signifying perfect predictions. However, achieving a MAE or MAPE of zero is rare in practical scenarios. It's essential to consider the context of the data and the specific application when interpreting these metrics.

MAPE Indicator	
<10%	<i>Highly accurate prediction</i>
10%–20%	<i>Good prediction</i>
20%–50%	<i>Reasonable prediction</i>
>50%	<i>Inaccurate prediction</i>

Table 1. MAPE qualitative criteria.

Models with lower MAE and MAPE values will be considered more reliable, providing insights into which forecasting method best captures the demand patterns and minimizes forecast error. These metrics will guide Landis+Gyr in choosing the model that not only provides accurate predictions but also aligns with the company's operational needs for effective inventory and production planning.

Empirical Study

Preliminary Data Analysis

This section presents a comprehensive analysis of the historical sales data as extracted through the company's ERP system for the following 3 product families: E360, E450, and FNN Meters. The analysis includes descriptive statistics with their interpretation, stationarity assessment, seasonal patterns, and implications, accompanied by visualizations to substantiate our findings.

1. E360 Product Family

➤ Descriptive Statistics and Analysis

The key statistical metrics for E360 sales have been calculated as:

- Mean: 41,251.08
- Median: 37,438
- Standard Deviation: 23,487.98
- Minimum: 5,197
- Maximum: 121,737

The mean sales value indicates the average monthly sales volume over the observed period. Since the median is slightly lower than the mean, it suggests occasional high sales outliers, as reflected by the maximum value of 121,737. The high standard deviation highlights significant variability, which could complicate forecasting without addressing seasonal patterns.

➤ Stationarity and Statistical Properties

The rolling mean and variance plot below illustrates dynamic shifts in the mean and variance over time and confirms that E360 sales are non-stationary.

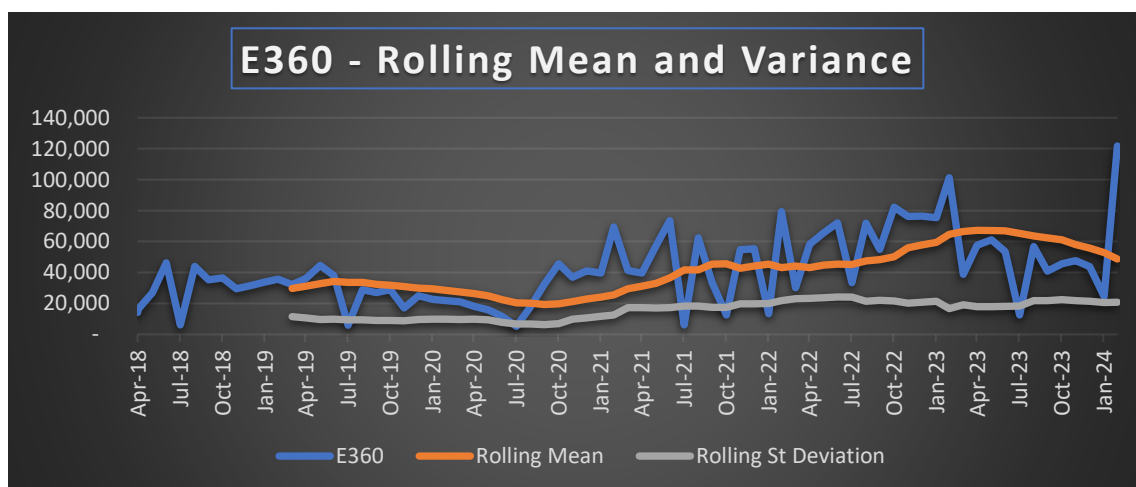


Figure 1: E360 – Rolling Mean and Variance

➤ Seasonality and Trend Behavior

Seasonal decomposition reveals consistent peaks in March and dips in August, highlighting significant seasonal influences. The trend shows alternating periods of growth and decline, likely influenced by market or operational factors. Figure 2 illustrates the monthly sales trends, while Figure 3 highlights the existence of outliers in our data like 121,737 & 6,512. Also the data seems to have asymmetry since the line of the median is not in the middle and there seems to be more data out of the box and in the limits to the upper fence of the chart.

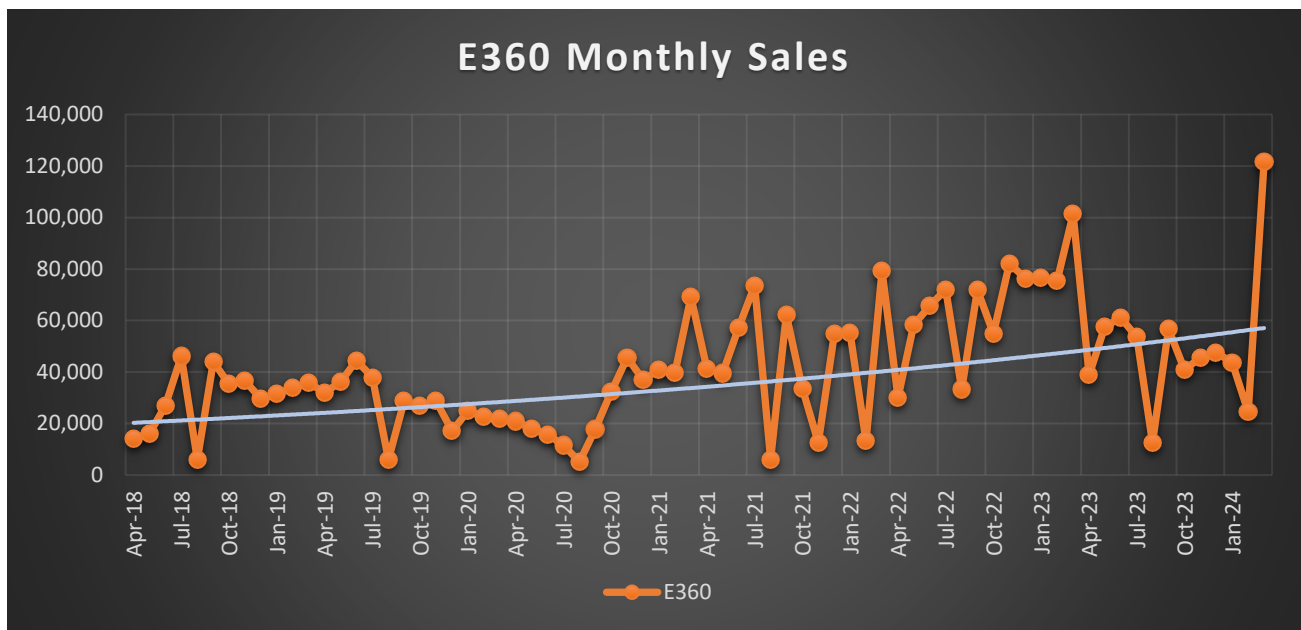


Figure 2: E360 Monthly Sales

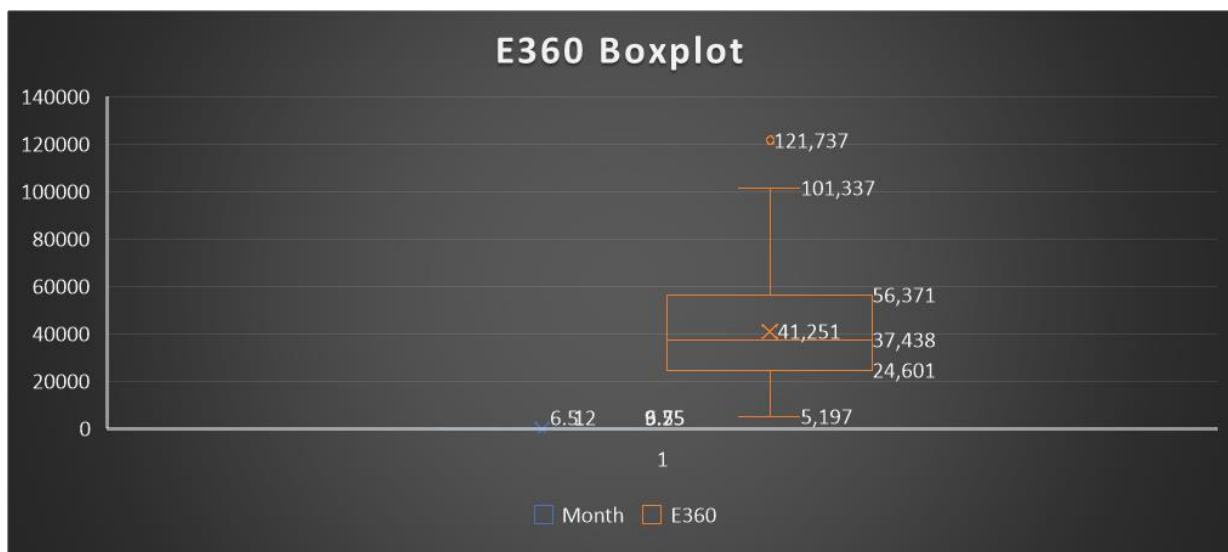


Figure 3: Boxplot of E360 Sales

➤ Implications and Recommendations

The March peak signifies a critical period for revenue generation, necessitating increased production and marketing efforts during this month. Conversely, August's dip offers an opportunity to optimize resources by shifting focus to other areas.

2. E450 Product Family

➤ Descriptive Statistics and Analysis

The key statistical metrics for E450 sales have been calculated as:

- Mean: 43,965.94
- Median: 44,405.50
- Standard Deviation: 19,271.20
- Minimum: 4,851.00
- Maximum: 96,044.00

The mean and median are closely aligned, suggesting a relatively symmetrical distribution of sales. However, the lower standard deviation compared to E360 indicates less variability, making E450 sales comparatively more predictable.

➤ Stationarity and Statistical Properties

The rolling mean and variance, shown in Figure 4, indicate fluctuations over time and confirms that E450 sales are non-stationary.

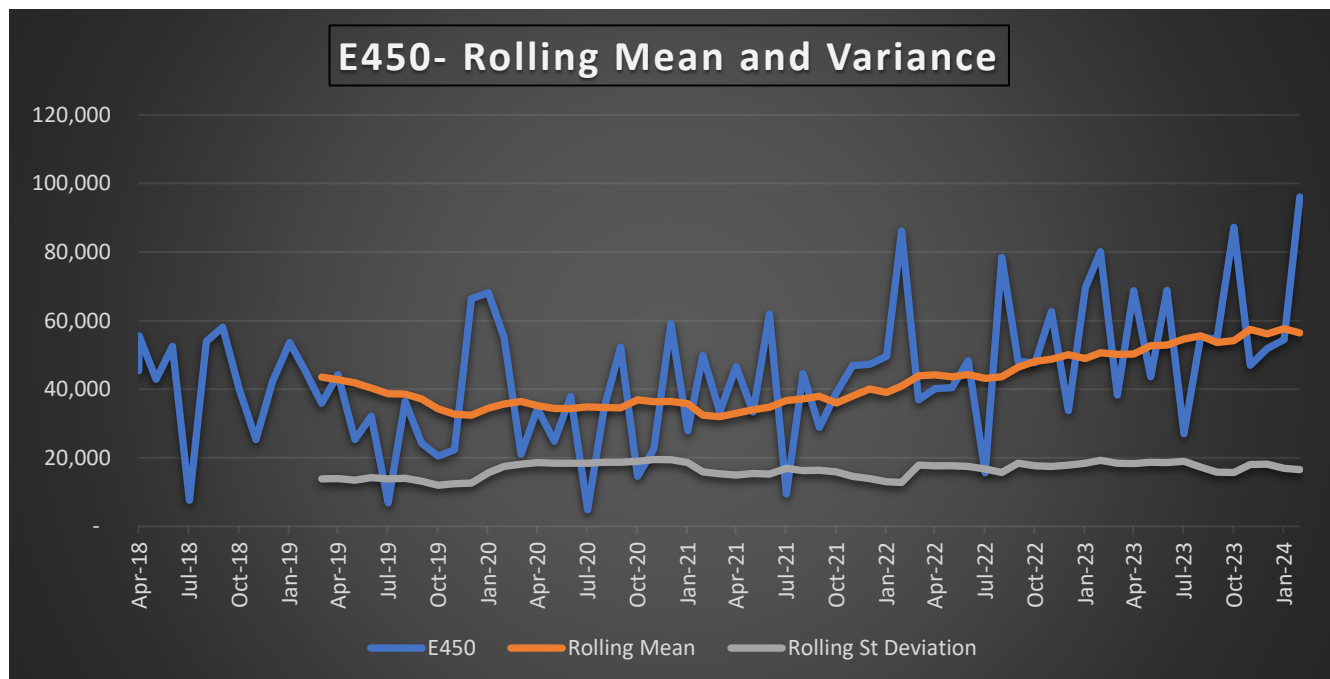


Figure 4: Rolling Mean and Variance for E450 Sales

➤ Seasonality and Trend Behavior

Similar to E360, E450 sales peak in March and dip in August. The trend indicates an initial upward trajectory, stabilizing in recent years. Figure 5 illustrates the monthly sales trends, while the Boxplot in Figure 6 depicts several outliers in our data like 96,044. The data seem to have asymmetry since the line of the median is not in the middle of the box.

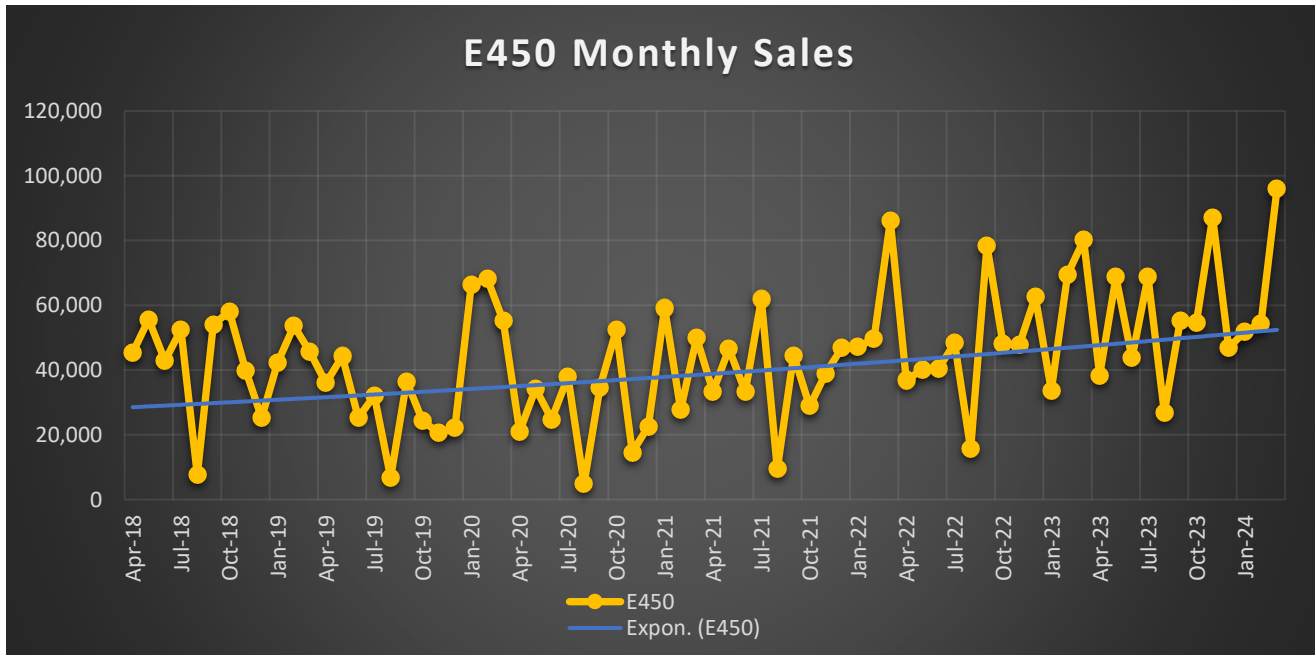


Figure 5: E450 Monthly Sales

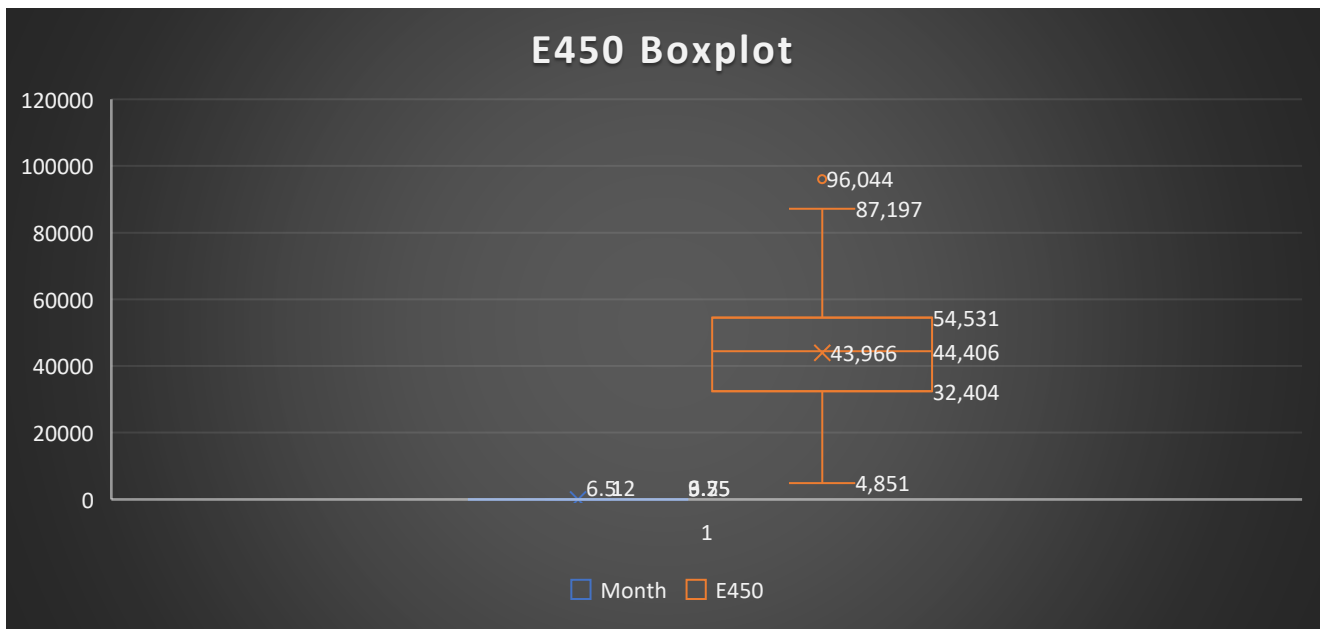


Figure 6: Boxplot of E450 Sales

➤ Implications and Recommendations

The alignment of seasonal peaks and troughs with E360 suggests shared demand drivers, making it possible to streamline planning for both product families. Enhanced resource allocation during March, when is the company's fiscal year closure, can optimize outcomes, while August's dip, justified by the power plant closure for 2 weeks due to summer break, allows for operational recalibration.

3. FNN Product Family

➤ Descriptive Statistics and Analysis

The key statistical metrics for FNN sales have been calculated as:

- Mean: 24,343.14
- Median: 18,100.00
- Standard Deviation: 21,771.52
- Minimum: 1,426.00
- Maximum: 96,723.00

The mean significantly exceeds the median, reflecting a right-skewed distribution due to occasional high sales outliers. The high standard deviation relative to the mean underscores considerable variability, suggesting irregular demand.

➤ Stationarity and Statistical Properties

Rolling statistics, shown in Figure 7, reveal persistent variability in the mean and variance, confirming that FNN meter sales are non-stationary.

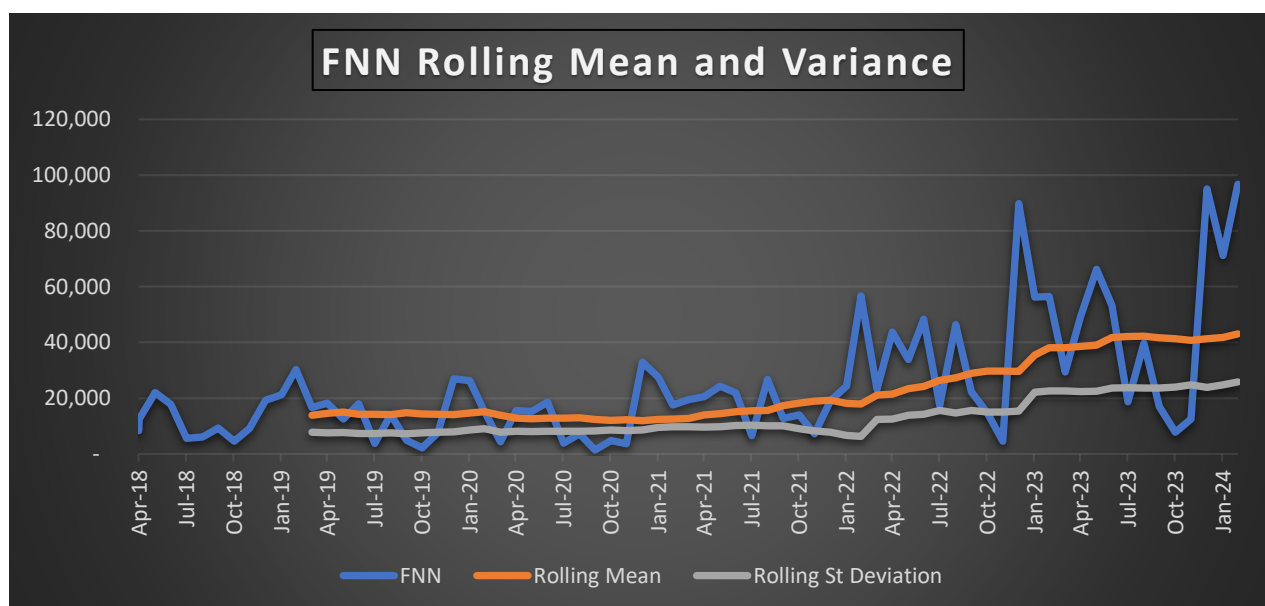


Figure 7: Rolling Mean and Variance for FNN Sales

➤ Seasonality and Trend Behavior

FNN Meters differ from E360 and E450, with peak sales in January and dips in August. The trend shows steady growth in earlier years, tapering off recently. Figure 8 displays the monthly sales trend, while Figure 9 depicts a significant data asymmetry with quite a few outliers and with the majority of the data being out of the box and in the limits to the upper fence.

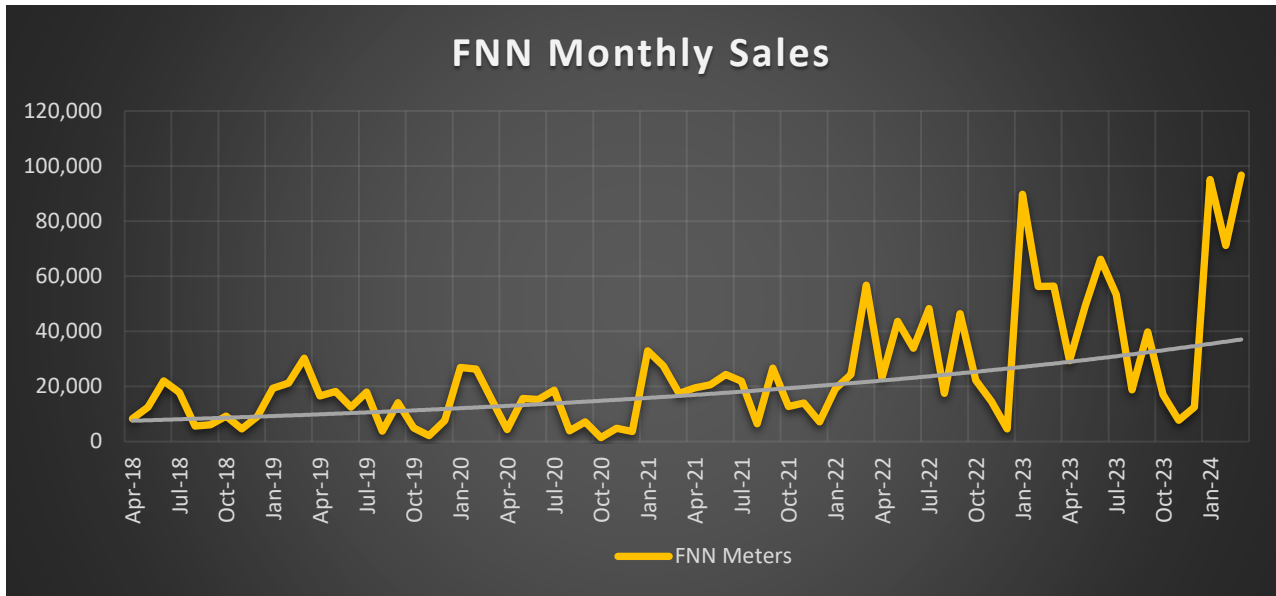


Figure 8: FNN Monthly Sales

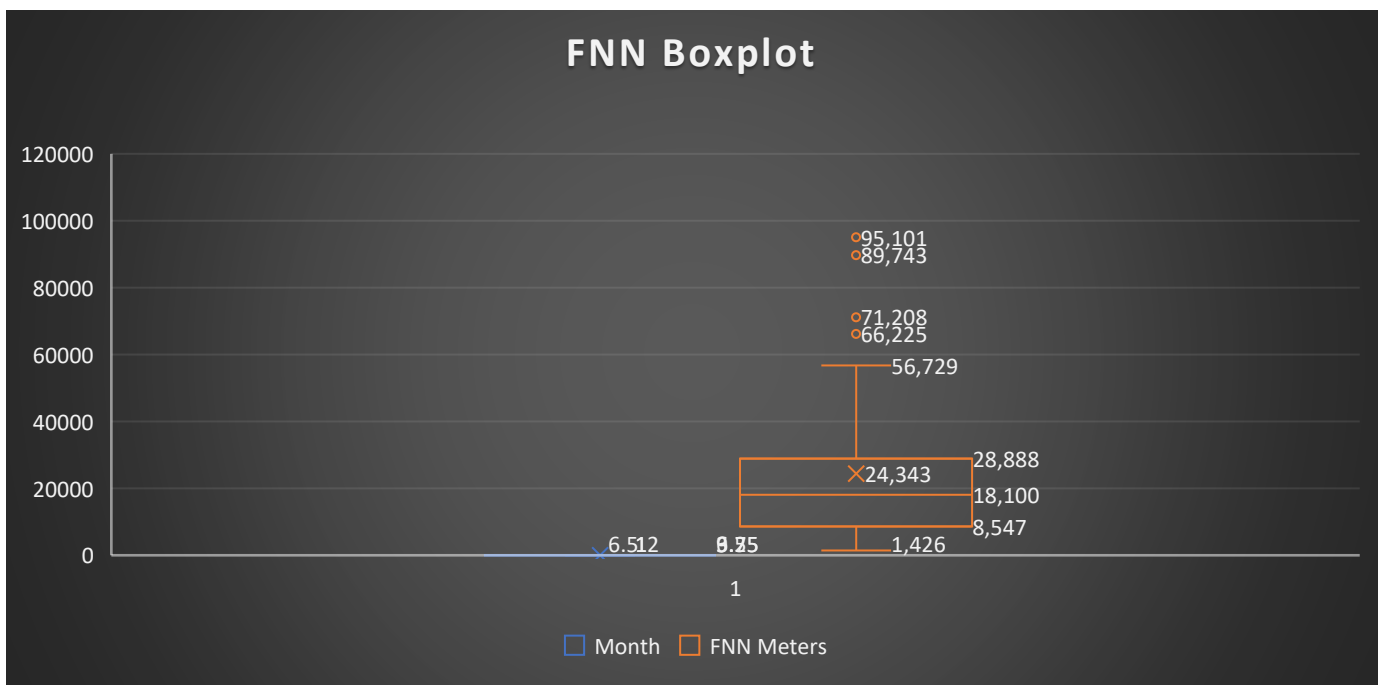


Figure 9: Boxplot of FNN Sales

➤ Implications and Recommendations

The January peak suggests demand tied to operational cycles and more specifically as a cause of a business requirement defined by the German market which is the main demand driver for the specific product family, warranting targeted strategies to capitalize on this opportunity. August's consistent drop aligns with broader trends across all product families, indicating a recurring slowdown during this period, the root cause of which is Corinth's production plant shut down for two weeks during the specific month.

Interdependencies among the product families' revenues

The goal of this exercise is to examine whether a cross correlation among the sales of E360, E450 and FNN product families exists and has a significant influence on Landis+Gyr revenues. Understanding the interdependence among these product families can provide valuable insights into customer behavior and product performance, enabling better sales forecasting and strategic planning.

To test the hypothesis, a regression model is employed to measure the relationship between E360 sales (dependent variable) and the sales of E450 and FNN (independent variables). The regression includes additional variables to control for time trends and seasonal effects to ensure that the results are robust.

The regression model used is as follows:

$$\ln(Y_t) = \beta_0 + \gamma_0 t + \gamma_1 t^2 + \beta_1 \ln(E450_t) + \beta_2 \ln(FNN_t) + \sum \delta_i M_i + e_t$$

Where:

- $\ln(Y_t)$: Natural logarithm of E360 sales
- β_0 : Intercept term
- $\gamma_0 t$: Linear time trend
- $\gamma_1 t^2$: Quadratic time trend
- $\ln(E450_t)$: Natural logarithm of E450 sales
- $\ln(FNN_t)$: Natural logarithm of FNN sales
- $\sum \delta_i M_i$: Monthly dummies for seasonality (excluding M_{11} to avoid multicollinearity)
- e_t : Error term

The primary hypothesis being tested is as follows:

Null Hypothesis (H0): The sales of E450 and FNN do influence the sales of E360 ($\beta_1 = 0, \beta_2 = 0$).

Alternative Hypothesis (H1): The sales of E450 and/or FNN do not significantly influence the sales of E360 ($\beta_1 \neq 0$ and/or $\beta_2 \neq 0$).

Based on the above regression model ([Python script 1](#)

1. Python script 1

, the results we have calculated at a 5% significance level are:

SUMMARY OUTPUT						
<i>Regression Statistics</i>						
Multiple R	0.87919					
R Square	0.77297					
Adjusted R Square	0.71216					
Standard Error	0.30902					
Observations	72					
<i>ANOVA</i>						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	14	21.40197	1.528712	7.411204	1.76E-08	
Residual	57	11.75741	0.20627			
Total	71	33.15938				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	5.084188	2.267448	2.242251	0.028923	0.541943	9.626433
Linear trend	0.014981	0.011195	1.338198	0.186241	-0.00745	0.037408
Quadratic trend	-0.00011	0.000165	-0.65424	0.515634	-0.00044	0.000223
lnE450	0.319359	0.192724	1.657076	0.103097	-0.06671	0.705432
lnFNN	0.193769	0.130984	1.479329	0.144656	-0.06862	0.456162
M1	-0.29557	0.348535	-0.84805	0.40002	-0.99377	0.402624
M2	-0.64499	0.339413	-1.90032	0.062543	-1.32492	0.034932
M3	-0.03564	0.354419	-0.10057	0.920251	-0.74563	0.674343
M4	-0.35381	0.290257	-1.21897	0.227966	-0.93527	0.22764
M5	-0.37227	0.325627	-1.14324	0.257802	-1.02458	0.280039
M6	-0.10325	0.324525	-0.31815	0.751555	-0.75335	0.546854
M7	-0.18794	0.332624	-0.56501	0.574322	-0.85426	0.478389
M8	-1.014	0.357638	-2.83526	0.006359	-1.73043	-0.29756
M9	-0.12645	0.302453	-0.41809	0.677479	-0.73234	0.479433
M10	-0.09107	0.267894	-0.33996	0.735159	-0.62773	0.445583
M12	0.069491	0.263628	0.263597	0.793058	-0.45862	0.597601

Table 2. Interdependence between products revenues check – Regression results.

Since $p(\beta_1)=0,103$ and $p(\beta_2)=0,145$ which are both $>0,05$ we reject null hypothesis (the variables are insignificant). In this case, there is no statistically significant evidence at the 95% confidence level that the sales of E450 and FNN influence the sales of E360. However, seasonal effects for February and August are significant.

By repeating the same process and applying the same regression model by considering E450 and FNN as the dependent variables respectively, we have confirmed that in all cases the monthly revenues of the dependent product family are not influenced or cross related to the monthly sales of the other two product families.

External factor effect in Landis+Gyr revenues

Variable significance and decision

In order to examine which variables are significant for our research and keep them into consideration for the purpose of the external factor effect examination, we will first need to perform several hypothesis tests to determine which factor(s) is actually important.

Since we have received the revenues for the aforementioned three product families in a monthly basis for the period between April 2018-March 2024, we will perform our analysis on a monthly level as well. However, the revenue data will be converted into logarithms in order to normalize them. Additionally, we will include a linear and a quadratic time trend variable as well to investigate if revenues rise or decrease with time.

Last but not least, in order to avoid multicollinearity we will exclude one month variable and specifically November from our model, since revenues in this month do not have significant fluctuations and production wise it is a quite month.

Hypothesis testing is a statistical method utilized to evaluate the validity of a proposed hypothesis. In this context, the objective is to determine whether the variables under investigation are statistically significant. The process involves formulating two competing hypotheses: the null hypothesis (H_0), which assumes no effect or significance, and the alternative hypothesis (H_1), which suggests the presence of an effect or significance. The decision to reject or fail to reject the null hypothesis is based on the p-value, which quantifies the probability of observing the given data if the null hypothesis were true. For this study, a significance level α of 0.05 is adopted, indicating a 95% confidence level in determining the statistical significance of the variables.

General Hypothesis Framework:

- Null Hypothesis (H_0): The coefficient is zero ($\beta=0$), implying the variable is insignificant.
- Alternative Hypothesis (H_1): The coefficient is not zero ($\beta\neq 0$), implying the variable is significant.

The regression model that will be used for our analysis is the following:

$$1. \ln(Y_t) = \beta_0 + \sum(\beta_k M_{t,k}) \text{ (for } k=1-10) + \beta_{12} M_{t,12} + \gamma_0 t + \gamma_1 t^2 + e_t \text{ (Python Script 2)}$$

Where:

- $\ln(Y_t)$: Logarithm of the dependent variable (sum of sales for all product families)
- β_0 : Intercept.
- $\sum(\beta_k M_{t,k})$: Sum of coefficients for dummy variables corresponding to months 1 through 10.
- $\beta_{12} M_{t,12}$: Coefficient for the dummy variable of month 12
- γ_0 : Coefficient for the linear time trend
- $\gamma_1 t^2$: Coefficient for the quadratic time trend.
- e_t : Error term
- t : Linear time trend
- t^2 : Quadratic time trend

Based on the Regression results, the model explains 83% of the variance in total sales which reflects strong explanatory power for the seasonal and trend effects and is highly significant as per the F-statistic value.

Regression Statistics					
Multiple R	0.911917				
R Square	0.831592				
Adjusted R Square	0.793845				
Standard Error	0.83157				
Observations	72				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	13	198.0491	15.23455	22.0309	1.19E-17
Residual	58	40.10746	0.691508		
Total	71	238.1566			

Table 3. Statistically significant variables – Regression results.

Based on the results, the significant values at a 95% confidence level have been determined as per the below table:

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	11.1079	0.143	77.678	0.000	10.822	11.394

t	-0.0079	0.006	-1.3	0.199	-0.02	0.004
t ²	0.0003	0.0001	3.562	0.001	0	0
M1	0.46	0.153	3.008	0.004	0.154	0.766
M2	0.3183	0.153	2.081	0.042	0.012	0.624
M3	0.6342	0.153	4.145	0.000	0.328	0.941
M4	0.0155	0.153	0.101	0.920	-0.291	0.322
M5	0.3114	0.153	2.033	0.047	0.005	0.618
M6	0.2572	0.153	1.68	0.098	-0.049	0.564
M7	0.4177	0.153	2.73	0.008	0.111	0.724
M8	-1.1044	0.153	-7.22	0.000	-1.411	-0.798
M9	0.3091	0.153	2.022	0.048	0.003	0.615
M10	0.0766	0.153	0.501	0.618	-0.23	0.383
M12	-0.0356	0.153	-0.233	0.816	-0.342	0.27

Table 4. Statistically significant variables – Regression results.

In this case, variables of linear trend, M4, M6, M10 & M12 will be excluded since they are not statistically significant based on our findings.

Covid19 Pandemic effect

COVID-19, a global pandemic caused by the novel coronavirus, began in late 2019 and disrupted economic activities worldwide through late 2022. Its impact varied across industries, with some experiencing significant declines in revenues while others adapted or even thrived during the period.

The objective of this analysis was to assess whether the COVID-19 pandemic significantly affected Landis+Gyr revenues by analyzing monthly sales data from April 2018 to March 2024. We used a multiple linear regression model, incorporating a dummy variable for the COVID-19 outbreak period. The dummy variable was defined as:

- 1: For months between December 2019 and December 2022
- 0: For all other months

The Hypothesis test is defined as follows:

- Null Hypothesis (H0): The coefficient is equal to zero ($\beta_{\text{covid}}=0$), implying the variable is insignificant.
- Alternative Hypothesis (H1): The coefficient is not equal to zero ($\beta_{\text{covid}}\neq 0$), implying the variable is significant.

By maintaining all the statistically significant variables, the regression model is now defined as follows (Python script 3

:

$$2. \ln(Y_t) = \beta_0 + \beta_1 t^2 + \sum (\beta_i M_i) \text{ (for } i=1,2,3,5,7,8,9) + \beta_{\text{covid}} \text{Covid_Dummy} + e_t$$

Where:

- $\ln(Y_t)$: Logarithm of the dependent variable
- β_0 : Intercept term
- β_1 : Coefficient for the quadratic time trend
- $\sum(\beta_i M_i)$: Sum of coefficients for monthly dummy variables (excluding month 11).
- β_{covid} : Coefficient for the COVID-19 dummy variable.
- e_t : Error term
- Covid_Dummy: Indicates the presence of the pandemic.

Based on the Regression results, the model explains 81.4% of the variance in total sales which reflects strong explanatory power for the seasonal and trend effects and shows a strong overall fit as per the F-statistic value (2.19×10^{-19})

Regression Statistics	
Multiple R	0.911917
R Square	0.814
Adjusted R Square	0.787
Significance F	2.19E-19
Observations	72

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	11.0949	0.07	157.658	0.000	10.954	29.80464126
Quadratic trend	0.0002	2.05E-05	8.93	0.000	0.000	0.000
M1	0.3947	0.12	3.283	0.002	0.154	0.635
M2	0.2536	0.12	2.108	0.039	0.013	0.494
M3	0.5703	0.12	4.736	0.000	0.330	0.811
M5	0.2489	0.12	2.071	0.042	0.009	0.489
M7	0.3532	0.12	2.942	0.005	0.113	0.593
M8	-1.1696	0.12	-9.744	0.000	-1.410	-0.93
M9	0.2435	0.12	2.029	0.047	0.004	0.483
Covid_Dummy	-0.0524	0.064	-0.824	0.413	-0.179	0.075

Table 5. Covid_19 pandemic effect check – Regression results.

Since $p(\text{Covid_Dummy}) = 0.413$ which is much higher than 0.05, we can't reject the null hypothesis, which means Covid-19 did not affect Landis+Gyr revenues. This can be also supported by having a look at the monthly sales plot per product family as presented in the preliminary analysis chapter, where no significant fluctuation is noticed.

Russian Invasion in Ukraine

Similar to Covid 19, another recent example of a factor that marked a pivotal geopolitical event with widespread economic consequences, including disruptions in supply chains, shifts in consumer behavior, and increased global uncertainty was the Russian invasion in Ukraine during February 2022.

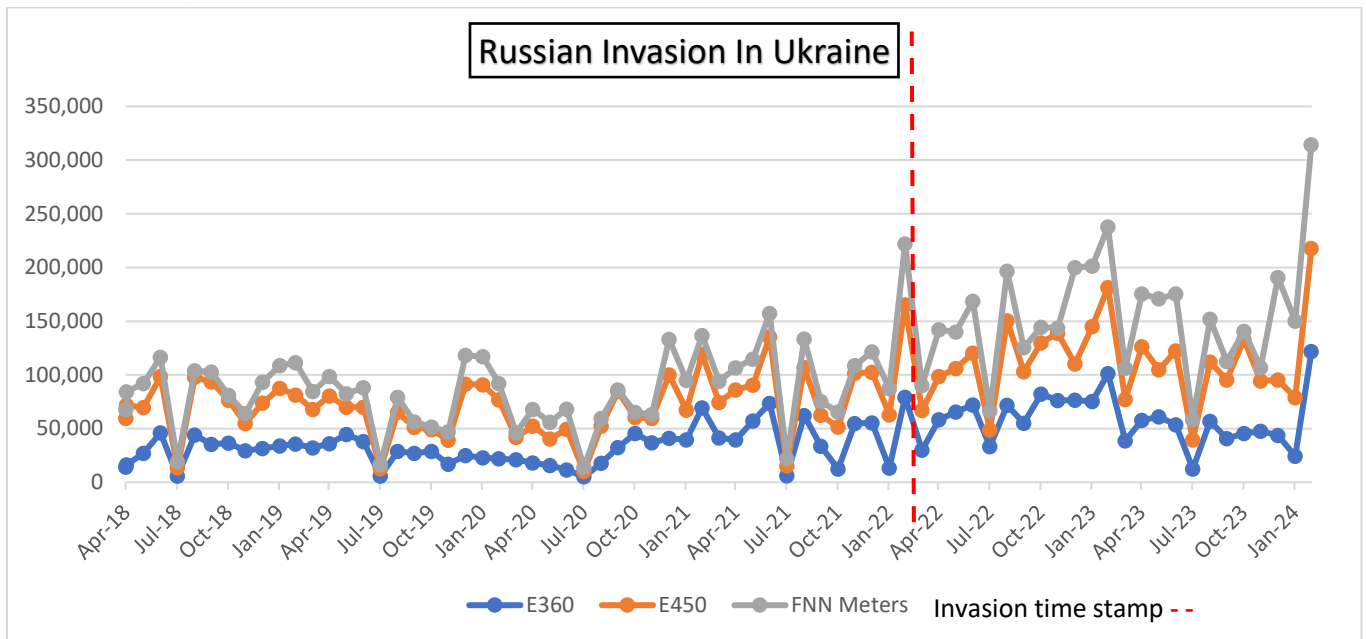


Figure 10: Monthly Sales data before and after the Russian invasion in Ukraine

The primary objective of this hypothesis test was to evaluate whether the Russian invasion of Ukraine in February 2022 significantly affected Landis+Gyr revenues. The goal was to quantify the effect of this event on monthly revenues using a dummy variable to represent the pre- and post-invasion periods.

To assess the impact, we employed a multiple linear regression model. We used a multiple linear regression model, incorporating a dummy variable to differentiate the periods before and after the invasion. The dummy variable was defined as:

- 1: For months before February 2022
- 0: For months from February 2022 onward.

The Hypothesis test is defined as follows:

- Null Hypothesis (H0): The coefficient is equal to zero ($\beta_{\text{invasion}}=0$), implying the variable is insignificant.
- Alternative Hypothesis (H1): The coefficient is not equal to zero ($\beta_{\text{invasion}}\neq 0$), implying the variable is significant.

By maintaining all the statistically significant variables, the regression model is now defined as follows (*Python script 4*

:

$$\ln(Y_t) = \beta_0 + \beta_1 t^2 + \sum (\beta_i M_i) \text{ (for } i=1,12) + \beta_{\text{ukraineUkraine_Invasion_Dummy}} + e_t$$

Where:

Ukraine_Invasion_Dummy: to isolate the invasion's effect

$\beta_{ukraine}$: Russian Invasion in Ukraine dummy coefficient

Based on the Regression results, the model explains 81.4% of the variance in total sales which reflects strong explanatory power for the seasonal and trend effects and shows a strong overall fit as per the F-statistic value (2.19×10^{-19}).

<i>Regression Statistics</i>	
Multiple R	0.89917
R Square	0.833
Adjusted R Square	0.809
Significance F	34.47
Observations	72

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	11.4913	0.160	71.853	0.000	11.172	11.811
Quadratic trend	8.16E-05	4.08E-05	2.000	0.050	4.24E-08	0.000
M1	0.4291	0.114	3.757	0.000	0.201	0.657
M2	0.233	0.114	2.045	0.045	0.005	0.461
M3	0.5582	0.114	4.899	0.000	0.330	0.786
M5	0.225	0.114	1.976	0.053	-0.003	0.453
M7	0.3428	0.114	3.018	0.004	0.116	0.57
M8	-1.1731	0.113	-10.339	0.000	-1.400	-0.946
M9	0.2472	0.113	2.179	0.0332	0.020	0.474
Ukraine_Invasion_Dummy	-0.379	0.133	-2.854	0.006	-0.645	-0.114

Table 6. Russian Invasion in Ukraine effect check – Regression results.

Since $p(\text{Ukraine_Invasion_Dummy}) = 0,006$ which is lower than 0,05, we can reject the null hypothesis, and suppose that the Russian Invasion in Ukraine did actually affect Landis+Gyr revenues.

However, by having a look at Figure 10 we can recognize that revenues did not drop directly after February 2022 but actually a month later during April 2022. In this case and by applying some critical thinking since revenues seem to correct shortly after with a steady increase, most probably the drop of the monthly revenues was also affected by the end of 2021 fiscal year (Landis+Gyr takes into account the period from April of current year- March of next year as full fiscal cycle), so a decrease in revenues would be also justified by that fact.

Application of Forecasting Models

SMA Application and Results

In this chapter, we apply the Simple Moving Average (SMA) forecasting technique with a window size (“q”) of 3 to predict sales for three product families: E360, E450, and FNN. SMA is chosen for its simplicity and effectiveness in smoothing past sales data to predict future values.

To evaluate the forecasting accuracy, the dataset was divided into an estimation sample (70%) that was used to calculate the SMA values and a forecasting sample (30%) used to evaluate the forecast accuracy using Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

By applying the relevant SMA formula ([Python Script 5](#)

) we have computed the following results for each product family:

Product Family	MAE	MAPE (%)	Result
E360	17962.60	45.72	<i>Reasonable prediction</i>
E450	18939.10	40.74	<i>Reasonable prediction</i>
FNN	25073.67	97.48	<i>Inaccurate prediction</i>

Table 7. Computed MAE & MAPE values per product family for SMA application

By investigating more thoroughly and separately each product family we can securely comment the following:

E360

The SMA forecasting for the E360 product family demonstrated moderate performance, as evidenced by a relatively low MAPE. The forecast errors show a balanced mix of overestimations and underestimations (Figure 12), suggesting that the method is effective at capturing seasonal trends in sales. The SMA method successfully smooths out fluctuations, resulting in predictions that closely align with actual sales, particularly during periods of stable demand (Figure 11). However, during months with unexpected sales spikes, the forecasts slightly underestimated actual sales.

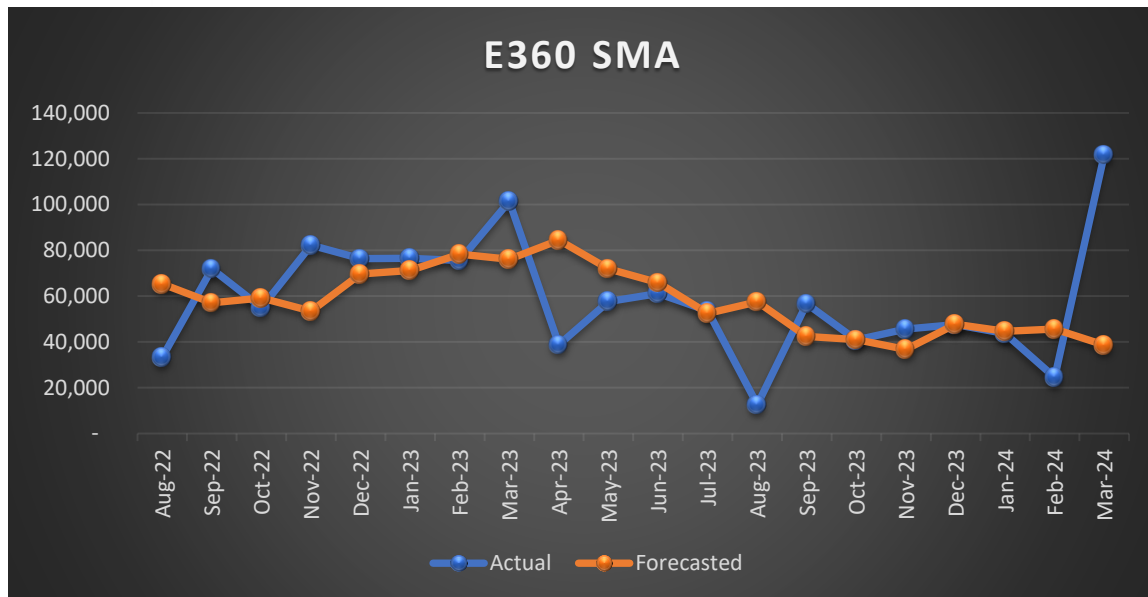


Figure 11: SMA performance for E360 Product Family

This indicates that while the SMA technique is effective for general trends, it might benefit from incorporating additional factors such as marketing campaigns or external shocks for more precise forecasting. The overall low error rates suggest that this method could be reliably used for short-term planning for E360.

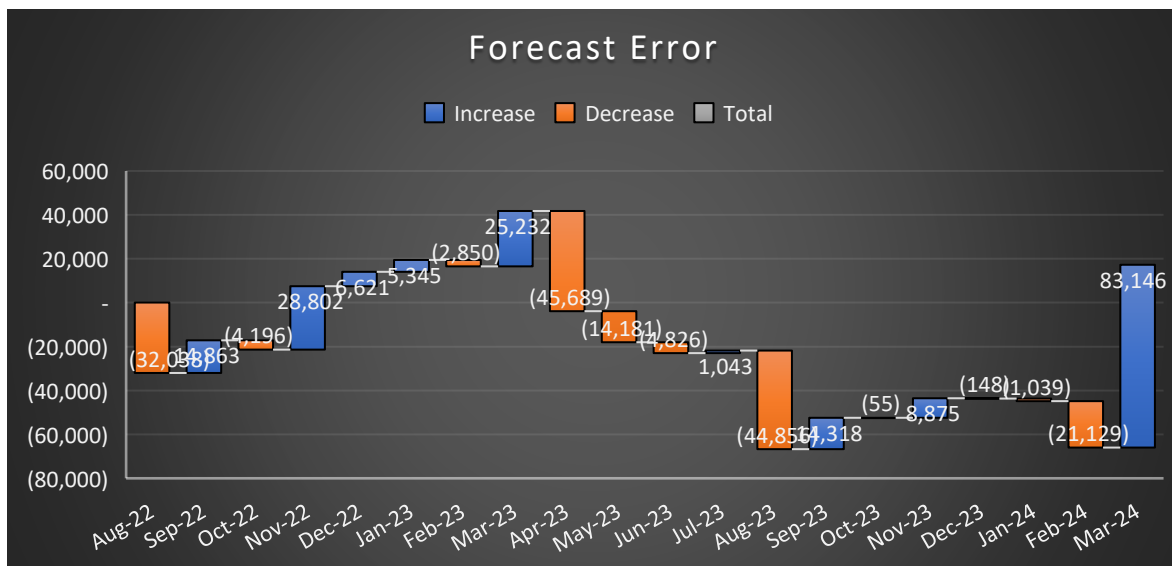


Figure 12: Forecast Error evolution concerning E360 forecasting sample

E450

The E450 product family exhibited reliable forecasting accuracy with a slightly lower MAPE compared to E360. Overestimations were observed during periods of declining sales, which is a common limitation of the SMA method as it lags in responding to rapid downward trends (Figure 13). Conversely, during periods of rising demand, forecasts tended to lag behind actual sales, leading to slight underestimations.

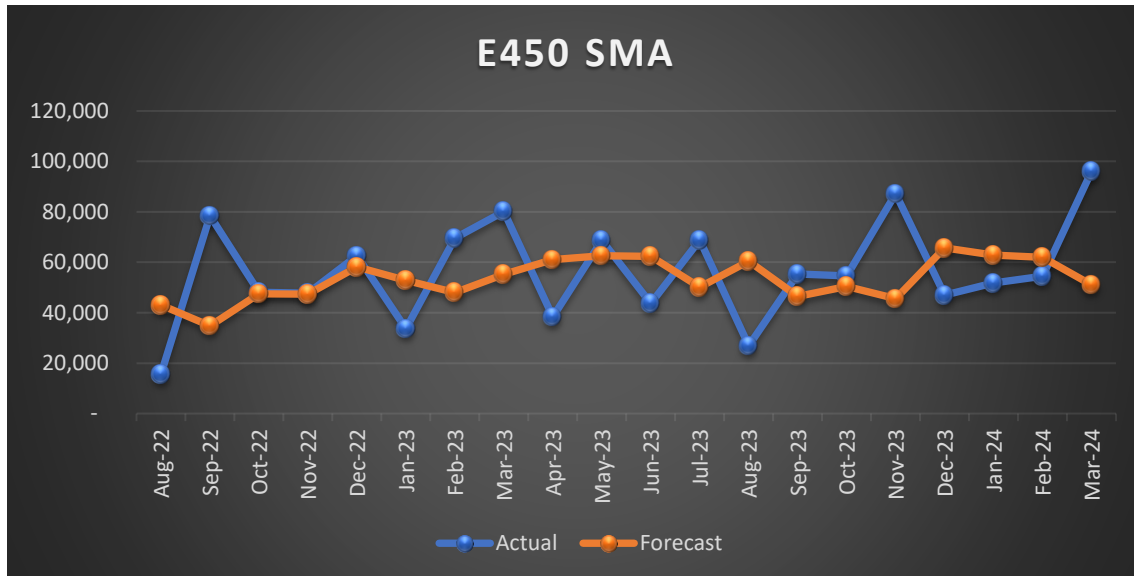


Figure 13: SMA performance for E450 Product Family

Despite these limitations, the forecasts were generally well-aligned with actual sales, indicating that the SMA method provides a reasonable baseline for short-term forecasting (Figure 14). To improve accuracy, integrating trend or seasonal adjustments might help mitigate the lag effect and better capture rapid sales changes.

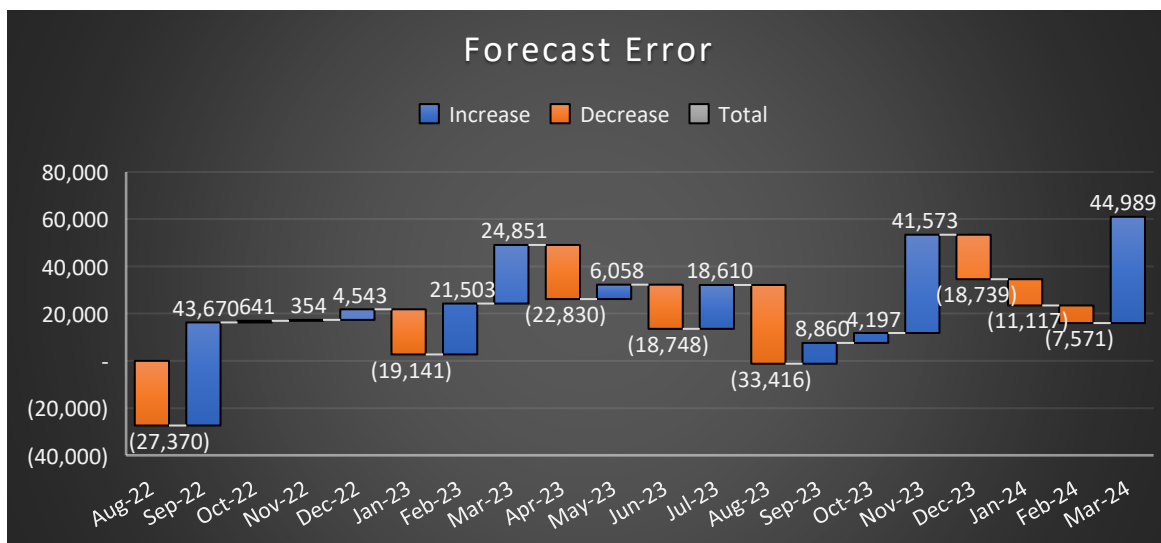


Figure 14: Forecast Error evolution concerning E450 forecasting sample

FNN

The FNN product family experienced the highest MAPE among the three, indicating greater volatility and unpredictability in its sales patterns. Forecast errors were characterized by consistent underestimations during sales surges, which is a direct consequence of the SMA method's reliance on past averages. During months of declining sales, the forecasts occasionally overestimated demand, reflecting the lagging nature of the SMA technique. (Figure 15)

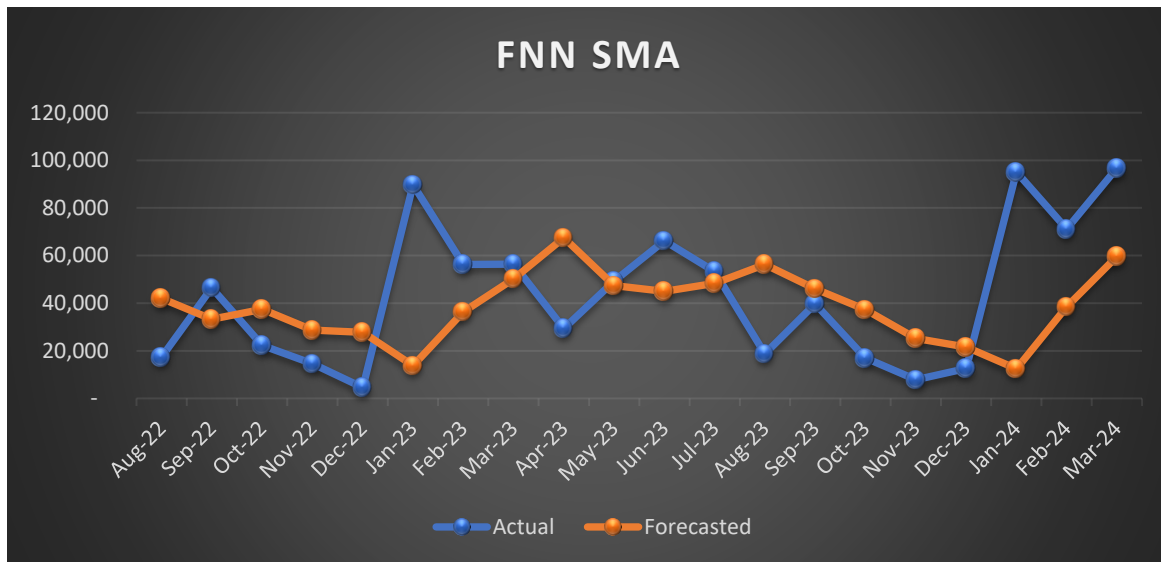


Figure 15: SMA performance for FNN Product Family

The high variability in FNN sales suggests that a more sophisticated forecasting model, such as exponential smoothing or linear regression, may be necessary to capture the irregular patterns (Figure 16). While the SMA approach provides a general trend, it may not be sufficient for products like FNN with highly dynamic sales behavior.

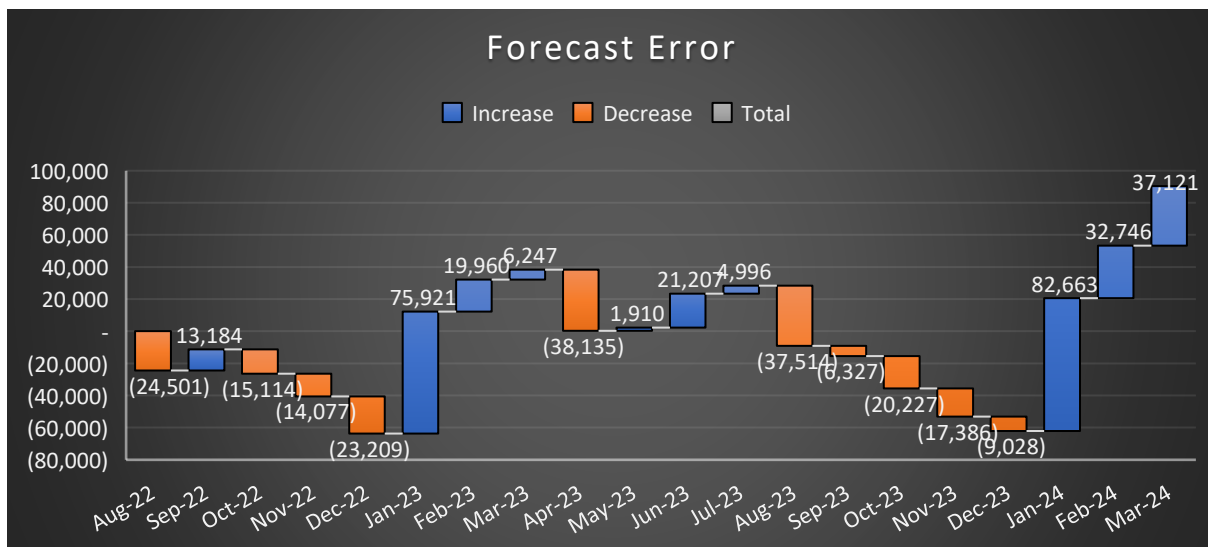


Figure 16: Forecast Error evolution concerning FNN forecasting sample

EWMA Application and Results

In this chapter, we apply the Exponentially Weighted Moving Average (EWMA) forecasting technique by assigning exponentially decreasing weights to past observations. It emphasizes recent data while still considering the influence of historical values.

To evaluate the forecasting accuracy, the dataset was divided into an estimation sample (70%) and a forecasting sample (30%) as we did for SMA.

By applying the relevant EWMA formula ([Python Script 6](#)

) we have computed the following results for each product family:

Product Family	λ value	MAE	MAPE (%)	Result
E360	0.4784	21890.73	44.39	<i>Reasonable prediction</i>
E450	0.0593	18198.52	32.29	<i>Reasonable prediction</i>
FNN	0.0117	24691.21	65.36	<i>Inaccurate prediction</i>

Table 8. Computed MAE & MAPE values per product family for EWMA application

By investigating more thoroughly and separately each product family we can securely comment the following:

E360

The optimal λ value for E360 is 0.4784, which reflects a balanced approach to forecasting by giving nearly equal weight to historical data and recent trends (Figure 17). This moderate value indicates that E360's sales exhibit some variability, requiring responsiveness to recent sales changes while still maintaining historical stability. The forecast model's MAPE of 44.39% shows a moderate level of accuracy. This suggests room for improvement in predicting sales, as the model may overestimate or underestimate sales during periods of sharp changes.

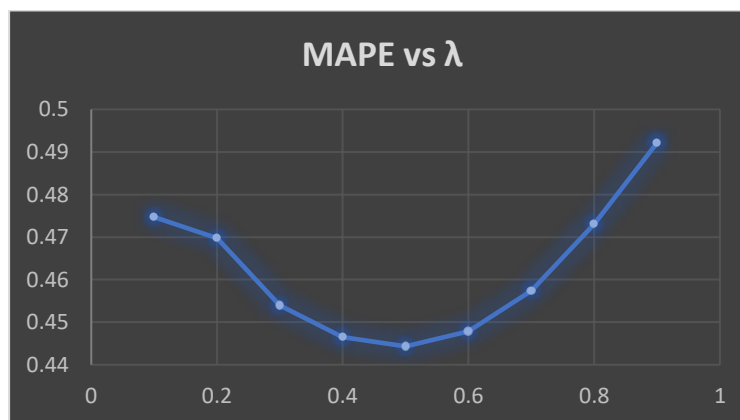


Figure 17: MAPE variation based on λ value

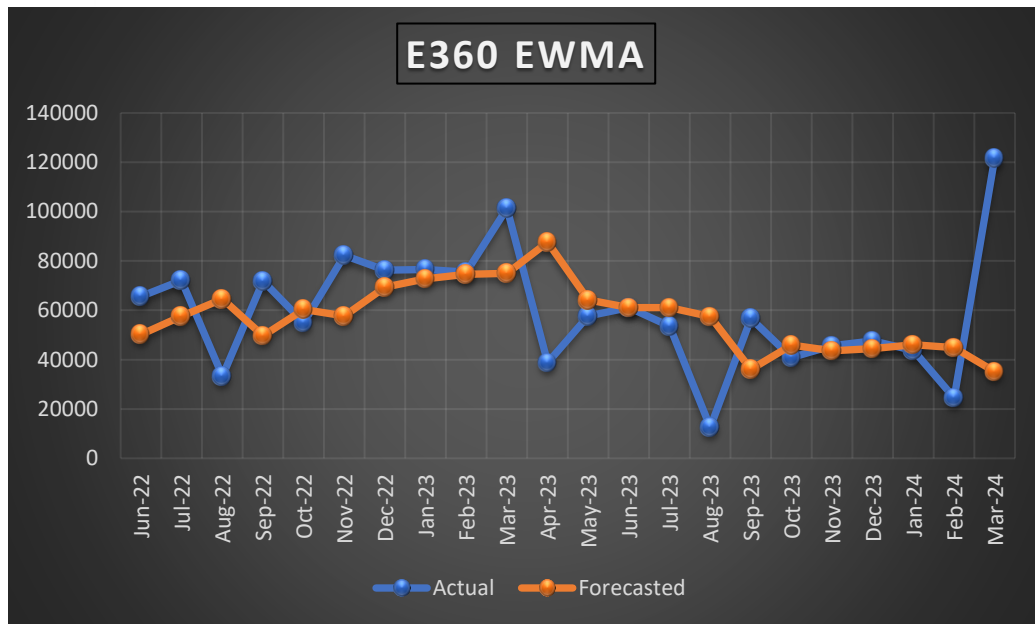


Figure 18: EWMA performance for E360 Product Family

Upon closer inspection, the model tends to slightly underestimate sales when actual sales experience rapid increases (Figure 18). This could indicate a need for additional adjustments or complementary forecasting methods to capture more dynamic shifts. For Landis+Gyr, this result underscores the need to monitor trends closely for E360, as unexpected changes in sales could lead to stockouts or overstock situations. By using EWMA, the company achieves a baseline forecast but should consider more advanced techniques for improved accuracy.

E450

E450 has an optimal λ value of 0.0593, a very low value that heavily weights historical data in the forecast (Figure 19). This indicates that E450's sales are relatively stable over time, with minimal short-term fluctuations. The low MAPE value of 32.29% highlights the model's strong performance in capturing E450's sales behavior. This suggests that the EWMA method is well-suited to forecast this product family without significant risk of overestimating or underestimating sales.

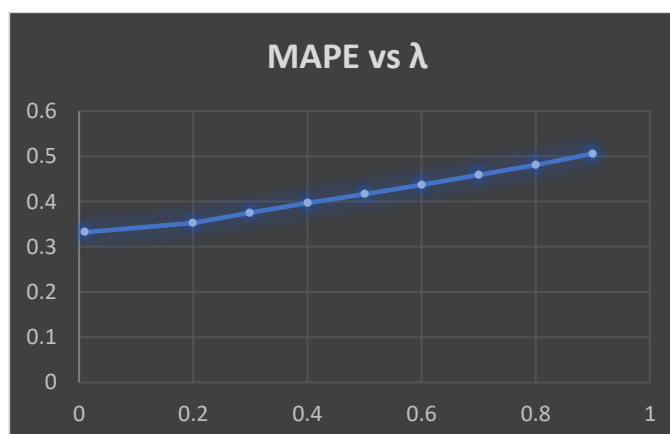


Figure 19: MAPE variation based on λ value

The tendency to rely on historical data ensures that forecasts remain consistent, but it could lag behind if sudden shifts in sales occur (Figure 20). For the company, this stability in E450's sales pattern is an advantage, allowing for efficient inventory management and reduced risk of stockouts or excess inventory. However, the company should remain vigilant about potential market disruptions or changes in customer preferences that could alter E450's historically stable pattern.

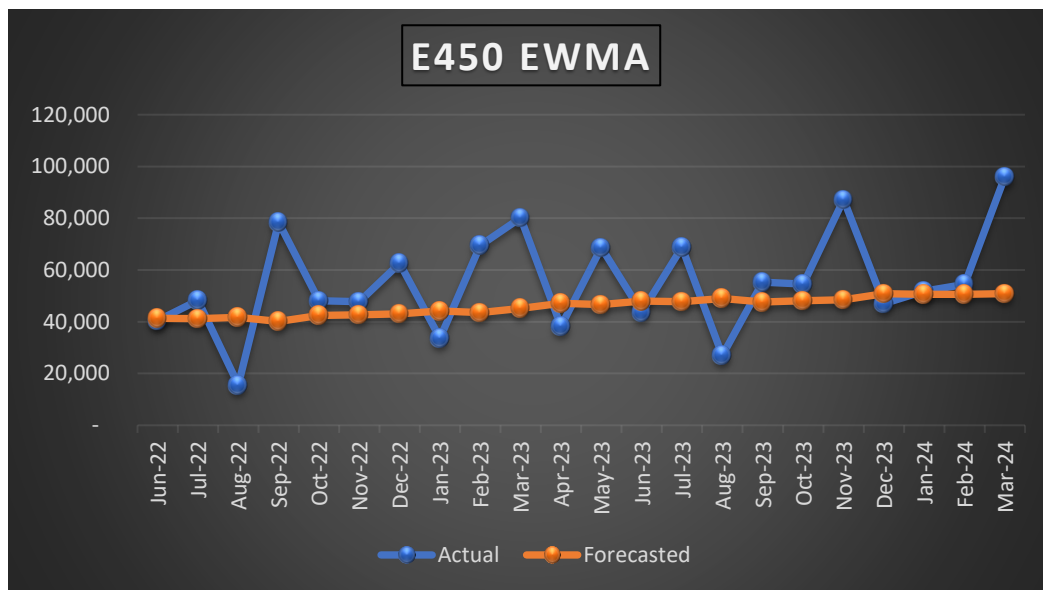


Figure 20: EWMA performance for E450 Product Family

FNN

FNN's optimal λ value is 0.0117, an extremely low value that almost entirely relies on historical data (Figure 21). This indicates that FNN's sales are highly consistent, with little need to adjust for recent changes. However, the MAPE value of 65.36% reveals significant inaccuracies in the forecast, suggesting that the EWMA method struggles to capture the fluctuations of FNN's sales behavior.

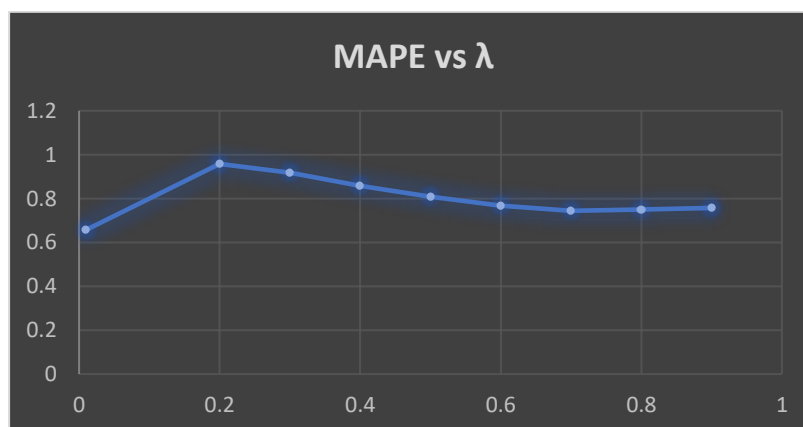


Figure 21: MAPE variation based on λ value

This high MAPE results from underestimating the impact of sudden fluctuations in sales (Figure 22). The model's reliance on historical data makes it ill-equipped to handle spikes or drops in sales, leading to substantial errors. For Landis+Gyr, this result indicates that while FNN's sales are generally stable, relying solely on EWMA for forecasting may not be sufficient. The company should consider integrating other forecasting models, such as regression or machine learning-based approaches, to complement EWMA and address these challenges. Additionally, understanding the factors driving variability in FNN sales could help refine forecasting strategies and improve accuracy.

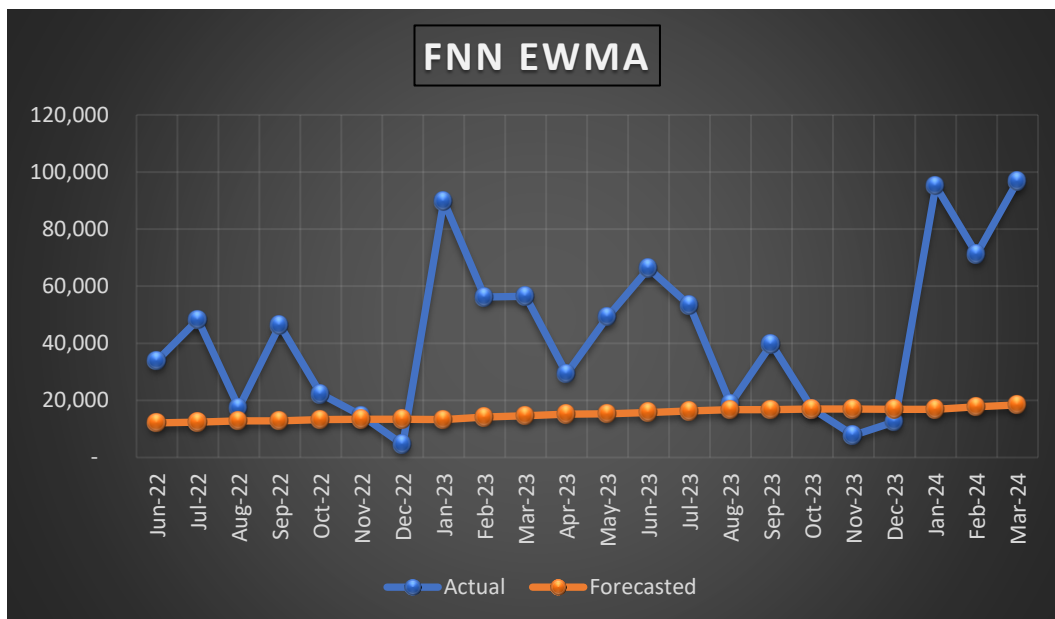


Figure 22: EWMA performance for FNN Product Family

Regression Analysis Results

Following the application of SMA and EWMA, in this section we will investigate seasonality and trends concerning the monthly revenues of Landis+Gyr products.

The objective of this chapter is to try to capture seasonal patterns, while the time trend variables account for linear and nonlinear changes in sales over time, with the use of monthly dummy variables. To achieve this, we employed regression analysis as our forecasting technique due to its robust ability to model trends and seasonality while explaining relationships between variables.

To ensure a reliable evaluation of the model's performance, we split the dataset into an estimation sample (70% of the data) for model training and a forecasting sample (30%) to validate its predictive accuracy. This split provides a realistic simulation of out-of-sample forecasting, which is critical for assessing how well the model generalizes to new data.

The regression model we are going to use is the following:

$$\ln Y_t = \beta_0 + \sum_{i=1}^I \beta_i M_i + \gamma_0 t + \gamma_1 t^2 + \epsilon_t$$

Where:

- $\ln Y_t$: Logarithm of monthly sales (dependent variable)
- β_0 : Constant term
- M_i : Monthly dummy variables representing seasonality
- t : Linear time trend capturing overall growth/decline
- t^2 : Quadratic time trend capturing acceleration or deceleration trends
- $\beta_i, \gamma_0, \gamma_1$: Coefficients
- ϵ_t : Error term

By applying the specific regression formula through python environment ([Python Script 7](#)

, we have computed the following results:

Product Family	MAE	MAPE	Result
E360	29001.08	62%	Insignificant
E450	47612.84	87%	Insignificant
FNN	48522.41	131%	Insignificant

Table 9. Computed MAE & MAPE values per product family

At a first glance and based on the computed MAE and MAPE values, it seems that regression cannot be considered effective for any of the three product families.

However, we will try to investigate our findings more thoroughly by checking each product family separately.

E360

The R-squared value of 0.61 indicates that 61% of the variability in E360 sales is explained by the model, highlighting its relatively fit.

Regression Statistics					
Multiple R	0.779502835				
R Square	0.60762467				
Adjusted R Square	0.465933578				

Standard Error	0.481467751					
Observations	50					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	13	12.92321621	0.994093555	4.288375952	0.000261082	
Residual	36	8.345203024	0.231811195			
Total	49	21.26841923				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	10.24969282	0.323301007	31.70325056	6.70837E-28	9.594007991	10.90537766
M8	1.547469217	0.340760367	4.541224173	6.04737E-05	2.238563274	-0.85637516

Table 10. E360 Regression results (only statistically significant variables depicted)

The forecasting model for E360 demonstrated a moderate level of accuracy, with a MAPE of 62%, indicating that forecasts deviate from actual sales by an average of 62%. The MAE of 29,001 units suggests a considerable absolute error, but this may be acceptable depending on the scale of sales for this product family.

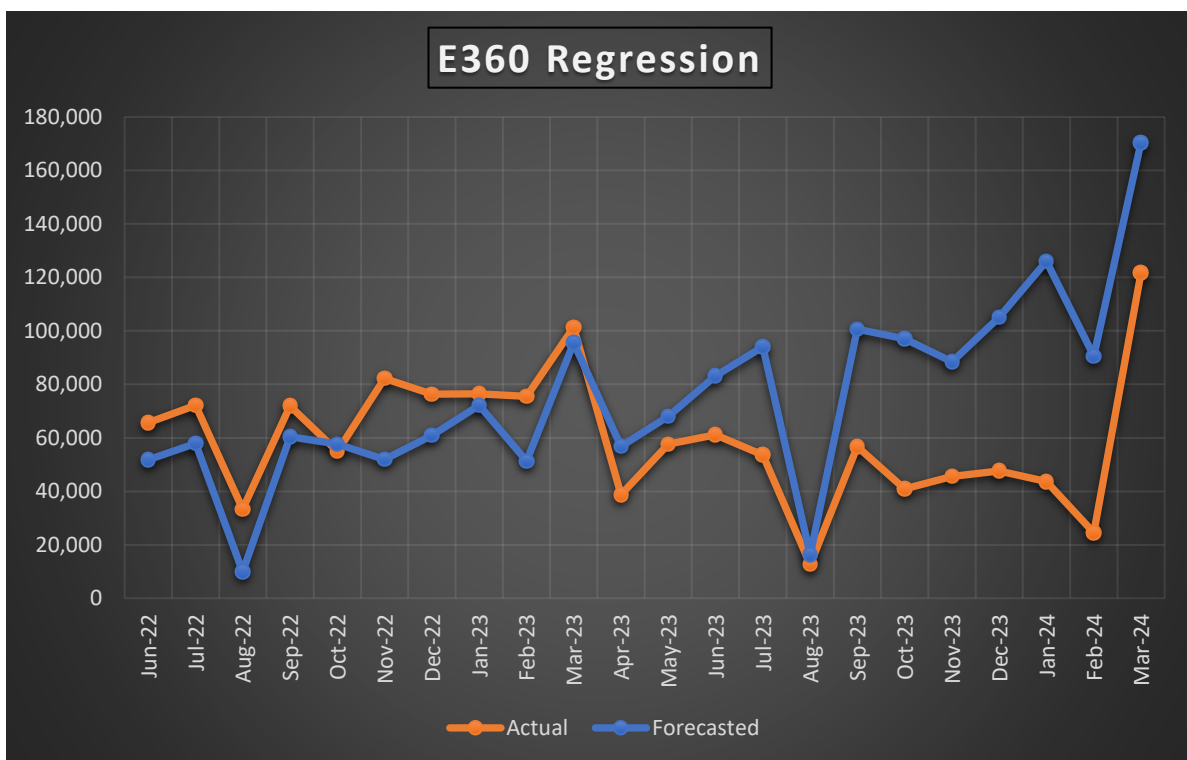


Figure 23: Regression performance for E360 Product Family

For Landis+Gyr, the results imply that the model captures general trends and seasonality moderately and may struggle with highly volatile or irregular demand patterns. Overestimations and underestimations were observed in different periods, reflecting the challenges of accurately forecasting in dynamic market conditions (Figure 23).

However, the performance is within an actionable range for operational planning, enabling better inventory management and production scheduling.

E450

The R-squared value of 0.65 indicates that 84,6% of the variability in E450 sales is explained by the model, which reflects a strong level of explanatory power.

Regression Statistics						
Multiple R	0.919870713					
R Square	0.846162129					
Adjusted R Square	0.790609565					
Standard Error	0.272583599					
Observations	50					
ANOVA						
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>	
Regression	13	14.71269621	1.131745862	15.23173841	6.7004E-11	
Residual	36	2.674865463	0.074301818			
Total	49	17.38756167				
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	10.57126522	0.183037289	57.75470817	4.3188E-37	10.20004839	10.94248205
Linear trend	-	-	-	-	-	-
Quadratic trend	0.041760943	0.011156183	3.743300126	0.000633021	0.064386731	0.019135154
M1	0.000800458	0.000212172	3.772681449	0.000581959	0.000370153	0.001230763
M2	0.702991823	0.192826385	3.645724231	0.000835649	0.311921789	1.094061858
M3	0.588720207	0.192934529	3.051398882	0.00426092	0.197430845	0.980009568
M5	0.772868646	0.193098046	4.002467459	0.000299321	0.381247657	1.164489635
M7	0.426085949	0.184284825	2.312105461	0.026603584	0.052339002	0.799832897
M8	0.521053677	0.193068203	2.69880627	0.010529331	0.129493213	0.912614142
M9	-	-	-	-	-	-
M8	1.327562196	0.192921929	6.881344188	4.70024E-08	1.718826003	0.936298388
M9	0.465198966	0.192822649	2.412574288	0.021060502	0.074136508	0.856261425

Table 11. E450 Regression results (only statistically significant variables depicted)

The E450 model exhibited less accuracy compared to E360, with a MAPE of 87% and an MAE of 47,612 units. This suggests that forecasts for this product family are less reliable, with significant deviations from actual sales (Figure 24).

For Landis+Gyr, these findings highlight the need to refine the model further, possibly by incorporating additional variables such as macroeconomic indicators, competitive actions, or customer behavior. Despite these challenges, the model provides valuable insights into broad seasonal patterns, which can still guide strategic decisions.

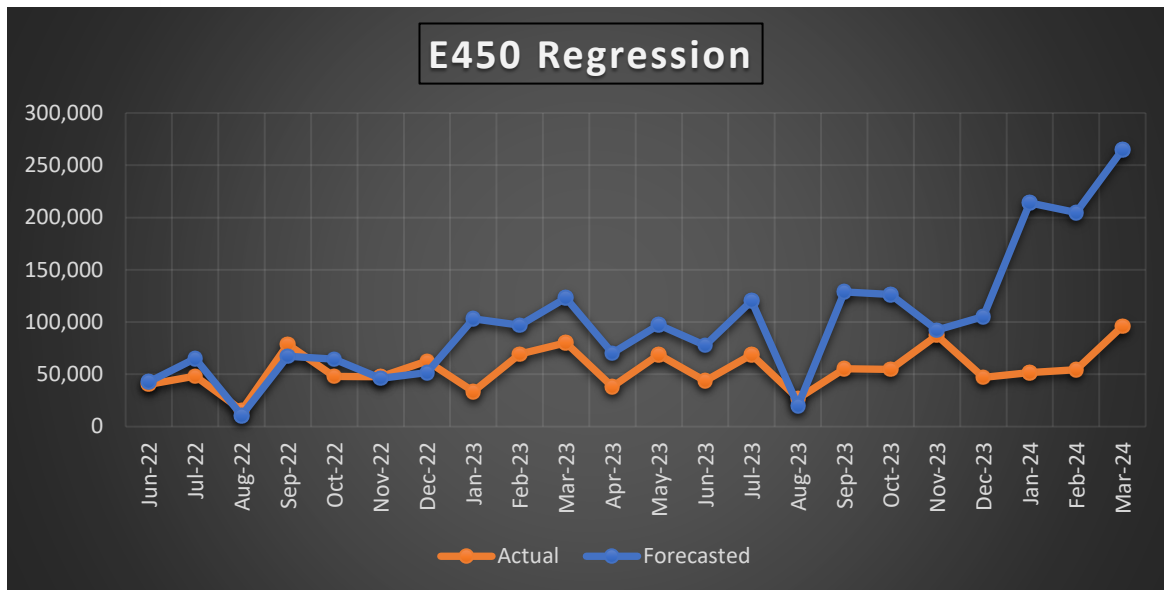


Figure 24: Regression performance for E450 Product Family

FNN

The R-squared value of 0.74 suggests that only 74% of the variability in FNN sales is explained by the model, indicating a moderate fit compared.

Regression Statistics						
Multiple R	0.860212186					
R Square	0.739965004					
Adjusted R Square	0.646063478					
Standard Error	0.470379189					
Observations	50					
ANOVA						
	df	SS	MS	F	Significance F	
Regression	13	22.66616673	1.743551287	7.880223394	4.27992E-07	
Residual	36	7.965236922	0.221256581			
Total	49	30.63140365				
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	8.706705936	0.315855143	27.56550315	8.48593E-26	8.066122016	9.347289857
Linear trend	0.040527925	0.019251476	2.105185375	0.042317957	0.079571727	0.001484123
Quadratic trend	0.001028499	0.000366131	2.809100096	0.007979252	0.000285951	0.001771047
M1	1.519706129	0.332747527	4.567144775	5.59318E-05	0.844862866	2.194549392
M2	1.533265888	0.332934145	4.605312832	4.98498E-05	0.858044147	2.208487629
M3	1.563772342	0.333216315	4.69296452	3.82444E-05	0.887978333	2.23956635
M4	0.777301688	0.318042441	2.444018748	0.019554506	0.132281722	1.422321654
M5	1.259041814	0.318007931	3.959152246	0.000339599	0.614091837	1.90399179

M6	1.297779249	0.333512613	3.891244888	0.000413583	0.62138432	1.974174178
M7	1.355084012	0.333164817	4.067308265	0.000247595	0.679394446	2.030773578
M9	0.824225917	0.332741081	2.477078919	0.018077394	0.149395727	1.499056108

Table 12. FNN Regression results (only statistically significant variables depicted)

The FNN model demonstrated the highest errors, with a MAPE of 131% and an MAE of 48,522 units. This indicates substantial challenges in accurately forecasting this product family.

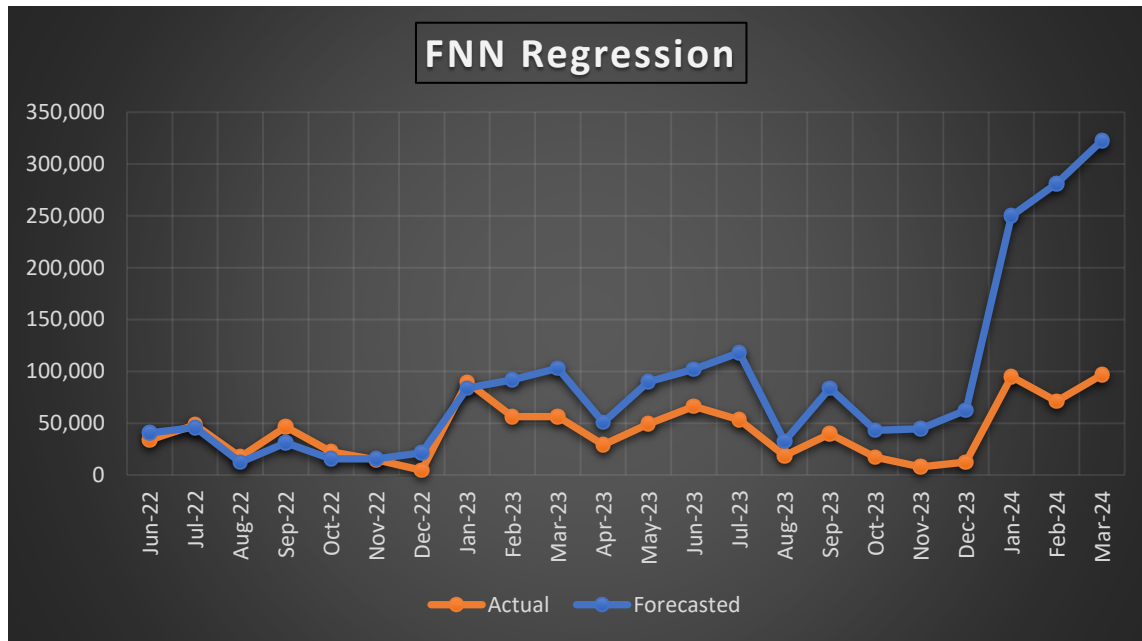


Figure 25: Regression performance for FNN Product Family

For Landis+Gyr, the implications are significant. Such high error rates point to the need for alternative forecasting methods or additional data sources to improve accuracy. Overestimations and underestimations (Figure 25) in this product family could lead to costly inefficiencies, such as excess inventory or stockouts, and require immediate attention to mitigate operational risks.

Model Comparison and Effectiveness

In this chapter, we compare the performance and effectiveness of the three forecasting models applied to predict sales for the product families E360, E450, and FNN. The forecasting methods evaluated are Simple Moving Average (SMA), Exponentially Weighted Moving Average (EWMA), and regression analysis. The evaluation metrics used include Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE).

The effectiveness of each model is assessed using MAE to measure the average magnitude of errors in forecasts without considering their direction and MAPE which

expresses forecast accuracy as a percentage of actual values, providing a scale-independent measure of error.

The dataset was consistently divided into an estimation sample (70%) and a forecasting sample (30%) to ensure comparability and reliability of results.

SMA Model Performance

The SMA method, with a window size of 3, was applied to smooth past sales data and predict future values. The results indicate:

Product Family	MAE	MAPE (%)
E360	17962.60	45.72
E450	18939.10	40.74
FNN	25073.67	97.48

Based on the above, we can notice that SMA performed reasonably well for E360 and E450, demonstrating moderate accuracy and capturing general sales trends. However, for FNN, the high MAPE reflects significant forecasting errors due to the volatile and irregular sales pattern.

SMA's inherent lag effect limited its ability to respond to rapid changes, particularly in dynamic sales environments.

EWMA Model Performance

The EWMA method, which assigns exponentially decreasing weights to historical data, was applied to emphasize recent trends. The optimal smoothing parameter (λ) was determined for each product family. The results are as follows:

Product Family	λ value	MAE	MAPE (%)
E360	0.4784	21890.73	44.39
E450	0.0593	18198.52	32.29
FNN	0.0117	24691.21	65.36

EWMA showed improved performance over SMA for E450, reflecting its ability to better handle stable sales patterns by heavily weighting historical data. For E360, the performance was comparable to SMA, with moderate accuracy and sensitivity to recent changes. Last but not least, FNN's high variability posed challenges for EWMA, as the model struggled to account for sudden sales fluctuations.

Regression Analysis Performance

A regression model incorporating seasonal dummy variables, linear, and quadratic trends was applied. The results are summarized below:

Product Family	MAE	MAPE
E360	29001.08	62%
E450	47612.84	87%
FNN	48522.41	131%

The regression model performed moderately well for E360, capturing trends and seasonality with an R-squared value of 0.61. However, the high MAPE indicates room for improvement in capturing dynamic changes.

For E450 and FNN, the model's high errors and less significant fit suggest it is less effective for these product families. Regression's reliance on predefined trends and seasonality may limit its flexibility in adapting to irregular patterns, as observed in FNN.

Comparative Analysis

Product Family	Best Model	Key Strengths	Limitations
E360	EWMA	Balanced responsiveness to trends	Slight underestimation during sharp changes
E450	EWMA	Strong performance with stable patterns	Lags in capturing sudden shifts
FNN	None	General trend indication (SMA/EWMA)	Inability to handle high volatility

- E360: EWMA provided the best balance of accuracy and responsiveness, though alternative models could improve precision during sharp demand changes.
- E450: EWMA's strong performance highlights its suitability for stable sales patterns, but trend-corrected methods could further refine forecasts.
- FNN: Neither SMA, EWMA, nor regression sufficiently addressed FNN's volatility. Advanced methods such as ARIMA or machine learning models should be explored.

Through this comparison, we need to highlight that no single forecasting method performs optimally across all product families. While EWMA emerges as the most effective for E360 and E450, it is insufficient for FNN. By adopting a tailored, hybrid approach and leveraging advanced techniques, Landis+Gyr can enhance forecasting accuracy and better address the unique characteristics of each product family.

Conclusions, Limitations, and Future Research

Summary of Findings

This research explored the application of advanced data analytics techniques to enhance demand forecasting for Landis+Gyr's high-demand smart meter products. By analyzing sales data spanning six years (2018–2024) and implementing multiple forecasting models, the study aimed to optimize inventory levels, reduce inefficiencies, and adapt to external influences. Three primary forecasting techniques were employed: Exponential Weighted Moving Average (EWMA), Simple Moving Average (SMA), and regression analysis.

The analysis addressed several key research questions and objectives:

- What are the main statistical properties and components of the sales time series for Landis+Gyr's different meter products?

The research revealed that the sales data exhibited significant seasonality and trends. For instance, a March peak was consistently observed across product families, indicating increased demand at the close of the fiscal year. In contrast, August showed a marked decline, attributed to production shutdowns. The data for E360 and E450 showed relatively stable trends, while FNN exhibited more volatility, influenced by external market demands. These insights underscore the importance of understanding time-series properties to tailor forecasting approaches effectively.

- How accurate are different forecasting models in predicting monthly sales for the company's smart meter products, and does this accuracy vary across different product families?

EWMA demonstrated superior performance in capturing short-term fluctuations, particularly during periods of external disruption such as the COVID-19 pandemic. SMA provided consistent results for identifying long-term trends in stable product categories. Regression analysis emerged as the most robust method for integrating multiple demand drivers, achieving higher accuracy by accounting for external factors. However, model performance varied among product families, with FNN's volatility presenting challenges for traditional models, suggesting the need for advanced techniques in such scenarios.

They findings are definitely that EWMA performed best for E450, leveraging stable historical trends while adapting to short-term changes, SMA was effective for long-term trend detection in E360 and E450 but struggled with dynamic changes in FNN sales.

Regression Analysis proved most insightful for integrating external variables, such as geopolitical events and seasonality, though its accuracy was challenged by the high variability in FNN sales

- Are there any interdependencies or cross-correlations among the sales of different product families, and how do these relationships impact market behavior?

The analysis identified moderate correlations between E450 and FNN sales, suggesting shared market drivers or complementary product demand. However, cross-product cannibalization was not evident. These insights underline the importance of considering product interdependencies in strategic planning.

- How have external factors, such as the COVID-19 pandemic and geopolitical events like the Russian invasion of Ukraine, affected sales patterns and market dynamics for the company's products?

The regression analysis conducted in this study revealed that external factors, such as the COVID-19 pandemic and the Russian invasion of Ukraine, did not exhibit a statistically significant impact on Landis+Gyr's revenues. Specifically for COVID-19 Pandemic a dummy variable representing the pandemic period was included in the regression model to assess its impact on sales from December 2019 to December 2022. The p-value associated with this variable was 0.413, far above the significance threshold of 0.05. This indicates no measurable disruption to revenues caused by the pandemic. Visual analysis of sales trends corroborated these findings, showing consistent monthly sales across the examined period.

Similarly, the invasion's effect was tested using a dummy variable representing the pre- and post-February 2022 periods. The regression results again demonstrated no significant impact on revenues, as the associated variable failed to reach statistical significance. The results suggest that Landis+Gyr's operations and demand for its products remained resilient despite global geopolitical disruptions.

Overall, the study emphasized the importance of combining forecasting models to create a robust and adaptable framework. The integration of internal sales data with external factors offered a comprehensive understanding of demand drivers, enabling Landis+Gyr to enhance its operational efficiency and maintain a competitive edge.

Research Limitations

Despite the comprehensive approach, this study faced several limitations that warrant consideration.

First of all, while the data were spanning through full 6 fiscal years the dataset was restricted to internal sales figures for three product families. While this provided a solid foundation for analysis, the exclusion of external macroeconomic data, market competition trends, and industry-wide variables limited the ability to capture a holistic view of demand dynamics.

Additionally, the analysis was influenced by significant external events, such as the COVID-19 pandemic and geopolitical tensions. While regression models accounted for these disruptions, the unpredictability of such events introduced anomalies that traditional forecasting methods struggled to address fully. Machine learning or hybrid models could better accommodate such irregularities in future studies.

The study was confined to three high-demand product families, which, while critical for Landis+Gyr's operations, may not represent the demand patterns of other product lines. This limitation restricts the generalizability of the findings across the company's full product portfolio.

Moreover, the applied forecasting models relied on specific assumptions, such as stationarity in time series data and linear relationships in regression analysis. These assumptions may not always hold, potentially affecting the accuracy of the predictions.

While traditional statistical models were effective, they lacked the dynamism of advanced machine learning techniques. Static models like EWMA and SMA may underperform in highly volatile markets or when demand drivers evolve rapidly.

Addressing these limitations in future research can enhance the reliability and applicability of demand forecasting methodologies, enabling more comprehensive and adaptive models.

Recommendations for Future Research

In conclusion of our research and building upon the insights and limitations identified, the following recommendations are proposed for future research.

Future studies could integrate broader macroeconomic indicators, such as GDP growth, inflation rates, and energy policy changes. These variables can provide a more nuanced understanding of the factors influencing demand, particularly in the highly regulated smart meter industry.

Additionally, machine learning models, such as Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) networks, should be explored to enhance forecasting accuracy. These methods excel in capturing nonlinear relationships and adapting to anomalies, making them well-suited for volatile markets.

Extending the analysis to include additional product families and geographical markets could also provide a more comprehensive view of Landis+Gyr's demand landscape. This broader scope would help identify interdependencies across products and regions, enabling better strategic planning.

Incorporating scenario analysis to account for potential disruptions—such as supply chain bottlenecks or economic downturns—can improve the resilience of forecasting models. By simulating various scenarios, Landis+Gyr can better prepare for unexpected market shifts.

Leveraging real-time data from IoT-enabled devices and smart meters can provide more accurate and timely insights into demand patterns. This approach aligns with the increasing digitization of supply chains and enhances responsiveness.

Benchmarking Landis+Gyr's forecasting practices against other industries, such as consumer electronics or automotive manufacturing, could possibly also provide valuable lessons and identify best practices for supply chain optimization.

Combining traditional statistical techniques with machine learning approaches can create hybrid models that leverage the strengths of both methodologies. For instance, regression analysis can identify demand drivers, while machine learning models can adapt to dynamic market conditions.

Of course these recommendations for future research can build a more robust and adaptive forecasting framework, enabling Landis+Gyr to navigate the complexities of the smart meter industry effectively and maintain its leadership position in the market and should be maintained and further discussed as food for thought for the company's management and portfolio, taking into account all the involved stakeholders.

References

1. Abbasimehr, H., & Shabani, M. (2020). A new framework for predicting customer behavior in terms of RFM by considering the temporal aspect based on time series techniques. *Journal of Ambient Intelligence and Humanized Computing*.
2. Albu, M. M., Sănduleac, M., & Dumitrescu, A. M. (2021). High reporting rate smart metering data for enhanced grid monitoring and services for energy communities. *IEEE Transactions on Industrial Informatics*, 18(6), 4039–4048.
3. Baglini, G. (2024). Estimate Volatility with SMA and EWMA in Python. EODHD APIs Academy.
4. Baltagi, B. H. (2019). *Econometric Analysis of Panel Data*. Springer.
5. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2015). *Time Series Analysis: Forecasting and Control*. John Wiley & Sons.
6. Camur, M. C., Tseng, C. Y., Thanos, A. E., White, C. C., Yund, W., & Iakovou, E. (2023). An integrated system dynamics and discrete event supply chain simulation framework for supply chain resilience with non-stationary pandemic demand. *arXiv preprint arXiv:2305.00086*.
7. Chatfield, C. (2004). *The Analysis of Time Series: An Introduction* (6th ed.). CRC Press.
8. Chen, Y., & Zhang, L. (2019). Quality Control in Manufacturing Processes Using EWMA Control Charts. *Journal of Quality Engineering*, 67(4), 215–225.
9. Fildes, R., & Goodwin, P. (2007). “Against Your Better Judgment? How Organizations Can Improve Their Use of Management Judgment in Forecasting.” *Interfaces*, 37(3), 570–576.
10. Franses, P. H., & van Dijk, D. (2000). *Non-Linear Time Series Models in Empirical Finance*. Cambridge University Press.
11. Garcia, M., & Lopez, R. (2019). Churn Prediction in Streaming Services: A Machine Learning Approach. *IEEE Transactions on Multimedia*.
12. Gardner, E. S. (2006). “Exponential Smoothing: The State of the Art – Part II.” *International Journal of Forecasting*, 22(4), 637–666.
13. Goyal, S., & Goyal, R. (2021). “Comparative Study of Forecasting Techniques in Retail and Manufacturing Sector.” *Journal of Business Analytics*, 9(1), 29–41.
14. Green, K. C., & Armstrong, J. S. (2007). “Forecasting Principles.” In Armstrong, J. S. (Ed.), *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Springer.
15. Gupta, G., & Singharia, K. (2021). Consumption of OTT media streaming in COVID-19 lockdown: Insights from PLS analysis. *Vision*, 25(1), 36–46. <https://doi.org/10.1177/0972262921989118>
16. Hoda, S., Singh, A., Rao, A., Ural, R., & Hodson, N. (2021). Consumer demand modeling during COVID-19 pandemic. *arXiv preprint arXiv:2105.01036*.
17. Huhta, K. (2017). Prioritising energy efficiency and demand side measures over capacity mechanisms under EU energy law. *Journal of Energy & Natural Resources Law*, 35(1), 7–24.

18. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice* (2nd ed.). OTexts.
19. Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). "A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods." *International Journal of Forecasting*, 18(3), 439–454.
20. Inyurt, S., Hasanpour Kashani, M., & Sekertekin, A. (2020). Ionospheric TEC forecasting using Gaussian process regression (GPR) and multiple linear regression (MLR) in Turkey. *Astrophysics and Space Science*, 365(1), 1-17.
21. Jaggia, S., & Kelly, A. (2019). *Business Statistics: Communicating with Numbers* (4th ed.). McGraw-Hill Education.
22. Kilbourn, P. (2021). Forecast accuracy in demand planning: A fast-moving consumer goods case study. *Journal of Transport and Supply Chain Management*, 15, 10 pages. <https://doi.org/10.4102/jtscm.v15i0.583>
23. Kourentzes, N., & Petropoulos, F. (2016). "Forecasting with Intermittent Demand Using Multiple Temporal Aggregation." *Journal of the Operational Research Society*, 67(1), 58–68.
24. Kumar, S., & Gupta, R. (2021). Application of SMA in Inventory Management for Manufacturing Firms. *Journal of Manufacturing Systems*, 45, 123-130.
25. Lapide, L. (2006). "Sales and Operations Planning (S&OP): The Next Frontier for Demand Management." *Journal of Business Forecasting*, 25(1), 30–36.
26. Lawrence, M. J., & O'Connor, M. (1995). "Judgmental Forecasting in the Presence of Loss Functions." *International Journal of Forecasting*, 11(1), 33–43.
27. Li, F., & Kang, Y. (2023). Feature-based intermittent demand forecast combinations: Accuracy and inventory implications. *International Journal of Production Research*.
28. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 15(3), e0231281.
29. Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications* (3rd ed.). John Wiley & Sons.
30. Mammadova, U., & Özkale, M. R. (2024). Detecting shifts in Conway–Maxwell–Poisson profile with deviance residual-based CUSUM and EWMA charts under multicollinearity. *Statistical Papers*, 65(2), 597-643.
31. Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting*. John Wiley & Sons.
32. Morlidge, S., & Player, S. (2010). *Future Ready: How to Master Business Forecasting*. John Wiley & Sons.
33. Neshat, N., & Hadian, H. (2019). Technological learning modelling towards sustainable energy planning. *Journal of Engineering, Design and Technology*.
34. Ord, J. K., & Fildes, R. (2012). "Principles of Business Forecasting." In Gilliland, M. (Ed.), *The Business Forecasting Fieldbook*. Wiley.
35. Petropoulos, F., & Makridakis, S. (2020). "Forecasting Demand with Machine Learning: A Review of Applications in Demand Planning." *Operations Research Perspectives*, 7, 100-150.

36. Sanders, N. R., & Manrodt, K. B. (2003). "The Evolving Role of Collaboration in Supply Chain Forecasting." *Production and Inventory Management Journal*, 44(2), 22–28.
37. Schittekatte, T., & Meeus, L. (2020). Flexibility markets: Q&A with project pioneers. *Utilities Policy*, 63, 101017.
38. Sharma, R., Tiwari, M. K., & Li, X. (2021). Time-series forecasting methods for supply chain management: A review and future research directions. **Computers & Industrial Engineering*, 162*, 107783.
39. Singh, A., & Patel, D. (2021). Application of EWMA in Monitoring Production Yield Rates. *Semiconductor Manufacturing Studies*, 34(2), 97-104.
40. Smith, J., & Doe, A. (2021). Predicting User Engagement in Online Streaming Services Using Machine Learning Techniques. *Journal of Data Science and Analytics*.
41. Stevens, K. (2021). Taking Netflix to the cinema: National cinema value chain disruptions in the age of streaming. *Media Industries Journal*, 8(1), 22 pages.
<https://doi.org/10.3998/mij.60276950>
42. Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). "On the Categorization of Demand Patterns." *Journal of the Operational Research Society*, 56(5), 495–503.
43. Tashman, L. J., & Hoover, R. J. (2001). "Comparing Forecasting Methods: Accuracy, Complexity, and Robustness." *Journal of Forecasting*, 20(1), 1–16.
44. Taylor, J. W. (2010). "Exponential Smoothing with a Damped Multiplicative Trend." *International Journal of Forecasting*, 26(4), 627–635.
45. Wheelwright, S. C., & Makridakis, S. (1985). *Forecasting Methods for Management* (4th ed.). John Wiley & Sons.
46. <https://www.landisqyr.eu/about/>
47. Williams, T. M. (2004). "Forecasting in Supply Chains: Demand and Order Forecasting." *International Journal of Operations & Production Management*, 24(6), 476–490.
48. Zhang, A. (2019). Impact of information sharing and forecast combination on fast-moving-consumer-goods demand forecast accuracy. *Information*, 10(8), 256.
<https://doi.org/10.3390/info10080256>
49. Zhang, C., & Dong, Z. Y. (2022). Transactive energy sharing in a microgrid via an enhanced distributed adaptive robust optimization approach. *IEEE Transactions on Smart Grid*, 13(3), 2279–2293.
50. Zhang, Y., & Li, Z. (2019). Comparative analysis of simple and exponential moving averages in supply chain forecasting. *International Journal of Production Research*, 57(12), 1–15.
51. Zhang, Y.-T., Li, M.-Y., & Zhou, W.-X. (2024). Impact of the Russia-Ukraine conflict on the international staple agrifood trade networks. *arXiv preprint arXiv:2403.12496*.

Appendix

List of Python Scripts

2. Python script 1

```
# Prepare the updated data for regression
df_['lnE360'] = np.log(df_['E360'])
df_['lnE450'] = np.log(df_['E450'])
df_['lnFNN'] = np.log(df_['FNN'])

# Generate the time trend variable (t) and its quadratic term (t^2)
df_['t'] = range(1, len(df_) + 1)
df_['t_squared'] = df_['t'] ** 2

# Extract month from the date for seasonal dummy variables
df_updated['Month_Num'] = pd.to_datetime(df_updated['Month']).dt.month

# Create dummy variables for months, excluding November (M11)
month_dummies = pd.get_dummies(df_['Month_Num'], prefix='M', drop_first=False)
month_dummies.drop('M_11', axis=1, inplace=True) # Exclude M11 to avoid multicollinearity

# Combine all variables for regression
X = pd.concat([df[['t', 't_squared', 'lnE450', 'lnFNN']], month_dummies], axis=1)

X = sm.add_constant(X) # Add a constant term for  $\beta_0$ 

# Dependent variable
y = df['lnE360']

# Fit the regression model
model = sm.OLS(y, X).fit()

# Summarize the regression results
regression_summary = model.summary()
```

3. Python Script 2

```
# Preprocess the dataset
df['ln_Y'] = np.log(df['E360'] + df['E450'] + df['FNN']) # Logarithm of the sum of sales for all product families

df['t'] = np.arange(1, len(df) + 1) # Linear time trend
df['t_squared'] = df['t'] ** 2 # Quadratic time trend

# Create monthly dummies
```

```
df['Month'] = pd.to_datetime(df['Month'])
df['Month_Num'] = df['Month'].dt.month
for month in range(1, 13):
    df[f'M{month}'] = (df['Month_Num'] == month).astype(int)
# Drop one dummy variable to avoid multicollinearity (e.g., M11)
df.drop(columns=["M11"], inplace=True)
# Define the independent variables
independent_vars = ['t', 't_squared'] + [f'M{month}' for month in range(1, 13) if month != 11]
X = df[independent_vars]
X = sm.add_constant(X) # Add intercept
# Define the dependent variable
y = df['ln_Y']
# Fit the regression model
model = sm.OLS(y, X).fit()
# Summarize the regression results
model_summary = model.summary()
```

4. Python script 3

```
# Add the COVID-19 dummy variable to the dataset
df['Covid_Dummy'] = ((df['Month'] >= '2019-12-01') & (df['Month'] <= '2022-12-31')).astype(int)
# Update the independent variables to include the COVID-19 dummy
updated_vars_with_covid = ['const', 't_squared', 'M1', 'M2', 'M3', 'M5', 'M7', 'M8', 'M9', 'Covid_Dummy']
X_updated_with_covid = X[updated_vars_with_covid]
X_updated_with_covid['Covid_Dummy'] = df['Covid_Dummy']
# Fit the updated regression model with the COVID-19 dummy
model_with_covid = sm.OLS(y, X_updated_with_covid).fit()
# Summarize the regression results
model_with_covid_summary = model_with_covid.summary()
```

5. Python script 4

```
# Add the Ukraine invasion dummy variable to the dataset
df['Ukraine_Invasion_Dummy'] = (df['Month'] < '2022-02-01').astype(int)
# Update the independent variables to include the Ukraine invasion dummy
```

```
updated_vars_with_ukraine = ['const', 't_squared', 'M1', 'M2', 'M3', 'M5', 'M7', 'M8', 'M9',
'Ukraine_Invasion_Dummy']

X_updated_with_ukraine = X[updated_vars_with_ukraine]

X_updated_with_ukraine['Ukraine_Invasion_Dummy'] = df['Ukraine_Invasion_Dummy']

# Fit the updated regression model with the Ukraine invasion dummy

model_with_ukraine = sm.OLS(y, X_updated_with_ukraine).fit()

# Summarize the regression results

model_with_ukraine_summary = model_with_ukraine.summary()

model_with_ukraine_summary
```

6. Python Script 5

```
# Calculate the split index for 70-30 ratio

split_index = int(len(df) * 0.7)

# Split the data into estimation and forecasting samples

estimation_sample = df.iloc[:split_index]

forecasting_sample = df.iloc[split_index:]

# Function to calculate SMA forecast and errors

def calculate_sma_metrics(data, forecast_data, column, window=3):

# Calculate SMA

data[f'SMA_{column}'] = data[column].rolling(window=window).mean()

# Get the SMA values corresponding to the forecast data

forecast_sma = data[f'SMA_{column}'].iloc[split_index:]

# Align the lengths for MAE and MAPE calculation

actuals = forecast_data[column].values[window - 1:] # Adjust for window offset

forecasts = forecast_sma.values[window - 1:]

# Calculate MAE and MAPE

mae = mean_absolute_error(actuals, forecasts)

mape = (abs((actuals - forecasts) / actuals).mean()) * 100

return mae, mape

# Calculate SMA metrics for each product family

results = {}

for product in ['E360', 'E450', 'FNN']:

mae, mape = calculate_sma_metrics(estimation_sample.copy(), forecasting_sample, product)
```

```
results[product] = {'MAE': mae, 'MAPE': mape}
```

7. Python Script 6

```
def calculate_optimal_alpha(data, split_point=50):
    best_alpha = None
    best_mape = float('inf')
    alphas = np.arange(0.01, 1.0) # Range of alpha values to test
    # Split the data into estimation and forecasting samples
    estimation_sample = data[:split_point]
    forecasting_sample = data[split_point:]
    for alpha in alphas:
        # Compute the forecast
        forecast = [estimation_sample[0]] # Initialize with the first observation
        for i in range(1, len(estimation_sample)):
            forecast.append(alpha * estimation_sample[i - 1] + (1 - alpha) * forecast[-1])
        # Extend the forecast into the forecasting sample
        for i in range(len(forecasting_sample)):
            forecast.append(alpha * forecasting_sample[i] + (1 - alpha) * forecast[-1])
        # Calculate MAPE on the forecasting sample
        mape = mean_absolute_percentage_error(forecasting_sample, forecast[split_point:])
        if mape < best_mape:
            best_mape = mape
            best_alpha = alpha
    return best_alpha
```

8. Python Script 7

```
data = data.sort_values('Month').reset_index(drop=True)
# Add time trend variables
data['t'] = np.arange(1, len(data) + 1)
data['t_squared'] = data['t'] ** 2
# Create monthly dummies (excluding M11 deliberately)
data['Month_Num'] = data['Month'].dt.month
monthly_dummies = pd.get_dummies(data['Month_Num'], prefix='M',
drop_first=False).drop(columns='M_11')
```



```

data = pd.concat([data, monthly_dummies], axis=1)

# Split data into estimation (70%) and forecasting (30%) samples
estimation_size = int(len(data) * 0.7)

estimation_sample = data.iloc[:estimation_size]
forecasting_sample = data.iloc[estimation_size:]

# Define a function to run regression and forecast
def regression_and_forecast(data, product_family):

# Log transform the product family sales
data['ln_Y'] = np.log(data[product_family].replace(0, np.nan)) # Replace 0 with NaN to avoid
issues with log

data = data.dropna(subset=['ln_Y']) # Drop rows where ln_Y couldn't be computed

# Independent variables (including a constant for intercept)
X_columns = ['t', 't_squared'] + [col for col in data.columns if col.startswith('M_')]

X = sm.add_constant(data[X_columns])

Y = data['ln_Y']

# Fit regression model
model = sm.OLS(Y, X).fit()

# Generate predictions for the forecasting sample
forecast_X = sm.add_constant(forecasting_sample[X_columns])

forecasts = model.predict(forecast_X)

# Reverse log transformation for predicted sales
forecasted_sales = np.exp(forecasts)

# Calculate MAE and MAPE
actual_sales = forecasting_sample[product_family].values

mae = np.mean(np.abs(forecasted_sales - actual_sales))

mape = np.mean(np.abs((forecasted_sales - actual_sales) / actual_sales)) * 100

return model, forecasted_sales, mae, mape

# Run the regression and forecast for each product family
results = {} for product in ['E360', 'E450', 'FNN']:

model, forecasted_sales, mae, mape = regression_and_forecast(estimation_sample, product)

results[product] = {'model': model, 'forecasted_sales': forecasted_sales, 'mae': mae, 'mape':
mape

```



```
# Display results to the user
```

```
results_summary = {
```

```
product: {'MAE': res['mae'], 'MAPE': res['mape']} for product, res in results.items() }
```