

MASTER
MODELING, DATA ANALYSIS AND DECISION SUPPORT SYSTEMS

Unveiling Retail Insights with Generative AI

Francisco Pinto Magalhães

M

2024



Unveiling Retail Insights with Generative AI

Francisco Pinto Magalhães

Internship Report

Master in Modeling, Data Analysis and Decision Support Systems

Supervised by

PhD João Manuel Portela da Gama

PhD Bruno Miguel Delindro Veloso

2024

Acknowledgements

First and foremost, I want to express my gratitude to Prof. João Gama, my supervisor, for all of his help and advice throughout the development of this work. I appreciate all of your time, expertise, and kind remarks. I also want to extend my thanks to my co-supervisor, Prof. Bruno Veloso, for his availability and efforts during the project. Your suggestions and ideas were extremely helpful in shaping this work.

A special thanks to my supervisor at Sonae MC, Ana Freitas, for her consistent support and mentorship. I was able to overcome my internship's obstacles thanks to your advice. To all my colleagues in the Advanced Analytics & Insights team at MC, thank you for creating a collaborative and supportive environment. Working with you has been a rewarding experience and unquestionably the finest one could have for a first professional experience.

I am grateful to my family for their continuous support and encouragement. To my mother and stepdad, thank you for always being there for me. Finally, to my closest friends, thank you for your friendship and companionship. Your support has been essential in helping me stay balanced and focused.

Abstract

In today's data-driven world, one of the main challenges faced by companies is the efficient interpretation of large volumes of complex data. Indeed, traditional methods which rely on manual human analysis are highly time-consuming. Thus, this internship report explores the potential of Generative Artificial Intelligence, specifically GPT-4, to transform data interpretation techniques within Sonae MC. The focus is on identifying changes in transactional data from the Continente Loyalty Card using change detection algorithms and leveraging GPT-4 to interpret these changes. Then, the goal is to automate the generation of a communication draft that highlights the detected change and is ready to be sent to the respective category director. By automating this process, the company improves its ability to quickly respond to important shifts in business metrics. The project follows a three-phased approach. The first phase establishes a solid understanding of GPT-4's ability to respond to data queries regarding key business indicators. The second phase involves applying change detection algorithms to four years of historical data. After validating the model's performance, a rolling window was introduced to include recent data dynamically, ensuring the model identifies significant recent change-points. In the third phase, the detected changes are interpreted with the aid of GPT-4, showcasing its ability to provide deeper insights. These insights are incorporated into a template email, highlighting the change and offering actionable insights, which streamlines decision-making at Sonae MC. Overall, the study demonstrates the potential of Generative Artificial Intelligence to transform data interpretation at Sonae MC, paving the way for autonomous, real-time data analysis in the retail sector.

Keywords: Generative Artificial Intelligence, Data Analysis, Retail, Sonae MC, Automation, Change Detection.

Resumo

No mundo atual orientado por dados, um dos principais desafios enfrentados pelas empresas é a interpretação eficiente de grandes volumes de dados complexos. De facto, os métodos tradicionais que dependem da análise manual humana são extremamente demorados. Assim, este relatório de estágio explora o potencial da Inteligência Artificial Generativa, especificamente o GPT-4, em transformar as técnicas de interpretação de dados na Sonae MC. O foco está na identificação de mudanças nos dados transacionais do Cartão Continente e na interpretação dessas mudanças com recurso ao GPT-4. De seguida, o objetivo é automatizar a criação de um rascunho de comunicação que destaque a mudança detectada e que esteja pronto para ser enviado ao diretor de categoria respetivo. Ao automatizar este processo, a empresa melhora a sua capacidade de responder rapidamente a alterações significativas nos indicadores de negócios. O projeto está dividido em três fases. A primeira fase averigua a capacidade do GPT-4 de responder a questões sobre indicadores-chave do negócio. A segunda fase envolve a aplicação de algoritmos de deteção de mudanças a quatro anos de dados históricos. Após validar o desempenho do modelo, foi introduzida uma janela rolante para incluir dados recentes de forma dinâmica, garantindo que o modelo identifica pontos de mudança recentes. Na terceira fase, as mudanças detectadas são interpretadas com o GPT-4, demonstrando a sua capacidade de fornecer *insights* mais profundos. Estes *insights* são incorporados num e-mail predefinido, em que se destaca a mudança e se oferecem *insights* para o negócio, agilizando o processo de tomada de decisão na Sonae MC. No geral, este estudo demonstra o potencial da Inteligência Artificial Generativa para transformar a interpretação de dados na Sonae MC, abrindo o caminho para uma análise de dados autónoma e em tempo real no setor de retalho.

Palavras-chave: Inteligência Artificial Generativa, Análise de dados, Retalho, Sonae MC, Automatização, Deteção de Mudança.

Glossary

AI – Artificial Intelligence

AIC – Akaike Information Criterion

ANNs – Artificial Neural Networks

API – Application Programming Interface

BIC – Bayesian Information Criterion

BinSeg – Binary Segmentation

CNNs – Convolutional Neural Networks

CPW – Cereal Partners Worldwide

CRISP-DM – Cross-Industry Standard Process for Data Mining

CUSUM – Cumulative Sum Control Chart

DL – Deep Learning

ELC – The Estée Lauder Companies Inc.

GANs – Generative Adversarial Networks

GenAI – Generative Artificial Intelligence

GPT – Generative Pre-Trained Transformer

LLMs – Large Language Models

ML – Machine Learning

NLP – Natural Language Processing

Opt – Optimal Partitioning

RBF – Radial Basis Function

RH – Real vs Homologous

RNNs – Recurrent Neural Networks

SIC – Schwarz Information Criterion

YTD – Year to Date

Table of Contents

1. Introduction	1
1.1 Context of the Internship	1
1.1.1 Sonae.....	1
1.1.2 Sonae MC.....	1
1.1.3 Loyalty Program Department.....	1
1.2 Problem Description	2
1.3 Internship Goals.....	3
1.4 Structure of the Internship Report.....	4
2. Literature Review.....	5
2.1 Generative Artificial Intelligence	5
2.1.1 From Artificial Intelligence to Deep Learning.....	5
2.1.2 The Rise of Generative Artificial Intelligence.....	7
2.1.3 Application in Retail.....	10
2.2 Time Series and Change Detection	16
2.2.1 Problem Definition.....	16
2.2.2 Types of Pattern Changes in Sequential Data	17
2.2.3 Methods Overview	18
2.2.4 Criteria for Model Selection	20
3. Methodology	22
3.1 Methodological Framework.....	22
3.1.1 Business Understanding.....	23
4. Project Execution	24
4.1 GenAI Performance Overview	24
4.2 Implementing Change Detection Algorithms	29
4.2.1 Model Selection Process	29
4.2.2 Implementation.....	32
4.2.3 Change Detection Model Results.....	36
4.2.4 Model Implementation for Recent Change-Points	38
4.3 Uncovering Retail Insights	40
4.3.1 Prompting Strategy.....	40
4.3.2 Output of GPT-4.....	41
4.3.3 Final Communication Draft.....	42

5. Conclusion	45
References	48
Appendix	54
Section A - Metrics Over Time	54
Section B - Seasonal Decomposition.....	56
Section C - PELT Results.....	64
Section D - Comparison of PELT vs CUSUM.....	70
Section E – Change Detection on Recent Data	81
Section F - Communication Template	84
Section G - Recommendation Letter.....	85

List of Figures

Figure 1.1 - The Spectrum of Artificial Intelligence	4
Figure 2.1 - Types of Changes	18
Figure 2.2 - Difference between Online and Offline Methods	21
Figure 3.1 - CRISP-DM Process Model.....	22
Figure 4.1 - Prompt Comparision	25
Figure 4.2 -Analysis of Two Prompts	27
Figure 4.3 - Multiplicative Decomposition of “vb_nf” in “Doçaria” (Sweets)	33
Figure 4.4 - Penalty Factor Comparison for "vb_nf" in "Doçaria" (Sweets)	34
Figure 4.5 - Penalty Factor Comparison for "pvp_medio" in "Pão Tradicional" (Traditional Bread)	35
Figure 4.6 - Comparison of PELT and CUSUM Methods.....	36
Figure 4.7 - PELT Results (pen=14)	37
Figure 4.8 - Key Insights Generated by GPT-4.....	41
Figure 4.9 - Final Communication Simulation.....	43

List of Tables

Table 2.1 - Applications and Impacts of GenAI Across Industries9

Table 2.2 - Comparative Overview of GenAI Applications in Four Companies 15

Table 2.3 - Unsupervised Methods for Change Detection 20

Table 4.1 - Descriptions of Key Metrics 24

Table 4.2 - Results of the First Testing Phase..... 26

Table 4.3 - Results of the Second Testing Phase..... 28

Table 4.4 - Description of Metrics 30

Table 4.5 - Criteria Evaluation for Change Detection Methods 32

Table 4.6 - Change Detection on Recent Data 39

1. Introduction

1.1 Context of the Internship

This internship report will be conducted based on a curricular internship at the Advanced Analytics & Insights team, at Sonae Modelo e Continente (MC) from Sonae Group. This internship will occur from February 2024 to July 2024 and will be supervised by the Advanced Analytics & Insights Manager - Ana Freitas.

1.1.1 Sonae

Sonae is a Portuguese multinational business group headquartered in Maia, Portugal. The company was founded in 1959 by Afonso Pinto de Magalhães and has sub-holdings such as Sonae MC, Zeitreel, Sparkfood, BrightPixel, Universo, Worten, Sonae Sierra, and NOS. This portfolio includes businesses in the areas such as retail, financial services, technology, real state and telecommunications. Throughout the years, Sonae has expanded their business and is present across all continents and geographies, totalizing 62 countries (Sonae, 2020).

1.1.2 Sonae MC

Sonae's sub-holding, Sonae MC, is the leading food retailer in Portugal. It was founded in 1985 and its beginning of activity was marked by a significant milestone: launching Portugal's first hypermarket - Continente de Matosinhos (SonaeMC, 2022a). It provides families living in Portugal with a broad range of goods and services through its multi-format retail portfolio. This includes food retail, health, wellness, beauty, and new business growth areas – such as BAGGA, Note! and ZU. In terms of MC's value proposition, this comprises competitive pricing, high-quality fresh products, a distinctive private label brand and a differentiated offer in healthy nutrition. Strategically, MC also offers a strong omnichannel strategy combining modern national retail stores with a leading e-commerce platform, and the largest loyalty program in Portugal, offering benefits and promotions to over 4 million families in Portugal (SonaeMC, 2022b).

1.1.3 Loyalty Program Department

This research was developed in the Continente Loyalty Program Department, which oversees all Continente loyalty card-related operations. This department creates everything from operationalizing marketing ideas to measuring campaign outcomes, forming and maintaining partnerships, and deriving insights from consumer data. Portuguese households are

represented among the roughly 4 million active accounts of the Continente Loyalty Program, which was introduced in 2007 and has been gradually growing its user base. This program presently includes 19 permanent partners from various industries, including food, fuel, health, and fashion, as a result of numerous agreements with both internal and external organizations. Wells, Nole!, ZU, Bagga, Meu Super, Zippy, Mo, and Continente stores are examples of internal partners. Galp gas stations are examples of external partners. Additional instances include Ibersol that oversees SOL, KFC, Ô Kilo, Roulotte, Pans & Company, Pizza Hut, Miit, and Pasta Café. Furthermore, the Continente Loyalty Program has occasionally formed alliances with businesses in various industries, including financial services, transportation, and culture and leisure.

1.2 Problem Description

Nowadays, one of the main challenges many companies face is the effective and efficient interpretation of large amounts of complex data. This is an issue that perpetuates given the common adoption of the traditional methods that rely on manual human analysis for understanding and drawing conclusions from the data – either in reports or dashboards. Sonae MC is no different from the rest and consistently tries to fight this problem.

Even though manual data interpretation is based on experience, it has one key disadvantage associated with it. Specifically, this process is extremely time-consuming, which often leads to delays in decision-making. Given that the retail sector is known for its rapid evolving environment, this delay can result in missed opportunities, and, consequently, in the adoption of reactive strategies, rather than proactive ones. In light of this, timely reactions are critical.

The high complexity and volume of data that analysts are expected to navigate through, further intensify these challenges. Indeed, companies like Sonae MC have numerous categories of products, each with extensive data that needs to be interpreted. Traditional methods are time-consuming, both in collecting and interpreting the data, making the process even more prolonged. This time consumption in the business world is prejudicial, as it can hinder the ability to make quick decisions necessary to stay competitive in a fast-paced market.

On top of that, this type of work is commonly required under pressing time constraints, which accentuates the need for an efficient and advanced data interpretation method. The reliance on conventional, human-centric methods of analysis is no longer suitable for the needs of modern retail.

1.3 Internship Goals

Stepping off from the premise that the traditional methods of data interpretation largely rely on human analysis, which often result in delayed and time-consuming outcomes, the purpose of this internship is to understand whether and how Generative Artificial Intelligence (GenAI) can revolutionize data interpretation techniques in Sonae MC's environment.

Therefore, to address that research question, GenAI will be used, specifically GPT-4 from OpenAI API in Azure, complemented by change detection techniques. The ultimate goal is to identify and interpret changes within transactional data from Continente Loyalty Card, demonstrating the power of GenAI to unveil retail insights.

To achieve that goal, this project will be divided into three phases, allowing for a deeper and more comprehensive approach. In the first phase, the goal is to assess the quality of GPT-4 on communicating business key performance indicators. Therefore, the focus consists on getting an overview of how GPT-4 behaves when confronted with questions regarding data from a dashboard. This phase is strategically designed to study the potential of GPT-4 to interpret key business indicators, potentially expanding the scope of specialized analyses within Sonae MC. In fact, this capability would enable near real-time communication to each category director regarding their respective categories, significantly reducing the time-consuming nature of interpreting data. This way, weekly reports could be sent, informing the state of the main key business indicators of a certain category. However, the risk of repetitive insights (e.g., reporting the same metric increase every week) creates a necessity of introducing change detection algorithms. These methods ensure that communication occurs only when there is a meaningful change detected.

This way, the second phase of the project expands beyond key business indicators to include a comprehensive dashboard with numerous metrics. This phase focuses on communicating new trends or changes in data through a thorough analysis of long-term data spanning four complete years (2020 until 2023). The initial step of this phase is to graphically represent the data, using time-series plots. Then, change detection techniques will be applied as planned, identifying the change-points that should be communicated.

Lastly, the third phase consists of combining GenAI's power with the results obtained with the change detection algorithms. Indeed, by complementing the obtained results with GenAI, the goal is to interpret the detected changes, showcasing the value of GPT-4. Once again, this will be executed recurring to the OpenAI API on Azure, which will be embedded

with base knowledge - some internal documentation so that the usual language of Sonae MC is reproduced by the model. As a final step of this phase, a simulation of a complete communication report based on a real change-point will be done, incorporating the GenAI output. This report will be prepared as if ready to be sent to a category director.

1.4 Structure of the Internship Report

This internship report is broken up into three sections. The setting of the internship, a description of the problem, and its primary goals are covered in this first portion, Chapter 1. An overview of the literature is given in Chapter 2. The terms Artificial Intelligence, Machine Learning, Deep Learning, and Generative Artificial Intelligence are introduced in this section. As seen in Figure 1.1, this structure was selected since the themes are all related to one another and build upon one another. In the latter section, the implications of GenAI's applicability to retail will be examined. To further consolidate the topic, real world examples are also described. Considering the described problem, this chapter also overviews “Time Series and Change Detection”, with an explanation of the several available methods and the criteria employed for their selection. In Chapter 3, the chosen methodology, CRISP-DM, is introduced and explained. Chapter 4 covers the project execution, with a detailed description of each project phase and their results. Finally, Chapter 5 provides the conclusion, discusses the limitations of the work, and offers suggestions for future research.

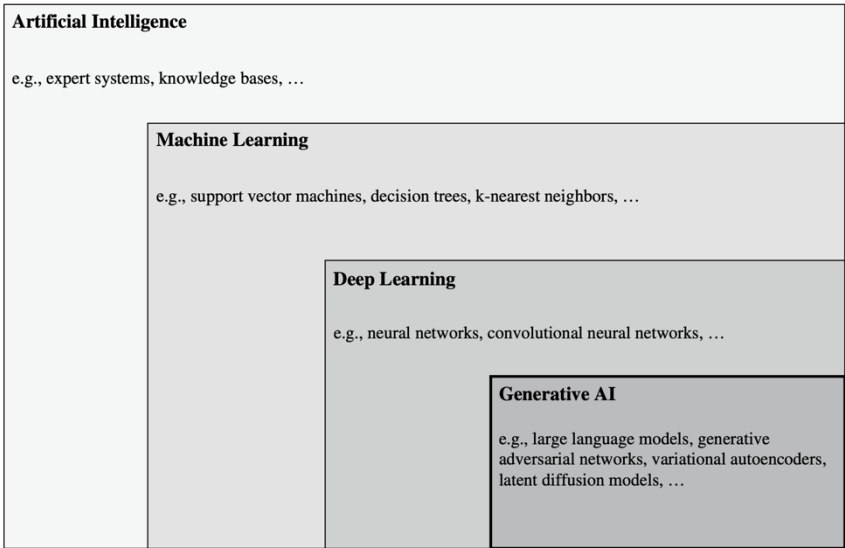


Figure 1.1 - The Spectrum of Artificial Intelligence

Source: Banh and Strobel (2023), p. 2.

2. Literature Review

2.1 Generative Artificial Intelligence

The emergence of Generative Artificial Intelligence stands as proof of the continuous evolution of technology in the realm of artificial intelligence. Delving into this new era, understanding the distinctions between its foundational concepts becomes paramount. Therefore, to navigate the full landscape of GenAI, considering the broader spectrum of Artificial Intelligence is essential.

2.1.1 From Artificial Intelligence to Deep Learning

Artificial Intelligence (AI), a concept that is often represented in science fiction movies, is now a reality in several aspects of the human daily life and in businesses. Indeed, this revolutionary technology has an impact on numerous sectors, including healthcare, finance, manufacturing, retail, and governments worldwide. AI's power to complement human tasks is revolutionary and opens new opportunities for innovation (Dwivedi et al., 2021).

Over the years, the concept of AI has been mentioned by several authors, but there is yet to come a consensual definition. These definitions, however, seem to always include the main concepts of “non-human intelligence” that have been programmed to develop specific tasks. In the words of Russell and Norvig (2016), AI mimics cognitive functions which are typically associated with the human intellect: “learning”, “speech” and “problem solving”. A more thorough explanation of AI is provided by Kaplan and Haelein (2019), who define it as a system's capacity to independently analyze and learn from external data. The authors emphasize that AI can achieve specific objectives through flexible adaptation, with the help of such learnings.

AI is increasingly being used in domains that have traditionally been dominated by human intelligence. This is due to AI's ability to aid in overcoming several challenges by expanding human capacities. AI's prominent applicability is noted by Russell and Norvig (2016) who explain how it can be applied in multiple subfields. Some examples of its applications are speech recognition, spam fighting and robotic vehicles (Russell & Norvig, 2016). In the business landscape, Miller (2018) states that a new human-machine symbiosis is “imperative”. The author underlines that this junction will allow organizations to increase their growth and productivity.

According to many researchers, AI can be seen as an umbrella term, with Machine Learning (ML) being a subset of it. ML is defined as the ability of a computer program to improve its own performance - in a certain class of tasks - based on experience (Jordan & Mitchell, 2015). For that reason, ML has become a critical aspect of intelligent systems, learning from problem-specific training data and automatizing the analytical model-building process (Janiesch et al., 2021). In line with Arthur Samuel, famous for the development of a checkers playing program, ML enables systems to learn without “explicit programming”, leading to the discover of hidden insights and complex patterns (Shrestha et al., 2021).

ML’s ability to solve high-dimensional tasks has led to its implementation on areas like fraud detection, speech recognition and Natural Language Processing (NLP). Indeed, machine learning has significantly evolved, becoming a part of various industries (Janiesch et al., 2021).

There are three different types of ML: supervised learning, unsupervised learning, and reinforcement learning. These approaches are selected according to the nature of the data and the preferred result. Based on Janiesch et al. (2021), in supervised learning it is necessary to train a model using data that includes both input features as well as corresponding labeled outputs. From these input-output pairs, the model is able to learn how to predict outputs for new data. There are two types of supervised learning which can be employed. Classification is chosen to predict categorical classes while regression is opted for in cases where the aim is to predict numerical values. On the other hand, unsupervised learning is the process of finding patterns in data without pre-existing labels. It is commonly done through the identification of clusters, dimensionality reduction and the creation of association rules. Lastly, reinforcement learning implies that the system learns by interacting with its environment, through trial and error. The goal is to learn optimal decision-making, by maximizing cumulative rewards over time. Logistic regression, Support vector machines, Artificial Neural Networks (ANNs) and K-nearest Neighbors are some examples of ML algorithms (Janiesch et al., 2021; Ray, 2019).

A significant subset of ML is Deep Learning (DL) which has its foundations in Artificial Neural Networks. These computational models are inspired by the human’s brain structure and function – although in a highly simplified form. Indeed, in deep learning systems, multiple hidden layers of interconnected neurons process information in a similar way to biological neural networks (Janiesch et al., 2021). The term “deep” itself is exactly due to the existence of multiple processing layers, which enable complex data transformations and feature learning. In other words, as data passes through these layers, the network learns to

recognize features and patterns (Sarker, 2021). It is this complex architecture that allows deep learning models to handle high-dimensional data, in a wide range of domains. LeCun et al. (2015) exposes as examples DL's capability to process images, videos, and audios.

In terms of its application, there are diverse fields that can benefit from implementing such models. For example, DL can be used on the creation of recommendation systems, as studied by Li et al. (2022). The authors explained the core concept of “within-basket recommendation”, to predict items to add to a consumer's basket – in the grocery and e-commerce landscapes. Usually, these problems are addressed using rule mining and association rules, but in this study a deep learning model was implemented to approach this task (Li et al., 2022). Still in the realm of e-commerce, DL's usage to detect fake consumer reviews is beneficial to guarantee authenticity and consumer trust (Zhang et al., 2023). Previous research on the topic mainly explores techniques that usually demand a lot of human effort and specialized knowledge, so Zhang et. al (2023) proposed a novel deep learning approach. This particular approach focuses on analyzing both the behavior and the textual content of the reviews, automatizing the process of detecting fake reviewers. Its accuracy was validated, and the results obtained drastically increased the performance in the detection of such reviewers.

2.1.2 The Rise of Generative Artificial Intelligence

Due to the advancements in Deep Learning techniques, deep generative models (DGMs) have emerged as a class of DL models, with the ability to generate new content based on existing data. These models end up creating a variety of new possibilities for AI applications (Lehmann & Buschek, 2020; Tomczak, 2022). This progress marked a significant milestone in the evolution of Generative Artificial Intelligence (GenAI), particularly with the appearance of Generative Adversarial Networks (GANs), firstly introduced by Goodfellow et al. (2014). This deep generative model consists of two parts: a generator and a discriminator. While the generator creates samples with the goal of mimicking a certain target data distribution, the discriminator evaluates these samples. The goal of the discriminator is to determine if these are real or generated. As their adversarial interaction continues, both improve their performance. The generator learns how to produce increasingly realistic data while the discriminator becomes better at detecting genuine data (Feuerriegel et al., 2023). Facial images and noise maps were some of the outputs created by GANs (Creswell et al., 2018).

Another landmark in GenAI was the introduction of the Transformer framework by Vaswani et al. (2017). Indeed, the Transformer architecture was an important advance in modern machine learning by shifting the paradigm from the Recurrent Neural Networks (RNNs) and Convolutional Neural Networks (CNNs). Even though it was initially aimed to machine translation (Vaswani et al., 2017), this design settled the foundations for several Large Language Models (LLMs), specifically, in the field of NLP. Examples are BERT, developed by Google, and the Generative Pre-Trained Transformer series (GPT, GPT-2, GPT-3, GPT-4) (Ray, 2023). As the term denotes, these are the models that OpenAI's ChatGPT was based on, specifically GPT-3 and its successors.

The fact that the model GPT-3 uses 175 billion parameters, and is trained on diverse datasets, exemplifies the potential of Generative AI in mirroring human-like interaction. Its capability to process and generate content such as text, images, and audio marks a significant shift in how machines understand and interact with data (Radford et al., 2018; Uc-Cetina et al., 2022). Additionally, the introduction of GPT-4 represents another major milestone in OpenAI's trajectory. Even though the official number of parameters has not been revealed, it is rumored to be significantly larger, building upon the successes of its predecessor, GPT-3. Indeed, GPT-4 is capable of exhibiting human-level performance in several cases, surpassing GPT-3, mainly because of its enhanced capability to process instructions. For instance, in a study made using simulated bar exams, GPT-4 consistently got scores in the top 10% of test takers, while GPT-3.5 typically scored in the bottom 10%. (OpenAI, 2023).

Moreover, the fact that GPT-4 has multimodal capabilities, meaning that it can accept both text and image inputs, expands its utility even more and across several tasks. However, as its predecessor, it also comes with limitations. Some examples are the fact that it may exhibit biases, hallucinate facts, or make reasoning errors (OpenAI, 2023).

Other technologies like DALL-E were also funded according to the Transformer framework. In this case, DALL-E is a “12-billion parameter version of GPT-3” (OpenAI, 2021) that is designed to create images from textual prompts. In fact, users just need to input descriptions and then the model generates images according to the requests.

One important note regarding the use of GenAI is the fact that effective prompting highly influences the obtained outcomes. To get around that, “Prompt Engineering” emerges as the process of designing prompts that accurately instruct LLMs how to elaborate highly relevant answers to the users' questions. Therefore, it can be concluded that prompt engineering maximizes the potential of these tools, leading to outputs with higher quality

(Wang et. al, 2024). For this process, OpenAI recommends several tactics such as writing clear instructions and specifying the steps necessary to complete the task. Also, asking the model to adopt a specific persona is also an important suggestion. The general idea is that the less the model has to guess, the better are the outputs it generates (OpenAI, n.d.).

These technologies not only have applicability in the creative field but also in aiding in tasks ranging from IT help desks to providing medical advice. Accordingly, Table 2.1 lists the variety of industries throughout which GenAI has been used, along with its opportunities and impacts.

Industry	Opportunities and Impact
Education	Supporting modern teaching-learning methods; Fostering practice-based training; Automating tasks like quiz checking and scoring.
Healthcare	Boosting digital health initiatives; Supporting medical staff in remote areas; Reducing learning cycle time for healthcare training.
Banking, Financial Services, and Insurance	Replacing conventional tasks such as customer support, fundamental financial analysis, and sentiment analysis in text data.
Hospitality and Tourism	Serving as an interactive interface for customers on websites, improving user experience, and reducing staffing costs.
Research & Development	Enhancing R&D pace; Supporting coding and debugging; Reducing duplication of R&D activities.
Legal Services	Providing advisory and support by text-mining legal databases, increasing productivity for individuals, regulators, and law firms.

Table 2.1 - Applications and Impacts of GenAI Across Industries

Source: Adapted from Dwivedi et al. (2023).

Concerns about content authorship and authenticity underline how important it is to handle this technology responsibly (Else, 2023; Stokel-Walker, 2023). In addition to these worries, there are other social, professional, and personal issues.

GenAI has begun to be utilized in the creation of political content, including deepfakes and fake news, which could diminish people's grasp of the truth and promote polarization. Such an outcome would be detrimental to political stability and democracy (Danaher & Sætra, 2022). On the other hand, GenAI could enhance democratic processes, but the risk of diluting the essence of democracy is an additional concern as underlined by Sætra, Borgebund, and Coeckelbergh (2022). In addition, these models may unintentionally produce inaccurate or misleading data that could be used to disseminate incorrect information. This is a serious problem in a time when misinformation and fake news can have negative real-

world repercussions (Patil et al., 2023). Another issue is the fact that this technology has the ability to alter work conditions and might even displace workers in companies. At a meso level, technological changes have impacted professions (Barley, 2020), but GenAI will not only transform them but also change power relationships between the different groups, which can potentially result in the redundancy of certain skills (Sætra, 2023). Moreover, the development and use of GenAI models, which require a considerable number of computational resources, may lead companies to be highly dependent on large tech companies (Widder et al., 2023; Patil et al., 2023).

Furthermore, GenAI models are based on historical data, which might hold back desired societal changes. Precisely, historical data is usually imbued with biases so systems based on that data will have the tendency to replicate those biases, disproportionately affecting marginalized groups (Bender et al., 2021; Patil et al., 2023). Other issues concerning the environmental landscape arise as these models consistently consume large amounts of energy, leading to a larger carbon footprint (Brevini, 2021; Patil et al., 2023). Simultaneously, by training these models with human-generated content without obtaining consent or offering some type of recompense raises significant legal and ethical issues (Sætra, 2023).

On the micro level, the use of Generative AI in complex mental tasks might contribute to cognitive atrophy, similar to the impact that calculators had on people's arithmetic skills (Sætra, 2019). Lastly, Sætra and Mills (2022) suggest as another point, the increasing ability of GenAI to persuade and potentially manipulate human behaviors and perceptions.

2.1.3 Application in Retail

In the early stages of e-commerce – the activity of buying and selling goods and services online – several companies expanded their business into the online realm in hopes of responding to the increased digitalization among consumers. Now, those days are long past, and recent advancements have allowed for further automation and innovation in retail (Ooi et al., 2023).

Integrating GenAI technologies enables the revolutionization of the retail industry, an industry with a prominent competitive landscape. As market dynamics become increasingly challenging, retailers are adopting GenAI to enhance customer engagement and streamline internal operations, thus gaining a crucial competitive advantage (Iskender, 2023).

Therefore, GenAI is the way to unlock new potentials in several areas of retail. For those retailers who seek to innovate and lead in their respective markets, this technology is

essential. Notably, in light of prevailing competition, its continued implementation will be an absolute prerequisite for the future success of retailing (Kumar, 2023).

A study by Google Cloud (2023a) outlines a summary of how GenAI is shaping the retail landscape through consistent results fashioned by innovative applications as well as customer engagement strategies. The applications are spread throughout the various facets of the retail industry, enabling retailers to change the way they operate as well as their connection with consumers.

One of the main use cases of GenAI in retail is in Creative Assistance. Indeed, this technology can be employed with the goal to empower creative teams by helping on the creation of custom-made images and creative content for marketing purposes. This way, new stages of innovation will be reached, and retailers will be able to deliver extremely engaging and 1:1 personalized content (Eschliman, 2023). As a result, the levels of customer engagement and conversion rates will be enhanced. Another benefit is that by using such technology, not only do retailers save time but also cut some costs (Google Cloud, 2023a).

Another possible utilization of GenAI is in Conversational Commerce by creating virtual bots that engage with consumers and help them find what they wish. In fact, this technology has the ability to provide real-time personalized product recommendations, influencing the customers' behavior and purchasing decisions. For instance, in a clothing store, GenAI-powered virtual stylists can be designed to recommend clothes that match the customers' style and size, while simultaneously showing pictures of digital influencers to provide more style inspiration (Eschliman, 2023). This is an excellent approach to solve concerns such as search abandonment (Google Cloud, 2023a).

In terms of Product Catalog Management, GenAI transforms the way retailers create and sustain their product catalogs. Indeed, with its utilization, simple textual prompts can be turned into high-quality images, the classification of the products into their category can be accelerated and appealing product descriptions that are focused on meeting the needs and interests of customers can be quickly written. These tasks are usually highly time-consuming and often affect the ease at which a customer finds the product they are looking for. In this manner, GenAI can enhance catalog management, leading to an increased likelihood of consumer purchases (Google Cloud, 2023a).

Regarding the business process of New Product Development, GenAI helps improve the existing consumer research efforts. It simplifies query processing and insight generation. Other than that, it helps generate product and package design concepts. With data-driven

insights, GenAI enables companies to concept test and refine product designs at a faster pace (Eschliman, 2023). This application accelerates the innovation cycle, cuts costs, and helps to identify winning product concepts more quickly (Google Cloud, 2023a).

Lastly, GenAI can be used in Customer Service Automation to assist customer service agents and quickly help customers find the answers they need. Undoubtedly, GenAI has the ability to support these teams in several tasks (Eschliman, 2023). As an example, virtual bots could be designed to generate summaries of previous customer enquiries and create a list of FAQs (Frequently Asked Questions) to update the website. A more complex example would be virtual bots capable of offering real-time assistance, therefore alleviating the workload in contact centers. In general, GenAI appears as an extremely valuable addition to customer service strategies (Google Cloud, 2023a).

2.1.3.1 Real-World Examples

In the quick-service restaurant industry, *Wendy's* stands out with the implementation of GenAI to automate its drive-through service. Due to a natural-language software, which was developed by Google and trained to understand the several possible order combinations made by the customers, *Wendy's* contributes to the automatization of customer service. Considering that most of this company's business is conducted over drive-through, improving this area represented an important challenge. Thus, the design of this GenAI system considered the need to efficiently handle complex menu options, as well as special requests, while still bearing in mind the possibility of ambient noise. However, the development of this new GenAI-powered chatbot was not an easy task and the main challenge was that it had to handle nuances in customer languages, including slang and regional variations (Google Cloud, 2023b). For instance, when ordering a "Frosty", customers might refer to it as a "chocolate shake" or even a "frappe", so the system must be able to understand what the client means regardless of the terminology used (Chaban, 2023).

Regarding goals, the main objective of implementing such technology is to improve the experience for clients, employees, and to the company itself. In truth, one can notice that, in this case, GenAI is used to streamline operations by simplifying the ordering process, which allows employees to focus more on delivering high-quality meals. Consequently, the overall customer service experience is enhanced (Google Cloud, 2023b).

Currently, this technology is being tested in a Columbus restaurant's drive-through and given the inclusion of certain advanced AI offerings, specifically, "Vertex AI Search and

Conversation”¹, the system is able to have full conversations with customers, understand their orders and answer to some questions that they might have. Such advancements are powered by Google’s foundational LLMs, which use data from *Wendy’s* menu, and incorporate predefined rules and logical frameworks to effectively guide its interactions with coherence, relevance and appropriateness. In addition, these LLMs are integrated with the physical hardware present in the restaurant as well as the “Point of Sale” (POS) system (Chaban, M., 2023; Google Cloud, 2023b). Thus, GenAI allows the company to differentiate itself from other companies in the restaurant industry, while streamlining its overall operations (Google Cloud, 2023b).

In the beauty industry, the American multinational company *The Estée Lauder Companies Inc. (ELC)* partners with Google Cloud to expand their retail strategy to include novel uses of GenAI. By doing so, the company aims to improve their customer experience and enhance operational efficiency. On a macro level, the goal is to revolutionize the luxury digital experience (Google Cloud, 2023c).

The first approach to GenAI that *ELC* adopted was in terms of personalization. Indeed, the company aimed to create personalized interactions on its brand websites, to match the quality and personalization that customers are offered in physical stores. Additionally, through the power of GenAI’s capabilities, *ELC* intends to monitor consumer sentiment and feedback in real-time. Particularly, the company uses PaLM 2, Google’s LLM, to deeply understand how the consumer feels on the various communication platforms. Immediately analyzing the sentiment expressed by a customer, allows *ELC*’s brands to quickly attend to their consumer needs and preferences. Consequently, the overall quality of the customer experience is enhanced. Moreover, new GenAI business applications are under development on Vertex AI in hopes of streamlining even more operations and simplify complex tasks. By doing so, lower operating costs and higher productivity are expected (Google Cloud, 2023c).

Like *ELC*, the intimate apparel company, *Victoria’s Secret & Co.*, teamed up with Google Cloud in order to use GenAI technologies that could add value to the online shopping experience and strengthen the relationship between the brand and the customers. Indeed, according to Chris Rupp, Chief Customer Officer at *Victoria’s Secret & Co.*, the company resorted to such technologies to develop a useful Virtual Assistant, powered by “Vertex AI”.

¹ Google Cloud feature that allows developers to build advanced AI-applications with conversational AI and enhanced search capabilities (Google Cloud, 2024).

As this technology is currently being designed, its main goal is to provide online customers with the same amount of personalisation and attention they would receive if they visited one of the physical stores. In addition, according to the company's vision, the chatbot is projected to provide personalized tips, as well as guidance and recommendations, thereby enhancing the shopping experience of those who wish to purchase via the brand's website or mobile application (Rupp, 2024).

In order to also improve internal corporate operations, *Victoria's Secret & Co.* particularly took advantage of GenAI-powered technologies that could assist the company's Human Resources department. The automation of job descriptions, immediate transcriptions of interview notes, and assistance with the onboarding of new collaborators represent some of the applications of GenAI that this retailer has benefited from (Rupp, 2024).

Cereal Partners Worldwide (CPW) and their use of the AnswerRocket ² GenAI Analytics Platform provide a last illustration of how GenAI can be used in retail. The company aimed to reduce the amount of time spent on data analysis to obtain competitive advantage. Indeed, due to the adoption of manual data interpretation processes, data analysis was extremely time-consuming. Consequently, the decision-making process within the company would often be adversely affected. By utilizing the AnswerRocket Platform, CPW was able to democratize data access across the organization, transforming data into a tool that other team members could use as well. The first opportunity resulting from these advancements is that teams may now communicate and query data directly within the platform, using natural language, instead of depending on analyst support or extracting information from a spreadsheet. With the automation of the analysis, CPW can now respond quickly to challenges or changes in the market, while being sure that all its decisions are based on data. Such an advantage is crucial in the quickly changing retail sector (AnswerRocket, 2023).

Table 2.2, as seen below, demonstrates the versatility of this technology by summarizing the ways in which these four businesses employ GenAI.

² Innovative Data Intelligence firm that transforms analytics with AI-driven solutions (AnswerRocket, n.d.).

Company	Industry Sector	GenAI Application	Key Features	Technology Used	Objective	Benefits	Author(s), Year
Wendy's	Fast food	Drive-through order automation	Conversational AI, custom order understanding, ambient noise adaptation	Google's foundational LLMs, Vertex AI Search and Conversation	Enhance customer and employee experience, improve service efficiency	Streamlined ordering process, improved customer service	Google Cloud (2023b); Chaban (2023)
The Estée Lauder Companies	Beauty and cosmetics	Personalization, real-time sentiment analysis	Personalized digital experiences, Sentiment analysis	PaLM 2, Vertex AI	Enhance online customer experience, monitor sentiment, streamline operations	Improved consumer insights, personalization, operational efficiency	Google Cloud (2023c)
Victoria's Secret & Co.	Apparel (intimate wear)	Online customer experience personalization, HR operations	Virtual Assistant for personalized service, HR process automation	Vertex AI	Personalization of online shopping experience, enhance customer service	Improved shopping experience, streamlined HR management	Rupp (2024)
CPW	Food manufacturing	Data democratization, real-time analysis	Natural language interaction with data, democratization of data analysis	AnswerRocket GenAI Analytics Platform	Speed up data analysis, facilitate data-driven decision-making	Faster decision-making, real-time market adaptability	AnswerRocket (2023)

Table 2.2 - Comparative Overview of GenAI Applications in Four Companies

2.2 Time Series and Change Detection

Over the past years, time series analysis has become an important tool across several fields. According to Aminikhanghahi and Cook (2017) time series data can be defined as “sequences of measurements” which describe the behavior of systems. Some external events or even internal systematic changes in dynamics or distribution can be the reason behind the changes of behavior observed over time (Montanez et al., 2015). Considering this dynamic nature of time series data, it is necessary to identify changes, which are defined as shifts from one condition to another over time or space (Roddick et al., 2000).

Change detection in time series data refers to the process used to identify changes in the underlying data generating model and has its roots in Page's original work concerning statistical process and quality control (Bhatt et al., 2022). It emerges as a widely applicable concept across various domains. Indeed, in healthcare, it plays a vital role in identifying emerging health issues by detecting changes in patient medical data, enabling early intervention and diagnosis (Liu et al., 2018). Similarly, in manufacturing, it assists in predicting potential failures in machinery by detecting changes in signal patterns. This way, operational efficiency is ensured, as well as, minimized downtime (Guo et al., 2019). In the financial sector, it aids in fraud detection by identifying changes in financial transactions, thereby safeguarding financial systems, and preventing monetary losses (Pourhabibi et al., 2020; Pástor & Stambaugh, 2001).

Moreover, when considering smart home systems, change detection optimizes energy efficiency and enhances comfort levels by detecting changes in home sensor data. By detecting such changes, automated adjustments become a possibility (Aminikhanghahi et al., 2018). When doing image analysis, it is also possible to formulate the detection of abrupt occurrences, including security breaches, as a change-point problem. In this case, the digital encoding of an image represents the observation at each time point (Radke et al., 2005). Lastly, in network security, it helps in identifying changes in network traffic patterns, which enhances the resilience of network systems and safeguards sensitive information (Li et al., 2019).

2.2.1 Problem Definition

The change detection problem is delineated as follows. Initially, defining a data stream as a potentially infinite sequence of elements, denoted as S , where each element comprises a pair (X_j, T_j) . Here, X_j represents a d -dimensional vector that arrives at time T_j (Tran, 2013), as denoted in Equation (1).

$$S = \{(X_1, T_1), \dots, (X_j, T_j)\} \quad (1)$$

Change detection can be reduced to the following problem of hypothesis testing (2) since it essentially entails finding differences in the state of an object or phenomenon observed at different periods and/or locations in space (Tran, 2013):

$$\begin{cases} \mathcal{H}_0 : & \neg \text{change} \\ \mathcal{H}_1 : & \text{change} \end{cases} \quad (2)$$

For a visual representation, Figure 2.1 below envisions a scenario where vectors originate initially from a data source labeled $S1$, later substituted by another data source $S2$. This figure perfectly underscores the primary aim of change detection: to precisely determine the instance when the transition occurred from $S1$ to $S2$ (Faithfull, 2018).

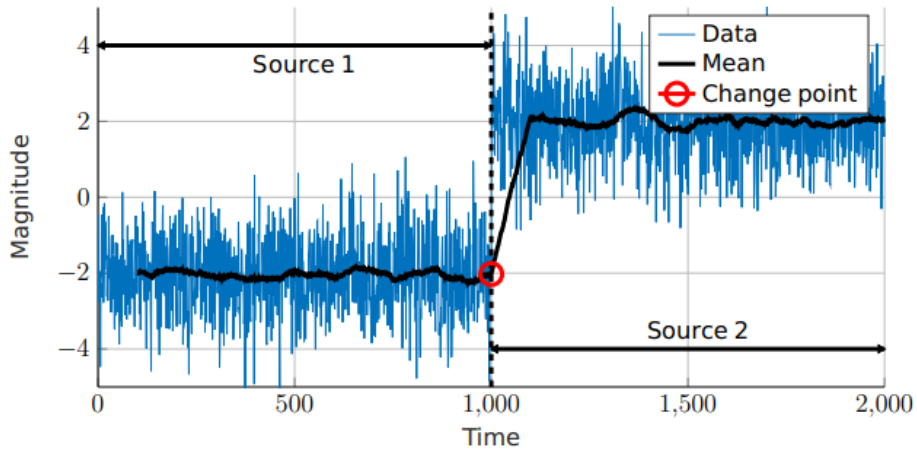


Figure 2.1 - Data Stream with Change Detection

Source: Faithfull (2018).

2.2.2 Types of Pattern Changes in Sequential Data

Time series data can alter in pattern due to the multiple causes that affect its statistical properties (Aminikhanghahi et al., 2018). Truong et al. (2020) explain that a shift in the mean is indicative of a change in the central tendency of the data. It is one of the most common alterations in sequential data, typically caused by a shift in the underlying trend or an external shock to the system (Gupta et al., 2024). Variance change is the term used to describe alterations to the distribution or scattering of data. For instance, in financial data these alterations are typically caused by shifts in the economy or market sentiment. A change in mean-variance (MV) denotes changes in the data's dispersion and central tendency at the same time. Changes of this kind occur often in climatic time series data when the returns' volatility and magnitude vary over time (Gupta et al., 2024). Auto-correlation changes reveal changes in the dependence structure of the data. Lastly, modifications

in the data's periodicity, commonly seen in seasonal data, happen when underlying factor changes cause the cycle frequency to fluctuate. To find these shifts, methods like wavelet analysis and spectrum analysis are used (Gupta et al., 2024).

Figure 2.2 below illustrates how these various types of changes in time series data are represented graphically, for a better understanding of their impact on data patterns and statistical features.

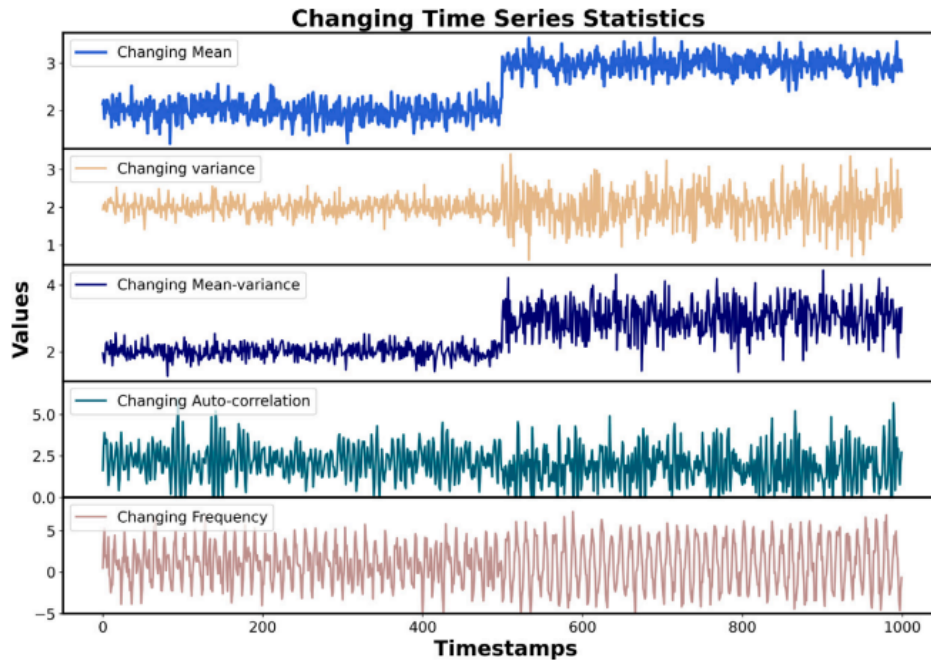


Figure 2.1 - Types of Changes

Source: Gupta et al. (2024).

2.2.3 Methods Overview

As suggested by Aminikhanghahi and Cook (2017), change detection methods can be divided into supervised and unsupervised. First, supervised methods include machine learning algorithms that create models that predict outputs based on labeled input data. These algorithms solve classification problems where instances are classified into predefined categories based on their similarities (Huang, 2015). In this context, supervised methods train ML algorithms as binary or multi-class classifiers (Ahmed et al., 2017). Multi-class classifiers are trained to identify each state boundary when the number of states is specified. While these methods require large amounts of training data to represent all classes accurately, they offer insights not only into the timing but also about the nature and duration of the change. Some examples of multi-class methods are SVM, Naïve Bayes, Bayesian Net, Hidden Markov Model, Conditional Random Field (Reddy et al., 2010; Zheng et al., 2008). Alternatively, change detection can be viewed as a binary classification problem, distinguishing between possible change points and within-state sequences. In this case,

support vector machine, Naïve Bayes, and Logistic regression models are usually implemented (Feuz et al., 2014).

Contrary to the supervised methods, the unsupervised methods are utilized to uncover hidden patterns in unlabelled data. Following the framework considered by Aminikhanghahi and Cook (2017), the authors divided these methods into six different categories: “likelihood ratio”, “subspace models”, “probabilistic methods”, “kernel-based methods”, “graph-based methods” and “clustering”.

The first category analyzes the probability distributions of data before and after a potential change point. If the two distributions differ statistically that data point can be identified as a change-point. Highly associated with the area of control theory, “subspace models”, detect changes by analyzing the structure of subspaces in time series sequences. The “probabilistic methods” focus on the data that has been observed after the last potential change-point and estimate the probability distributions of the new interval, while “kernel-based methods” transform data onto a higher dimensional space using kernel functions and try to detect change-points by comparing each subsequence’s homogeneity. “Graph-based methods” are a more recent approach in the field of change detection. In these methods, the time series gets plotted, and changes are detected based on structural changes in the graph itself, according to statistical techniques. Lastly, when “clustering” methods are utilized, groups of similar data are created and the changes are detected by identifying transitions between clusters (Aminikhanghahi and Cook, 2017).

In addition, exact search-based methods are also important techniques that can be used for change detection. These methods repeatedly examine the possible segmentations of the data in order to identify points where major changes happen. In fact, the way these methods work is by optimizing an objective function (cost function) that basically evaluates how well the segmentation fits the data. This mathematical function measures the "homogeneity" of segments within a signal (Truong et al., 2020). Truong et al. (2020) also approach approximate methods as change detection models to be employed when the computational complexity of the exact methods is too elevated for practical applications. These methods provide a single change point estimate at each iteration and can be used to solve both fixed and variable number of change points problems approximately.

Table 2.3 provides an overview of the unsupervised techniques that may be used for the change-point identification problem according to the authors Aminikhanghahi and Cook (2017) and Truong et al. (2020).

Category	Method
Likelihood Ratio	Cumulative Sum Control Chart (CUSUM); Autoregressive (AR); Kullback-Leibler Importance Estimation Procedure (KLIEP); Unconstrained Least-Squares Importance Fitting (uLSIF); Relative Unconstrained Least-Squares Importance Fitting (RuLSIF); Self-Predicting Log Likelihood (SPLL)
Subspace Models	Subspace Identification (SI); Singular Spectrum Transformation (SST)
Probabilistic Methods	Bayesian; Gaussian Processes (GP)
Kernel Based Methods	Kernel Change Point Analysis (KcpA)
Clustering	Sliding Window and Bottom-Up (SWAB); Minimum Description Length (MDL); Shapelet; Model Fitting
Graph Based Methods	-
Exact Search Methods	Pruned Exact Linear Time (PELT); Optimal Partitioning (Opt)
Approximate Search Methods	Window-based (Win); Binary Segmentation (BinSeg); Bottom-Up (BotUp)

Table 2.3 - Unsupervised Methods for Change Detection

Source: Aminikhanghahi and Cook (2017); Truong et al. (2020).

2.2.4 Criteria for Model Selection

The practical application of change detection involves choosing one of the reviewed methods according to some criteria. Aminikhanghahi and Cook (2017) suggest looking at three criteria to make this selection.

2.2.4.1 Online vs Offline

Change detection algorithms are traditionally classified as “online” or “offline”. Offline algorithms are usually used for retrospective analysis, requiring the entire dataset to identify changes. Their conventional procedure consists of data modeling and detecting change by using the model’s residuals. The time interval that separates an algorithm's predicted change from the actual change's occurrence is known as the Timestamp Delay. First half of Figure 2.3 shows how offline algorithms can detect changes in behavior in a time series very quickly and with little latency. Unfortunately, in order to effectively detect changes, these models frequently need to be run through numerous iterations on the full dataset (Gupta et al., 2024).

On the other hand, online algorithms are designed for real-time data processing. The goal is to detect a change-point as soon as it happens - before the next change-point happens. To do so, these methods process each data point as it becomes available (Aminikhanghahi and Cook 2017). A major challenge for these methods is to detect changes, while quickly reducing the false positives.

As shown in the second half of Figure 2.3, the Timestamp Delay exists between the actual and predicted changes, inherent to the real-time processing nature of the algorithm. Hence, a primary goal for online detection algorithms is to minimize this delay period, while making sure that the detection of changes is made accurately (Gupta et al., 2024). In practice, none of these algorithms operates in perfect real-time as it must examine new data before determining if a change point occurred between the old and new data points (Aminikhanghahi and Cook, 2017).

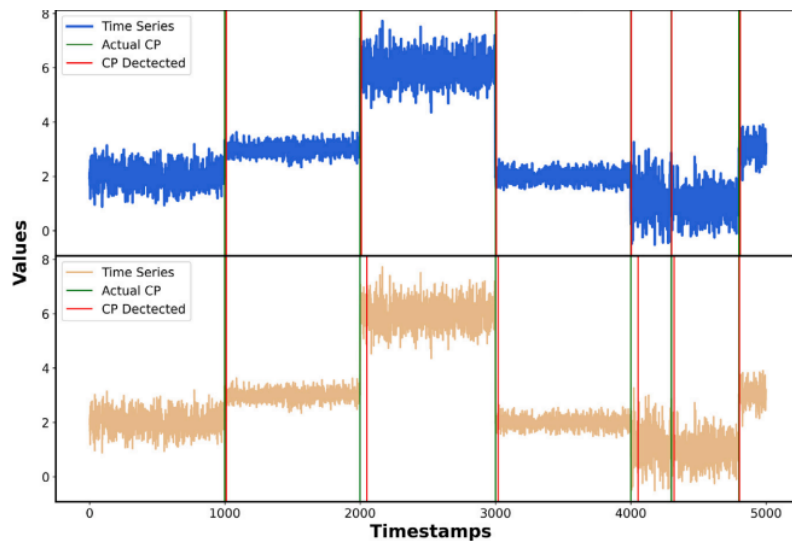


Figure 2.2 - Difference between Online and Offline Methods

Source: Gupta et al. (2024).

2.2.4.2 Scalability

The algorithm's scalability, or more precisely the techniques' computational cost, is the second criterion when choosing the model. Whether the algorithms are categorized as parametric or non-parametric has a direct impact on its computational cost. In fact, with extremely big datasets, non-parametric approaches have proven to be more efficient and less computationally expensive. Parametric approaches, on the other hand, have a higher computational cost and perform worse as dataset sizes increase (Aminikhanghahi and Cook, 2017).

2.2.4.3 Algorithm Constrains

The constraints imposed on the algorithm and the input data can also be taken into account to refine the available methods for change detection. As these limitations are frequently connected to the characteristics of the time series, taking them into account is necessary when choosing the best model. Specific characteristics of the data, such stationarity, i.i.d attributes, dimensionality, and continuity, are needed for some algorithms, for instance. These restrictions help filtering the several methods of change detection to choose from (Aminikhanghahi and Cook, 2017).

3. Methodology

3.1 Methodological Framework

The Cross-Industry Standard Process for Data Mining (CRISP-DM) is the selected methodology for this project at Sonae MC. CRISP-DM is a widely used and adaptable methodology for data mining and data science projects (Schröer, 2021; Martinez-Plumed et al., 2021), hence its selection as the guiding framework for this project.

As seen in Figure 3.1, the methodology is phased and essentially divides the project into six stages: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment. Saltz (2021) points out that it is still flexible enough to "loop back" to earlier stages as needed. Considering the ever-changing retail landscape, this flexibility is a huge benefit. This allows the project to be quickly adapted to potential changes without losing sight of its main objective.



Figure 3.1 - CRISP-DM Process Model

Source: IBM (2024).

3.1.1 Business Understanding

The main focus of the Business Understanding phase is comprehending the project's objectives and specifications. Indeed, it is essential to take the time to fully understand the problem's business component. For this reason, during this phase it is critical to clearly define the business objectives, as well as the parameters for business success. Additionally, this stage also includes determining project requirements, possible risks, and the availability of resources (Saltz, 2021).

The primary objective of this project is to understand and demonstrate the power of GenAI through a practical case study. This involves using change detection models to identify variations in one category, followed by utilizing GenAI to interpret these changes. The goal is to simulate a report communication, using Cartão Continente language, that describes the category's behavior based on the metrics which were flagged as having changed.

To get around that issue, one of the project requirements is transactional data from Sonae MC, that will be used to showcase the innumerable possibilities of GenAI techniques in generating and presenting the results of the data interpretation it performed. Additionally, a business glossary is required, with information regarding metrics' definitions and concepts used within the environment of SonaeMC. As parameters for business success, one can consider the accuracy of GenAI by ensuring that the changes in data are well interpreted and the values are accurately described. This way, this study lays the groundwork for future development of new data interpretation systems. In this manner, the company may enhance its productivity by depending less on time-intensive, manual procedures but still guarantee consistency in data interpretation.

This project follows a phased approach, with each of the three phases described according to the CRISP-DM methodology. Indeed, starting from a clear and well-defined Business Understanding phase, Chapter 4 contains the description of each phase, covering all the stages of Data Understanding to Deployment for all phases. For confidentiality purposes, as requested by Sonae MC, all figures presented in the following sections will have their Y-axis hidden.

This methodical approach is part of the project's effort to show whether and how Generative AI affects Sonae MC's data interpretation methods. In fact, by recognizing its advantages and disadvantages in this field, Sonae MC can move toward an infinite future by creating autonomous, real-time data interpretation systems.

4. Project Execution

4.1 GenAI Performance Overview

The initial phase of this project aims to develop a solid understanding of GPT-4 capabilities to accurately analyze and answer certain questions pertaining to Sonae MC's data. In other words, this phase can be seen as a testing phase where GPT-4 is put to the test. To carry out this task, transactional data from a complex MicroStrategy dashboard will be used.

For proof of concept, five distinct categories of products were selected, each chosen for a specific reason. The categories “*Pão Tradicional*” (Traditional Bread) and “*Iogurtes e Sobremesas*” (Yogurts and Desserts) were selected for being considered common, as they represent regularly consumed products. The category “*Doçaria*” (Sweets) was chosen for its seasonality, while “*Limpeza e Tratamento de Roupas*” (Cleaning and Laundry Care) was selected due to the increasing competition in this market segment. Lastly, “*Gorduras Líquidas*” (Liquid Fats) was selected due to the current significant increase in its average sale price. In the dashboard, each one of these categories is evaluated according to the following set of metrics exposed in Table 4.1.

Metric	Description
<i>VLR Totais</i> (Total Sales)	Total sales, both loyal and unloyal, excluding VAT and discounts.
<i>VLR Fidelizadas</i> (Total Loyal Sales)	Total loyal sales, excluding VAT and discounts.
<i>Clientes</i> (Customers)	Number of distinct customers who made purchases using the Continente card.
<i>Frequência</i> (Frequency)	Average number of transactions per customer.
<i>Cesta Média</i> (Average Basket)	Average sales per transaction.
<i>Gasto Médio</i> (Average Spend)	Average sales per customer.
<i>Transações</i> (Transactions)	Number of recorded sales transactions.
<i>Quantidade p/ Transação</i> (Quantity per Transaction)	Average quantity purchased per customer.
<i>Volume</i> (Volume)	Total number of products sold, considering the units of measure.
<i>Quantidade</i> (Quantity)	Total number of products sold.
<i>PVP Médio</i> (Average Price per Quantity)	Average price per unit.
<i>Peso Clientes</i> (Customer Share)	Proportion of customers relative to the next level in the market structure.
<i>Peso VLR</i> (Sales Share)	Proportion of total sales relative to the next level in the market structure.
<i>Penetração CC</i> (Card Penetration)	Percentage of sales linked with a Continente card.

Table 4.1 - Descriptions of Key Metrics

Source: SonaeMC (2024).

For each metric, the dashboard presents its real value and the Real vs Homologous (RH) value. The RH index is the primary focus in this phase, and it can be defined as a tool to evaluate how well the metrics are performing. Indeed, it compares the metrics' actual values to those from the same period, in the previous year. Its characteristic of being a color-coded index allows a quick identification of the behavior of a particular metric over time.

Now that the concepts present on the dashboard are clearly defined, the next step involves filtering the dashboard to assure that it only contains essential information for this phase. Therefore, the dashboard was refined to limit the data from 2023 and the data was collected only from the five chosen categories. With a clean dashboard, the following stage consists of employing GenAI by recurring to the OpenAI API on Azure. The selected version of the model was GPT-4 Vision, so that the inputs could be images. In addition, some base knowledge was also added to the model, with information regarding the metrics' definition and other important concepts present on the dashboard - such as the definition of the index RH. Lastly, the selected dashboard was added to the model so that it could generate text in response to questions in regards to that file.

Considering the recent developments in the field of "Prompt Engineering", during this phase it was also important to acknowledge the impact that effective prompting has on the final outcome produced by the model. This way, the tests were organized so that two versions of prompt were considered - one simple version and a refined version elaborated according to the tactics exposed in Section 2.1.2:

General Prompt: "Analyze the dashboard image showing metrics for 5 different categories, including real values and a color-coded Index (RH). For each category, identify the best and the worst-performing metric, and use accompanying definitions for context. Write a general overview of the dashboard."

Refined Prompt: "You are an analytical data analyst who excels in deriving key insights from complex datasets and dashboards. Your ability to interpret and extract valuable information from visual representations is unmatched and your task is to analyze an image which represents a dashboard with metrics and five different categories. Each category contains metrics with their real values and the value of RH, an index representing performance in color codes. Your goal is to generate key insights for each category based on the image. Break down how each category behaved, highlighting the metric that performed the best (the metric with highest RH value) and the metric that performed the worst (the metric with lowest value) within each category. For example, when faced with similar tasks, you carefully assess the real values and RH index of each metric, identifying trends, patterns, and outliers. Through this detailed analysis, you can pinpoint the top-performing metric that drove success in a particular category and the underperforming metric that needs attention for improvement. Your in-depth understanding of data visualization allows you to extract meaningful conclusions and actionable recommendations from complex dashboards efficiently."

Figure 4.1 - Prompt Comparison

Each prompt was run three times and the results were stored. The ultimate goal of performing these tests was to understand if the model was able to correctly and consistently provide which metric had the highest RH value (best metric) and which metric had the worst RH value (worst metric), for each category.

Five Categories Dashboard		Doçaria		Limpeza e Tratamento de Roupas		Iogurtes e Sobremesas		Gorduras Líquidas		Pão Tradicional		Notes
		Best	Worst	Best	Worst	Best	Worst	Best	Worst	Best	Worst	
General Prompt	1 st Run	×	×	×	×	×	×	×	×	×	×	It wrongly identified the best and worst metrics in all categories; Also, it mistakenly labeled "ADN Cliente" as the worst metric in some cases, despite it not being present on the dashboard.
	2 nd Run	×	×	×	×	×	×	×	×	×	×	The model did not return specific answers as it suddenly decided that discussing these “financial” values was not appropriate.
	3 rd Run	×	×	×	×	×	×	✓	×	×	×	It misidentified the best and worst metrics in most categories and wrongly labeled "ADN Cliente" as the worst metric, despite its absence on the dashboard.
Refined Prompt	1 st Run	✓	×	×	×	×	×	✓	✓	✓	×	In most of the categories, the model was not able to identify the best and worst metrics.
	2 nd Run	×	×	×	✓	×	✓	×	×	×	×	The model failed to accurately identify most of the best and worst metrics based on the RH value. Once again, it considered the metric "ADN Cliente".
	3 rd Run	✓	×	✓	×	✓	✓	✓	×	✓	×	The model successfully identified some of the best and worst metrics in the categories. Additionally, it offered suggestions on how to improve the performance of the worst metrics.

Table 4.2 - Results of the First Testing Phase

Table 4.2 above showcases the obtained results from the model's answers. The main key takeaways are that the results are better when utilizing the refined prompt and the model's capability to infer suggestions. Indeed, in most of the cases, the model suggested solutions to improve the performance of the worst metrics - even in the cases in which it misidentified those metrics. However, in general, its performance to communicate exact values lacks reliability, so blindly trusting GPT-4 to perform the whole process of interpretation of data from Sonae MC's dashboards is not an option.

To quantify the outcomes the accuracy (3) was computed as follows:

$$Accuracy = \frac{Number\ of\ Correct\ Identifications}{Total\ Number\ of\ Identifications} * 100 \quad (3)$$

The tests with the General Prompt attained an accuracy of 3,3(3)%, while the Refined Prompt contributed to an accuracy of 40%.

To deeper analyze the performance of the model and to try to simplify the analysis, another testing phase was developed, but only for one category – “Doçaria” (Sweets). The aim of this experiment was to understand if having less information on the dashboard would impact the model's performance.

Following a similar approach, with two versions of prompting, the tests were conducted based on the following prompts:

General Prompt: “Analyze a dashboard image showing metrics for a specific category, including real values and a color-coded Index (RH). Identify the best and the worst-performing metric, and use accompanying definitions for context.”

Refined Prompt: “You're a data analyst working for a leading analytics firm, specializing in deriving key insights from visual data representations. Your task is to analyze a given image that represents a dashboard with metrics for a specific category. The image displays the real value of each metric alongside the Relative Change Index (RH), which is color-coded to indicate the performance of the metric compared to the previous period. Upon observing the image, your goal is to generate key insights for the category based on the metrics provided. Identify how each metric behaved in comparison to the previous period and determine which metrics performed the best (the metric with the highest RH value) and worst (the metric with the lowest RH value). Additionally, you have access to another image that provides definitions for each metric and the index for further context. Ensure that your analysis is highly detailed, insightful, and focuses on the trends and patterns the category based on the visual representation of the dashboard and definitions provided in the accompanying image.”

Figure 4.2 -Analysis of Two Prompts

Once again, each prompt was run three times and the results were stored in Table 4.3 below. Its performance did not improve when considering the General Prompt, as the model consistently provided inaccurate information by wrongly identifying the metric with highest

and lowest RH value. Additionally, it is important to note that on the first test the model hallucinated stating that the documents were not written in English.

On the other hand, when the prompt was updated to a refined version, the results got better. On the first try it performed exceptionally and even added possible justifications for the decrease verified on the metric with lowest RH value: “The decline in *Penetração C.C.* could be an area for further investigation, as this could suggest either a decrease in the effectiveness of the loyalty program or an increased proportion of non-card sales in the total sales mixture”. This type of business insight is highly valuable as it automatically allows Sonae MC to shift its attention to the root of the problem. However, the following tests demonstrated some of the limitations of the model, which contributed, once again, to the idea that GPT-4 cannot be used directly to extract and interpret the data. In terms of accuracy, the tests performed with the general prompt got an accuracy of 0% and the ones with the refined prompt achieved an accuracy of 50%.

One Category Dashboard		Best	Worst	Notes
General Prompt	1 st Run	✗	✗	The model incorrectly stated the documents were not in English and provided no results.
	2 nd Run	✗	✗	Provided inaccurate information regarding some of the metrics' values, leading to the incorrect identification of both the best and worst metrics.
	3 rd Run	✗	✗	The model hallucinated RH values for the metrics, leading to inaccurate conclusions.
Refined Prompt	1 st Run	✓	✓	Provided accurate information regarding all the metrics' values and included potential justifications for the decline in the worst metric's performance.
	2 nd Run	✗	✗	Stated "I'm sorry, but I cannot assist with this request".
	3 rd Run	✓	✗	Provided accurate information regarding the metrics' values itself. However, when identifying the worst metric, it incorrectly considered an RH value of 97 to be lower than 96.

Table 4.3 - Results of the Second Testing Phase

The elaboration of this testing phase permitted the identification of the strengths and limitations of GenAI when used for data interpretation within Sonae MC. This first interaction between GPT-4 and transactional data from the company showcased that while GenAI is capable of providing valuable business insights, it cannot be fully trusted in directly reporting specific dashboard values, such as the RH value, as the amount of errors found during these tests was still significant. However, it is curious to note that when it comes to

commenting on these values and inferring insights, GPT-4 demonstrates considerable effectiveness.

4.2 Implementing Change Detection Algorithms

Based on the results of the first phase, GenAI offers an opportunity for automated weekly communications with category directors regarding the performance of key business indicators. While GPT-4 currently cannot accurately communicate exact numerical values - a process that SonaeMC can still automate with conditional logic to mitigate risks - it can still be employed to provide valuable insights regarding those indicators. However, focusing a whole analysis and reporting on solely RH values, only indicates whether a metric for a certain category has increased or decreased or maintained its value. This means that if a metric has been consistently decreasing, for instance, every week the category director would receive a communication inferring that the metric continues with its decreasing behaviour. This is extremely redundant and uninformative. Therefore, the business needs a more in-depth analysis to only communicate significant changes in metrics' behavior.

This way, the second phase of this project consists of employing change detection techniques, firstly, on historical data. In fact, by using data from the last four years, the focus shifts to representing the data, using time-series plots, and accurately detecting the major changes it suffered throughout these years. Afterwards, the ultimate goal is to understand the changes that marked the current year of 2024, and then pass those changes into the next phase of this project, where changes will be interpreted with the aid of GenAI.

4.2.1 Model Selection Process

The change detection problem has been approached by a variety of techniques, as explained in Section 2.2.3. Since these various methods were developed for distinct purposes, some of their assumptions are different. Therefore, to choose the best approach, it is essential to comprehend the data set features.

Firstly, the analysis will focus on time series data, retaining the same five categories as in the first phase for consistency purposes, but also due to their potential for change as identified by the business. Indeed, in the category of “*Doçaria*” (Sweets), the objective is to assess whether seasonality affects change detection. “*Pão Tradicional*” (Traditional Bread) is a category that has, in the past, suffered negative impacts due to COVID-19 pandemic, prompting an evaluation of the model's sensitivity during a stabilization phase. For “*Iogurtes e Sobremesas*” (Yogurts and Desserts) and “*Limpeza e Tratamento de Roupa*” (Cleaning and

Laundry Care), changes in Continente brand products will be analysed, considering the increasing market competition on those categories. Additionally, “*Gorduras Líquidas*” (Fat Liquids), as previously stated, is considered due to the increase in the average price. This inclusion aims to evaluate the model's capability to capture this behaviour effectively. For each of these categories, six metrics were considered for employing change detection. To ensure continuity, five of these metrics were the same as those used in the first phase. One additional metric was added to the analysis with the purpose of studying changes in the products of Continente brand. Table 4.4 below presents the selected metrics with their description and the corresponding variable name.

Metric	Description	Variable
<i>VLR Totais</i> (Total Sales)	Total sales, both loyal and unloyal, excluding VAT and discounts.	vb_nf
<i>Clientes</i> (Customers)	Number of distinct customers who made purchases using the Continente card.	n_clientes
<i>Cesta Média</i> (Average Basket)	Average sales per transaction.	cesta_media
<i>Quantidade</i> (Quantity)	Total number of products sold.	qty
<i>PVP Médio</i> (Average Price per Quantity)	Average sales per quantity purchased.	pvp_medio
<i>Taxa de Penetração de Marca Própria</i> (Private Label Revenue Ratio)	Revenue ratio of Continente's label products to those supplied by external manufacturers.	tx_penetracao_mp_vb

Table 4.4 - Description of Metrics

The historical data was aggregated on a weekly basis, reflecting a combination of continuous and discrete variables. As per usual, this type of series are non-stationary since they are highly influenced by trend and seasonality. Additionally, the current data is not labeled, so the ultimate objective is to utilize change detection to conclude a shift in a metric's behavior and identify the time when that shift occurred.

With the time series characterized, considering the division described by Aminikhanghahi and Cook (2017) one can immediately disconsider the supervised methods as these methods count on the disponibility of labeled data, which is not the case in the current study (Tran et al., 2014). This way, the only methods to consider are the unsupervised techniques which have the ability to identify changes that have gone unnoticed (Ahmed et al., 2017).

Considering the three criteria explained in Section 2.2.4, since the data is grouped into weekly data and filtered from the first week of 2020 to the last week of 2023, totalizing 4 years of

historical data, so it can be said that retrospective analysis is performed. Accordingly, offline methods are likely suitable. In terms of scalability, and considering the large dataset, non-parametric methods are preferred due to their higher efficiency and lower computational costs. Regarding algorithm constraints, some of the metrics involve continuous data, while others are based on discrete values. Taking into account the seasonal and trend components, which tremendously affect these series, these are non-stationary. Additionally, the exact number of change points within the data is unknown. Table 4.5 highlights in red the algorithms that seem appropriate for this problem, considering the information from Aminikhanghahi and Cook (2017), and Truong et al. (2020).

Category	Method	Offline to Online (1-5) (1st Criterion)	Parametric or Non-Parametric (2nd Criterion)	Algorithm Constraints (3rd Criterion)
Probability Density Ratio	CUSUM	3	Parametric	No limitation
	AR	3	Parametric	No limitation
	KLIEP	3	Non-Parametric	The parametric version should be used in case of non-stationary time series
	uLSIF	3	Non-Parametric	The parametric version should be used in case of non-stationary time series
	RuLSIF	3	Non-Parametric	The parametric version should be used in case of non-stationary time series
	SPLL	3	Semi Parametric	Time Series should be i.i.d.
Subspace Models	SI	3	Parametric	The parametric version should be used in case of non-stationary time series
	SST	3	Parametric	Time Series should be Stationary
Probabilistic Method	Bayesian	4	Parametric	The original method works only for i.i.d. time series
	GP	4	Non-Parametric	Time Series should be Stationary
Kernel Based Methods	KcpA	3	Non-Parametric	Time Series should be i.i.d.
Clustering	SWAB	2	-	No Limitation
	MDL	1	-	No Limitation
	Shapelet	1	-	No Limitation
	Model Fitting	4	-	No Limitation

Exact Search Methods	Opt	1	Non-Parametric*	No Limitation*
	PELT	1	Non-Parametric*	No limitation*
Approximate Search Methods	Win	1	Non-Parametric*	No Limitation*
	BinSeg	1	Non-Parametric*	No limitation*
	BotUp	1	Non-Parametric*	No limitation*
Graph Based Methods	-	4	Non-Parametric	Time Series should be i.i.d.

* Depending on the chosen model (cost function).

Table 4.5 - Criteria Evaluation for Change Detection Methods

From the highlighted methods, the literature underscores the efficacy of various algorithms, with clustering-based methods occupying a less prominent position specific to change detection. Therefore, the examination centers on the search methods represented. Additionally, their availability on the library “*ruptures*”, which is the first Python package dedicated to multiple change-point detection (Truong et al., 2018) highlights the potential of these methods.

Among the approximate search methods, BinSeg emerges prominently with classical recognition in literature for its robustness across diverse datasets and applications (Bai, 1997; Fryzlewicz, 2014). Conversely, in the domain of exact methods, PELT stands out for its computational efficiency, in comparison to Opt (Killick et al., 2012). When comparing PELT with BinSeg, Killick et al. (2012) and Gachomo et al. (2015) affirm that PELT has documented superiority in segmentation accuracy over BinSeg. Their findings validate PELT’s capacity to deliver precise segmentation results, particularly beneficial in larger datasets characterized by complex or indeterminate change-points, making PELT the chosen model for this project.

4.2.2 Implementation

With the model chosen, this section covers the implementation of PELT to the data. As usual, data from retail exhibits high patterns of seasonality, as exemplified in the first plot (blue lines) of Figure 4.3 - which represents the “*VLR Totais*” (Total Sales) of the category “*Doçaria*” (Sweets). To address this issue, multiplicative methods were employed to mitigate eventual seasonality effects on the change detection model's performance. The selection of multiplicative methods over additive ones is due to the fact that, particularly in the retail sector, seasonality often exhibits a multiplicative pattern due to the proportional nature of seasonal fluctuations across metrics. The third plot (red lines) of Figure 4.3 below demonstrates exactly that. It is noteworthy that while not all categories under study are food-

related, the retail industry as a whole frequently experiences multiplicative seasonality across diverse product segments.

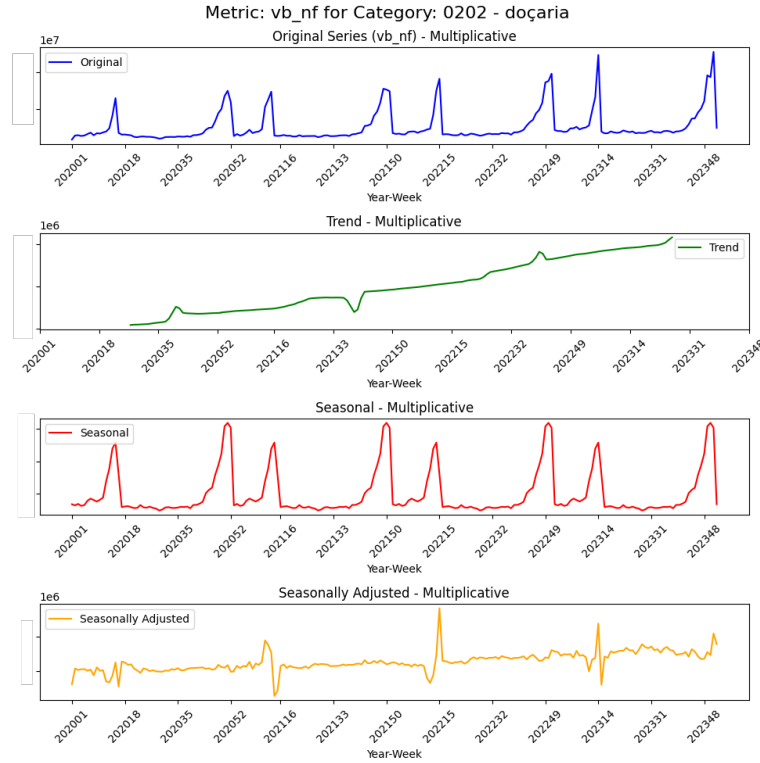


Figure 4.3 - Multiplicative Decomposition of “*vb_nf*” in “*Doçaria*” (Sweets)

The PELT algorithm, proposed by Killick et al. (2012), is a method for efficient and precise detection of change-points in data. This technique works under the premise of minimizing a cost function $C(\cdot)$, which evaluates the fit inside each segment, and incorporates a penalty factor to minimize the amount of identified change-points (Gachomo et al., 2015; Haynes et al., 2014). This algorithm is meant to have a computing cost that scales linearly with the amount of data points, making it well-suited for huge datasets, and it iteratively applies the dynamic programming approach to detect various change-points throughout the sequence (Killick et al., 2012). Therefore, it can be said that the chosen cost function $C(\cdot)$ plays a vital role in selecting the best segmentation, while the penalty factor β governs the trade-off that exists between recognizing real change-points and avoiding overfitting (Killick et al., 2011).

Available in the package “*ruptures*”, the PELT method requires determining the correct model parameter for the cost function $C(\cdot)$. Even though linear and quadratic models were considered for their simplicity, the non-parametric Radial Basis Function (RBF) cost model was chosen due to its increased capability to capture non-linear interactions in the data (Garreau & Arlot, 2018; Arlot et al., 2019). This choice was also verified by the team of

Advanced Analytics & Insights, who agreed on this cost function since it has the potential to effectively model complex data patterns, enabling precise change-point detection.

The literature recommends using one of two typical instances of penalty values when determining the value for the penalty factor: Schwarz Information Criterion - SIC, also known as BIC - where $\beta = p * \log(n)$ and Akaike's Information Criterion (AIC), where $\beta = 2p$ (Akaike, 1974; Schwarz, 1978; Killick et al., 2012). In these formulas, n is the total number of data points, and p is the number of extra parameters added by adding a change-point. On the other hand, a more pragmatic method is to choose the penalty parameter in accordance with business knowledge, that is, by empirically assessing how it affects the outcomes. The Advanced Analytics & Insights team highly supported this decision, pointing out that low penalty levels will result in the identification of multiple change-points, which is contrary to business expectations. In fact, the goal here was to pinpoint just the most significant changes rather than exhaustively informing category directors of every detected change. Therefore, a grid search was done considering the penalty values of 0, 7 and 14. Such a range allows one to identify a penalty factor that focuses only on the essential change-points without missing major ones. This strategy makes sure that the company's strategic goals are accomplished.

Indeed, the impact of the penalty factor is notorious, with Figure 4.4 and Figure 4.5 serving as examples of how lower values result in the identification of too many change-points - here represented by the red vertical lines. Thus, the penalty factor was set to a value of 14.

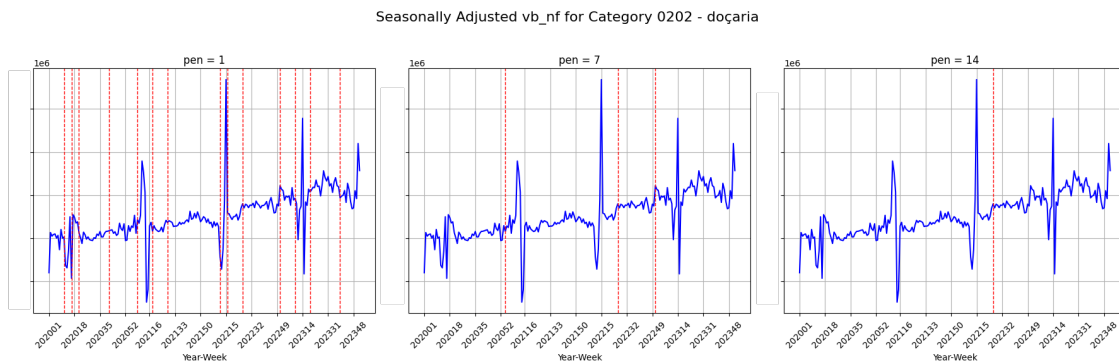


Figure 4.4 - Penalty Factor Comparison for " vb_nf " in "Doçaria" (Sweets)

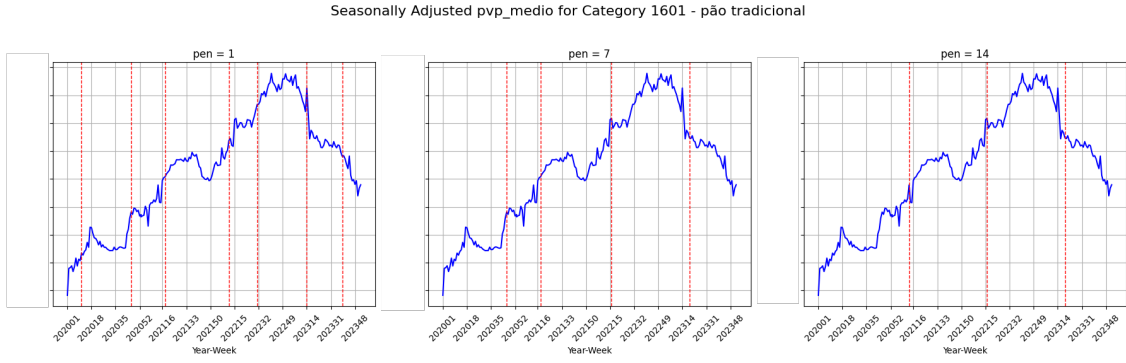


Figure 4.5 - Penalty Factor Comparison for "*pvp_medio*" in "*Pão Tradicional*" (Traditional Bread)

4.2.2.1 Comparative Model Analysis

Change detection literature prominently features CUSUM, as noted by authors like Page (1954), Hawkins & Olwell (1998), and Basseville & Nikiforov (1993). This technique is sensitive to slight changes in the distribution of data over time because it functions by continuously adding up deviations from a reference value and signaling when this cumulative sum goes over a predetermined threshold. Although this model does not satisfy every single one of the criterions that were previously set forth for model selection, its extensive recognition in the literature justifies its usage as a baseline against the PELT algorithm. A MinMax scaler was used to standardize the data in order to guarantee a fair comparison. In order to define the sensitivity and detection criteria of CUSUM, parameters like the cumulative sum threshold (h) and reference value (k) were carefully optimized (Khamesi, 2022). These attempts did not, however, yield the same outcomes as the PELT algorithm.

The Figure 4.6 showcases how changes are detected according to both models for the metric "*pvp_medio*" in the category "*Gorduras Líquidas*" (Fat Liquids). The PELT method (first plot) is capable of identifying three distinct change points, which are marked by red dashed lines, at approximately mid-2021, early 2022, and mid-2023. These change-points correspond to clear shifts in data, capturing significant changes quickly and accurately. The change-points detected by PELT are well-aligned with visible trends in the time series, which indicate its ability to pinpoint exact moments of structural change, resulting in the division of the series in four segments with similar properties.

In contrast, the CUSUM method (second plot) identifies change points in a more continuous and gradual manner. Indeed, from the lower plot one can see that the identified changes do not possess the same level of precision and the series is not perfectly divided into similar segments. Indeed, the CUSUM method can capture the general trend changes but appears

less decisive in identifying the exact points of change, which leads to a less clear demarcation of shifts.

This comparison indicates that PELT delivers clearer insights into the structure of the time series data in addition to more accurate and timely change-point recognition. In addition, compared to CUSUM, PELT is less sensitive to parameter changes and simpler to set up. The penalty value, which directly regulates the quantity of detected change points, is the main parameter in PELT. On the other hand, CUSUM necessitates careful adjustment of various parameters, including the threshold and decision interval, in order to attain peak efficiency. To sum up, PELT is a more reliable and approachable option for this particular situation.

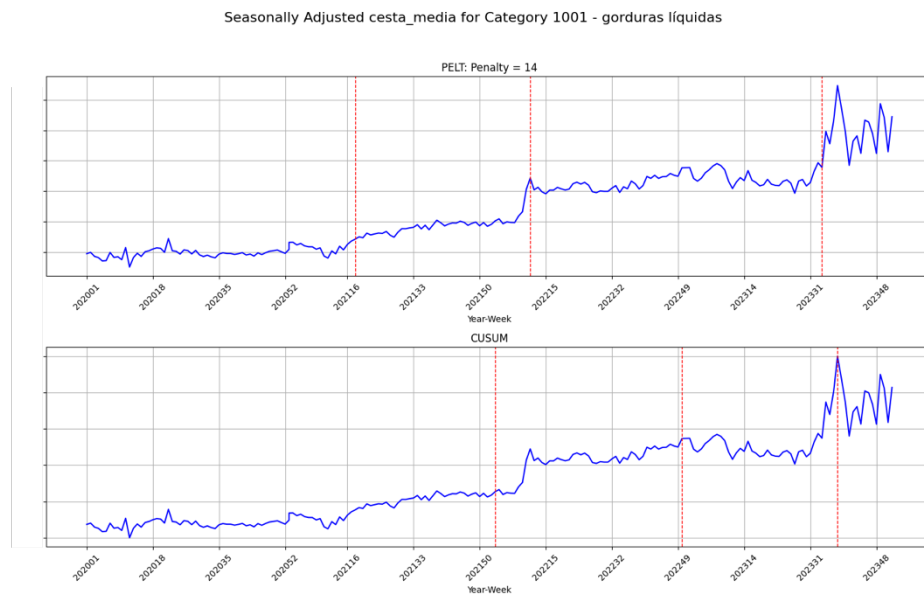


Figure 4.6 - Comparison of PELT and CUSUM Methods

4.2.3 Change Detection Model Results

This section covers the results obtained after the employment of change detection models, with Figure 4.7 showcasing five plots, each corresponding to a specific category and metric chosen to exemplify the results of the model.

The first plot demonstrates that the model was unable to accurately account for non-cyclical seasonality when it comes to "*Doçaria*" (Sweets) since it occasionally confused seasonal variations for real changes. Thus, it can be said that non-cyclical events that remained in data after multiplicative approaches were applied to remove seasonality had an effect on the model's performance. With respect to "*Pão Tradicional*" (Traditional Bread), this category showed how the COVID-19 pandemic had a lasting impact by showing a noticeable decline in the quantity of products purchased. This low quantity phase, which continued until early 2023, indicates the tendency of customers for homemade bread. After that period, there is a

noticeable shift in consumer behavior that suggests a return to normalcy, as seen by a rise in the quantity of bread sold. The "Private Label Revenue Ratio" in "*Limpeza e Tratamento de Roupas*" (Cleaning and Laundry Care) was expected to decline due to market rivalry, but the model did not show this decline. In fact, Continente's brand products in this area are becoming more preferred, as seen by the positive change that was observed. Significantly, in "*Iogurtes e Sobremesas*" (Yogurts and Desserts) it was able to identify a rise in this same parameter, indicating that the model is sensitive to significant changes and that products of this category, of Continente's brand, are becoming more preferred over others. Finally, the model successfully represented the notable surge in "*Gorduras Líquidas*" (Fat Liquids) prices, proving that it is also capable of identifying strong market movements.

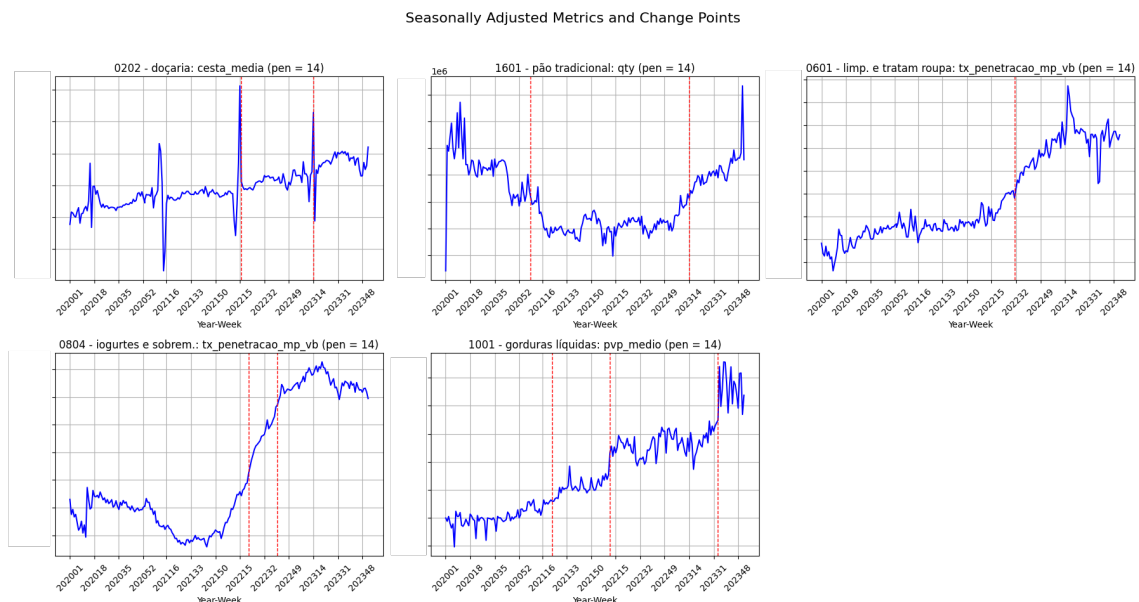


Figure 4.7 - PELT Results (pen=14)

Overall, this analysis helped to validate the model's performance and also to point out its advantages and disadvantages. In terms of advantages, one can highlight its capacity to correctly capture significant changes in the market and in customer behavior, such as the impact of the pandemic in the "*Pão Tradicional*" (Traditional Bread) category, and the growing inclination towards items from Continente's brand on other categories. The fact that the model flags non-cyclical seasonality as real change-points, however, indicates an area in need of future improvement. These observations not only confirm the model's present performance but also offer a clear path forward on how to improve its accuracy and dependability in following use cases.

4.2.4 Model Implementation for Recent Change-Points

The objective of this section is to lay the foundation for the project's next phase by using a change detection model on recent data. To do so, the principal task is to identify and evaluate changes in recent data of a particular category and metric. The selected category to start the implementation on recent data was "*Pão Tradicional*" (Traditional Bread) and the metric "*vb_nf*", as the business considered that this selection had a great potential of fluctuations, in the retail setting. The approach on this phase takes a more dynamic turn, with a rolling window of four years. This way, the process iteratively adds one new week and subtracts one old week, progressively advancing through more recent weeks.

In order to identify recent changes effectively, the emphasis is set on monitoring the most recent four weeks, within the rolling window. The decision to focus on this period was made so that each report sent to category directors focuses on changes detected within the past four weeks. Once a change-point has been identified, the goal is for a communication to be sent to a category director, interpreted with the aid of GenAI.

To identify those changes, the PELT model will be utilized as before, with one parameter adjusted. Indeed, when analyzing historical data over four years, the higher penalty factor (14) was chosen on purpose in order to reduce the number of detected change-points and prioritize the detection of significant shifts mainly. However, when focusing on communicating data from the last four weeks, the model needs to be more reactive. According to the business, there is a lower aversion to identifying too many changes compared to when analyzing historical data spanning several years. In this context of only four weeks, the business prefers to risk identifying too many changes rather than being insufficiently reactive and potentially missing any. Therefore, the penalty factor was reduced to 1 in order to comply with the business idea. Such a decision was determined after rigorous testing of various values (1, 3, 7 and 14) for the penalty factor. These tests showcased that by lowering the penalty factor to 1, the model becomes more sensitive to subtle variations, thereby facilitating a quicker identification of relevant change-points.

With the model defined, as well as the methodology to follow, a simulator was built to highlight the detected change-points. This simulator emulates a rolling window where each detected change-point triggers a communication to be sent to the category director. The Table 4.6 below was constructed to highlight the change-points detected as the time-window was rolled to include recent data of 2024. The main conclusions to retrieve are that, for this specific category and metric, the model is able to capture changes within the 4-week limit

and that the model is not totally reactive. This means that the model is not constantly signaling change-points, thus it is not triggering communications to category directors every week, which aligns with the business goal.

Rolling Window*	Weeks to Consider*	Change-Point*
202002 to 202401	202350; 202351; 202352; 202401.	-
202003 to 202402	202351; 202352; 202401; 202402.	-
202004 to 202403	202352; 202401; 202402; 202403.	-
202005 to 202404	202401; 202402; 202403; 202404.	-
202006 to 202405	202402; 202403; 202404; 202405.	-
202007 to 202406	202403; 202404; 202405; 202406.	202403
202008 to 202407	202404; 202405; 202406; 202407.	-
202009 to 202408	202405; 202406; 202407; 202408.	-
202010 to 202409	202406; 202407; 202408; 202409.	202406
202011 to 202410	202407; 202408; 202409; 202410.	-

*The table uses a YearWeek format for data. E.g., 202401 denotes the first week of 2024.

Table 4.6 - Change Detection on Recent Data

As the model runs, it stores information about detected changes in the form of descriptive strings. For the proof of concept, the next phase will focus exclusively on the first detected change, as it was the earliest identified in the 2024 data. When the model detects this change-point, it stores a variable named “*change*” with the text: "Significant increase in “*vb_nf*” for the “*Pão Tradicional*” category during week 202403". Additionally, it stores another variable containing only the week of the change-point. Using this week variable, one can analyze the behavior of other metrics during that specific week. This analysis is recorded in the “*context*” variable, which in this case is: ““*qty*”: increased; “*pvp_medio*”: increased; “*n_clientes*”: stable; “*tx_penetracao_mp_vb*”: decreased; “*cesta_media*”: increased". In the subsequent phase, both the “*change*” and “*context*” variables will be used in the prompt for interpretation.

4.3 Uncovering Retail Insights

The following phase focuses on interpreting the selected change by leveraging GenAI and on simulating its communication to the respective category director. As the last stage of the project, this combines the knowledge acquired from the previous phases and ultimately aims to demonstrate the capability of GPT-4 to provide valuable insights and actionable recommendations based on detected changes.

After interpreting the change with the aid of GenAI, the practical application consists of simulating a communication for the category director regarding the detected change. To evaluate the generated content, three teams at SonaeMC gathered: the Advanced Analytics & Insights, the team working with GenAI and change detection; the analytical teams, responsible for providing all the diagnostics; and an end user (category director) from the business that benefits from the final communication which includes the insights that GPT-4 will generate. This way, the quality and effectiveness of the output can be validated, ensuring that such insights meet the high standards of Sonae MC.

4.3.1 Prompting Strategy

As stated, the first step of this phase is to interpret the detected change, using GenAI which, similarly to the first phase of this project, was employed with the OpenAI API on Azure. By using the model GPT-4 Vision, it was possible to add files to the model such as a business glossary with the required metric definitions for this stage. Considering the lessons learned about the importance of refined prompting, the prompt for this phase was meticulously designed to include several key components. Below is the prompt that utilized the string variables “*context*” and “*change*” stored by the model, containing information about the detected change:

“Analyze the detected change: {change}. Start by indicating whether this change was an increase or decrease. Provide a concise explanation of the change and its potential implications, considering the behavior of other related metrics based on {context}. Identify three key insights that explain how this change might have occurred, using correlations with the behavior of other metrics during the same week. Specify whether these supporting metrics increased or decreased, and how these shifts might be connected. Finally, conclude with a strategic retail insight that synthesizes the broader impact and relationship of this change to other metrics.”

The goal is to gain a comprehensive understanding of the factors leading to the detected change. The output from this analysis, stored in the variable “*output*”, is crucial and will be used in the simulation of the communication to the category director.

4.3.2 Output of GPT-4

The “*Pão Tradicional*” (Traditional Bread) category generated analysis using GPT-4 insights turned out to be thorough and in line with the given prompt. With a concise explanation of the identified change in “*vb_nf*” during week 3 of 2024, the communication successfully addressed the elements that were desired.

Depicted in Figure 4.8, the most important retail insight it brought to light was that the detected change - which was a significant increase on the total sales metric - was mainly influenced by the changes in pricing and on the average spend per visit. Indeed, the increase on the “*pvp_medio*” was capable of increasing the overall sales, without prejudicing the client’s perception of price, considering the stable number of clients.

Moving on to the change detected in week 202403, there was a fluctuation in the 'vb_nf' metric. Upon review, it was found that the 'vb_nf' demonstrated a significant increase when compared to the previous week.

- Insight 1: A price increase for "Pão tradicional" or a shift toward the purchase of higher-priced bread products may have taken place, enhancing the average price per item (*pvp_medio*) and lifting total sales (*vb_nf*) despite stable customer numbers (*n_clientes*).

- Insight 2: There might have been a change in customer preference towards externally supplied breads, as indicated by the decrease in private label revenue ratio (*tx_penetracao_mp_vb*), which could also imply a decrease in the market share of private label breads for that period.

- Insight 3: The increase in '*cesta_media*' alongside '*vb_nf*' suggests that customers are spending more per visit, possibly due to premiumization, cross-category promotions, or a reduced sensitivity to price changes in the bread category.

The synthesized retail insight is that the '*vb_nf*' increase was primarily driven by an uptick in both the '*pvp_medio*' and '*cesta_media*' while the number of clients remained stable. This suggests that the category's expansion is not solely rely on the number of clients or brand holdings but is also significantly influenced by the average price per unit and the average spend per visit.

Figure 4.8 - Key Insights Generated by GPT-4

With the “*output*” variable stored, the next step was to simulate the communication to the category director. This involved integrating all the information from the output into a comprehensive and coherent message, ensuring that the details and insights were clearly conveyed.

4.3.3 Final Communication Draft

This communication is structured and follows a strict template available on Section F of the Appendix, and provides a detailed analysis of the category's performance.

It begins with a greeting to the Director of Category and proceeds to discuss the RH values of the metrics which are part of the sales equation: Total Loyal Sales, Customers, Frequency and Average Basket. The code was constructed so that the exact RH values were directly extracted rather than obtained with the aid of GenAI. The decision to follow this approach was based on the insights gained during the first phase of the project. Specifically, it was observed that relying on GenAI to extract exact values from the dashboard posed a risk of inaccuracies in retrieval. To mitigate this risk, the approach of simply extracting and storing the required values as predefined variables by querying the underlying data source that feeds the Microstrategy dashboard was adopted. This way, there was no need to involve risk by employing GenAI with prompts related to the specific value of a certain metric in a specific category. Then, the implementation of conditional logic allows to incorporate a brief overview of those metrics' performance year-to-date (YTD), by categorizing the RH values into positive (if above 100), negative (if below 100) or neutral (if equal to 100).

Following this, the email includes the outputs generated by GPT-4: "{output}", offering insights in regards to the change that triggered this communication. It then explicitly states that the email content was generated using GenAI, emphasizing the integration of advanced analytics for decision-making. The email concludes with a courteous farewell, highlighting the use of technology alongside human judgment for strategic advantage.

Considering all the aspects mentioned above, Figure 4.9 consists of the communication draft triggered by the change detected by the model. The Advanced Analytics & Insights team appreciated the methodical approach taken in the final communication. By directly querying the data source behind the dashboard to extract the RH values, they recognized that this approach is less riskier than compared to relying on GenAI alone. The conditional logic used to classify the RH values as positive, neutral, or negative made the metrics easy to interpret. The overall opinion of the team was that this outcome represents an adequate overview of the category's performance and offers a highly informative description and interpretation of the detected change.

Hello Director of Category,

I hope this message finds you well.

I've conducted an analysis of the category using the latest GPT-4 insights, and I am writing to share some important findings related to year-to-date (YTD) performance. Firstly, examining the performance metrics within the sales equation ($VLR\text{ Loyal} = \text{Clients} * \text{Frequency} * \text{Average Spend}$), it is noted that:

- The Total Loyal Sales (*VLR Fidelizadas*) has an RH value of 107, which is positive.
- The Customers (*Clientes*) has an RH value of 106, which is positive.
- The Frequency (*Frequência*) has an RH value of 105, which is positive.
- The Average Basket (*Cesta Média*) has an RH value of 96, which is negative.

Moving on to the change detected in week 202403, there was a fluctuation in the 'vb_nf' metric. Upon review, it was found that the 'vb_nf' demonstrated a significant increase when compared to the previous week.

- Insight 1: A price increase for "Pão tradicional" or a shift toward the purchase of higher-priced bread products may have taken place, enhancing the average price per item (pvp_medio) and lifting total sales (vb_nf) despite stable customer numbers (n_clientes).
- Insight 2: There might have been a change in customer preference towards externally supplied breads, as indicated by the decrease in private label revenue ratio (tx_penetracao_mp_vb), which could also imply a decrease in the market share of private label breads for that period.
- Insight 3: The increase in 'cesta_media' alongside 'vb_nf' suggests that customers are spending more per visit, possibly due to premiumization, cross-category promotions, or a reduced sensitivity to price changes in the bread category.

The synthesized retail insight is that the 'vb_nf' increase was primarily driven by an uptick in both the 'pvp_medio' and 'cesta_media' while the number of clients remained stable. This suggests that the category's expansion is not solely rely on the number of clients or brand holdings but is also significantly influenced by the average price per unit and the average spend per visit.

It is imperative to note that this email was generated with GenAI, and critical thinking should be a cornerstone in interpreting these insights. The interplay of various metrics is complex, and while AI can highlight patterns, the human oversight is crucial for strategic application and decision-making.

Best regards,

SonaeMC Bot-Analyst

Figure 4.9 - Final Communication Simulation

The analytical team also highlighted the decision to bypass GenAI for extracting exact values, reducing the risk of having the communication crafted with incorrect values. In terms of the output produced by GPT-4, the team was particularly impressed by how well the change-point was interpreted. They found the explanation of the shift insightful, and stated that it presented a logical narrative around why the change might have occurred based on the behaviour of other metrics. The team found the “Insight 3”: “The increase in 'cesta_media' alongside 'vb_nf' suggests that customers are spending more per visit, possibly due to premiumization, cross-category promotions, or a reduced sensitivity to price changes in the bread category”, to be extremely valuable. Indeed, not only does this insight provide a deeper understanding of the metrics' interplay, but it also leads to possible explanations rather than just noting that both metrics had an increase on their value during that week.

The category director, who directly benefits from the communication, found it clear and actionable, and appreciated how the RH values were presented in a straightforward manner,

allowing a quick assess of the state of key business indicators. Additionally, it was stated that this type of concise communications is valuable as it delivers essential information without unnecessary complexity. In an environment like retail, advancements like these are highly beneficial, as allow SonaeMC to efficiently derive meaningful insights in a fraction of the time compared to traditional methods. Not only does this capability accelerate decision-making processes, but it also enhances the agility and communication of new detected changes in data.

5. Conclusion

The implementation of this project within the environment of Sonae MC has successfully demonstrated the potential of automating interpretation procedures through the use of GenAI paired with change detection models. As a result, Sonae MC increases its ability to swiftly respond to significant changes that might occur in business metrics.

In the aim of this internship, and for proof of concept, the execution of this project focused solely on one specific category “*Pão Tradicional*” (Traditional Bread) and one metric “*VLR Totais*” (Total Sales). The goal was to assess whether the selected change detection model could accurately flag changes, which then trigger the generation of a communication alert for the category director. This goal was achieved through the power of GenAI, which was employed to interpret the detected change, unveiling key retail insights. These insights are then integrated into a template communication report, which results in the final output of a formal communication draft ready to be sent out for the director of category.

The main strength of the project consists in its ability to provide timely alerts when significant changes are detected in data. Indeed, by automating the detection of important changes and communicating them directly to category directors, it equips them with the information they need to act quickly and make informed decisions. In the fast-paced retail market, this improvement enhances adaptability and offers a competitive advantage, as quicker responses to emerging trends reduce the risk of missed opportunities.

Moreover, the overall positive feedback from the teams which evaluated the project’s results emphasised the importance of such a system. Indeed, its usefulness was highly recognized, reinforcing the idea that automation is the future for business analytics. In addition, Ana Freitas, the Advanced Analytics & Insights Manager at Sonae MC, provided me with a recommendation letter, in which she emphasised the value of my contributions and underscored the significance of the work accomplished during my internship. For further details, one may refer to Section G of the appendix, where the full letter is included.

Considering the approval, there is now an open opportunity to include new product categories and metrics, making this internship’s project a stepping stone for broader retail applications. Overall, it laid a solid foundation for future development and proved that GenAI’s advanced capabilities combined with change detection hold a great potential for future implementations.

From an academic standpoint, the project was rooted in the CRISP-DM methodology which ensured that each stage of this project was meticulously planned and executed. The first methodological contribution was the integration of GPT-4 into a business context, as GenAI was leveraged to interpret data and generate retail insights from that transactional data. When combined with the change detection model PELT, this innovative approach provided deeper insights into business metrics by highlighting both the changes and the underlying reasons behind them. The use of PELT as an unsupervised change detection model allowed the project to analyse large datasets without requiring labelled data, making the process scalable and adaptable. This type of implementation advances the academic discourse on automating data interpretation procedures in the retail sector and offers a methodology that could be replicated or further developed in similar sectors.

While the project achieved its main goals, there were some limitations which emerged during its implementation. Therefore, for future improvements these challenges require full attention and should be addressed. One of the most significant limitations relates to the inherent risks associated with GenAI, specifically the possibility of it having hallucinations. Indeed, during the first phase of the project, these hallucinations were observed in instances where the model incorrectly identified the best and worst metrics, or produced suggestions that were not fully grounded in the actual data. As such, it was recognised that there would be a significant risk involved in relying on GenAI to retrieve exact values from the dashboards.

Additionally, the change detection model itself presented some limitations, particularly in its handling of non-cyclical seasonality. Indeed, despite the application of seasonal decomposition methods, some non-cyclical seasonality remained on the data. As a result, it was observed that the model sometimes interpreted these variances as changes that require attention, rather than identifying those fluctuations as non-cyclical seasonality. This misinterpretation would lead to the generation and communication of unnecessary alerts to the category director, which negatively impacts the system's efficiency. Therefore, further investigation into the effectiveness of the seasonal decomposition process is needed to refine the system and reduce false positives, ensuring that only meaningful changes are flagged.

Despite the identified limitations, the project has opened up several promising avenues for future research, as stated throughout this chapter. One of the most immediate areas for exploration is improving the handling of non-cyclical seasonality in the data. In fact, retail data is highly seasonal, and some false positives were observed in the historical data analysis,

when the change detection model misinterpreted residual seasonal patterns as significant changes. Future research could focus on enhancing the seasonal decomposition process to ensure that non-cyclical seasonality is properly removed before the data is analysed by the change detection model. Additionally, refining the PELT algorithm or exploring other advanced models capable of distinguishing between residual seasonal patterns and genuine structural changes could further enhance the system's precision.

Another important area of research would be addressing the limitations of GenAI, specifically reducing the risks of hallucinations. While GPT-4 has demonstrated its ability to provide insightful recommendations, there is a potential risk of it producing inaccurate information. It is essential to keep this risk in check by monitoring the outputs carefully. Therefore, future research could focus on improving the prompt engineering techniques to better control the output, ensuring that it stays aligned with the factual data and the relevant context. This way, when analysing detected changes with GPT-4, the accuracy and depth of the insights could be further enhanced.

Another key area for future exploration would be scaling this project to cover more metrics and categories. Sonae MC has a vast category range meaning that there is an opportunity to further develop this system. Further research could explore how to adapt the change detection model to the different metrics and categories, as some refinement will most likely be needed to ensure the system maintains its performance.

Lastly, another interesting research direction is the exploration of real-time change detection. Indeed, in the scope of this internship, only offline models were considered. However, implementing online change detection algorithms, which analyse data as it is generated, would allow for even faster decision-making and give Sonae MC a competitive edge. Research into reducing the latency of online algorithms, while maintaining accuracy, could greatly enhance the system's responsiveness.

In conclusion, this project has laid a solid foundation for automating data interpretation in retail using GenAI and change detection models. By addressing the limitations identified and expanding its scope, future implementations could unlock even greater potential for Sonae MC.

References

- Ahmed, M., Choudhury, N., & Uddin, S. (2017). Anomaly detection on big data in financial markets. In *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017 - ASONAM '17*(pp. 998-1001). New York, NY, USA: ACM Press. <https://doi.org/10.1145/3110025.3119402>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Aminikhanghahi, S., & Cook, D. J. (2017). A survey of methods for time series change point detection. *Knowledge and Information Systems*, 51(2), 339–367. <https://doi.org/10.1007/s10115-016-0987-z>
- Aminikhanghahi, S., Wang, T., & Cook, D. J. (2018). Real-time change point detection with application to smart home time series data. *IEEE Transactions on Knowledge and Data Engineering*, 31(5), 1010-1023. <https://doi.org/10.1109/TKDE.2018.2850347>
- AnswerRocket. (2023). How Cereal Partners Worldwide Used AI Augmented Analytics to Unlock Insights. Retrieved [January 17, 2024], from <https://answerrocket.com/customers/cpw-case-study/>
- AnswerRocket. (n.d.) About. Retrieved [January 17, 2024], from <https://answerrocket.com/about/>
- Arlot, S., Celisse, A., & Harchaoui, Z. (2019). A kernel multiple change-point algorithm via model selection. *Journal of Machine Learning Research*, 20(162), 1–56.
- Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory*, 13(3), 315–352.
- Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33(1), 63. <https://doi.org/10.1007/s12525-023-00680-1>
- Barley, S. R. (2020). *Work and Technological Change*. Oxford University Press.
- Basseville, M., & Nikiforov, I. V. (1993). *Detection of abrupt changes: Theory and applications*. Eaglewood Cliffs, NJ: Prentice Hall.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bhatt, S., Fang, G., & Li, P. (2022). Proceedings of the Thirty-Eighth Conference on Uncertainty in Artificial Intelligence. *Proceedings of Machine Learning Research*, 180, 191-201.
- Brevini, B. (2021). *Is AI Good for the Planet?* John Wiley & Sons.
- Chaban, M. A. V. (2023). Going from repeated orders to a side of AI speed, Wendy's is reinventing the drive-thru. Retrieved [January 4, 2024], from <https://cloud.google.com/blog/transform/wendys-generative-ai-drive-thru-reinvention-worker-freedom>
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative Adversarial Networks: An Overview. *IEEE Signal Processing Magazine*, 35(1), 53–65. <https://doi.org/10.1109/MSP.2017.2765202>

- Danaher, J., & Sætra, H. S. (2022). Technology and moral change: The transformation of truth and trust. *Ethics and Information Technology*, 24(3), 35. <https://doi.org/10.1007/s10676-022-09661-y>
- Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., Duan, Y., Dwivedi, R., Edwards, J., Eirug, A., Galanos, V., Ilavarasan, P. V., Janssen, M., Jones, P., Kar, A. K., Kizgin, H., Kronemann, B., Lal, B., Lucini, B., ... Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, 57, 101994. <https://doi.org/10.1016/j.ijinfomgt.2019.08.002>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). Opinion Paper: “So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Else, H. (2023). Abstracts written by ChatGPT fool scientists. *Nature*, 613(7944), 423–423. <https://doi.org/10.1038/d41586-023-00056-7>
- Eschliman, A. (2023). Generative AI for retail: How to keep pace and get ahead. Google Cloud Blog. Retrieved [January 22, 2024], from <https://cloud.google.com/blog/topics/retail/top-gen-ai-use-cases-for-retail-executives>
- Faithfull, W. (2018). Unsupervised change detection in multivariate streaming data. Retrieved from <https://doi.org/10.13140/RG.2.2.25121.66409>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2023). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Feuz, K. D., Cook, D. J., Rosasco, C., Robertson, K., & Schmitter-Edgecombe, M. (2014). Automated detection of activity transitions for prompting. *IEEE Transactions on Human-Machine Systems*, 45(5), 1–11.
- Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, 42(6), 2243–2281.
- Gachomo, D. W., Gichuhi, A. W., & Wanjoya, A. (2015). The power of the pruned exact linear time (PELT) test in multiple changepoint detection. *American Journal of Theoretical and Applied Statistics*, 4(6), 581–586. <https://doi.org/10.11648/j.ajtas.20150406.30>
- Garreau, D., & Arlot, S. (2018). Consistent change-point detection with kernels. *Electronic Journal of Statistics*, 12(2), 4440–4486.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
- Google Cloud. (2023a). The executive’s guide to generative AI. Retrieved [January 22, 2024], from https://services.google.com/fh/files/misc/exec_guide_gen_ai.pdf

- Google Cloud. (2023b). Wendy's Taps Google Cloud to Revolutionize the Drive-Thru Experience with Artificial Intelligence. Retrieved [January 9, 2024], from <https://www.googlecloudpresscorner.com/2023-05-09-Wendys-Taps-Google-Cloud-to-Revolutionize-the-Drive-Thru-Experience-with-Artificial-Intelligence>
- Google Cloud. (2023c). The Estée Lauder Companies Inc. and Google Cloud Partner to Transform the Online Consumer Experience with Generative AI. Retrieved [January 18, 2024], from <https://www.googlecloudpresscorner.com/2023-08-29-The-Estee-Lauder-Companies-Inc-and-Google-Cloud-Partner-to-Transform-the-Online-Consumer-Experience-with-Generative-AI>
- Google Cloud. (2024). What is Vertex AI Search and Conversation? Retrieved [January 18, 2024], from <https://cloud.google.com/generative-ai-app-builder/docs/introduction>
- Guo, S., Guo, W., Abolhassani, A., Kalamdani, R., Puchala, S., Januszczak, A., et al. (2019). Manufacturing process monitoring with nonparametric change-point detection in automotive industry. *Journal of Manufacturing Science and Engineering*, 141(7), Article 071013.
- Gupta, M., Wadhvani, R., & Rasool, A. (2024). Comprehensive analysis of change-point dynamics detection in time series data: A review. *Expert Systems with Applications*, 227, 123342. <https://doi.org/10.1016/j.eswa.2024.123342>
- Hawkins, M. D., & Olwell, D. H. (1998). *Cumulative sum charts and charting for quality improvement (Statistics for engineering and physical science)*. Berlin/Heidelberg: Springer Verlag.
- Huang, D. T. (2015). Change mining and analysis for data streams. Retrieved [May 17, 2024] from <https://researchspace.auckland.ac.nz/handle/2292/27746>
- IBM. (2024). CRISP-DM help overview. Retrieved [January 22, 2024], from <https://www.ibm.com/docs/en/spss-modeler/18.5.0?topic=dm-crisp-help-overview>
- Iskender, A. (2023). Holy or Unholy? Interview with Open AI's ChatGPT. *European Journal of Tourism Research*, 34, 3414. <https://doi.org/10.54055/ejtr.v34i.3169>
- Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3), 685–695. <https://doi.org/10.1007/s12525-021-00475-2>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Khamesi, V. (2022). ocpdet: A Python package for online changepoint detection in univariate and multivariate data (Version v0.0.5). Zenodo. <https://doi.org/10.5281/zenodo.7632721>.
- Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107(500), 1590–1598. <https://doi.org/10.1080/01621459.2012.737745>
- Killick, R., Jonathan, P., & Eckley, I. A. (2011). Efficient detection of multiple changepoints within an oceanographic time series. *Department of Mathematics & Statistics, Lancaster University*. Retrieved [May 20, 2024] from https://www.lancaster.ac.uk/people/jonathan/2011_ISI_ChangePoints_Talk.pdf

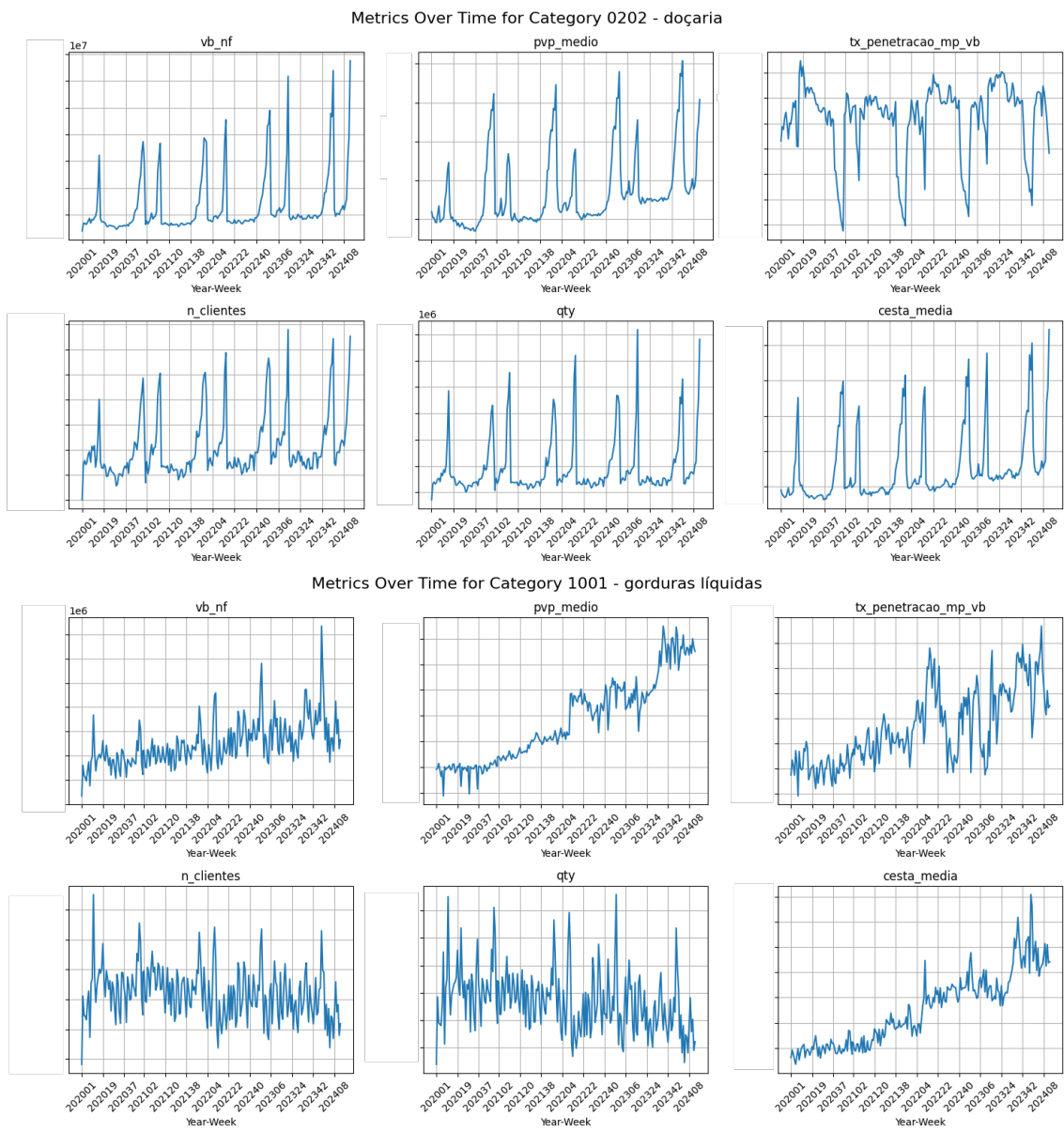
- Kumar, A., Gupta, N., & Bapat, G. (2023). Who is making the decisions? How retail managers can use the power of ChatGPT. *Journal of Business Strategy*. <https://doi.org/10.1108/JBS-04-2023-0067>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Li, E., Zeng, L., Zhou, Z., & Chen, X. (2019). Edge AI: On-demand accelerating deep neural network inference via edge computing. *IEEE Transactions on Wireless Communications*, 19(1), 447–457. <https://doi.org/10.1109/TWC.2019.2954577>
- Li, M., Bao, X., Chang, L., & Gu, T. (2022). Modeling personalized representation for within-basket recommendation based on deep learning. *Expert Systems with Applications*, 192, 116383. <https://doi.org/10.1016/j.eswa.2021.116383>
- Liu, S., Wright, A., & Hauskrecht, M. (2018). Change-point detection method for clinical decision support system rule monitoring. *Artificial Intelligence in Medicine*, 91, 49–56.
- Martinez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernandez-Orallo, J., Kull, M., Lachiche, N., Ramirez-Quintana, M. J., & Flach, P. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8), 3048–3061. <https://doi.org/10.1109/TKDE.2019.2962680>
- Miller, S. (2018). AI: Augmentation, more so than automation. *Asian Management Insights*, 5(1), 1–20.
- Montanez, G. D., Amizadeh, S., & Laptev, N. (2015). Inertial hidden Markov models: Modeling change in multivariate time series. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*. Retrieved from <https://cdn.aaai.org/ojs/9457/9457-13-12985-1-2-20201228.pdf>
- Ooi, K.-B., Tan, G. W.-H., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., Dwivedi, Y. K., Huang, T.-L., Kar, A. K., Lee, V.-H., Loh, X.-M., Micu, A., Mikalef, P., Mogaji, E., Pandey, N., Raman, R., Rana, N. P., Sarker, P., Sharma, A., ... Wong, L.-W. (2023). The Potential of Generative Artificial Intelligence Across Disciplines: Perspectives and Future Directions. *Journal of Computer Information Systems*, 1–32. <https://doi.org/10.1080/08874417.2023.2261010>
- OpenAI. (2021). DALL·E: Creating images from text. Retrieved January 20, 2024, from <https://openai.com/research/dall-e>
- OpenAI. (2023). Introducing ChatGPT Plus. Retrieved [January 4, 2024], from <https://openai.com/blog/chatgpt-plus>
- OpenAI. (n.d.). Prompt engineering. In OpenAI API Documentation. Retrieved [May 19, 2024] from <https://platform.openai.com/docs/guides/prompt-engineering>
- Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, 41(1-2), 100–115. <https://doi.org/10.1093/biomet/41.1-2.100>
- Pástor, L., & Stambaugh, R. F. (2001). The equity premium and structural breaks. *The Journal of Finance*, 56(4), 1207–1239. <https://doi.org/10.1111/0022-1082.00365>
- Patil, D. D., Dhotre, D. R., Gawande, G. S., Mate, D. S., Shelke, M. V., & Bhoje, T. S. (2023). Transformative Trends in Generative AI: Harnessing Large Language Models for Natural Language Understanding and Generation. *International Journal of Intelligent Systems and Applications in Engineering*, 12(4s), 309–319.

- Pourhabibi, T., Ong, K.-L., Kam, B. H., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems*, 133, Article 113303.
- Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). Improving language understanding by generative pre-training. Preprint. pp. 1–12. Retrieved [January 20, 2024], from <https://openai.com/research/language-unsupervised>
- Radke, R. J., Andra, S., Al-Kofahi, O., & Roysam, B. (2005). Image change detection algorithms: A systematic survey. *IEEE Transactions on Image Processing*, 14(3), 294–307. <https://doi.org/10.1109/TIP.2005.843255>
- Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121–154. <https://doi.org/10.1016/j.iotcps.2023.04.003>
- Ray, S. (2019). A quick review of machine learning algorithms. In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)* (pp. 35-39). IEEE. <https://doi.org/10.1109/COMITCon.2019.8862451>
- Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., & Srivastava, M. (2010). Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks*, 6(2), 1–27.
- Roddick, J. F., Al-Jadir, L., Bertossi, L., Dumas, M., Estrella, F., Gregersen, H., Hornsby, K., Lufter, J., Mandreoli, F., Minnist, T., Mayol, E., & Wedemeijer, L. (2000). Evolution and change in data management - Issues and directions. Proceedings of the 14th Australasian Database Conference. <https://dl.acm.org/doi/pdf/10.1145/344788.344789>
- Rupp, C. (2024). Five ways Victoria's Secret & Co. Is using AI to make shopping more personalized and inclusive. Retrieved [January 17, 2024], from <https://blog.google/products/google-cloud/victorias-secret-google-cloud-ai/>
- Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (Third edition, Global edition). Pearson.
- Sætra, H. S. (2019). The Ghost in the Machine: Being Human in the Age of AI and Machine Learning. *Human Arenas*, 2(1), 60–78. <https://doi.org/10.1007/s42087-018-0039-1>
- Sætra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society*, 75, 102372. <https://doi.org/10.1016/j.techsoc.2023.102372>
- Sætra, H. S., & Mills, S. (2022). Psychological interference, liberty and technology. *Technology in Society*, 69, 101973. <https://doi.org/10.1016/j.techsoc.2022.101973>
- Sætra, H. S., Borgebund, H., & Coeckelbergh, M. (2022). Avoid diluting democracy by algorithms. *Nature Machine Intelligence*, 4, 804–806. <https://doi.org/10.1038/s42256-022-00537-w>
- Saltz, J. S. (2021). CRISP-DM for Data Science: Strengths, Weaknesses and Potential Next Steps. *2021 IEEE International Conference on Big Data (Big Data)*, 2337–2344. <https://doi.org/10.1109/BigData52589.2021.9671634>
- Sarker, I. H. (2021). Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(6), 420. <https://doi.org/10.1007/s42979-021-00815-1>

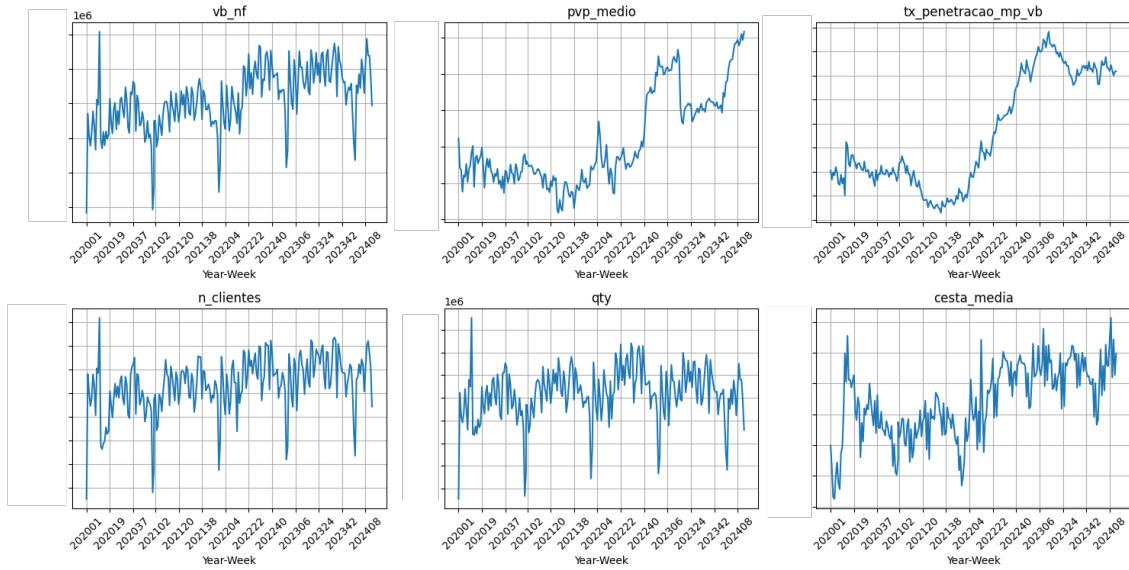
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). A Systematic Literature Review on Applying CRISP-DM Process Model. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2), 461–464.
- Shrestha, Y. R., Krishna, V., & von Krogh, G. (2021). Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges. *Journal of Business Research*, 123, 588–603. <https://doi.org/10.1016/j.jbusres.2020.09.068>
- Sonae. (2020). Sonae in the World. Retrieved [December 18, 2023], from <https://www.sonae.pt/en/sonae/where-we-are-in-the-world/>
- SonaeMC. (2022a). História. Retrieved [December 18, 2023], from <https://mc.sonae.pt/historia/>
- SonaeMC. (2022b). Negócios. Retrieved [December 18, 2023], from <https://mc.sonae.pt/negocios/>
- Stokel-Walker, C. J. N. (2023). ChatGPT listed as author on research papers: many scientists disapprove. *Nature*, 613, 620–621. <https://doi.org/10.1038/d41586-023-00107-z>
- Tran, D.-H. (2013). Automated change detection and reactive clustering in multivariate streaming data. Retrieved from <http://arxiv.org/abs/1311.0505>
- Tran, D.-H., Gaber, M. M., & Sattler, K.-U. (2014). Change detection in streaming data in the era of big data. *ACM SIGKDD Explorations Newsletter*, 16(1), 30–38. <https://doi.org/10.1145/2674026.2674031>
- Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, 167, 107299.
- Uc-Cetina, V., Navarro-Guerrero, N., Martín-Gonzalez, A., Weber, C., & Wermter, S. (2022). Survey on reinforcement learning for language processing. *Artificial Intelligence Review*, 1–33. <https://doi.org/10.1007/s10462-022-10205-5>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is All you Need*.
- Wang, L., Chen, X., Deng, X., Wen, H., You, M., Liu, W., Li, Q., & Li, J. (2024). Prompt engineering in consistency and reliability with the evidence-based guideline for LLMs. *NPJ Digital Medicine*. <https://doi.org/10.1038/s41746-024-01029-4>
- Widder, D. G., West, S., & Whittaker, M. (2023). Open (for business): Big tech, concentrated power, and the political economy of open AI. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4543807>
- Zhang, D., Li, W., Niu, B., & Wu, C. (2023). A deep learning approach for detecting fake reviewers: Exploiting reviewing behavior and textual information. *Decision Support Systems*, 166, 113911. <https://doi.org/10.1016/j.dss.2022.113911>
- Zheng, Y., Liu, L., Wang, L., & Xie, X. (2008). Learning transportation mode from raw GPS data for geographic applications on the web. In *Proceedings of the 17th international conference on World Wide Web (WWW '08)* (pp. 247). ACM.

Appendix

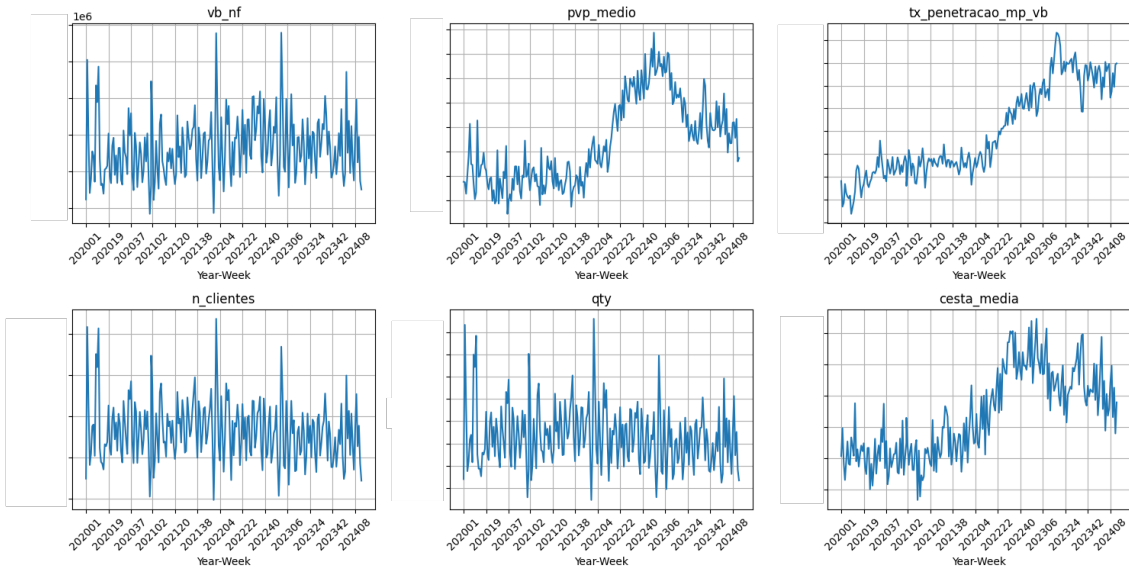
Section A - Metrics Over Time



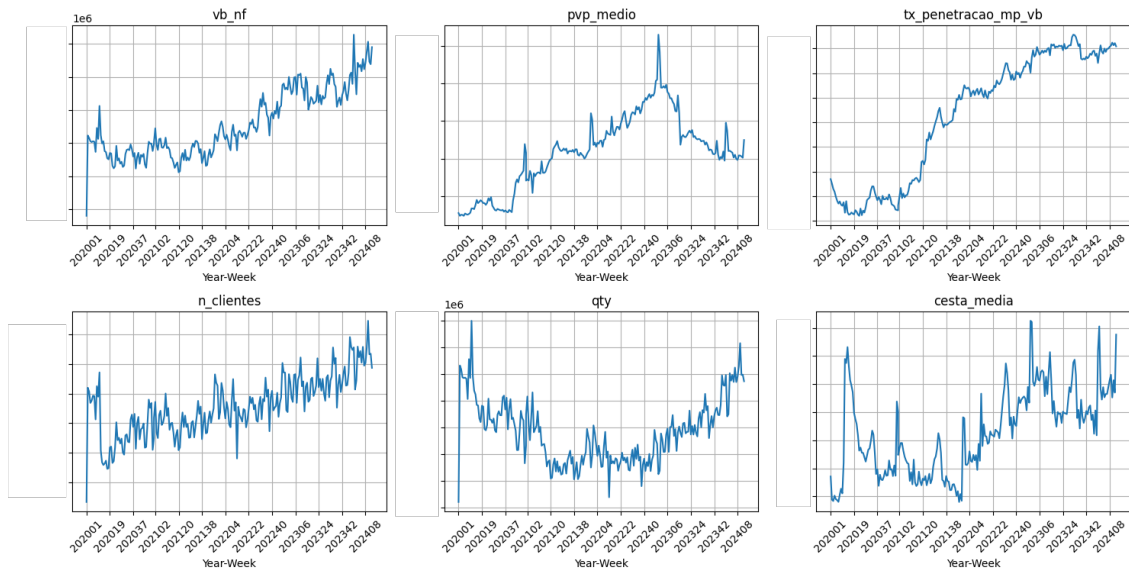
Metrics Over Time for Category 0804 - iogurtes e sobrem.



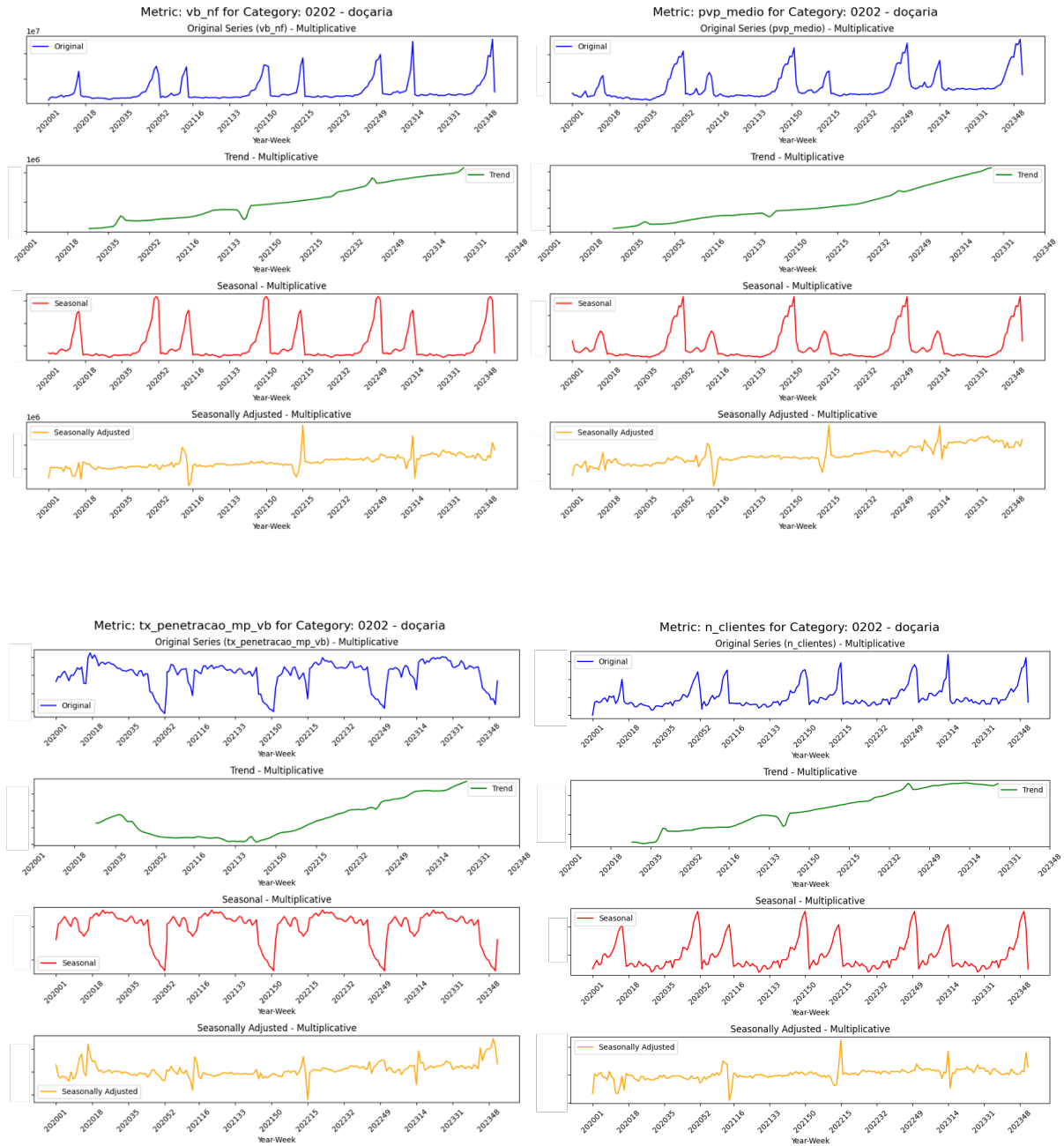
Metrics Over Time for Category 0601 - limp. e tratam roupa

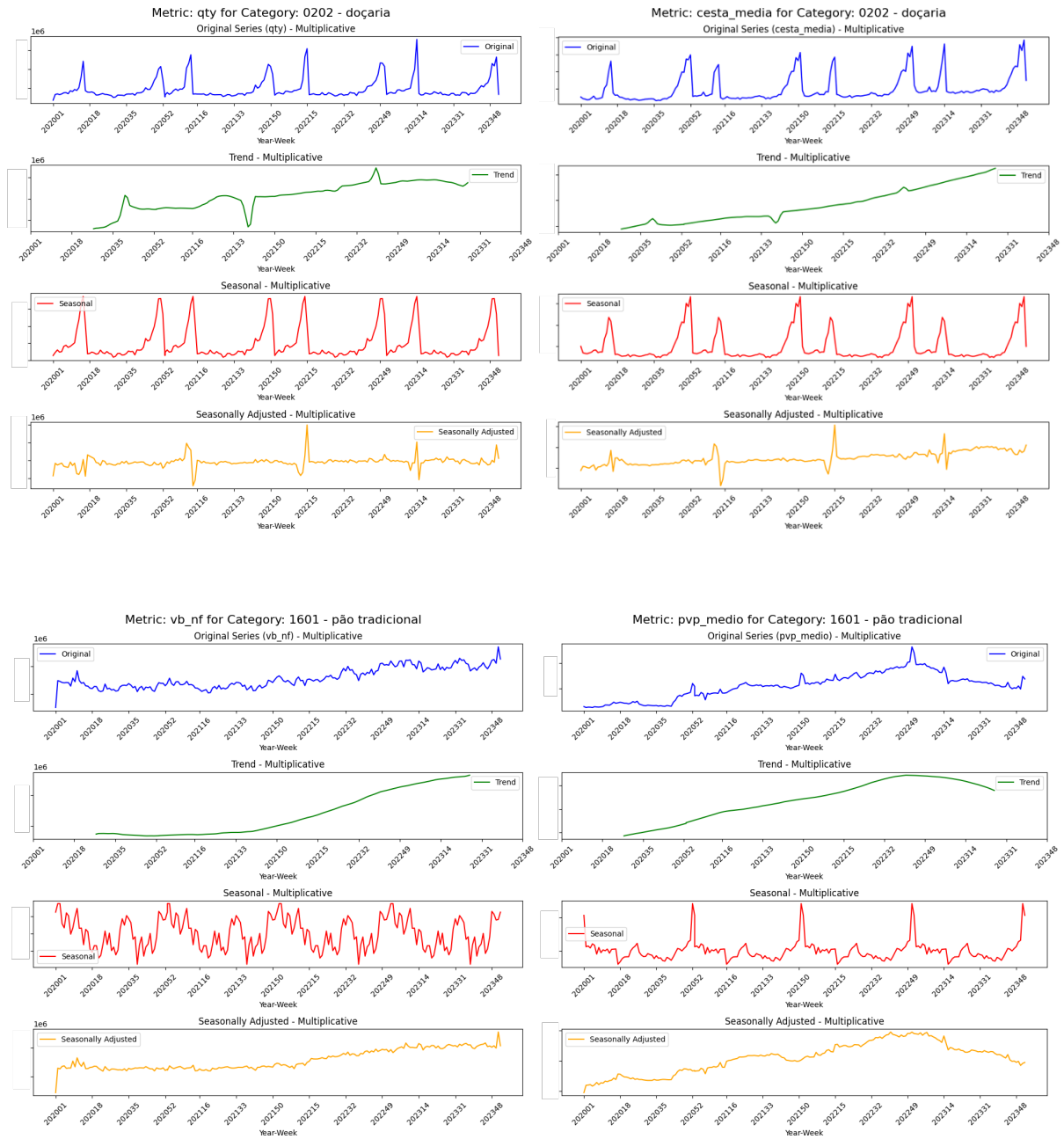


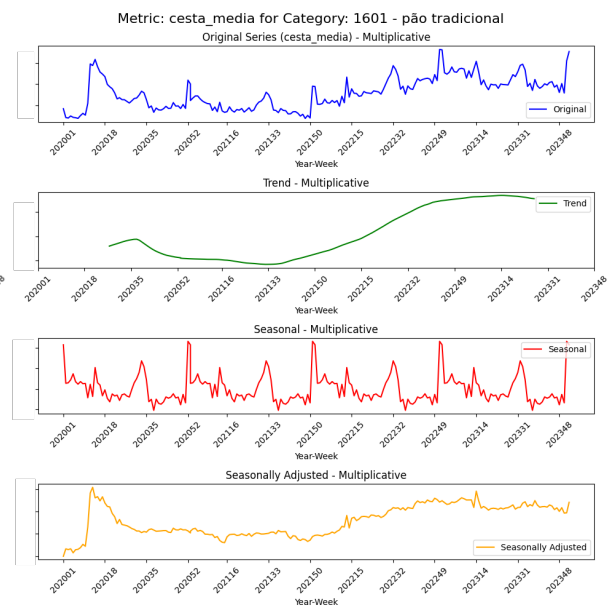
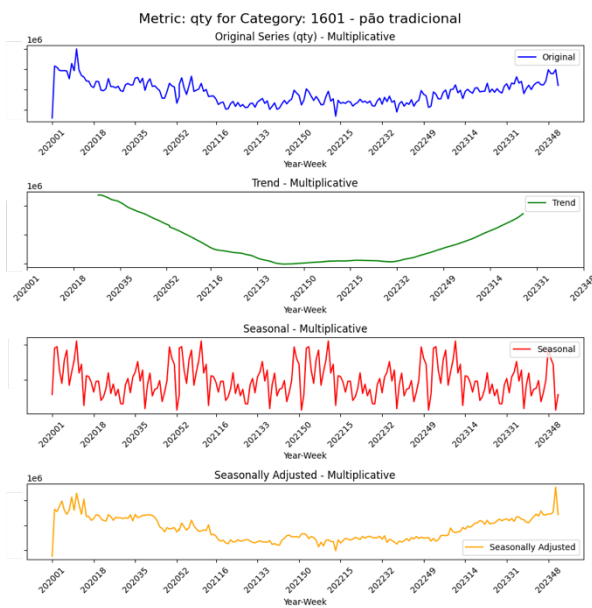
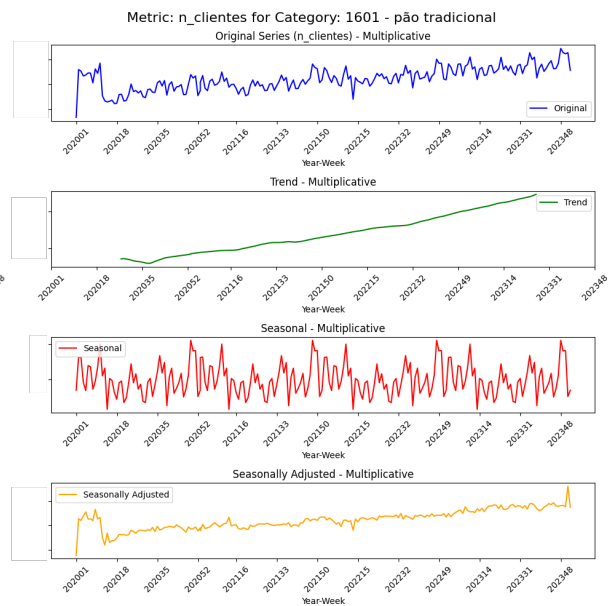
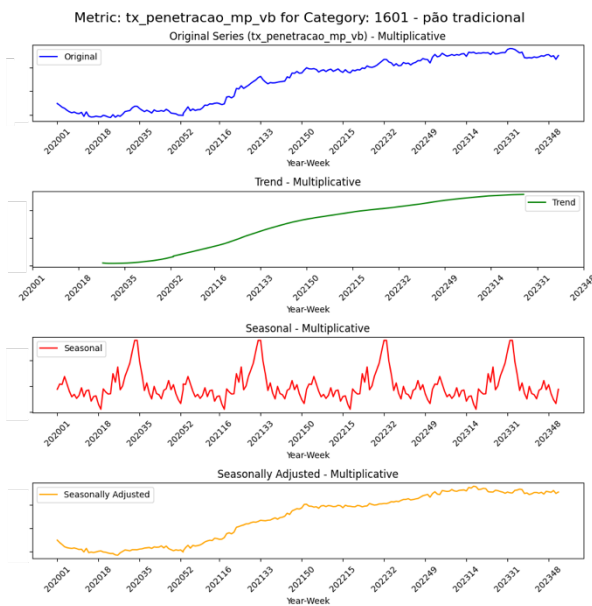
Metrics Over Time for Category 1601 - pão tradicional

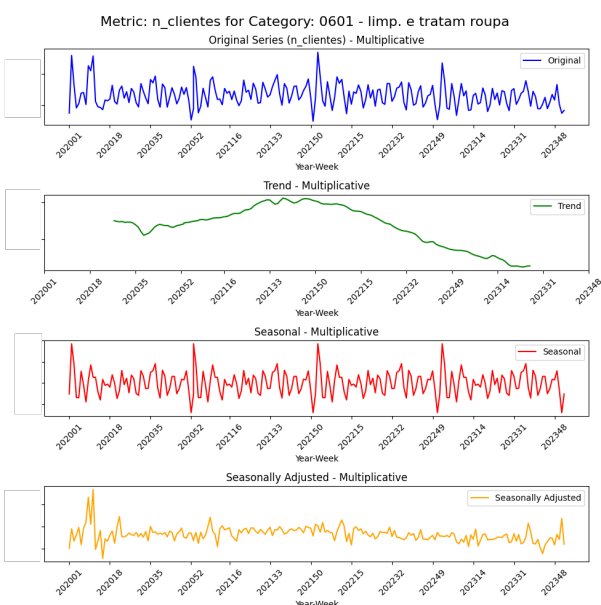
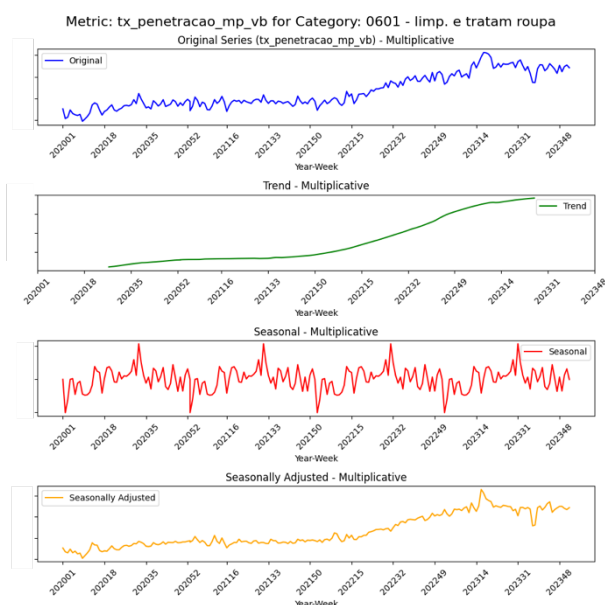
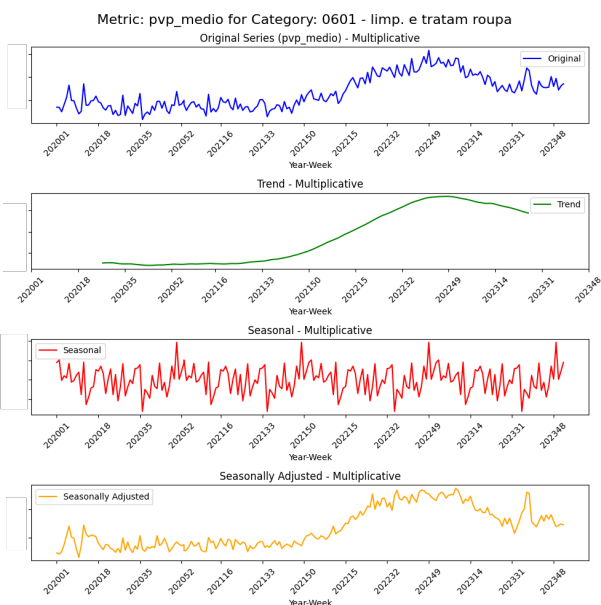
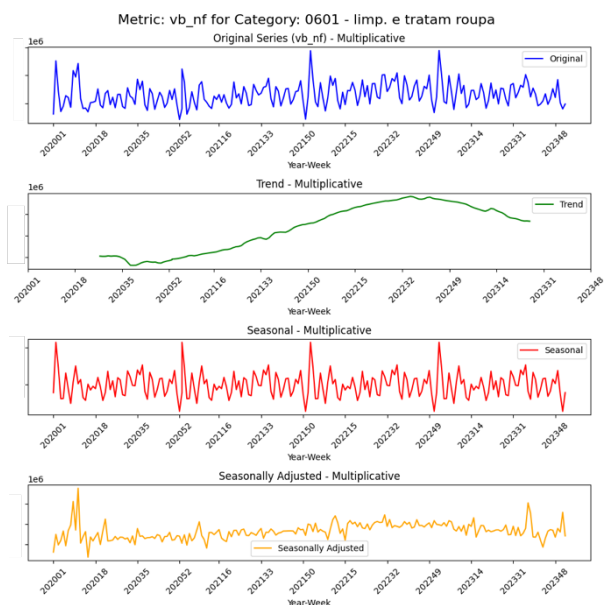


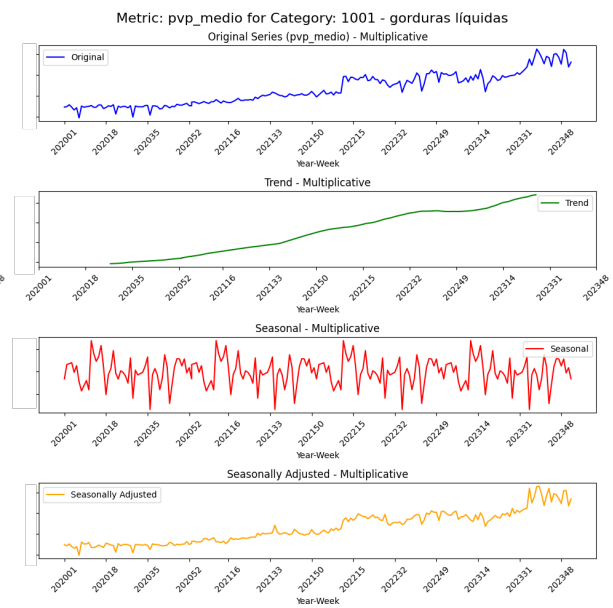
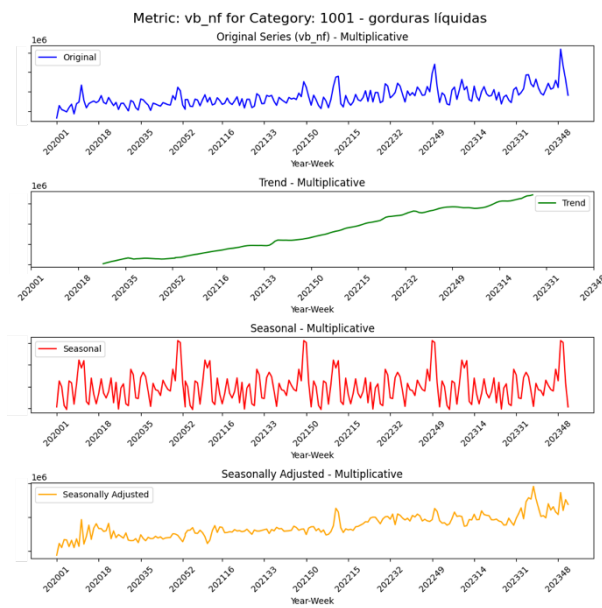
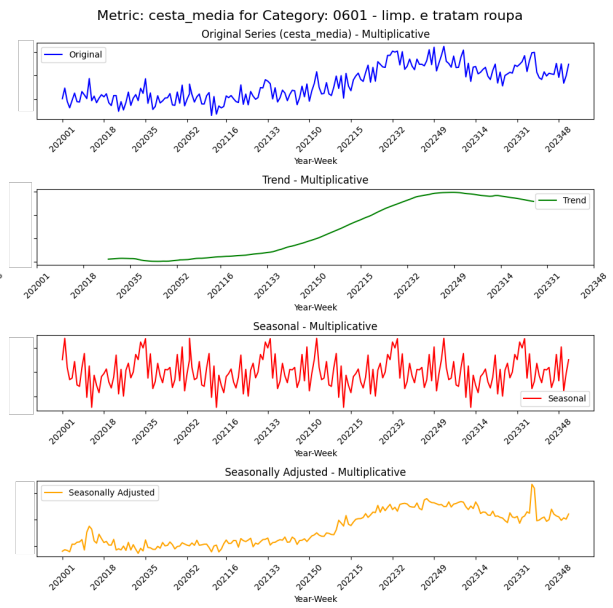
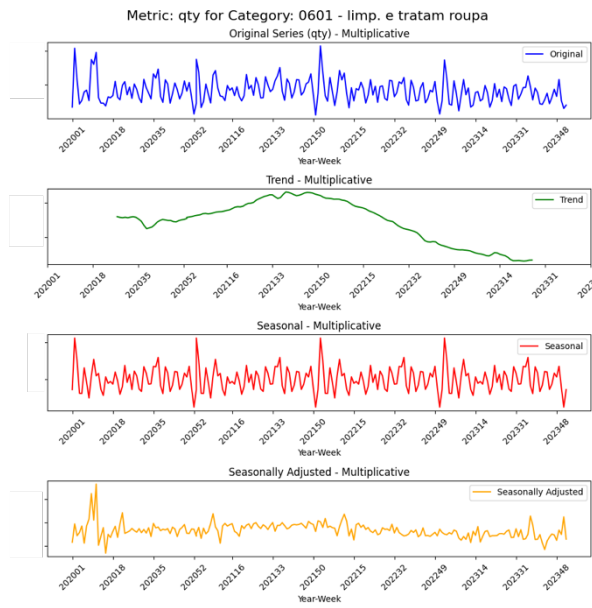
Section B - Seasonal Decomposition

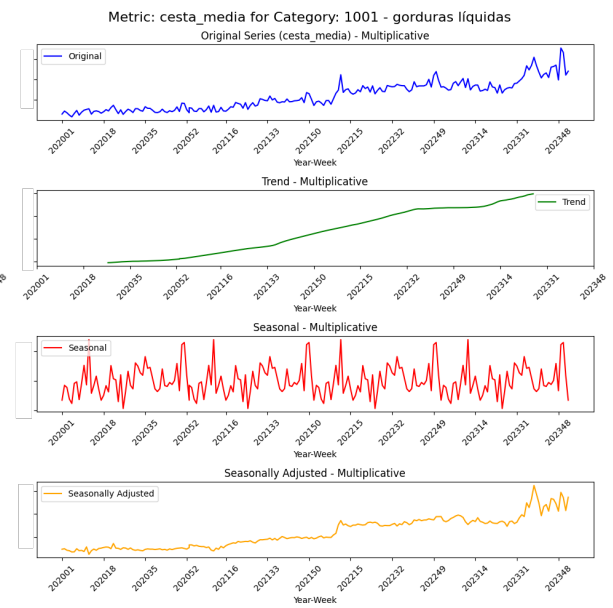
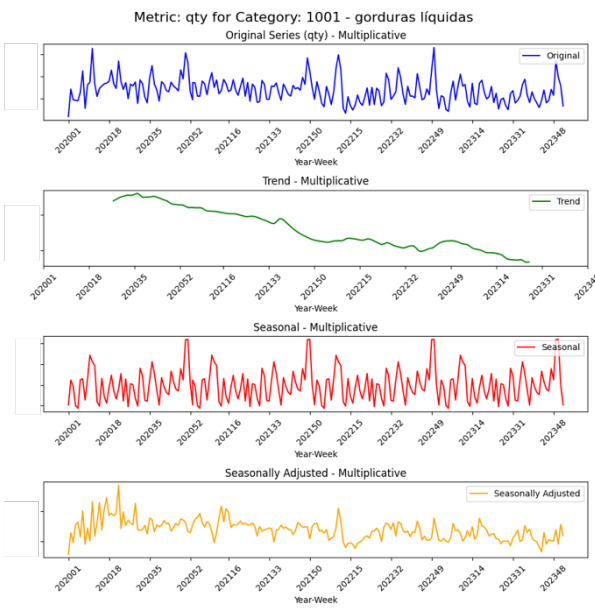
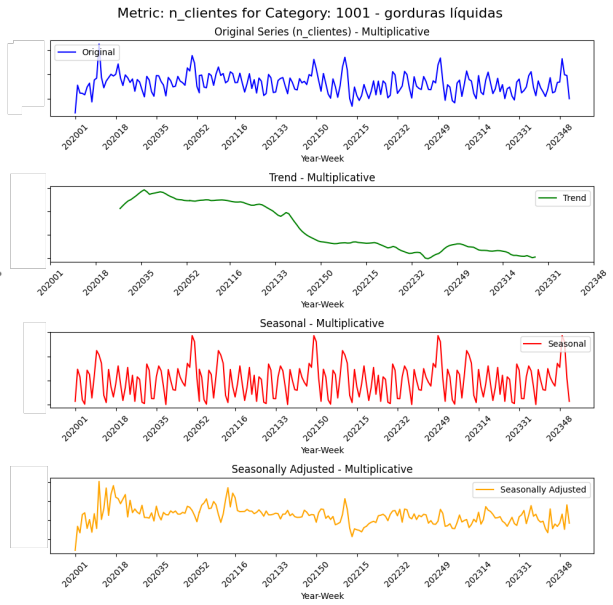
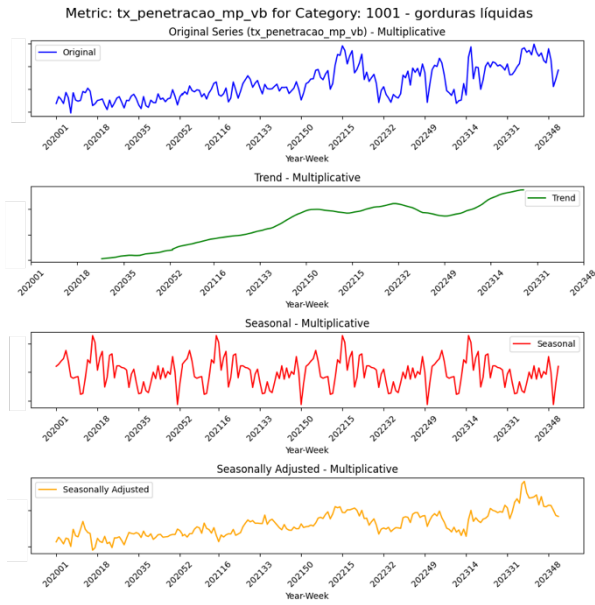


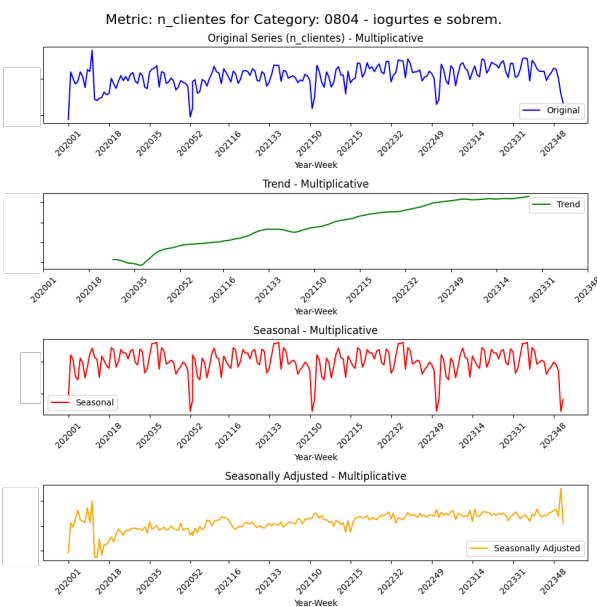
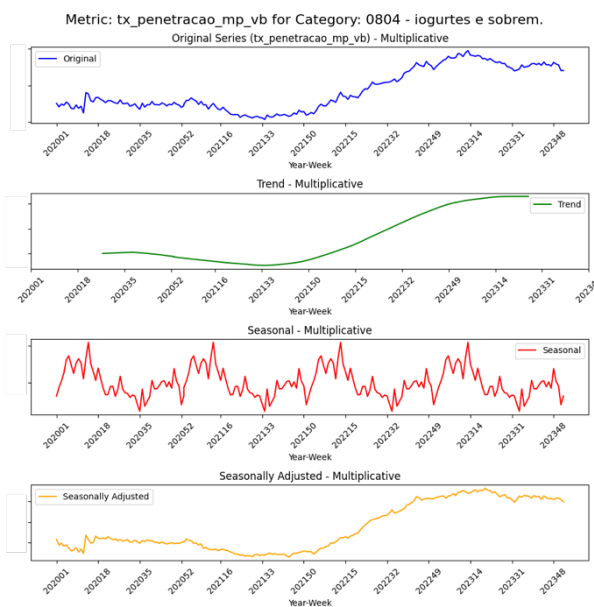
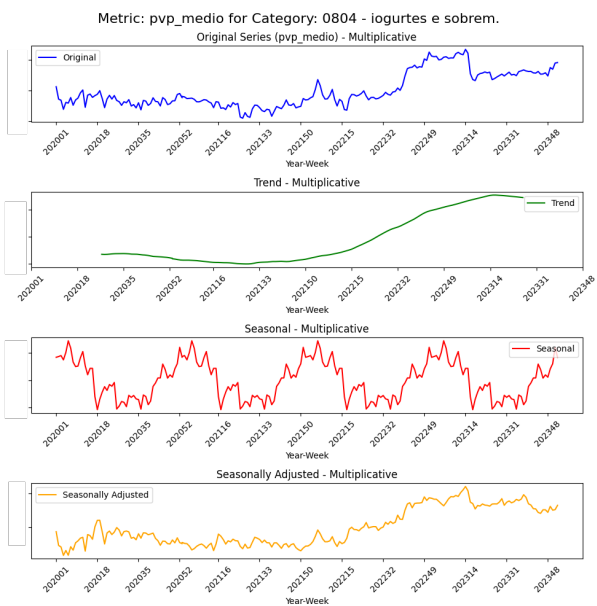
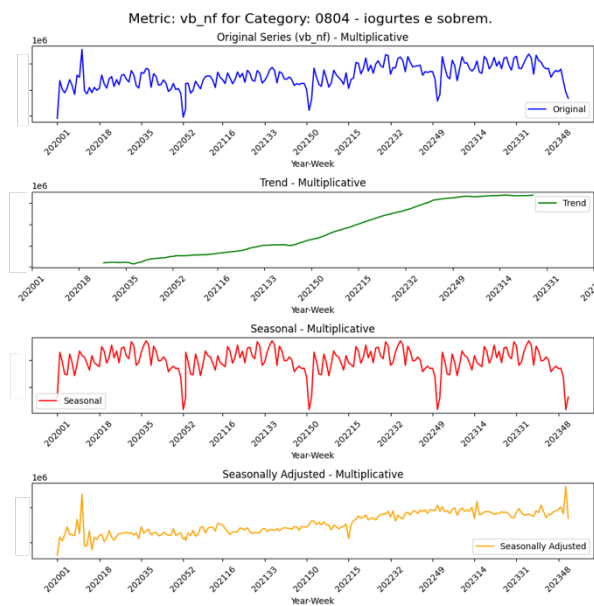


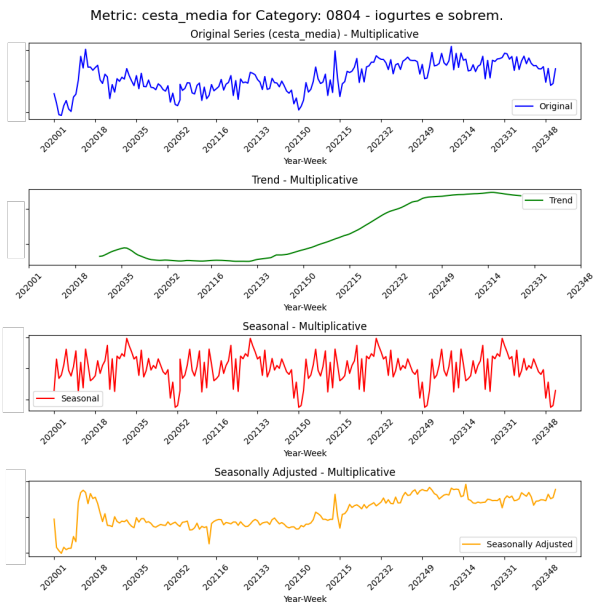
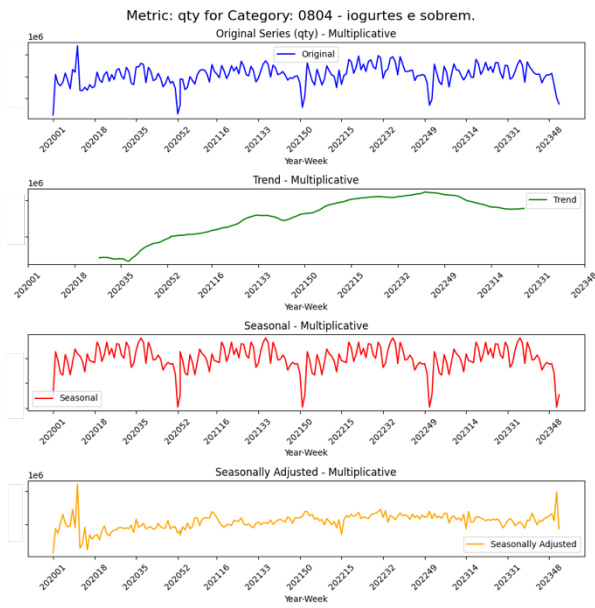




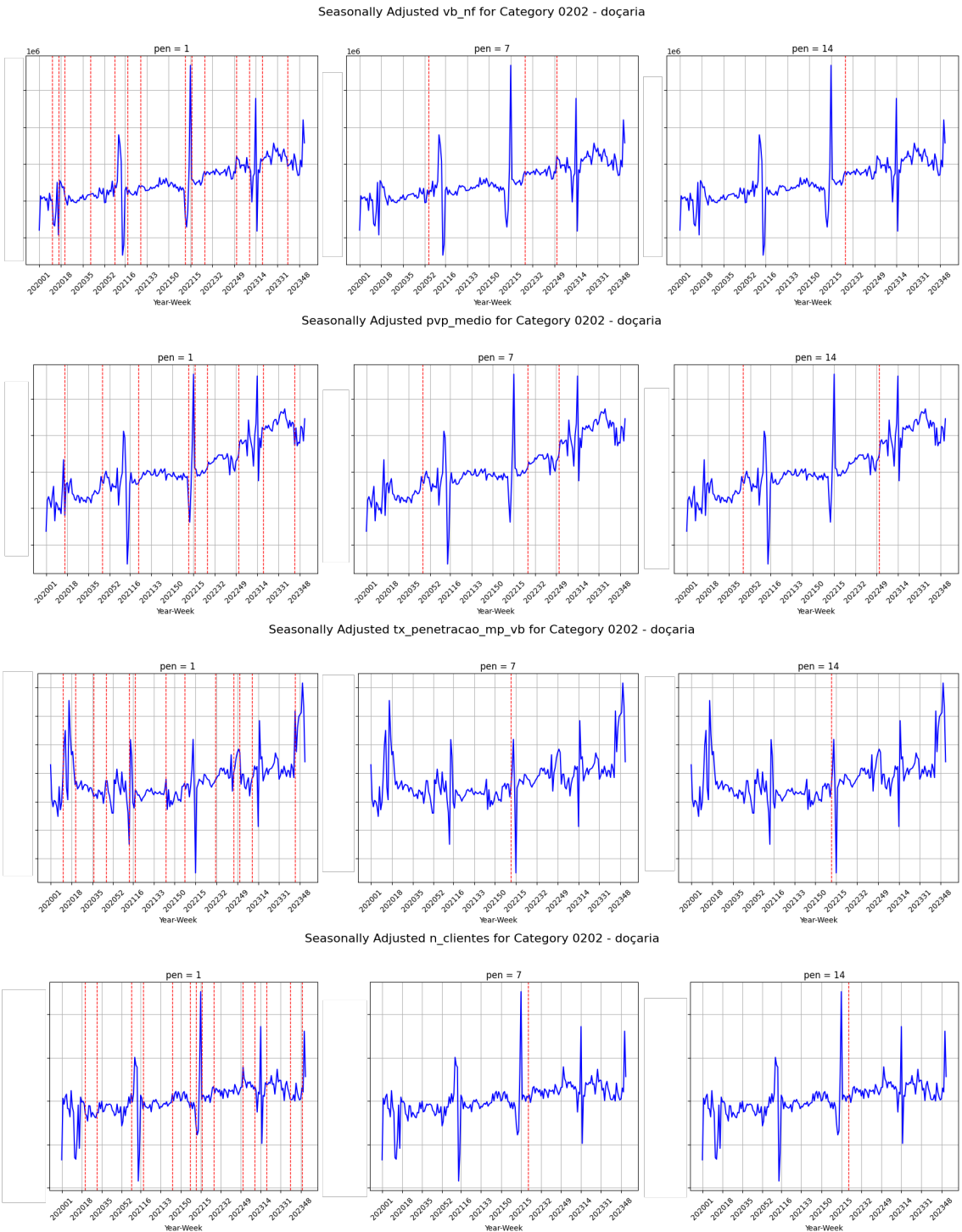




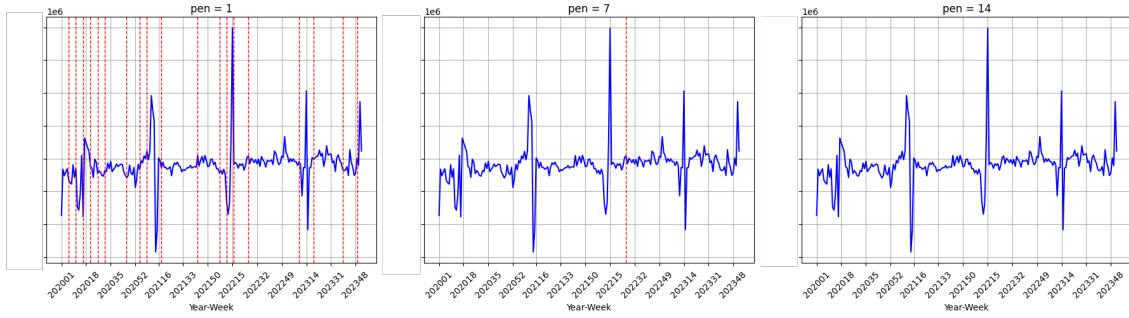




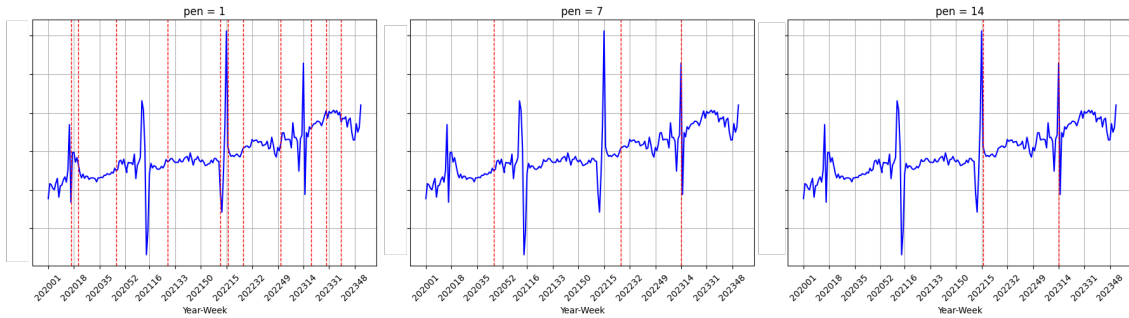
Section C - PELT Results



Seasonally Adjusted qty for Category 0202 - doçaria



Seasonally Adjusted cesta_media for Category 0202 - doçaria



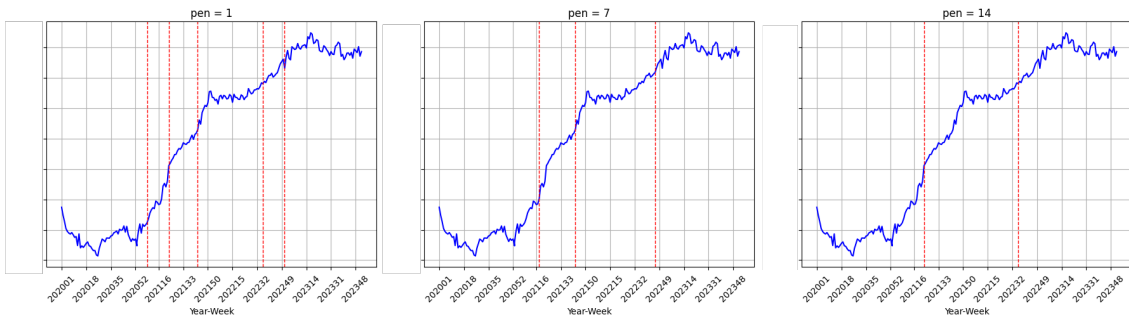
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional



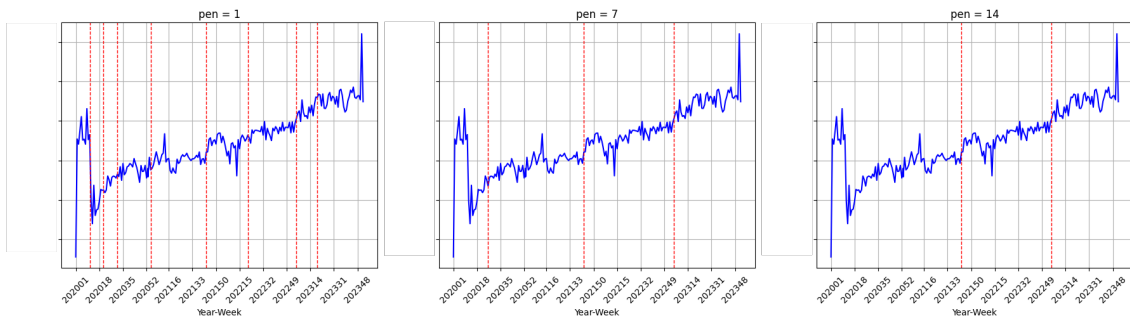
Seasonally Adjusted pvp_medio for Category 1601 - pão tradicional



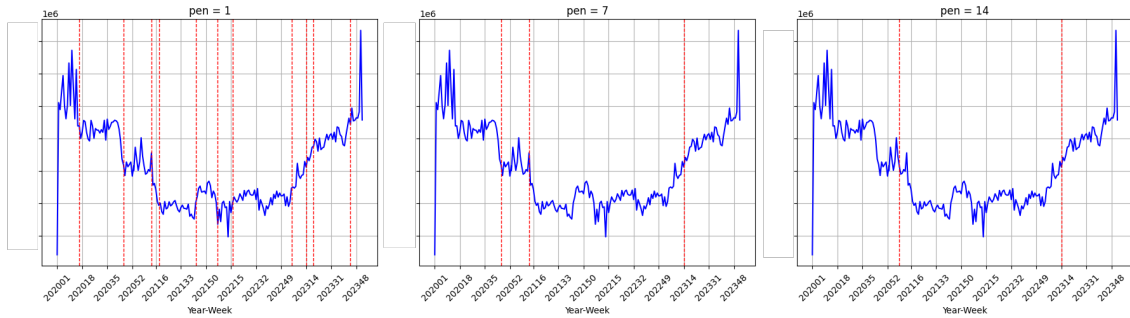
Seasonally Adjusted tx_penetracao_mp_vb for Category 1601 - pão tradicional



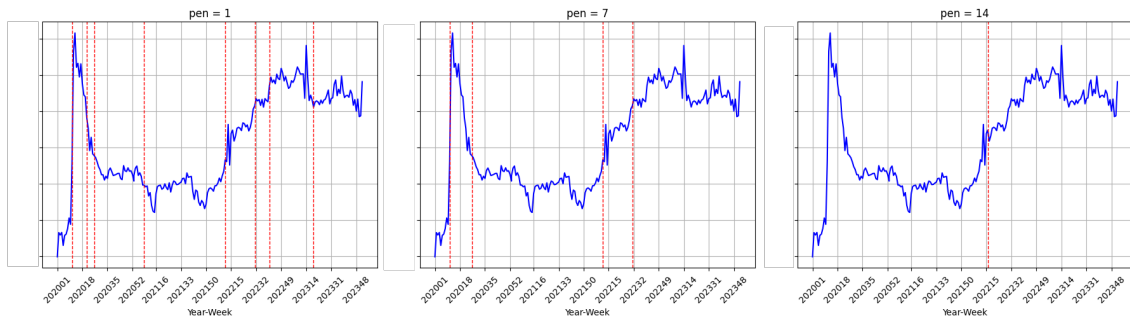
Seasonally Adjusted n_clientes for Category 1601 - pão tradicional



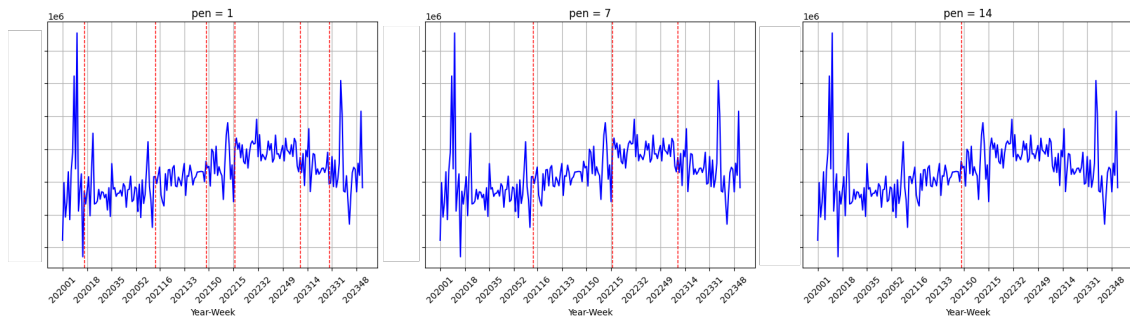
Seasonally Adjusted qty for Category 1601 - pão tradicional



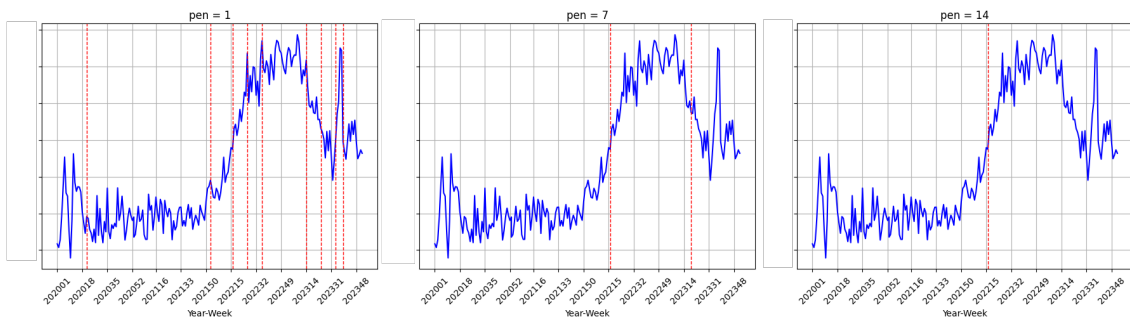
Seasonally Adjusted cesta_media for Category 1601 - pão tradicional



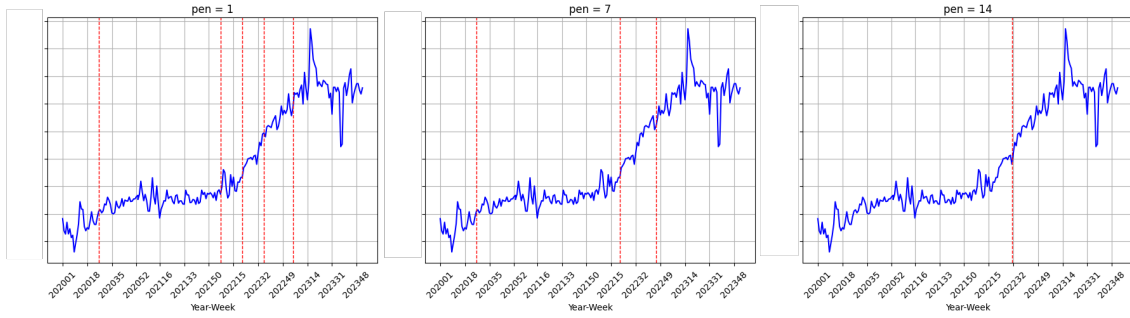
Seasonally Adjusted vb_nf for Category 0601 - limp. e tratam roupa



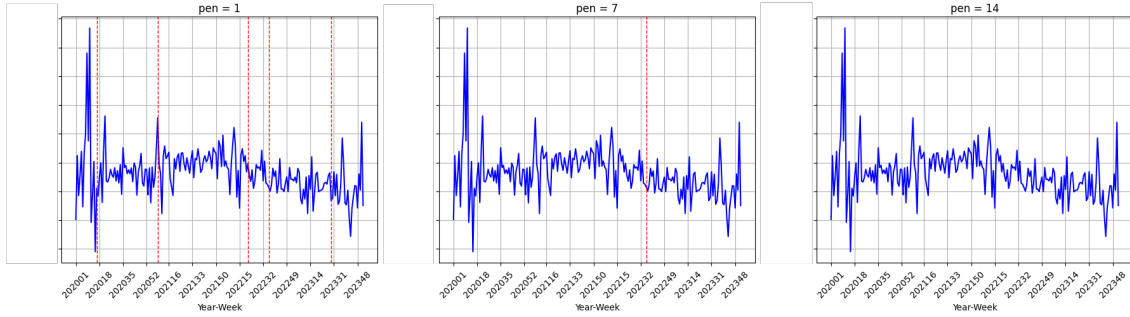
Seasonally Adjusted pvp_medio for Category 0601 - limp. e tratam roupa



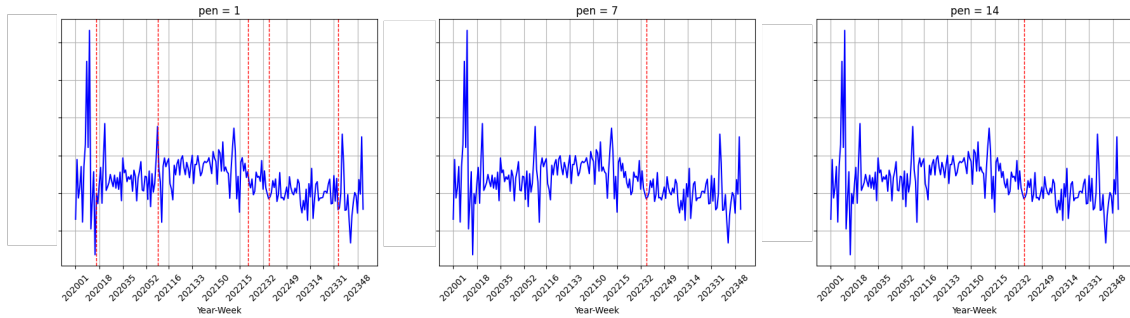
Seasonally Adjusted tx_penetracao_mp_vb for Category 0601 - limp. e tratam roupa



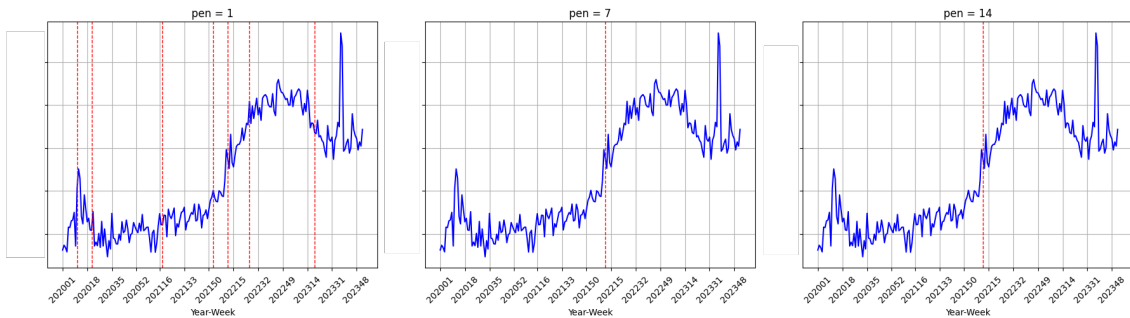
Seasonally Adjusted n_clientes for Category 0601 - limp. e tratam roupa



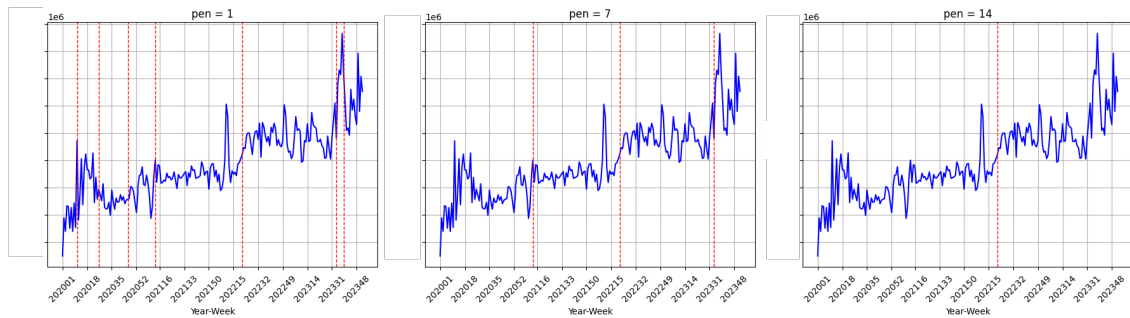
Seasonally Adjusted qty for Category 0601 - limp. e tratam roupa



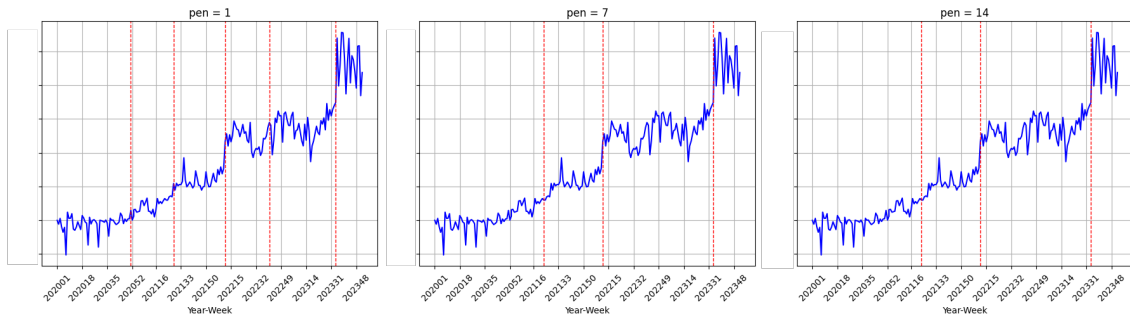
Seasonally Adjusted cesta_media for Category 0601 - limp. e tratam roupa



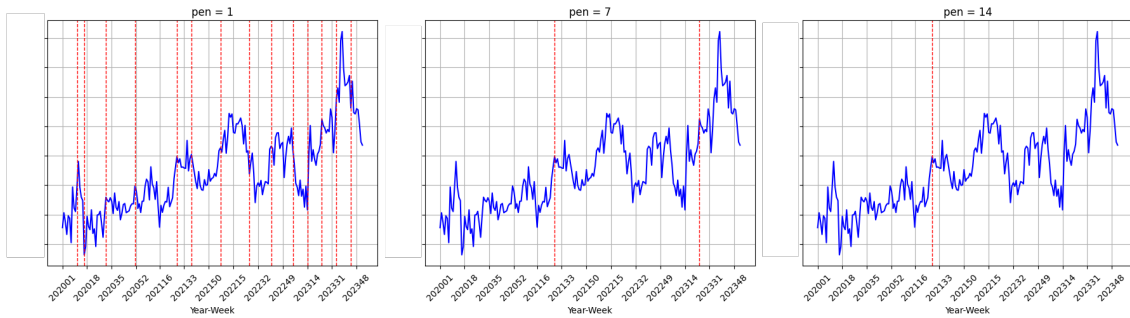
Seasonally Adjusted vb_nf for Category 1001 - gorduras líquidas



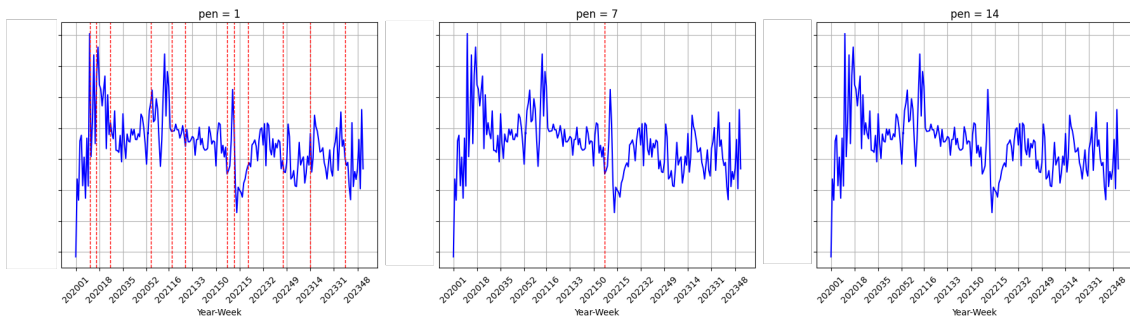
Seasonally Adjusted pvp_medio for Category 1001 - gorduras líquidas



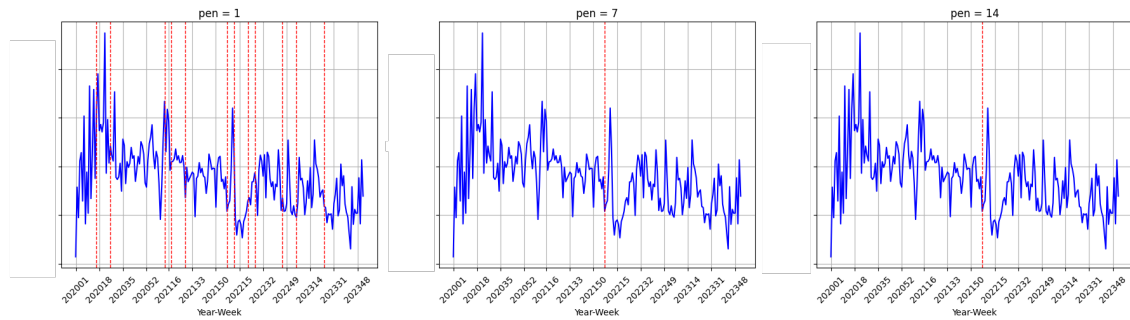
Seasonally Adjusted tx_penetracao_mp_vb for Category 1001 - gorduras líquidas



Seasonally Adjusted n_clientes for Category 1001 - gorduras líquidas



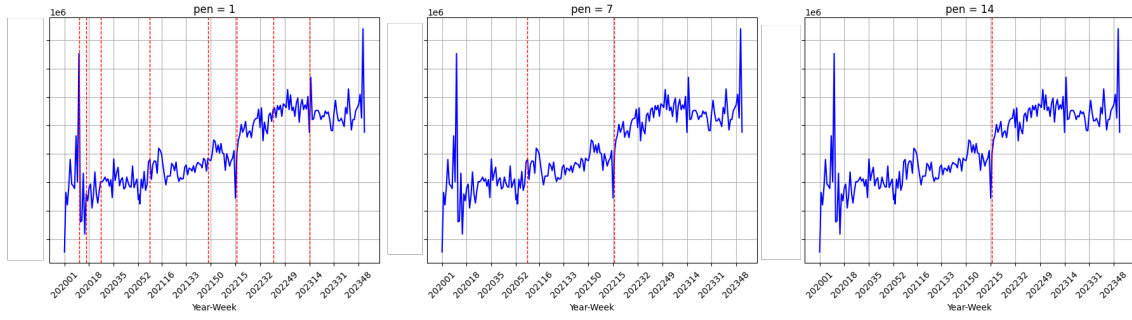
Seasonally Adjusted qty for Category 1001 - gorduras líquidas



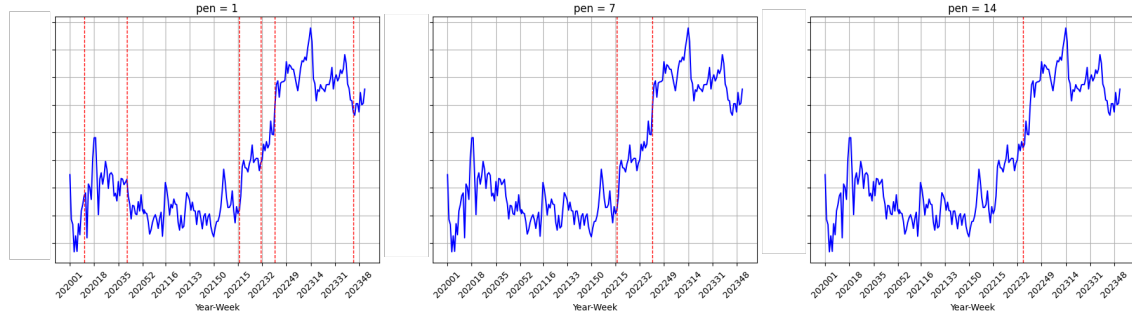
Seasonally Adjusted cesta_medio for Category 1001 - gorduras líquidas



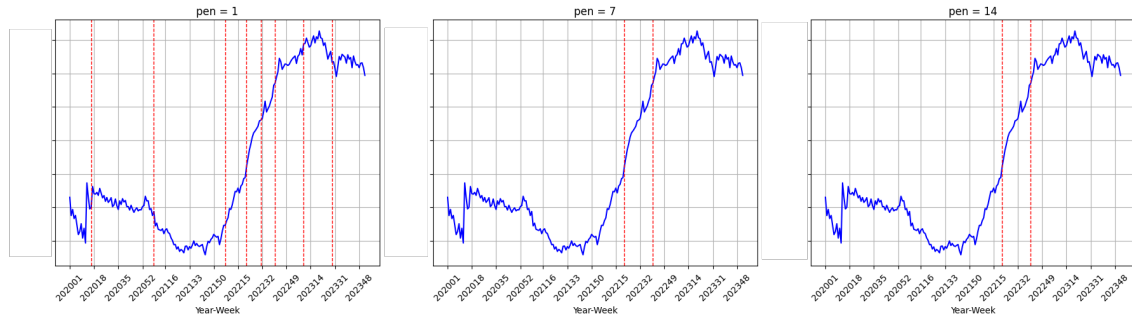
Seasonally Adjusted vb_nf for Category 0804 - iogurtes e sobrem.



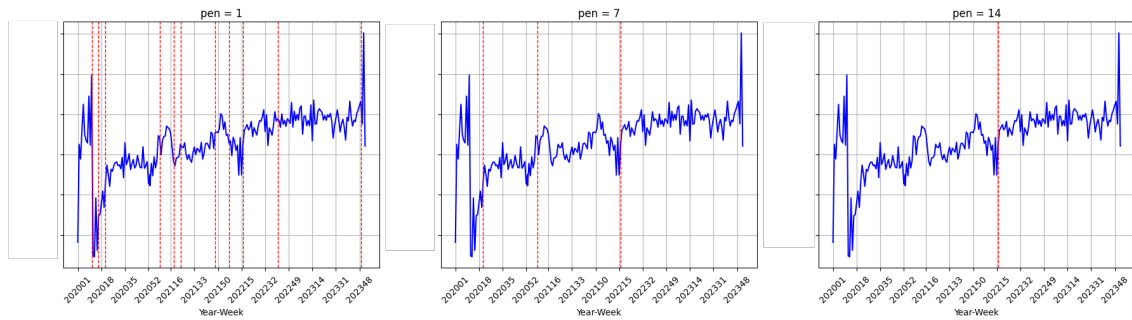
Seasonally Adjusted pvp_medio for Category 0804 - iogurtes e sobrem.



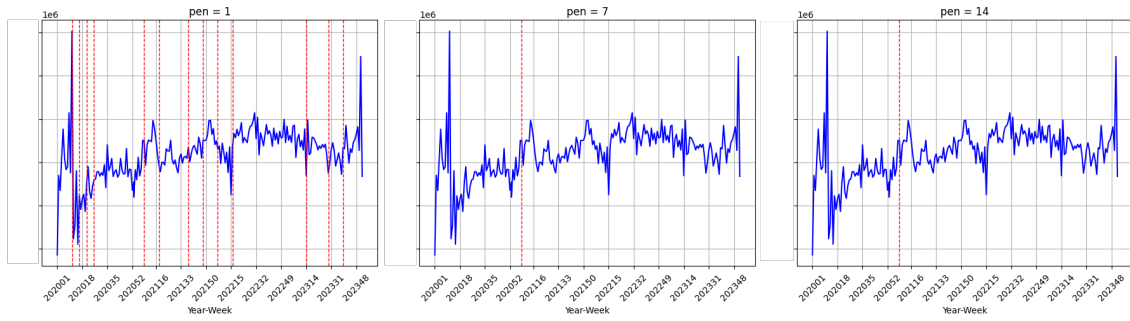
Seasonally Adjusted tx_penetracao_mp_vb for Category 0804 - iogurtes e sobrem.



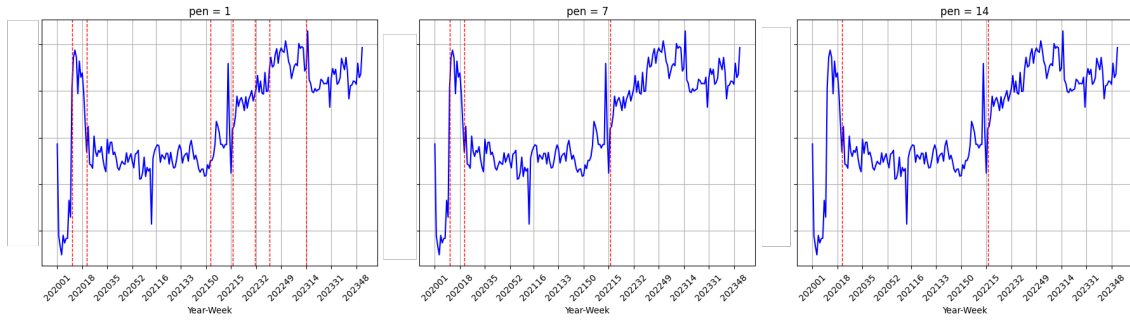
Seasonally Adjusted n_clientes for Category 0804 - iogurtes e sobrem.



Seasonally Adjusted qty for Category 0804 - iogurtes e sobrem.

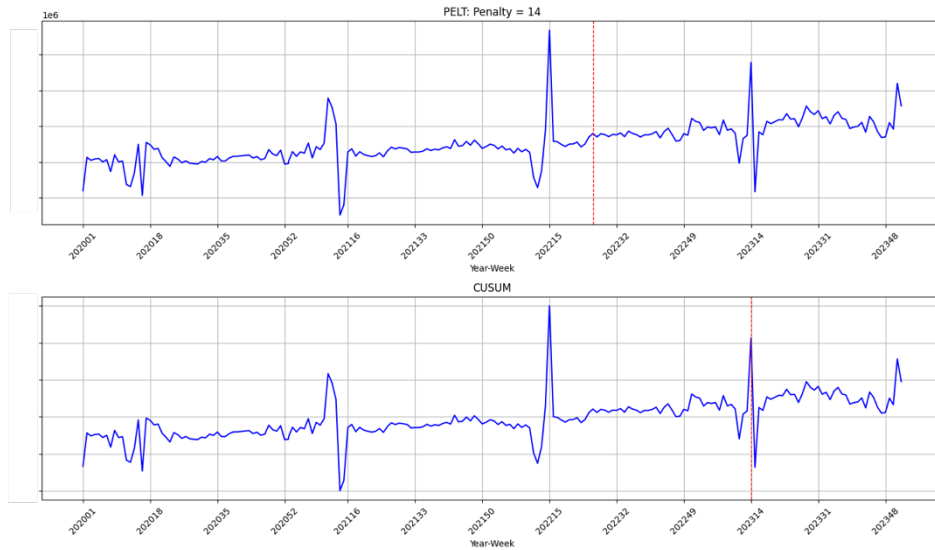


Seasonally Adjusted cesta_media for Category 0804 - iogurtes e sobrem.

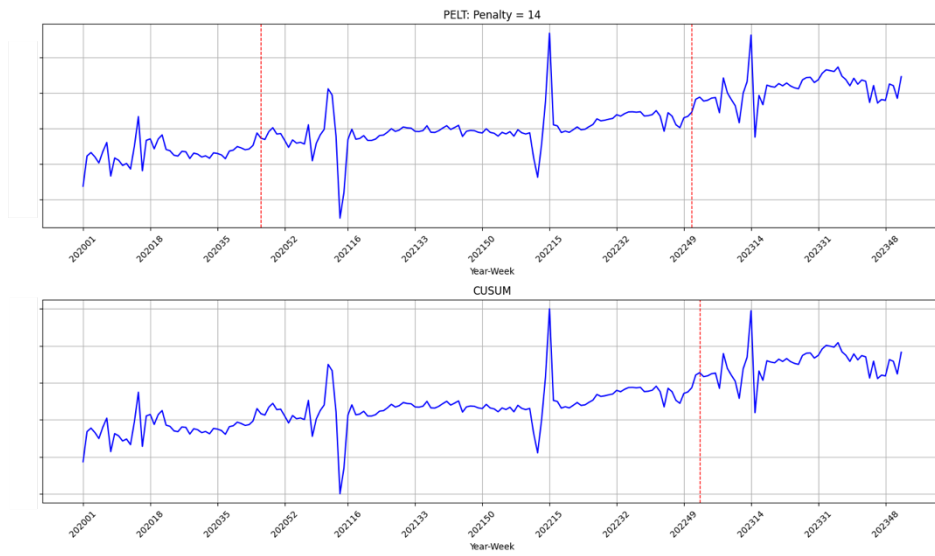


Section D - Comparison of PELT vs CUSUM

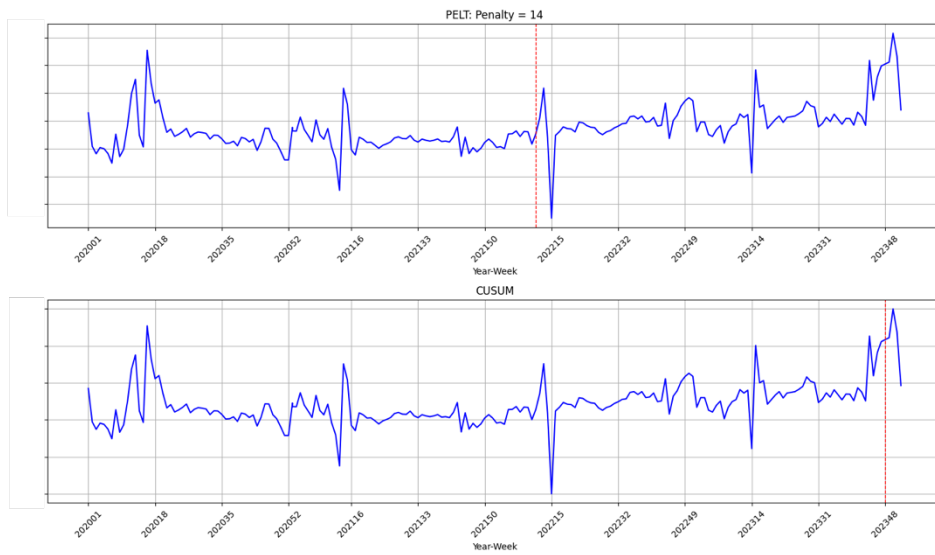
Seasonally Adjusted vb_nf for Category 0202 - doçaria



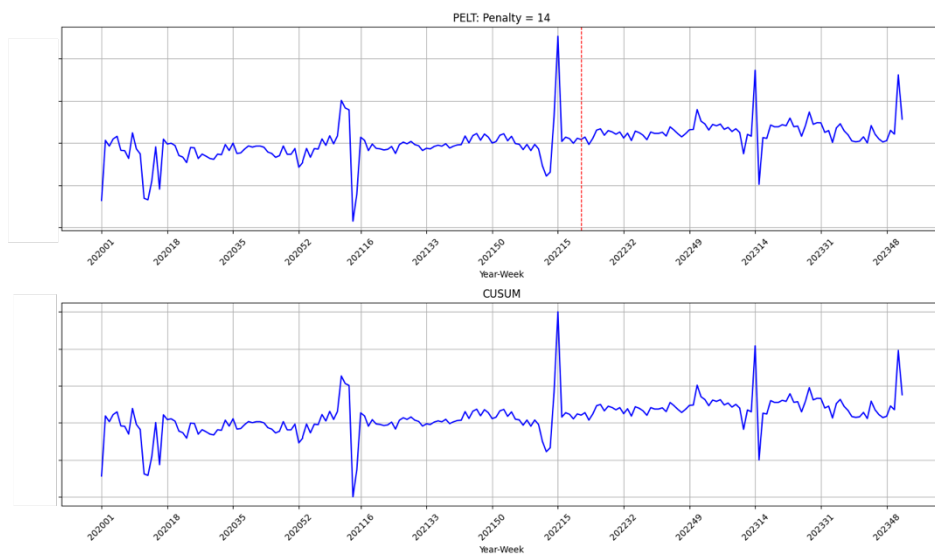
Seasonally Adjusted pvp_medio for Category 0202 - doçaria



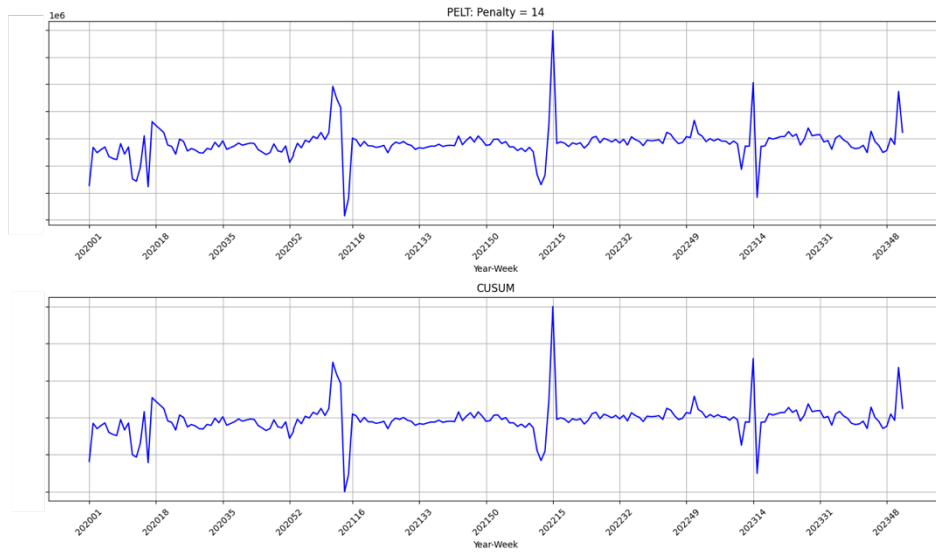
Seasonally Adjusted tx_penetracao_mp_vb for Category 0202 - doçaria



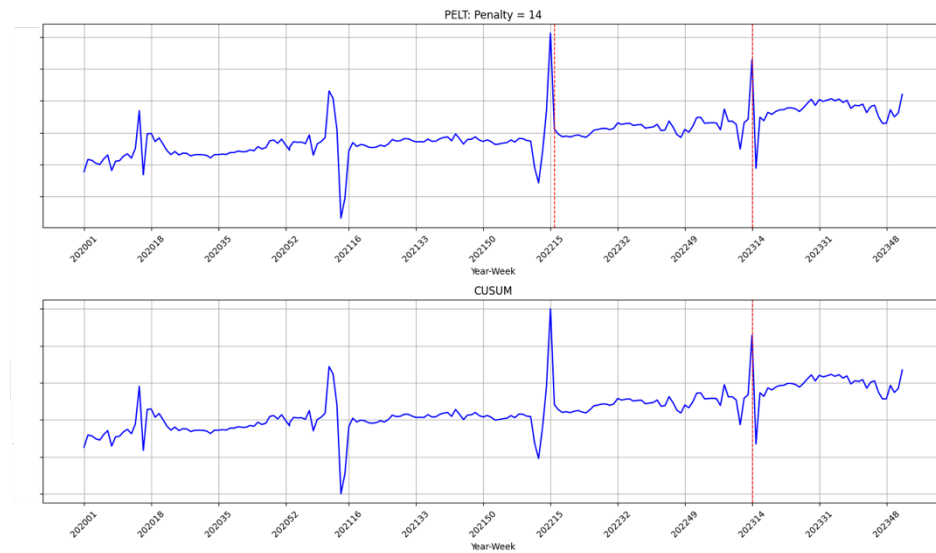
Seasonally Adjusted n_clientes for Category 0202 - doçaria



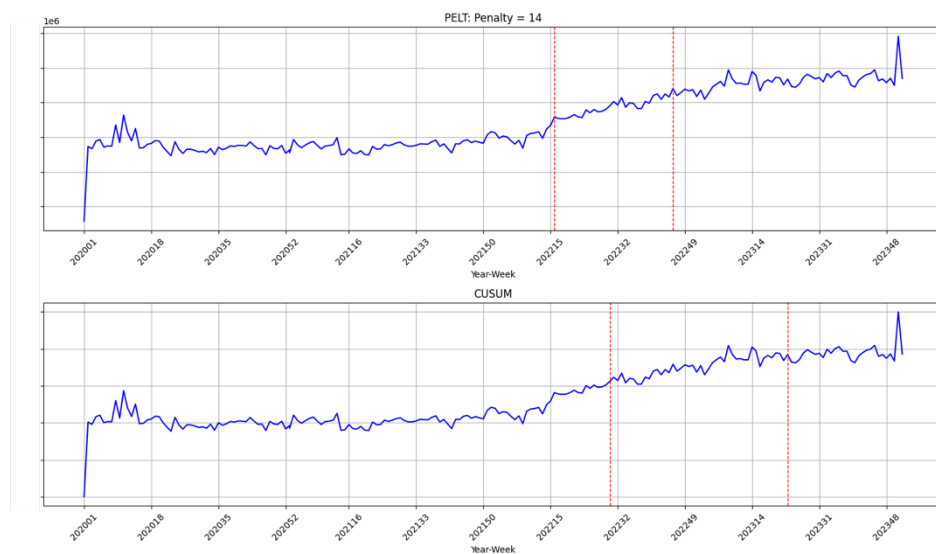
Seasonally Adjusted qty for Category 0202 - doçaria



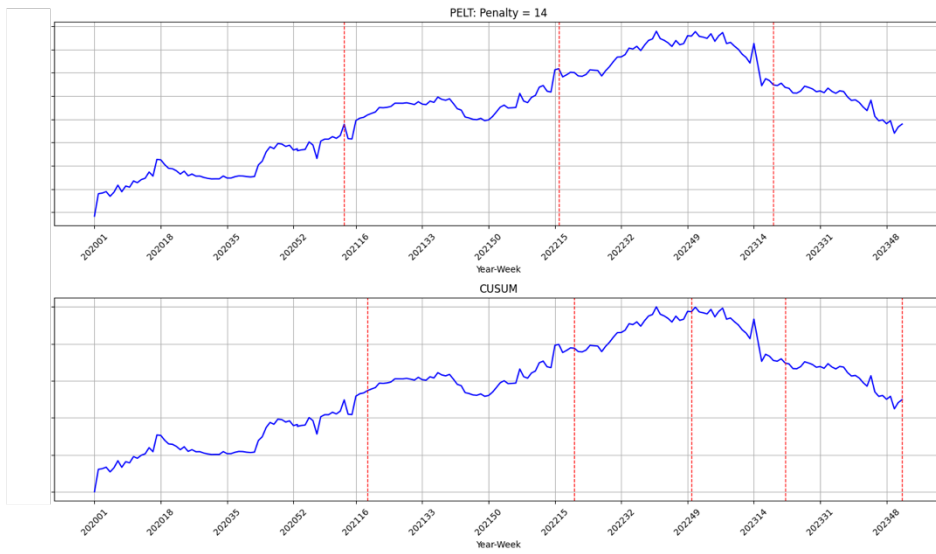
Seasonally Adjusted cesta_medio for Category 0202 - doçaria



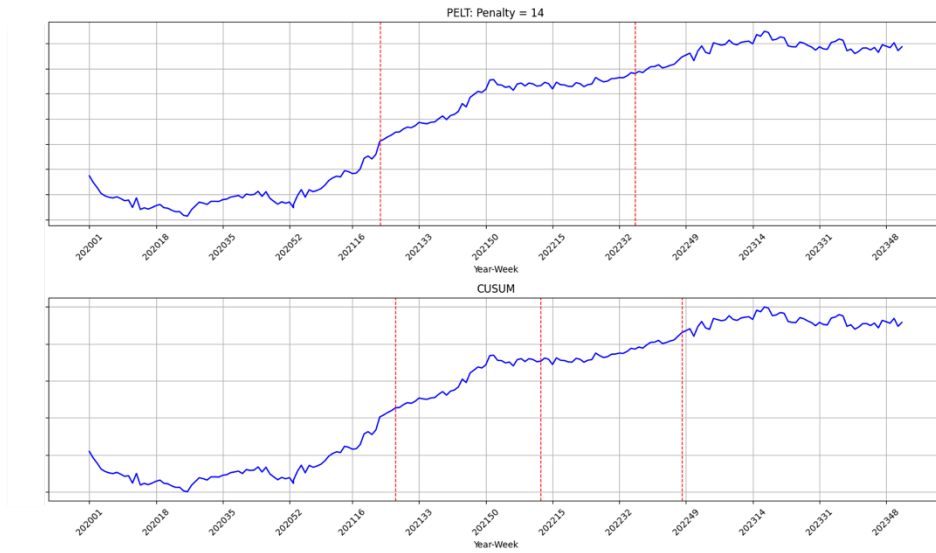
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional



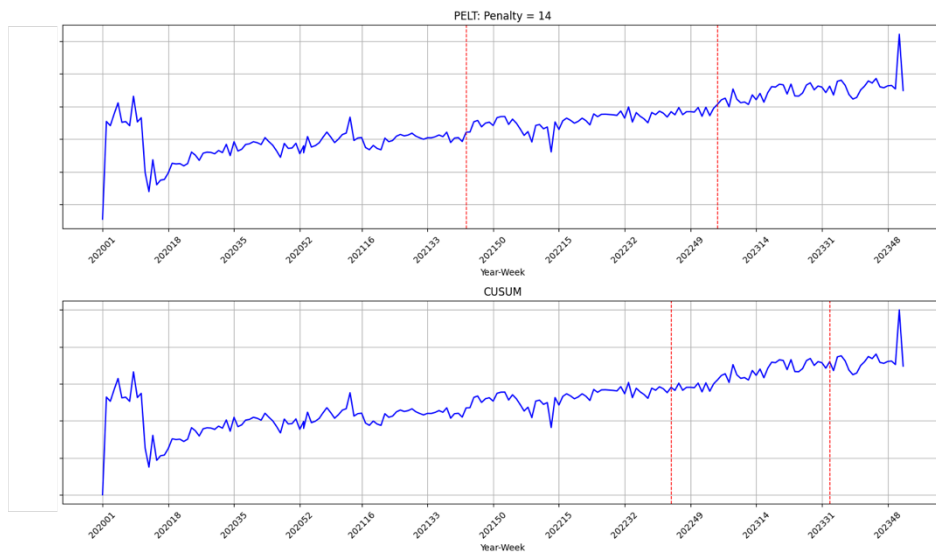
Seasonally Adjusted pvp_medio for Category 1601 - pão tradicional



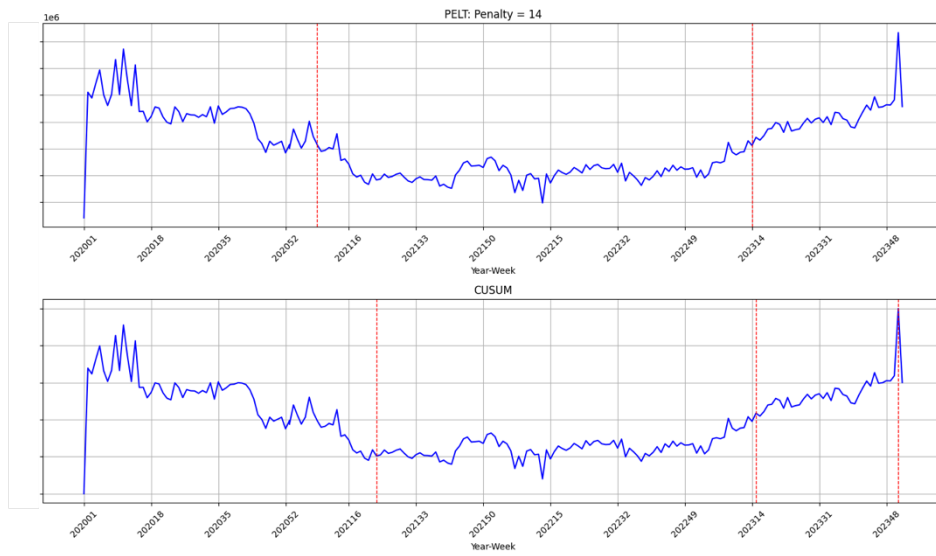
Seasonally Adjusted tx_penetracao_mp_vb for Category 1601 - pão tradicional



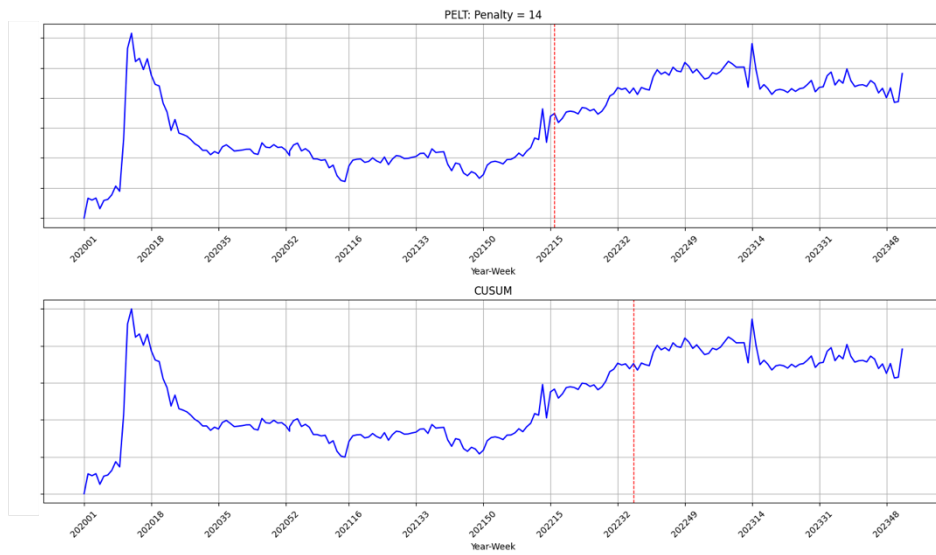
Seasonally Adjusted n_clientes for Category 1601 - pão tradicional



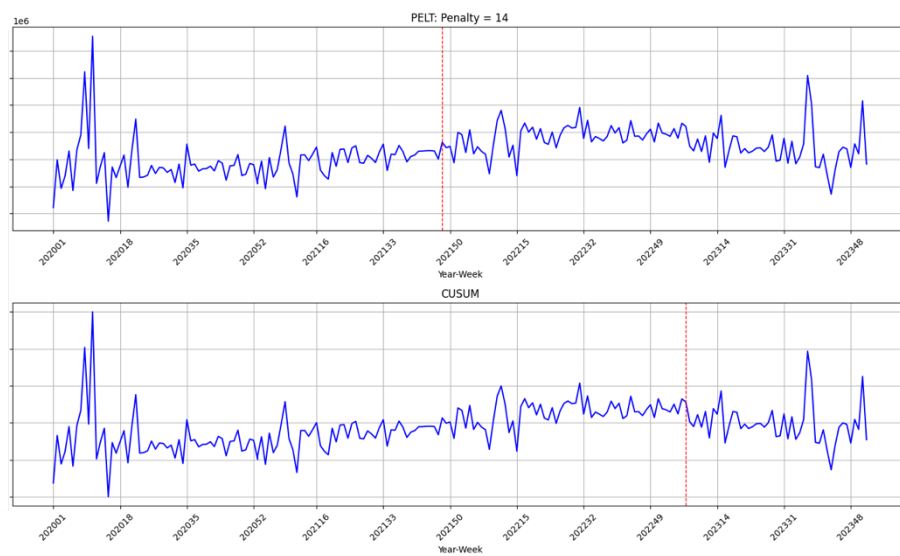
Seasonally Adjusted qty for Category 1601 - pão tradicional



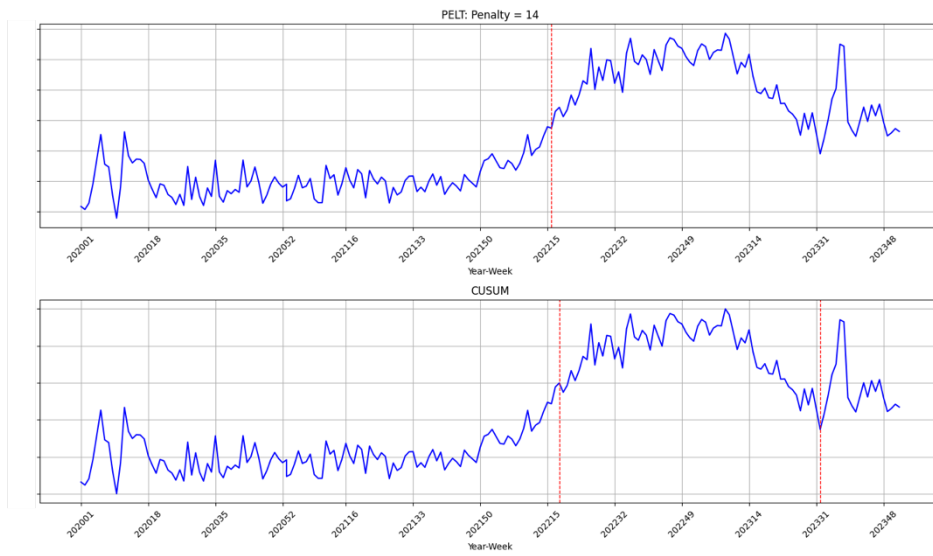
Seasonally Adjusted cesta_medio for Category 1601 - pão tradicional



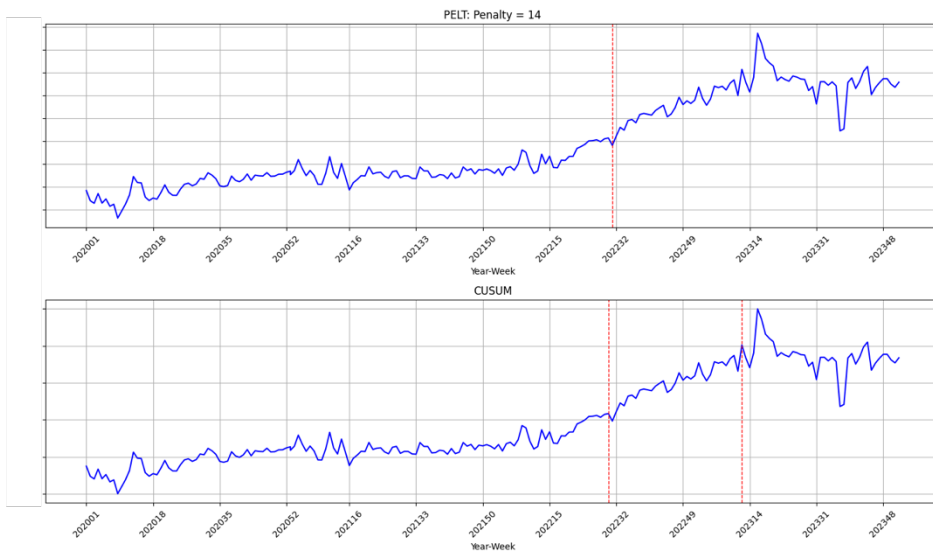
Seasonally Adjusted vb_nf for Category 0601 - limp. e tratam roupa



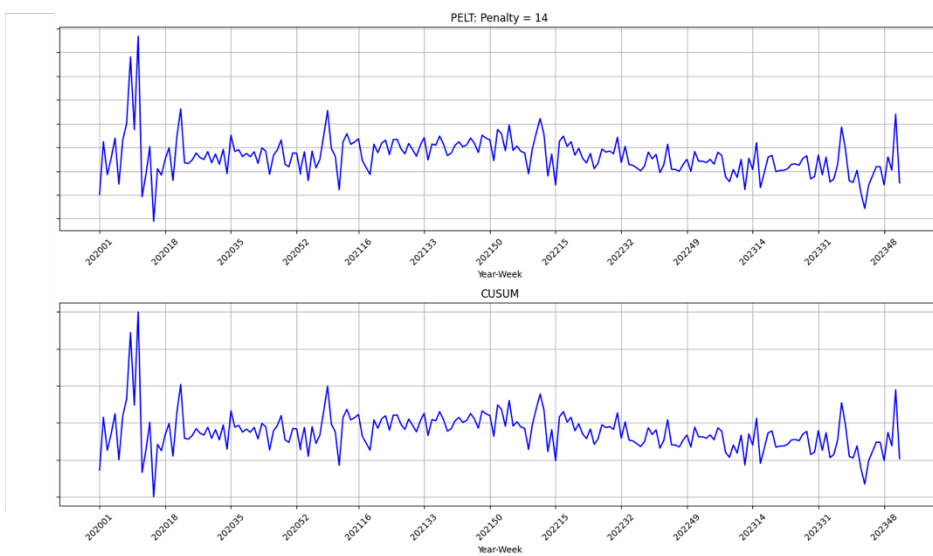
Seasonally Adjusted pvp_medio for Category 0601 - limp. e tratam roupa



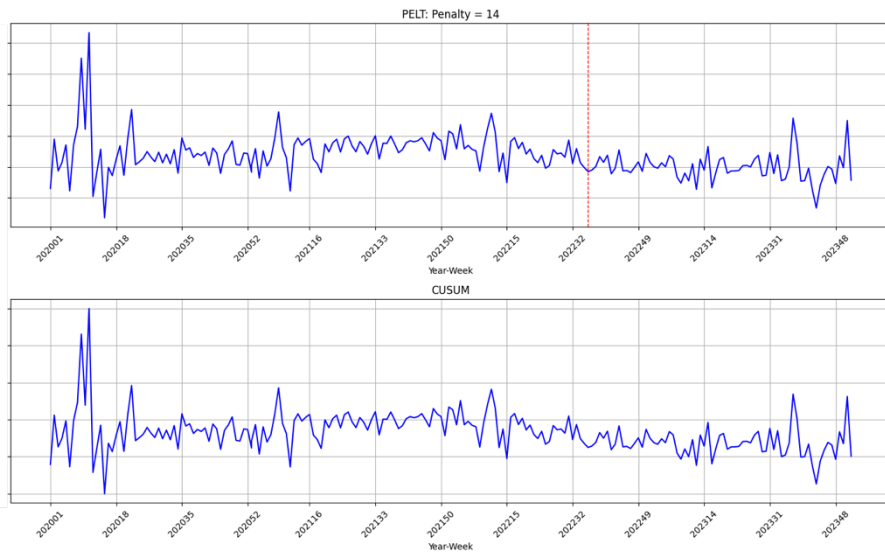
Seasonally Adjusted tx_penetracao_mp_vb for Category 0601 - limp. e tratam roupa



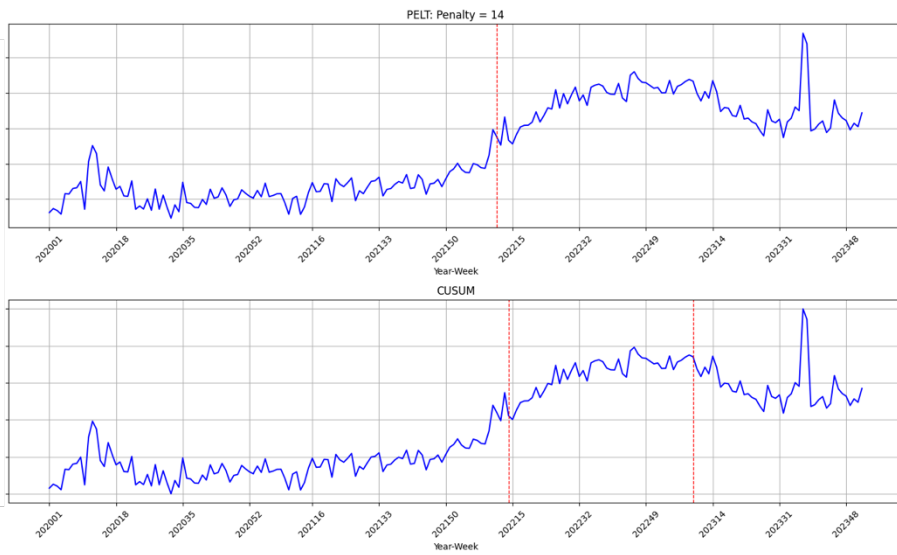
Seasonally Adjusted n_clientes for Category 0601 - limp. e tratam roupa



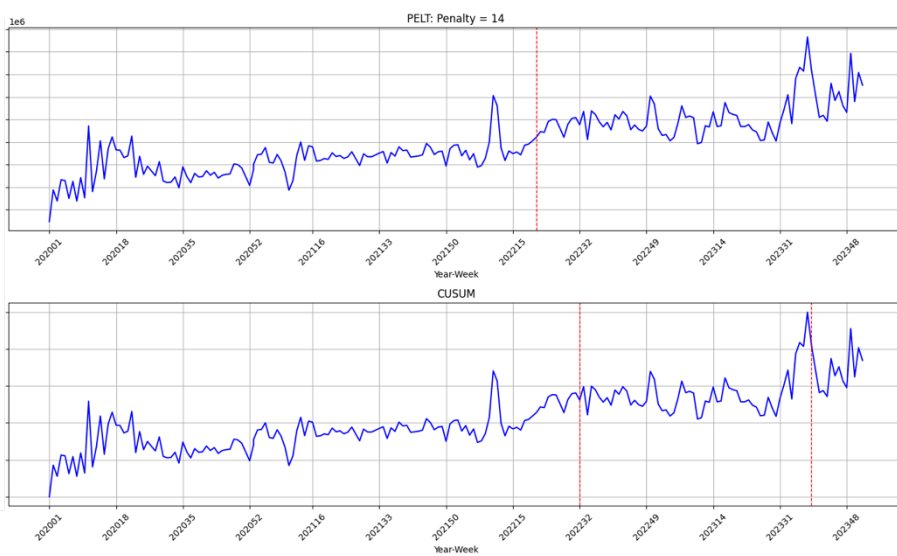
Seasonally Adjusted qty for Category 0601 - limp. e tratam roupa



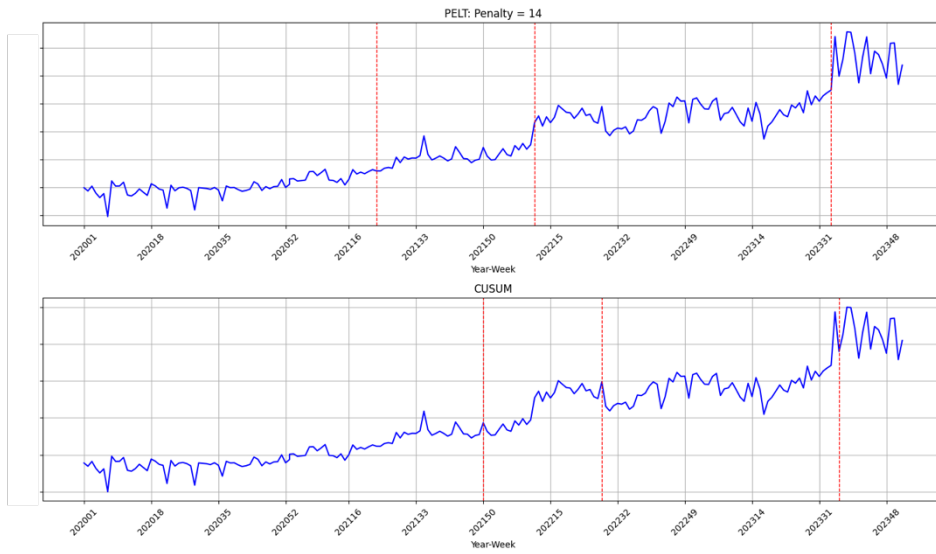
Seasonally Adjusted cesta_media for Category 0601 - limp. e tratam roupa



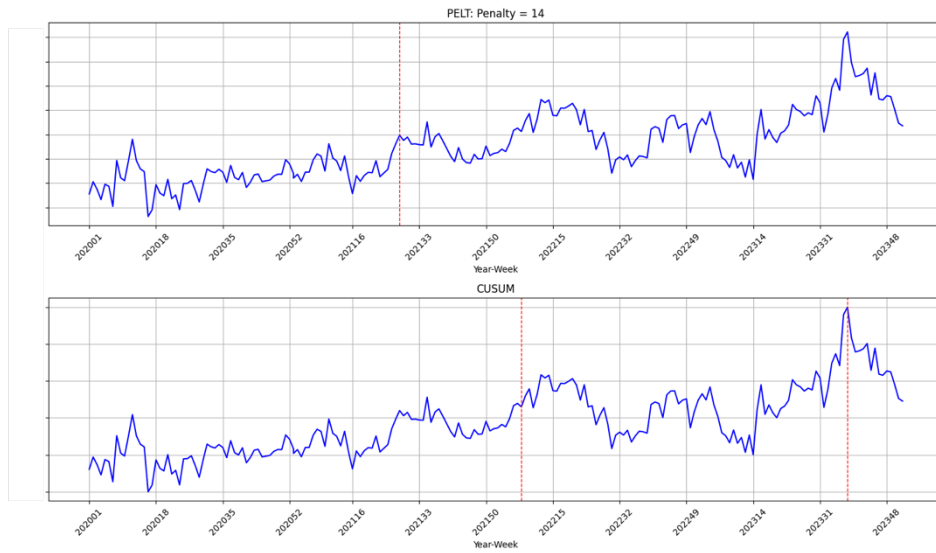
Seasonally Adjusted vb_nf for Category 1001 - gorduras líquidas



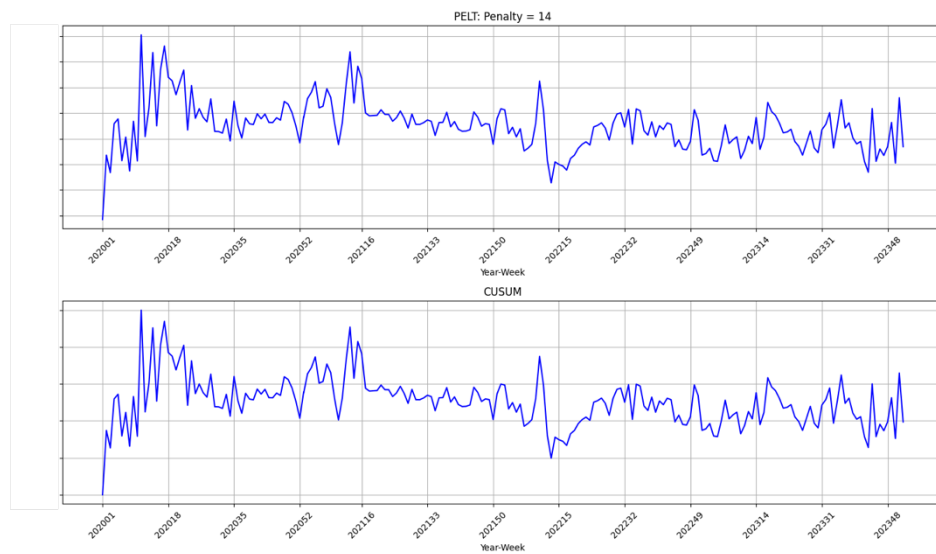
Seasonally Adjusted pvp_medio for Category 1001 - gorduras líquidas



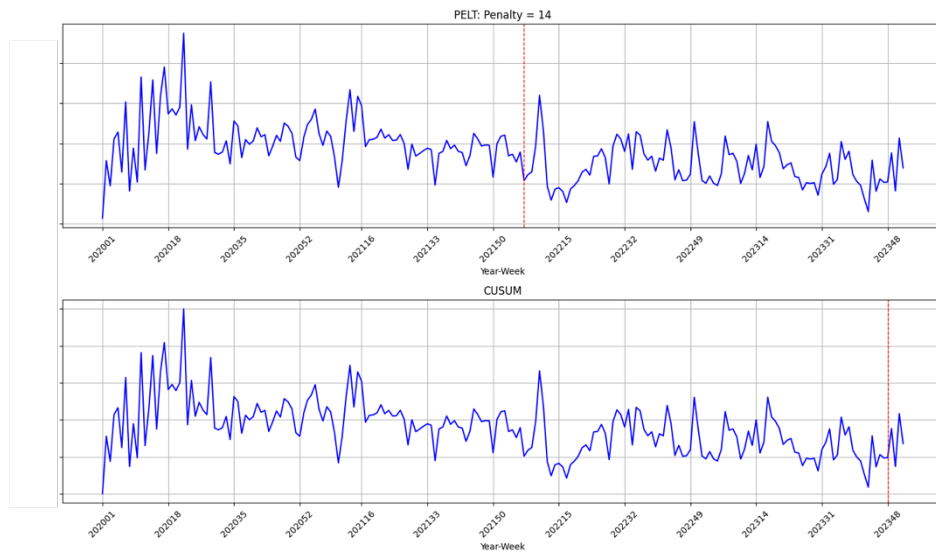
Seasonally Adjusted tx_penetracao_mp_vb for Category 1001 - gorduras líquidas



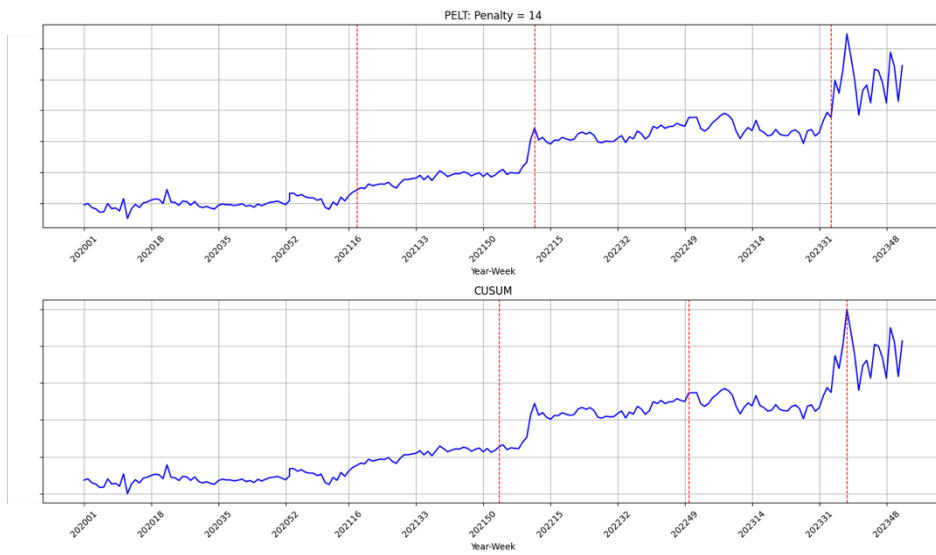
Seasonally Adjusted n_clientes for Category 1001 - gorduras líquidas



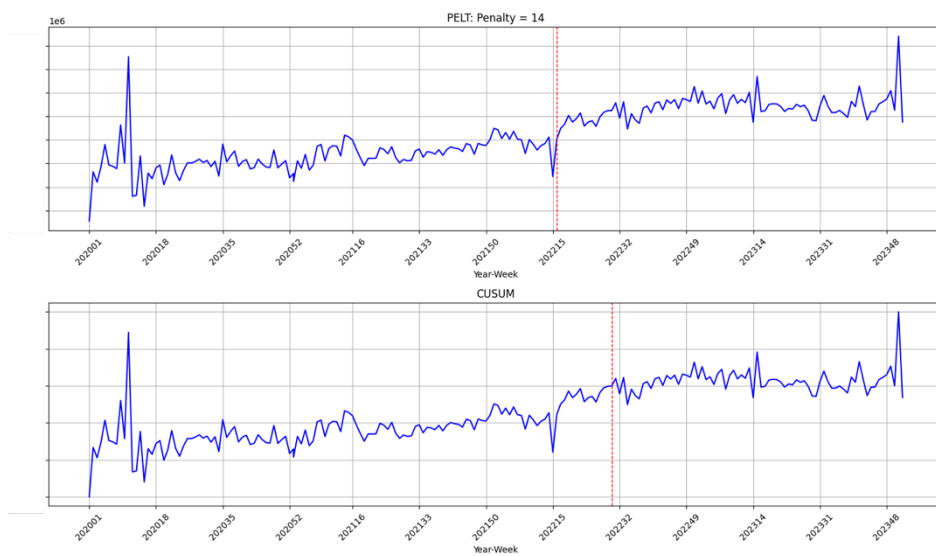
Seasonally Adjusted qty for Category 1001 - gorduras líquidas



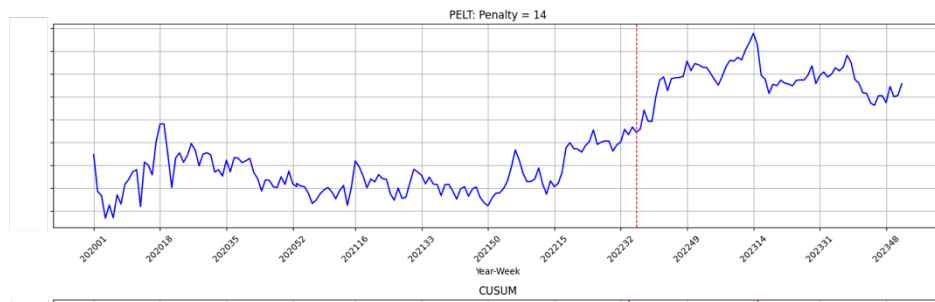
Seasonally Adjusted cesta_medio for Category 1001 - gorduras líquidas



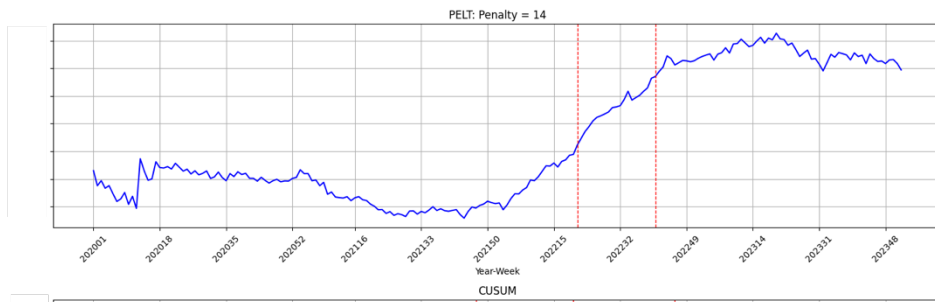
Seasonally Adjusted vb_nf for Category 0804 - iogurtes e sobrem.



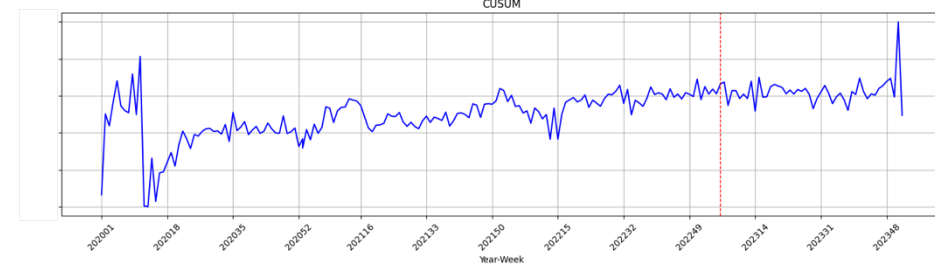
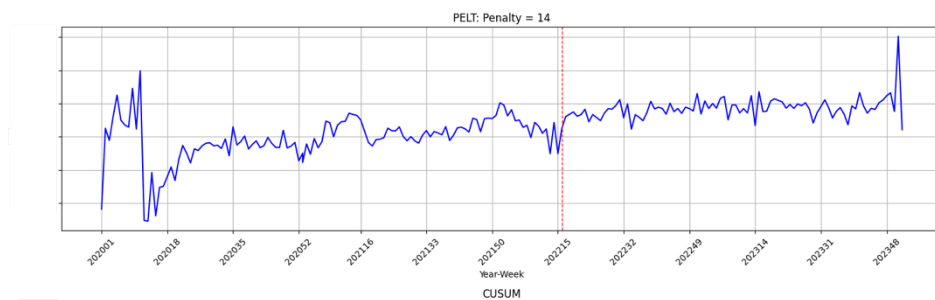
Seasonally Adjusted pvp_medio for Category 0804 - iogurtes e sobrem.



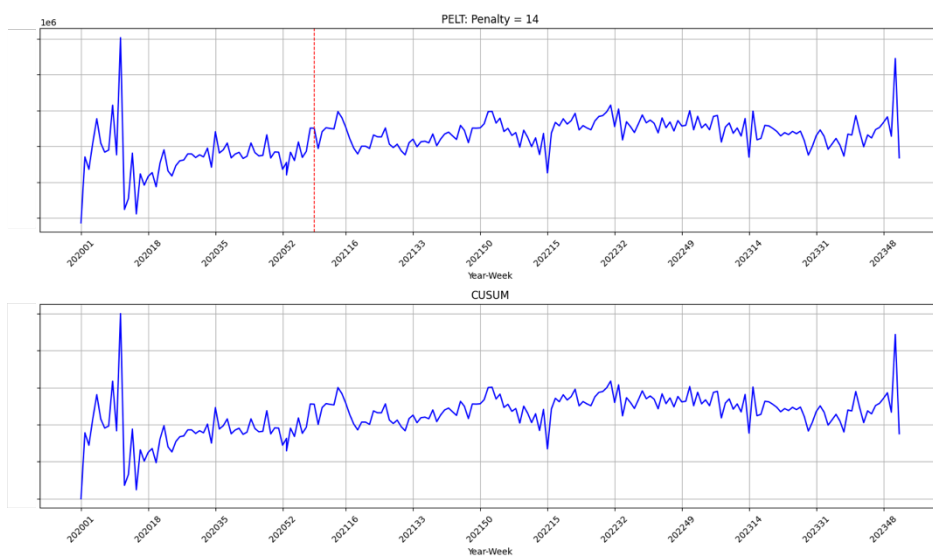
Seasonally Adjusted tx_penetracao_mp_vb for Category 0804 - iogurtes e sobrem.



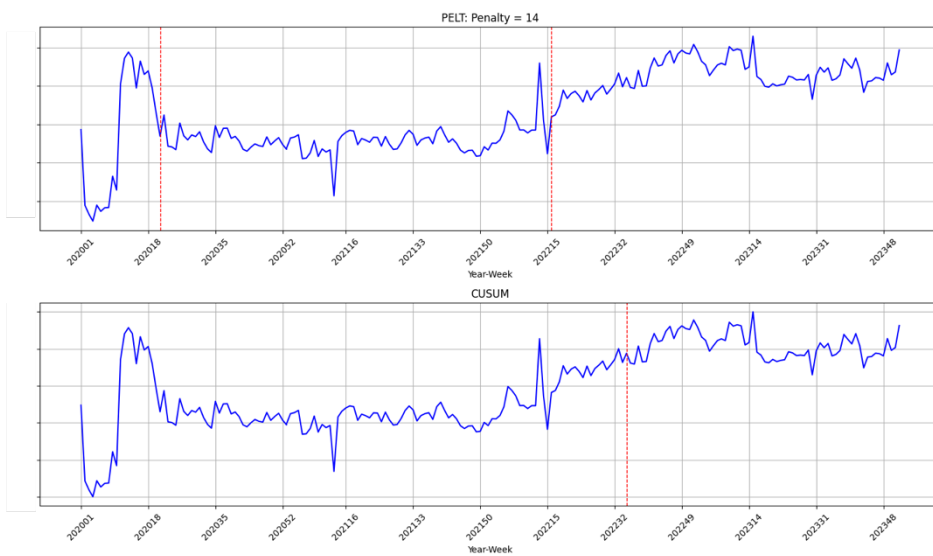
Seasonally Adjusted n_clientes for Category 0804 - iogurtes e sobrem.



Seasonally Adjusted qty for Category 0804 - iogurtes e sobrem.

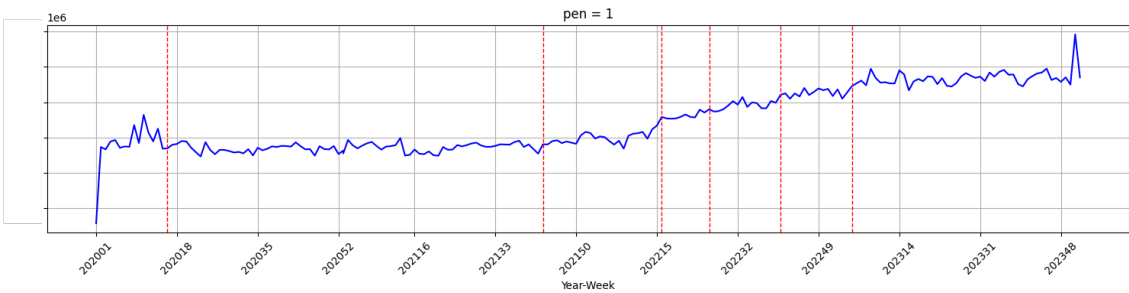


Seasonally Adjusted cesta_media for Category 0804 - iogurtes e sobrem.

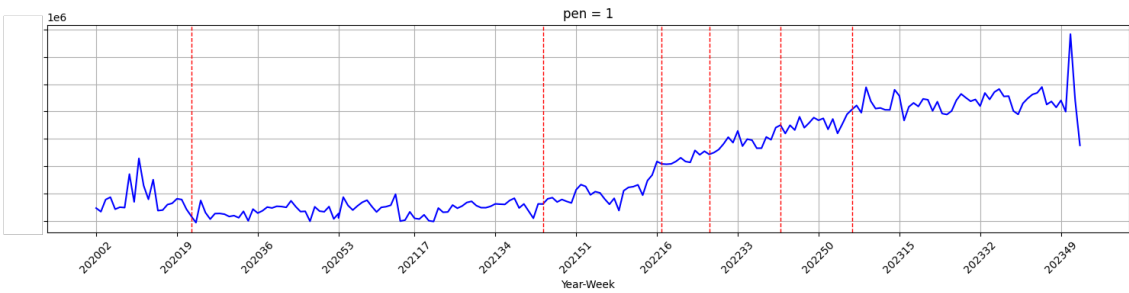


Section E – Change Detection on Recent Data

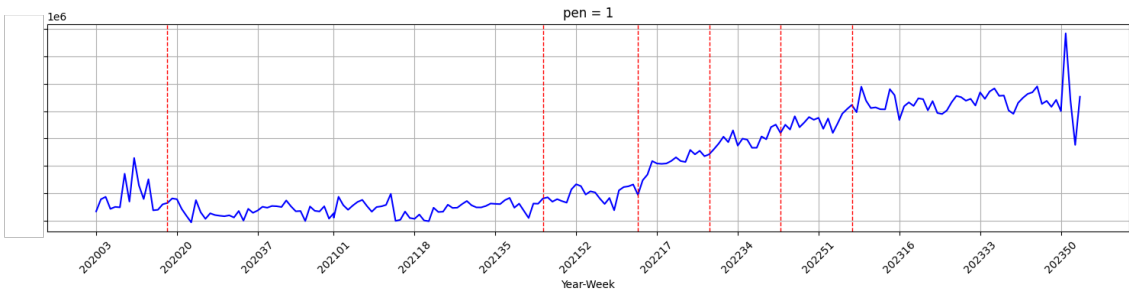
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202001 to 202352



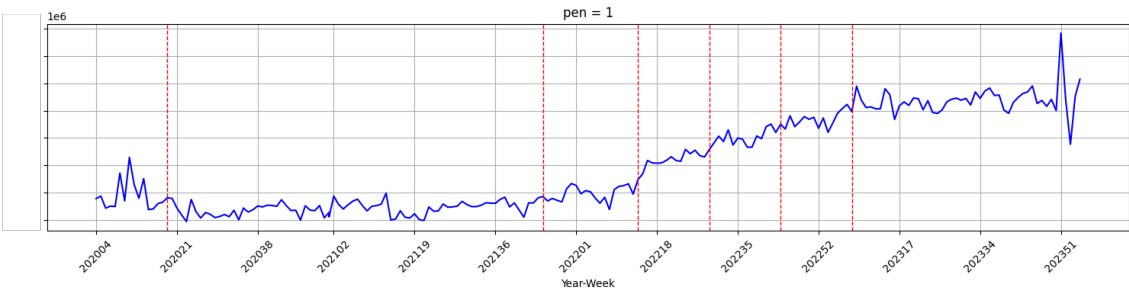
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202002 to 202401



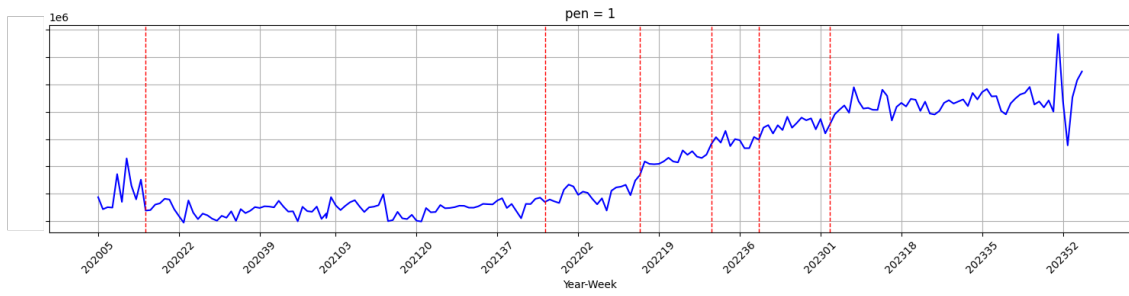
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202003 to 202402



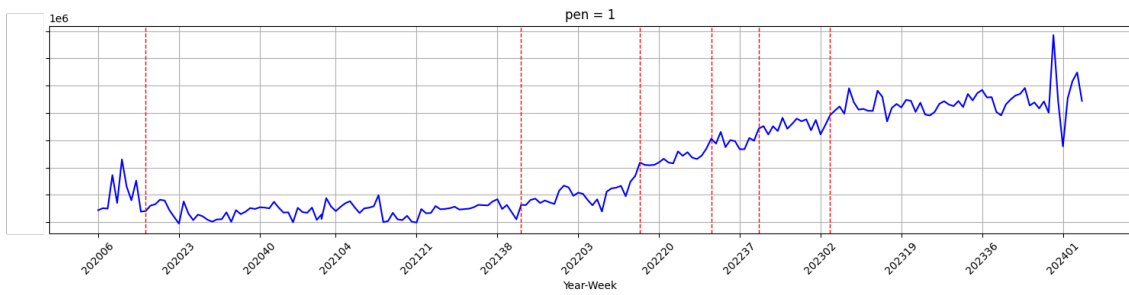
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202004 to 202403



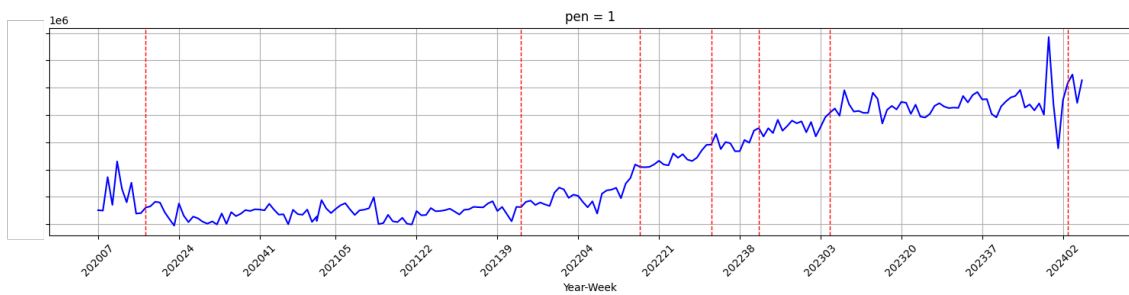
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202005 to 202404



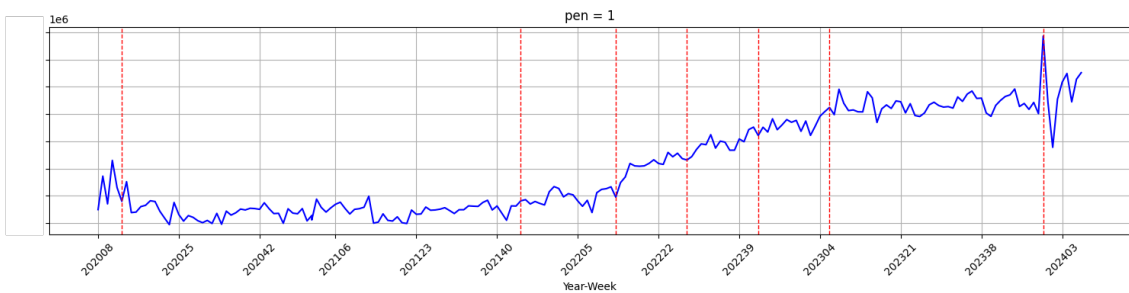
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202006 to 202405



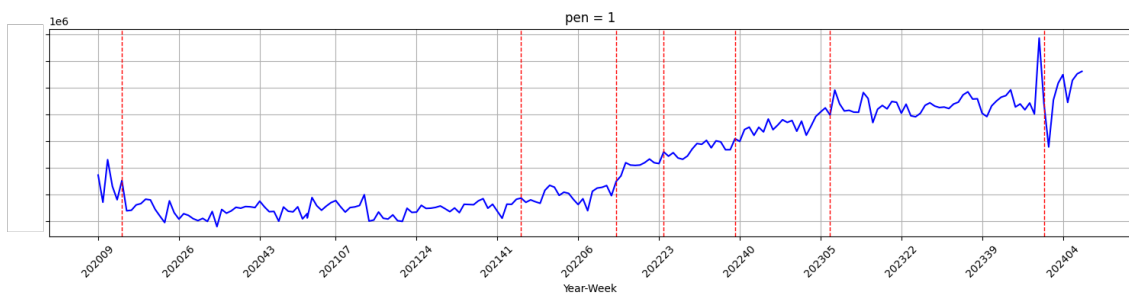
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202007 to 202406



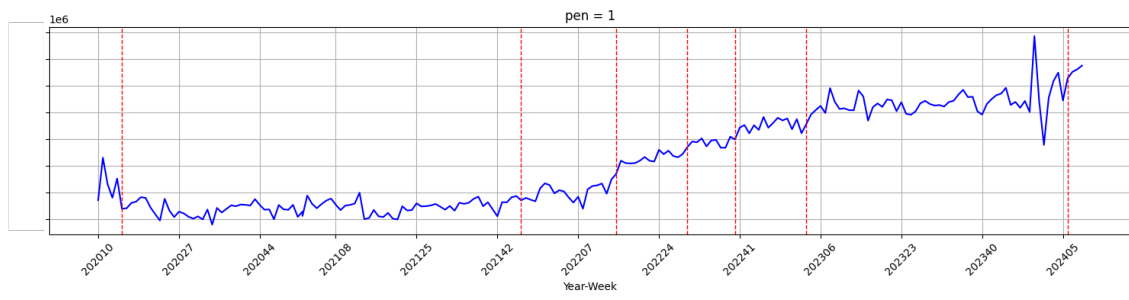
Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202008 to 202407



Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202009 to 202408



Seasonally Adjusted vb_nf for Category 1601 - pão tradicional
Year-Week: 202010 to 202409



Section F - Communication Template

```
email_content = f"""
```

Hello Director of Category,

I hope this message finds you well.

I've conducted an analysis of the category using the latest GPT-4 insights, and I am writing to share some important findings related to year-to-date (YTD) performance. Firstly, examining the performance metrics within the sales equation (Total Loyal Sales = Customers * Frequency * Average Basket), it is noted that:

- The Total Loyal Sales (*VLR Fidelizadas*) has an RH value of {value_1}, which is {'positive' if value_1 > 100 else 'neutral' if value_1 == 100 else 'negative'}.
- The Customers (*Clientes*) have an RH value of {value_2}, which is {'positive' if value_2 > 100 else 'neutral' if value_2 == 100 else 'negative'}.
- The Frequency (*Frequência*) has an RH value of {value_3}, which is {'positive' if value_3 > 100 else 'neutral' if value_3 == 100 else 'negative'}.
- The Average Basket (*Cesta Média*) has an RH value of {value_4}, which is {'positive' if value_4 > 100 else 'neutral' if value_4 == 100 else 'negative'}.

```
{output}
```

It is imperative to note that this email was generated with GenAI, and critical thinking should be a cornerstone in interpreting these insights. The interplay of various metrics is complex, and while AI can highlight patterns, human oversight is crucial for strategic application and decision-making.

Best regards,

SonaeMC Bot-Analyst

```
"""
```

```
# Print or use the email content as needed
```

```
print(email_content)
```

Section G - Recommendation Letter

Ana Freitas

Area Manager Advanced Analytics

MC Sonae – Cartão Continente

To whom it may concern,

I have had the pleasure of supervising Francisco Magalhães during his masters' internship at MC Sonae, where he has demonstrated exceptional skills and dedication.

He successfully achieved the master thesis objectives, being very autonomous since day one, challenging the framework and proposing new methodologies.

Francisco is an avid learner, with a hands-on test and learn mindset.

The output of his internship combines the business needs and an analytical framework, resulting in an actionable analytical product that automatically summarizes the main insights for the reports' end users.

Sincerely,

Ana Freitas

FACULDADE DE ECONOMIA

