Lisbon School
of Economics
& Management
Universidade de Lisboa

Master

DATA ANALYTICS FOR BUSINESS

Master's Final Work

PROJECT

Competing Forecasting Models to Study Crisis Periods: The Case of Sweet Snacks Sales

SARA CRISTINA COELHO BARRADAS

MARCH - **2024**

Lisbon School of Economics & Management
Universidade de Lisboa

MASTER
DATA ANALYTICS FOR BUSINESS

MASTER'S FINAL WORK
PROJECT

COMPETING FORECASTING MODELS TO STUDY CRISIS PERIODS: THE CASE OF SWEET SNACKS SALES

SARA CRISTINA COELHO BARRADAS

SUPERVISION:
PROF. PATRÍCIA ALEXANDRA LAGARTO MARTINS

MARCH - **2024**

ACKNOWLEDGMENTS

First and foremost, I would like to start by expressing my deepest gratitude to my parents Ana and Pedro for their unconditional love, support, and trust. I owe my academic success to them for their involvement in my education.

In addition, I would like to thank my grandparents for the endless care and encouragement along this journey, without being physically present.

A special mention to Armando Mateus and *TouchPoint Consulting* for welcoming me so generously at the very beginning of my professional career. Their contribution was essential in providing access to the data used to elaborate this MFW.

My work colleagues Michael Delgado and Ivan Borges also deserve recognition for the valuable knowledge they have shared with me over the last few months.

Lastly, I would like to express my sincere appreciation to my thesis advisor, Professor Patrícia Martins, for her patience, dedication, and guidance during the entire course of the project.

Glossary

ACF – Autocorrelation Function.

AIC – Akaike Information Criterion.

AICc – Corrected Akaike Information Criterion.

ARIMA – Autoregressive Integrated Moving Average.

BI – Business Intelligence.

BIC – Bayesian Value of Information Criterion.

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

DAX – Data Analysis Expressions.

DM – Data Mining.

DW – Data Warehouse.

ELT – Extract, Load and Transform.

ETL – Extract, Transform and Load.

KD – Knowledge Discovery.

KDD – Knowledge Discovery in Databases.

KDDM – Knowledge Discovery and Data Mining.

MAE – Mean Absolute Error.

MAPE – Mean Absolute Percentage Error.

ML – Machine Learning.

MLE – Maximum Likelihood Estimation.

MSE – Mean Square Error.

PACF – Partial Autocorrelation Function.

RMSE – Root Mean Square Error.

SARIMA – Seasonal Autoregressive Integrated Moving Average.

SSBI – Self-Service Business Intelligence.

STL – Seasonal-Trend Decomposition Procedure Using Loess.

ABSTRACT

The COVID-19 pandemic significantly affected the purchasing behaviour of Portuguese families, compelling them to reduce their shopping expenditures. This socioeconomic crisis necessitated that food retailers adapt their strategies to evolving consumer preferences, emphasizing digitalization, sustainability, and safety. This study examines the sales evolution of the Sweet Snacks category at two major Portuguese retail banners[1] from January 2018 to June 2023, segmented into three forecasting periods: pre-crisis, crisis and post-crisis. The project's primary objective is to infer forecasting models for these periods, using the ARIMA and Prophet time series models, and compare them to assess consumer preference changes. Additionally, this work forecasts Sweet Snacks sales beyond June 2023 to extend the appraisal of sales performance in the post-crisis and detect potential anomalies in the sector. Using the CRISP-DM methodology, the research developed an integrated BI solution, employing Power BI for data preparation and R Studio for multidimensional data modelling and forecasting analysis. In the pre-crisis period, Sweet Snacks sales progressively increased until the onset of lockdown, declining in the crisis period. At the end of the crisis, consumption patterns normalised, but post-crisis, retailers diverged due to their adaptability to new trends. The results indicate that while ARIMA models generally offer higher accuracy, Prophet provides more precise future forecasts. ARIMA predicts a steady future trend, whereas Prophet captures post-crisis sales patterns more effectively. This project's main contribution is the development of a BI solution and a comprehensive forecasting report for a Consulting organisation in the Food Retail sector.

KEYWORDS: Time Series Forecasting; Crisis periods; Sweet Snacks sales; ARIMA; Prophet; CRISP-DM.

JEL CODES: C22; C53; C87; L81; M21.

[1] Refers to a specific brand or chain owned by a food retailer.

TABLE OF CONTENTS

LIST OF APPENDICES

List of Figures

List of Tables

1. INTRODUCTION

1.1. Project context and motivation

The outbreak of the COVID-19 pandemic in early 2020 caused significant disruptions worldwide, with a slowdown in the global economy, particularly in private consumption. Like many other countries, Food Retail was heavily impacted by lockdowns, social distancing and public health concerns. Consequently, this period of socioeconomic crisis reshaped the behaviour of consumers regarding food consumption, namely Portuguese consumers, in response to the unprecedented challenges posed by the pandemic. Therefore, it is essential to understand if these effects indicate a permanent structural shift in consumers' purchasing patterns or a temporary change. In this context, the present project focuses on analysing the impact of crisis on the Sweet Snacks retail category in Portugal, by identifying differences in purchasing patterns over three periods: pre-crisis, crisis and post-crisis. Additionally, the analysis intends to assess sales performance in the post-crisis period and recognise potential anomalies within the sector.

The study's application of time series data mining aims to extract valuable business insights from extensive sales data. By employing time series forecasting techniques, organisations can reassess processes based on historical patterns to enhance decision-making regarding resource allocation, capacity planning, and demand forecasting, especially in the face of future crisis scenarios and unexpected events (Faloutsos *et al.*, 2019). However, producing accurate predictions requires analytical expertise, which is often not embedded in the entrepreneurial context. Therefore, this research is conducted in collaboration with *TouchPoint Consulting*, a Category Management and Shopper Marketing consulting company, to provide the BI department with an advanced predictive analysis tool for application in internal business analytics projects. Automated forecasting processes such as Autoregressive Integrated Moving Average (ARIMA) models have historically been a very popular approach, while the Prophet model has recently emerged as a viable alternative.

1.2. Objectives definition

The main objective of this project is to study differences in food retail purchasing patterns of Portuguese consumers in the Sweet Snacks category resulting from the

COVID-19 crisis. To accomplish this, this work develops forecasting models to identify the time series forecasting models that most accurately describe the evolution of sales in the Sweet Snacks category. Therefore, this project compares competing time series forecasting models across three forecasting periods:

- Pre-crisis period: January 2018 to February 2020
- Crisis period: March 2020 to December 2021
- Post-crisis period: January 2022 to June 2023

The time frame's division was established based on the official start of the COVID-19 pandemic crisis in Portugal on 2 March 2020. The crisis period ended in December 2021, coinciding with the end of mandatory remote working in the national panorama. During the crisis period, brick-and-mortar stores except for food and health retail units, faced closure due to lockdown. The post-crisis period is marked by a socioeconomic crisis resulting from the pandemic and aggravated by the after-effects of the Russia-Ukraine military conflict that began on 24 February 2022.

The complementary objective of this work is to produce future predictions of the Sweet Snacks category sales for the period after June 2023, to incorporate the developed BI solution into an actionable context. The analysis used sales data from two market-leading Portuguese food retailers to detect distinct sales trends and compare sales behaviour between the two retailers during three different sub-periods. For this purpose, the methodology adopted was CRISP-DM (Costa and Aparício, 2020), commonly used in the field of Data Mining, to organise data ELT and Modelling processes conducted in the Power BI and R Studio software tools. After an exploratory data analysis, models were inferred to generate forecasts and extract insights using specialized R packages, among the most widely known, *forecast, tsibble, tidyverse, tseries, tidyquant, gridExtra, ggplot2* and *prophet.* Finally, the best predictive models provided numerical outputs and graphs returned by R for a robust analysis of performance metrics. This interactive approach has proved the added features of ARIMA and Prophet models.

1.3. Document structure

The following sections of this report are organised as outlined below. The Literature Review covers two different topics: Portuguese Food Retail Banners, the changes in the Sweet Snacks purchasing patterns arising from the crisis period and an overview of

Manufacturer WOW Portugal in this segment; and an in-depth time series data description, its positioning in the context of Data Mining and the steps for time series forecasting process with relevant models. The CRISP-DM Methodology and the software solutions adopted for the approach, including a detailed description of its phases. The Time series Results Analysis in R with evaluation and comparison of outputs for the three periods. The Conclusions and key takeaways from the project, with mentions of some limitations experienced during the course and possible future research prospects.

2. LITERATURE REVIEW

2.1. The Portuguese Food Retail Banners

The Food industry is a critical trade sector in Portugal, including both retail and wholesale markets. In turn, the Food Retail market is a subsector of the Retail industry that focuses on selling consumer goods, mainly groceries and non-food products such as personal care and household cleaning products. Food retail companies play a vital role in the national economy by connecting producers and consumers and providing benefits to both. Retailers enable producers to access a broader market through multiple points of sale and distribution channels, while consumers enjoy a wide range of products in a single commercial place.

Between 2016 and 2020, national revenue in the Food industry increased steadily. However, the first half of 2020 witnessed a decline of 3.48%, placing revenues at 13.3 billion euros. In 2021, the industry rebounded with a 9.42% increase, reaching 14.6 billion euros by year-end. As the year 2022 approached, the food industry experienced a significant growth of 26.71%, amounting to almost 18.5 billion euros (*Portugal: food industry revenue 2022*, 2022). The Portuguese food retail landscape is highly competitive, with both national and international banners vying for market share through dynamic pricing and periodic promotional strategies. From the fourth quarter of 2021 to the fourth quarter of 2023, the Food Retail sector was mostly dominated by five main players (Figure 1): *Continente* (*Sonae*), *Pingo Doce* (*Jerónimo Martins*), *Lidl, Intermarché* and *Mercadona,* which collectively held 74% of the market share in 2023. The two retail domestic groups leading the Portuguese preference are *Continente* and *Pingo Doce*, which controlled nearly half of the national food market (48.2% in 2022). *Auchan*, *MiniPreço* and *Aldi* group chains account for the remaining 9.6%, with a further 16.4% allocated to other minor players (*Portugal: food retailers market share 2023*, 2023).

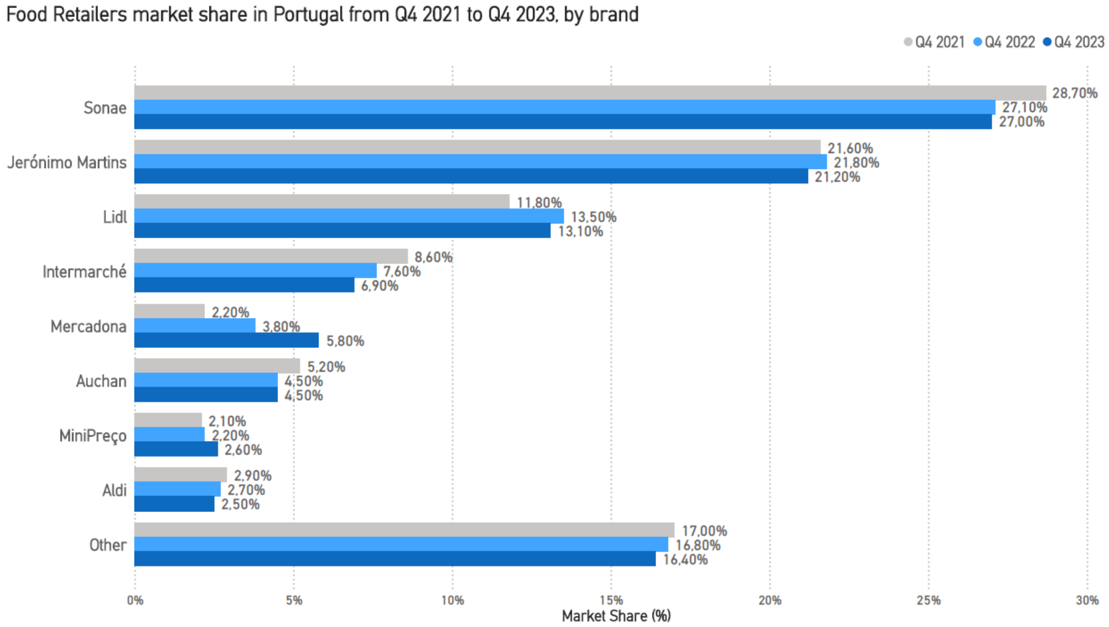


Figure 1. Food retailers market share in Portugal from 4th quarter 2021 to 4th quarter 2023, by brand (adapted from *Statista*, 2023; own elaboration)

Portuguese Food Retailers are classified into four types depending on the physical store format, characteristics, size, and the variety and extent of their assortment, namely: Hyper, Super, Discount, and Convenience (Appendix 1). According to Statista on the market share of Food retailers in Portugal by Category between 2019 and 2022, hypermarkets (*Auchan* and *Continente*) and supermarkets (*E.Leclerc, El Corte Inglés, Intermarché, Mercadona*, *MiniPreço* and *Pingo Doce*) held the highest market share at 66.8% in 2022. Meanwhile, *Aldi* and *Lidl* have gained increasing relevance in the Discounts category, achieving a market share of 17.9%. Convenience stores such as *Amanhecer, Coviran, Meu Super* and *SPAR* are also growing in popularity for covering emergency needs, accounting for 15.3% of the sector in 2022 (*Portugal: food retailers market share 2022, by category*, 2022).

Retail companies are diversifying their businesses in response to challenges posed by the external environment. This includes investing in adjacent retail formats such as restaurants, coffee shops, health and wellness services, home furnishings, and telecommunications. This diversification strategy has helped mitigate the negative impacts of the COVID-19 pandemic and has the added benefit of building customer loyalty to the various services offered by a specific retailer, thereby satisfying a wide range of customer needs (*Portugal: The Portuguese Food Retail Sector | USDA Foreign Agricultural Service*, 2021). Another important aspect is the need for innovation in the Food Retail sector to address the challenges of climate change and natural calamities. Consequently, sustainability has garnered increased investor attention over the past years. Portuguese food retailers are adopting the ESG (Environmental, Social, and Corporate Governance) framework to assess their performance on sustainability and ethical practices. In practice, they are implementing initiatives to reduce their environmental footprint, promote responsible consumption, lower operating costs, and improve business viability in the face of shifting regulations. Furthermore, retailers are also working towards reducing food waste and its associated societal costs by donating unsold food to charities and food banks. Currently, Portuguese food retailers have efficient solutions in place, such as partnerships with NGOs and discounts on products nearing their sell-by date. During the COVID-19 pandemic, retailers supported institutions by distributing food and non-food products to vulnerable individuals and health and safety professionals. These initiatives aimed to enhance brand image, attract environmentally conscious customers, and consequently, increase business revenues (Cereja, 2018).

2.1.1. Changes in Food Retail purchasing patterns during the crisis period

COVID-19 revealed weaknesses in the global food supply chain. Disruptions in transportation, shortages in labour and a surge in demand for specific products posed challenges to the Food Retail industry in Portugal. Two significant trends emerged from this situation: a shift in consumer behaviour and a growing dependence on e-commerce for grocery shopping (Roggeveen and Sethuraman, 2020; Faghih and Forouharfar, 2022).

At the beginning of the pandemic, panic buying and stockpiling led to temporary shortages of essential food and household items. Retailers were forced to prioritize hygiene and safety measures such as contactless payment, curbside pickup and store sanitation. Consumer spending was significantly affected by the crisis, with rising unemployment rates and increased price sensitivity leading to an 8% reduction in European households’ expenditure in 2020 (Zwanka, 2022). Portuguese consumers have reduced the frequency of dining out and increased eating at home, leading to a preference for private labels due to their lower prices compared to manufacturer brands (Pinto *et al.*, 2022). Additionally, there was a growing demand for healthier and more sustainable food choices, such as organic and locally sourced products. Retailers responded by innovating their own-brand product range, including new “free-from” products (no sugar, additives/preservatives, and allergens), meat substitutes and options for specific diet regimes (da Costa *et al.*, 2023). Also, a segment of consumers with higher disposable income occasionally indulge themselves by buying non-essential items such as gourmet and premium quality goods.

In Portugal, the retail food categories most affected by COVID-19 were tree nuts, due to increased snack consumption at home; alcoholic beverages, whose sales only increased in the first months, given the ad hoc government restrictions on purchases after 8 pm; and fresh produce, which initially saw greater demand for packaged goods due to uncertainty surrounding virus transmission. In contrast, ready-to-eat and ready-to-cook meals suffered a drop in sales as people had more time to prepare meals at home (USDA Foreign Agricultural Service, 2021). According to the publication “Estatísticas do Comércio – 2020” (INE, 2021), INE identifies the best-selling categories in food retail in 2020, as presented in Appendix 2.

Moving to e-commerce has advantages in terms of efficiency and convenience, and the omnichannel strategy enables retailers to capture the benefits of both online and offline channels. The COVID-19 pandemic led to a significant increase in e-commerce activity in the Food Retail sector, particularly in online grocery shopping and home delivery services (Gomes & Lopes, 2022). Moreover, fully online retailers as *360hyper* and *Mercadão*, have leveraged their share in the Portuguese industry. In 2020, 44.5% of Portuguese residents aged between 16 and 74 made an online purchase in the previous 12 months, with clothing retail being the main product category (60.4%), followed by food takeaway and home delivery (38.2%), and IT retail (1.4%) (INE, 2021).

When not purchasing online, consumers opted for more frequent shopping trips to smaller stores that are easily accessible from home, as opposed to larger hypermarkets and supermarkets. This growing demand for convenience and proximity has led to investments in new services, such as click-and-collect and pick-up points, which offer the benefits of online shopping with immediate access to products (Gomes et al., 2023). Food retail chains have been increasingly integrating omnichannel experiences into their strategies, allowing for data collection from online interactions and enhancing the relationship between retailers and shoppers in physical stores. Despite the rise of online sales, it is expected that physical stores remain the largest and most significant sales channel for retailers in the next few years.

2.1.2. Impacts on the Portuguese Sweet Snacks Market

The Portuguese Confectionery and Snacks market (commonly referred to as the Sweet Snacks market), is characterized by the distinctive taste, texture, appearance, and significant sugar content of its food items. The market comprises two main segments: Confectionery and Snacks. The Confectionery segment is divided into four subsegments: Chocolate (boxed chocolates, count lines, straight lines, and moulded bars); Gum (bubble and chewing gum); Sugar confectionery (hard-boiled sweets, mints, caramels, toffees, and marshmallows); and Preserved pastry goods and cakes (pies, tarts, doughnuts, croissants, and scones). The Snacks segment includes Sweet (cookies and crackers) and Salted snacks (tortilla chips, potato chips, and pretzels) ('Chocolate Confectionery in Portugal', 2023; *Confectionery - Portugal | Statista Market Forecast*, 2023).

Portuguese supermarkets offer a wide variety of sweet snacks from popular international brands, to which the Portuguese public has been very receptive. The Sweet Snacks category grew globally as it satisfied the snacking cravings of many families confined at home, positively impacting many international players in the packaged food industry. However, in the Portuguese market, the demand for sweet snacks was expected to decrease due to the imposed lockdown.

In this context, five trends have been identified that reflect the post-COVID-19 behaviour of Portuguese consumers and remain relevant in the current sweet snacks market. These trends include snacking at home, the "premiumisation" of purchases, a shock to loyalty and a shift toward private label brands, sustainability and ethical consumerism, and an alignment with health and wellness choices. Firstly, the Sweet Snacks category has seen a boost in consumption due to the increasing trend of enjoying snacks and pastries at home with a focus on "premiumisation". This includes the occasional purchase of premium quality products like chocolates and bonbons, as more people seek innovation, diversity, and indulgence in their at-home dining experiences. After the global recession, private labels have gained prominence in the Portuguese market offering affordable alternatives targeted at budget-conscious consumers. Manufacturer brands have seen a decline in market share (*World Market for Packaged*

Food, 2021). A McKinsey survey revealed that 73% of Portuguese consumers have adopted new shopping behaviours, including trying own-brand products (*Consumer sentiment in Portugal during the coronavirus crisis | McKinsey*, 2020). In this sequence, by 2023 the *Continente* private label was leading the category with 17.2% market share, followed by *Pingo Doce* with 13.6% (*Sweet Biscuits, Snack Bars and Fruit Snacks in Portugal*, 2023). The pandemic has raised awareness of environmental issues, leading to an increase in the number of eco-active consumers who adopt ethical and sustainable consumerism (Dimitris Skalkos and Zoi C. Kalyva, 2023). “Green consumers” not only consider the sustainability of ingredients but also of the packaging material, favouring materials such as paper, cardboard, and aluminium over plastic. Consumers are now relying more on official certifications to guide their purchasing decisions since they are not experts. Portugal is aligned with the global health and wellness trend, with the pandemic strengthening demand for healthier sweet snack options and transparent labelling (*Health and Wellness in Portugal*, 2022). Brands have been incorporating solutions with cleaner attributes, such as “free-from” and organic snacks. After the COVID-19 pandemic, manufacturers must re-evaluate their recipes to balance both physical and emotional well-being considerations.

2.2. Manufacturer WOW Portugal in the Sweet Snacks Market

Established in 2012, Manufacturer WOW Portugal is a subsidiary of Manufacturer WOW International, one of the world’s largest snack and confectionery companies with a presence in more than 150 countries. The subsidiary manages the distribution and marketing of WOW International’s popular brands in Portugal.

WOW International is a US-based company with global net revenues close to $32 billion in 2022. The Biscuits category accounts for half of total revenue, followed by the Chocolate contributing with 30%, and Gum and Candy making a further 11% of total sales. The Cheese and Grocery, and Beverages categories represent 6% and 3% of global revenue, respectively. Europe is the most important market. WOW International has a network of manufacturing plants in several geographical regions, ranked by its percentage of revenues: Europe (36%), North America (31%), Asia, Middle East and Africa (21%), and Latin America (12%) (Manufacturer WOW International Annual Report 2022).

In Portugal, WOW is known for its prominent presence in the Sweet Snacks segment among manufacturer brands. WOW Portugal is the market leader in Biscuits, Chocolates and G&C, operating with 19 portfolio brands. In the Biscuits category, the company sells cookies, crackers, and snack bars under ten brands: *Bite, Carousel, Crispy, Lucky, MammaMia Biscuits, Minty, OhSugar, Plum, Pop,* and *Temptation*. The organisation produces its chocolate in various formats including bars, tablets, and wafers using trademarks like *Cocoa, Craving, MammaMia*, and *Treat*. As for Gum and Candy, it offers chewing gums, candies, and oral health mints and markets these products as *Bliss, Bubble, Cotton, Honey,* and *Tidbit* brands. The broad range of products is sold through modern (hypers, supers and discount stores), traditional channels (convenience stores and gas stations), and Cash and Carries (Manufacturer WOW Portugal Website 2023). The distribution network is handled through own and satellite warehouses, direct store deliveries, and distribution centres to ensure timely deliveries.

2.3. Time Series analysis in Data Mining

In today's era of abundant and diverse information access, there is an increasing urgency to structure time series data. Thus, there is a rising demand for new data-driven analysis tools that can extract valuable information and deliver actionable insights to users. These techniques fall under the field of Knowledge Discovery in Databases (KDD) (Klösgen and Zytkow, 2002), a term often used in the literature interchangeably with Data Mining (DM), although there remains a lack of consensus on the distinction between the two. KDD is a multi-disciplinary concept that emerges from the intersection of research fields such as databases, statistics, machine learning, artificial intelligence, data visualisation, and knowledge acquisition for expert systems (Armstrong, 2001; Siguenza-Guzman *et al.*, 2015).

The Knowledge Discovery (KD) can be framed in a sequence of activities aimed at extracting value from data, with Data Mining (DM) placed at the core of this process. DM applies specific algorithms to identify patterns in time series data, given the computational limitations (Hand, Mannila and Smyth, 2001). While DM constitutes one step in this process, KDD refers to the entire non-trivial and interactive process of identifying valid and useful patterns in databases, evaluating the outcomes produced by DM (Fayyad, Piatetsky-Shapiro and Smyth, 1996). The Knowledge Discovery and Data

Mining (KDDM) Models strive to ensure that the generated output is useful for end-users (Kurgan and Musilek, 2006), with the CRISP-DM model being a prominent methodology used in this project. The KDD process flow is illustrated in Figure 2.

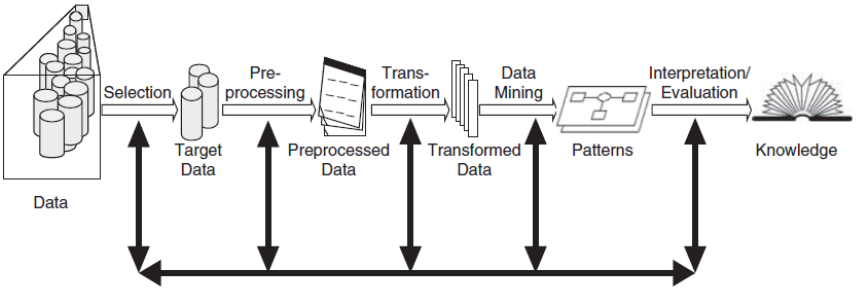


Figure 2. Overview of the steps of Knowledge Discovery in Databases Process (Mariscal, Marbán and Fernández, 2010, p. 8)

In brief, the process begins with the selection and integration of a set of data, which can come from various sources. This is followed by data cleaning i.e. techniques for handling missing values and removing noise during the pre-processing phase. Subsequently, the data is transformed to align with the requirements of the DM algorithm, which looks for patterns in the representational form of data. The relevance of these extracted patterns is assessed against predefined criteria, with visualisation techniques aiding in their interpretation. The knowledge obtained is then practically consolidated either by incorporating it into existing systems or by its documentation and release to interested parties (Fayyad, Piatetsky-Shapiro and Smyth, 1996).

Indeed, DM tasks can be classified based on the ultimate purpose of the KDD process as descriptive or predictive (Gibson *et al.*, 2007). Descriptive tasks focus on uncovering patterns that summarise hidden relationships in data such as associations, correlations, clusters, segmentation, and anomaly detection. In contrast, predictive tasks attempt to predict the value of a particular attribute based on observed patterns. Forecasting is a predictive DM task and is particularly relevant in this project. In this way, time series forecasting is applied across various business sectors inside organisations. It serves as a crucial tool for monitoring and optimising resources, refining business processes, and aiding medium and long-term strategic decision-making, thereby providing more accurate projections (Faloutsos, Flunkert, *et al.*, 2019; Petropoulos *et al.*, 2022). To conclude, the current dynamic business landscape demands higher quality forecasts, which poses a challenge for new solutions to improve existing algorithms.

2.3.1. Time Series Forecasting Process

The Time Series Forecasting process makes use of historical time series data collected at regular intervals over time and a forecasting model. The forecasting model reflects data patterns through a statistical relationship between past and current values of a given variable, intending to project these data patterns into the future. Time series forecasting is a quantitative forecasting technique that can be applied to prediction problems when two conditions are met: there is numerical information regarding the past and it is assumed that some past events' features will continue into the future (Montgomery, Jennings and Kulahci, 2008; Petropoulos *et al.*, 2022). Formally, the Time Series Forecasting process is expressed as

$$\hat{y}_{T+h|T} = y_{T+h} + \varepsilon_{T+h|T} \,, \tag{1}$$

where y_{T+h} is the $(T+h)th$ observation with $h = 1, \dots, H$ based on a set of variables observed at period T; H is the forecast horizon, i.e. number of future periods for which forecasts are produced; period T is the forecast interval, i.e. frequency of new forecasts; $\hat{y}_{T+h|T}$ is an h-step forecast y_{T+h} for the value of variable Y taking into account all observations up to time T; $\varepsilon_{T+h|T}$ is the error associated with the forecast, i.e. $\varepsilon_{T+h|T} = y_{T+h} - \hat{y}_{T+h|T}$, the difference between the observed value and the predicted value for each $(T+h)th$ observation.

Choosing an appropriate forecasting model for Y depends significantly on the available information and the accuracy of selected models. Predictor variables are often used for time series forecasting when it is known that they impact the time series target, although this relationship is not always exact and might not cover all the effects of variation. These models are called explanatory models and incorporate external variables that can affect the system besides the historical data to generate forecasts, however, time series models should be preferred as they are more accurate in this task. When it comes to the latter, it is challenging and time-consuming to forecast the future values of different predictor variables to obtain the forecast of a response variable. As such, these models are useful for forecasters interested in the impact of external factors on the time series instead of solely predicting future values (Shumway and Stoffer, 2011; John E. Hanke and Dean Wichern, 2014; Keith Ord, Robert Fildes, and Nikolaos Kourentzes, 2017).

The Time Series Forecasting process, illustrated in Figure 3, starts with collecting historical data relevant to the variables to be forecast. Next, the data is processed into the proper format through pre-processing tasks, such as identifying missing values, detecting outliers, and filtering the time series. The visual inspection of time series plots is an essential step in identifying recognizable patterns including trends, seasonal, or cyclical components. Subsequently, a set of forecasting models is selected and fitted to the data, meaning that model parameters are estimated using the Least Squares method. These models are then evaluated and compared using accuracy measures to determine their performance on the data. At last, once the model has been validated, forecasts are produced attending to the appropriate forecast horizon (Armstrong, 2001; John E. Hanke and Dean Wichern, 2014). This process is explained in more detail in subsections *2.3.1.1.* to *2.3.1.4.*.

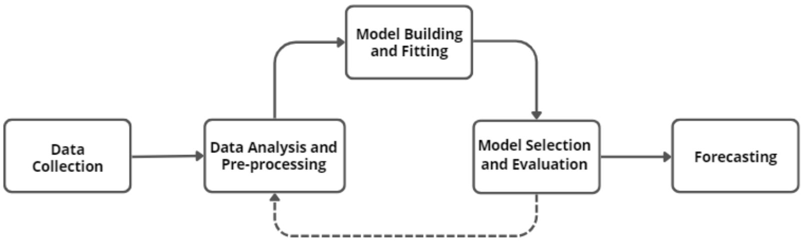

Data Collection
Data Analysis and Pre-processing
Model Building and Fitting
Model Selection and Evaluation
Forecasting

Figure 3. Time Series Forecasting Process (adapted from Montgomery, Jennings and Kulahci, 2008; own elaboration)

2.3.1.1. Data analysis and pre-processing

The second step of the process can be divided into two parts. First, visualising the time series plots and using auxiliary metrics to identify patterns in the data. Second, performing pre-processing tasks to ensure the time series data is in the correct format to be ingested into the forecasting models. The Data analysis and pre-processing step is covered in Appendix 3.

2.3.1.2. Model building and fitting

In the current practice of forecasting, statistical forecasting methods are predominantly used among practitioners, while ML and DM methods are increasingly mentioned in the literature as alternatives. According to researchers, the latter are less accurate and require greater dependence on computer science, underlining the supremacy of statistical methods (Makridakis, Spiliotis and Assimakopoulos, 2018, p. 2). This project focuses on statistical time series forecasting methods, in particular, the non-seasonal Autoregressive Integrated Moving Average (ARIMA) model. In addition, it covers Prophet, an algorithm that accommodates additional components such as seasonality, trend, and holidays.

Forecasting Models: ARIMA and Prophet

A widely recognised contribution to the field of statistical methods in time series forecasting is the Box-Jenkins Method. The authors proposed a practical approach for modelling linear time series, which can either be stationary or non-stationary. The method constitutes a valid inference in contrast to the misleading inference that might be obtained from simple regression (George E. P. Box *et al.*, 2015). At an early stage, models with stationary behaviour were predominant however, non-stationarity models became the dominant ones in diverse areas of application (Fildes and Makridakis, 1995). Box-Jenkins's popularity derives from its versatility in handling different time series patterns across a wide range of models and provides high accuracy in short and medium-term forecasts (Makridakis *et al.*, 2021). The Box-Jenkins class of models are composed of three key components, each of which helps to model a certain type of pattern:

- The Autoregressive (AR) component is denoted by $AR(p)$. It corresponds to the autocorrelation component of a time series and uses a linear combination of previous values to predict future values (Shumway and Stoffer, 2011).
- The Integration (I) component is denoted by $I(d)$. It refers to the number of differences required to obtain a stationary time series. When differencing, it is necessary to ensure adequate modelling and the time series is labelled as an integrated process of order d (George E. P. Box *et al.*, 2015).
- The Moving Average (MA) component is denoted by $MA(q)$. It connotes the moving average component of a time series and uses a linear combination of past prediction errors, which constitute unknown factors that affect the time series but are not explained by their past values (George E. P. Box *et al.*, 2015).

The ARIMA-based models result from the combination of these components. Accordingly, the non-seasonal Autoregressive Integrated Moving Average (ARIMA) model takes p, d and q as parameters and its formula is given in Eq. (2). When dealing with time series data with seasonal patterns, an extension to the standard model is used instead (Shumway and Stoffer, 2011; George E. P. Box *et al.*, 2015). The seasonal Autoregressive Integrated Moving Average (SARIMA) Model is parameterised with three additional orders P, D and Q that correspond to the seasonal AR, I and MA components and it is defined by Eq. (3).

$$ARIMA(p,d,q): \qquad \phi p(L)\ \Delta^d y_t = c + \theta q(L)\ \varepsilon_t \tag{2}$$

$$SARIMA(p,d,q)x(P,D,Q)s: \qquad \phi p(L)\ \Phi P(L^s)\ \Delta^d \Delta_s^D\ y_t = c + \theta q(L)\ \Theta Q\ (L^s)\ \varepsilon_t \tag{3}$$

where $\phi p(L)$ corresponds to the non-seasonal autoregressive polynomial of order p; $\theta q(L)$ corresponds to the non-seasonal moving average polynomial of order q; $\Phi P(L^s)$ corresponds to the seasonal autoregressive polynomial of order P; $\Theta Q\ (L^s)$ corresponds to the seasonal moving average polynomial of order Q; $\Delta^d \Delta_s^D\ y_t$ is the time series with $d$differentiations and D seasonal differentiations; s is the seasonal period; c is a constant; and $\varepsilon_t \sim N(0, \sigma^2)$ is a white noise series.

Appendix 4 shows the process of building and fitting an ARIMA-based model for forecasting, presenting two alternative procedures: Box-Jenkins method and Automatic ARIMA modelling. The Box-Jenkins method comprises three steps that culminate in the effective forecasting step as described in Appendix 5. The alternative procedure was proposed by Hyndman and Khandakar (2008), which involves generating the model automatically using the *auto.arima()* function from the *forecast* package in R. This function automates the selection process of the best fit for the p, d, q parameters of an ARIMA model, based on a combination of unit root tests and minimisation of information criterion (Awan and Aslam, 2020). The benefits of the Automatic Modelling method point to more accurate model performance and suitable predictions, besides the evident fact of being simpler and more efficient in obtaining results.

Prophet model is a non-parametric time series forecasting algorithm developed by Facebook in 2017 for business-related applications. It extends the classic decomposition model by incorporating additional components such as seasonality, trend and holidays (Taylor and Letham, 2017). Prophet was created to optimise the wide range of business forecasting tasks, which follow all or some of the characteristics outlined: Regular nature, which implies hourly, daily or weekly observations of several months and at least a year of history; Strong repeated seasonality associated with day of the week or time of the year; Significant holidays, which are known in advance such as New Year, Valentine's Day, Carnival, Easter and Christmas; Not too many missing observations or significant outliers; Historical trend changes due to internal organisational procedures or external factors; Stochastic process, working well with non-regular spaced measurements and thus does not require stationary time series (Taylor and Letham, 2017). Several studies conducted in the field of time series forecasting, show that Prophet provides significant improvements and greater accuracy for variables under study over ARIMA-based models, particularly in the presence of seasonality (Duarte, Walshaw and Ramesh, 2021).

The Prophet time series model is represented by three main model components: trend, seasonality and holidays, which are combined as in Eq. (4) (Taylor and Letham, 2017).

$$y(t) = g(t) + s(t) + h(t) + \varepsilon_t \tag{4}$$

where $y(t)$ corresponds to the value observed in the time series in period t; $g(t)$ corresponds to the general trend of the time series; $s(t)$ corresponds to the seasonality component of the model; $h(t)$ represents the influence of holidays and special events; and ε_t represents the idiosyncratic changes, assuming a normal distribution. Prophet's main forecast component is the trend term $g(t)$ which defines how the time series has developed previously and how it is expected to continue. There are two types of models, depending on data characteristics: a non-linear saturating growth model, where the growth is non-linear and expected to saturate at a carrying capacity, and a piecewise logistic growth model, where the growth rate is stable and linear. The model applies linear fitting to ensure it is not affected by outliers or missing data. Also, business time series often have multiple seasonal patterns occurring at weekly, monthly, or yearly intervals due to the repeated human actions. Depending on how seasonality $s(t)$ affects the time series, it can be modelled as additive or multiplicative. To capture periodic data patterns, Prophet employs the Fourier series (Harvey and Shephard, 1993). Holidays and special events $h(t)$ provide a relatively significant and predictable influence in business time series. Since they do not follow a periodic pattern, their effects are not well-modelled by a smooth cycle. For this reason, Prophet offers functionality to include a custom list of holidays identified by their unique name, both past and future, based on the assumption that their effects are independent.

In short, the Prophet forecasting approach offers several practical advantages for analysts over other time series models, which are mentioned below (Wang *et al.*, 2023). The model is simple to use and works efficiently with large volumes of data, facilitating model tuning as it often operates well with default parameters. Prophet enables the incorporation of component effects, by adjusting or including new parameters to impose different assumptions. Users can directly specify changepoint dates if they have prior knowledge of relevant trend periods, and input holiday dates and appropriate seasonal time scales that will impact growth in certain locations. It also has less sensitivity to outliers and missing data, allowing for more accurate forecasts. Unlike other models such as ARIMA, Prophet does not explicitly account for temporal dependence structure in the data and for the need to interpolate missing data and remove outliers.

2.3.1.3. Model selection and evaluation

In DM, an algorithm has two types of parameters, model parameters and hyperparameters, both essential for defining the best-performing model in a data set. The parameters are estimated using the training data set, whereas hyperparameters are adjusted according to the data specificities before the training phase (Ding, Tarokh and Yang, 2018). Table 1 displays these two types of parameters for the ARIMA-based Models. The Prophet model is non-parametric and is therefore not included.

TABLE 1. HYPERPARAMETERS AND PARAMETERS FOR ARIMA-BASED MODELS

	Hyperparameters	**Parameters**
ARIMA	p, d, q	$\phi p, \theta q, c, \sigma^2$
SARIMA	p, d, q, P, D, Q	$\phi p, \Phi P, \theta q, \Theta Q, c, \sigma^2$

where $\phi 1, \ldots., \phi p$ indicates the number of autoregressive terms; $\Phi 1, \ldots, \Phi P$ indicates the number of seasonal autoregressive terms; $\theta 1, \ldots, \theta q$ indicates the number of lagged forecasting errors; $\Theta 1, \ldots, \Theta Q$ indicates the number of seasonal lagged forecasting errors; c is a constant; and σ^2 is the variance of residuals.

The ARIMA-based Models are estimated by maximum likelihood. Thus, information criterion can be applied to these models when carrying out the model selection procedure. Information criterion are measures of the relative quality of candidate models, considering the trade-off between their goodness of fit based on the MLE and their complexity assessed through the number of parameters (Nakamura *et al.*, 2006; Dziak *et* *al.*, 2019; Zhang, Yang and Ding, 2023). Different criterion are distinguished given the latter penalty term k. The most frequently used are the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in Eq. (5) and (6), respectively (Bozdogan, 1987; deLeeuw, 1992; Neath and Cavanaugh, 2012). When dealing with small sample sizes, it is recommended to use AICc, a corrected version of AIC as defined by Eq. (7) (Hurvich and Tsai, 1989).

$$AIC = -2log(L) + 2K \tag{5}$$

$$AICc = AIC + (2K(K+1))/(n-K-1) \tag{6}$$

$$BIC = -2\,log(L) + Klog(n) \tag{7}$$

where L is the logarithm of the maximum likelihood estimate, $K = p + 2$ is the number of parameters in the model; and n is the number of time series data points.

Typically, AIC and AICc are preferred, but any of the three criteria are valid as long as these are minimised (Kuha, 2004; Medel and Salgado, 2012). These criteria only apply to selecting the values of p and q, since models with different orders of differencing (d) are not comparable. These measures simplify model selection for forecasters, by reducing the in-depth split of training data for the test set and facilitating the identification of overfitting models (Lever, Krzywinski and Altman, 2016). It is important to emphasise that information criterion apply only to statistical models based on likelihood estimation and are therefore not applicable when selecting Prophet models.

Once the best model(s) have been retained, their performance on the test set is evaluated using the minimisation of predictive accuracy measures. Unlike previously, these measures can be applied to ARIMA and Prophet time series models, in addition to other ML algorithms. This allows for a proper comparison to select the best forecasting method (Hyndman and Koehler, 2006). The set of accuracy measures is centred on squared or absolute errors, where the most used being Mean Squared error (MSE) in Eq. (8), Root Mean Square Error (RMSE) in Eq. (9), Mean Absolute Error (MAE) in Eq. (10), and Mean Absolute Percentage error (MAPE) in Eq. (11).

$$MSE = mean(e^2) \tag{8}$$

$$RMSE = \sqrt{mean(e^2)} \tag{9}$$

$$MAE = mean(|e|) \tag{10}$$

$$MAPE = 100\,mean(|e|/|y_{T+h}|) \tag{11}$$

where $e = y - \hat{y}$, being y the actual value and $\hat{y}$ the predicted value; y_t denotes the time series.

2.3.1.4. Forecasting

In the forecasting task, analysts must also consider two aspects, namely the forecast horizon and the inherent uncertainty of forecasts. As defined in Eq. (1), an h-step forecasting horizon is established in the forecasting process. Thus, two possible cases are considered: one-step-ahead forecasting ($H = 1$), which is the default behaviour, and multi-step-ahead forecasting ($H > 1$), which is preferred by specialists since single-step forecasting is very short-sighted and provides insufficient information about the future. The recursive method is the most popular method for multi-step forecasting, which uses the previous step's forecast to predict the immediate next future step repeatedly until the end of the forecast horizon (Kline, 2004). This approach applies to several forecasting models, including the ARIMA and Prophet.

To overcome the uncertainty underlying the forecasting process, it is necessary to generate prediction intervals in addition to point forecasts (John E. Hanke and Dean Wichern, 2014). For ARIMA models, prediction intervals are calculated based on the standard deviation of residuals (σ^2), assuming they are uncorrelated and normally distributed. ARIMA-based intervals tend to be too narrow and generally widen as the forecast horizon increases (Tayman, Smith and Lin, 2007; Fanoodi, Malmir and Jahantigh, 2019). In Prophet Models, the uncertainty is incorporated into the model components and calculated through a Monte Carlo simulation, generating several possible future scenarios from random samples. The uncertainty of the estimates can be controlled by adjusting the prediction interval width using the parameter *'interval_width'*. A wider interval reflects a higher level of uncertainty, while a narrower interval implies higher confidence in the prediction (Kennedy *et al.*, 2011).

3. CRISP-DM METHODOLOGY

The main objective of this project is to understand the impact of COVID-19 on retailers' sales by comparing pre-pandemic and post-pandemic Sweet Snacks sales trends to identify potential deviations attributed to the crisis. Therefore, this project employs two main theoretical approaches: Business Intelligence (Appendix 6 presents an overview of

BI concepts essential for the understanding of Data Modelling techniques) and a DM project that comprises the forecasting models.

The CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology was selected to structure the phases of this research and integrate business intelligence within the DM framework. This methodology was created in 1996 by a group of leading companies (Daimler-Benz, Integral Solutions Ltd. (ISL), NCR, and OHRA) and provides a standardized framework that is industry-agnostic and technology-independent (Costa and Aparício, 2020, p. 3). Its design allows use by a broad range of users, from data mining experts to those with less technical expertise, facilitating its application across various projects (Wirth and Hipp, 2000, p. 1).

CRISP-DM consists of six main stages: Business Understanding, Data Understanding, Data Preparation, Modelling, Evaluation, and Deployment (Schröer, Kruse, and Gómez, 2021). The lifecycle of a DM project is depicted in Figure 4. This hierarchical and cyclical process incorporates a strong iterative component, allowing flexibility in the sequence of phases based on the project's specific context and needs (Wirth and Hipp, 2000, p. 4). According to the CRISP-DM Guide, these phases are described as in Table 2.

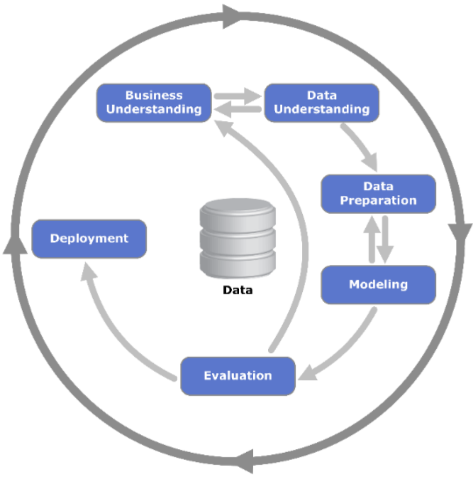


Figure 4. Phases of CRISP-DM Methodology (Costa and Aparício, 2020, p. 3)

TABLE 2. DESCRIPTION OF CRISP-DM PHASES (Chapman, 2000)

CRISP-DM Methodology	Description of Phases
Business Understanding	The initial phase involves an in-depth study of the business area to identify the main and secondary objectives as well as the project requirements. After this, it is possible to precisely formulate the DM problem to be handled and design the preliminary project plan.
Data Understanding	Starts with initial data collection, followed by data exploration. An exhaustive analysis is conducted to describe the variables and test the data quality.
Data Preparation	Consists of selecting the final data set to which modelling techniques will be applied. This involves processes such as data cleaning, integration and formatting.
Modelling	Selects a set of algorithms and modelling techniques through an interactive process. Hyperparameter optimisation is performed to attempt to obtain one or more reliable models for the proposed business objective.
Evaluation	Involves review and assess the accuracy of results returned by the models, considering potential business issues. The final model is then elected.
Deployment	Refers to the strategy of implementing results in an organised way and making them accessible to the end user in the form of a system. Subsequently, ensure its monitoring and maintenance.

The overall objectives were introduced in section *1.2.* and are further extended in the section *3.1.*, into more practical and concrete goals. In this sense, the six steps described in the methodology are framed to the sales data set which constitutes the object of analysis. To develop this project, Microsoft Power BI was the core tool for data processing, with further integration of the R advanced analytics software to leverage its modelling capabilities as detailed in Appendix 7.

3.1. CRISP-DM Phases

3.1.1. Business Understanding

This project used sales data from Manufacturer WOW Portugal concerning two Portuguese food retail banners, the Retailers Miam and Munch. The data set was provided by the Consulting company. Accordingly, the data had to be properly anonymised to comply with privacy requirements.

This study aims to contribute to the repository of literary research undertaken in the context of a socio-economic crisis focused on the national Food Retail sector. As such, it provides an analysis of the evolution of the Sweet Snacks food retail category between January 2018 and June 2023. The primary user of the final BI methodology is the company *TouchPoint Consulting*. Nevertheless, it can be consulted by various entities such as stakeholders and industries operating in the Sweet Snacks category, as well as in other food categories.

3.1.2. Data Understanding

The second phase of the process refers to data collection and analysis, description, exploration, and a set of crucial quality verification tasks. The Extract, Transform and Load (ETL) process can be defined as the storage process that consists of extracting, transforming, and loading integrated and clean data into the DW, acting as an intermediary between operational source systems and the BI presentation area (Joe Caserta and Ralph Kimball, 2004). The Data Understanding stage corresponds to 'Extract' in the ELT process, suited to project requirements.

The source data was collected from the retailers' data management platforms with restricted access. From these platforms, the respective daily sales reports were extracted in Excel format for the period under analysis from January 2018 to June 2023. For Miam, the daily data was compiled in a single Excel file, while for Munch it was divided into four files. As regards Data Requirements, they comprise a single Excel file updated to 30th June 2023, provided by Manufacturer WOW Portugal. After data collection, the Excel files were aggregated into twelve folders organised by year and retailer for more efficient loading into the Power BI application.

The first set of data contains daily sales for the sweet snacks category of Miam, while the second data set covers Munch for the period from January 2018 to June 2023. Both include information on the sale date, store name and identification code where the transaction took place, and the geographic region of Portugal where the store is located (covering North, South, and Islands). In addition, information was gathered on the item's identification code, brand, and category segmentation. Furthermore, the unit sales quantity data, which represents the number of units sold for each item, also includes records with negative quantities indicating product returns. The Data Requirements were compiled across various Excel sheets, focusing solely on relevant information essential for model construction. This involved retrieving details on retailer identification, banner, store, store region, and product scope, including its identification and hierarchical categorisation. The item hierarchy follows a structured depth as business, category, segment, brand, subcategory, and product code. Consequently, the analysis of the collected data identified ten distinct variables and their type. This information is summarised in Table 3, along with a brief description of the variables.

TABLE 3. DESCRIPTION OF VARIABLES FOR RETAILERS' SALES TABLE

Variable Name	Description	Variable Type
SalesDate	The period under analysis.	Numeric Discrete
Retailer	Designation of the Portuguese retailer.	Categorical Nominal
Banner	Designation of the banner owned by a particular retailer.	Categorical Nominal
Region_Store	Store location in Portugal by geographical region.	Categorical Nominal
Store	Store identification code.	Categorical Nominal
Category	Designation of product category associated with a particular item.	Categorical Nominal
SubCategory	Product subcategory code associated with a particular item.	Numeric Discrete
Brand	Designation of the brand associated with a particular item.	Categorical Nominal
Item	Item code.	Numeric Discrete
SalesUnits	Number of units sold of a particular item.	Numeric Discrete

Data quality concerns the level of accuracy, completeness, and consistency of data in a data set. High data quality is critical for meaningful analysis and accurate data-driven decisions. This includes data profiling, cleansing, enrichment techniques, and ongoing quality control (Heinrich *et al.*, 2017; Günther *et al.*, 2019). The reliability of the sources was assured since they concern two leading national retailers and were obtained from a major private organisation in this market segment. Moreover, the data sources are local i.e. collected per store, making them more accurate. However, data from different sources had inconsistent formats, necessitating the alignment of retailers' data sets to a common format. In this way, data profiling examined the data structure to detect potential patterns and anomalies. Data cleansing was then performed to correct gaps and incompleteness by removing missing values, duplicates, and irrelevant data. In addition, the data set was enriched using the multi-source merging technique. Data quality procedures were applied whenever deemed necessary throughout the project. The routine assurance of data quality was guaranteed through frequent updates, best-practice data handling, and error recognition. There has been an ongoing process of data quality monitoring by the manufacturer and updates to Data Requirements during the selected period of analysis.

3.1.3. Data Preparation

The third phase of the process focuses on loading and subsequently, integrating and cleaning the extracted data, which corresponds to the 'Load' and 'Transform' of the ELT process. To note the project data set was anonymised to protect the Manufacturer's identity and to guarantee confidentiality of source data. The source data was masked with fictitious names using the Data Substitution technique, which was executed in Power BI.

In Power BI, data sets were organised into tables. Appendix 8 shows how data was imported from the Excel data requirements file *(1),* followed by selecting the relevant tabs *(2)* and loading the data *(3).* A similar process was conducted for each of the twelve sales folders shown in Appendix 9. In the same way, it starts with importing the folder *(1),* identifying the folder path *(2)* and loading its data *(3).* After the 'Load' of data, a sequence of transformations was applied to the various data tables to obtain the final tables that will be part of the Dimensional Model. To do this, the 'Transform Data' option was selected from the BI tool's Home tab and transformations were performed using the Power Query Editor functionality.

To create the Snowflake model, various tasks including data selection, cleansing and enrichment, formatting, exploration, and integration were performed. For a detailed description of the data preparation tasks, refer to Appendix 10.

3.1.4. Modelling

Once the Dimensional Model is created, the data can be analysed according to the Time Series Forecasting Process covered in section *2.3.1* of the Literature Review.

3.1.4.1. Time series Forecasting Process analysis in R

After reviewing the diverse literature on business series forecasting in the context of the COVID-19 pandemic, it was concluded that ARIMA, Holt-Winters additive and Prophet models are the predominant methods for time series forecasting. Among these, ARIMA models consistently demonstrate the best performance. The most commonly used accuracy metrics to evaluate model performance are RMSE and MAPE. In this sequence, both ARIMA and Prophet time series forecasting models will be applied to the sales data set.

The Modelling phase includes the steps of Data analysis and Pre-processing and Model building and fitting, which will be described for the three forecasting horizons. The DAX Studio tool connected to the Power BI model and, by executing a DAX query, it was possible to export the Sales Fact table data in a CSV file format. This provided flexibility for further analysis on the target fields in R. Firstly, the CSV file for each period was loaded into R Studio using the *read.csv()* function.

ARIMA: Data analysis and pre-processing, Model building and fitting

To obtain a readable sales data structure for ARIMA models, the data set was converted into a *tsibble* object. The *tsibble* was grouped by keys, and 'total_Sales' was calculated by applying the *summarise()* function to the sum of the measure 'SalesUnits'. Thus, the *tsibble* index is the attribute 'SalesDate', and the keys consist of the remaining columns: 'BannerID', 'Region_StoreID', 'StoreID', 'BrandID', and 'ItemCODE'. Each row in the *tsibble* represents the unit sales of a specific product from a brand, sold in a particular store owned by one of the retailers mentioned, located in a certain region, and associated with a date of sale. The Sales *tsibble* was divided into three subsets corresponding to the three periods under analysis: pre-crisis period (2018/01/01 – 2019/02/29); crisis period (2020/03/01 – 2021/12/31); and post-crisis period (2022/01/01 – 2023/06/30). In turn, the Period Sales *tsibble* was filtered by 'BannerID' to obtain three *tsibbles* for each retailer. As such, Retailer Miam *tsibble* considers Banners 'B09', 'B10', 'B14', and 'B17' (corresponds to Miam Supers, Miam Hipers, Miam&GO and Miam Wellness), while Retailer Munch *tsibble* considers Banners 'B01', 'B02', 'B03', 'B05', 'B06', and 'B08' (Munch Hub, Munch Central, Munch Street, Munch HomeTech, Munch Office and Munch Online). The dimension of the six *tsibbles* is included in Appendix 11.

To visually identify patterns in the data, a Time series Plot of total Sales was produced for each retailer over the different time horizons. Furthermore, the time series was decomposed, and the presence of potential trends and seasonal patterns was analysed, although in the absence of any evident seasonal pattern for the crisis and post-crisis period. It should be noted that the detection of seasonality was restricted to the two data sets mentioned covered less than two complete annual periods, making the analysis only accurate for the pre-crisis period. The trend component varies according to the model and will be covered in section *4* of the Results Analysis. Afterwards, it was applied a Yeo-Johnson transformation to stabilise the variance and make data distribution more symmetric and closer to a Gaussian distribution during the respective period (Riani, Atkinson and Corbellini, 2023). The transformation was proposed by Yeo and Johnson (2000) as an extension of the Box-Cox transformation method, allowing for a wider range of input data including zero and negative values and making it useful for skewed data As such, this transformation was considered the most suitable for the sales data set under analysis since it covers zero and negative values that a Logarithmic or Box-Cox transformation would not handle. Along with this, the Yeo-Johnson Transformed Time series was plotted to visualise the changes resulting from the stabilisation of sales behaviour throughout the periods. The unit root tests (ADF and KPSS tests) were computed to evaluate the stationarity of the time series, considering a significance level of 0.05. The results show that data is non-stationary, evidenced by a very small p-value, which indicates strong evidence against the null hypothesis. In this connection, the *unitroot_ndiffs()* function was applied to determine the recommended number of differences needed to achieve stationarity, which returned one order of differencing for yj-transformed time series data. This procedure was also complemented by visual inspection of the First Differenced Time series Plot of total Sales. To identify the potential values for lags *p* and *q*, the ACF and PACF for the first differenced sales time series were plotted. For the AR process, the PACF plot will show a sharp cutoff after lag *p*, and the ACF will decay gradually or remain significant for several lags. As for the MA process, the ACF plot will show a sharp cutoff after lag *q*, and the PACF will decay gradually or remain significant for several lags. Overall, positive values indicate a positive correlation between current and lagged observations, and negative values imply the opposite.

In the first instance, the *auto.arima()* function was employed to facilitate the selection of the best ARIMA model based on the lowest AIC and BIC values for the yj sales time series. However, while this function automates the selection process, it might not always find the actual best model, especially in cases where time series data exhibits complex patterns or requires specialised modelling techniques. Therefore, automatic selection was complemented by further manual evaluation of candidate models. In this sense, the three data subsets were divided into a training set with the first 500,000 observations and a test set with the last 200,000 observations (Appendix 12). During the model selection, the training and test sets were restricted to allow the comparison of models with different parameter orders, given the computational limitations of R software. The candidate ARIMA models were fitted on the training set and predictions were made on the test set. For each of the ARIMA obtained, the candidate models were chosen by adjusting the number of autoregressive and moving average terms by increasing or reducing the term by one or two, while maintaining the same differencing order ($d=1$).

Prophet: Data analysis and pre-processing, Model building and fitting

To implement the Prophet model, the data set was converted into a *data frame* and structured into two columns: ‘ds’ assigned to the index ‘SalesDate’ and ‘y’ attributed to the measure ‘SalesUnits’. Then the *data frame* was grouped by ‘ds’, and summarised according to the sum of the measure ‘y’. Each row in the *data frame* represents the total sales for a retailer associated with a date of sale. The procedure applied to the ARIMA was replicated to obtain three *data frames* for each retailer associated with temporal subsets. The dimension of the six *data frames* is indicated in Appendix 11. Moreover, Time series Graphs of total Sales were produced, as well as the decomposition of time series, to validate the consistency of seasonality and trend patterns observed in ARIMA.

For model building, the three periods were split into a training set of 80% and a test set of 20% of the data set (Appendix 13). The Prophet Model was fitted attending to holidays and special events added as regressors. The joint national holidays of both retailers are ‘Valentines_Day’, ‘Carnival’ and ‘Easter’, since Munch does not provide sales data for 25/12 (‘Christmas’) and 01/01 (‘New_Year’). As such, the parameter ‘extra_regressors’ was set to 3. The remaining model parameters returned when the *fit.prophet()* function was applied are briefly described below. A ‘growth’ rate of 1 denotes a linear growth where the trend increases or decreases steadily over time. In all models, ‘changepoints’ were set to 24, meaning 24 specific time points at which changes in the trend of time series were detected. In this sequence, ‘changepoint_prior_scale’ can be adjusted with a higher value making the trend more flexible, or a lower value imposing more trend regularisation. This parameter can be used to tackle two well-known issues of forecasting: overfitting and underfitting. As a final note, Prophet has detected two types of ‘seasonalities’ (components of ‘yearly.seasonality’ and ‘weekly.seasonality’ are set to 1), which implies that models include seasonal patterns occurring on an annual and weekly frequency within the time series periods. However, this does not necessarily mean that time series data inherently contains observed seasonality patterns, as stated above.

3.1.5. Evaluation

The Evaluation phase includes the steps of Model selection and assessment, which will be described below for the three forecasting horizons and the final Forecasting conducted in the post-crisis period.

ARIMA: Model selection, evaluation and Forecasting

The *glance()* function was used to compute the information criterion of fitted alternative models. This function provides a concise summary of various model fit statistics including R-squared, AIC, BIC, log-likelihood, and estimated residual standard deviation (sigma). Both AIC and BIC statistics perform well on large data sets, and therefore the best ARIMA model was chosen given the lowest values returned by two. Additionally, the accuracy metrics of MSE, RMSE, MAE, and MAPE were computed to compare simulated and observed values among the ARIMA. Unlike the model selection process in which a limited training and test set were applied, the model fitting considered a training set of 80% and a test set of 20% of the total observations (Appendix 12), to allow a final comparison between ARIMA and Prophet models, which will be detailed in section *4* of the Results Analysis. Once the best model was obtained, the pattern of the residuals was analysed using ACF plot and Portmanteau tests, specifically the *Box.test()* function with the ‘Ljung-Box’ and ‘Box-Pierce’ types specified. The function provided test statistics, degrees of freedom, and p-values for residuals, determining the presence of autocorrelation between residuals at lags 10, 30, 50 and 70 for all models, since these are random character data. ARIMA models with autocorrelated residuals indicate that there are patterns or dependencies present in the residuals, implying that the model might not capture all the underlying period data patterns. Nevertheless, it is important to note this does not invalidate the ARIMA for forecasting purposes.

For the forecasting phase, 366 period-ahead predictions were generated to cover the period from July 2023 to June 2024. A 95% confidence interval was used, in which the actual value of total sales is expected to be within the range estimated by the model. For a complete analysis, a plot of the forecasted values was executed to monitor the ARIMA projection of unit sales over the post-crisis period.

Prophet: Model evaluation and Forecasting

Prophet Models were evaluated by computing the accuracy metrics using two alternative methods, namely cross-validation and making predictions on test set. The outputs considered are from the second method, which was also employed in the ARIMA. Prophet includes functionality for automatic cross-validation across a range of historical cutoff points. It fits the model using data up to each cutoff point and then compares forecasted values to the actual values. For model fitting, an 80% training set and a 20% test set were specified, with the 'initial' parameter set to the length of the training set and the 'horizon' set to the length of the test set. The 'period' argument refers to the spacing between cutoff points and is not required. In this case, Prophet inferred a period of '1' for daily granularity time series data. For instance, during the pre-crisis period for Miam, the data set consisted of 790 data points: hence the 'initial' is 632 days, with 'period' set to 1 by default, and 'horizon' set to 158 days. During iterations, Prophet trained on different subsets of the data. Thus, in the first iteration, it trains on days 1 to 632 and forecasts for days 633 to 634. Similarly, in the second iteration, it trains on days 2 to 633 and forecasts for days 634 to 635 (Taylor and Letham, 2017). In subsequent iterations, it adjusted the training window accordingly. As an alternative approach, predictions were generated using the trained Prophet model on a single test set for which the performance was directly assessed. Then, manually calculated error metrics for the entire test set.

In the forecasting step, the *make_future_dataframe()* function was used to create a *data frame* with a column 'ds' containing the next 366 future timestamps for which forecasts will be generated. The *head()* and *tail()* functions displayed the first and last observations of the period between July 2023 and June 2024. After this, the *predict()* function returned a *data frame* containing the forecasted values, where the following parameters are worth noting: 'ds' refers to the datestamp of the forecasted value; 'yhat' is the forecasted value of metric 'y'; 'yhat_lower' is the lower bound; and 'yhat_upper' refers to the upper bound of forecasts. As a final step, the Prophet forecast with its components (date, day of week, and day of year) were visualised in a plot.

3.1.6. Deployment

The present report was developed with support from Power BI Desktop and R Studio software environment. This work encapsulates the analytical insights arising from the advanced analysis conducted in R, including actionable visualisations, summaries, and key findings that stakeholders can readily access for decision-making. Thus, the BI solution was specifically designed as an internal tool for *TouchPoint Consulting,* intended for application in the organisation's business analytics projects within the Food Retail sector. The aim is to integrate the predictive models derived from R analysis back into the business workflow. This integration will support further predictive research across other retail segments, beyond the Sweet Snacks category studied, broadening awareness about categories in crisis conjectures. As a final point, it is worth mentioning that the methodology adheres to anonymised data practices, ensuring continuous compliance with data protection and privacy regulations.

4. RESULTS ANALYSIS

This section presents a detailed analysis and discussion of the outputs of the time series forecasting process covered in sections *3.1.4.* and *3.1.5.*.

4.1. Time series analysis in R for the Pre-crisis period: January 2018 to February 2020

The total sales of Retailers Miam and Munch behave similarly throughout the Pre-crisis period (Figure 5). Sweet snacks sales showed fluctuations over time, following a clear upward trend. Between January 2018 and June 2018, there was an increase in sales until March 2018 (ranging from 43,750 to 87,500 for Miam, and from 45,000 to 90,000 for Munch), followed by a noticeable decline until the end of the semester (from 37,500 to 75,000, and from 36,250 to 75,000, respectively). During this first subperiod, Miam had its pre-crisis sales peak, with sales reaching 97,938 on 2018/01/27. The following semester showed another change in trend with an increase in values until the end of 2018. The demand for products in this category continued to behave similarly in 2019, with sales rising in the first two months of 2020. Munch reached its maximum value of 118,059 units on 2020/01/18. The pre-crisis period ended with a range of values between 50,000 and 87,500 for the first retailer, and between 60,000 and 120,000 for the second. Easter 2018 and 2019 (2018/04/01 and 2019/04/21) marked the lowest point in sales recorded for both players during the pre-crisis period. The first outliers assume sales figures of

19,201 and 14,093, and the second outliers of 18,448 and 14,110 for Miam and Munch, correspondingly. As a final remark, analysing two consecutive periods of 2018 and 2019 revealed the presence of yearly seasonality in the pre-crisis period.

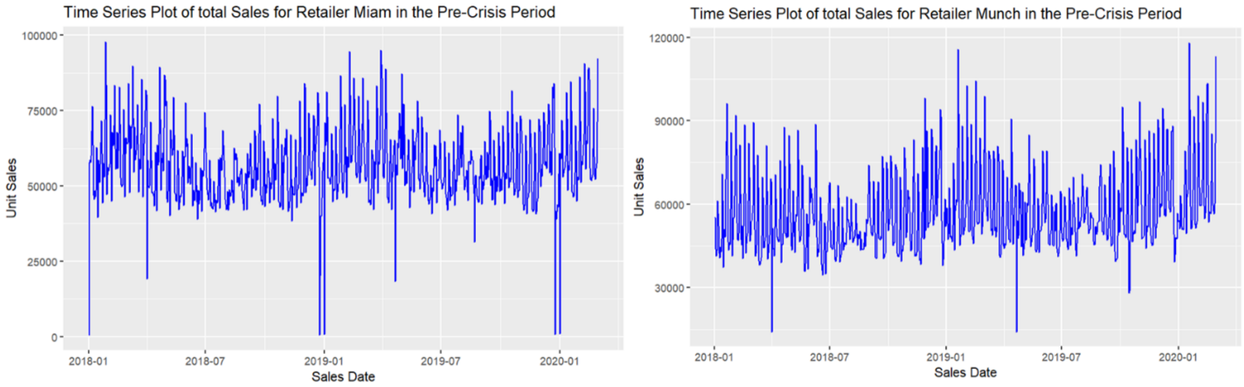


Figure 5. Time Series Plot of total Sales in the Pre-Crisis Period [ARIMA Output]

Tables 4 and 5 present information criterion and accuracy measures for ARIMA models in the pre-crisis period. Table 4 shows the outputs for two ARIMA models, with the first returned by the *auto.arima()* function and the second represents the best model selected with the lowest information criterion. Therefore, ARIMA(2,1,1) was allocated to Miam, and ARIMA(3,1,5) to Munch for the pre-crisis period. In addition, the accuracy measures presented in Table 5 were computed to compare simulated and observed ARIMA values, with Retailer Miam having a lower error margin.

TABLE 4. INFORMATION CRITERION FOR ARIMA MODELS BETWEEN JAN 2018-FEB 2020

	***auto.arima*(1,1,1)**		**ARIMA(2,1,1)**			***auto.arima*(2,1,5)**		**ARIMA(3,1,5)**	
	AIC	**BIC**	**AIC**	**BIC**		**AIC**	**BIC**	**AIC**	**BIC**
Miam	1,023,795	1,023,828	1,023,398	1,023,443	**Munch**	1,132,962	1,133,051	1,132,838	1,132,938

TABLE 5. STATISTICAL COMPARISON BETWEEN ARIMA MODELS FOR PREDICTING TOTAL SALES VARIABLE BETWEEN JAN 2018-FEB 2020

	ARIMA(2,1,1)					**ARIMA(3,1,5)**			
	MSE	**RMSE**	**MAE**	**MAPE**		**MSE**	**RMSE**	**MAE**	**MAPE**
Miam	-0.0640	0.6221	0.4555	Inf	**Munch**	0.5559	0.8707	0.7015	Inf

Similarly, Prophet Model was also perceived as more accurate for Retailer Miam, as evidenced by lower precision errors for the MSE, RMSE and MAE metrics, except for MAPE in Table 6. Table 7 summarises the results of models for the pre-crisis data set, reflecting their performance and accuracy. ARIMA achieved very competitive results when compared to Prophet for both retailers. The low error measures of both ARIMA models produced on the retailer's data set, made them the best predictive models for the pre-crisis period. Nevertheless, it is worth noting the 'Inf' output for MAPE in the ARIMA implies that in some cases, actual values might be zero, leading to division by zero and an infinite metric result. This can result in potentially unrealistic metrics in ARIMA, whereas the Prophet model has higher error values, but it provides finite and more interpretable error metrics.

TABLE 6. STATISTICAL COMPARISON BETWEEN PROPHET MODELS FOR PREDICTING TOTAL SALES VARIABLE BETWEEN JAN 2018-FEB 2020

	Prophet			
	MSE	**RMSE**	**MAE**	**MAPE**
Miam	88,480,474	9,406.406	6,312.42	86.0198
Munch	239,240,104	15,467.39	10,851.28	16.6182

TABLE 7. STATISTICAL COMPARISON BETWEEN ARIMA AND PROPHET MODELS FOR PREDICTING TOTAL SALES VARIABLE BETWEEN JAN 2018-FEB 2020

	ARIMA(2,1,1)/ARIMA(3,1,5)				**Prophet**			
	MSE	**RMSE**	**MAE**	**MAPE**	**MSE**	**RMSE**	**MAE**	**MAPE**
Miam	-0.0193	0.6597	0.4345	Inf	88,480,474	9,406.406	6,312.42	86.0198
Munch	0.0149	0.7819	0.4732	Inf	239,240,104	15,467.39	10,851.28	16.6182

For robustness check, a different train and test split was tested for the pre-crisis period, given that for model performance, the data set used to select the ARIMA models was a subset of the period's data set. The same training set was considered with the first 500,000 observations taking place in January 2018 and a test set with the first 200,000 observations in August 2019 (2019/08/01 – 2019/08/09 for Miam and 2019/08/01 – 2019/08/10 for Munch). Despite this adjustment, similar information criterion were obtained, although with slightly lower error metrics for the models. These results demonstrate that the data subset used for model computation does not impact the ARIMA model selection process, as similar results were provided.

4.2. Time series analysis in R for the Crisis period: March 2020 to December 2021

During the crisis period, Retailer Munch assumed a similar pattern as Retailer Miam with emphasis from July 2020 (Figure 6). Overall, Miam followed an upward trend characterised by the projection of sequential growing cycles where the market tried to counter the declines throughout the entire period. Both retailers suffered a sharp drop in

March 2020, after the high sales figures recorded at the end of the pre-crisis period, which were only maintained in the first half of this month at around 75,000 and 105,000 units. Between the second half of March and June 2020, Miam's sales dropped to 25,000 but progressively increased to 65,500. In contrast, Munch continued to show a downward trend in sales, with a range of 37,500 and 75,000. From the second half of 2020, the retailer's sales balance gradually improved until the end of the crisis period. The sales range for Miam and Munch are presented below, in accordance: from July 2020 to December 2020 (37,500 to 75,000 and 45,000 to 90,000) and from January 2021 to June 2021 (37,500 to 62,500 and 45,000 to 75,000). In the first semester of 2021, Miam signalled its high sales point of 89,627 units on 2021/05/01. The last month of the period revealed markedly higher sales, with figures between 50,000 and 87,500 for Miam, and between 60,000 and 105,000 for Munch. Moreover, there are two common peaks observed in the last half of the crisis period, registered on 2021/10/30 (83,943 units for Miam and 16,045 units for Munch), and in 2021/12/23 (88,745 and 113,550 units). As in the previous two years, it was at Easter 2020 (2020/04/12) that sales reached the worst level recorded during the crisis period. This represents unit sales of 183 for Miam, and 10,703 for Munch.

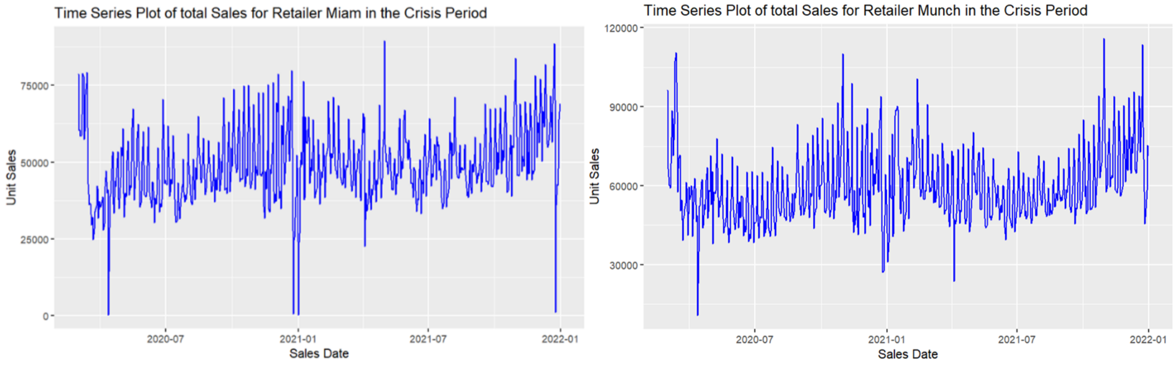


Figure 6. Time Series Plot of total Sales in the Crisis Period [ARIMA Output]

Tables 8 and 9 show information criterion and accuracy measures for ARIMA models in the crisis period. Table 8 reveals ARIMA(4,1,1) was assigned to Miam and ARIMA(2,1,3) to Munch for the crisis period. In addition, the accuracy measures presented in Table 9 indicate that Retailer Munch has a lower error margin.

Table 8. Information Criterion for ARIMA Models between March 2020-Dec 2021

	auto.*arima*(3,1,1)		**ARIMA(4,1,1)**			**auto.*arima*(2,1,2)**		**ARIMA(2,1,3)**	
	AIC	**BIC**	**AIC**	**BIC**		**AIC**	**BIC**	**AIC**	**BIC**
Miam	1,193,570	1,193,626	1,193,528	1,193,595	**Munch**	1,259,780	1,259,835	1,259,617	1,259,684

Table 9. Statistical comparison between ARIMA Models for predicting total Sales variable between March 2020-Dec 2021

	ARIMA(4,1,1)					**ARIMA(2,1,3)**			
	MSE	**RMSE**	**MAE**	**MAPE**		**MSE**	**RMSE**	**MAE**	**MAPE**
Miam	0.0277	0.7981	0.5384	Inf	**Munch**	-0.00035	0.7888	0.5327	Inf

In contrast, Prophet Model was perceived as a more accurate model for Retailer Miam, as evidenced by lower precision errors for the MSE, RMSE and MAE metrics in Table 10. Table 11 summarises the results of models for the crisis data set. Once again, the ARIMA models proved to be the best predictive models of the crisis period.

Table 10. Statistical comparison between Prophet Models for predicting total Sales variable between March 2020-Dec 2021

	Prophet			
	MSE	**RMSE**	**MAE**	**MAPE**
Miam	100,675,107	10,033.7	7,043.55	53.76767
Munch	131,775,185	11,479.34	8,248.461	11.64143

Table 11. Statistical comparison between ARIMA and Prophet Models for predicting total Sales variable between March 2020-Dec 2021

	ARIMA(4,1,1)/ARIMA(2,1,3)				**Prophet**			
	MSE	**RMSE**	**MAE**	**MAPE**	**MSE**	**RMSE**	**MAE**	**MAPE**
Miam	0.0735	0.9881	0.5384	Inf	100,675,107	10,033.7	7,043.55	53.76767
Munch	0.0422	0.8021	0.5355	Inf	131,775,185	11,479.34	8,248.461	11.64143

4.3. Time series analysis in R for the Post-crisis period: January 2022 to June 2023

The post-crisis period was characterised by contrasting trends among the retailers, with Miam showing a decreasing pattern and Munch an ascending tendency (Figure 7). Regarding Miam, the player assumes unit sales values between 50,000 and 87,500 for the first semester of 2022, followed by a range of 37,500 and 81,250 for the subsequent middle of the same year. In 2023, sales were reported between 50,000 and 75,000 units, whereas in June 2023 sales slightly fell between 50,000 and 62,500. As for Munch, the sales volume started at 50,000 and 112,500 in the first half of 2022, and then 62,500 and 125,000 in the second half of 2022. In the last five months of analysis, the banner recorded a unit interval between 62,500 and 125,000, while in June 2023 it showed a lower upper limit between 62,500 and 100,000, although still revealing an increasing pattern in sales. It's worth noting that both retailers experienced a peak in sales on 2022/12/23, close to Christmas Day. This increase in demand may be attributed to the attractiveness of sweet snacks as gift options. On this day, 97,534 observations were registered for Miam, and 126,280 for Munch. Similar to the pre-crisis and crisis periods, the post-crisis period also saw a decline in sales during the Easter holiday. Sales figures were 24,215 and 23,430 for Easter 2022 (2022/04/17), and 25,240 and 24,075 for Easter 2023 (2023/04/09). Both retailers consistently experience their lowest sales during the Easter holidays across all three periods, except for 2021. This could be due to concerns about a national public holiday potentially leading to limitations or gaps in sales recording.

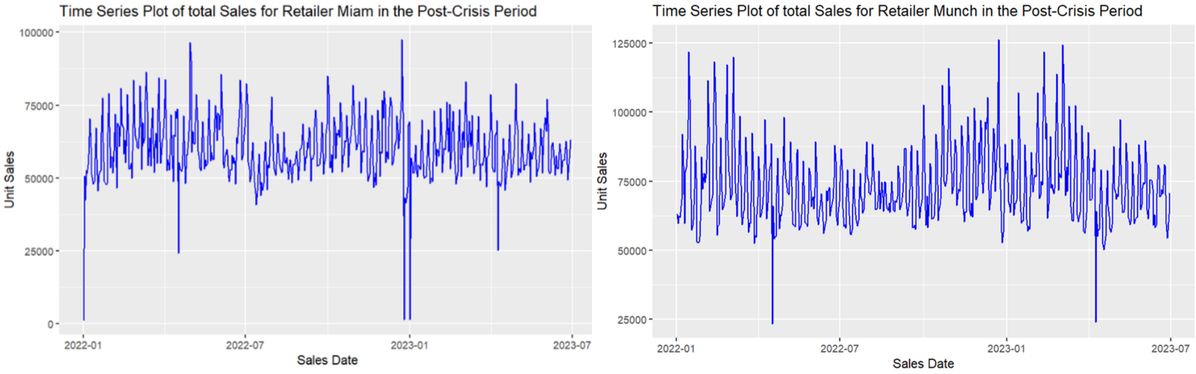


Figure 7. Time Series Plot of total Sales in the Post-Crisis Period [ARIMA Output]

Tables 12 and 13 show information criterion and accuracy measures for ARIMA models in the post-crisis period. Table 12 reveals that ARIMA(3,1,1) was assigned to Miam and ARIMA(1,1,2) to Munch for the post-crisis period. In addition, the accuracy measures presented in Table 13 indicate that Retailer Munch has a lower error margin.

TABLE 12. INFORMATION CRITERION FOR ARIMA MODELS BETWEEN JAN 2022-JUN 2023

	auto.*arima*(1,1,3)		**ARIMA(3,1,1)**			***auto.arima*(2,1,2)**		**ARIMA(1,1,2)**	
	AIC	**BIC**	**AIC**	**BIC**		**AIC**	**BIC**	**AIC**	**BIC**
Miam	915,526.9	915,582.5	915,475.4	915,531	**Munch**	800,904.1	800,959.7	800,618.2	800,662.7

TABLE 13. STATISTICAL COMPARISON BETWEEN ARIMA MODELS FOR PREDICTING TOTAL SALES VARIABLE BETWEEN JAN 2022-JUN 2023

	ARIMA(3,1,1)					ARIMA(1,1,2)			
	MSE	RMSE	MAE	MAPE		MSE	RMSE	MAE	MAPE
Miam	-0.1355	0.6088	0.4639	Inf	**Munch**	-0.0155	0.5114	0.3605	Inf

In contrast, Prophet Model was perceived as a more accurate model for Retailer Miam, as evidenced by lower precision errors for the MSE and RMSE metrics, except for MAE and MAPE in Table 14. In this case, the discrepancy in the performance between Prophet models was narrower. Table 15 summarises the results of models for the post-crisis data set, concluding that ARIMA models are the most competitive choice in predicting the post-crisis period.

TABLE 14. STATISTICAL COMPARISON BETWEEN PROPHET MODELS FOR PREDICTING TOTAL SALES VARIABLE BETWEEN JAN 2022-JUN 2023

	Prophet			
	MSE	RMSE	MAE	MAPE
Miam	52,109,041	7,218.659	5,825.868	10.88
Munch	65,573,201	8,097.728	5,563.366	9.618567

TABLE 15. STATISTICAL COMPARISON BETWEEN ARIMA AND PROPHET MODELS FOR PREDICTING TOTAL SALES VARIABLE BETWEEN JAN 2022-JUN 2023

	ARIMA(3,1,1)/ARIMA(1,1,2)				Prophet			
	MSE	RMSE	MAE	MAPE	MSE	RMSE	MAE	MAPE
Miam	0.0161	0.5975	0.4182	Inf	52,109,041	7,218.659	5,825.868	10.88
Munch	0.0413	0.5233	0.3753	Inf	65,573,201	8,097.728	5,563.366	9.618567

The previous accuracy metrics results suggest a solid foundation for future projections considering overall model quality. In conducting the forecasting process, note that only measure ‘total_Sales’ in the data object was considered to enable a proper comparison between ARIMA and Prophet models. The 366-period projections for ARIMA and Prophet models were performed, covering the period from July 2023 to June 2024, depicted in Figures 8 and 9, respectively.

The upper row of Figure 8 depicts the time plot outputted by the ARIMA(3,1,1) forecasting model for Miam, showing the forecasted values along with their 95% prediction intervals. The next 366-period projections indicate a slowing of the downward post-crisis trend, towards steady values. Accordingly, it is expected to reach values around 62,050 units, subjugated in the 95% confidence interval (minimum expected values of 19,942 and maximum expected to reach 104,157). In the bottom row, the graph reflects projections for Munch using the ARIMA(1,1,2) forecasting model. After a period of continuous growth in the post-crisis period, the 366-ahead forecasts point to trend neutrality, setting sales around 80,462 units, with a 95% lower limit of 45,974 and an upper limit of 114,950.

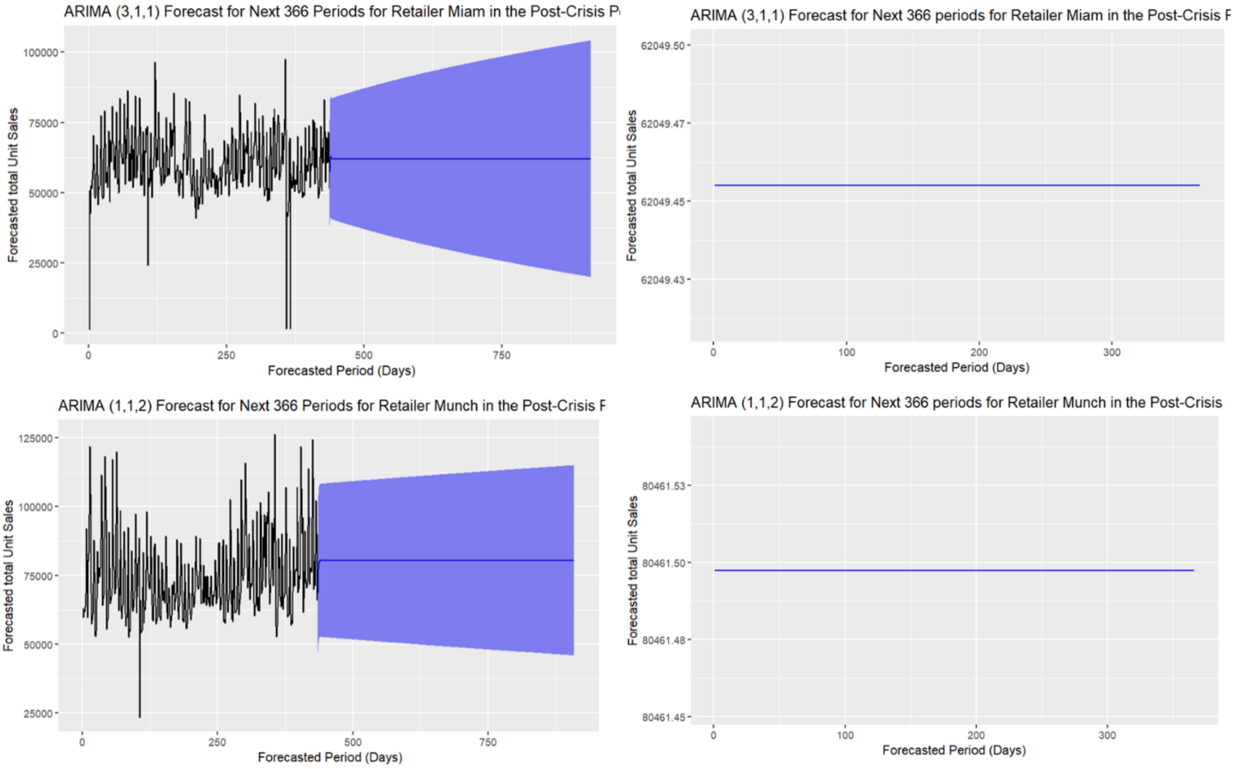


Figure 8. ARIMA Forecast of total Sales between July 2023-June 2024 [ARIMA Output]

Figure 9 shows the 366-period Prophet forecast for Miam, following the downward trend in sales experienced during the recent post-crisis period. The forecasted values are expected to be around 59,778, the 95% lower limit being 49,719 and the upper limit 69,832. In seasonal terms, the sales display a first semester with a minimally uniform monthly behaviour and a second semester with a volatile behaviour (Appendix 14). The biggest drop is recorded at the beginning of the first month of the year, followed by declines in the middle of July and August, while the peaks with the greatest impact on the category’s sales are recorded at the beginning of November and December. Furthermore, Figure 9 suggests continued growth projections for Munch in the recent post-crisis period, evolving to values around 76,968 with a 95% confidence interval between 66,098 and 87,832. Regarding the seasonal component, sales show a changing pattern of declines and recoveries throughout the months (Appendix 14). The first sharp rise of the year is registered in the middle of February, followed by a series of months with pronounced declines in sales, mainly in the middle of July and August, and at the beginning of April and October. However, these falls were cancelled out by the increases observed in November and December.

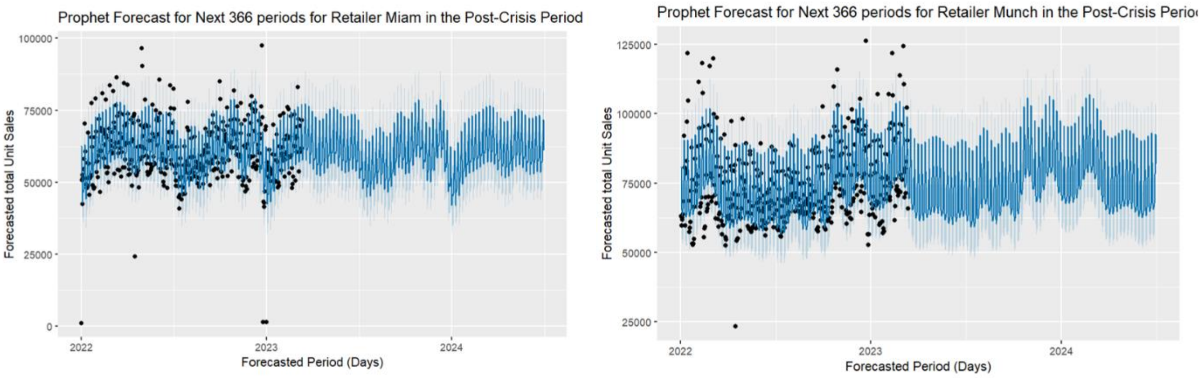


Figure 9. Prophet Forecast of total Sales between July 2023-June 2024 [Prophet Output]

5. CONCLUSIONS, LIMITATIONS AND FUTURE WORK

5.1. Conclusions

This project examined the sales evolution of the Sweet Snacks category for two market-leading Portuguese food retailers over a five-and-a-half-year period, including the COVID-19 crisis conjecture, and inferred the best forecasting model for the three subperiods under analysis.

During the pre-crisis period, both retailers experienced a progressive increase in sales until February 2020. At the very beginning of the crisis period (in March 2020, the month in which the national lockdown was declared), there was a sudden drop in sales. However, the market quickly recovered slightly. The first three semesters of the crisis period recorded lower sales levels compared to the pre-crisis period, with Miam experiencing a more pronounced decline in values. Despite this, both retailers managed to recover in the latter half of the crisis, with sales volumes reaching similar levels to those before the crisis by December 2021. This demonstrates a quick responsiveness and ability to adapt to the challenges posed by periods of crisis. In the post-crisis period, Munch's sales balance continued to improve and even exceeded pre-crisis numbers, indicating an ongoing recovery. On the other hand, Miam witnessed a significant decline, suggesting a setback in its recovery process.

Comparing ARIMA models returned for both retailers across different periods, provides insights into their sales strategy dynamics and recovery trajectories during various market conditions. In the pre-crisis period, Miam (ARIMA(2,1,1)) exhibited a comparatively simpler model compared to Munch (ARIMA(3,1,5)). Miam's model indicated fewer lagged observations influencing sales, whereas Munch's sales dynamics appeared more intricate with higher dependency on lagged observations and error terms, suggesting a potentially more volatile or less predictable sales pattern. During the crisis period, Miam (ARIMA(4,1,1)) adopted a more complex model compared to its pre-crisis approach, reflecting adjustments to the economic uncertainty, shifting consumer behaviours, supply chain disruptions, and changes in market demand. Meanwhile, Munch (ARIMA(2,1,3)) also adjusted its model structure, albeit maintaining a slightly simpler configuration. The adaptation to different best-fit models during the crisis underscores the limitations of static models in capturing dynamic changes during periods of external shocks or significant global events. In the post-crisis phase, Miam (ARIMA (3,1,1)) transitioned back to a simpler model after the turbulent period, suggesting potential stabilisation or normalisation in sales behaviour. Similarly, Munch (ARIMA(1,1,2)) also indicated a move towards a simpler model post-crisis. This evolution in sales dynamics reflects the establishment of new patterns and trends in response to the changing market landscape.

It should be remarked both retailers experienced a recovery in sales at the end of the crisis period. However, while Munch sustained growth, Miam faced a setback in post-crisis period. This divergence in performance can be attributed to the different sales strategies employed by each retailer. Miam opted for a more complex ARIMA model during the crisis period and transitioned to a simpler model in the post-crisis period, changing only the AR component between periods. Conversely, Munch adopted a strategy of continuity since the pre-crisis period, reflected in progressively less complex models. This suggests that Miam's sales strategy was not aligned with the new post-crisis consumption patterns. In this way, both retailers demonstrated adaptability by adjusting their forecasting models, mirroring their efforts to optimise sales strategies and operational approaches during different economic phases.

Through model fitting and performance evaluation over periods, it was possible to identify the time series model that best adjusts to the data and minimises precision errors. In this respect, ARIMA outperformed Prophet in the three horizons. Some factors have been identified that could contribute to the significantly higher accuracy metrics of the

Prophet model compared to the ARIMA. Firstly, there might be insufficient historical data in adjacent periods for the model to effectively learn and generalise underlying trends and seasonal patterns. Additionally, the data sets exhibit high volatility with abrupt changes, making the model sensitive to periods of unique behaviour or unexpected fluctuations. Moreover, Prophet assumes "smooth seasonalities", while the data set under study may not demonstrate a regular and continuous seasonal pattern. Besides this, two overall findings emerged regarding the models' performance. First, ARIMA revealed greater accuracy in capturing Munch's sales patterns, while Prophet performed better for Miam throughout the analysis period. Second, the post-crisis period returned the highest predictive ability, as evidenced by the lowest error metrics obtained by both models.

In the future forecast period between July 2023 and July 2024, the models indicate divergent patterns. While ARIMA points to a stabilisation of sales patterns, Prophet model reflects the sales trend of the post-crisis period. The difference in forecasted values between the two models is not significant on average, differing by approximately 2,000 units for Miam and 3,000 units for Munch. The ARIMA model produces higher absolute sales predictions than Prophet. The wide prediction intervals for Miam imply high uncertainty in the ARIMA(3,1,1) model regarding predicted future values. The same occurs for Munch, but the forecast intervals are narrower, indicating greater confidence in the ARIMA(1,1,2). Unlike ARIMA, the Prophet model provided considerably narrower confidence intervals for both retailers, enhancing the reliability of future forecasts and reaffirming its superiority in forecasting Sweet Snacks sales. In summary, while ARIMA showed superior overall performance based on statistical comparison, the Prophet model achieved more accurate predictions of the observed sales pattern.

5.2. Limitations and Future work

The first limitation encountered in the project was a data availability issue during the data collection phase. The analysis was conducted in unit sales, as only Miam provided unit sales data, while Munch provided both unit and value sales data. For comparison purposes, quantity was chosen as the preferred unit of measurement, given that retail prices fluctuate over time and may be affected by promotional periods. Other main limitation concerns software performance constraints when handling large data sets. Firstly, when trying to integrate tools to visualise R outputs into a Power BI environment, and subsequently into R Studio, limited training and test sets were used during the model selection process. Nevertheless, experiments conducted still indicate good overall system performance, as proven by the robustness check analysis. Hence, further research should be performed using different BI tools, such as SQL, Tableau or Python, to assess the performance of alternative BI solutions to overcome computational restrictions. Regarding the forecasting models, it is important to acknowledge that external factors impacting sales in the Sweet Snacks category were not incorporated in the forecasting period and should not be disregarded.

Finally, the study of sales data from online channel would provide valuable insights to complement our findings, given the substantial increase in e-commerce during the crisis period, and this data was not available in this work. Furthermore, future work could investigate the Sweet Snacks category of private label brands and compare these results with those obtained for a manufacturer's brand in this project, to follow the growing trend towards private label loyalty witnessed since the pandemic.

REFERENCES

Ahn, H., Sun, K. and Kim, K. (2021) 'Comparison of Missing Data Imputation Methods in Time Series Forecasting', *Computers, Materials and Continua*, 70, pp. 767–779. Available at: https://doi.org/10.32604/cmc.2022.019369.

Allen, S. and Terry, E. (2005) *Beginning Relational Data Modeling*. Apress.

Armstrong, J.S. (2001) *Principles of Forecasting: A Handbook for Researchers and Practitioners*. Springer Science & Business Media.

Asesh, A. (2022) 'Normalization and Bias in Time Series Data', in C. Biele et al. (eds) *Digital Interaction and Machine Intelligence*. Cham: Springer International Publishing (Lecture Notes in Networks and Systems), pp. 88–97. Available at: https://doi.org/10.1007/978-3-031-11432-8_8.

Awan, T.M. and Aslam, F. (2020) 'Prediction of daily COVID-19 cases in European countries using automatic ARIMA model', *Journal of Public Health Research*, 9(3), p. 1765. Available at: https://doi.org/10.4081/jphr.2020.1765.

Becker, L.T. and Gould, E.M. (2019) 'Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data', *Serials Review*, 45(3), pp. 184–188. Available at: https://doi.org/10.1080/00987913.2019.1644891.

Bozdogan, H. (1987) 'Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions', *Psychometrika*, 52(3), pp. 345–370. Available at: https://doi.org/10.1007/BF02294361.

Campbell, M. (2019) 'RStudio Projects', in M. Campbell (ed.) *Learn RStudio IDE: Quick, Effective, and Productive Data Science*. Berkeley, CA: Apress, pp. 39–48. Available at: https://doi.org/10.1007/978-1-4842-4511-8_4.

Cereja, C.G. (2018) 'Effects of corporate social responsibility in Portuguese food retail own brands' trust: Millennials vs. Generation X perceptions'.

Chapman, P. (2000) 'CRISP-DM 1.0: Step-by-step data mining guide', in. Available at: https://www.semanticscholar.org/paper/CRISP-DM-1.0%3A-Step-by-step-data-mining-guide-Chapman/54bad20bbc7938991bf34f86dde0babfbd2d5a72 (Accessed: 5 October 2023).

Chatfield, C. (2003) *The Analysis of Time Series: An Introduction, Sixth Edition*. 6th edn. New York: Chapman and Hall/CRC. Available at: https://doi.org/10.4324/9780203491683.

'Chocolate Confectionery in Portugal' (2023) *Marketline*. Available at: https://store.marketline.com/report/chocolate-confectionery-in-portugal/ (Accessed: 19 November 2023).

Consumer sentiment in Portugal during the coronavirus crisis | McKinsey (2020).

Available at: https://www.mckinsey.com/capabilities/growth-marketing-and-sales/our-insights/survey-portuguese-consumer-sentiment-during-the-coronavirus-crisis (Accessed: 23 November 2023).

Costa, C. and Aparício, J.T. (2020) *POST-DS: A Methodology to Boost Data Science*, p. 6. Available at: https://doi.org/10.23919/CISTI49556.2020.9140932.

Davey, A.M. and Flores, B.E. (1993) 'Identification of seasonality in time series: A note', *Mathematical and Computer Modelling*, 18(6), pp. 73–81. Available at: https://doi.org/10.1016/0895-7177(93)90126-J.

Dégerine, S. and Lambert-Lacroix, S. (2003) 'Characterization of the partial autocorrelation function of nonstationary time series', *Journal of Multivariate Analysis*, 87(1), pp. 46–59. Available at: https://doi.org/10.1016/S0047-259X(03)00025-3.

deLeeuw, J. (1992) 'Introduction to Akaike (1973) Information Theory and an Extension of the Maximum Likelihood Principle', in S. Kotz and N.L. Johnson (eds). New York, NY: Springer New York (Springer Series in Statistics), pp. 599–609. Available at: https://doi.org/10.1007/978-1-4612-0919-5_37.

Dimitris Skalkos and Zoi C. Kalyva (2023) *Exploring the Impact of COVID-19 Pandemic on Food Choice Motives: A Systematic Review*. Available at: https://www.mdpi.com/2071-1050/15/2/1606 (Accessed: 19 November 2023).

Ding, J., Tarokh, V. and Yang, Y. (2018) 'Model Selection Techniques -- An Overview', *IEEE Signal Processing Magazine*, 35(6), pp. 16–34. Available at: https://doi.org/10.1109/MSP.2018.2867638.

Duarte, D., Walshaw, C. and Ramesh, N. (2021) 'A Comparison of Time-Series Predictions for Healthcare Emergency Department Indicators and the Impact of COVID-19', *Applied Sciences*, 11(8), p. 3561. Available at: https://doi.org/10.3390/app11083561.

Dziak, J.J. *et al.* (2019) 'Sensitivity and specificity of information criteria', *Briefings in Bioinformatics*, 21(2), pp. 553–565. Available at: https://doi.org/10.1093/bib/bbz016.

Faghih, N. and Forouharfar, A. (eds) (2022) *Socioeconomic Dynamics of the COVID-19 Crisis: Global, Regional, and Local Perspectives*. Cham: Springer International Publishing (Contributions to Economics). Available at: https://doi.org/10.1007/978-3-030-89996-7.

Faloutsos, C., Gasthaus, J., *et al.* (2019) 'Classical and Contemporary Approaches to Big Time Series Forecasting', *Proceedings of the 2019 International Conference on Management of Data*, pp. 2042–2047. Available at: https://doi.org/10.1145/3299869.3314033.

Faloutsos, C., Flunkert, V., *et al.* (2019) 'Forecasting Big Time Series: Theory and Practice', in. *KDD '19: Proceedings of the 25th ACM SIGKDD International*

Conference on Knowledge Discovery & Data Mining, pp. 3209–3210. Available at: https://doi.org/10.1145/3292500.3332289.

Fanoodi, B., Malmir, B. and Jahantigh, F.F. (2019) 'Reducing demand uncertainty in the platelet supply chain through artificial neural networks and ARIMA models', *Computers in Biology and Medicine*, 113, p. 103415. Available at: https://doi.org/10.1016/j.compbiomed.2019.103415.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996a) 'From Data Mining to Knowledge Discovery in Databases', *AI Magazine*, 17(3), pp. 37–37. Available at: https://doi.org/10.1609/aimag.v17i3.1230.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996b) 'Knowledge Discovery and Data Mining: Towards a Unifying Framework'.

FENG, C. *et al.* (2014) 'Log-transformation and its implications for data analysis', *Shanghai Archives of Psychiatry*, 26(2), pp. 105–109. Available at: https://doi.org/10.3969/j.issn.1002-0829.2014.02.009.

Ferrari, A. and Russo, M. (2016) *Introducing Microsoft Power BI*. Microsoft Press.

Fildes, R. and Makridakis, S. (1995) 'The Impact of Empirical Accuracy Studies on Time Series Analysis and Forecasting', *International Statistical Review / Revue Internationale de Statistique*, 63(3), pp. 289–308. Available at: https://doi.org/10.2307/1403481.

George E. P. Box *et al.* (2015) *Time Series Analysis: Forecasting and Control, 5th Edition*. Wiley-Blackwell. Available at: https://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,shib&db=a9h&AN=117000109&lang=pt-pt&site=eds-live&scope=site (Accessed: 11 October 2023).

Gibson, J. *et al.* (2007) 'Data mining for social scientists', in.

Glynn, J., Perera, N. and Verma, R. (2007) 'Unit root tests and structural breaks: a survey with applications', *Faculty of Commerce - Papers (Archive)* [Preprint]. Available at: https://ro.uow.edu.au/commpapers/455.

Gomes, S., Lopes, J.M. and Oliveira, J. (2023) 'Online Food Shopping: Determinants and Profile of Portuguese Buyers in the Pandemic Context', 33(87), pp. 73–91. Available at: https://doi.org/10.15446/innovar.v33n87.105507.

Günther, L.C. *et al.* (2019) 'Data quality assessment for improved decision-making: a methodology for small and medium-sized enterprises', *Procedia Manufacturing*, 29, pp. 583–591. Available at: https://doi.org/10.1016/j.promfg.2019.02.114.

Han, J. and Kamber, M. (2013) *Data mining: concepts and techniques*. 2. ed. Elsevier, Morgan Kaufmann Publishers (The Morgan Kaufmann series in data management systems).

Hand, D., Mannila, H. and Smyth, P. (2001) *Principles of Data Mining*.

Harris, R.I.D. (1992) 'Testing for unit roots using the augmented Dickey-Fuller test: Some issues relating to the size, power and the lag structure of the test', *Economics Letters*, 38(4), pp. 381–386. Available at: https://doi.org/10.1016/0165-1765(92)90022-Q.

Harvey, A.C. and Shephard, N. (1993) '10 Structural time series models', in *Handbook of Statistics*. Elsevier (Econometrics), pp. 261–302. Available at: https://doi.org/10.1016/S0169-7161(05)80045-8.

Health and Wellness in Portugal (2022). Available at: https://www.marketresearch.com/Euromonitor-International-v746/Health-Wellness-Portugal-30817017/ (Accessed: 23 November 2023).

Heinrich, B. *et al.* (2017) 'Requirements for Data Quality Metrics', *Journal of Data and Information Quality*, 9(2), pp. 1–32. Available at: https://doi.org/10.1145/3148238.

Herranz, E. (2017) 'Unit root tests', *WIREs Computational Statistics*, 9(3), p. e1396. Available at: https://doi.org/10.1002/wics.1396.

Horton, N.J. and Kleinman, K. (2015) *Using R and RStudio for Data Management, Statistical Analysis, and Graphics*. CRC Press.

Hurvich, C.M. and Tsai, C.-L. (1989) 'Regression and time series model selection in small samples', *Biometrika*, 76(2), pp. 297–307. Available at: https://doi.org/10.1093/biomet/76.2.297.

Hyndman, R.J. and Khandakar, Y. (2008) 'Automatic Time Series Forecasting: The forecast Package for R', *Journal of Statistical Software*, 27, pp. 1–22. Available at: https://doi.org/10.18637/jss.v027.i03.

Hyndman, R.J. and Koehler, A.B. (2006) 'Another look at measures of forecast accuracy', *International Journal of Forecasting*, 22(4), pp. 679–688. Available at: https://doi.org/10.1016/j.ijforecast.2006.03.001.

INE (2021) *Instituto Nacional de Estatística - Estatísticas do Comércio : 2020*. Available at: https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_publicacoes&PUBLICACOESpub_boui=280814598&PUBLICACOESmodo=2 (Accessed: 20 November 2023).

Joe Caserta and Ralph Kimball (2004) *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*. Wiley India Pvt. Limited.

John E. Hanke and Dean Wichern (2014) *Business Forecasting: Pearson New International Edition*. Available at: https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=1418286

&site=eds-live (Accessed: 11 October 2023).

Kamani, G., Parmar, R. and Ghodasara, Y. (2022) ‘Data Normalization in Data Mining using Graphical User Interface: A Pre-Processing Stage’.

Keith Ord, Robert Fildes, and Nikolaos Kourentzes (2017) *Principles of Business Forecasting, 2nd ed.* Available at: https://wessexlearning.com/products/principles-of-business-forecasting-2nd-ed (Accessed: 19 October 2023).

Kennedy, O.A. *et al.* (2011) ‘Fuzzy prophet: parameter exploration in uncertain enterprise scenarios’, in *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*. New York, NY, USA: Association for Computing Machinery (SIGMOD ’11), pp. 1303–1306. Available at: https://doi.org/10.1145/1989323.1989482.

Kimball, R. and Ross, M. (2013) *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*. John Wiley & Sons.

Kline, D. (2004) ‘Methods for Multi-Step Time Series Forecasting with Neural Networks’, in *Neural Networks in Business Forecasting*, pp. 226–250. Available at: https://doi.org/10.4018/978-1-59140-176-6.ch012.

Kritzman, M. (1994) ‘What Practitioners Need to Know…About Serial Dependence’, *Financial Analysts Journal*, 50(2), pp. 19–22. Available at: https://doi.org/10.2469/faj.v50.n2.19.

Kuha, J. (2004) ‘AIC and BIC: Comparisons of Assumptions and Performance’, *Sociological Methods & Research*, 33(2), pp. 188–229. Available at: https://doi.org/10.1177/0049124103262065.

Kurgan, L. and Musilek, P. (2006) ‘A survey of Knowledge Discovery and Data Mining process models’, *Knowledge Eng. Review*, 21, pp. 1–24. Available at: https://doi.org/10.1017/S0269888906000737.

Kwiatkowski, D. *et al.* (1992) ‘Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?’, *Journal of Econometrics*, 54(1), pp. 159–178. Available at: https://doi.org/10.1016/0304-4076(92)90104-Y.

Lever, J., Krzywinski, M. and Altman, N. (2016) ‘Model selection and overfitting’, *Nature Methods*, 13(9), pp. 703–704. Available at: https://doi.org/10.1038/nmeth.3968.

Ljung, G.M. and Box, G.E.P. (1978) ‘On a Measure of Lack of Fit in Time Series Models’, *Biometrika*, 65(2), pp. 297–303. Available at: https://doi.org/10.2307/2335207.

Lopes da Costa, R. *et al.* (2023) ‘Hábitos Alimentares Saudáveis para a Indústria Alimentar em Portugal’, *RPER*, (66), pp. 133–152. Available at:

https://doi.org/10.59072/rper.vi66.80.

Lopez, J.H. (1997) 'The power of the ADF test', *Economics Letters*, 57(1), pp. 5–10. Available at: https://doi.org/10.1016/S0165-1765(97)81872-1.

Lütkepohl, H. and Xu, F. (2012) 'The role of the log transformation in forecasting economic variables', *Empirical Economics*, 42(3), pp. 619–638. Available at: https://doi.org/10.1007/s00181-010-0440-1.

Makridakis, S. *et al.* (2021) 'The Future of Forecasting Competitions: Design Attributes and Principles', *INFORMS Journal on Data Science*, 1. Available at: https://doi.org/10.1287/ijds.2021.0003.

Makridakis, S., Spiliotis, E. and Assimakopoulos, V. (2018) 'Statistical and Machine Learning forecasting methods: Concerns and ways forward', *PLOS ONE*, 13(3), p. e0194889. Available at: https://doi.org/10.1371/journal.pone.0194889.

Mariscal, G., Marbán, O. and Fernández, C. (2010) 'A survey of data mining and knowledge discovery process models and methodologies', *Knowledge Eng. Review*, 25, pp. 137–166. Available at: https://doi.org/10.1017/S0269888910000032.

Medar, R., Rajpurohit, V.S. and Rashmi, B. (2017) 'Impact of Training and Testing Data Splits on Accuracy of Time Series Forecasting in Machine Learning', in *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA). 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, pp. 1–6. Available at: https://doi.org/10.1109/ICCUBEA.2017.8463779.

Medel, C. and Salgado, S. (2012) 'Does BIC Estimate and Forecast Better than AIC?', *Revista de análisis económico*, 28. Available at: https://doi.org/10.4067/S0718-88702013000100003.

Montgomery, D., Jennings, C. and Kulahci, M. (2008) *Introduction to Time Series Analysis and Forecasting*.

Moody, D. and Kortink, M. (2003) 'From ER Models to Dimensional Models: Bridging the Gap between OLTP and OLAP Design, Part I', *Journal of Business Intelligence*, 8.

Moody, D.L. (2000) 'From Enterprise Models to Dimensional Models: A Methodology for Data Warehouse and Data Mart Design'.

Moritz, S. and Bartz-Beielstein, T. (2017) 'imputeTS: Time Series Missing Value Imputation in R', *The R Journal*, 9(1), p. 207. Available at: https://doi.org/10.32614/RJ-2017-009.

Nakamura, T. *et al.* (2006) 'A Comparative Study of Information Criteria for Model Selection.', *I. J. Bifurcation and Chaos*, 16, pp. 2153–2175. Available at: https://doi.org/10.1142/S0218127406015982.

Neath, A.A. and Cavanaugh, J.E. (2012) ‘The Bayesian information criterion: background, derivation, and applications’, *WIREs Computational Statistics*, 4(2), pp. 199–203. Available at: https://doi.org/10.1002/wics.199.

Petropoulos, F. *et al.* (2022) ‘Forecasting: theory and practice’, *International Journal of Forecasting*, 38(3), pp. 705–871. Available at: https://doi.org/10.1016/j.ijforecast.2021.11.001.

Pinto, J. *et al.* (2022) ‘Managerial Practices and (Post) Pandemic Consumption of Private Labels: Online and Offline Retail Perspective in a Portuguese Context’, *Sustainability*, 14. Available at: https://doi.org/10.3390/su141710813.

Portugal: food industry revenue 2022 (2022) *Statista*. Available at: https://www.statista.com/statistics/1399454/portugal-food-industry-turnover/ (Accessed: 24 May 2024).

Portugal: food retailers market share 2022, by category (2022) *Statista*. Available at: https://www.statista.com/statistics/1396480/portugal-food-retailers-market-share-by-category/ (Accessed: 20 November 2023).

Portugal: food retailers market share 2023 (2023) *Statista*. Available at: https://www.statista.com/statistics/1396506/portugal-food-retailers-market-share-by-brand/ (Accessed: 24 May 2024).

Portugal: The Portuguese Food Retail Sector | USDA Foreign Agricultural Service (2021). Available at: https://fas.usda.gov/data/portugal-portuguese-food-retail-sector (Accessed: 7 December 2023).

Powell, B. (2017) *Microsoft Power BI Cookbook: Creating Business Intelligence Solutions of Analytical Data Models, Reports, and Dashboards*. Packt Publishing Ltd.

Powell, B. (2018) *Mastering Microsoft Power BI: Expert techniques for effective data analytics and business intelligence*. Packt Publishing Ltd.

Quenouille, M.H. (1949) ‘Approximate Tests of Correlation in Time-Series’, *Journal of the Royal Statistical Society: Series B (Methodological)*, 11(1), pp. 68–84. Available at: https://doi.org/10.1111/j.2517-6161.1949.tb00023.x.

Ralph Kimball *et al.* (2008) *The Data Warehouse Lifecycle Toolkit*. Available at: https://search.ebscohost.com/login.aspx?direct=true&db=edsebk&AN=413352&site=eds-live (Accessed: 10 October 2023).

Riani, M., Atkinson, A.C. and Corbellini, A. (2023) ‘Automatic robust Box–Cox and extended Yeo–Johnson transformations in regression’, *Statistical Methods & Applications*, 32(1), pp. 75–102. Available at: https://doi.org/10.1007/s10260-022-00640-7.

Robert B. Cleveland, William S. Cleveland, and Irma Terpenning (1990) ‘STL: A Seasonal-Trend Decomposition Procedure Based on Loess’, in. *Journal of*

Official Statistics, pp. 3–33.

Roggeveen, A.L. and Sethuraman, R. (2020) 'How the COVID-19 Pandemic May Change the World of Retailing', *Journal of Retailing*, 96(2), pp. 169–171. Available at: https://doi.org/10.1016/j.jretai.2020.04.002.

Santos and Ramos (2006) *Business intelligence: tecnologias da informação na gestão de conhecimento*. FCA. Available at: https://books.google.pt/books?id=WuqTPgAACAAJ.

Schröer, C., Kruse, F. and Gómez, J.M. (2021) 'A Systematic Literature Review on Applying CRISP-DM Process Model', *Procedia Computer Science*, 181, pp. 526–534. Available at: https://doi.org/10.1016/j.procs.2021.01.199.

Shumway, R.H. and Stoffer, D.S. (2011) *Time Series Analysis and Its Applications*. New York, NY: Springer New York (Springer Texts in Statistics). Available at: https://doi.org/10.1007/978-1-4419-7865-3.

Siguenza-Guzman, L. *et al.* (2015) 'Literature Review of Data Mining Applications in Academic Libraries', *The Journal of Academic Librarianship*, 41, pp. 499–510. Available at: https://doi.org/10.1016/j.acalib.2015.06.007.

Sofia Gomes and João M. Lopes (2022) 'Evolution of the Online Grocery Shopping Experience during the COVID-19 Pandemic: Empiric Study from Portugal', *Journal of Theoretical and Applied Electronic Commerce Research*, 17(47), pp. 909–923. Available at: https://doi.org/10.3390/jtaer17030047.

Stock, J.H. and Watson, M.W. (1999) 'Chapter 1 Business cycle fluctuations in us macroeconomic time series', in *Handbook of Macroeconomics*. Elsevier, pp. 3–64. Available at: https://doi.org/10.1016/S1574-0048(99)01004-6.

Sweet Biscuits, Snack Bars and Fruit Snacks in Portugal (2023) *Euromonitor*. Available at: https://www.euromonitor.com/sweet-biscuits-snack-bars-and-fruit-snacks-in-portugal/report (Accessed: 23 November 2023).

Taylor, S.J. and Letham, B. (2017) 'Forecasting at scale'. Available at: https://doi.org/10.7287/peerj.preprints.3190v2.

Tayman, J., Smith, S.K. and Lin, J. (2007) 'Precision, bias, and uncertainty for state population forecasts: an exploratory analysis of time series models', *Population Research and Policy Review,* 26(3), pp. 347–369. Available at: https://doi.org/10.1007/s11113-007-9034-9.

Verzani, J. (2011) *Getting Started with RStudio*. O'Reilly Media, Inc.

Wang, X. *et al.* (2023) 'Multilevel Residual Prophet Network Time Series Model for Prediction of Irregularities on High-Speed Railway Track', *Journal of Transportation Engineering, Part A: Systems*, 149(4), p. 04023012. Available at: https://doi.org/10.1061/JTEPBS.TEENG-7437.

Wang, X., Smith, K. and Hyndman, R. (2006) ‘Characteristic-Based Clustering for Time Series Data’, *Data Mining and Knowledge Discovery*, 13(3), pp. 335–364. Available at: https://doi.org/10.1007/s10618-005-0039-x.

Wirth, R. and Hipp, J. (2000) ‘CRISP-DM: Towards a Standard Process Model for Data Mining’, *Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining* [Preprint].

World Market for Packaged Food (2021) *Euromonitor*. Available at: https://www.euromonitor.com/world-market-for-packaged-food/report (Accessed: 23 November 2023).

Yeo, I.-K. and Johnson, R.A. (2000) ‘A New Family of Power Transformations to Improve Normality or Symmetry’, *Biometrika*, 87(4), pp. 954–959.

Zhang, J., Yang, Y. and Ding, J. (2023) ‘Information criteria for model selection’, *WIREs Computational Statistics*, 15(5), p. e1607. Available at: https://doi.org/10.1002/wics.1607.

Zwanka, D.R.J. (2022) ‘COVID-19 Impacts to the Supermarket Industry: A Framework of Major Long-Term and Short-Term Changes in Consumer Behavior’.

APPENDICES

Appendix 1. Classification of Food Retail Stores

Table A- 1. Characteristics of Different Food Retail Stores Formats (adapted from *Portugal: The Portuguese Food Retail Sector | USDA Foreign Agricultural Service*, 2021; own elaboration)

Food Retail Stores	**Store Size**	**Main Characteristics**
Hypermarket	Biggest store format. > 2,500 m²	Offers a wide variety of food and non-food products.
Supermarket	Medium store format. 400 m² – 2,499 m²	Mainly provides food products and a limited range of non-food products.
Discount	Do not have a defined statutory size.	Sells a more restricted number of products and brands at low prices.
Convenience	Small store format. < 400 m²	Provides emergency products, located in neighbourhoods closest to customers.

Appendix 2. Best-selling Categories in Food Retail in 2020

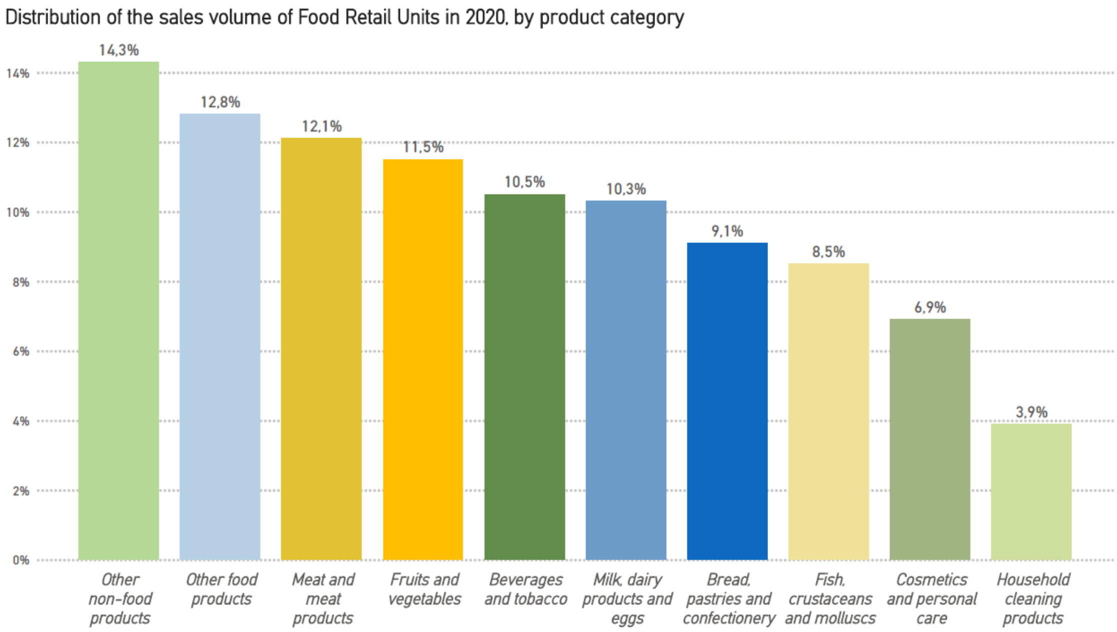


Figure A- 1. Distribution of the sales volume of Food Retail Units in 2020, by product category (adapted from INE, 2021; own elaboration)

- The main category 'food, beverages and tobacco' represented 74.8% of total sales of food retail establishments (+1.8% compared to 2019), with a value of 10.3 billion euros.
- The 'other food products' (rice, pasta, cereals, among others) generated the highest revenue of 12.8% of total sales (+0.4% than in 2019), followed by 'meat and meat products' with 12.1% (+0.3%), and 'fruits and vegetables' with 11.5% (+0.6%).

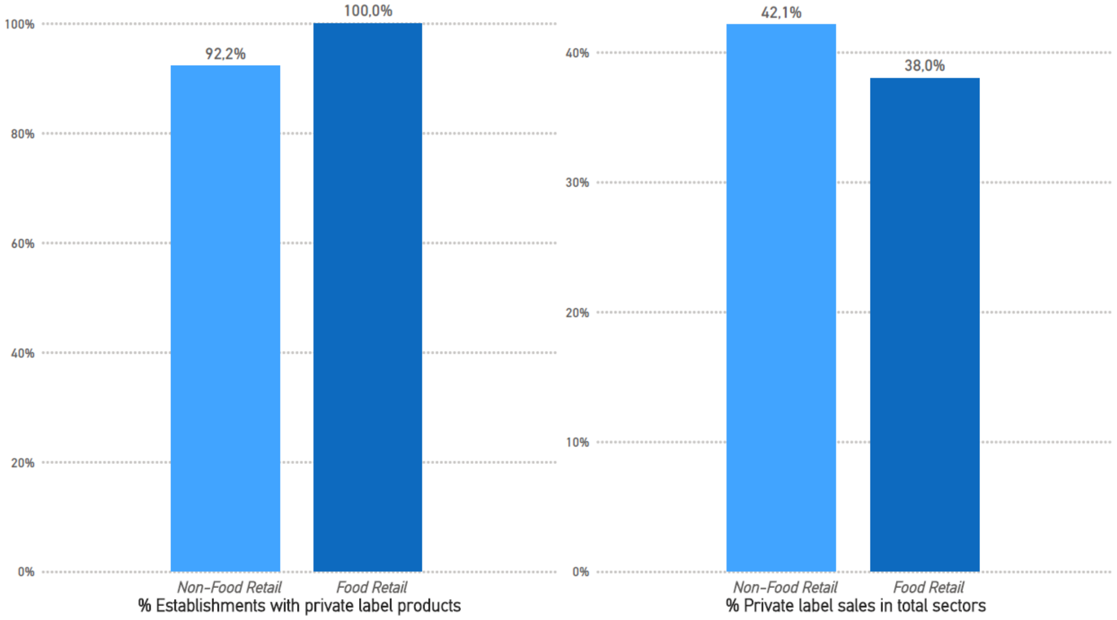


Figure A- 2. Private label products of Retail Units in 2020 (adapted from INE, 2021; own elaboration)

- Private label products were offered by all established food retail units and accounted for 38% of total sales, totalling 5.2 billion euros (+10.6% than in 2019).

Appendix 3. Data Analysis and Pre-processing

This appendix is devoted to the Data Analysis and Pre-processing step of the Time Series Forecasting process included in section *2.3.1.1.* of the Literature Review.

1. Data Analysis and Visualisation

Time series data has three characteristics: common patterns, serial dependence on previous values, and how its statistical properties vary over time (Wang, Smith and Hyndman, 2006).

1.1. Identification and Decomposition of Time Series Patterns

Firstly, it is known time series often include trend, seasonality and cycles which are recognized components that reflect the behaviour of the values over time. A trend reflects either a long-term increase or decrease in the time series (Chatfield, 2003), while seasonality is a period of regular and predictable variations that occurs at specific time intervals due to the effect of seasonal factors, such as time of year or day of the week (Davey and Flores, 1993). A cycle occurs when there are fluctuations in time series values that are not of a fixed frequency and typically last for at least two years (Stock and Watson, 1999). Seasonality and trends are seen as patterns of particular interest, compared to cycles that are not so predictable for science. However, visually identifying these components is not straightforward. For this reason, a time series decomposition method called STL (Seasonal and Trend decomposition using Loess method) developed by R.B. Cleveland (Robert B. Cleveland, William S. Cleveland, and Irma Terpenning, 1990) is employed, which separates trend (T_t) and seasonal (S_t) patterns from the remainder patterns (R_t). These techniques are useful for subsequently de-trend or de-seasonalise the time series and assess the impact of these patterns on the series as a whole.

1.2. Analysis of Serial Dependence and Autocorrelation in Time Series

Secondly, time series are usually correlated with their previous values and thus have serial dependence (Kritzman, 1994). This can be investigated through the autocorrelation coefficient $(r_k),$ which measures the sample correlation between the observation in time period $t\,(y_t),$ and the observation k periods prior (y_{t-k}) assuming values between -1 and 1. The Autocorrelation Function (ACF) represents the graph plot of the autocorrelation coefficients as a function of the lag k (George E. P. Box *et al.*, 2015) and is a useful complementary tool that reflects common patterns of trend and seasonality in the time series. As a rule, for trend data autocorrelations tend to be large and positive for small lags, decreasing slowly as lags increase. In the case of seasonal data, autocorrelations tend to be larger at seasonal lags (John E. Hanke and Dean Wichern, 2014). In addition, upper and lower bounds help to assess the significance of the autocorrelations. If one or more autocorrelations exceed the bounds, it means the coefficients are different from zero, which indicates the presence of patterns in the time series. On the other hand, if all autocorrelations remain within the bounds, the time series is given as serial independent and considered a good model for forecasting. In the latter case, the time series is called white noise and contains normally distributed variables with a mean of zero, which implies a lack of correlation between residuals (Quenouille, 1949). Another related statistical metric is the partial autocorrelation coefficient, which measures the direct correlations between y_t and y_{t-k}. The Partial Autocorrelation Function (PACF) is the graph plot of the partial autocorrelation coefficients as a function of the lag k (Dégerine and Lambert-Lacroix, 2003). When the ACF and PACF are plotted, the significance of each autocorrelation coefficient is analysed separately. Note that this analysis must be supported by the computation of Portmanteau tests, such as the Box-Pierce test or the Ljung-Box test (Ljung and Box, 1978). Both test a group of autocorrelation coefficients of a time series, assessing the null hypothesis of residuals with no autocorrelation.

1.3. Evaluation of Stationarity in Time Series and Unit Root Tests

Finally, the last feature is translated into the concept of stationarity, which refers to a time-invariant time series, in terms of their statistical properties of mean, variance, and autocorrelation structure. This simply means the time series changes around a fixed level, always independent of time (John E. Hanke and Dean Wichern, 2014). In practice, time series prove to be non-stationary since they are impacted by meaningful trend and seasonality patterns, as already pointed out (Chatfield, 2003). Unit root tests, among the most popular being the Augmented Dickey-Fuller (ADF) test and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are used to determine the required order of differencing for a given time series. A unit root measures the recurrence of non-constant patterns in a time series (Glynn, Perera and Verma, 2007; Herranz, 2017). The ADF test considers as a null hypothesis that time series has a unit root and is therefore non-stationary, while the KPSS test operates oppositely, with a null hypothesis that data is stationary (Harris, 1992; Kwiatkowski *et al.*, 1992; Lopez, 1997).

2. Pre-processing Tasks

Regarding pre-processing, this involves handling missing values and applying a series of transformations to the time series data to be ingested into the forecasting models.

2.1. Methods for Missing Value Imputation in Time Series

There are various methods for missing value imputation, including imputation by the mean, mode, or median of the series, imputation by the last observation carried forward, imputation by linear interpolation, imputation by a random sample, replacement by a defined value, or removal of missing values (Moritz and Bartz-Beielstein, 2017; Ahn, Sun and Kim, 2021).

2.2. Transformations Applied to Time Series Data

Transformations are applied to time series data based on their characteristics, as described in the first part. A frequent one is the logarithmic transformation, which aims to stabilise the variance of the time series, resulting in substantial improvements in the forecasting of economic variables (Lütkepohl and Xu, 2012; FENG *et al.*, 2014). Another is data normalisation, which reduces the values to a smaller common scale, usually between 0 and 1, without changing their original spacing. The most widely used techniques of data normalisation are MinMax and Z-score (Asesh, 2022; Kamani, Parmar and Ghodasara, 2022). One last transformation to mention is differencing, which stabilises the mean of a time series by removing changes in its level, eliminating trend and seasonal effects. The differenced series is obtained from the change between consecutive observations in the original series, which turns it into a stationary time series. Differenced data often requires a second-order differencing, if the first order does not achieve stationarity. Seasonal differencing is considered when data shows strong seasonal patterns, which compares the difference between observations in the same season. Sometimes, both types of differencing are combined to achieve stationarity (Shumway and Stoffer, 2011).

3. Data set splitting for Model fitting and Evaluation

Lastly, split the data set into a training and a test set. The training set is used to fit the models and compute the parameters, and the test set to evaluate their performance. Typically, one-step forecasts are considered in the training data and multiple forecasting horizons are in the test data (Medar, Rajpurohit and Rashmi, 2017).

Appendix 4. Summary of Box-Jenkins Method and Automatic ARIMA for forecasting process

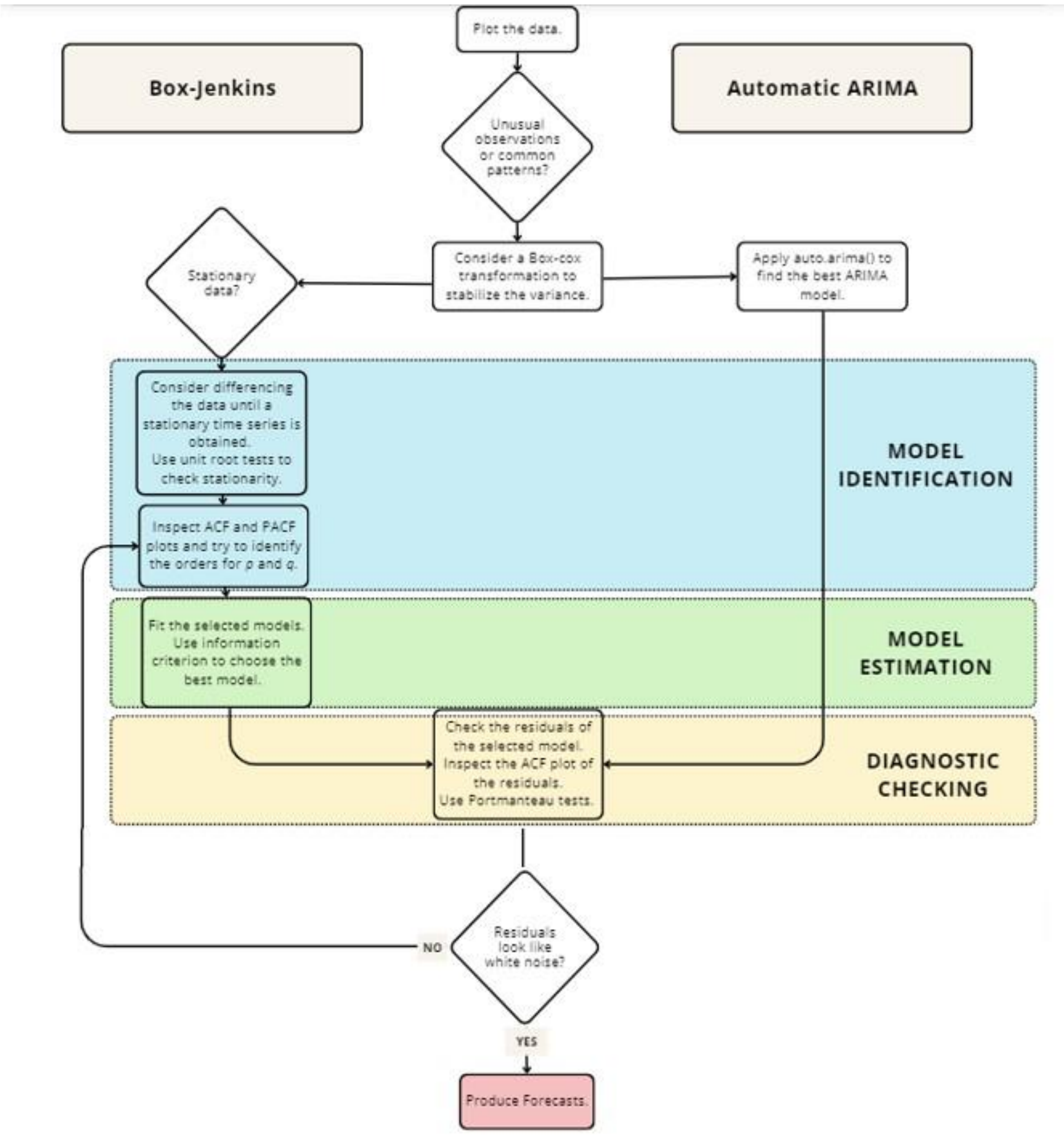

Plot the data.
Box-Jenkins
Automatic ARIMA
Unusual observations or common patterns?
Stationary data?
Consider a Box-cox transformation to stabilize the variance.
Apply auto.arima() to find the best ARIMA model.
Consider differencing the data until a stationary time series is obtained. Use unit root tests to check stationarity.
MODEL IDENTIFICATION
Inspect ACF and PACF plots and try to identify the orders for p and q.
Fit the selected models. Use information criterion to choose the best model.
MODEL ESTIMATION
Check the residuals of the selected model. Inspect the ACF plot of the residuals. Use Portmanteau tests.
DIAGNOSTIC CHECKING
NO
Residuals look like white noise?
YES
Produce Forecasts.

Figure A- 3. Box-Jenkins Method and Automatic ARIMA (adapted from Hyndman and Khandakar, 2008; George E. P. Box *et al.*, 2015 ; own elaboration)

Appendix 5. Description of Box-Jenkins Method Steps

Table A- 2. Steps of the Box-Jenkins Method (George E. P. Box *et al.*, 2015)

Box-Jenkins Method	**Description of Steps**
Model Identification	Check the stationarity of the time series and, in the case of a non-stationary pattern, determine the required number of differences. The stationarity condition is inspected by analysing ACF, PACF and time series plots to identify seasonal patterns or possible trends. In addition, unit root tests can be used if stationarity is uncertain. Once the time series has been transformed to a stationary format, another plot of ACF and PACF is generated to attempt to identify the correct orders of the autoregressive and moving average components for the set of candidate models.
Model Estimation	After identifying the possible orders and number of differences for the components, the model coefficients are estimated using the MLE method. To this end, information criterion measures are used to select the most accurate model.
Diagnostic Checking	Check the residuals behaviour in the chosen model. The residuals of a suitable model should be white noise, with no autocorrelation between them or evidence of any significant pattern. To validate the model, the ACF of the residuals is plotted, supported by the computation of the Portmanteau tests.

Appendix 6. Definitions regarding Data Modelling

This appendix contains a set of definitions related to Data Modelling.

1. Data Modelling:

- Multidimensional Modelling: The most widely used data model for DW. An intuitive technique for structuring data with the ability of optimising the performance of system queries (Moody and Kortink, 2003; Ralph Kimball *et al.*, 2008).
- Dimensional Model: It is composed by a central fact table and a set of dimension tables, which are related through a data model and represented in the form of a schema (Kimball and Ross, 2013).

- Data Model: It consists of a set of tables interconnected through relationships between identifying keys, i.e. columns shared between each set of tables (Ferrari and Russo, 2016).

2. Key Concepts in Database Design:

- Identifiers: These comprise a primary key, which labels each row in the table, and a foreign key, which references a value that is found in another table.
- Relationship: This operates by matching data in key columns or fields with the same name. The relationship between tables is determined by a certain direction and cardinality (Allen and Terry, 2005).
- Direction: This is established between a source table and a target table, where the attribute to be investigated is obtained from the source table, and the search is conducted in the destination table.
- Cardinality: This refers to the number of times each value in the key column appears in the related columns, encompassing three possible types of cardinality: one-to-one, one-to-many (or many-to-one), and many-to-many (Allen and Terry, 2005).

3. Tables of Dimensional Model:

- Fact Table: It is at the centre of the schema, related to the business subject to be analysed and is composed of a set of attributes, or facts, and a group of foreign keys, that relate the table to different dimension tables (Santos and Ramos, 2006).
- Dimension Tables: They provide background information for fact tables and are therefore formed by attributes that answer questions such as: “Who?”, “What?”, “When?”, “Where?”, “How?”, and “Why?” (Moody and Kortink, 2003).

4. Types of Multidimensional Models:

- Star Schema: It is a visual representation with a central fact table and several dimension tables, whose attributes are linked to the fact. This structure denormalises the data, which means adding redundant columns to some dimensions to improve query performance and enable quick access to aggregated data. Also, it allows for a straightforward and intuitive design for business users (Ferrari and Russo, 2016).

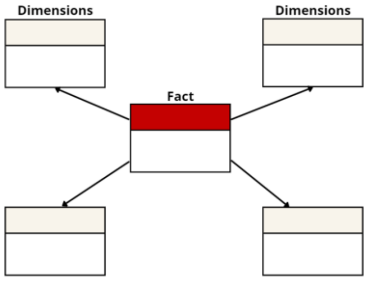


Figure A- 4. Generic Structure of a Star Schema (adapted from Moody, 2000; own elaboration)

- Snowflake Schema: It relies on the star schema structure, but it normalises some dimension tables, dividing data into additional related tables. This results in more efficient storage space and less data redundancy (Han and Kamber, 2013).

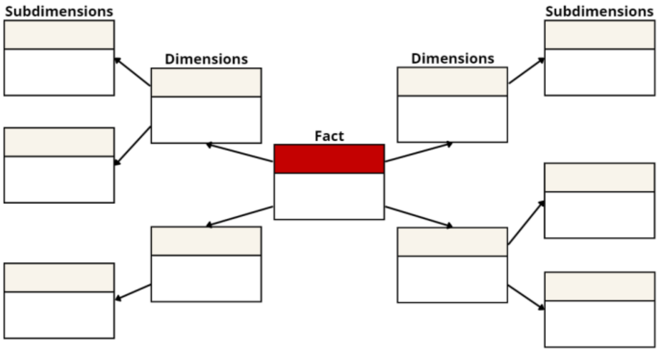


Figure A- 5. Generic Structure of a Snowflake Schema (adapted from Moody, 2000; own elaboration)

- Fact Constellation (or Galaxy Schema): It refers to a scenario where multiple fact tables are interconnected through shared dimensions in a combination of several star schemas. This design pattern is suitable for modelling complex business scenarios, where different types of business events or measures are related to common dimensions (Santos and Ramos, 2006; Han and Kamber, 2013).

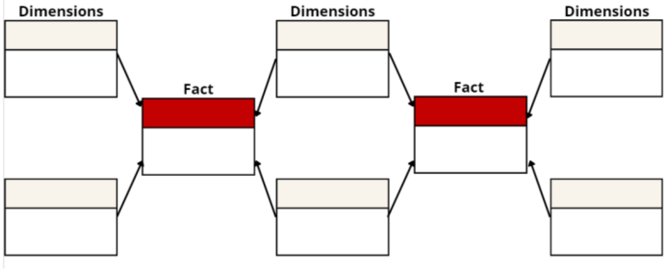


Figure A- 6. Generic Structure of a Fact Constellation (adapted from Moody, 2000; own elaboration)

Appendix 7. Software Integration - Power BI with R

This appendix provides an overview of the software solutions employed in section *3.1.* of the CRISP-DM Phases.

1. Microsoft Power BI

Power BI is a Self-Service BI (SSBI) solution developed by Microsoft that was released in July 2015. SSBI is an approach that allows business users to easily develop solutions without a technical background in BI or statistical analysis to inform their decision-making (Ferrari and Russo, 2016). Hence, the tool enables the connection of cloud and local data sources to transform them into coherent, interactive, and easy-to-visualise insights. These insights are systematised in dashboards which can be shared as BI Reports in the final deployment phase. Within data analytics, Power BI uses the DAX (Data Analysis Expressions) programming language to execute formulas and calculations on the underlying data infrastructure (Powell, 2018; Becker and Gould, 2019).

2. RStudio Integrated Development Environment (IDE)

RStudio Integrated Development Environment (IDE) is a programming language for statistical computing and graphics initially launched in February 2011 by Posit PBC (Verzani, 2011; Horton and Kleinman, 2015).

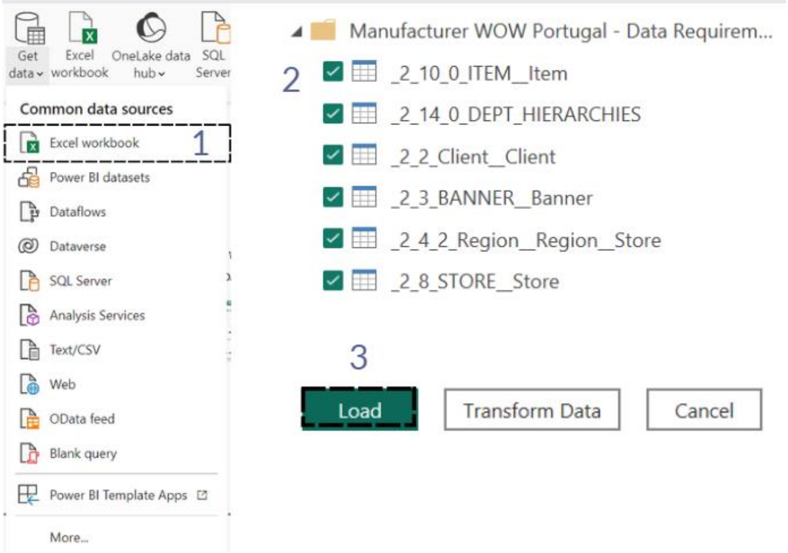
3. Integration Constraints and Considerations

There are constraints to consider when integrating Power BI with R scripts for custom visuals and advanced analytics. The input data is limited to 150,000 rows and 250 MB with a 60-second calculation timeout. R visuals are not interactive, do not support tooltips, and cannot be selected to cross-filter other visuals. These are the main factors which are limiting the full range of R capabilities, script execution performance, and resource consumption. In this way, the direct integration between the two tools was not a viable approach, and time series analysis was conducted in the standalone R Studio Desktop environment.

4. Combined Approach and Advantages

Combining both platforms offers numerous advantages, allowing data analysts to obtain more advanced forecasting models by using R’s extensive collection of packages and customising them to suit the analysis needs (Powell, 2017; Campbell, 2019). In conclusion, this BI integration solution was seen as a complete approach that met the project’s forecasting and visualisation requirements.

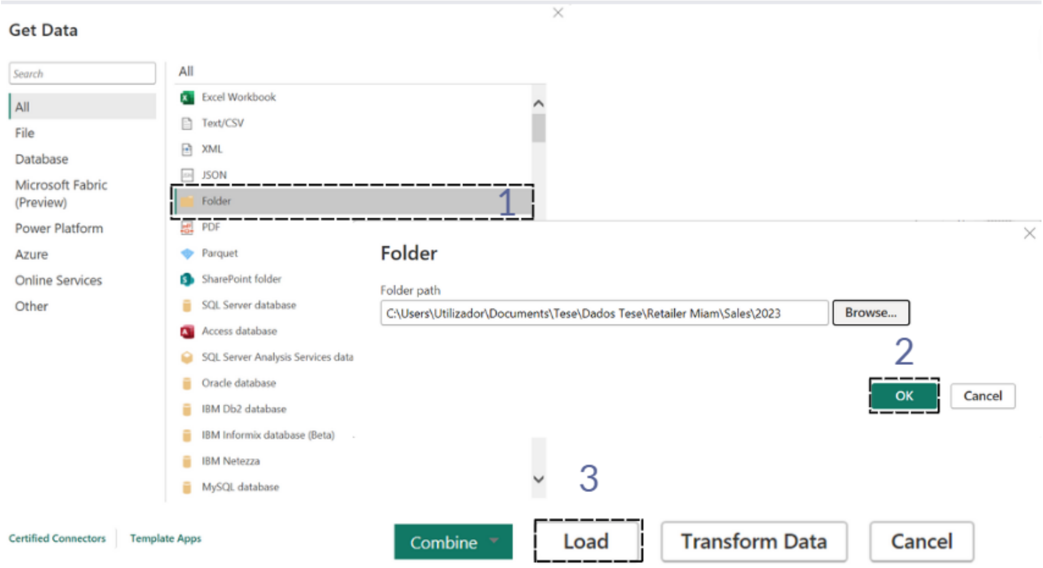
Appendix 8. Import, Select, and Load data from Excel files into Power BI



☑ _2_10_0_ITEM__Item
☑ _2_14_0_DEPT_HIERARCHIES
☑ _2_2_Client__Client
☑ _2_3_BANNER__Banner
☑ _2_4_2_Region__Region__Store
☑ _2_8_STORE__Store

Figure A- 7. Steps to Import, Select, and Load data from Excel files into Power BI

Appendix 9. Import, Select, and Load folder data into Power BI



Folder path: C:\Users\Utilizador\Documents\Tese\Dados Tese\Retailer Miam\Sales\2023

Figure A- 8. Steps to Import, Select, and Load folder data into Power BI

Appendix 10. Detailed Data Preparation Tasks

This appendix provides a detailed description of the Data Preparation tasks conducted in section *3.1.3.* of the CRISP-DM Phases.

1. Transformations in the Power Query Editor:

- Data Selection: Remove three top null rows by transposing the table twice and promoting headers (e.g. Excel sales files of Miam); Filter by the column 'Category' to include only relevant segments to the sweet snacks category (e.g. Excel sales files of Munch).
- Data Cleansing and Enrichment: Renaming tables to aid in identification during the transformation process; Check data type of each column for entire data set (e.g. changed data type for column 'StoreCustomerID' from 'whole number' to 'text'); Eliminate blank rows and columns in some tables (e.g. blank fields in sales table for Munch); Remove null values (e.g. column 'EAN' and five NA rows in column 'Region_StoreName' due to an oversight by Miam and WOW Portugal in providing the information); Remove irrelevant columns from retailers' sales table (e.g. columns 'EAN' and 'StoreCustomerID', which were initially needed to establish relationships between tables).
- Data Exploration: Create a new column (e.g. new column 'StoreName' resulting from various transformations, including a column split by space and a merge of columns 'BannerName' and 'StoreID').
- Data Integration (Fact Table): Append queries to combine daily sales data from various folders by year into two tables relating to sales by retailer; Rename fields to align data dictionary among retailers; Build Sales Fact table with left join of two daily sales tables of Miam and Munch through the column 'EAN' to obtain attribute 'SalesUnits', with remaining attributes being foreign keys of dimension tables.
- Data Integration (Dimension Tables): Perform inner join between Item and Item Hierarchies tables through common field 'Subcategory' for single-dimension table containing item information; Carry out same data integration process between Client and Banner dimensions to incorporate retailer profile information into a single banner dimension table; Create new dimension table entitled Brand

from the Item dimension as it was considered a relevant attribute to incorporate into the final model.

- Data Integration (Final Model): Execute two left joins to establish core relationships between the Sales Fact table and Item, and Store dimensions.
- Select the ‘Close & Apply’ option on the Home tab to apply changes made in Power Query Editor to reflect them in the Power BI model.

2. Power BI Model View Interface:

- Power BI advantage: The ability to create connections and cardinality between tables automatically; the user needs to manually set up missing links in the Model view interface.
- Model Relationships: All relationships are established automatically; except for the connection between Sales fact and MasterCalendar dimension via attribute ‘SalesDate’, which was subsequently established manually.
- MasterCalendar table: Created to enrich time-based analysis with a complete and consistent set of time-related attributes.
- Snowflake Schema: It refers to the multidimensional model with a Sales Fact table clearly identifiable at the centre, and the remaining six dimension tables.

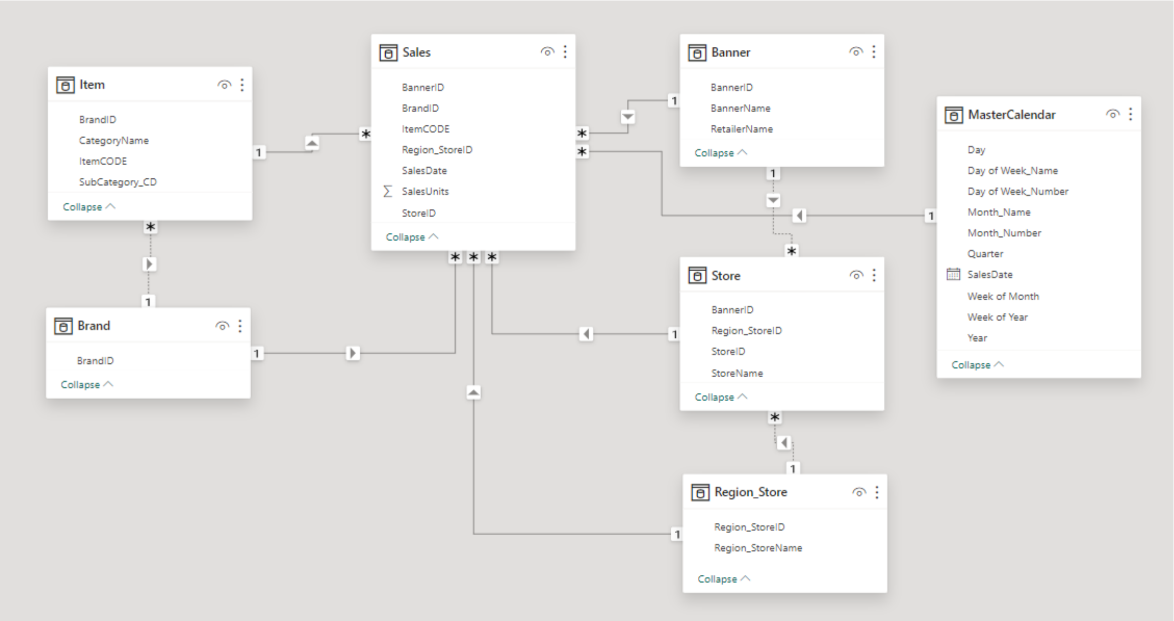


Figure A- 9. Dimensional Model

- Active and inactive relationships: Manage active connections between the Sales fact table and dimensions; Ensure inactive connections between different levels of dimension hierarchy.

Table A- 3. Active Relationships in Dimensional Model

Fact Table	Dimension Table	Common Attribute
Sales	**MasterCalendar**	SalesDate
	Banner	BannerID
	Region_Store	Region_StoreID
	Store	StoreID
	Brand	BrandID
	Item	ItemCODE

Table A- 4. Inactive Relationships in Dimensional Model

Dimension Table	Dimension Table	Common Attribute
Banner	**Store**	BannerID
Store	**Region_Store**	Region_StoreID
Item	**Brand**	BrandID

Appendix 11. Dimension of the ARIMA *tsibbles* and Prophet *data frames* for each Period in R analysis

Table A- 5. ARIMA and Prophet Observations across pre-crisis, crisis and post-crisis periods

	# ARIMA observations			**# Prophet observations**		
	Pre-Crisis	**Crisis**	**Post-Crisis**	**Pre-Crisis**	**Crisis**	**Post-Crisis**
Miam	18,067,574	15,095,833	14,149,497	790	671	546
Munch	16,165,194	15,653,551	15,493,526	785	668	543

Appendix 12. ARIMA Train and Test Sets for Model Selection and Fitting

Table A- 6. ARIMA Train and Test Sets across pre-crisis, crisis and post-crisis periods

ARIMA	**Training set: first 500,000 observations Test set: last 200,000 observations**		
	Pre-Crisis	**Crisis**	**Post-Crisis**
Miam	2018/01/01 – 2018/01/26 2020/02/22 – 2020/02/29	2020/03/01 – 2020/03/22 2021/12/22 – 2021/12/31	2022/01/01 – 2022/01/22 2023/06/23 – 2023/06/30
Munch	2018/01/02 – 2018/01/29 2020/02/21 – 2020/02/29	2020/03/01 – 2020/03/21 2021/12/23 – 2021/12/31	2022/01/02 – 2022/01/20 2023/06/23 – 2023/06/30
ARIMA	**Training set: 80% Test set: 20%**		
	Pre-Crisis	**Crisis**	**Post-Crisis**
Miam	2018/01/01 – 2019/10/01 2019/10/02 – 2020/02/29	2020/03/01 – 2021/08/30 2021/08/31 – 2021/12/31	2022/01/01 – 2023/03/12 2023/03/13 – 2023/06/30
Munch	2018/01/02 – 2019/10/08 2019/10/09 – 2020/02/29	2020/03/01 – 2021/09/02 2021/09/03 – 2021/12/31	2022/01/02 – 2023/03/14 2023/03/15 – 2023/06/30

- Note: Retailer Munch does not provide sales data for 25/12 and 01/01. As such, the training and test sets do not correspond precisely to those of the Retailer Miam.

Appendix 13. Prophet Train and Test Sets for Model Fitting

Table A- 7. Prophet Train and Test Sets across pre-crisis, crisis and post-crisis periods

Prophet	**Training set: 80% Test set: 20%**		
	Pre-Crisis	**Crisis**	**Post-Crisis**
Miam	2018/01/01 – 2019/09/24 2019/09/25 – 2020/02/29	2020/03/01 – 2021/08/18 2021/08/19 – 2021/12/31	2022/01/01 – 2023/03/12 2023/03/13 – 2023/06/30
Munch	2018/01/02 – 2019/09/23 2019/09/24 – 2020/02/29	2020/03/01 – 2021/08/18 2021/08/19 – 2021/12/31	2022/01/02 – 2023/03/13 2023/03/14 – 2023/06/30

Appendix 14. Prophet Trend and Seasonality Components for Retailer Miam and Retailer Munch

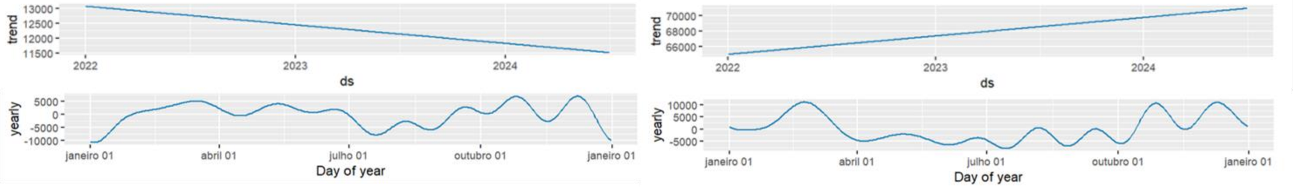


Figure A- 10. Prophet Trend and Seasonality Components of total Sales for Retailer Miam and Retailer Munch between July 2023-June 2024 [Prophet Output]