Politecnico di Torino

Master Degree in Engineering and Management: Innovation Management and Entrepreneurship
Master Thesis
2023-2024

In-depth study and development of a planning and forecasting application to generate forecasting scenarios in support to annual and monthly planning activities

Relator:
Prof.ssa Tania Cerquitelli
Dott. Daniele Mangano

Candidate:
Lorenzo Costa

Abstract

This thesis work aims to explore the use of IBM Planning Analytics, an integrated business planning solution, declined to the firm's function of Corporate Performance Management in the context of the large-scale distribution sector to develop a solution that integrates forecasting techniques into traditional planning activities. The main steps involved in this work relate to the analysis of the market situation for software in Corporate Performance Management, selection of the most suitable software, development of the solution and integration with real use-case. The potential offered by forecasting techniques is immense, but carrying out planning and forecasting activities as separate entities introduces friction that could diminish the potential benefit while, instead, merging into the planning process the tools offered by forecast could trigger highly valuable results. The necessity to combine the two activities was found in planners' difficulty in completing the activities on time and without impediments. By following the logics and adhering to the boundaries of Corporate Performance Management, a solution has been developed in order to obtain quicker, safer and more precise planning processes with the added possibility to exploit forecasting techniques.

Summary

List of Figures

1. Introduction

The present work is the result of a six-month project made possible by the company "Mediamente Consulting s.r.l.", operating in the Business Analytics field. The company is divided into five business units, although each area is not tightly separated and, indeed, there is frequent mixing and sharing of knowledge and projects between them. These are Technological Infrastructure, Data Integration and Management, Corporate Performance Management, Business Intelligence and Advanced Analytics.

Corporate Performance Management (CPM), the business unit which my work belonged to, is an umbrella term that describes the methodologies, metrics, processes and systems used to monitor and manage the business performance of an enterprise (Gartner, s.d.). A well-structured CPM strategy is able to identify problems and provide solutions with the aim of reviewing outdated methods, improving efficiency and increasing productivity. Applications that enable CPM translate strategically focused information to operational plans and send aggregated results that trigger the real value of CPM: fast analysis in many different situations that allows fast responses to changes in markets. These applications are also integrated into many elements of the planning and control cycle and for CPM to be useful, organizations should implement analytical applications that can support the processes, methods and metrics used in business performance management (Formula, s.d.).

By integrating business planning, sales, marketing, forecasting, budgeting, human resources and operations organizations can link their business goals and strategies to their plans and execution. Aligning the business around its strategic priorities allows it to focus attention on the key drivers of business operations, as well as the key business metrics that must be maintained to improve revenues and increase profits. Corporate Performance Management often includes the following important management processes:

- Creation of a business model and identification of corporate objectives
- Budget, planning and forecasting (BP&F)
- Merge results and consolidate accounting books on a regular basis
- Sharing results with internal and external stakeholders
- Analysis of business performance about plans, previous years and across products and divisions
- Remodeling or replanning based on results and new forecasts

By introducing forecasting activities into common CPM practices firms may gain advantages and avoid sources of human error or undue influences that could result critical in key processes and methodologies. This introduction drove the necessity to find solutions that would give the possibility to merge existing applications with the classical needs of CPM and the new implementation of a forecasting framework. The boundaries set by the CPM's range of operations defined clear requirements to be satisfied. The software considered in the following analysis needed to primarily satisfy the requirements imposed by CPM techniques and then be appropriate to forecasting specifications.

Studies and analyses conducted for this work lead to a definition of the thesis' structure as follows:

- Chapter 2 provides an overview of the theoretical background on planning including its importance, benefits and drawbacks, types of different planning, and impact of implementation on firms.
- Chapter 3 also give a wide basic knowledge about forecasting. It is about the methods, applications, and challenges of forecasting, and the tools and steps involved in the overall process providing a deeper analysis of available algorithms that can be adopted to the considered case.
- Chapter 4 introduces IBM Planning Analytics and its competitors in the market. It describes the main components, functionalities, and benefits of IBM Planning Analytics, as well as its data management and model development features.
- Chapter 5 presents a case study about the development of an application for planning and forecasting using IBM Planning Analytics. It describes the steps and processes involved in creating the application, such as data management, model definition, forecast generation, budget planning, and variance analysis. It also demonstrates the application's features and capabilities using real-world data from a large-scale distribution firm.
- Chapter 6 summarizes the main findings and contributions of the thesis. It highlights the benefits and challenges of using IBM Planning Analytics for planning and forecasting activities. It also provides some suggestions for future research and development in this field.

2. Planning

In today's dynamic and competitive business environment, effective planning plays a crucial role in the success and sustainability of firms. Planning is the formulation of a program or a scheme or a project that defines a series of action or behaviors that need to be executed based on external factors. “Planning is deciding in advance what to do, how to do it, when to do it and who is to do it”, according to Harold Koontz and Cyril O'Donnell (1972). The purpose of business planning is not simply to create a plan – it is to make better decisions, better resource allocation and in general better business activities execution. This chapter aims to delve the importance of planning for firms, the different type of plans utilized in business management, and examine the impact of planning on firm performance. By understanding the importance of planning and its implications on organizational outcomes, managers and practitioners can improve their strategic decision-taking process and steer the company toward long term competitiveness.

Organizations require a lot of planning - some formal, but much of it informal and almost always in business silos. When business managers are asked to prepare a budget, they first create a business plan, informally, and then translate that into financial terms. Then, all these individual siloed plans, generally developed by managers, need to be merged into one general company plan that will account for every aspect of each single plan and therefore will make every need coincide. The impact that a well-structured process can have on the overall performance of a firm might be of no negligible importance and this define the need of people, processes and tools that allow to achieve the result.

People plan sales, how to produce products and deliver services, how to purchase, they plan for the headcount they'll need and how to organize distribution and their supply-chain. This logical sequence of tasks defines a different approach to organizational planning that involve a structured and comprehensive plan of sales from which the depending plans are derived. Sales are the determinant of production and purchase, which in turn determine the need for human capital and supply-chain planning. This top-down type of plans definition follows a path in which the management has more control: higher-level plans are determined upfront and lower-level plans must reflect decisions taken. The latter can be seen as a more integrated approach as the goal is, for the different business functions, to work together so that plans are designed to be integrated with each other. This finds application in Customer Performance Management approaches: plans, operations and functions are designed in such a way that the subsequent integration is as smooth as possible.

Though different kind of planning (strategic, operational, financial, environmental etc.) with different time horizons lead to various externalities, the ultimate objectives and goals of drawing plans can be considered to be:

- achievement of organizational goals and vision.
- facilitate decision making in various circumstances.
- provide overall control over an organization.
- reduce uncertainty and unexpected situations.
- optimum resources utilization.
- Innovation promotion in management techniques.

2.1. The importance of planning for a firm

To establish the importance of planning for firms, various studies have shown its role in rising organizational effectiveness. For instance, Mintzberg, Ahlstrand, & Lampel (Mintzberg, Ahlstrand, & Lampel, 1998) highlight that effective planning enables firms to anticipate future challenges, make informed choices, and align resources to achieve strategic objectives. Similarly, Hill & Jones, (Hill & Jones, 2010) found that planning assists firms in identifying market opportunities, evaluating potential threats, and formulating strategies to gain a competitive advantage. These studies underscore the critical role of planning as a proactive process that enables firms to adapt to dynamic market conditions and capitalize on new and emerging trends.

Firms that aim to growth and development need to put efforts in the definitions of clear and feasible plans and, at the same time, ensure that these plans are respected and implemented systematically by workers and employees. The act of planning can be the trigger of many benefits into a business such as:

- lead the organization in becoming proactive rather than reactive by setting objectives and benchmarks.
- instills a shared sense of responsibility in individuals and groups.
- stimulate the recognition of past mistakes.
- improves staff satisfaction and retention.
- manages expectations and bolsters trust.
- allows flexibility in planning and spending.

A clear and defined path to be followed along a period, help employees to self-organize their work and eventually discuss about it with managers, to modify schedules and deadlines. Accepting a plan to follow gives individuals the responsibility to carry it out and make them feel satisfied when achieved. The recognition of the work done is a potential trigger for rewards and this creates a perception of interest in the individual, leading to his retention. At a higher organizational level, a plan, allows to improve the firm's processes by recognizing problems happened in past and then take action to avoid its repetition, at the same time it creates the necessity to set goals and objectives to be reached in a certain time frame, driving a company in the direction of action instead of reaction. Being in control of future events injects an overall sense of trust that can boost performances.

As any business decision, even planning cannot come without drawbacks, which in comparison with the benefits are negligible but still it's useful to study and understand them to mitigate and settle actions that can prevent some. Certain situations related to planning activity that might happen are: action prevention and creativity inhibition because individuals might focus on fulfilling the activities defined and not go beyond, leading to complacency for achieving the expected results, flexibility prevention that might even be discouraged in perspective of fulfillment of the assigned tasks. These can even be encouraged and considered as positive for certain firms but led to the extreme they often generate unfavorable circumstances. From an organizational standpoint, planning is a time and resources consuming activity, not easy to be performed if not completely understood and that can generate misleading plans because a raw planning process, based on rough procedures and gut, defines unreal goals.

Depending on the firm size the plan definition process can be more or less structured and can be performed one or multiple times along the year. Smaller firm might define long term plans that are yearly revised and modified, while big companies could be considered more prone to more frequent plans definition and review such yearly plans with quarterly or monthly checks. The review process involves controlling whether the estimates done are respected and if the plan is being actually implemented and, if necessary, some adjustments may be approved to re-target goals.

2.2. Type of planning

As previously mentioned, all business areas need a particular plan to define the activities that are going to be undertaken in future and so different plans are declined for different business units. Each area will merge the requirements of the specific unit and the general requirements of the business organization, to reach a consensus that meets management's expectations and predictions of "businesspeople". A plan can be categorized based on several dimensions, depending on the specific context and purpose.

1. Scope, each business area requires a different plan that can be differentiated by its purpose.
2. Level, mostly in bigger companies a certain degree of aggregation might be necessary so high-level plans are used for strategy and direction while lower levels may involve operations.
3. Time, based on the duration covered.
4. Objective, depending on goals and outcomes a plan can be focused on revenue growth, cost reduction, market expansion etc.
5. Audience, end-users define a further classification which outlines the intended audience that will interact with the plan.

To further analyze planning along the time dimension, a sub-classification can be given by:

- Long-term plans: typically covering a period of three (fast moving markets) to five years or more (ordinary markets). This horizon is focused in defining strategic goal and general direction of the organization but in such a long period socio-economic and environmental and market conditions changes might occur, in fact adjustments are often required. Top-level management is involved in this activity and is responsible of defining the firm's path for the coming years.
- Medium-term plans: generally, span one to three years for fast moving markets and one to five years for other markets, its focus is to translate strategic direction into actions, in fact they often involve setting targets, allocating resources, and implementing initiatives that contribute to the organization's long-term vision.
- Short-term plans: range from months to a year and even when a full year is covered it is broken down in quarters or shorter time periods. They are more tactical in nature and focus on specific objectives, in fact this kind of plans generally cover more on-the-field activities such as operational goals, sales targets, and marketing campaign. This kind of plans involves middle and first-level management which are entitled to update and monitor the implementation of the plan.
- Rolling plans: continuously updated and revised at regular intervals, allow organizations to adapt and respond to changes in the business environment more effectively.

The time dimension in planning may vary depending on the industry in which the firm operates. Fast moving/developing markets require shorter plans because it might be too difficult to look far ahead in time, while more established markets allow longer time frame because of more predictable market trends.

Defining this type of plans establish the strategy to be followed by the firm so it's a crucial activity that need to be monitored. It's important to periodically revise plans to determine if adherence to the plan is respected and if, or not, take actions to modify the plan or the general behavior of the firm.

The general classification that involves the various business areas includes:

- Strategic plan: used by and developed for the top-management, sets the overall direction and strategy of a company including vision and mission statements (David, 2017).
- Business plan: a document including information on the business concept, market analysis, financial projections, and marketing strategies.
- Operational plan: focused on day-to-day operations of a business, outlines specific actions, processes, and resources required to implement the specific strategic plan (Slack, Chambers, Johnston, & Betts, 2019). It can cover areas such as production, cost control, purchase, distribution, staffing, and quality control.
- Financial plan: outlining financial goals and strategies, it includes revenue projections, expense forecasts, budgeting, and funding strategies. This plan helps in effective finances management and insurance of financial stability of the organization. (Ross, Westerfield, & Roberts, 2020)
- Marketing plan: establish marketing objectives, target market, and strategies to promote a product or service and generally include market research, competitive analysis, branding, pricing, and promotional activities (Kotler, Keller, Brady, Goodman, & Hansen, 2017).
- Sales plan: it focuses on achieving sales targets and revenue goals and includes sales strategies, sales team structure, sales forecast, customer relationship management, and sales performance measurement.
- Human resource plan: it defines human resource strategies, such as recruitment, training and development, performance management, and employee retention. It is aimed to ensure that the organization has the right people with the required skills to achieve its objectives.
- Crisis management plan: includes risk assessment, communication protocols, and steps to mitigate potential damages that are mandatory to execute during emergencies or unexpected events that may disrupt business operations.
- Project plan: it defines the objectives, scope, timeline, and resources required for a specific project by including task breakdown, milestones, responsibilities, and project monitoring mechanisms.
- Contingency plans: are designed to address unforeseen events or potential risks that could impact business operations. They are typically developed as a response to specific threats such as natural disasters, economic downturns, or disruptions in the supply-chain. They provide guidelines and procedures for managing crises and minimizing its impact.

Depending on the industry and specific needs of an organization, additional plans may be developed to address unique challenges and opportunities.

2.3. Impact of planning on firm performances

Strategic planning has been an integral part of organizations while its impact on their performance has been debated for many years. A meta-analysis of 87 correlations from 31 empirical studies reveals that strategic planning has a positive, moderate, and significant impact on organizational performance (George, Walker, & Monster, 2019). The positive impact of strategic planning on organizational performance is strongest when performance is measured as effectiveness and when strategic planning is measured as formal strategic planning. This impact holds across sectors (private and public) and countries (U.S. and non-U.S. contexts) (George, Walker, & Monster, 2019). The research states that simply having a plan is not enough, an entire process defined and proven to be actually working need to be put in place in order to get the benefits offered from planning. Planning effectiveness is reached after successful implementation of strategies paved by management and this requires human and financial resources. According to the authors, "The findings suggest that strategic planning should be part of the standard managerial approaches in contemporary organizations and contradict many of the critiques of strategic planning".

Understanding the impact of planning on firm performance is crucial for organizations seeking to optimize their strategic decision-making processes. Empirical studies have consistently shown positive relationships between planning and firm's key performance indicators. For instance, a study by Dess et al. (Dess, Lumpkin, & Covin, 2003), conducted across a sample of manufacturing firms, demonstrated that firms with well-developed planning processes tend to outperform their counterparts in terms of profitability and growth. Similarly, a meta-analysis conducted by Powell found a positive correlation between planning and firm performance across various industries (Powell, 1992). These findings highlight the significant influence of planning on driving firm success and achieving superior performance outcomes.

Miller and Cardinal also found that strategic planning positively influences firm performance (Miller & Cardinal, 1994) and furthermore the higher the amount of strategic planning the higher the impact on financial performance. Strategic planning (SP) is one of the most popular management approaches in contemporary organizations, and it is consistently ranked among the five most popular managerial approaches worldwide (Rigby & Bilodeau, 2013) (Wolf & Floyd, 2017). The formality of the strategic planning processes (i.e., the extent to which strategic planning includes internal and external analyses and the formulation of goals, strategies, and plans) is important to enhancing organizational performance and SP is particularly potent in enhancing organizational effectiveness (i.e., whether organizations successfully achieve their goals), but it should not necessarily be undertaken in the hope of achieving efficiency gains (George, Walker, & Monster, 2019).

In a review on conceptual and practice perspectives by Jayawarna & Dissanayake, they argue that future research should shift from a focus on "Does planning lead to performance?" to "How, when, and why does strategic planning lead to performance?". They also suggest that it is equally important to identify internal and external factors affecting strategic planning and performance relationship as previous research were limited to finding directional relationship (Jayawarna & Dissanayake, 2019).

3. Forecasting

Forecasting is a fundamental aspect of strategic planning and decision-making within organizations. It is a process that involves predicting future outcomes based on historical data, current market trends and other external and internal factors of a company. The process of forecasting is not merely a guesswork but a systematic and scientific approach that uses statistical and mathematical techniques to predict future events. It is a forward-looking perspective that helps organizations prepare for the future rather than merely reacting to events as they occur. The importance of forecasting cannot be underestimated, as it plays a pivotal role in various aspects of an organization's operations, from production planning to financial management. The ability to accurately forecast future trends and events can significantly enhance a firm's performance, providing a competitive edge in the marketplace (Armstrong, 2001).

Forecasting is not just about predicting the future; it's about understanding the past and present to make informed decisions about the future. It involves analyzing historical data, identifying patterns and trends, and using this information to predict future outcomes. This process is crucial in helping organizations anticipate changes in the market, manage risks, and make strategic decisions. The accuracy of these forecasts can significantly impact on a firm's performance. Therefore, it is essential for organizations to invest in accurate and reliable forecasting methods.

Moreover, forecasting is not a one-time activity but a continuous process. As new data become available, forecasts are updated and refined to reflect the most updated information. This iterative process ensures that the forecasts remain relevant and accurate, providing the most reliable basis for decision-making. It is a critical tool used in various sectors to anticipate future trends and make reality-based decisions, in fact the benefits to firms are manifold. By accurately predicting future trends, firms can align their strategies and resources to meet anticipated changes, thereby gaining a competitive advantage (Hyndman & Athanasopoulos, 2018).

The mastery of forecasting can significantly improve a firm's risk management capabilities. By predicting potential risks and uncertainties, firms can develop strategies to mitigate these risks, thereby reducing their potential impact on business operations. This proactive approach can significantly enhance a firm's resilience and agility, enabling it to navigate through uncertain and volatile market conditions. Furthermore, forecasting can also improve a firm's strategic planning process. By providing insights into future trends and events, it helps firms to draw more effective and robust strategic plans, thereby enhancing their long-term performance and sustainability.

However, there are some drawbacks to forecasting that are worth considering (Galt, 2019):

1. Forecasts will never be 100% accurate: Despite advances in forecasting techniques and technologies, it is important to recognize that forecasts are inherently uncertain and subject to errors. The future is influenced by numerous factors, such as changes in customer preferences, market dynamics, economic conditions, and unforeseen events. Even with sophisticated models and expert analysis, it is impossible to predict future outcomes with absolute precision. Businesses need to acknowledge the inherent limitations of forecasting and incorporate a margin of error in their decision-

making processes. By understanding the potential range of forecast outcomes, businesses can make more reliable and flexible plans to adapt to changing circumstances.

2. It can be time-consuming and resource-intensive: Developing accurate forecasts requires significant time, effort, and resources. Forecasting involves collecting and organizing relevant data, analyzing historical patterns, and identifying relevant drivers and factors that influence demand. This process often requires close collaboration and coordination among different departments, such as sales, marketing, operations, and finance. Additionally, maintaining accurate and up-to-date data can be a challenge, especially when dealing with large data sets or complex product lines. Organizations need to allocate adequate resources, both in terms of personnel and technology, to ensure effective forecasting practices. Investing in automation tools and advanced forecasting software can streamline the process and free up valuable time for decision-makers to focus on analyzing and interpreting the results.

3. It can be costly: Implementing a robust forecasting system can involve significant upfront costs. This includes investing in skilled personnel, training, and acquiring advanced forecasting tools and software. Additionally, forecasting accuracy often improves with access to quality data sources, which may require investments in data collection and technological infrastructure. While these investments can be substantial, they are essential for generating reliable forecasts and facilitating better decision-making. By aligning forecasting investments with the potential benefits of improved accuracy and optimized inventory management, businesses can justify the costs over time. It is important to view forecasting activities as a strategic investment that can yield to long-term cost savings through improved operational efficiency, reduced stockouts, and minimized excess inventory levels.

Forecasts, despite their inherent limitations, are valuable tools for informed decision-making and proactive planning in an uncertain future. Executives should approach forecasts with a critical mindset, recognizing their potential insights while acknowledging their imperfections. By understanding and addressing the drawbacks of forecasting, businesses can adopt a realistic perspective. Integrating forecasts into the decision-making process enhances the ability to navigate uncertainties and make strategic choices. By uncovering hidden truths within forecasts, executives gain a comprehensive understanding of the present landscape. This awareness enables them to identify early indicators of future developments, giving a competitive advantage. Embracing forecasts as valuable tools unlocks insights and empowers executives to stay ahead in today's dynamic business landscape. (Reeves, Ramaswamy, & O'Dea, 2022)

3.1. Tools and Steps

Forecasting involves the use of various algorithms, each with its unique strengths and weaknesses. Some of the most commonly used includes time series methods, regression models, and machine learning algorithms.

Regression models are used when there is a known relationship between the variable being forecasted and other variables. These models can be particularly effective in forecasting outcomes that are influenced by multiple factors, such as market demand or sales revenue.

Machine learning algorithms, such as neural networks or random forests, are increasingly being used in forecasting due to their ability to handle large data sets and complex relationships. These algorithms can learn from historical data and adapt to changes in trends and patterns, thereby improving the accuracy of forecasts from the comparison of predicted data and actual values.

Time series methods will be object of further explanation because of their use in the development of the application object of this thesis.

Time series forecasting, used when historical data is available, involves fitting a model from past data and using it to make predictions beyond the known data range. The interdependence between successive observations necessitates careful consideration of their order during analysis to maintain data integrity. A time series refers to a collection of observations recorded in chronological order, or a set of observations measured sequentially through time in a process that stores information continuously in time or at discrete time points.

Main objectives of time series analyses are:

1. Description of data through summarization and graphical methods

2. Model the process that generates the data through a statistical model. Univariate time series only depend on past values of the variable defining the series while multivariate time series depend upon past values and other variables describing the process (predictors).

3. Forecasting to estimate future values.

4. Control actual values to take actions and give directions to processes.

A deterministic time series is the one for which future values can be predicted exactly from past values, but most series are stochastic, or random, and for these the future is only partly determined by past values. For these random series, finding an appropriate model to capture their random behavior enables accurate forecasting. It is important to note that time series analysis differs from other statistical problems in that the observed time series is typically the only realization available for analysis (Chatfield, 2000).

To generate a prediction, it’s necessary to use different techniques that may involve the analysis of the historical series or the analysis of factors that may characterize that series. In general, a forecast involves the analysis of historical data of the variable of interest, the creation of a model that estimate the behavior of the time series and the judgment of industry experts that should approve/validate the prediction.

Before analyzing a time series, what the analyst needs to perform is the IDA, Initial Data Analysis, that allows to gain confidence and to clean (removing or adjusting errors) the raw data set. The one represented below in Figure 1 is an example of a data set read and managed in Python from a ".csv" file while Figure 2 displays the same values in a tabular, Excel-like, representation.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36 entries, 0 to 35
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Date    36 non-null     datetime64[ns]
 1   Sales   36 non-null     float64
dtypes: datetime64[ns](1), float64(1)
memory usage: 704.0 bytes
(        Date  Sales
 0 2020-01-01  102.0
 1 2020-02-01  148.0
 2 2020-03-01   98.0
 3 2020-04-01   97.0
 4 2020-05-01  152.0,
          Date     Sales
 31 2022-08-01  177.0269
 32 2022-09-01  112.8695
 33 2022-10-01  117.6219
 34 2022-11-01  193.6603
 35 2022-12-01  108.1171,
```

Figure 1 - Data set Representation

sales_data

Date	Sales
01/01/2020	102
01/02/2020	148
01/03/2020	98
01/04/2020	97
01/05/2020	152
01/06/2020	101
01/07/2020	100
01/08/2020	149
01/09/2020	95
01/10/2020	99
01/11/2020	163
01/12/2020	91

Figure 2 - Tabular data set Format

Along the same process a decomposition is performed: the time series is analyzed, and the various component are singularly examined. The decomposition of a time series leads to 4 components:

- *Seasonal variation* that represents a cyclical behavior of the time series on a defined time frame. Data points should be measured on a frequency lower than the year because otherwise no seasonality could be detected.

- *Trend* defines an upward/downward shift of the mean value of the time series over several successive time periods.
- *Cyclic variation* includes regular variation on time frames longer than one year.
- *Irregular fluctuations*, or *residuals*, that include any other variation left over after other variations are removed, they may be completely random or exhibit short-term correlation or include discontinuities.

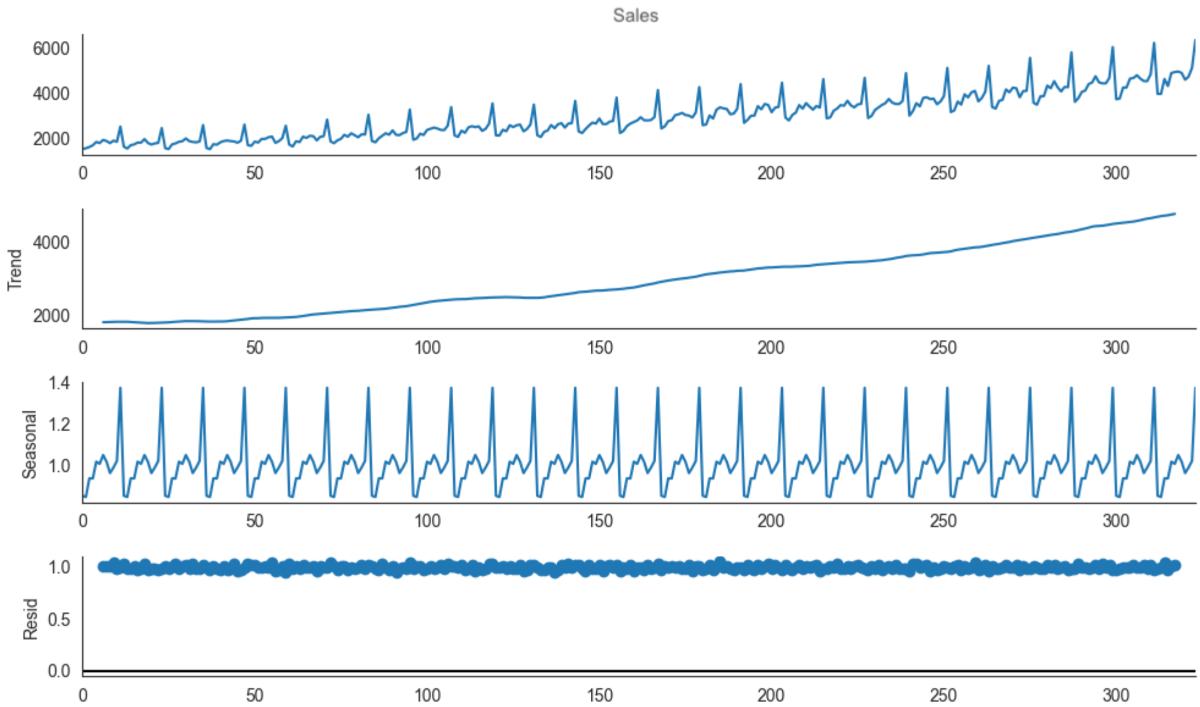


Figure 3 – Time series Decomposition

The first step to perform time series analysis is to plot the observations against time. This action allows to visualize the path of the data points and then identify eventual trend, seasonality, cycles, outliers, turning points and discontinuities. It's fundamental to choose an appropriate visualization of the graph in order to grant readability and comprehensibility of the data represented: scales play an important role and they should appropriately suit the interval of representation. The context is crucial in deciding how to modify data, if at all (Chatfield, 2000). This explains why it is essential to get background knowledge about the problem, and to clarify the objectives. It is essential to combine statistical theory with sound common sense and knowledge of the particular problem being tackled while performing data cleaning.

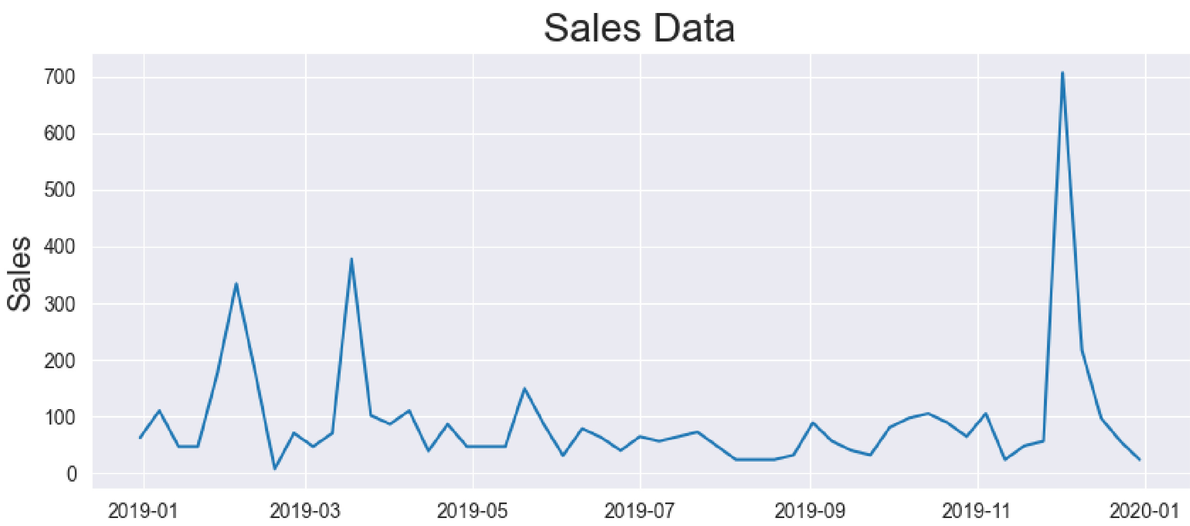


Figure 4 - Sales plot

The second step in preparing time series data for forecasting involves transformations, which may be necessary to stabilize variance, detrend, or de-seasonalize the series. Stabilizing the variance refers to reducing or eliminating the variability in the series over time. This is important because many forecasting methods assume that the series has a constant variance, also called stationary. If the variance changes significantly across the data, it can adversely affect the accuracy of the forecast. There are several methods commonly used to stabilize the variance of a time series including Logarithmic Transformation, Box-Cox Transformation, Square Root Transformation, and other Power Transformation. The choice of transformation method depends on the nature of the time series and the patterns observed. A widespread method used to obtain a stationary time series is differencing: it involves taking the difference between consecutive observations in a time series. The resulting series, called a differenced series, represents the changes or fluctuations between successive data points rather than the absolute values themselves.

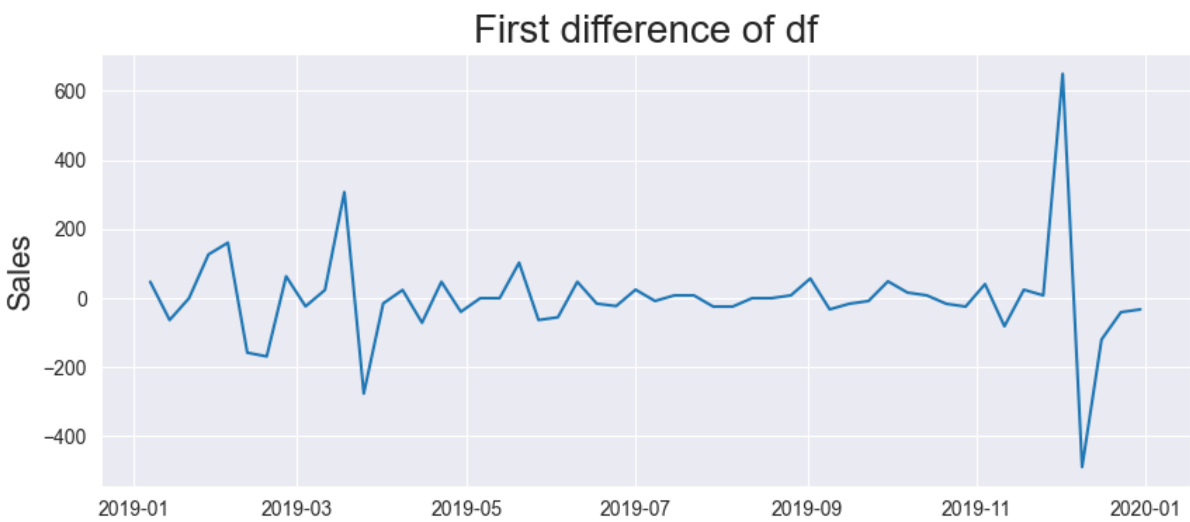


Figure 5 - Differenced Time series

This transformation helps eliminate or reduce trends and seasonality present in the original data, making the series stationary. It is important to note that after applying the transformations and the forecast, the forecasted values need to be back transformed to their original scale to obtain meaningful predictions. By stabilizing the variance of the time series, we ensure that the forecasting models can make reliable predictions without being influenced by excessive variability.

Third step is about trend analysis. A description of it is “the long-term change in the underlying mean level per time-unit”. This may appear reasonable, but what is meant by “long-term” and what is “the current mean level”? The perception of trend, and the understanding of “long-term”, depends on the length of the observed time series, too. It is also important to realize that the treatment of trend depends on whether or not seasonality is present. If seasonality is present, the analyst must decide whether to measure and/or remove the seasonality before or after measuring the trend and it is usual to adopt an iterative procedure. Preliminary estimates of trend and seasonality are found, typically with a fairly simple method and these estimates are then revised using a more sophisticated one, until more refined estimates are obtained. Thus, the treatment of trend and seasonality are inextricably related, reflecting the fact that there is no unique decomposition of variation into trend and seasonality (Chatfield, 2000). The measurement and removal of seasonal variation is complicated by the possible presence of calendar (or trading day) effects (ex. variations in business days in a month between years).

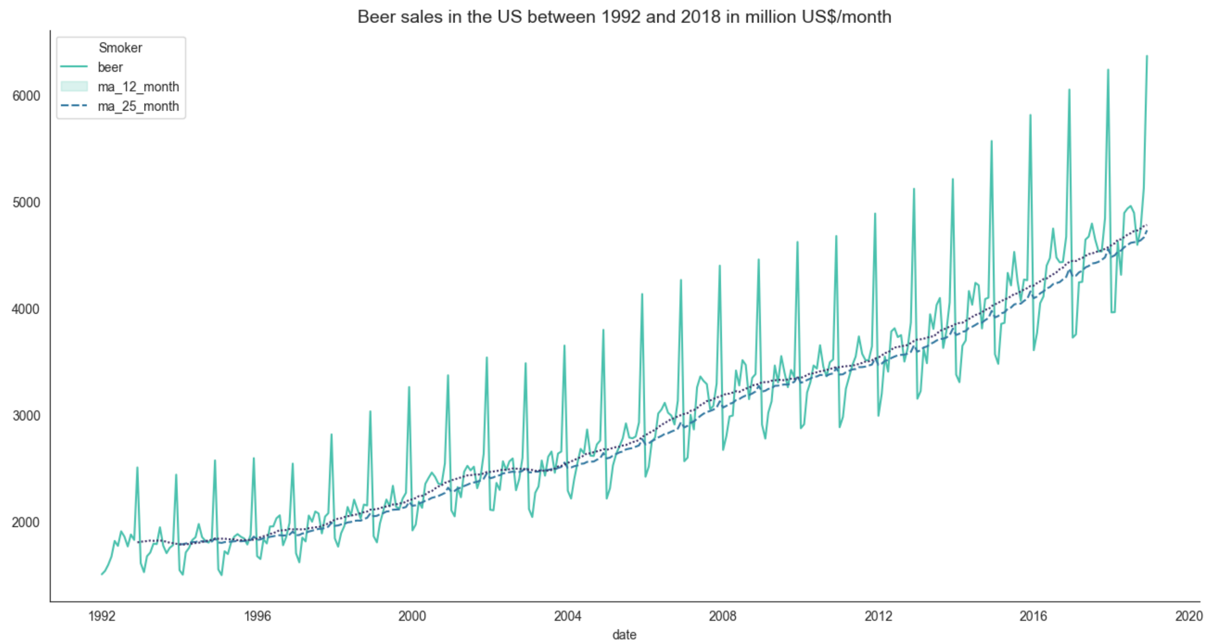


Figure 6 - Trend and Seasonality Analysis Example

The fourth step is related to the prediction generation. Various algorithms can be used and the choice is to be made on the basis of the data. In-sample predictions give the possibility to analyze the goodness of a model and compare different models to choose the most appropriate and then, with the best model, perform predictions out of the test time frame that will be the real forecasted data.

The results of the previous step will define the best algorithm to be applied. A broader view on algorithms will be given in following sections.

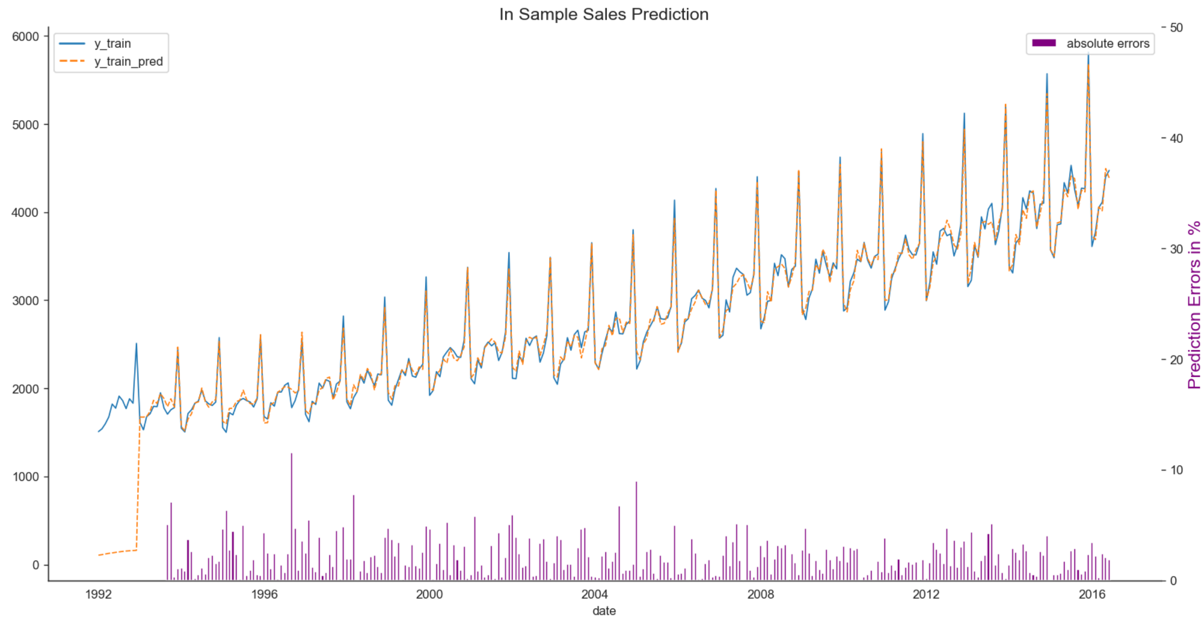


Figure 7 - In-Sample Sales Prediction

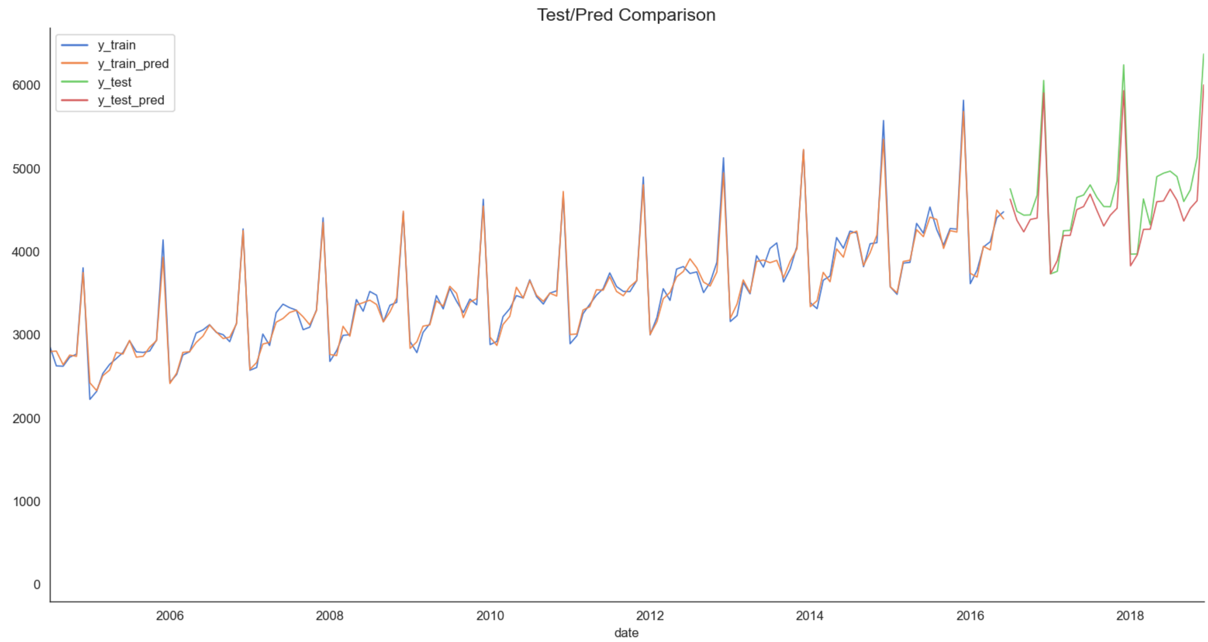


Figure 8 - In-Sample and Prediction Comparisons

3.2. Classification and Algorithms

A forecasting method is a procedure for computing forecasts from present and past values. Forecasting methods may be classified into three types:

1) *Qualitative* or *Judgmental* forecasts based on subjective judgement, intuition, "inside" commercial knowledge, and any other individual private relevant information.
2) *Quantitative* methods which can be subdivided into:
   a) *Univariate* methods where forecasts depend only on present and past values of the single series being forecasted, possibly augmented by a function of time such as a linear trend.

b) *Multivariate* methods where forecasts of a given variable depend, at least partly, on values of one or more additional time series variables, called predictor or explanatory variables. Multivariate forecasts may depend on a multivariate model involving more than one equation if the variables are jointly dependent.

Qualitative forecasting methods comprehend Delphi Method, Consumer Surveys, Bass diffusion model and Panel of Experts. These methods require industry experts which provide opinions based on their knowledge and experience and it can be useful when historical data are not available as in the case of new product launches or radical/innovative changes in an industry.

Quantitative methods are instead based on data patterns and patterns change, in fact the whole data set is analyzed to understand the behave of the variable (which is presumed to be only time-dependent) and the result is a model that contain all the necessary information to create forecasts. In the multivariate field the prediction is driven by one or more predictors that are proved to be influent on the values of the time series and which may have behavior that are easier to predict. An idea could be the case of forecasting sales that are strongly influenced by weather; it’s easier to generate prediction on air temperature because the seasonal component is strong and historical data are massive so the forecast on sales could be driven by the prediction of future values for air temperature in addition to the model generated by the analysis of historical sales.

Another classification of forecasting methods is given by the automation. Automatic methods require no human intervention, and an example can be provided by inventory control: it’s impossible to manually fit a model for each individual time series of sales and then an automatic model is used for the whole products range. Non-automatic methods require human intervention like in economic planning: the model should be built by an analyst to accurately describe relationships between variables and then the forecast can be performed.

After having categorized the different types of forecasts, a list of well-known algorithms is provided to offer practitioners a clear overview of the available prediction methods. This compilation aims to assist practitioners in understanding and choosing the most suitable forecasting technique for their specific needs.

The Moving Average model is the most straightforward approach to time series modelling. This model simply states that the next observation is the mean of m, number of observations, past observations. High m values give more stability to the model but it will be less reactive too. The parameter m defines the number of past time periods that influence the mean to be forecasted. The same weight is given to all past m observations.

$$MA_k = \frac{1}{k} \sum_{i=n-k+1}^{n} p_i$$

Exponential Smoothing model uses a similar logic to Moving Average but a different decreasing weight is assigned to each observation. In other words, if the value of the parameter α ($\alpha \in [0,1]$), that represents the smoothing factor or the weight given to the most recent observation, is higher than 0.5 then less importance is given to further observations. In case of α<0.5 more importance is given to further in the past observations than to more recent ones. If α=0.5 the same weight is attributed to past and more recent observations.

$$s_t = \alpha x_t + (1-\alpha)s_{t-1} = s_{t-1} + \alpha(x_t - s_{t-1})$$

Double Exponential Smoothing or Holt's linear method is used when there is a trend in the time series. In that case, we use this technique which is simply a recursive use of Exponential Smoothing twice. β is the trend smoothing factor and has the same impact of α, it assumes values between 0 and 1 and the higher β values give more importance to the trend nearer in time while lower values of β attribute more weight to further trend values.

Triple Exponential Smoothing or Holt-Winters' method extends Double Exponential Smoothing, by adding a seasonal smoothing factor. Two parameters are added: γ that represents the seasonal smoothing factor and L that represents the length of the season. Triple Exponential Smoothing applies Exponential Smoothing algorithm for three times, this is commonly used when three high frequency signals are to be removed from a time series. Two types of seasonality can be easily identified: "multiplicative" and "additive". For instance, if in every time period 10,000 more units are sold than the period before the seasonality is additive in nature. However, if the units sold are 10% more in the next period than in the current, then the seasonality is multiplicative in nature.

SARIMA or Seasonal Autoregressive Integrated Moving Average represents a class of algorithms which is the combination of simpler techniques to make a complex model that can estimate time series exhibiting non-stationary properties and seasonality.

A key concept necessary to understand SARIMA is Autocorrelation. This concept represents the similarity between observations as a function of the time lag between them. When there is a trend in the data, the autocorrelations for small lags tend to be large and positive, slowly decreasing over time as the lags increase. When there is a seasonality component within the time series, the autocorrelations will be larger for the seasonal lags (at multiples of the seasonal frequency) than for other lags.

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

Figure 9 - Autocorrelation Formula
Source: (Hyndman & Athanasopoulos, 2018)

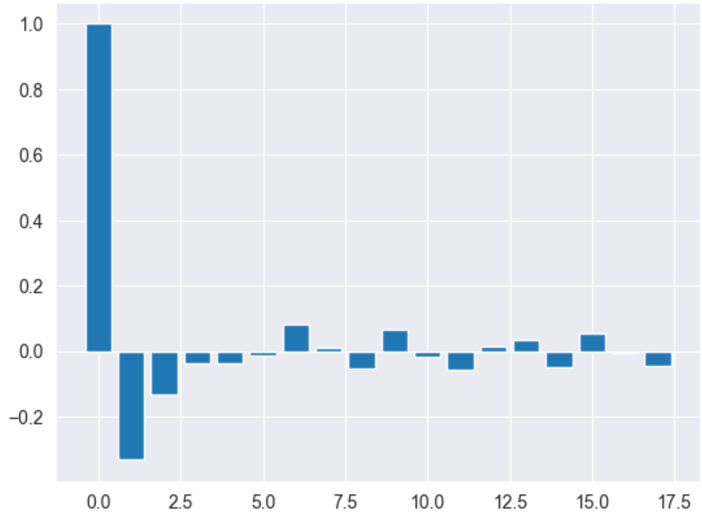


Figure 10 - Autocorrelation Plot

The first algorithm of SARIMA class is the Autoregression model AR(p). This is basically a regression of the time series onto itself that focuses on capturing the relationship between the observed values and their lagged values. It models the dependence on past observations to predict future values and it is assumed that the current value depends on its previous values with some lag. The "p" in AR(p) represents the order of the autoregressive component and it indicates the number of lagged values that are included in the model. The partial autocorrelation plot can provide the value of the parameter p by identifying the lag after which most lags are not significant.

$$y(t) = c + \phi_1 y(t-1) + \phi_2 y(t-2) + \cdots + \phi_p y(t-p) + e(t)$$

Where:

- y(t) represents the observed value at time t.
- c is a constant term or the mean of the time series.
- $\phi_1, \phi_2, \ldots, \phi_p$ are the AR coefficients corresponding to the p lagged values.
- e(t) represents the error term at time t.

Then, the Moving Average model, MA(q), focuses on capturing the relationship between the observed values and the residual errors (or "shocks") from previous time periods. It models the dependence on past errors rather than directly on past observations. The "q" in MA(q) represents the order of the moving average component. The order (q) indicates the number of lagged error terms that are included in the model and is identified by the biggest lag after which other lags are not significant on the autocorrelation plot. Each lagged error term is multiplied by a corresponding parameter known as a "moving average coefficient" or "MA coefficient". These coefficients determine the weights assigned to the past error terms when generating the forecast.

$$y(t) = c + e(t) + \theta_1 e(t-1) + \theta_2 e(t-2) + \cdots + \theta_q e(t-q)$$

Where:

- y(t) represents the observed value at time t.
- c is a constant term or the mean of the time series.
- e(t) represents the error term at time t.
- θ_1, θ_2, ..., θ_q are the MA coefficients corresponding to the q lagged error terms.

The order of Integration, I(d), is then added. It focuses on transforming a non-stationary time series into a stationary one by differencing the data. It models the differenced series to make it suitable for analysis using autoregressive and moving average components. The parameter "d" represents the number of differences required to make the series stationary. The ideal differencing parameter is the least number of differencing steps to achieve stationarity while avoiding over-differencing, which can result in a series that is too noisy or devoid of useful information. The differenced series is then used as the input for the AR and MA components of the SARIMA model.

The final component is Seasonality, S(P, D, Q, s). It captures the systematic variations that repeat over known seasonal periods, such as daily, weekly, monthly, or yearly cycles. It allows the model to account for the predictable patterns and fluctuations observed in the data at specific time intervals. In S(P, D, Q, s) notation:

- P represents the order of the seasonal autoregressive (SAR) component.
- D represents the order of seasonal differencing.
- Q represents the order of the seasonal moving average (SMA) component.
- s represents the length of the seasonal period or the number of time periods in one complete cycle.

The SARIMA model is the combination of all the previous techniques defining then SARIMA(p, d, q, P, D, Q, s). This family models are effective for forecasting time series data that exhibit both non-seasonal and seasonal patterns because parameters can eventually be set to 0 to disregard the modeling of the time series behavior offered by that specific parameter. They provide a flexible and comprehensive framework to capture the complex dynamics and variations in the data, leading to accurate and reliable forecasts. This kind of complex models are generally implemented in applications and software ready-to-use in order to offer a smooth user experience. More skilled users can build forecasting models by themselves from scratch or exploiting available libraries in Python or R.

The set of forecasting algorithms explored have been chosen for comparison reasons: the set of application on which the module, object of this thesis work, has been developed, implement basic forecasting algorithms and then the use of a model belonging to the same family would be appropriate. This has allowed the use of a script, developed in Python, to check the correctness and appropriateness of the obtained results.

3.3. Approaches

Following studies and new fields of expansion of forecasting, businesses often employ a combination of approaches depending on the specific context, data availability, and forecasting requirements for each use case. The adaptation of the various techniques, methods and models to the specific needs creates an approach to the forecasting activity which therefore is defined from the specific requests. These approaches do not belong to the core forecasting activity but are in support of it and are helpful to tailor the results obtained on the objectives set. By incorporating these complementary approaches, businesses can enhance the accuracy and relevance of their forecasts, ultimately aiding in better decision-making and planning.

Top-down forecasting is a method where the forecast is created at an aggregate level and then broken down into lower-level forecasts. This approach is often used when the organization has a clear strategic direction and the upper management's decisions drive the operational activities. The advantage of this method is its simplicity and the ability to maintain a consistent strategic direction across the organization. However, it may overlook the nuances and variations at the lower levels of the organization, leading to potential inaccuracies. In environments with low demand variability and ample production capacity, the top-down strategy may be more effective (Handik, 2005).

On the other hand, bottom-up forecasting involves creating forecasts at a lower level and then aggregating the detailed forecasts to create a higher-level forecast. This method is often used when few lower-level units have a significant impact on the overall performance of the organization. The advantage of this method is its ability to capture the variations and nuances at the lower levels, leading to potentially more accurate forecasts. However, it can be more complex and time-consuming than the top-down approach. In environments with high demand variability and limited production capacity, the bottom-up strategy tends to perform better (Handik, 2005) (Zotteri, Kalchschmidt, & Caniato, 2005).

In both cases an explicative scenario can be imagined: an established firm owning various store spread around one or more geographic area sell a huge variety of products. In such a case many options can be considered for the development of forecasts:

- Maximum level: a prediction will be generated for each product in each store. Depending on the specific case this can be time and resource consuming to a level where drawbacks are higher than benefits, and above all this decision might lead to incorrect or devoid of useful information forecasts. Some products in some store with very little sales won't generate useful information but still require the same amount of resources to generate a forecast. This is the application of a pure bottom-up approach.
- Minimum level: at corporate-level sales data are aggregated without considering any difference between stores and products and then a forecast is generated. This process will require very little resources, being classified as the most efficient, but the "blind" aggregation of sales data lead to loss of information content on lower-level data. Trend and variance of store or product-level sales, if aggregated and summed at higher levels, are smoothed and represent total values.
- Mixed approach: after having analyzed the set of data, some of the time series are forecasted at maximum level while others can be forecasted at a higher level. This mix allows to maintain relevant trend and variance information, where necessary, and to optimize the forecast process for the data that, if aggregated, do not lose any of the characterizing properties.

- Fixed-level approach: a preliminary analysis establishes the best data level to forecast (ex. category-level, geographic area-level) and then data are aggregated up to the predefined level and the forecast is generated. After a prediction is obtained, the aggregated data is spread to lower levels with the help of drivers that can be past sales data or values defined by users.

Collaborative forecasting is an approach to forecasting that regard the active involvement of multiple stakeholders and departments within an organization. In collaborative forecasting, representatives from various departments, such as sales, marketing, operations, finance, and supply-chain, come together to contribute their perspectives and inputs to the forecasting process. They share their domain-specific knowledge, market insights, and data-driven information to collectively develop forecasts that reflect a more comprehensive view of the future demand or other forecasting variables. The collaboration can be asynchronous, too. Among firms, well-defined processes might implement strict steps to be followed which include various corporate figures and require them to approve or modify an ongoing forecast process.

Ensemble forecasting is an approach that combines the forecasts generated by multiple individual models or techniques to produce a more accurate and robust prediction. Instead of relying on a single model, ensemble forecasting leverages the collective wisdom of multiple models to improve the overall forecast accuracy and reliability. The basic principle behind ensemble forecasting is that different models may have strengths and weaknesses in capturing various aspects of the complex relationships and patterns in the data. By aggregating the forecasts from diverse models, the ensemble approach aims to reduce biases, errors, and uncertainties associated with individual models and provide a more comprehensive and reliable prediction.

Scenario-based forecasting involves creating and analyzing multiple scenarios to forecast future outcomes. It recognizes that the future is inherently uncertain and subject to various potential events, trends, or changes in the business environment. Scenario-based forecasting aims to provide insights into different possible futures by exploring a range of plausible scenarios and their potential impacts on the organization. Scenario-based forecasting enables organizations to anticipate and prepare for a range of potential outcomes, reducing uncertainty and enhancing resilience. It helps decision-makers avoid overreliance on a single deterministic forecast.

3.4. Reliability controls

Forecasts undergo various methods and evaluation metrics to assess their accuracy, reliability, and overall performance. Evaluation can be performed on both the model and the predictions themselves.

Model evaluation typically occurs concurrently with the forecast process and serves as a reliability test for the data obtained. This evaluation involves splitting the data into a training set and a test set. The training set is utilized to develop the model, while the test set contains data points that are compared against the initial predictions to determine reliability. If the predicted values closely align with the reference values, the model is considered acceptable. However, if there are significant disparities, adjustments in the model formulation may be necessary.

Forecast evaluation, on the other hand, takes place during a year-end analysis called "Analysis of Variance". The objective is to analyze the estimates made and calculate the absolute or percentage difference between the predicted values and the actual values. Substantial discrepancies between the estimates and the observed values suggest the need to revise the model to achieve more reliable predictions.

These evaluations provide insights into the quality of the model and predictions, enabling organizations to refine their forecasting techniques, enhance accuracy, and make informed decisions. Different kinds of approaches can be used in the evaluation of a forecast and among them are: error metrics, visual inspection, variance analysis, out-of-sample testing, cross-validation and multiple models comparison.

Being a tool that can be adapted to most use cases and which allows comparability between unrelated scenarios, the error metrics will be further examined.

3.4.1. Error Measures

Forecasts are assessed using error metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE). These metrics quantitatively measure the disparity between the forecasted values and the actual values observed after they occur. By evaluating the error metrics, we gain insights into the performance and effectiveness of the forecasting model. Lower error values indicate a superior model that generates forecasts that closely align with the actual observations. This desirable outcome is crucial for obtaining accurate and valuable estimates.

During the analysis of time series data and the development of an appropriate forecasting model, several metrics can be utilized to evaluate and select the best-performing model. These metrics serve as evaluation criteria, guiding the selection process to identify the most suitable model for a given data set. By considering various metrics, such as MAE, MAPE, and RMSE, analysts can comprehensively assess the forecasting accuracy, capture different aspects of forecast errors, and make informed decisions on model selection. The choice of error metric depends on the specific characteristics of the data, the forecasting objectives, and the relative importance of different types of forecast errors. Each metric offers unique insights into the model's performance, highlighting different aspects of accuracy and bias.

MAE provides a straightforward measure of the average forecast error magnitude. It is calculated as the average absolute difference between the values fitted by the model and the observed historical data.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

MAPE expresses the forecast error as a percentage relative to the actual value. It is calculated as the average of the absolute percentage errors and is based on the same values used to obtain MAE but express errors in percentages. This allows broader comparisons as the errors obtained are not linked to a specific unit of measure.

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{|y_i|}$$

On the other hand, RMSE takes into account the squared errors, giving more weight to larger deviations. To obtain the error value it is necessary to compute the norm of residuals for each data point, compute the mean of residuals and calculate the square root of that mean.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \|y_i - \hat{y}_i\|^2}{n}}$$

Akaike Information Criterion, AIC, is a mathematical method used to evaluate how well a model fits the data it was generated from. In the forecast application field it can be used to determine between several possible models which one best fits the data and it can be done by calculating and comparing the different AIC scores of each model. The parameter is calculated as

$$AIC = 2K - 2\ln(L)$$

having K the number of model parameters and L the log-likelihood of the model that, given the data, determines how likely a model is. This metric is useful to understand, between a set of models, which one fit best the data set, hence the absolute value of the metric is not meaningful or comparable with other set of data.

By employing robust evaluation metrics and comparing different models, analysts can make informed decisions, improve forecasting accuracy, and enhance the reliability of future predictions. Continuous evaluation and refinement of forecasting models based on these metrics contribute to the advancement of forecasting techniques and the generation of valuable insights for effective decision-making in various domains.

3.5. Impact on firms' performances

In a business context, firms use various performance measures to assess their overall performance and track progress towards their goals. These performance measures are often tailored to the specific industry, company size, and objectives. Some common performance measures used by firms can be financial, sales and revenue, customer satisfaction and loyalty, operational efficiency and productivity, quality, employee performance and engagement, and innovation and research & development (R&D).

There are many factors that can affect a firm's performance. Some of these factors include the characteristics of the industry in which the firm competes, the firm's position relative to its competitors, the quality or quantity of the firm's resources, CEO remuneration, gender equality at a board/executive level, family/founder ownership, debt leverage, liquidity, capitalization, investment, size, age, location, export performance and managerial efficiency. According to a study published in the International Journal of Production Economics, forecasting techniques alone are not enough to improve companies' performance (Danese & Kalchschmidt, 2011). Companies should also consider several other issues associated with forecasting process management, such as how companies collect and use information on the market, or how the forecast is used in different decision-making processes. The study found that when companies intend to improve cost and delivery performances, they should devote their attention to all the different forecasting variables.

The impact of accurate forecasting on a firm's performance cannot be underrated. It can lead to improved operational efficiency, better financial management, and increased customer satisfaction. For instance, accurate sales forecasting can help firms manage their inventory more effectively, reducing costs and improving customer service. By accurately predicting demand, firms can ensure that they have the right amount of inventory at the right time, thereby reducing stockouts and overstocks. This can lead to significant cost savings and improved customer satisfaction (Lawrence, Goodwin, O'Connor, & Önkal, 2006). By predicting potential financial risks, firms can develop strategies to mitigate these risks, thereby enhancing their financial stability and resilience.

In the current market firms that can be considered as important players in their niche, generally adopt structures that includes several business units or branches for each sector they operate in. The forecasting activity, whether done at higher or lower level, might trigger unexpected benefits on different business units not even involved in the process. The definition of a pathway to be followed establish in individuals a clear way to proceed, avoiding doubts and getting over limitations that can occur in personal relations.

4. Tools

In today's dynamic business landscape, effective planning and accurate forecasting are essential for organizations to navigate uncertainty, make informed decisions, and stay ahead of the competition. Fortunately, the advancements in technology and data analytics have provided businesses with a wealth of tools and software solutions to streamline their planning processes and elevate the accuracy of their forecasts. This chapter aims to explore the wide range of tools available in the market for planning and forecasting, shedding light on their significance and practical applications.

Software solutions have become indispensable in planning and forecasting activities due to several compelling reasons. Firstly, they enable businesses to harness the power of automation and advanced algorithms, allowing for efficient and speedy data analysis, modeling, and scenario planning. This not only saves valuable time and resources but also improves the accuracy and reliability of forecasts. Moreover, software tools facilitate collaboration and information sharing among different teams and departments, fostering a cohesive and integrated approach to planning. By providing a centralized platform for data management and analysis, these tools enhance communication, coordination, and decision-making processes within organizations.

The considered applications are mainly intended to provide typical Customer Performance Management functionalities, detailed below, such as planning, budgeting, closing and scenario analysis and so the need was to find a solution offering these basic requirements and the possibility to develop frameworks that would have allowed the implementation of forecasting workflows. These two pillars and some other technical requirements (multidimensionality management, OLAP cubes, availability to develop custom workflows etc.) drove the software selection and analysis.

4.1. Market situation on software

The market for Customer Performance Management (CPM) software tools presents a plethora of alternatives, providing organizations with a range of options to choose from. CPM encompasses various processes, methodologies, and tools aimed at managing customer relationships, analyzing customer data, and driving business growth through enhanced customer satisfaction, loyalty, and profitability. In light of this, a comprehensive assessment has been conducted to identify the software solution that best aligns with the required characteristics and functionalities. This analysis focuses on the current market landscape of CPM software tools, specifically those that offer capabilities for forecasting and planning. The evaluated tools in this first software selection process encompass IBM Planning Analytics, Jedox, CCH Tagetik, Anaplan, Board, Oracle Hyperion Planning, SAP Business Planning and Consolidation, and Workday. All of these applications provide the basic set of functionalities necessary to implement classic CPM workflows. By examining these industry-leading solutions it is aimed to provide valuable insights into their features, capabilities, and suitability for supporting effective forecasting and planning activities within organizations.

The second step of the selection focused on the identification of applications that satisfies forecasting needs, too. The definition of the dimensions of analysis has been performed before focusing on the analysis itself of each tool and the requirements related to forecasting needs that the software should satisfy are:

- Automatic Algorithm Selection: Automatic choice of the algorithm that returns the most accurate prediction. The application must be able to search among the various algorithms implemented, for the one that offers the greatest forecasting accuracy and then choose that one to train a model.
- Basic Algorithms: Implementation of basic algorithms (e.g. exponential smoothing, moving average, seasonal naïve) that are not composed by the union of several algorithms.
- Complex Algorithms: Implementation of complex algorithms (e.g. ARIMA, SARIMAX, neural networks, machine learning) whose behavior is not always directly understandable and of which it is not always possible to reconstruct the choices and the path taken to obtain the forecast model.
- Upper and Lower Limits: Generation of upper and lower bounds for each forecast point that outline a forecast window and not a single value.
- Seasonality Management: Seasonality recognition and management for forecast generation.
- Multiple Forecast Generation Simultaneously: By selecting the time series of interest, the application must be able to simultaneously generate a forecast for each selected series, e.g. by selecting more than 1 time series (sales relating to multiple different products) the application must simultaneously generate a forecast for each product.
- Outliers Detection and Correction: The application must be able to recognize data that deviate significantly from the seasonal trend and the trend of the data series considered, in order to be able to signal them and possibly propose corrective actions.
- Driver-based forecasting: Ability to define regressors that drive the prediction.
- Multiple Scenarios Management: Possibility to create alternative scenarios (e.g. actual, budget, forecast) aimed at what-if evaluations.
- Integrated Data Spread: Dissemination of forecast data native to the application. The spread can be proportional or relative proportional and the data on which to base the spread must be selectable.
- Confidence Level and Accuracy Metrics: Presence of prediction accuracy estimation metrics that allows to understand the goodness of the model and the resulting data.
- Error Metrics Calculation: The application must provide error metrics associated with the prediction made.

To assess the availability of specific functionalities in a product, it is not feasible to directly observe and test them without acquiring a license. Therefore, a comprehensive analysis was conducted by studying vendors' performance documents, audit documents prepared by third parties (Gartner Magic Quadrant, BARC The Planning Survey) and gathering users' feedback trough availability of comments and opinions on the web or online forums. The findings of this analysis are presented in a summary table in Appendix A, where the presence of technical features in the software is represented by the symbol "x" at the intersection of the respective software and functionality. This summary table provides an overview of the features offered by each product to the public.

			CCH Tagetik	Anaplan	IBM Planning Analytics	Board	Oracle EPM Cloud + Oracle Hyperion Planning + Oracle Analytics Cloud + Oracle Analytics Server + Oracle Analytics for Applications	SAP Business Planning and Consolidation + SAP 4/HANA + SAP Analytics Cloud + SAP BusinessObjects BI	Workday Adaptive Insights	Jedox
Forecast	FORECAST SCORE		2	9	10	8	7	8	2	9

Figure 11 - Software Analysis Summary

To summarize the results obtained in the analysis of the various solutions, a summary table has been drafted with an index describing the suitability of the solution. The score represents the number of dimensions analysis that the solution satisfies, then higher scores indicate a higher relevance to the requirements imposed for the analysis. The score cannot be the sole evaluation metrics, and this visualization has the only purpose of drawing conclusions, since other factor related to specific use cases and technical requirements should necessarily be evaluated.

After carefully analyzing the comparative table in appendix A, it has been determined that IBM Planning Analytics is the suitable tool to further enhance and expand the planning and forecasting activities. This decision is based on the suite's exceptional customizability, which is a crucial factor in developing a unique tool that is not currently available in the market. With IBM Planning Analytics, it will be possible to establish a target definition model, engage in strategic planning, define scenarios for each planning and forecasting horizon, control key performance indicator to set objectives to be reached, perform forecasts based on chosen variables and analyze the variances between the different expected outcomes.

The remaining compared software offer different combinations of features such as in the case of CCH Tagetik that provided roughly every functionality offered from IBM Planning Analytics but the forecast level was completely unsatisfactory. Anaplan, Oracle Hyperion Planning, SAP Business Planning and Consolidation, Board, Workday Adaptive Insights and Jedox, in comparison to the chosen application, miss some key technical specification like integrability and modularity with external tools and do not provide features related to sustainability planning and end-user experience. Furthermore the Oracle product does not provide full financial consolidation while Tagetik and Anaplan do not offer dashboarding capability.

4.2. IBM Planning Analytics: Data Management and Model

The choices made for the analysis of the software selected are related to CPM functionalities (such as planning, closing, reporting, performance analysis, sustainability and ESG), technical specifications, end-user experience and forecast. Regarding CPM functionalities, technical specifications and forecast the sources of information has been vendors' performance documents provided as product presentation and audit documents, prepared by third parties, supplied by Mediamente Consulting. End-user experience has been assessed trough the consultation of online forums, articles and reviews available on the web. As a conclusion to the drafted document, some applications were comparable with the chosen one, while others did not satisfy minimum requirements for the use in productive environments.

The chosen solution offered by IBM is an advanced planning and analytics platform developed to help businesses to streamline their financial planning, budgeting, forecasting, and performance management processes. With its powerful features and capabilities, it empowers organizations to make data-driven decisions, improve their financial performance, and achieve their strategic objectives. It combines the flexibility of spreadsheets with the power and control of a multidimensional database (IBM, s.d.).

At its core, IBM Planning Analytics provides a unified environment where businesses can bring together financial and operational data from various sources, enabling a holistic view of their organization's performance. This integrated approach allows for more accurate and insightful planning and forecasting, leading to improved decision-making and better business outcomes.

One of the key features of IBM Planning Analytics is its ability to create dynamic, driver-based models. These models allow businesses to link financial data to operational drivers, such as sales volume, pricing, or resource utilization. By incorporating these drivers into the planning process, organizations can simulate different scenarios and assess the impact of various factors on their financial performance. This flexibility enables better sensitivity analysis and helps businesses identify opportunities and mitigate risks effectively. On the technical side, the possibility to define on-demand views, requested by users, is made possible by the data-structure used: OLAP cubes.

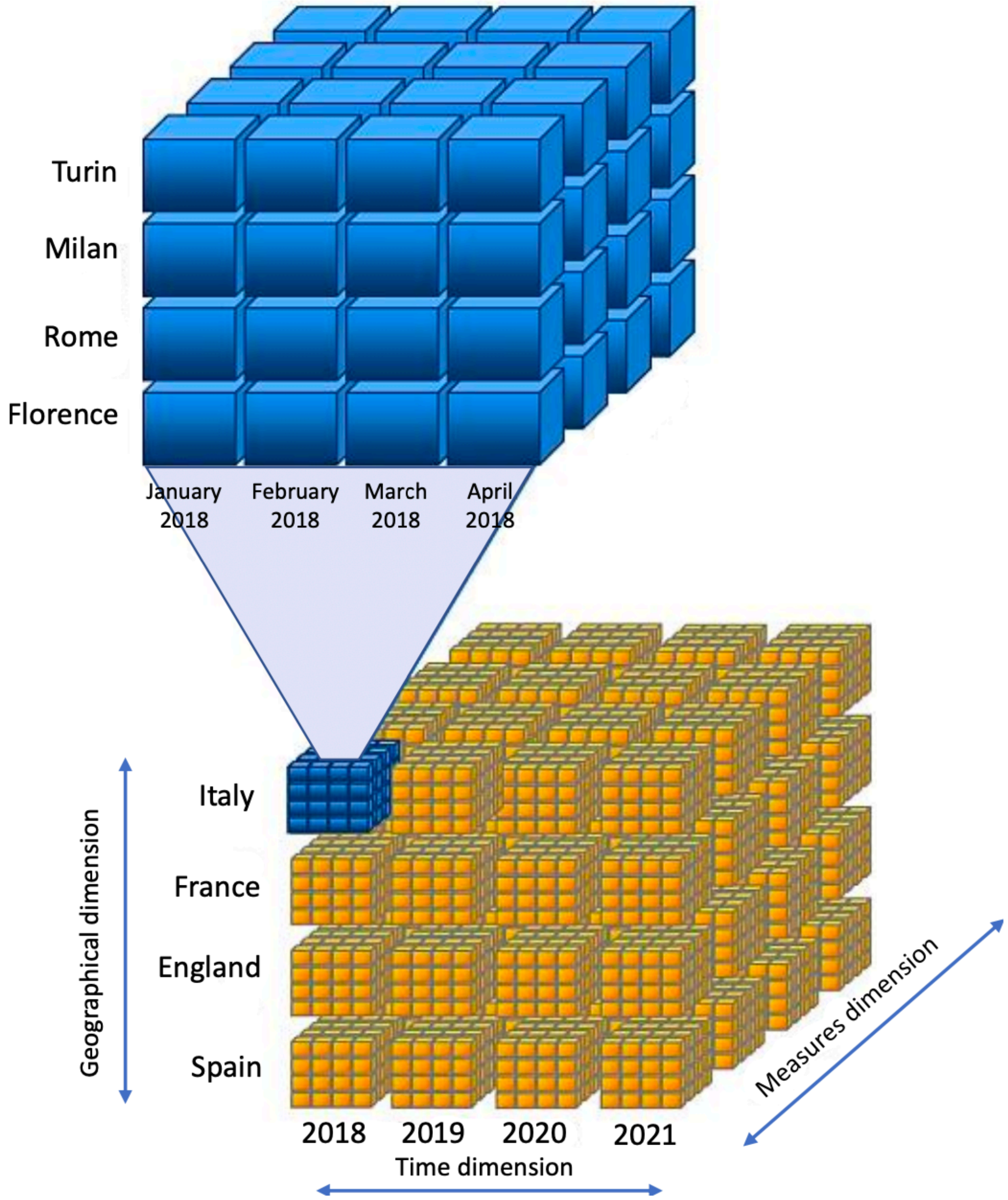


Figure 12 - OLAP cube Data Structure

Online Analytical Processing cubes are a data management technology used to organize and analyze large volumes of multidimensional data. In the context of IBM Planning Analytics, OLAP cubes play a crucial role in enabling efficient and effective planning, budgeting, and forecasting processes. This tool provides a structured framework for storing and manipulating data in a multidimensional format. Unlike traditional relational databases that store data in rows and columns, OLAP cubes organize data hierarchically across multiple dimensions, such as time, geography, products, or customers.

This multidimensional structure allows for more complex and in-depth analysis of data, facilitating better insights and decision-making. IBM Planning Analytics utilizes OLAP cubes to enable slice, dice, drill-down, and aggregate data across different dimensions. This flexibility allows users to explore data from various perspectives and levels of granularity, gaining a comprehensive understanding of their organization's performance. Users can analyze historical data, compare scenarios, and perform what-if analysis to simulate different business situations and assess their impact on financial outcomes. Furthermore, OLAP cubes in IBM Planning Analytics support dynamic calculations and calculations based on user-defined business rules. This capability enables users to create complex financial models and perform calculations on the fly, such as profitability analysis, variance analysis, or allocation of costs and revenues. The calculations can be performed interactively in real-time, providing instant insights and facilitating agile decision-making. Overall, OLAP cubes in IBM Planning Analytics serve as a powerful tool for organizing, analyzing, and manipulating multidimensional data. They enable users to gain a comprehensive view of their organization's performance, perform sophisticated analysis, and drive better planning and forecasting processes. By leveraging OLAP cubes, organizations can unlock the full potential of their data and make data-driven decisions to achieve their strategic objectives.

Another notable capability of IBM Planning Analytics is its robust forecasting functionality. The platform leverages advanced statistical techniques, including time series analysis and predictive modeling, to generate accurate and reliable forecasts. By considering historical data, trends, and external factors, organizations can forecast key metrics such as sales, demand, or revenues with a high level of precision. This enables proactive planning, optimal resource allocation, and effective inventory management.

Furthermore, IBM Planning Analytics offers powerful data visualization and reporting capabilities. The platform allows developers to create interactive dashboards, reports, and scorecards that present financial and operational insights in a visually compelling and easily understandable manner. These visualizations facilitate data exploration, trend analysis, and performance monitoring, enabling stakeholders at all levels of the organization to access and interpret information in real-time. The visualization tool is called Planning Analytics Workspace, or PAW, and act as the front-end of the hidden engine running on IBM Planning Analytics. PAW is a web browser interface to interact with the functionalities offered by the application. Data-management and workflows are defined by developers that design their visualization too and, thanks to this interface, all the rules and restrictions of the model can be respected. Graphs, lists and exploration views are just some examples of the wide arrays of tool that come in help to developers while building applications. Every analysis tool offered can be linked to the underlying data structure and can provide snippets of analysis that might be relevant to understand key concept and performances of a business.

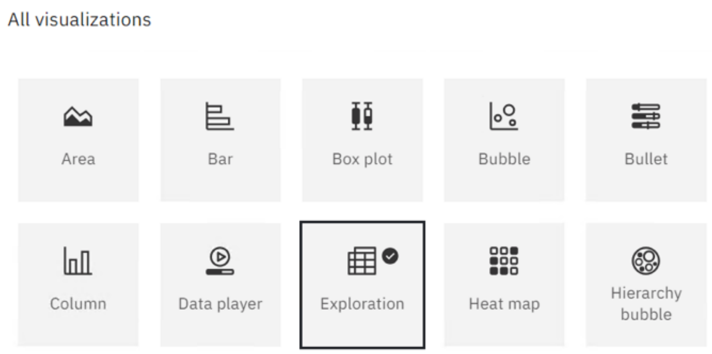


Figure 13 - Available Visualization Tool 1

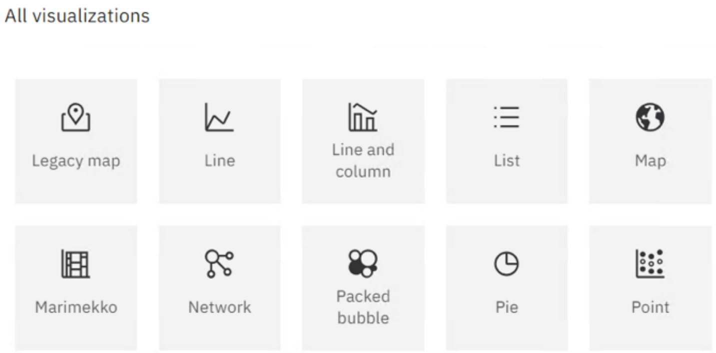


Figure 14 - Available Visualization Tool 2

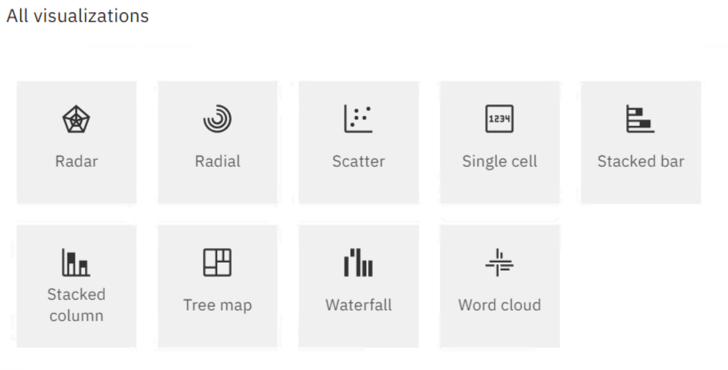


Figure 15 - Available Visualization Tool 3

The analysis made by users are available thanks to "processes", which are sets of instruction written in TM1 language and run by the TM1 server integrated in IBM Planning Analytics. The TM1 processes are run by users to perform specific actions during an ongoing workflow and they perform specific predetermined instruction aimed at data-management and transformation. This sort of instructions can be repeated as many times as needed and they generally involve:

- Data copy, when necessary for data integrity or redundance, some values can be copied in different scenarios or just cloned with the same variables to keep a trustable source as reference.
- Data aggregation or disaggregation, used in case of need to copy data and aggregate or disaggregate values on different levels.
- Data-structure management, involving creation and definition of the structure of dimensions and deletion of elements and structures.
- Data-load and structures and dimensions population.
- Technical maintenance of the entire project.

IBM Planning Analytics also supports collaborative planning and decision-making. The platform provides a secure and collaborative workspace where teams can collaborate on plans, budgets, and forecasts. This fosters cross-functional alignment, improves communication, and ensures that all stakeholders are working towards a shared set of objectives. Moreover, the platform enables workflow automation, approval processes, and version control, enhancing efficiency and governance in the planning process.

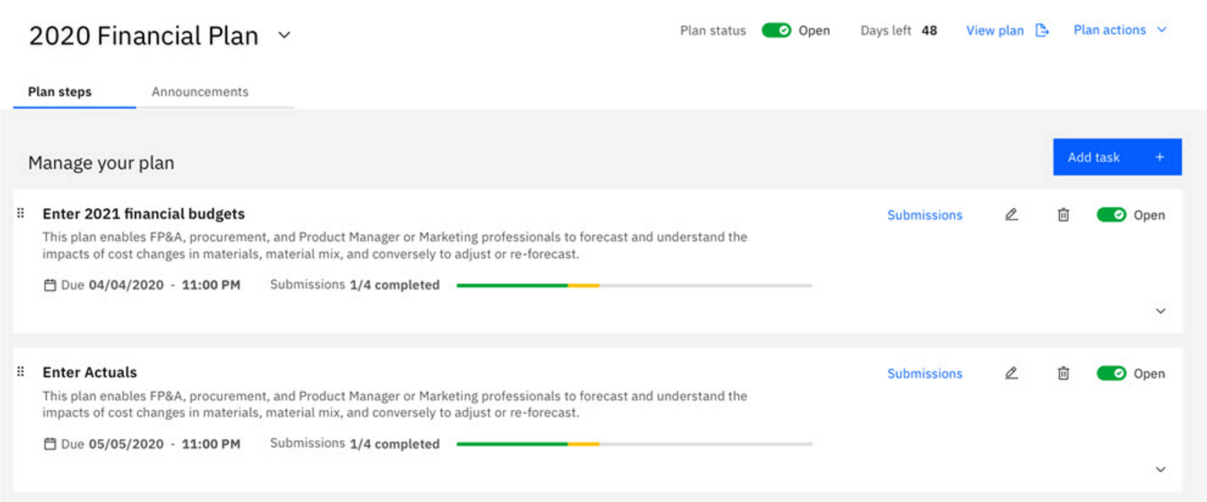


Figure 16 - Submission Example

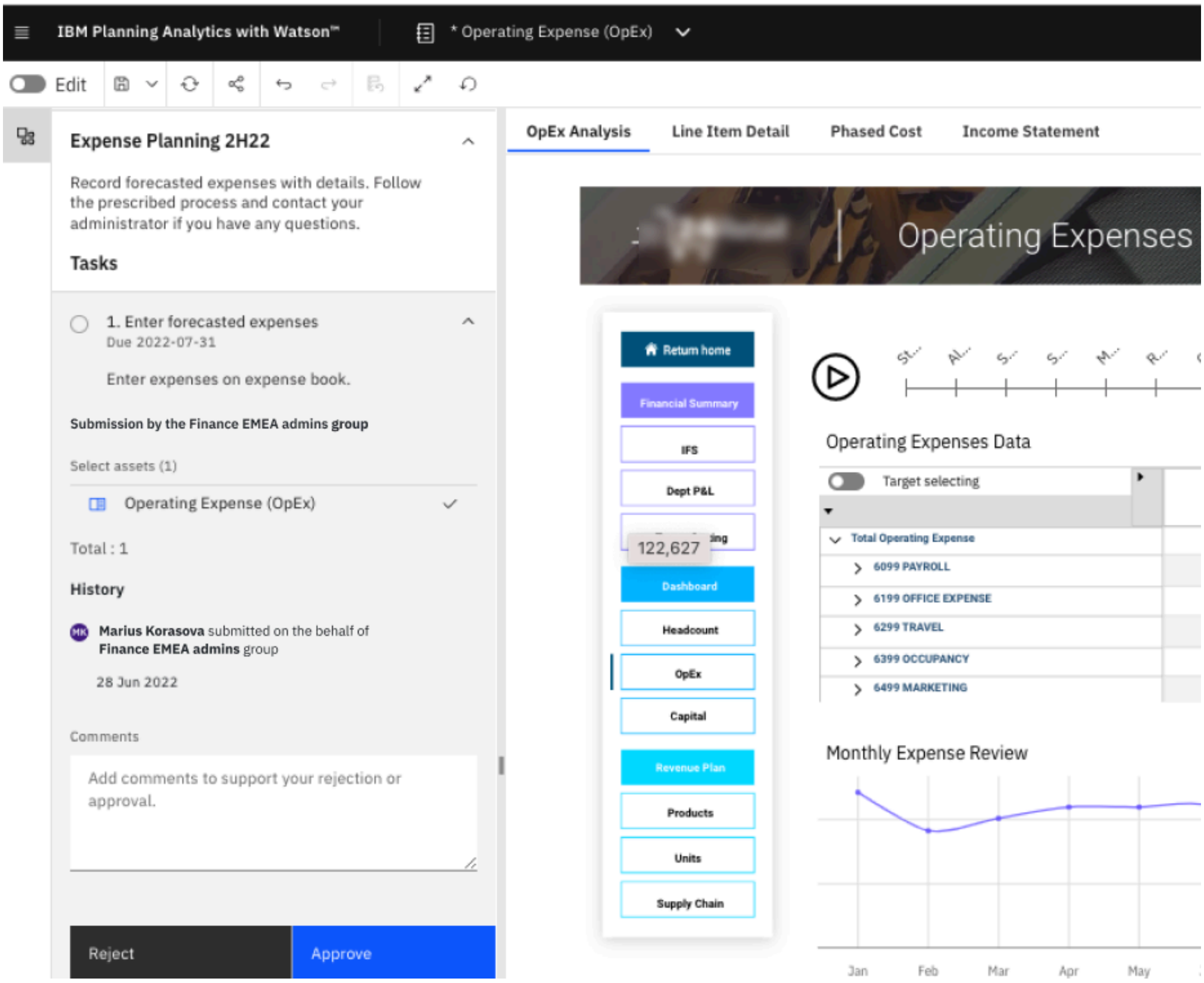


Figure 17 - Approval Process Example

To ensure a smooth implementation and ongoing support, IBM provides comprehensive training, consulting, and customer support services. This ensures that organizations can maximize the value of IBM Planning Analytics and effectively utilize its features to drive business growth and success. Even though IBM itself develop applications and modules to the availability of end-users, generally this kind of interfaces are developed by third-party companies that, in advance, set up dashboards, visualizations, workflows and applications on the IBM Planning Analytics tools in a such a way that end-users will receive a custom product tailored on their specific business needs.

By harnessing the power of data and analytics, organizations can gain a competitive edge in today's dynamic business landscape and drive sustainable growth and profitability. The high-level customizability of IBM Planning Analytics, presenting himself as a “white canvas”, drive development towards specific implementation of application that satisfies every users need and allows users to further customize, in complete autonomy, every further aspect that might arise in time.

5. Case study

The development of the application has been followed by an assessment phase. The results obtained from the forecasting algorithms of the platform chose has been compared to results obtained from a custom implementation in Python of the SARIMA set of algorithms. This analysis allowed to assess the goodness of data obtained and whether or not they were reasonable in comparison to a surely correct result obtained with code designed to perform forecasts.

Mediamente Consulting allowed the possibility to test the solution thanks to the availability of real use-case data coming from one of their clients in the large-organized distribution sector. Conducting a test on a real-world use-case required, as expected, adjustments to the model to completely integrate the predeveloped standard application and the application pertaining to the real use-case.

In the implementation of the module a choice has been made in relation to the specific large-scale distribution case, in the approach to the planning process. When performing planning processes, objectives to be reached in future can be set through Gross Sales or Margins. By defining one of the two metrics, a business can, in cascade, let the objectives set also fall on other KPIs, planning then the entire set of metrics necessary for the various aspects of the business. In large-scale distribution planning, generally, objectives are set with Sales and, in the considered case, this has been respected. Future goals are set in Sales term and the same happen with forecasts that are predicted using historical sales and then produce values for the Sales metric. Ad-hoc processes are implemented to transform, at an appropriate moment in the planning and forecasting workflow, the results obtained and set in term of Sales to margins, quantities, and prices. The two approaches are interchangeable and then shifting from one to another is easy and feasible even with the solution already implemented in production environments.

5.1. The firm

Mediamente Consulting is a player in the market of solutions offered on IBM Planning Analytics and develops application of wide range and use for different kinds of users. Having managed many differentiated projects, built from nothing or just from internally developed basis, capabilities and knowledge of the developing platform are an asset for the company. One of the strengths of their developments is the implementation of a standard set of modules that can be used over any project and be adapted to the specific situation and requests of a client. This flexibility allows to offer services to many different clients ranging from medium to big corporations and involving sectors such as Retail, Food & Beverage, Manufacturing, Fashion, Banking, Insurance. The use made from Mediamente Consulting of IBM's software contemplate the addition of further features to pre-developed modules and applications. By defining a standard they can build on top of it, adding functionalities and customizing solutions for specific requests of clients. The forecast application can be added to pre-existing solutions and so its development and integration with prior developments is free of impediments.

The development of the application followed a need arose from Mediamente Consulting: trends in the market served showed the growing willingness to implement forecasts into the planning process and then the necessity to provide clients with an all-around solution that could satisfy them. With these assumptions the company decided to develop a general implementation that could, in future, be adapted to specific business requirements. The tool developed focused then on the integration of the forecasting activity into the yearly planning and budgeting processes typical of Customer Performance Management, to optimize time required and concentrate the managerial efforts in a dedicated timeframe.

5.2. Minimum Viable Product

The aim, while developing a Minimum Viable Product, was to give final users the possibility to:

- Define forecast scenarios from the availability of actual data, this implementation embeds the analysis of past data in order to draw models that depict the trend of key metrics to be forecasted.
- Constantly interact with budget scenarios by creating and adjusting values in relation to customized forecast that can be performed up to a monthly basis.
- Perform budgeting activity with different sets of starting data, create many what-if scenarios that can be analyzed to discover differences between different type of data (planned, forecasted, actual) and the budget object of analysis.

Then, being clear the scope of the MVP, the proposed solution allows to:

- Gather data from different sources (CSV, XLSX, relational databases).
- Organize and prepare data to be managed in the structures offered by IBM's application.
- Perform forecasts using embedded algorithms and revise the results obtained with the possibility to:create different scenarios and customize them, compare side by side the different available scenarios, feed new type of scenario and spread results, obtained at higher levels, to maximum dimension detail.
- Prepare a plan for future timeframes, setting goals and KPIs at different levels of aggregation.
- Analyze variances with past years at different levels of aggregation.
- Modify the created budget scenario.

The application development required a first part in which processes of data management has been handled. This phase involved the definition of procedures to perform critical operations that represent core activities, such as the delicate spread processes. In order to obtain reasonable performances, data spread should be implemented in such a way that large data sets do not cause performance issues and then a specific technique has been used.

After having implemented the processes related to the management of data and workflow, the second phase concentrated on the user interfaces. What users see and interact with, is a web-based browser page on which developers set up dashboards, books and graphs to define a path to be followed in its order and carry out the intended activity. The interface tool is called Planning Analytics Workspace (PAW) and the third phase has been related to the creation of the link between PAW and the processes defined in the first phase. Starting with actual values data load, the implementation connected processes with visual tools to run processes and set variables for these last ones or, in other cases, to analyze results obtained and intervene on the data to modify them according to criteria set by the user.

5.3. Forecast's technical aspects

The analysis dimensions involved in the use-case were time, product, store, scenario and various measures and each dimension could be aggregated at different levels depending on the specific needs. The time dimension was at week detail and this choice has been taken because more detailed data would not have added any benefit since the combination of products and stores at a daily level introduce a lot of data without relevant values. Aggregating data by week allowed to have as many significant cells as possible. Higher aggregation levels such as months, quarters and years are offered, by design, from the application on which the model has been developed. Hierarchically defining elements of a dimension allows the application to "understand" which elements should be calculated as sum of lower-level elements and which are leaf-level elements. Similarly, the store dimension was grouped by brands and the lowest level elements represented single stores. Lastly, the product dimension contained leaf-level elements representing single products which were grouped by categories, sectors, and departments. This grouping triggered a forecasting level by category in light of the following reasons:

- Introduce much better performances than single product. The single-product single-store forecast introduces lots of data with just few of them effectively relevant, then grouping at least on one of the two dimensions, eventually reducing the number of cells to be considered ignoring null values, determines better performances.
- Better figures understandings by managers. This kind of activities are generally assigned to managers, being them product, category, sector, or department managers, and to operate on data with higher aggregation levels enables better general understandings of the matter being analyzed.
- Improved task assignment. Those responsible for one or more categories can see assigned the control, over plans and forecasts, for one or more categories and the definition of hierarchies adapts to hierarchical structure of organizations. Higher-level managers can control the progress of categories/sectors managers.

The dimension of measures included metrics like sales, costs, margins, units, unitary margins, unitary costs and unitary prices. Some elements were calculated from other values (e.g. margin resulting from units, unitary prices and unitary costs) while others were directly related to data-sources.

A functionality not developed, but included in the module, is the forecast tool offered by IBM Planning Analytics itself. To enable the use of the tool some preparation was needed in the way data are presented and structured so that the application can identify which dimension represents time and which are the metrics to be forecasted. The tool offers an automatic seasonality detection but, at the same time, the possibility to set it manually, too. The built-in algorithms are all variations of Simple Exponential Smoothing, Holt's linear method, and Holt-Winters' method depending on the presence of trend and seasonality.

Trend component	Seasonal component		
	None	Additive	Multiplicative
None	Simple exponential smoothing	Simple exponential smoothing	
Additive	Holt’s linear method	Additive Holt-Winters’ method	Multiplicative Holt-Winters’ method
Additive damped	Additive damped trend method	Additive Holt-Winters’ method with damped trend	Multiplicative Holt-Winters’ method with damped trend

Figure 18 - Built-in Available Algorithms

Before running the forecast process, a preview is offered with plenty of detail on the time series considered. The application performs a first analysis aimed at establishing the applicability of the forecast algorithms, eventually displaying error messages in case of data not compliant with requests and showing results in a window called “Forecast Preview”.

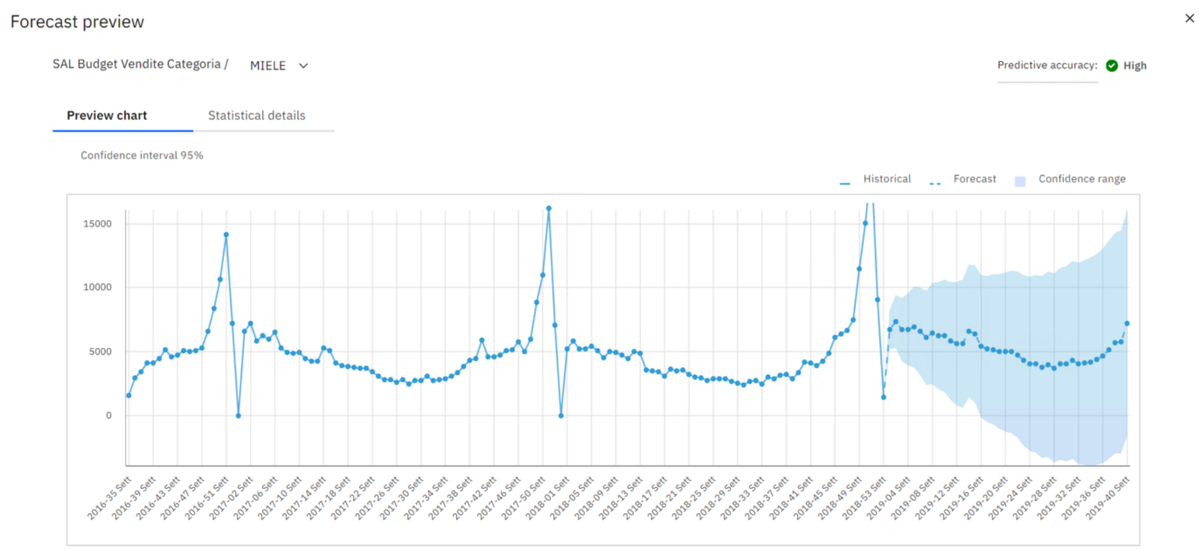


Figure 19 - Forecast Preview

This window offers the answer to the applicability of the forecast, by classifying the results obtained in term of accuracy: if the forecast performed returned valuable results, which respected trend or seasonality or both, then the accuracy is “High” and the model used to generate test predictions can be used to forecast. When the accuracy is defined as “Low” some issues might be influencing the process like incorrect data formatting, too few historical data, or data set is not being understood in its meaning by the model. Values between these two extremes are classified as “Medium” forecast accuracy. Statistical details, also provided, are divided in:

- The type of forecasting method used.
- The accuracy details, which include the number that is used to derive the overall accuracy of the forecast.
- Parameters.

- Information about the forecasting method that was used along with the detected trend and seasonality components.
- Trend and seasonality strength and seasonality period. Period indicates information that relates to the cycles detected in the data along the time series.

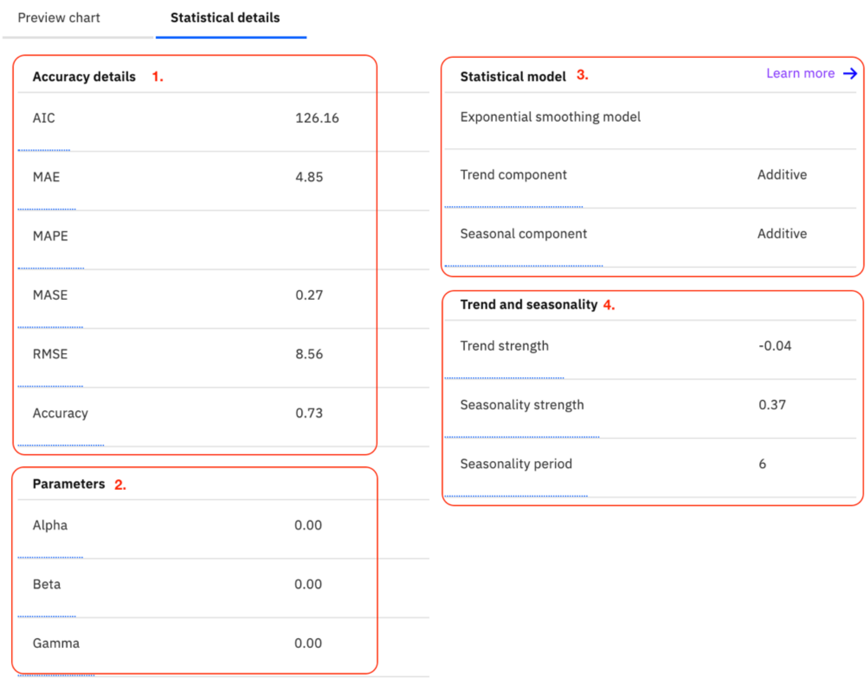


Figure 20 - Statistical Detail Visualization

Accuracy details are showed in “Statistical Detail” tab and they deep dive into an analysis of the model selected to forecast values and extend the accuracy metrics to have a broader understanding of what happened behind the generation of the prediction. The code, that implements the algorithms used to forecast, is closed-source then only estimations and official guides can be used to understand the logic. AIC, MAE, MAPE, MASE, RMSE, and MSE are the accuracy and error metrics used to evaluate results. The section “Parameters” deepens detected seasonal period and estimates for other parameters that are used in the selected exponential smoothing. α, β, and γ, are the level state, trend, and seasonality smoothing factors respectively, while φ represent the damping coefficient. These values are technical and provided for further analysis and assessments.

While verifying and controlling the choices made by the forecasting tool, the user has the possibility to fully customize and control the process: before launching the preview or the forecast process, a menu allows for customization of different parameters (source of time series, forecast horizon, seasonality of the series, outlier detection, confidence intervals, and ignoring specifics data-points) that can impact the results. While being a closed-source model, the customizability allows to properly interact with the tool and, if wanted or necessary, to fine tune specific parameters so that they can be adapted to practical scenarios.

5.4. The proposed solution

A complete description of the developed application will be presented to offer clear understanding, with the support of images for clarity. Four sections characterize the application:

- Overview
- Forecast Definition
- Budget Plan
- Variance Analysis

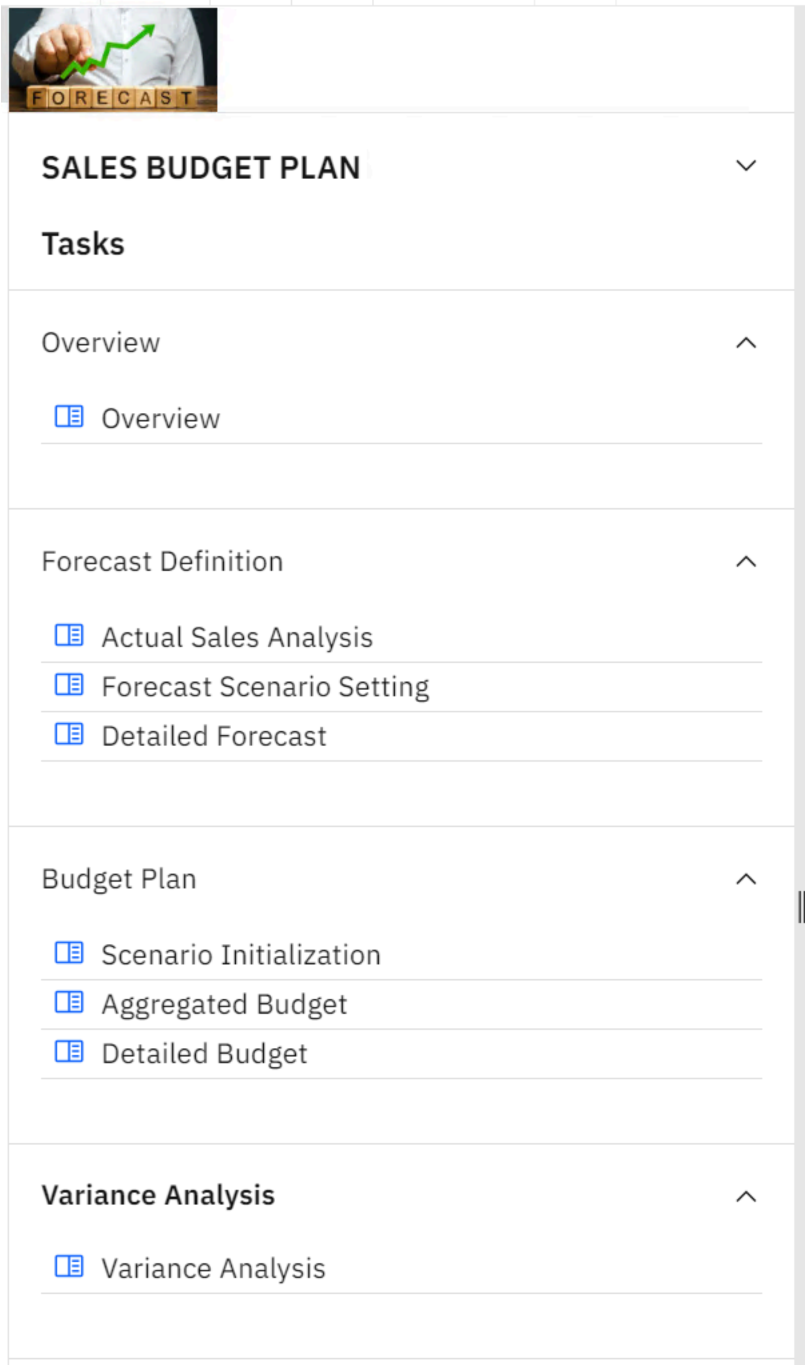


Figure 21 - Application Overview

Every section contains a collection of “books” and every book comprehends different dashboards with which the user will interact. The order of sections and books has been defined to direct users into a series of steps to be followed. If the flow is respected the application will not display any error related to data-structures or the logic of the process.

5.4.1. Overview

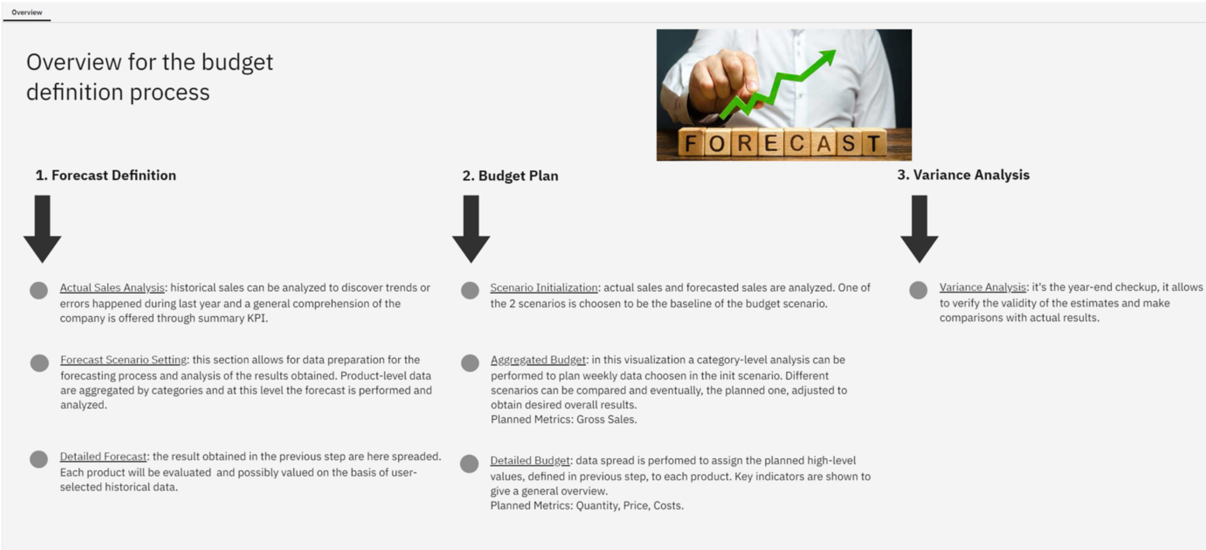


Figure 22 - Overview visualization

This section offers a brief description of the whole process, explaining steps and activities involved in each of them. This can be useful to start building a general comprehension of the model in less expert users and then help them to understand the sequentiality of actions that should lead to the final result. Single books are presented and the immediate accessibility of this section, being it always available on the left side of each visualization, allow for multiple uses over time to better elucidate complex concepts.

5.4.2. Forecast Definition

The forecasting process starts with a data-load. This is necessary to make available historical data into the data structures of the application. As shown in the figure below, the button “Load Actual Sales” triggers a process that will request the selection of custom parameters and then will load data into predefined structures. The integrative comment on the side, outlines the limit to upload one product category for each run of the process, but this is related to the specific use-case and further implementations may overcome this limitation to enable the upload of multiple categories at the same time.

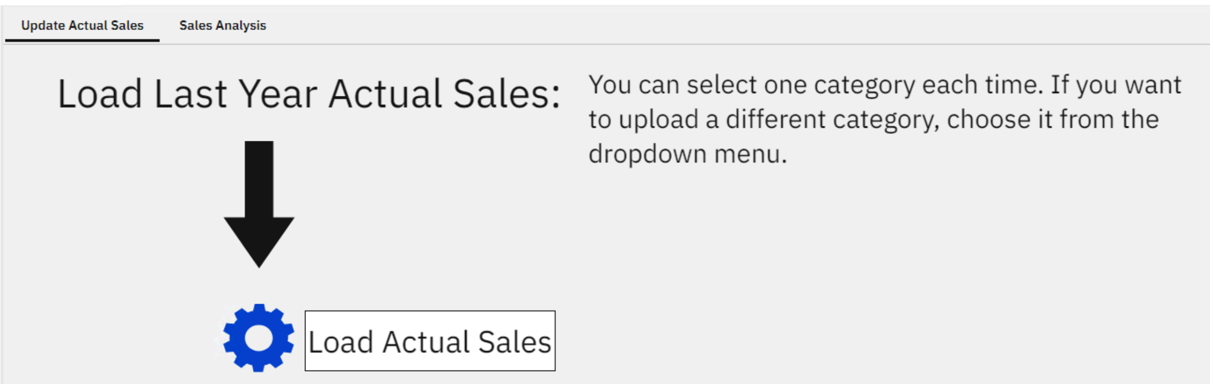


Figure 23 - Forecast Data-load

The following dashboard gives an overview on past (actual) data.

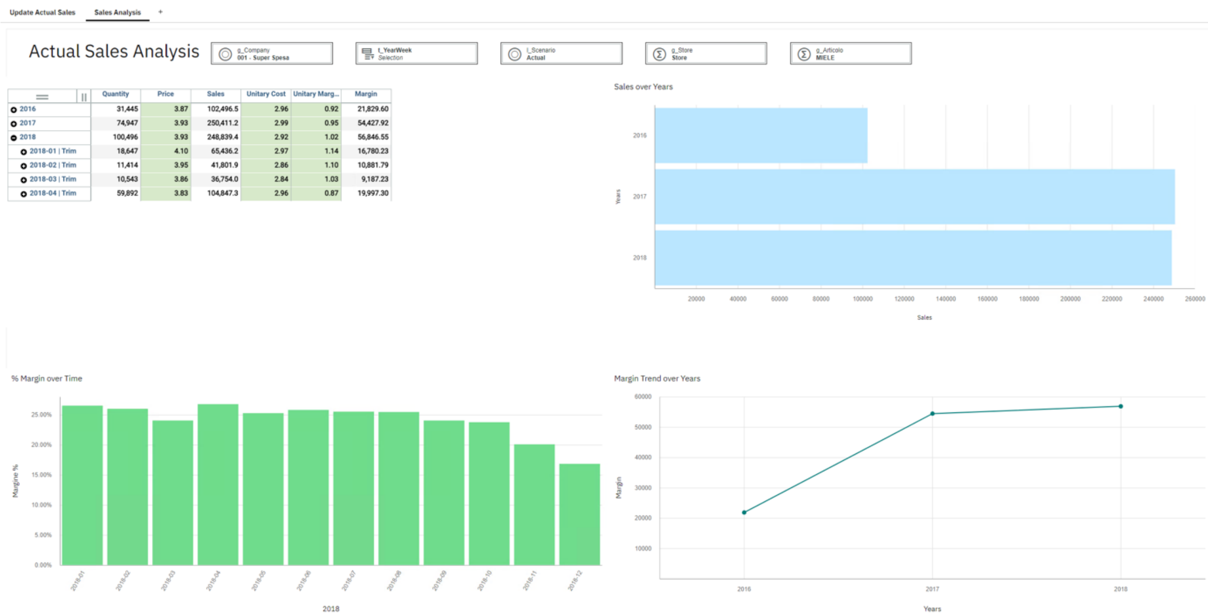


Figure 24 - Sales Analysis

On the upper side of the image the rectangular white boxes are filters, they allow to select specific elements in case of deeper analysis. The filters mirror the dimensions of the model and the selection of an element, or a subset of elements, from one of the filters will update graphs and values in the visualization to update data with the selection made.

	Quantity	Price	Sales	Unitary Cost	Unitary Marg...	Margin
2016	31,445	3.87	102,496.5	2.96	0.92	21,829.60
2017	74,947	3.93	250,411.2	2.99	0.95	54,427.92
2018	100,496	3.93	248,839.4	2.92	1.02	56,846.55
2018-01 | Trim	18,647	4.10	65,436.2	2.97	1.14	16,780.23
2018-02 | Trim	11,414	3.95	41,801.9	2.86	1.10	10,881.79
2018-03 | Trim	10,543	3.86	36,754.0	2.84	1.03	9,187.23
2018-04 | Trim	59,892	3.83	104,847.3	2.96	0.87	19,997.30
2018-10	6,908	3.87	21,329.1	2.88	1.00	5,068.25
2018-11	16,648	3.86	26,655.7	3.00	0.86	5,354.44
2018-12	36,336	3.76	56,862.5	3.00	0.76	9,574.61
2018-49 Sett	9,096	3.81	11,485.1	3.01	0.81	1,943.27
2018-50 Sett	9,352	3.85	15,046.5	3.07	0.79	2,530.58
2018-51 Sett	12,171	3.84	19,843.5	3.09	0.76	3,236.73
2018-52 Sett	5,035	3.70	9,046.8	2.96	0.75	1,594.12
2018-53 Sett	682	3.34	1,440.5	2.68	0.66	269.91

Figure 25 - Exploration View Actual Metrics

The previous figure zooms into the "Exploration view", a particular tool that offers tabular view of specific portions of data. By selecting dimensions and placing them in columns or rows, the information is displayed accordingly. As clearly understandable from the figure, depending on the structure of the dimension, elements can be drilled-down, meaning that higher hierarchy elements can be split in their lower-level components in such a way that

the values that compose the aggregated value can be more specifically analyzed. This functionality can be exploited for dimensions placed on rows or columns only, while dimensions that compose the context can define no more than one element (aggregated or leaf level).
Provided that cells can be assigned to particular, developer-selected, colors or conditional formatting (creating an exception to the next sentence), the usual difference in colors is due to cells content: green cells are calculated, this includes calculations coming from other elements on the same row (e.g. Sales calculated from values of quantity and price, all on the same row) and hierarchical calculations (e.g. the case of Price in which an aggregation to higher hierarchical levels by summing the underlying ones would not have any practical meaning and then aggregated levels are the mean values of the underlying cells), and grey/white cells represents real static values, modifying this values will modify in cascade the calculated cells.

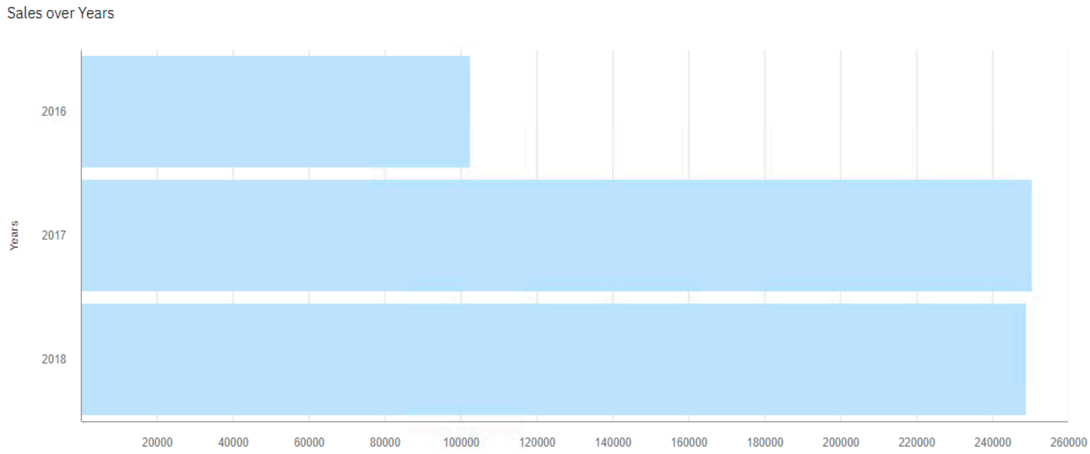


Figure 26 - Sales over Years

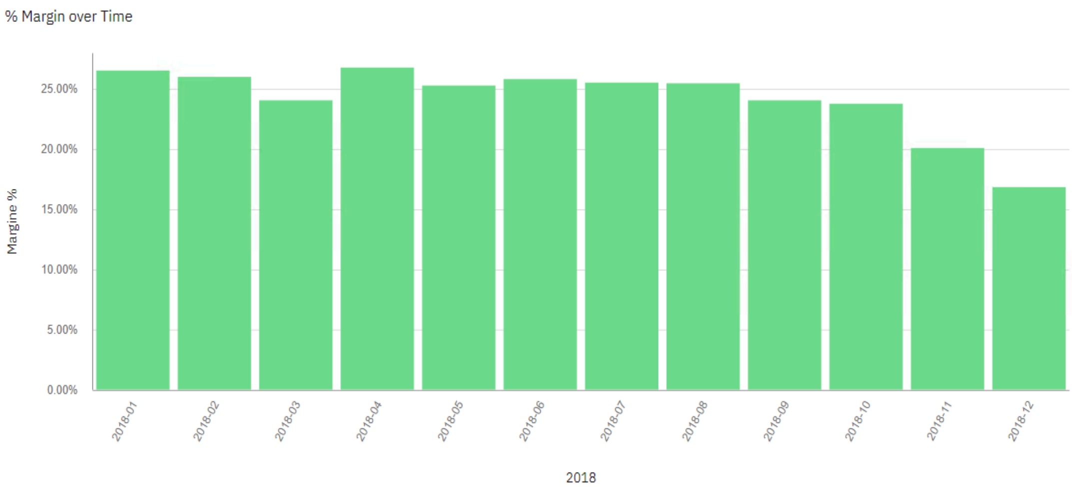


Figure 27 - % Margin trend

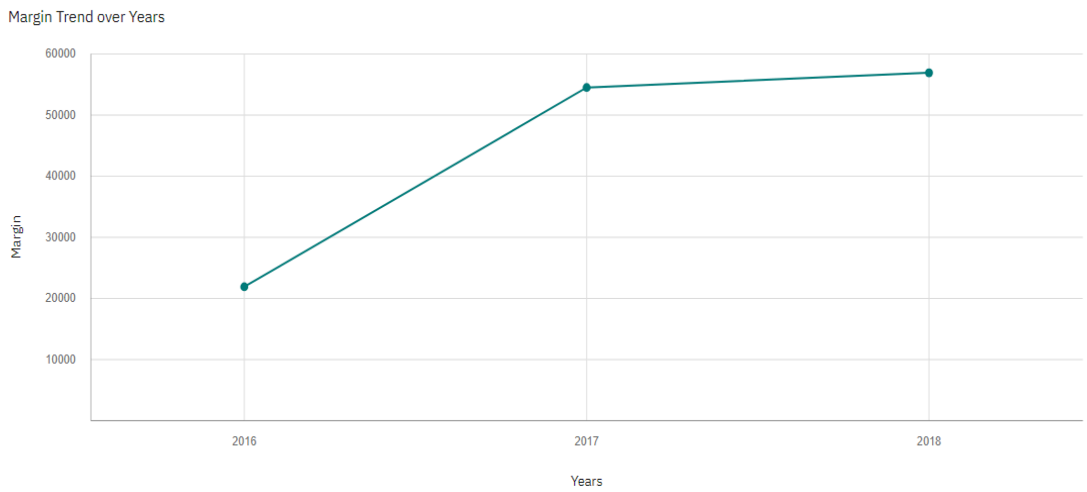


Figure 28 - Total Margin trend

The last images zoom the first overall visualization. These graphs are built to provide an overview on past years and the last actuals information. Sales are aggregated to understand the general trend of the business and have an idea of what to expect from future figures. Margins, in Figure 27, are instead analyzed as monthly percentages of the last available year in order to capture eventual seasonality on data and identify key factors that can influence profitability of a company. A year-based visualization of margins is deemed useful for the purpose of general assessment of the company's performance. The analysis of sales and margins can provide valuable insights into a company's financial performance, for example in identifying potential issues related to cost control. It enables the detection of situations where a company may be generating sufficient sales but experiencing low profit margins. This underlines the significance of going beyond a superficial examination of graphs and numbers. Instead, a proactive understanding of the underlying meaning of the data becomes essential in drawing accurate conclusions, even when explicit information may not be readily apparent. This perspective emphasizes the importance of interpreting data holistically and employing analytical techniques to uncover hidden patterns and relationships that contribute to a comprehensive understanding of a company's financial dynamics. By delving deeper into the data and embracing a proactive mindset, decision-makers can gain valuable insights that aid in identifying and addressing underlying issues, ultimately fostering improved financial performance and sustainable business growth.

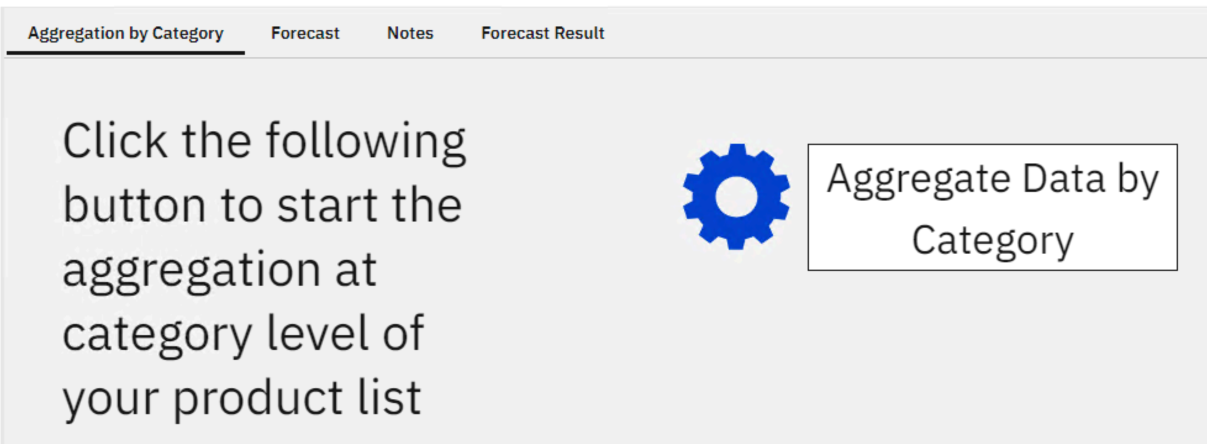


Figure 29 - Data Aggregation

The first operative part of the process consists in the data aggregation. From this point on, the sole metric that will be considered to generate forecasts and set up plans will be Sales. This choice is due to the nature of the business case selected and, as stated before, may vary from firm to firm. Further in the process all metrics will be updated. By selecting the appropriate button, product-level sales are aggregated by summing them to category-level. The overall data structure remains the same but instead of considering single product items in the product dimension, data is aggregated and the maximum detail is represented by category sales.

The following visualization offers an overview on the forecast process.

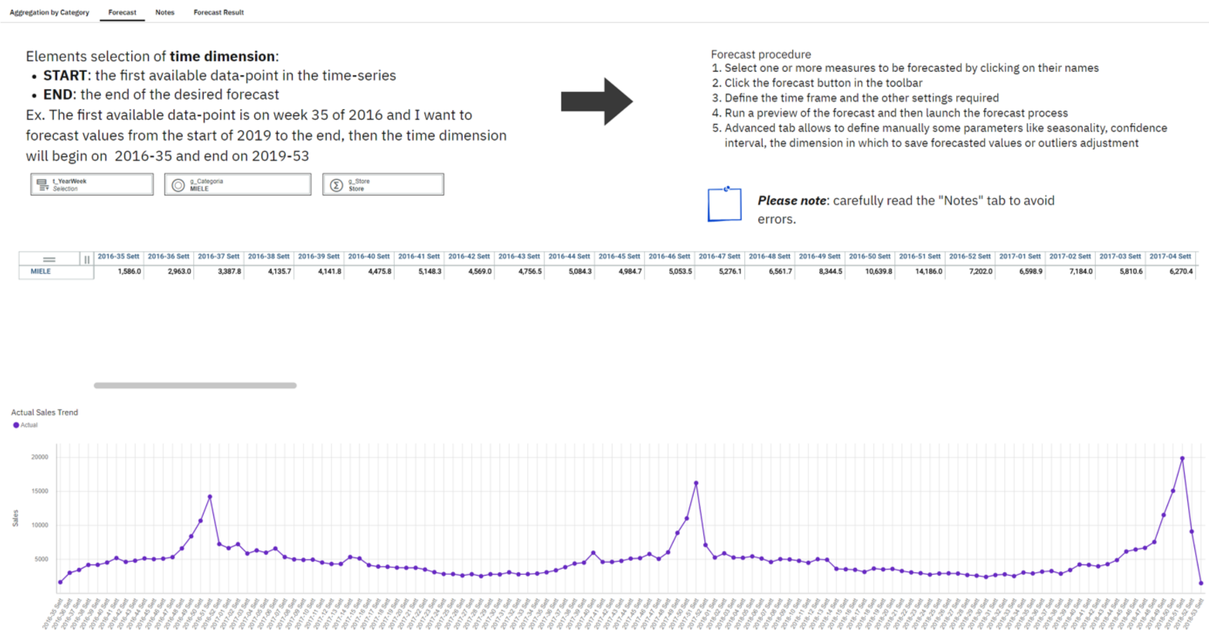


Figure 30 - Forecast Setup

Filters are available to select the proper elements of a dimension. It is necessary to pay attention to the time dimension since it is one of the most important parameters while forecasting. All available historical time-points should be included, avoiding null o zero values which do not provide any information, and the desired forecasting period should be added. Then, up to 25 time series can be selected and in the considered case, it means that up to 25 categories can be identified to be forecasted exploiting their historical data. A dedicated “Note” tab and comments in the dashboard give the user a clear understanding of the steps that must be taken and the details to which refer attention. By selecting the time series of interest, the forecast button will appear on the toolbar and after clicking on it, a window will be displayed.

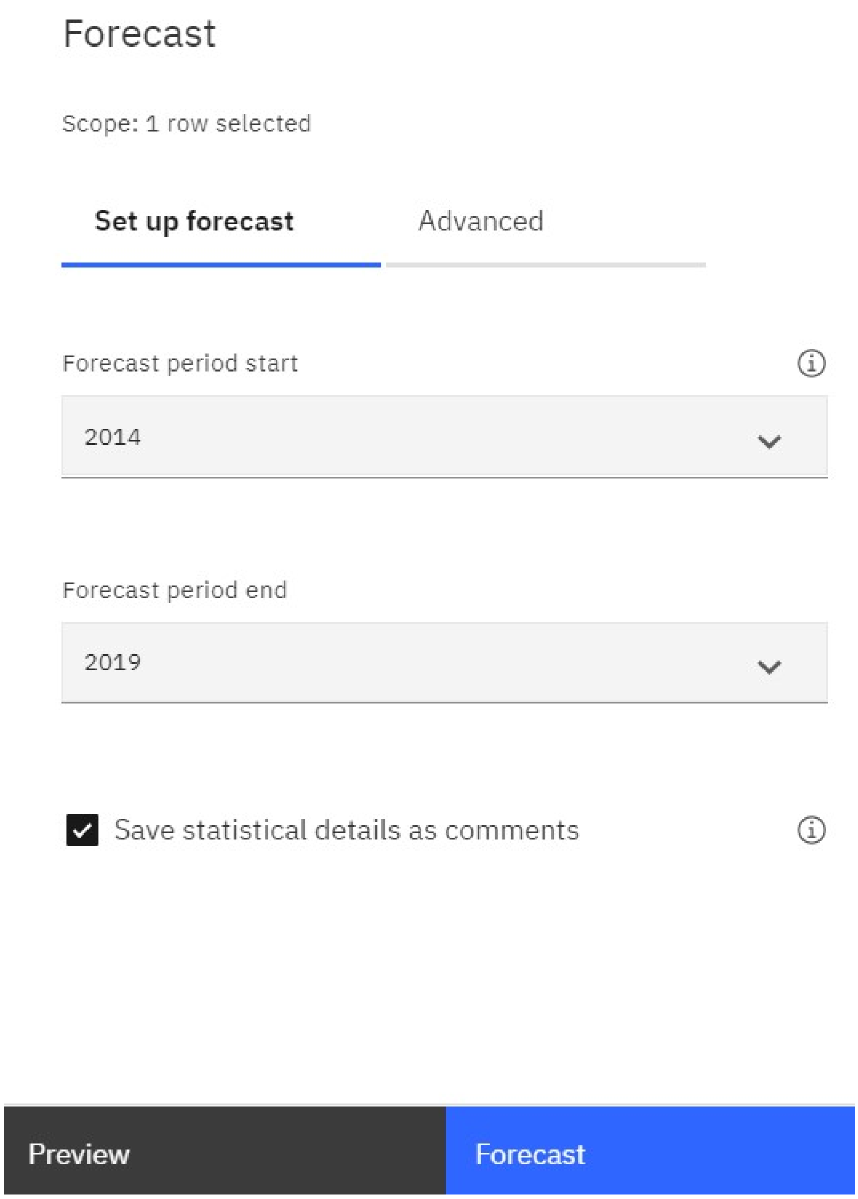


Figure 31 - Forecast settings

Figure 31 represents the first tab involved in forecast's variables definition: the time horizon should be defined and it is important to note that for each requested data point there should be at least 3 historical data points. This restriction implicates that past data should be at least 3 times longer than the requested prediction.

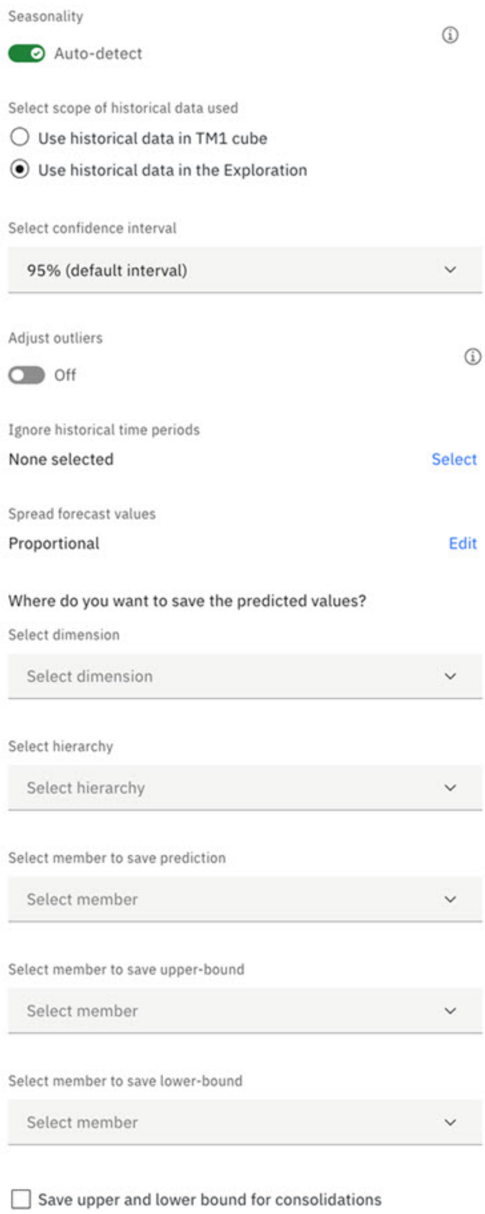


Figure 32 - Advanced Settings

The Advanced tab offers selection for different parameters:

- Seasonality can be set to be detected automatically or manually predefined. Automatic detection is generally capable in understanding trends but a manual set might be choose as result of a deep analysis of the time series.
- Source data to be used for forecasting purposes might be found inside the data-structure or the exploration view visualized, in case of manually entered variations.
- 90%, 95% and 99% are the available choice for confidence intervals.
- Outliers detection regards the analysis of the time series to find possible outstand values that can mislead the comprehension of the time series. When the parameter is enabled, outliers found in the time series are adjusted to more “appropriate” values.
- Dimensions in which to save the real predictions, upper bounds and lower bounds are defined with lower selectors.

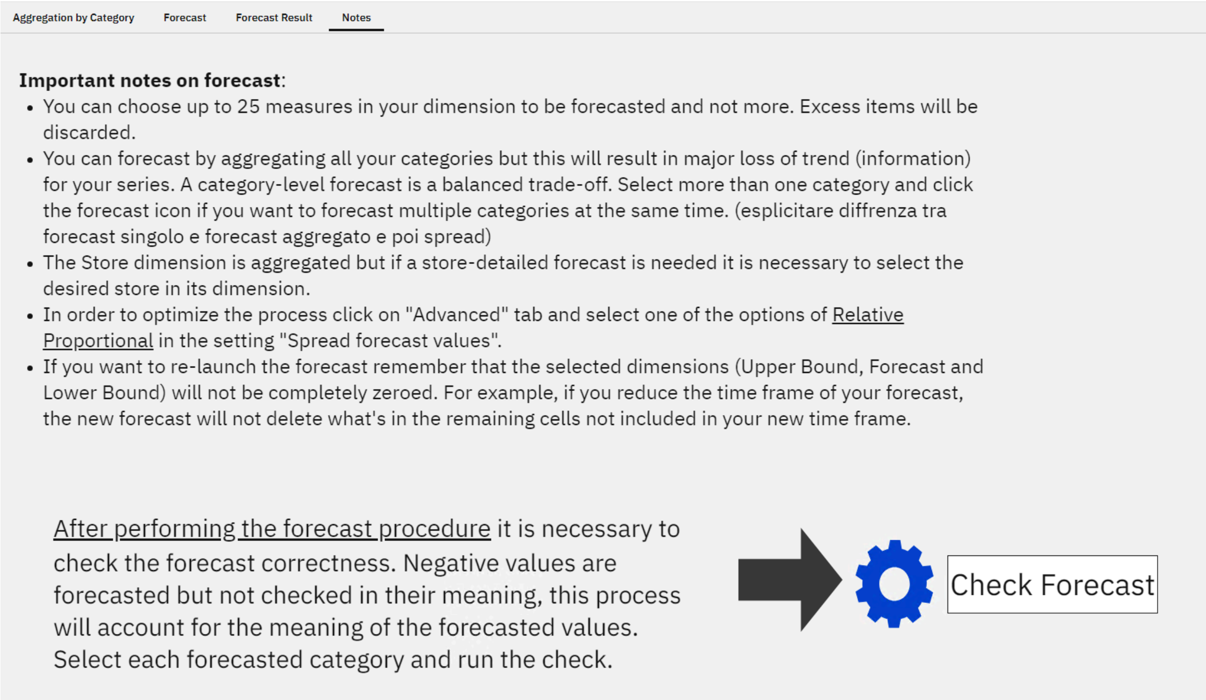


Figure 33 - Detailed Notes

Hints are offered to assist the choice of forecast variables, to improve performance and reliability of data obtained. The lower part of the page contains a button that triggers a process in which, due to the logic implementation of IBM Planning Analytics forecast process, an assessment of the forecast obtained is implemented. By launching a forecast process, the choice of saving confidence bounds would write, into the selected elements, values coming from pure calculations and this include negative values. In the context of sales analysis, it is important to note that negative numbers are not meaningful or applicable. Negative sales values defy the fundamental concept of revenue generation and represent a contradiction of terms. Therefore, when examining sales data, it is crucial to exclude any negative values from the analysis as they do not provide valuable or meaningful insights. For this reason, the "Check forecast" process has the purpose of analyzing the result obtained and acting on the negative numbers by replacing them with zeros.

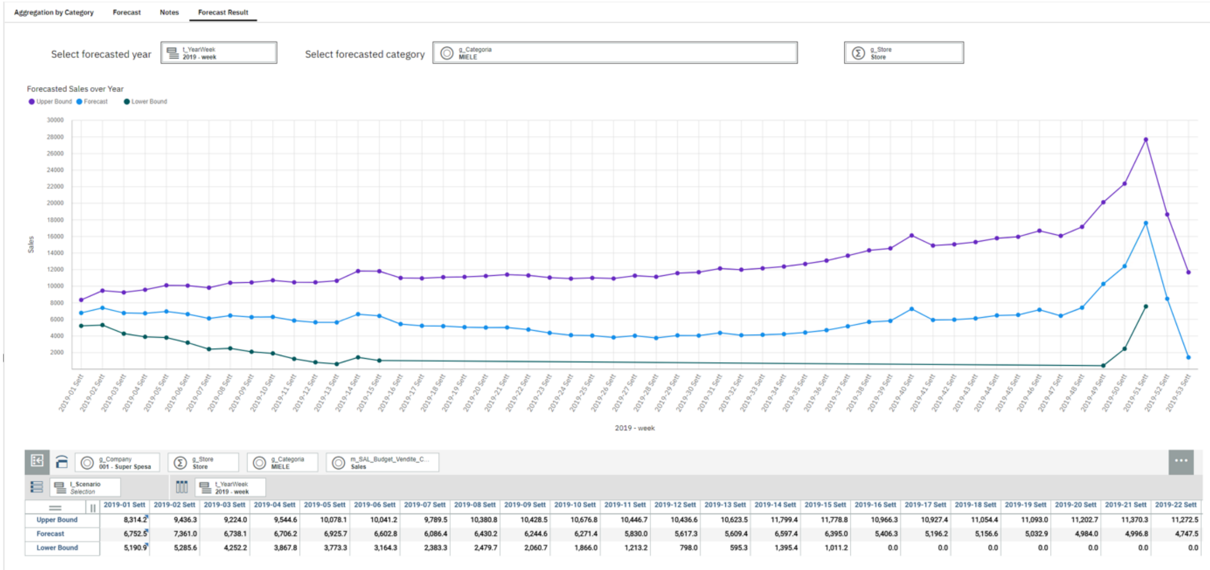


Figure 34 - Forecast Results Analysis

The last dashboard represents an analysis of results obtained. Filters placed in the upper section of the dashboard allow to select the desired category, store, or time window while in the central section a graph outlines the trend of the result. Negative values were set to zero in accordance to previous reasonings and the lower Exploration view offers the possibility to manually modify single values of the forecast for customization reasons. Even if provided as functionality, manually modifying data is not recommended for a logical model-related reason: the scenario containing forecast information should always represent the forecast result and not the result of what a user might believe sales will be. Scenarios dimension implementation is intended to make comparisons between the different what-if situations that a firm seek to analyze, therefore a “Forecast” scenario should always be representative of that type of data. There can be situations in which the forecasts generated might lack of insights that only professionals can have and so this option has been considered to be of valuable help.

5.4.3. Budget Plan

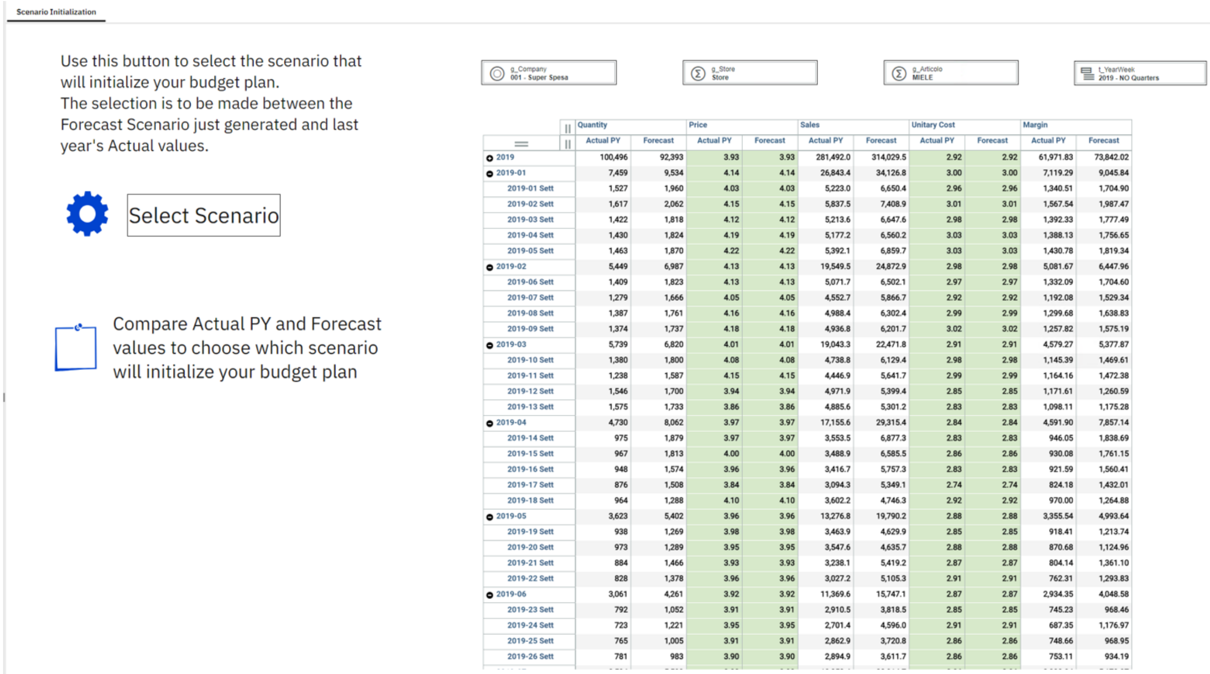


Figure 35 - Scenario Plan Aggregation

The budgeting function starts with the same process of aggregation set-up for the forecast, but this case involves the preliminary choice of which scenario will initialize the Budget. The choice is to be made between Actual PY (Previous Year) sales data, representing the sales of the last available year, and Forecast, representing the previously predicted sales. This offered possibility is a key enhancer of the budget process. By choosing between the two options an analyst might notice that one of the two sets gives better results and thus brings its budget closer to what will actually occur. The availability of this option is not widely diffused and should not be disregarded. Standard procedures typically dictate that the reference values are those of the last available year and then giving a meaningful option to consider while defining budgets is an innovative approach to the budgeting activity.

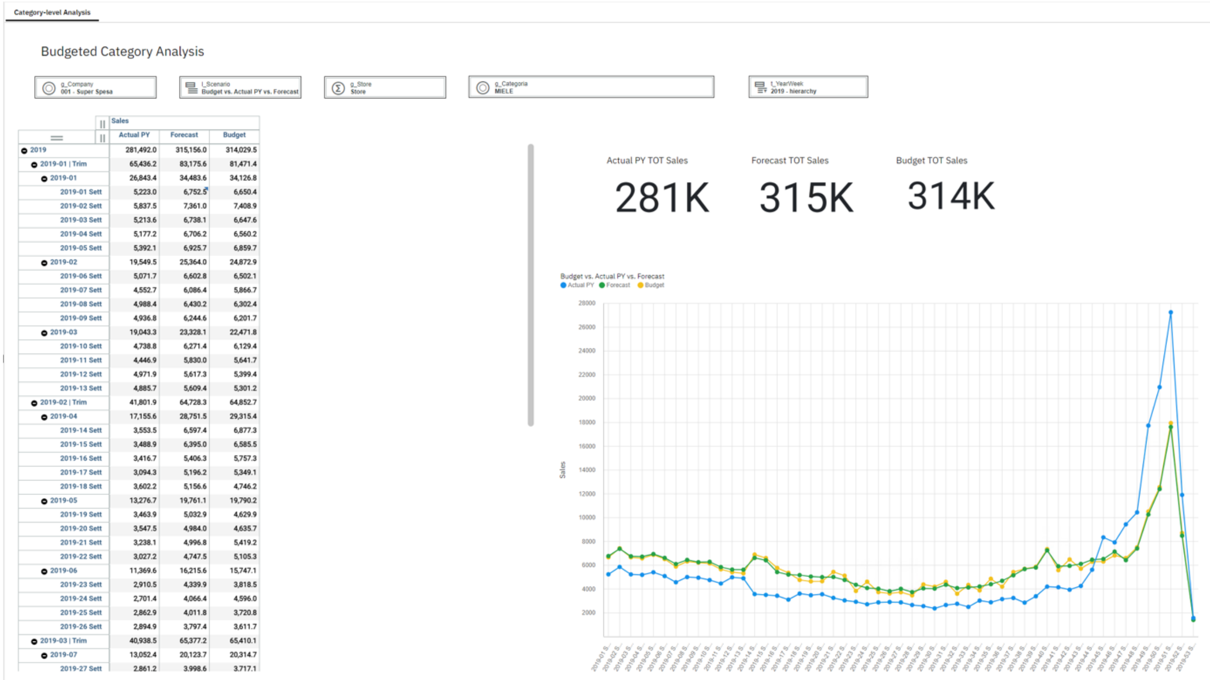


Figure 36 - Budget Analysis

After having chosen the scenario to initialize the budget, a comparison dashboard is set to analyze overall KPIs and single values across time. The graph compares scenarios giving better and immediate understanding in trends and spikes included in each scenario.

	Sales		
	Forecast	Budget	% Forecast v...
2019	315,156.0	314,029.5	(0.36%)
2019-01	34,483.6	34,126.8	(1.03%)
2019-02	25,364.0	24,872.9	(1.94%)
2019-03	23,328.1	22,471.8	(3.67%)
2019-04	28,751.5	29,315.4	1.96%
2019-05	19,761.1	19,790.2	0.15%
2019-06	16,215.6	15,747.1	(2.89%)
2019-07	20,123.7	20,314.7	0.95%
2019-08	16,759.3	16,630.0	(0.77%)
2019-09	28,494.3	28,465.4	(0.10%)
2019-10	24,372.7	24,005.8	(1.51%)
2019-11	27,435.4	27,184.0	(0.92%)
2019-12	50,066.8	51,105.6	2.07%
2019-49 Sett	10,243.2	10,497.1	2.48%
2019-50 Sett	12,379.4	12,544.5	1.33%
2019-51 Sett	17,599.2	17,930.0	1.88%
2019-52 Sett	8,460.7	8,694.9	2.77%
2019-53 Sett	1,384.3	1,439.0	3.95%

Figure 37 - % Scenario Comparisons

Variance analysis can be set to calculate differences between values in columns. Percentages or absolutes differences can be set to be automatically calculated and users can directly modify the resulting values to set the percentage or absolute difference between numbers that will be autonomously modified by IBM Planning Analytics.

The process goes on with the data spread.

In a classical spread activity, having the reference values (x_i and X) and the value to be spread (Z), what is done is the calculation of the ratio between the past values that drive the spread, $d = x_i / X$, and then the product with the value to be spread, $D = d * Z$. D will represent the spread value from Z that has been obtained by the driver d. Adopting this procedure would have meant to accept computational times not compatible with business requirements. Nowadays, firms that aim to introduce these kinds of processes into their businesses generally have to cope with large amount of data and this would have meant to implement a process that may take time in the order of hours, clearly not acceptable. The reason for such a long time is in the logic of the process: for each value at the maximum detail, it is necessary to read the values, calculate the ratio, multiply the latter by the value to be spread, which has previously been read, and then write the result in the appropriate cell. It might be the case that a large number of cells results empty or zero, but with this implementation the calculation would still be performed regardless of the value of the read cell. This problem has been overcome by adopting a different approach to the problem. A preliminary analysis of the values excludes meaningless cells to permit the avoidance of useless calculation, that worsen performances, and then instead of a basic scrolling of all values, only the meaningful ones are considered. The adoption of this method led to major time improvements. The case of large data sets mostly filled with relevant data cannot obviously be improved, since the cells to be calculated are so many that a different approach won't improve time required.

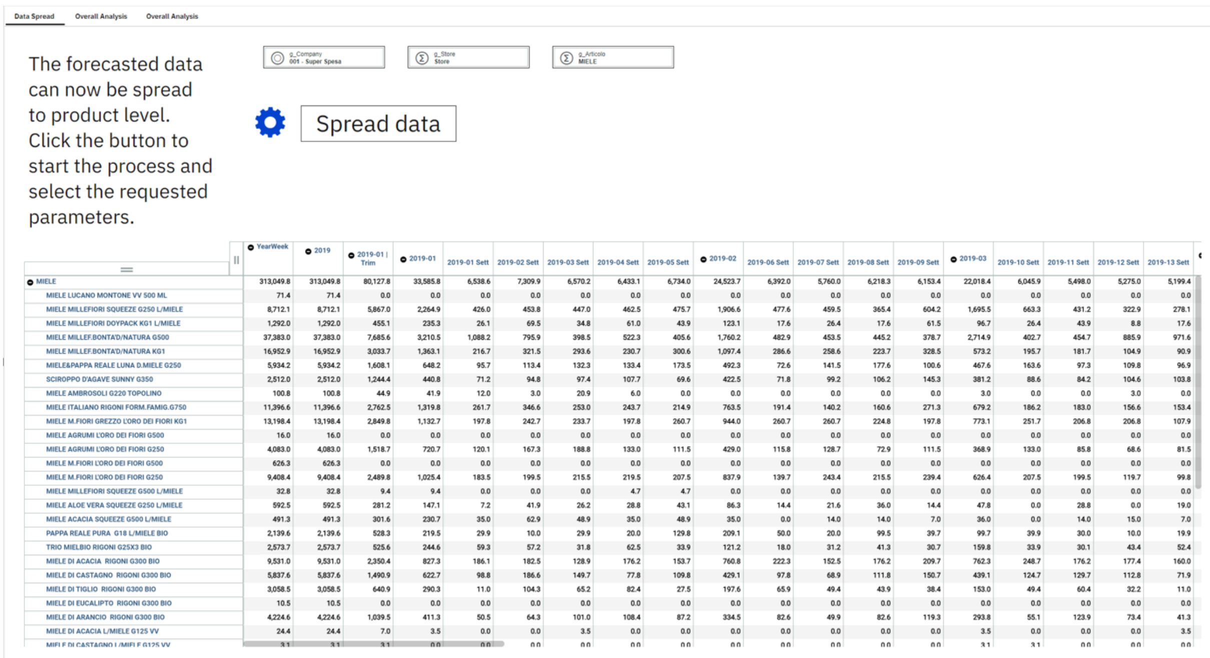
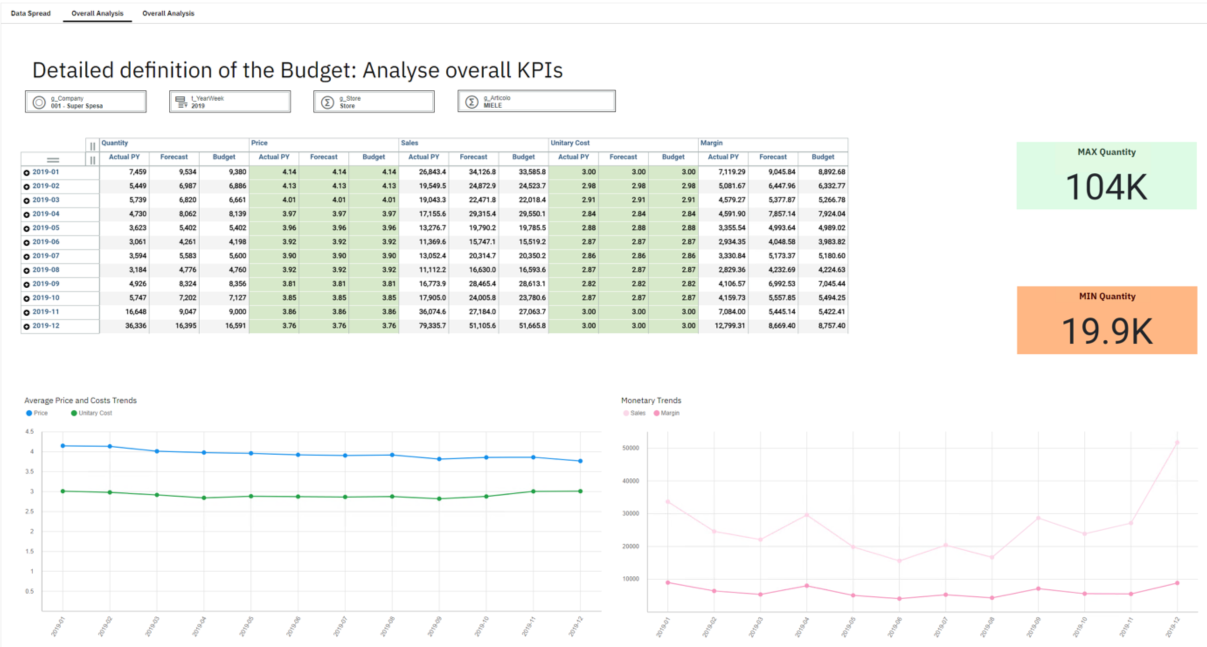


Figure 38 - Budget Spread

Aggregated category-level values for the budget defined need to be spread over each product. This has the purpose to diffuse the aggregated data to single-products by following a defined driver provided at the set-up of the process. Selecting the button "Spread data" and entering the required parameters will start the spread of values to the appropriate products.



	Quantity			Price			Sales			Unitary Cost			Margin		
	Actual PY	Forecast	Budget	Actual PY	Forecast	Budget	Actual PY	Forecast	Budget	Actual PY	Forecast	Budget	Actual PY	Forecast	Budget
2019-01	7,459	9,534	9,380	4.14	4.14	4.14	26,843.4	34,126.8	33,585.8	3.00	3.00	3.00	7,119.29	9,045.84	8,892.68
2019-02	5,449	6,987	6,886	4.13	4.13	4.13	19,549.5	24,872.9	24,523.7	2.98	2.98	2.98	5,081.67	6,447.96	6,332.77
2019-03	5,739	6,820	6,661	4.01	4.01	4.01	19,043.3	22,471.8	22,018.4	2.91	2.91	2.91	4,579.27	5,377.87	5,266.78
2019-04	4,730	8,062	8,139	3.97	3.97	3.97	17,155.6	29,315.4	29,550.1	2.84	2.84	2.84	4,591.90	7,857.14	7,924.04
2019-05	3,623	5,402	5,402	3.96	3.96	3.96	13,276.7	19,790.2	19,785.5	2.88	2.88	2.88	3,355.54	4,993.64	4,989.02
2019-06	3,061	4,261	4,198	3.92	3.92	3.92	11,369.6	15,747.1	15,519.2	2.87	2.87	2.87	2,934.35	4,048.58	3,983.82
2019-07	3,594	5,583	5,600	3.90	3.90	3.90	13,052.4	20,314.7	20,350.2	2.86	2.86	2.86	3,330.84	5,173.37	5,180.60
2019-08	3,184	4,776	4,760	3.92	3.92	3.92	11,112.2	16,630.0	16,593.6	2.87	2.87	2.87	2,829.36	4,232.69	4,224.63
2019-09	4,926	8,324	8,356	3.81	3.81	3.81	16,773.9	28,465.4	28,613.1	2.82	2.82	2.82	4,106.57	6,992.53	7,045.44
2019-10	5,747	7,202	7,127	3.85	3.85	3.85	17,905.0	24,005.8	23,780.6	2.87	2.87	2.87	4,159.73	5,557.85	5,494.25
2019-11	16,648	9,047	9,000	3.86	3.86	3.86	36,074.6	27,184.0	27,063.7	3.00	3.00	3.00	7,084.00	5,445.14	5,422.41
2019-12	36,336	16,395	16,591	3.76	3.76	3.76	79,335.7	51,105.6	51,665.8	3.00	3.00	3.00	12,799.31	8,669.40	8,757.40

Figure 39 - Budget Analysis

Another dashboard offers analysis of the budgeted figures. Comparisons between scenarios and metrics can be analyzed in the explorative tab showing a view on data and more direct impact is given by KPIs, which in this case represents the maximum and minimum quantity budgeted for the "next" year, meaning the year subject to the definition of the budget.

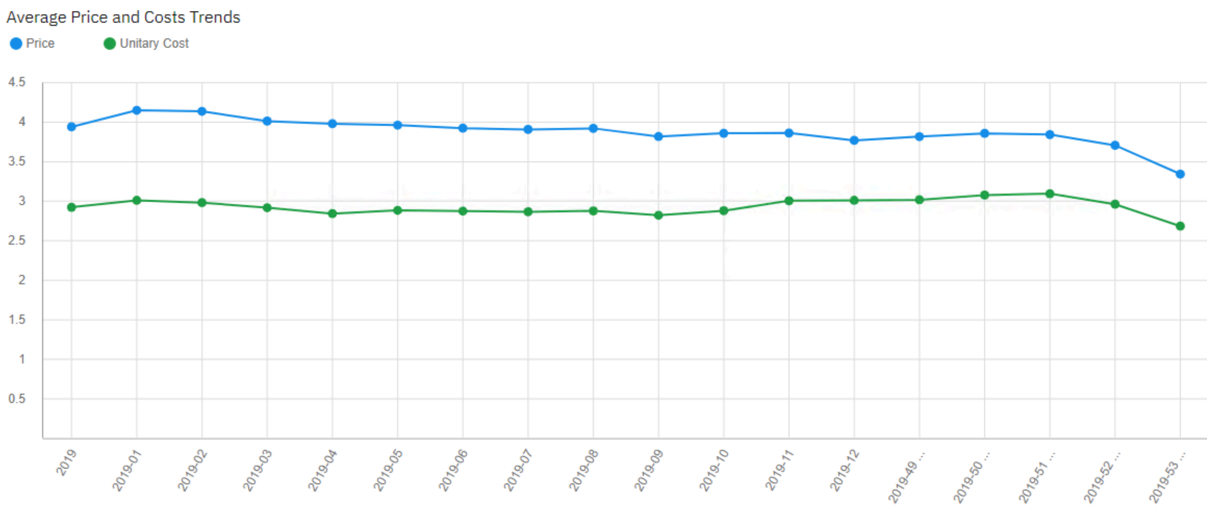


Figure 40 - Budgeted Price Trends

The previous graph represents an analysis of average prices and costs over time. Time selected in the exploration view will affect the time represented in graphs: the elements selected for the exploration are then displayed in different graphs that "depend" on the same data.

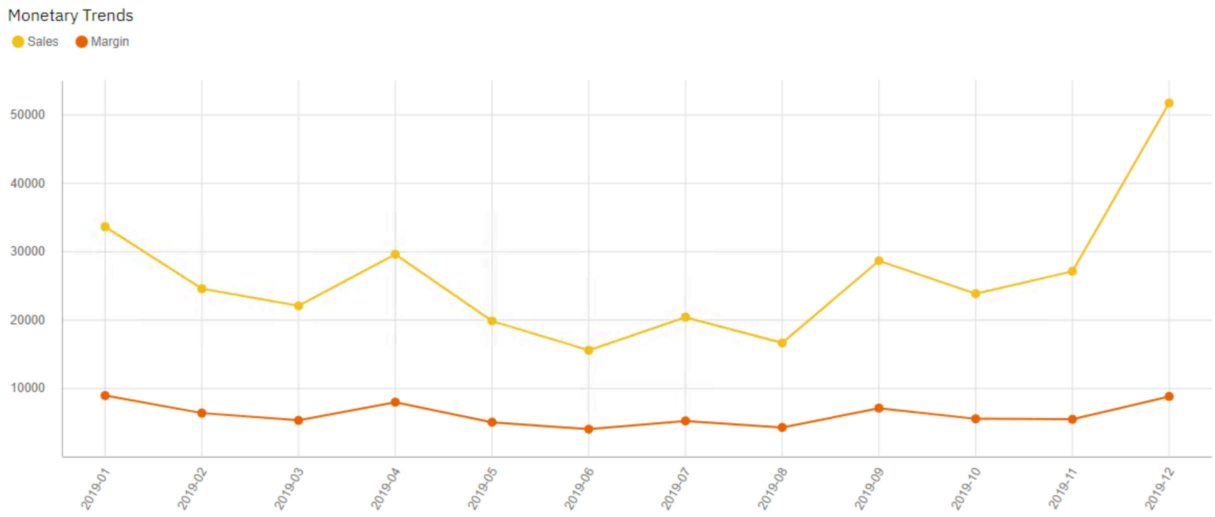


Figure 41 - Budgeted Monthly Monetary Trends

Sales and Margins are key metrics to assess the health state of a company. Fixing objectives on Sales and Margins might be necessary or required to plan the liquidity a firm will have, or the ability to repay short-term debt then an accurate planning will trigger an improved cash management.

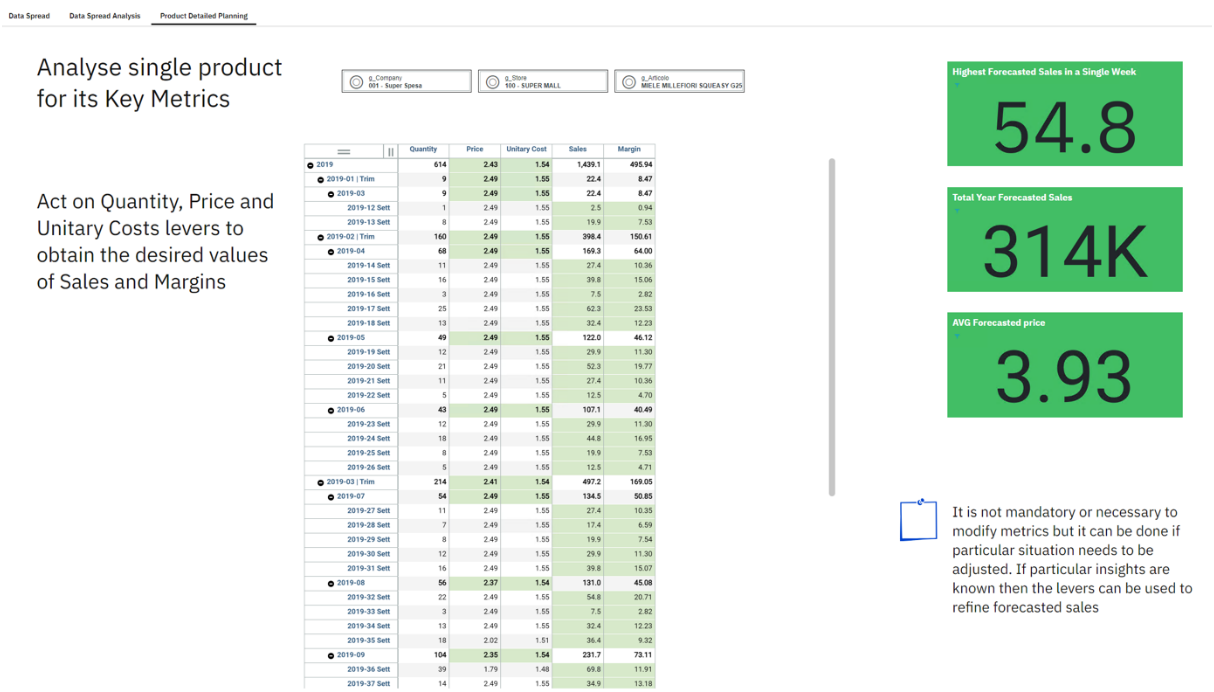


Figure 42 - Product Detailed Planning

The last visualization is optional. It is offered the possibility to act on quantity, price, and unitary cost levers to modify single-product, and eventually single-store, sales levels. This step is not mandatory and will provide benefit only in case a manager can provide better understanding of a market or a market-niche. With the knowledge of a market expert and eventually other insights information, it may be necessary to update specific values in order to reflect a situation not predicted before. KPIs still offers an immediate understanding of a firms' metrics in relation to sales and prices.

5.4.4. Variance Analysis

The final phase of the planning and forecasting process is known as "Variance Analysis". It serves as a critical evaluation of the accuracy of the estimates made one year prior. This essential step aims to enhance the effectiveness of planning and forecasting activities. By incorporating the latest information, businesses can compare the actual results with the previously forecasted and budgeted values and identify any variances or deviations. This analysis allows for a comprehensive assessment of the precision of the forecasting models and provides insights into the effectiveness of the planning processes. By leveraging the capabilities of IBM Planning Analytics, organizations can gain valuable insights into the factors that contributed to the variances, enabling them to refine their future forecasting and planning strategies. Overall, the Variance Analysis phase serves as a crucial feedback loop, facilitating continuous improvement and informed decision-making in the dynamic business environment. The process begins by loading the most up-to-date data into the IBM Planning Analytics application.

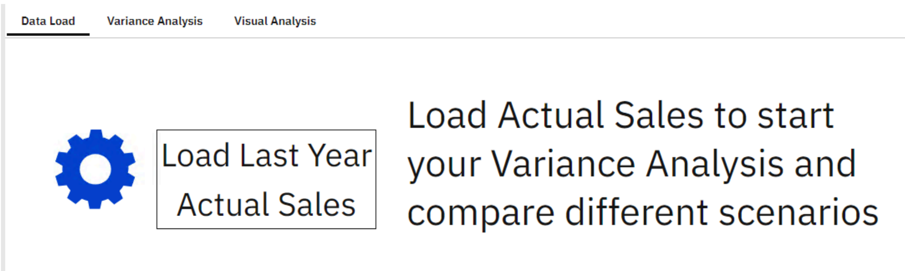
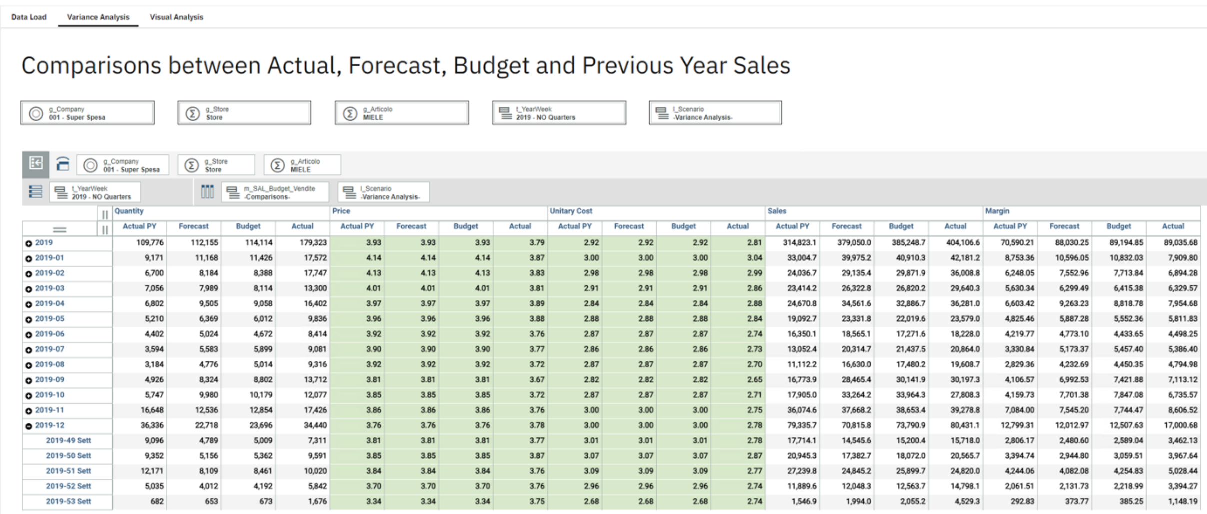


Figure 43 - Variance Data Load

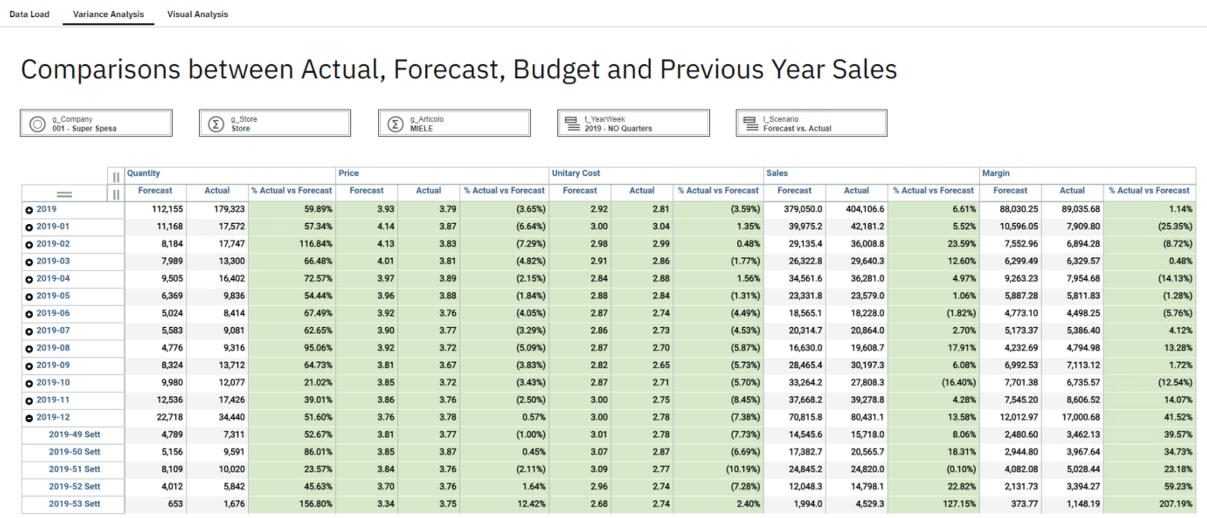


	Quantity				Price				Unitary Cost			
	Actual PY	Forecast	Budget	Actual	Actual PY	Forecast	Budget	Actual	Actual PY	Forecast	Budget	Actual
2019	109,776	112,155	114,114	179,323	3.93	3.93	3.93	3.79	2.92	2.92	2.92	2.81
2019-01	9,171	11,168	11,426	17,572	4.14	4.14	4.14	3.87	3.00	3.00	3.00	3.04
2019-02	6,700	8,184	8,388	17,747	4.13	4.13	4.13	3.83	2.98	2.98	2.98	2.99
2019-03	7,056	7,989	8,114	13,300	4.01	4.01	4.01	3.81	2.91	2.91	2.91	2.86
2019-04	6,802	9,505	9,058	16,402	3.97	3.97	3.97	3.89	2.84	2.84	2.84	2.88
2019-05	5,210	6,369	6,012	9,836	3.96	3.96	3.96	3.88	2.88	2.88	2.88	2.84
2019-06	4,402	5,024	4,672	8,414	3.92	3.92	3.92	3.76	2.87	2.87	2.87	2.74
2019-07	3,594	5,583	5,899	9,081	3.90	3.90	3.90	3.77	2.86	2.86	2.86	2.73
2019-08	3,184	4,776	5,014	9,316	3.92	3.92	3.92	3.72	2.87	2.87	2.87	2.70
2019-09	4,926	8,324	8,802	13,712	3.81	3.81	3.81	3.67	2.82	2.82	2.82	2.65
2019-10	5,747	9,980	10,179	12,077	3.85	3.85	3.85	3.72	2.87	2.87	2.87	2.71
2019-11	16,648	12,536	12,854	17,426	3.86	3.86	3.86	3.76	3.00	3.00	3.00	2.75
2019-12	36,336	22,718	23,696	34,440	3.76	3.76	3.76	3.78	3.00	3.00	3.00	2.78
2019-49 Sett	9,096	4,789	5,009	7,311	3.81	3.81	3.81	3.77	3.01	3.01	3.01	2.78
2019-50 Sett	9,352	5,156	5,362	9,591	3.85	3.85	3.85	3.87	3.07	3.07	3.07	2.87
2019-51 Sett	12,171	8,109	8,461	10,020	3.84	3.84	3.84	3.76	3.09	3.09	3.09	2.77
2019-52 Sett	5,035	4,012	4,192	5,842	3.70	3.70	3.70	3.76	2.96	2.96	2.96	2.74
2019-53 Sett	682	653	673	1,676	3.34	3.34	3.34	3.75	2.68	2.68	2.68	2.74

	Sales				Margin			
	Actual PY	Forecast	Budget	Actual	Actual PY	Forecast	Budget	Actual
2019	314,823.1	379,050.0	385,248.7	404,106.6	70,590.21	88,030.25	89,194.85	89,035.68
2019-01	33,004.7	39,975.2	40,910.3	42,181.2	8,753.36	10,596.05	10,832.03	7,909.80
2019-02	24,036.7	29,135.4	29,871.9	36,008.8	6,248.05	7,552.96	7,713.84	6,894.28
2019-03	23,414.2	26,322.8	26,820.2	29,640.3	5,630.34	6,299.49	6,415.38	6,329.57
2019-04	24,670.8	34,561.6	32,886.7	36,281.0	6,603.42	9,263.23	8,818.78	7,954.68
2019-05	19,092.7	23,331.8	22,019.6	23,579.0	4,825.46	5,887.28	5,552.36	5,811.83
2019-06	16,350.1	18,565.1	17,271.6	18,228.0	4,219.77	4,773.10	4,433.65	4,498.25
2019-07	13,052.4	20,314.7	21,437.5	20,864.0	3,330.84	5,173.37	5,457.40	5,386.40
2019-08	11,112.2	16,630.0	17,480.2	19,608.7	2,829.36	4,232.69	4,450.35	4,794.98
2019-09	16,773.9	28,465.4	30,141.9	30,197.3	4,106.57	6,992.53	7,421.88	7,113.12
2019-10	17,905.0	33,264.2	33,964.3	27,808.3	4,159.73	7,701.38	7,847.08	6,735.57
2019-11	36,074.6	37,668.2	38,653.4	39,278.8	7,084.00	7,545.20	7,744.47	8,606.52
2019-12	79,335.7	70,815.8	73,790.9	80,431.1	12,799.31	12,012.97	12,507.63	17,000.68
2019-49 Sett	17,714.1	14,545.6	15,200.4	15,718.0	2,806.17	2,480.60	2,589.04	3,462.13
2019-50 Sett	20,945.3	17,382.7	18,072.0	20,565.7	3,394.74	2,944.80	3,059.51	3,967.64
2019-51 Sett	27,239.8	24,845.2	25,899.7	24,820.0	4,244.06	4,082.08	4,254.83	5,028.44
2019-52 Sett	11,889.6	12,048.3	12,563.7	14,798.1	2,061.51	2,131.73	2,218.99	3,394.27
2019-53 Sett	1,546.9	1,994.0	2,055.2	4,529.3	292.83	373.77	385.25	1,148.19

Figure 44 - Variance Analysis

Annual and monthly data provide a deep understanding of the difference between previous years, budgeted, forecasted, and actual metrics. Different sets of metrics and scenarios are predefined to give the user a responsive and comparative dashboard to interact with.



Data Load | Variance Analysis | Visual Analysis

Comparisons between Actual, Forecast, Budget and Previous Year Sales

g_Company 001 - Super Spesa | g_Store Store | g_Articolo MIELE | t_YearWeek 2019 - NO Quarters | l_Scenario Forecast vs. Actual

	Quantity			Price			Unitary Cost			Sales			Margin		
	Forecast	Actual	% Actual vs Forecast	Forecast	Actual	% Actual vs Forecast	Forecast	Actual	% Actual vs Forecast	Forecast	Actual	% Actual vs Forecast	Forecast	Actual	% Actual vs Forecast
2019	112,155	179,323	59.89%	3.93	3.79	(3.65%)	2.92	2.81	(3.59%)	379,050.0	404,106.6	6.61%	88,030.25	89,035.68	1.14%
2019-01	11,168	17,572	57.34%	4.14	3.87	(6.64%)	3.00	3.04	1.35%	39,975.2	42,181.2	5.52%	10,596.05	7,909.80	(25.35%)
2019-02	8,184	17,747	116.84%	4.13	3.83	(7.29%)	2.98	2.99	0.48%	29,135.4	36,008.8	23.59%	7,552.96	6,894.28	(8.72%)
2019-03	7,989	13,300	66.48%	4.01	3.81	(4.82%)	2.91	2.86	(1.77%)	26,322.8	29,640.3	12.60%	6,299.49	6,329.57	0.48%
2019-04	9,505	16,402	72.57%	3.97	3.89	(2.15%)	2.84	2.88	1.56%	34,561.6	36,281.0	4.97%	9,263.23	7,954.68	(14.13%)
2019-05	6,369	9,836	54.44%	3.96	3.88	(1.84%)	2.88	2.84	(1.31%)	23,331.8	23,579.0	1.06%	5,887.28	5,811.83	(1.28%)
2019-06	5,024	8,414	67.49%	3.92	3.76	(4.05%)	2.87	2.74	(4.49%)	18,565.1	18,228.0	(1.82%)	4,773.10	4,498.25	(5.76%)
2019-07	5,583	9,081	62.65%	3.90	3.77	(3.29%)	2.86	2.73	(4.53%)	20,314.7	20,864.0	2.70%	5,173.37	5,386.40	4.12%
2019-08	4,776	9,316	95.06%	3.92	3.72	(5.09%)	2.87	2.70	(5.87%)	16,630.0	19,608.7	17.91%	4,232.69	4,794.98	13.28%
2019-09	8,324	13,712	64.73%	3.81	3.67	(3.83%)	2.82	2.65	(5.73%)	28,465.4	30,197.3	6.08%	6,992.53	7,113.12	1.72%
2019-10	9,980	12,077	21.02%	3.85	3.72	(3.43%)	2.87	2.71	(5.70%)	33,264.2	27,808.3	(16.40%)	7,701.38	6,735.57	(12.54%)
2019-11	12,536	17,426	39.01%	3.86	3.76	(2.50%)	3.00	2.75	(8.45%)	37,668.2	39,278.8	4.28%	7,545.20	8,606.52	14.07%
2019-12	22,718	34,440	51.60%	3.76	3.78	0.57%	3.00	2.78	(7.38%)	70,815.8	80,431.1	13.58%	12,012.97	17,000.68	41.52%
2019-49 Sett	4,789	7,311	52.67%	3.81	3.77	(1.00%)	3.01	2.78	(7.73%)	14,545.6	15,718.0	8.06%	2,480.60	3,462.13	39.57%
2019-50 Sett	5,156	9,591	86.01%	3.85	3.87	0.45%	3.07	2.87	(6.69%)	17,382.7	20,565.7	18.31%	2,944.80	3,967.64	34.73%
2019-51 Sett	8,109	10,020	23.57%	3.84	3.76	(2.11%)	3.09	2.77	(10.19%)	24,845.2	24,820.0	(0.10%)	4,082.08	5,028.44	23.18%
2019-52 Sett	4,012	5,842	45.63%	3.70	3.76	1.64%	2.96	2.74	(7.28%)	12,048.3	14,798.1	22.82%	2,131.73	3,394.27	59.23%
2019-53 Sett	653	1,676	156.80%	3.34	3.75	12.42%	2.68	2.74	2.40%	1,994.0	4,529.3	127.15%	373.77	1,148.19	207.19%

Figure 45 - Variance Analysis Forecast vs. Actual

The previous figure depicts a comparison between forecast and actual values on all available metrics. Time-periods can be drilled down to the lowest granularity and with the aim of filters, placed at the top of the page, single-stores or single-product can be chosen. The filter aimed at scenario selection offers the possibility to choose between predefined sets that compare scenarios and implement the automatic calculation of variances between them.

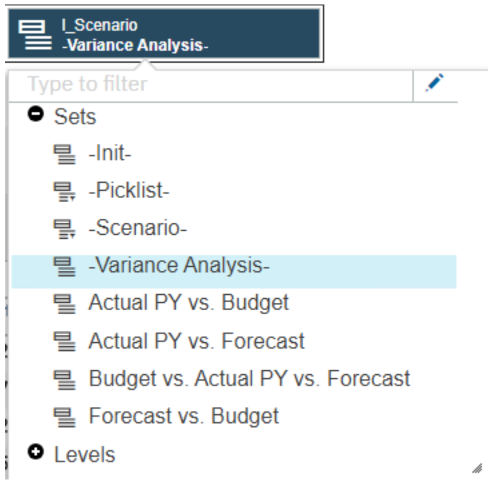


Figure 46 - Scenario Filter

The choice of one set will display measures and variances between scenarios. This is useful for immediate comparisons and to jump between scenarios to be analyzed.

Data Load | Variance Analysis | Visual Analysis

Comparisons between Actual, Forecast, Budget and Previous Year Sales

g_Company: 001 - Super Spesa | g_Store: Store | g_Articolo: MIELE | t_YearWeek: 2019 - NO Quarters | l_Scenario: Actual PY vs. Actual

	Quantity			Price			Unitary Cost			Sales			Margin		
	Actual PY	Actual	% Actual PY vs Actual	Actual PY	Actual	% Actual PY vs Actual	Actual PY	Actual	% Actual PY vs Actual	Actual PY	Actual	% Actual PY vs Actual	Actual PY	Actual	% Actual PY vs Actual
2019	109,776	179,323	63.35%	3.93	3.79	(3.65%)	2.92	2.81	(3.59%)	314,823.1	404,106.6	28.36%	70,590.21	89,035.68	26.13%
2019-01	9,171	17,572	91.60%	4.14	3.87	(6.64%)	3.00	3.04	1.35%	33,004.7	42,181.2	27.80%	8,753.36	7,909.80	(9.64%)
2019-02	6,700	17,747	164.89%	4.13	3.83	(7.29%)	2.98	2.99	0.48%	24,036.7	36,008.8	49.81%	6,248.05	6,894.28	10.34%
2019-03	7,056	13,300	88.49%	4.01	3.81	(4.82%)	2.91	2.86	(1.77%)	23,414.2	29,640.3	26.59%	5,630.34	6,329.57	12.42%
2019-04	6,802	16,402	141.13%	3.97	3.89	(2.15%)	2.84	2.88	1.56%	24,670.8	36,281.0	47.06%	6,603.42	7,954.68	20.46%
2019-05	5,210	9,836	88.79%	3.96	3.88	(1.84%)	2.88	2.84	(1.31%)	19,092.7	23,579.0	23.50%	4,825.46	5,811.83	20.44%
2019-06	4,402	8,414	91.14%	3.92	3.76	(4.05%)	2.87	2.74	(4.49%)	16,350.1	18,228.0	11.49%	4,219.77	4,498.25	6.60%
2019-07	3,594	9,081	152.67%	3.90	3.77	(3.29%)	2.86	2.73	(4.53%)	13,052.4	20,864.0	59.85%	3,330.84	5,386.40	61.71%
2019-08	3,184	9,316	192.59%	3.92	3.72	(5.09%)	2.87	2.70	(5.87%)	11,112.2	19,608.7	76.46%	2,829.36	4,794.98	69.47%
2019-09	4,926	13,712	178.36%	3.81	3.67	(3.83%)	2.82	2.65	(5.73%)	16,773.9	30,197.3	80.03%	4,106.57	7,113.12	73.21%
2019-10	5,747	12,077	110.14%	3.85	3.72	(3.43%)	2.87	2.71	(5.70%)	17,905.0	27,808.3	55.31%	4,159.73	6,735.57	61.92%
2019-11	16,648	17,426	4.67%	3.86	3.76	(2.50%)	3.00	2.75	(8.45%)	36,074.6	39,278.8	8.88%	7,084.00	8,606.52	21.49%
2019-12	36,336	34,440	(5.22%)	3.76	3.78	0.57%	3.00	2.78	(7.38%)	79,335.7	80,431.1	1.38%	12,799.31	17,000.68	32.82%
2019-49 Sett	9,096	7,311	(19.62%)	3.81	3.77	(1.00%)	3.01	2.78	(7.73%)	17,714.1	15,718.0	(11.27%)	2,806.17	3,462.13	23.38%
2019-50 Sett	9,352	9,591	2.56%	3.85	3.87	0.45%	3.07	2.87	(6.69%)	20,945.3	20,565.7	(1.81%)	3,394.74	3,967.64	16.88%
2019-51 Sett	12,171	10,020	(17.67%)	3.84	3.76	(2.11%)	3.09	2.77	(10.19%)	27,239.8	24,820.0	(8.88%)	4,244.06	5,028.44	18.48%
2019-52 Sett	5,035	5,842	16.03%	3.70	3.76	1.64%	2.96	2.74	(7.28%)	11,889.6	14,798.1	24.46%	2,061.51	3,394.27	64.65%
2019-53 Sett	682	1,676	145.75%	3.34	3.75	12.42%	2.68	2.74	2.40%	1,546.9	4,529.3	192.79%	292.83	1,148.19	292.10%

Figure 47 - Overall Analysis Actual PY vs. Actual

Data Load | Variance Analysis | Visual Analysis

Comparisons between Actual, Forecast, Budget and Previous Year Sales

g_Company: 001 - Super Spesa | g_Store: Store | g_Articolo: MIELE | t_YearWeek: 2019 - NO Quarters | l_Scenario: -Picklist-

	Sales						
	Actual	Forecast	Budget	Actual PY	% Forecast vs Budget	% Actual vs Forecast	% Actual PY vs Actual
2019	404,106.6	379,050.0	385,248.7	314,823.1	1.64%	6.61%	28.36%
2019-01	42,181.2	39,975.2	40,910.3	33,004.7	2.34%	5.52%	27.80%
2019-02	36,008.8	29,135.4	29,871.9	24,036.7	2.53%	23.59%	49.81%
2019-03	29,640.3	26,322.8	26,820.2	23,414.2	1.89%	12.60%	26.59%
2019-04	36,281.0	34,561.6	32,886.7	24,670.8	(4.85%)	4.97%	47.06%
2019-05	23,579.0	23,331.8	22,019.6	19,092.7	(5.62%)	1.06%	23.50%
2019-06	18,228.0	18,565.1	17,271.6	16,350.1	(6.97%)	(1.82%)	11.49%
2019-07	20,864.0	20,314.7	21,437.5	13,052.4	5.53%	2.70%	59.85%
2019-08	19,608.7	16,630.0	17,480.2	11,112.2	5.11%	17.91%	76.46%
2019-09	30,197.3	28,465.4	30,141.9	16,773.9	5.89%	6.08%	80.03%
2019-10	27,808.3	33,264.2	33,964.3	17,905.0	2.10%	(16.40%)	55.31%
2019-11	39,278.8	37,668.2	38,653.4	36,074.6	2.62%	4.28%	8.88%
2019-12	80,431.1	70,815.8	73,790.9	79,335.7	4.20%	13.58%	1.38%
2019-49 Sett	15,718.0	14,545.6	15,200.4	17,714.1	4.50%	8.06%	(11.27%)
2019-50 Sett	20,565.7	17,382.7	18,072.0	20,945.3	3.97%	18.31%	(1.81%)
2019-51 Sett	24,820.0	24,845.2	25,899.7	27,239.8	4.24%	(0.10%)	(8.88%)
2019-52 Sett	14,798.1	12,048.3	12,563.7	11,889.6	4.28%	22.82%	24.46%
2019-53 Sett	4,529.3	1,994.0	2,055.2	1,546.9	3.07%	127.15%	192.79%

Figure 48 - Sales-based Scenario Comparison

The last step concerns a graphical analysis of the available information. The dashboard in Figure 49 has been developed with the aim to reach immediate understandings and conclusions. As in any other dashboard, filters allow to guide the selection of data and graphs react accordingly by updating and reflecting selections made.

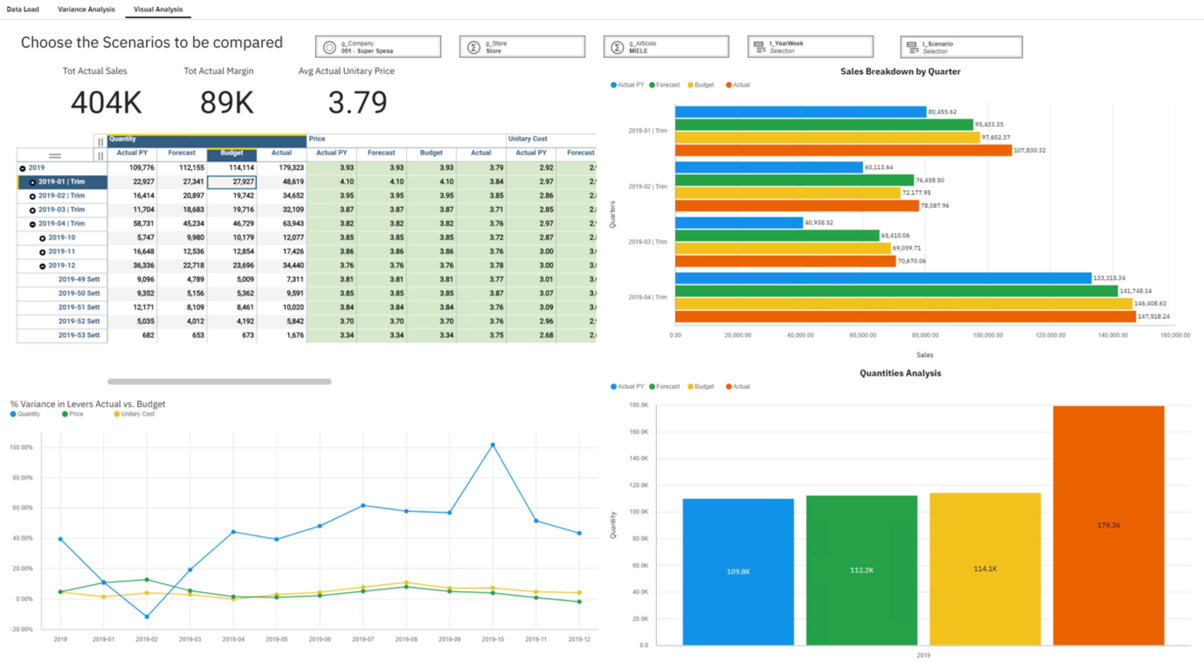


Figure 49 - Graphical Variance Analysis

Terms for comparison are still the various metrics considered and the different scenarios involved in the whole process.

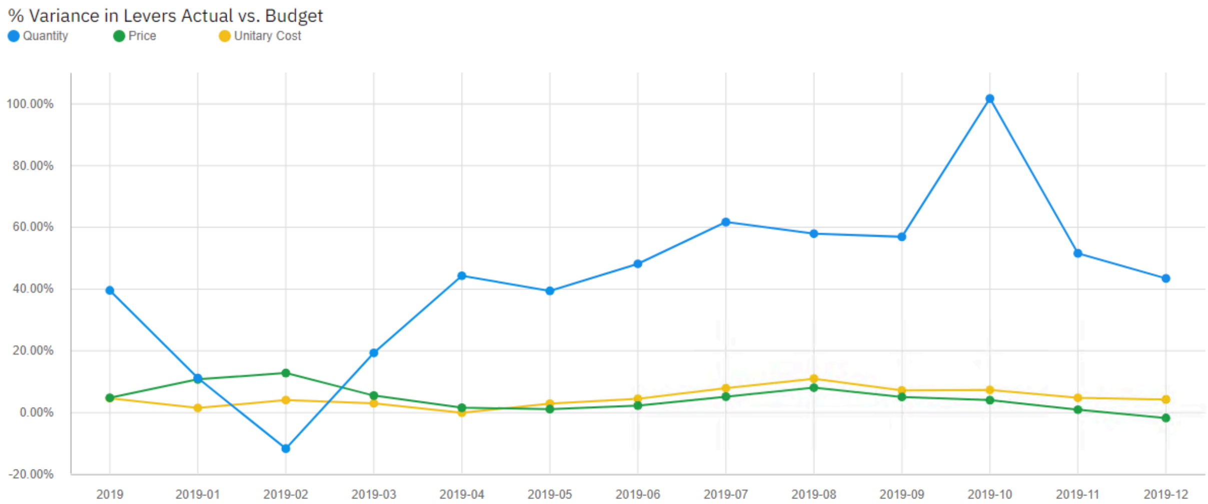


Figure 50 - % Variance in Levers Actual vs. Budget

Trends in variance analysis enable the comparison of different metric variances at the same point in time, then crossing information on different metrics might permit to take conclusions on business development and forecasting and budget models goodness. An example could be the case of February 2019 in Figure 50: in that month prices went up as a consequence of slight increases in costs and this caused a remarkable decrease in quantities sold. This shows how sales and prices are linked to costs, for the firm considered, and a little increase in prices trigger major decreases in prices. This type of consideration triggers deeper analysis of a firm's behaviors, leading then to more accurate understanding of the market and the ability to anticipate market movements instead of reacting to it.

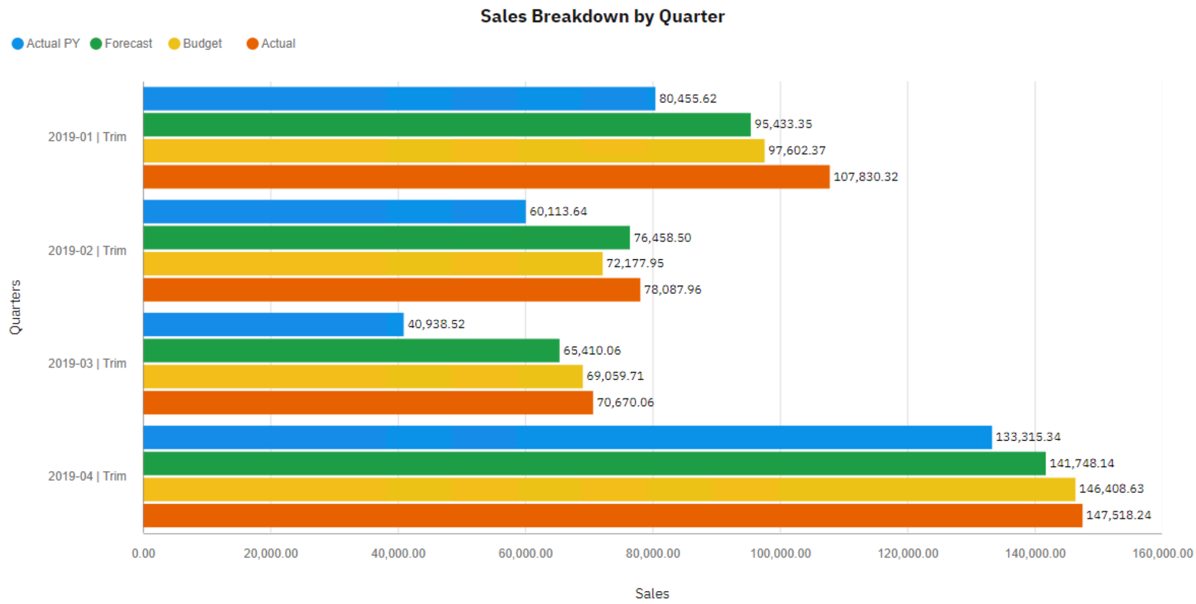


Figure 51 - Sales Breakdown by Quarter

The analysis of sales broken down by quarters and split for scenario is a powerful tool available. The possibility to visually recognize which component reflect a specific behavior is fundamental in avoiding errors and locate data to understand trends. In this graph, quarters can be compared and the focus can be set on specific scenarios or on the comparison of different scenarios. Bars in the graph represent the amount of sales for each scenario and their length is an immediate indicator of results obtained or forecasted. Comparisons between scenarios in the same quarters allows for identification of how well the forecasted and budgeted scenarios were estimated while comparisons of scenarios between different quarters allows identifications of repetitive patterns that might be recurring in the business. If a firm is systematically underestimating forecasts or budgets, it means that forecast models should be revised or that budgeting functions should be supervised by more experienced figures.

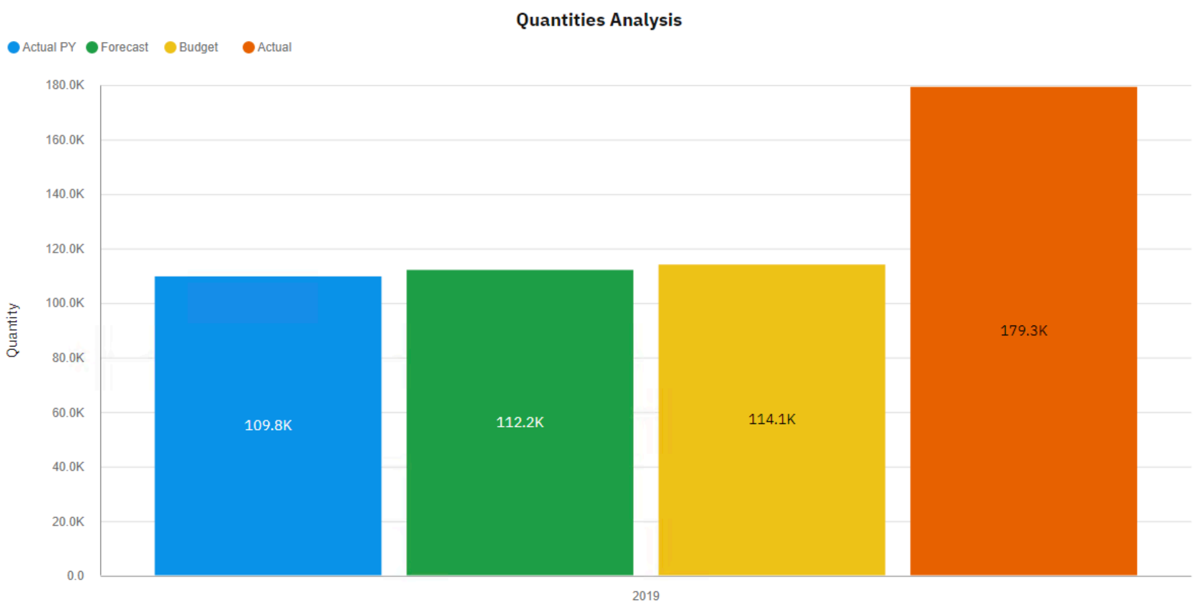


Figure 52 - Quantities Analysis

Ultimately, a yearly quantities analysis gives the possibility to compare growth or degrowth with past years and, as before, with forecasted and budgeted quantities. Appropriate corrective actions should be taken in case of large deviations from the expected quantities.

In conclusion, the presented graphs serve as a valuable recap tool for analyzing sales and performance of a firm over the course of a year. They provide a concise summary of the data and insights obtained through the forecasting and budgeting application developed. By reviewing these graphs, it becomes possible to validate or refute any initial conclusions drawn from earlier analysis. The use of these graphs in the conclusion phase of the workflow enables a holistic assessment of the firm's performance throughout the year. It allows for a comprehensive review of the key findings, trends, and patterns observed in the data. Moreover, these graphs provide a visual representation of the firm's sales trajectory, highlighting any noteworthy fluctuations, successes, or areas of concern.

Building upon the earlier analysis, the conclusion phase goes beyond mere confirmation or denial of suspected conclusions. It aims to synthesize the findings into a coherent narrative, drawing connections between different aspects of the firm's performance and shedding light on the underlying factors influencing sales and overall performance. By considering the implications of the observed trends and patterns, this phase offers insights into potential areas for improvement or strategic actions that can be taken to drive future growth. It serves as a guide for decision-makers in leveraging the knowledge gained from the forecasting and budgeting process to inform strategic planning, resource allocation, and performance management in the coming year.

5.5. Assessment of applicability of forecast and processes

During the testing phase of forecast models and algorithms on IBM Planning Analytics application, an evaluation of the achieved results was conducted. To facilitate this process, a Python module was developed to analyze historical data and generate predictions. This module exploits various packages that leverage the forecasting techniques provided by the ARIMA set of techniques. While the development and implementation of the Python module yielded significant results, it is important to note that sharing specific pieces of code or presenting the obtained results without providing the necessary technical context would be of limited value. Merely showcasing the code or results without accompanying explanations would fail to convey the rationale and supporting evidence behind the findings.

The results obtained from comparing IBM Planning Analytics and the developed Python module, which served as a reference due to its demonstrated capability to generate accurate forecasts on test data sets, indicated comparable outcomes. This implies that the forecasts generated by IBM Planning Analytics were accurate and aligned with the performance of algorithms belonging to the ARIMA class. The evaluation involved a comprehensive analysis of the forecast outputs produced by both IBM Planning Analytics and the Python module. By examining their respective forecasts against the actual data, it was observed that the predictions generated by both approaches exhibited similar levels of accuracy and reliability. This finding highlights the effectiveness of IBM Planning Analytics in generating reliable forecasts that are on par with the performance of the ARIMA algorithms employed by the developed Python module. The comparable results validate the credibility and usefulness of IBM Planning Analytics as a forecasting tool, reinforcing its potential to provide accurate insights for decision-making processes within the scope of this study.

It is important to note that the comparison was conducted on data sets specifically designed for testing purposes, which allowed a controlled and standardized evaluation. However, further analysis and validation using real-cases data sets would be beneficial to confirm the generalizability and robustness of these findings. Special attention and thorough analysis are required when dealing with historical time series data that exhibit unique or peculiar trends. These exceptional patterns can potentially mislead forecasting models, leading to inaccurate predictions and therefore, it is crucial to carefully examine and test such data to ensure the reliability and effectiveness of the forecasting models in capturing and interpreting extreme behaviors.

Overall, the evidence suggests that IBM Planning Analytics delivers reliable forecasting capabilities comparable to the ARIMA algorithms implemented in the developed Python module. This supports the notion that IBM Planning Analytics can serve as a valuable tool for forecasting and decision-making processes in the context of the study, providing accurate and meaningful insights for effective planning and resource allocation.

The specific data set used to evaluate the developed application on IBM Planning Analytics for forecasting and budgeting purposes exhibited a lack of unusual trends, resulting in favorable outcomes. Concurrently, the budgeting and forecasting processes employed by market segments within the large-scale distribution industry aligned seamlessly with the application's guidance. This successful integration was made possible by Mediamente Consulting's profound expertise, which facilitated a comprehensive understanding of the business requirements and the needs of the application's end-users. The synergy between Mediamente Consulting's knowledge and the application's built capabilities ensured a harmonious interaction, meeting the demands of the users effectively.

6. Conclusions

The aim of the last chapter is the analysis of the results achieved within the scope of this thesis. Finally, possible next steps for this project are outlined.

6.1. Achievements

The successful realization of the application has met its intended objectives, resulting in the creation of a valuable solution that can be offered to Mediamente Consulting's customers. This solution caters to the needs of businesses seeking to incorporate planning and forecasting capabilities for their key performance metrics. By harnessing the power of their own business data, the proposed solution implements a structured and defined process that guides users through the creation of comprehensive plans and accurate forecasts for the upcoming years. Additionally, the application facilitates a final analysis stage, enabling businesses to compare and understand their progress over time. Data-aggregation level can be customized and this possibility makes the application a valuable tool in assigning tasks to specific roles and figures inside a firm, higher levels managers can assign sub-processes to lower levels managers improving then the overall planning and forecasting process.

In conclusion, this thesis has demonstrated the value and practicality of developing an advanced planning and forecasting application using IBM Planning Analytics. The integration of the forecast process into a typical set of CPM procedures has been reached excluding the necessity of forecast dedicated software, using an integrated planning and forecasting approach within a CPM solution. The successful implementation of the application showcases its effectiveness in helping businesses navigate the complexities of their planning processes and make informed decisions based on accurate forecasts. The obtained solution could be suitable for various use cases unrelated to the application sector. The high modularity allows to implement minimal changes to fine-tune the application for specific purposes, then most of the sectors could be reached. Regarding the scale of business, IBM Planning Analytics itself is mainly intended for bigger companies due to its licenses costs, then medium to large corporations are the main target but the ability to “develop once and redeploy” permits to lower costs and potentially offer in future the solution to clients with lower ability to spend.

The collaboration between Mediamente Consulting's expertise and the capabilities of IBM Planning Analytics has proven to be a winning combination, empowering businesses to gain deeper insights, improve their financial performance, and ultimately achieve their strategic goals. The results of this thesis highlight the importance of leveraging cutting-edge technologies and data-driven approaches to enhance business planning and forecasting processes, paving the way for future advancements in the field.

6.2. Further developments

An area of potential improvement in the developed application is the inclusion of clustering techniques before performing predictions. Currently, the application aggregates all product-level data belonging to a category into a single time series, which may result in the loss of important patterns and nuances inherent in individual product time series. To address this limitation, a promising approach is to apply clustering algorithms at the product level, grouping products with similar time series patterns together. By doing so, a more accurate forecast can be generated for each cluster, taking into account the unique characteristics of the products within that cluster.

This clustering-based forecasting approach offers several advantages. Firstly, it allows for a more granular and precise forecast at the product level, capturing the specific demand patterns and behaviors of each product group. This is particularly important when dealing with diverse product portfolios, where different products may exhibit distinct seasonality, trends, or other time series patterns. By capturing and leveraging these patterns, the accuracy and reliability of the forecasts can be significantly improved.

Furthermore, the concept of clustering can be extended to the store dimension within the application. Stores that exhibit similar aggregated seasonality or aggregated demand patterns across product categories can be grouped together, forming clusters of stores. By aggregating the demand data at the cluster level, valuable insights can be gained without sacrificing the information embedded in individual store-level data. This approach ensures that the forecasting process considers the shared characteristics and behaviors of stores, leading to more accurate and reliable demand predictions.

Incorporating clustering techniques into the developed application on IBM Planning Analytics opens up new avenues for enhancing the planning and forecasting capabilities. By leveraging the power of clustering algorithms, businesses can unlock valuable insights from their data and make punctual decisions. This future development not only improves the precision of forecasts but also enables businesses to better understand and utilize the inherent patterns within their data. The utilization of clusters in forecasting represents a significant step forward in harnessing the full potential of the application, offering businesses a comprehensive and tailored approach to planning and forecasting that aligns with their unique characteristics and needs.

Appendix A

			CCH Tagetik	Anaplan	IBM Planning Analytics	Board	Oracle EPM Cloud + Oracle Hyperion Planning + Oracle Analytics Cloud + Oracle Analytics Server + Oracle Analytics for Applications	SAP Business Planning and Consolidation + SAP 4/HANA + SAP Analytics Cloud + SAP BusinessObjects BI	Workday Adaptive Insights	Jedox
Forecast	Automatic choice of the algorithm that returns the most accurate forecast: the application must be able to search among the various algorithms implemented for the one that offers the greatest forecast accuracy and then choose that one as the forecast model	Automatic Algorithm selection		x	x	x	x	x		x
	Implementation of basic algorithms that are not composites of multiple algorithms such as exponential smoothing, moving average, seasonal naïve	Basic Algorithms		x	x			x		x
	Implementation of complex algorithms such as ARIMA, SARIMAX, neural networks, machine learning whose behavior is not always directly understandable and of which it is not always possible to reconstruct the choices and the path taken to obtain the forecast model	Complex Algortihms		x		x				x
	Generation of upper and lower bounds relating to each forecast point which outline a forecast window and not a single point	Upper & Lower Limits			x	x		x		x
	Seasonality recognition and management for forecast generation	Seasonality Management		x	x	x	x	x		x
	The application must be able, by selecting the time series of interest, to simultaneously generate a forecast for each selected series, e.g. by selecting more than 1 time series, for example the sales relating to two different products, the application must generate, simultaneously, a forecast for each product	Multiple forecast generation at the same time			x		x			
	The application must be able to recognize data that deviate significantly from the seasonal trend and the trend of the data series in question in order to be able to report them and possibly propose a correction	Outliers detection and correction			x	x		x		x
	Ability to define regressors that drive the prediction	Driver-based forecasting		x		x	x			x
	Possibility of creating alternative scenarios (actual, budget, forecast) aimed at what-if evaluations	Multiple scenarios management	x	x	x	x	x	x		x
	Spread of forecast data native to the application. The spread can be proportional, relatively proportional, and the data on which to base the spread should be selectable	Integrated data spread	x	x	x		x	x		
	Presence of prediction accuracy estimation metrics that allow to understand the goodness of the model and the resulting data	Confidence level and accuracy metrics		x	x			x		x
	The application must provide error metrics associated with the prediction made	Error metrics calculation		x	x	x	x			

Table 1 - Software Selection Table

Bibliography

Armstrong, J. S. (2001). *Principles of forecasting: a handbook for researchers and practitioners.* Kluwer Academic Publishers.

Chatfield, C. (2000). *Time-series Forecasting.* CHAPMAN & HALL/CRC.

Danese, P., & Kalchschmidt, M. (2011). The impact of forecasting on companies' performance: Analysis in a multivariate setting. *International Journal of Production Economics*(133), pp. 458-469.

David, F. R. (2017). *Strategic management: Concept and cases.* Pearson.

Dess, G., Lumpkin, G. T., & Covin, J. G. (2003). *Entrepreneurial strategy-making and firm performance: Tests of contingency and configurational models.* Strategic Management Journal.

Formula. (n.d.). *Scegliere il software di Corporate Performance Management adatto alla tua impresa*. Retrieved from Formula: https://www.formula.it/scegliere-il-software-di-corporate-performance-management-adatto-alla-tua-impresa/

Galt, J. (2019, 05 23). *3 Advantages and 3 Disadvantages of Forecasting*. Retrieved from John Galt Blog: https://johngalt.com/learn/blog/3-advantages-disadvantages-of-forecasting

Gartner. (n.d.). *Gartner Glossary*. Retrieved from Gartner: https://www.gartner.com/en/information-technology/glossary/cpm-corporate-performance-management?_its=JTdCJTIydmlkJTIyJTNBJTIyZTEyOGVkYWItYWU2YS00ZWZhLThmNjgtMjJmODhkNGVmMjBlJTIyJTJDJTIyc3RhdGUlMjIlM0ElMjJybHR%2BMTY5MzU2Mzg1Mn5sYW5kfjJfMTY0NjVfc2VvXzlhY2Iw

George, B., Walker, R. M., & Monster, J. (2019). *Does strategic planning improve organizational performance? A meta-analysis* (Vol. 79). Wiley Online Library.

Goodwin, P., & Wright, G. (2010, 05 07). The limits of forecasting methods in anticipating rare events. *Technological forecasting and social change, 77*(3), pp. 355-368.

Handik, W. (2005). *Analytical evaluation of top-down versus bottom-up forecasting strategies in a production planning environment.* Singapore: Doctoral thesis, Nanyang Technological University.

Hill, C. W., & Jones, G. R. (2010). *Strategic management theory: An integrated approach.* Cengage Learning.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice.* OTexts.

IBM. (n.d.). *IBM Planning Analytics*. Retrieved from IBM: https://www.ibm.com/products/planning-analytics

Jayawarna, S., & Dissanayake, R. (2019). *Strategic Planning and Organization Performance: A Review on Conceptual and Practice Perspectives.* University of Kelaniya.

Kotler, P., Keller, K. L., Brady, M., Goodman, M., & Hansen, T. (2017). *Marketing management.* Pearson.

Lawrence, M., Goodwin, P., O'Connor, M., & Önkal, D. (2006). *Judgmental forecasting: A review of progress over the last 25 years* (Vol. 22). International Journal of Forecasting.

Miller, C. C., & Cardinal, L. B. (1994). *Strategic Planning and Firm Performance: A Synthesis of More Than Two Decades of Research* (Vol. 37). Academy of Management Journal.

Mintzberg, H., Ahlstrand, B., & Lampel, J. (1998). *Strategy safari: A guided tour trough the wilds of strategic management.* Free Press.

Powell, T. C. (1992). *Organizational alignment as competitive advantage.* Strategic Management Journal.

Reeves, M., Ramaswamy, S., & O'Dea, A. (2022, 03 08). *Business Forecasts Are Reliably Wrong — Yet Still Valuable*. Retrieved from Harvard Business Review: https://hbr.org/2022/03/business-forecasts-are-reliably-wrong-yet-still-valuable

Rigby, D., & Bilodeau, B. (2013). *Management Tools & Trends 2013.* Boston: Bain & Company.

Ross, S. A., Westerfield, R. W., & Roberts, G. S. (2020). *Fundamentals of corporate finance.* McGraw-Hill Education.

Slack, N., Chambers, S., Johnston, R., & Betts, A. (2019). *Operations and process management: Principles and practice for strategic impact.* Pearson.

Wolf, C., & Floyd, S. W. (2017). Strategic Planning Research: Toward a Theory-Driven Agenda. *43(6)*, pp. 1754–88.

Zotteri, G., Kalchschmidt, M., & Caniato, F. (2005). The impact of aggregation level on forecasting performance. *International Journal of Production Economics*, pp. 479-491.