المجلة الدولية للأداء الاقتصادي
International journal of economic performance
ISSN: 2661-7161 EISSN:2716-9073
Publications are Open Access in this journal

Using Microsoft Power BI for sales forecasting as a data mining technique

Laifa Assala*	Hadouga Hassiba
assala.laifa@univ-constantine2.dz	hassiba.hadouga@univ-constantine2.dz
laboratory GMES university of constantine2 (Algeria)	University of constantine2 (Algeria)

Submitted:03/03/2023 Accepted:18/04/2023 Published:04/06/2023

Abstract:

This study aims to predict the sales of a commercial organization in order to know the role that modern information technology plays in achieving accurate and rapid processing of data based on the data mining tool represented in the Microsoft Power BI business intelligence program, through a theoretical and applied study. The significant role played by the estimated future sales information in the planning process as well as guiding and rationalizing the decisions of the sales manager to improve the performance of the organization.

Key words: sales forecasting, data mining, business intelligence, Microsoft Power BI.
JEL Classification Codes: C13 ; E2.

*Corresponding author

Introduction:

Economic forecasts are considered one of the important topics in most planning fields, especially since most institutions operate in a competitive environment, which necessitates the adoption of modern scientific and technological methods to predict the future status of the institution, which in turn calls for information-aware management, given that information today is a major resource for any institution, regardless of the nature of its activity.

Sales forecasting is an essential way for the organization to achieve its long-term goals, through a deep understanding of the business and facilitating the planning process by developing a marketing strategy that helps in optimal sales management based on the sales expected to be obtained in the future.

With the development of information technology and the development of mathematical methods, data mining techniques, especially modern ones, have contributed to building accurate forecasts in an automatic way, depending on the historical and current time series of the organization, through smart systems, artificial intelligence programs, and others such as business intelligence tools that specialists are keen to develop continuously in line with The needs of each field, perhaps the most prominent of which is Microsoft Power BI, which has been widely applied in decision-making processes

Accordingly, we devoted the study to an Algerian commercial enterprise to come up with a clear vision to formulate the basic problem represented in:

How effective is the application of data mining technology to predict future sales?

Which resulted in several partial questions, as follows:

-What is the importance of data mining for businesses?

-What are the most important methods used in the sales forecasting process?

-What are the requirements to do the forecasting process in Microsoft Power BI?

-Hypotheses:

In order to be able to answer the aforementioned questions, we relied on the following hypotheses:

The first hypothesis: data mining rationalizes business decisions;

The second hypothesis: quantitative methods are considered the best and most reliable means in the sales forecasting process;

The third Hypothesis: Microsoft Power BI processes all time series without specific limitations.

-Objectives of the study:

Through this research, we aim to reach the following points:

-Introducing data mining and its role in various fields in order to keep pace with technological development to improve the performance of institutions;

-Clarifying the importance of modern quantitative methods applied in the sales forecasting process;

-Developing our knowledge about the algorithms used by Microsoft Power BI in the sales forecasting process.

-The importance of the study:

The importance of the research lies in the need for sales managers to pay attention to the sales forecasting process to strike a balance between the needs of customers and the quantity that the organization must provide in a timely manner, using modern methods of information technology represented in various data mining techniques, especially advanced ones such as business intelligence programs, as well as knowledge of models included in the program in order to perform the forecasting process properly.

-Methodology applied:

In order to answer the questions and hypotheses that we formulated, we relied in this research on a combination between the analytical descriptive approach in the theoretical aspect related to the conceptual framework of data mining as well as the methods used in the sales forecasting process and the standard approach in the applied side when performing the organization's sales forecasting process by applying Microsoft Power BI program.

Theoretical framework of the study :

In this section we will look at the conceptual side of both data mining and sales forecasting.

First: Data mining:

Data mining is one of the recent topics that have witnessed great interest, whether by specialists or non-specialists in information technology sciences. Contributes to support the decision-making process in a timely manner.

1.The concept of data mining:

There are many definitions related to data mining due to the multiplicity of sciences that make up it, such as statistics, expert systems, machine learning programs, business intelligence, as well as artificial intelligence and other information technology sciences, as well as due to the many fields applied to data mining, and among the definitions that seem to us more appropriate are the following:

Data mining is a process based on interaction with the aim of discovering valid, useful, new and understandable patterns from large databases (Zerabi, 2021) that help provide appropriate information to support the decision-making process at the right time;

Data mining is the use of a set of algorithms to extract useful and previously unknown information from huge amounts of data, so it is considered the core of the knowledge eradication process (Nemiche, 2014/2015) ;

Accordingly, it can be said that data mining is based on a set of algorithms and programs for automatic detection of tacit knowledge in large, organized and unorganized data by processing it and extracting the relationships between the variables under study in order to benefit from them in the decision-making process to raise the effectiveness and efficiency of performance, and among the things that led to the development Ongoing data mining includes: (GK Gupta, 2014)

-Availability of data mining programs;

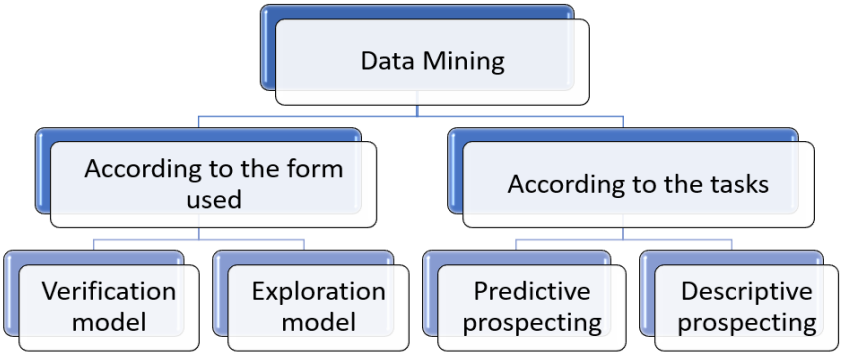
- the growth of data storage capacity and data volume;

- The competitive environment that requires continuous development in information technology;

2. Types of data mining:

The divisions differed by the researchers, each according to his point of view, but most of them fall into two important aspects: the tasks to be performed and the model used, as illustrated by the following summary figure:

Title of figure :Types of data mining



Source: Prepared by the two researchers, based on (Arun, 2008).

1- According to the tasks: it is divided into two main tasks:

-Descriptive prospecting: It focuses on finding patterns that describe the data related to the studied variables, analyzing and interpreting them, and presenting them to the user.

-Predictive prospecting: it provides future predictions based on the behavior of the phenomenon in the past.

2- According to the form used: it is divided into:

-Verification model: The focus in this model is on the user who formulates the hypothesis, because in this model the techniques test the hypothesis developed by the user in order to confirm or deny it.

- Exploration model: This model differs from the validation model in its focus, so that various technologies automatically discover useful information hidden in the data without user intervention or guidance. (Arun, 2008)

Second: Sales forecast:

In the past, institutional managers relied on personal intuition to predict sales, and the biggest example of that was Henry Ford when he launched the famous Ford "T" car when he believed and felt the sales that would be achieved without relying on an official device, and events proved the validity of his prediction, but this may be due to the fact that the investments were Less important, and equipment and machines are less specialized than it is today, and thus the lack of competition, so the technological development and economic maturity that the world is witnessing today has led to the necessity of making predictions by scientific methods keeping pace with technology by using tools and methods to extract knowledge in a timely manner, considering the field of error is limited. (Renaud, 1985)

Through the above, sales forecasting can be defined as determining the volume of estimated future sales based on the values achieved in the past, through techniques, including those that rely on scientific methods, and some that rely on the experience and intuition of users who perform the forecasting process, which can be explained as follows:

1. Sales Forecasting Methods:

It is divided into two basic types: qualitative techniques and quantitative techniques, so that one of the techniques cannot be judged to be effective unless it fulfills a set of conditions represented in: cost, providing the necessary data, accuracy, availability of the necessary material, human and moral capabilities to carry out the forecasting process .(Qadri and Ben Bozian, No year)

1.1. "Descriptive" qualitative methods: rely on intuition, emotions away from calculation, that is, in short, they can be said to be subjective rather than objective. This type of technique is recommended for those situations in which managers, or the sales team are particularly adept at forecasting sales revenue. Many of the methods include Among them:

- Expert committee method: This type of forecasting technique is called "top-down" as it is based on a meeting between senior executives and recording their opinions regarding future sales and considering it as a prediction of future sales values;

- Composite sales force: It is based on combining the individual expectations of sales representatives so that the sales representative makes an expectation for each product separately for his selling point and calls it the term “down to up”.(Martinovic & Damnjanovic, 2006).

1.2. Quantitative methods: It is represented in objective and mathematical techniques because it relies more on quantitative analytical methods and on past and current data(Mojekwu, Ganiyu, & ghomereho, No year)

From the previous definition of quantitative methods, it is possible to deduce a condition that must be available in the data under study, represented in the necessity of providing historical and current data for the phenomenon under study, in order to then apply one of the quantitative prediction techniques, which we mention in particular:

- exponential smoothing: a set of prediction models based on weighted averages of previous values in order to predict new values that combine: error, trend, and season. Therefore, it is called ETS (Franco, 2022), so that:

- Analyze the data to trend, level and seasonality: the level is the average value during the period under study

Trend: It is the direction to which the level turns, is it increasing, decreasing, or constant

Seasonality: expresses the type of data, is it monthly or annual...etc

The error consists in subtracting how much from the level, seasonality and trend from the actual values.

ETS consists of a large number of models, including:

-ETS (A.A.A): "Additive" error, "Additive" trend and "Additive" seasonality

It is used when the trend of the time series is linear with the presence of seasonality and it is called as Holt Winter's triple exponential smoothing model;

ETS (A.A.N) - "Additive" error, "Additive" trend and "No" seasonality:

It is used when there is no seasonality in the time series and when the error and trend are “additive”). It can also be used when the data has no trend (it is called the linear Holt model) (Sandeep, 2020).

-Data mining: One of the most important modern techniques that combined many of the previously mentioned quantitative prediction algorithms, such as time series analysis, classification and regression, but today's technological development calls for immediate sales forecasting, especially since not all managers are proficient in mathematics skills and The composition of algorithms, and accordingly advanced data mining techniques sought to facilitate this task because they may need to be trained on how to analyze data by knowing how to use artificial intelligence and machine learning and business intelligence tools that contain most of the prediction algorithms implicitly to improve the process of extracting knowledge to be accessed To it (Campos, 2020), among which we mention:

-Artificial intelligence: Artificial intelligence programs are relied upon by many institutions at the present time, especially in developed countries, and this is due to the following reasons: (Menon, 2022)

• Forecasting sales in a timely manner: due to fluctuations in demand for products and services, the organization's artificial intelligence works by predicting customer needs and behavior in a confidential manner, ensuring the availability of products when customers want them, and thus improving both the organization's performance and

customer satisfaction when good control over forecasting actual sales for the future period;

• Decision Making: Enables organizations to analyze huge amounts of data and provide timely answers, thus helping to plan and make decisions by seeing patterns and trends in the future based on past sales behaviors and results.

-Business intelligence: includes a set of tools and techniques based on machine learning that have been developed to predict the required variables through statistical modeling, mathematical calculation and simulation of results with the aim of supporting managers in the decision-making process, and among the most prominent business intelligence tools that have been widely applied in operations Business Decision Making Microsoft Power BI, Google Analytics, and Models can be accessed from the GitHub repositories. (Yili, Congdong, & Han, 2022).

-Applied framework:

Forecasting future sales is considered one of the main resources for commercial enterprises, as are human and financial resources, given that the information obtained from the forecast contributes to the establishment's achievement of the largest possible revenues, due to its knowledge of the future volume of actual sales that customers need and providing them with it in a timely manner. The forecasting process requires a comparison between the many technologies available in the field of forecasts and choosing the most appropriate method that saves time, effort and costs. Theoretical study on an Algerian commercial enterprise based on the historical data available to it on the volume of actual sales in the period from 2012 to 2022, expressed in Algerian dinars, based on the Microsoft Power BI business intelligence program in predicting the future sales of the enterprise and analyzing its position.

First: Introduction to Microsoft Power BI :

One of Microsoft's programs is a business intelligence platform that includes many mathematical and statistical methods, artificial intelligence, and machine learning algorithms, combining advanced analytics with an intuitive user experience.(Microsoft, No year)

The prediction model used in Microsoft Power BI: According to Sandeep Pawar, a specialist in data analysis and data science, Power BI uses the ETS(AAA) and ETS(AAN) models to predict seasonal and non-seasonal data, which have been dealt with in the theoretical aspect so that it works according to The type of time series to be

predicted automatically, and to obtain accurate results in the forecasting process, some program limitations must be taken into account, which are:

-The trend must be linear and not basic or weak;

- It predicts monthly, quarterly, semi-annual and annual data;

-The time series must not be less than 36 points for monthly data and 12 points for quarterly data, and the upper limit is not specified;

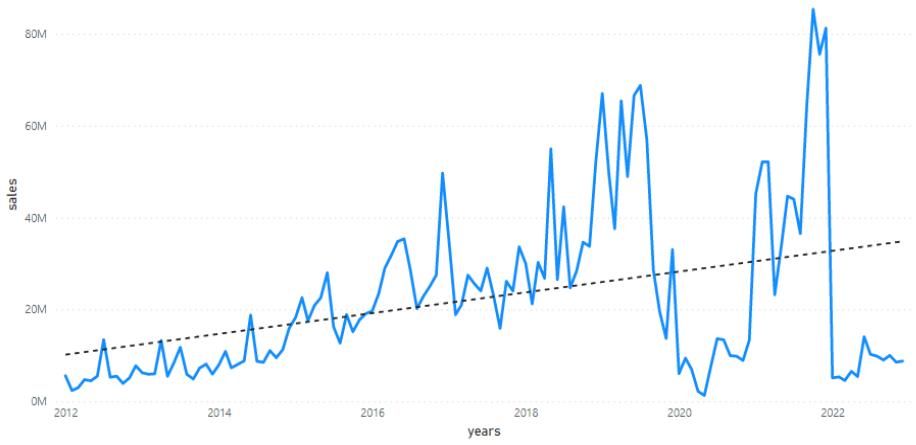
-It detects outliers by monitoring the trend and adjusts forecasts accordingly;

- It takes into account only one variable, and does not address additional external variables that may affect the phenomenon. (Sandeep, 2020)

Second:The time series of the actual sales of the commercial establishment from 2012 to 2022:

The historical time series is represented in the actual monthly sales of the organization from 2012 to 2022, which we express in a graphic curve as follows:

Title of figure: The evolution of actual sales from 2012 to 2022



Source: Prepared by the two researchers based on Microsoft Power BI

From the graph, depending on the actual sales of the organization in each month, we notice the general trend of the time series appears as a slightly elevated linear

So that we see fluctuations in the values, sometimes low and sometimes high, from one month to another, so that the amount of actual sales recorded an increase from the month of January 2012 with a value of 5 575 600A.D until the month of July 2019, when sales in this month amounted to 68 764 900 A.D., and then began to decline To reach its lowest level with a value of 1 294 950 A.D in May of the year 2020, to return to a gradual rise again, reaching a peak in October 2021 to reach 85 336 000 A.D, to decline after that

by a significant value in January 2022, which was estimated at 5 123 200 A.D, and it returns to the increase and decrease in slight values throughout the year 2022.

From the analysis of the organization's sales reports and plan, and despite the fact that the general trend of actual sales is increasing, the reasons that led to fluctuation and failure to follow a steady pace in the period under study are as follows:

- Poor performance of the salesmen because they were unable to identify the needs of customers and sell the products, which led to a great fluctuation in the actual sales of the institution;

- Replacing the sales manager responsible for managing sales and activating groups of salesmen in the field four times in a two-year period from 2020 to 2022, which led to a change in the management map, so that we notice a significant drop in sales values during that period;

- The organization does not rely on quantitative methods in the forecasting process, but is satisfied only with the best sales manager and salesmen in determining expected sales;

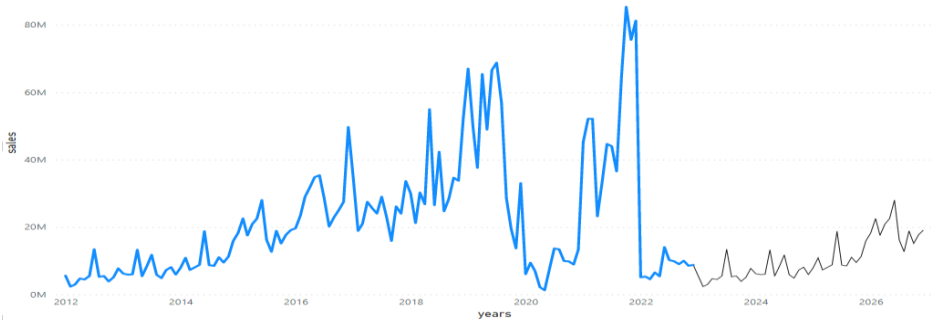
- The institution's failure to follow a marketing policy that responds to modern marketing methods and strong competition by institutions that are active in the same field, which led to a decline in demand from customers from time to time.

Third: Sales forecast from 2023 to 2026:

Since the time series is monthly, exceeds the minimum of 36 sales points, its trend is linear and does not contain outliers or exogenous variables, the limitations of applying the ETS(AAA) model in Microsoft Power BI are realized in order to give accurate forecasts of future sales.

And based on the aforementioned historical data, we will determine the values expected to be obtained in the future, with the aim of developing a cognitive vision for the institution about the situation that it will reach if it follows the same path of its sales management in the next four years, after downloading the historical time series of the actual sales of the institution in the program and at a level of confidence 95% We were able to obtain the following results:

Title of figure: Historical actual sales and expected sales from 2023 to 2026



Source: Prepared by the two researchers based on Microsoft Power BI

From the results of the ETS (AAA) model, we see a discrepancy in the sales values expected to be achieved in the next four years, but in general they started to rise from the level they reached in the year 2022, especially in the years 2025 and 2026, but this does not cover the expected performance of the institution in achieving sales The future is lower than what it achieved in previous years, especially in 2018 and 2021, and it may be due to its impact on the low level of sales that it reached in 2022.

Accordingly, if the establishment continues to follow the same commercial policy, then the sales values according to the ETS (AAA) model will be as explained above, and although the expected sales are trending high, the establishment must review the performance of the salesmen because selling requires special skills and competencies, the quality of the products that You display it and if it is in line with the needs of customers, as well as the appropriateness of the marketing and promotional method to what the market has reached in modern marketing strategies.

Conclusion:

From the foregoing, the most important findings we have reached can be summarized in the following points:

First: On the theoretical level:

- Data mining is applied in many areas, especially in the commercial field, in order to extract knowledge that helps in the decision-making process in a timely manner to improve the performance of the enterprise, This confirms the first hypothesis that states on "data mining rationalizes business decisions";

- Sales forecasting is considered the basis for measuring the efficiency of a commercial enterprise's performance by comparing expected and achieved sales and identifying errors;

-The technological development that the world is witnessing today requires carrying out the forecasting process in contemporary quantitative ways using data mining techniques such as machine learning, business intelligence, etc., when it achieves a set of conditions represented in cost, accuracy, providing the necessary data, availability of the necessary material, human and moral capabilities to carry out the forecasting process, and this This negates the validity of the second hypothesis, which states that "quantitative methods are the best reliable method in the sales forecasting process.".

Second: at the applied level:

- Business intelligence programs contribute to the forecasting process in a short time without specializing in information technology, as Microsoft works to make dealing with smart systems directed at all administrative levels through the Microsoft Power BI program;

- The Microsoft Power BI program works on processing the data and performing the sales forecasting process using exponential smoothing through the ETS (AAA) and ETS (AAN) model, and by imposing a set of constraints that must be taken into account in order to obtain accurate forecasts, and this negates the third hypothesis that states on "Microsoft Power BI processes all time series without specific restrictions" ;

- By analyzing the sales plan of the organization, it is clear that the sales manager has a major role in directing the salesmen and maintaining achieving a stable or high pace in sales values. marketing policy from time to time;

- The institution's lack of interest in quantitative forecasting methods that help in the decision-making process objectively and accurately, and its restriction to using qualitative methods only may lead to poor vision and performance, especially since the functional structure of the institution is not stable;

-One of the drawbacks of the Microsoft Power BI program is that in the forecasting process, it processes according to one variable and does not take into account the internal and external variables that may affect the expected sales values, through which the organization may discover the reason that led to a decrease or increase in its sales, such as environmental and marketing variables and costs. changing...etc.

Referrals and references:

- Arun, K. (2008). *Data Mining techniques.* hyderabad, India : universities press.
- G.K, G. (2014). *Introduction to data mining with case studies.* PHI Learning.
- Nemiche , M. (2014/2015). *Data Mining.* Agadir, Faculté des science d'Agadir, Morocco.
- Renaud, d. (1985). *La prévision des ventes.* Paris, France: Presses universitaires de france.
- Zerabi, S. (2021). *Contribution à la fouille de données volumineuses: Validation du clustering de données dans un contexte Big Data* (Thése de doctorat). Constantine, Département de L'informatique Fondamentale et ses Applications (IFA), Algeria.
- Martinovic, J., & Damnjanovic, V. (2006). THE SALES FORECASTING TECHNIQUES. *INTERNATIONAL SCIENTIFIC DAYS(4).*
- Mojekwu, J., Ganiyu, A., & ghomereho, O. (No year). A Review and Application of Quantitative Sales Forecasting Techniques. Lagos, Department of Actuarial Science,University of Lagos, Nigeria.
- Qadri, R., & Ben Bouziane, M. (No year). Sales Forecasting Models ALGAL Aluminum Case Study. *North African Economics Journal*.
- Yili , C., Congdong , L., & Han , W. (2022). Big Data and Predictive Analytics for Business Intelligence: A Bibliographic Study (2000–2021). *Forecasting 2022, 4*.
- Campos, L. (2020). *A Complete Guide to Data Mining and How to Use It.* Retrieved from blog.hubspot: https://blog.hubspot.com/
- Franco, D. (2022). *Exponential smoothing methods for Time Series Forecasting.* Retrieved from encora: https://www.encora.com/
- Menon, V. (2022). *How to Use AI to Improve Sales Forecasting Accuracy?* Retrieved from geekyants: https://geekyants.com
- Microsoft. (No year). *Présentation (Présentation de Power BI)(Présentation de Power BI)(What is Business Intelligence)*. Retrieved from Microsoft Power BI: https://powerbi.microsoft.com/
- Sandeep , P. (2020). *Time series Forecasting in Power BI.* Retrieved from pawarbi.github: https://pawarbi.github.io/