Power BI and Time Series in Budget Management of Business Support Services for a Large Brazilian Company

Geraldo Batista da Silva[1], Allan César Silva[1], Priscila Paloma Malta de Medeiros Coelho[1], Luiz Melk de Carvalho[1], Igor Júlio Pimenta[1], Diva de Souza e Silva Rodrigues[1], Flávio Henrique Batista de Souza[1], Vladmir Alexei Rodrigues Rocha[2]

[1]Centro Universitário de Belo Horizonte UNIBH, Brazil

[2]Universidade Federal de Minas Gerais UFMG, Brazil

Abstract.

The strategic use of data is one of the prerogatives of the wealth of a company. This premise is true, especially in large companies and in critical periods, as in the pandemic of COVID-19, where each decision is crucial for the organization. Thus, this research demonstrates the case study of a data analysis, with the Power BI® tool, focused on the budget management of business support services of a large Brazilian company. Information demands, the allocation of relevant data and forms of connection between datasets that could generate relevant information were analyzed. Relevant issues presented themselves as hindering the process, demanding strategies for resolution, such as: non-standardized reports; reporting errors; integrations; external reports; access restrictions. In summary, the results obtained were: an objective analysis methodology compatible with Enterprise Relationship Planning (ERP) tools available on the market with results in .xls, .xlsx and .csv formats; the diagnosis of an optimized control of 3.2 million Brazilian reais for the company under study, showing savings of up to 354.59 thousand Brazilian reais; a trend analysis, based on time series (ARIMA) was carried out, in order to demonstrate viability and directions to the business, with millions of reais to be managed and invested, with a prediction between 3 to 6 months.

Keywords: Business Support Services, Business Intelligence, Power BI®, ARIMA, Decision Making.

1. Introduction

In recent years, many companies have become aware that they increasingly need technology to excel in such a competitive market. Through the implementation and use of Business Intelligence (BI) tools, it is possible to correct and optimize resource management for business support services. In line with this concept, the use of time series can bring not only the mastery of the behavior of the actions of a company, but also to outline expectations of values of expenses and returns.

For Lira et al (2019), the contemporary world requires companies to be able to meet market demands in a timely manner and constantly prospect scenarios, making them increasingly competitive in order to guarantee the development and the permanence in the current scenario. After all, the pressure of the market falls on management and demands a fast process of capturing, integrating and interpreting. Therefore, software that performs this function is very useful in a company. According to Gartner (2018), the concept of BI is based on the transformation of data into information, then decisions and, finally, actions. The use of BI can optimize work processes, budget management and business support services. What characterizes a valuable strategic support, since the lack of accuracy in the budget projection generates unexpected expenses, in the same way that increasing the budget also impacts the freezing of current assets, consequently investment is lost (Ain et al., 2019).

According to Božič & Dimovski (2019), amidst almost unrestricted access to information, it becomes necessary to be able to filter it, and BI can help in two fundamental issues: reduce costs and increase revenues. With BI as an available resource, in addition to an efficient data management, analyzes with a focus on time series can be made feasible.

Pereira et al. (2019) infers that "time series can be conceptualized as a set of sequential observations of a given variable, expressed in numbers and obtained in a regular period of time". Through them, it is possible to trace behavioral trends, resulted from the analysis of current values in a given period.

Thus, the main objective of this research is to demonstrate a real case study of a process of collecting, analyzing and optimizing data availability, so that they would optimize the budget management process of the business support services of a large company. These tools focused on transforming raw data into coherent and relevant insights for the context of the company, which cooperates with decision making and supports the evaluating trends in resource spending.

As specific objectives, it aims at first, to make a diagnosis of the environment and analyze how separated and how dispersed is the data and how expensive is this analysis for the business. Then, perform an analysis of how a data relocation technique and methodology would be helpful in order to generate competitive reactions. And also, analyze how the data availability optimization process would impact the business. Finally, an assessment, supported by time series, with modeling based on ARIMA (autoregressive integrated moving-average model), to predict budget trends in the expenditure plan. This work is justified, because thousands of Brazilian reais (or even millions) can be lost in an inefficient management.

2. Methods

2.1 Concepts Evaluation

Two concepts were the basis for this work: BI and Time Series. BI is a term that combines architectures, tools, databases, analytical tools, applications and methodologies applied to analyze and cross a huge amount of information from various data sources, which support decision making (De Castro & Da Silva , 2018; Richards et al., 2019). Time Series, on the other hand, are a set of observations obtained over time, in a given period (Pereira et al,

2019). Chatfield and Xing (2019), stated that the classification of processes as to stationarity is the basis in the analysis of time series, such that a stationary series can be fully described by its mean, variance and correlation. According to the authors, the time series can be demonstrated as in equation 1:

$$\{y_k\}_{k=1}^{N_t} \tag{1}$$

Where k is the index and N_t the total number of samples. In this case, equation 1 is a realization of a stochastic process $\{Y_t\}$, composed of a sequence of random variables Y. A stochastic process is considered to be strictly stationary (probability distribution is invariant under time displacement) or stationary in the broad sense (or weak - if its average and covariance are invariant over time). Thus, the series average remains constant, with the addition of new samples, while in a non-stationary series it will present trends. For modeling time series, the ARIMA is widely used, which was the premise of this work.

2.2 Analysis Method

This case study presents a quantitative analysis and a scientific approach to managerial decision making. The approach starts with data, which is manipulated or processed into valuable information for those who make decisions. The actions taken in this were based on the following activities: an environment diagnosis and a data adaptation process were carried out in order to optimize them (the monitoring was carried out at company X from January to June 2019); an experimental process was carried out that analyzed the conjunction of the data service in order to generate strategic information; a return on investment analysis was obtained from the new data service organization. Such analyzes were presented graphically to optimize the representation of the information.

3. Results

In order to preserve the integrity of the information of the company, which is the source of this research data, fictitious nomenclatures were created for platforms. Systems, accounting classifications, branches and their respective cities where the company operates and also the name of the company, that is being referred to as Company X, were changed. A Base of different data sources was considered, such as:

- Three company partner platforms (Travel Service Providers): TPK, a travel booking platform (airline tickets, road tickets, accommodation, vehicle rentals and executive transportation services); XYZ Taxi, a municipal transport service operating throughout the country via APP; AMBER Executive, a national and international taxi transport service;

- A budget management platform called here ORCKY: ORCKY, during the budget cycle, is used to download and upload templates to fill in the expenses of the company, capex, social charges and revenue data. This is because the platform has a standard dataset format for input to its system (Table 1); subsequently ORCKY is used to post purchase requisitions, expenses and CAPEX (funds for the acquisition of capital goods); launches are automatically integrated from ORCKY to SUFIX, a

financial ERP integrating the budgeting / launching, tax, accounts payable, accounts receivable and accounting departments.

Table 1: Budget Model

| Code Cost Center	Cost Center	Observation	Jan/19	Feb/19	Mar/19
MG-101.1001	MG Matriz BH	Maintenance of the ERP SUFIX system			
RS-301.3001	RS Matiz Sul	Travel snacks and meals			

Source: (Authors)

- Then, DEPLIS, which is a platform for registering suppliers, customers, items and products, is also used to assist priority payment protocols, as well as payments in foreign currencies. The platform also integrates automatically with ERP SUFIX and of course with ORCKY, as this data is a reference for integration.

3.1 Processes Analysis

An analysis of the information generation process was performed (Figure 1). Firstly, it was identified the data sources and performed their extraction, according to their rules, such as, only extraction of series is allowed for the travel platform, over a period of no more than 90 days, which also applies to executive transportation platforms, but each with different dataset standards (reports), to the extent compatible with Excel ® software (xls, xlsx and csv).

Figure 1: Flowchart of the information generation process
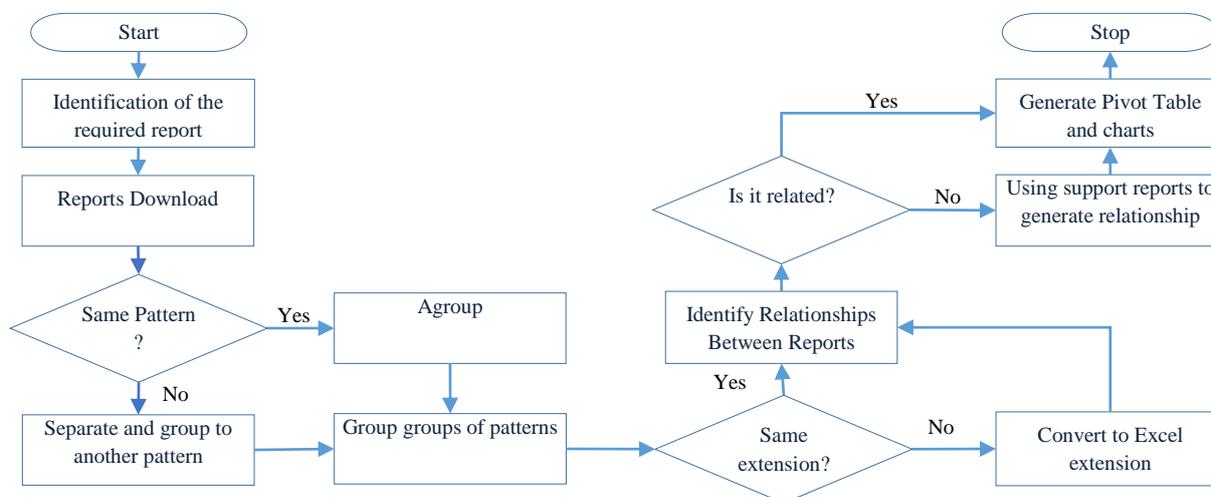
Source: (Authors)

As for the ERP system, the monthly extraction is adopted as a rule and in the text.csv extension, otherwise it takes a long time to download the reports (approximately 6 hours). Besides, as the system has a parameterization to disconnect the user when there is no action in the operation of the system for more than 2 hours, this procedure prevents the increase of the period. Then, the datasets that have the same standards are grouped, all are converted to the Excel extension, and the columns that are not necessary for the process are removed. In order to obtain the desired information, it is necessary to pass the reports through treatments generating relationships between them.

3.1.1 Complexity of Activities

The budget is built with an annual projection, but it has rules for quarterly use, that is, the funds that are not used in the current quarter, cannot be transferred to the next quarter, which is justified, because in theory it was not necessary. Thus, assertiveness is expected in the managers' forecasts, regarding the use of the resources during the period, because for any company to keep money in cash is a huge loss and the ideal would be to invest the money with a predetermined term, with redemption of it, only one day before the commitments foreseen in the budget cycle. However, in reality, this is not what happens, as it is very difficult to have this assertiveness in the face of difficulties related to: Errors in the Reports; Reporting Without Standards; Access Restrictions; External Reports (without integration); Integrations between ERP systems and platforms.

The datasets are diverse, which contain incomplete data, with little or no correlation between them, requiring specific treatments. This denotes that the analyst needs to have great knowledge of the procedures processes of the company. In view of the problems mentioned, the analyst needs to download from SUFLIX, several datasets that contain scattered and incomplete data, cross-check the information, mining this data in order to find or create relationships between them, building a dataset that presents a managerial view closing the use of the budget by periods, by cost centers, by companies, in short, making it understandable to interested parties (Figure 2).

The analysis is necessary so that the managers have information about the resources that still remain, facilitating the decision making, justifying with the history of the data, the budget composition for the next periods. The company uses this analysis to take preventive actions, improvements, in relation to the use of resources provided in the budget, assessing the successes and failures achieved by the company, by specific sectors or projects. They also serve to assess whether there are sufficient resources for their assumed commitments and whether it is possible to start other projects, which may be necessary in the process of improving results.

Figure 2: Data crossing diagram
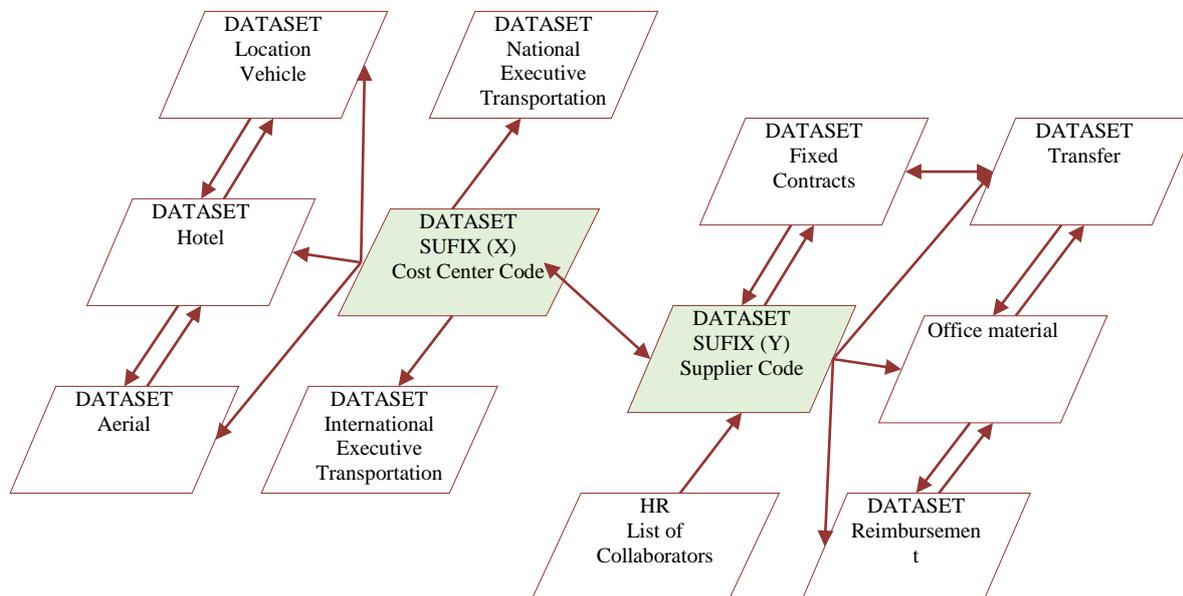
Source: (Authors)

3.2 Data Processing

An important phase of the budgeting process is the preparation of the different reports in the search for understanding, organization, treatment and control of the data obtained. This is necessary since each set of them, presents themselves with specific standards, some with unnecessary information for the results that are at this point, others are missing information, which is only found in a third report used to compose that information. For this study, it was determined that only the data on travel expenses will be used for the results presented in this work, so it is said that there will be five sources of data, presented as provided in Table 2.

In the TPK datasets it contains all the information needed for the result, however concatenated or in another format, making it necessary to use the resources of the Excel® tools, to carry out the modifications on them to adjust to the needs of the process proposed in this work. The XYZ and AMBER reports also contain this information, however with completely different standards and configurations from TPK, as do SUFIX and ORCKY. Thus, the first action is to study the datasets to understand the dispositions of the data. Defining how the data presentation standardization will be, the treatment must be performed using Excel® formulas and resources, eliminating the columns that are not necessary, creating others to cross data between the datasets.

Table 2: Data source relationship

Data type	Data source	Data type	Data source
flight tickets	TPK	Transfer	SUFIX - ERP Financial
Hotels	TPK	Reimbursements	SUFIX - ERP Financial
Rental car	TPK	Suppliers informations	SUFIX - ERP Financial
National taxi	XYZ	Cost center informations	SUFIX - ERP Financial
National and international taxi	AMBER Executive	Validation total	ORCKY - Budget platform

61

Source: (Authors)

Finally, various actions, resources and formulas are used in order to build a pattern, which is understood by managers and thus, facilitating decision making. Some of these actions taken to construct the patterns are shown in the following Tables. Note that the data presented contains the date and also the time, which it isnot needed in the process and because of that it is necessary to remove it. So, it was used the 'Text for Columns' feature, located on the 'Data' tab in Excel ®, to separate the date from the time and delete unnecessary information.

For the result, it is necessary to create a new column, called Deadline, which will provide information on how many days in advance the reservation was made in relation to the day of the trip, which is important for measuring the planning of the employee's agenda to fulfill his commitments. To check what is the budget impact, for this action, in the column created, it was performed the difference operation from the travel date column to the request date column, making it possible now to create another column containing a reservation advance range. In Table 3 it is possible to see an overview of how this part of the report was organized. After this implementation, it is possible to compare and analyze the average value of the reserves, which is an important indicator for the forecast of the budget for the next period, as shown in Table 4.

Table 3: Creating Anticipation Strip

Request Date	Deadline	Travel Date	Advance Range (days)
27 JAN 2019	17	13 FEB 2019	from 16 to 20
28 JAN 2019	18	15 FEB 2019	from 16 to 20
02 JAN 2019	12	14 JAN 2019	from 0 to 15
02 JAN 2019	14	16JAN 2019	from 0 to 15

Source: (Authors)

Table 4. Average Value Composition by Range

Advance Range (days)	Value	Mean Value
from 16 to 20	R$ 179.05	R$ 196.86
from 16 to 20	R$ 214.66	R$ 196.86
from 0 to 15	R$ 625.06	R$ 554.02

Source: (Authors)

The treatment, cleaning, organization and manipulation of the datasets, takes a long time and a lot of knowledge of the analyzed data to be able to extract information. Even then, only with Excel® and its various resourcesit is not enough to obtain efficient, fast and dynamic information visualizations, , because there are several sources and from very extensive databases, demanding thousands and even millions of lines in the software. Based on that, it is proposed to use the Power BI tool, as a BI solution, to facilitate the creation of relationships between the different reports.

3.2.1 Business Intelligence Application

The use of a BI solution does not mean that data processing will be eliminated from tasks. These actions will only be optimized and modified, for example, in the creation of specific datasets to generate relationships within the Power BI tool (see Table 5 and Table 6). It was noted that in Power BI a relationship was determined between the columns that contain the same data, making the visualization dynamic and optimizing the use of Excel® formulas and resources. Table 5 represents a dataset with only one column, which contains only one

occurrence of data from one of the columns of the dataset with all the information represented in Table 6. The relation of an occurrence of Table 5 was observed for many occurrences of Table 6. The relationship created between the datasets allows to generate clear visualizations, obtaining instant information for decision making. In Figure 3 it is possible to see the average value of the days bands in advance, as well as the percentage that it represents of the total value of the reserves.

Table 5: Dataset Advance Days Range Reference

Advance Range (days)
from 16 to 20
from 41 and above
from 0 to 15
from 21 to 30
from 31 to 40

Source: (Authors)

Table 6 - Dataset com Todas as Informações

| Advance Range (days)	Complete Trip	Airline	Total (R$)
from 41 above	BH/Singapura - India/BH	Emirates	31,921.50
from 16 to 20	São Paulo - Belo Horizonte - São Paulo	Azul	325.28

Source: (Authors)

Figure 3: Average Value of the Advance Range

Average Value of Advance Ranges

R$ 398.6(16.52%)
R$ 627.02 (25.98%)
R$ 409.0(16.95%)
R$ 450.39(18.66%)
R$ 528.55(21.9%)

Advance Range (Days)
From 0 to 15
From 16 to 20
From 41 above
From 21 to 30
From 31 to 40
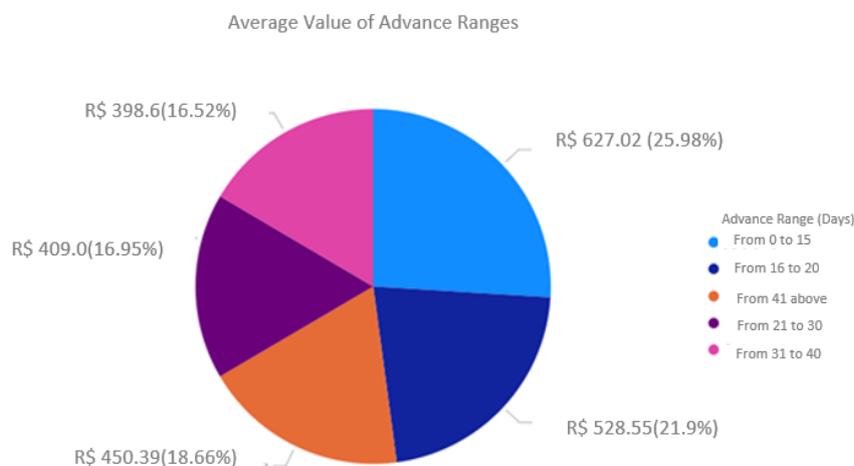
Source: (Authors)

Another advantage of using BI is to obtain several visualizations in different dimensions, where all the evolution of the information is observed, so that a more accurate and assertive analysis of the use of resources foreseen in the budget plan is possible. In Figure 4 it is presented the evolution mentioned. The visualization of Figure 4 allows the user to quickly analyze the average amount per band in advance, the average amount paid per airline, the total amount paid per month, a quarterly view of the expense and also the gross amount spent in the period from January to June 2019 , all of this on the same screen, thus streamlining the possibility of making a more assertive decision.

Figure 4: Evolution of Air expenses

63

Analysis of Travel Expenses from January to June 2019
Average value by advance ranges and main cities
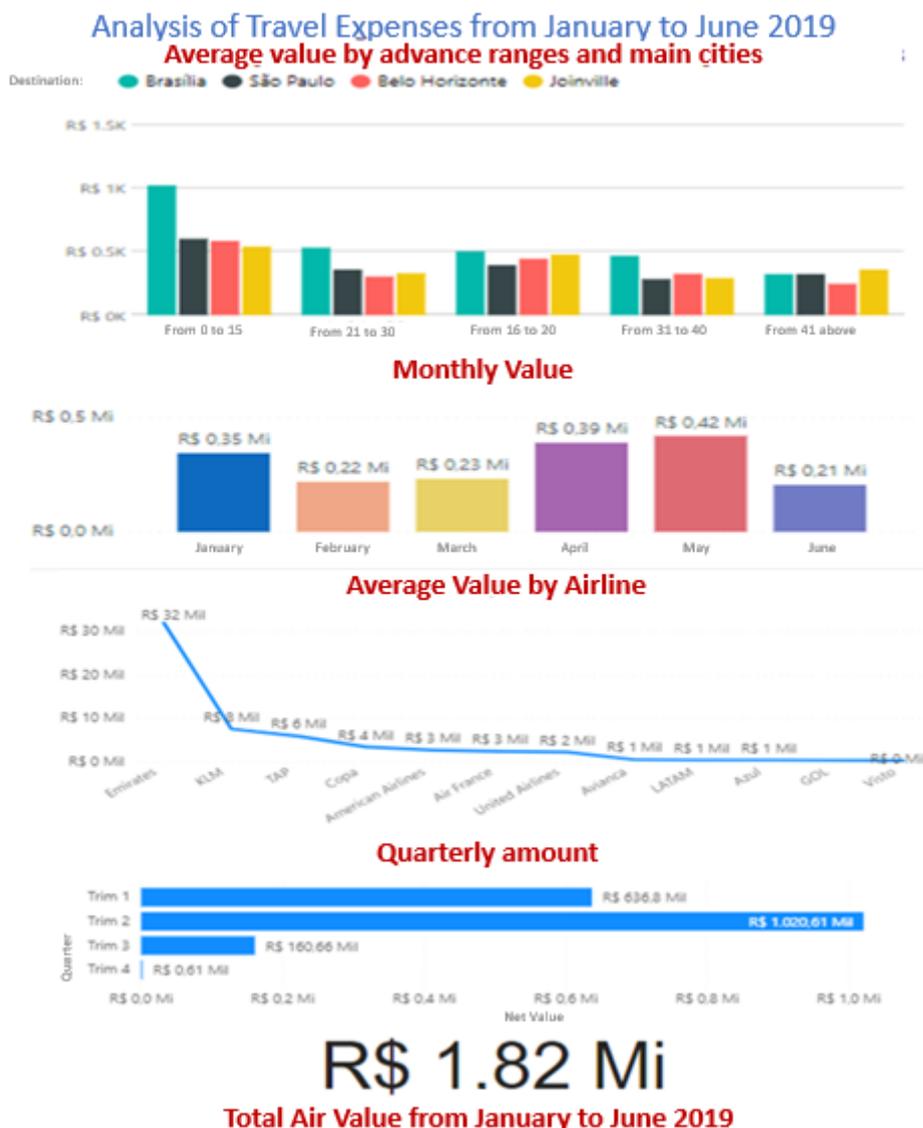
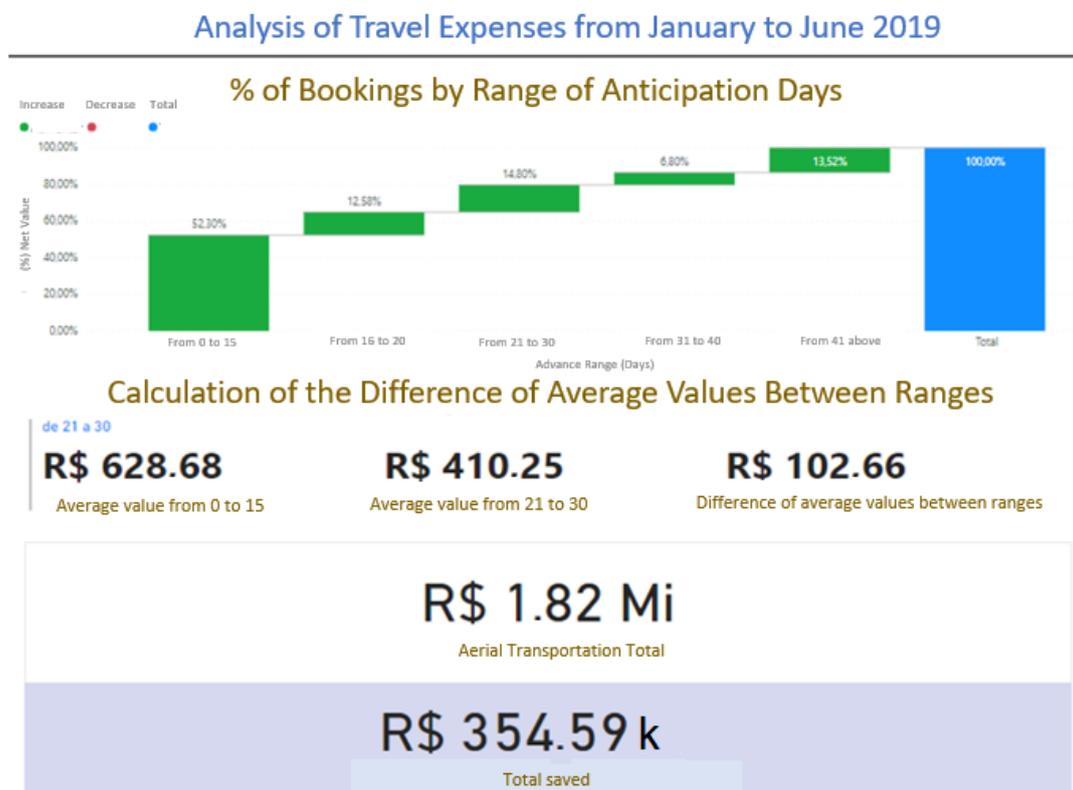*Mi = R$ 1 Million;*
Source: (Authors)

3.3 Financial Results

The financial results that can be obtained, with the possibility of better visualizing the information, are many. The following figures illustrate some of these cases and decisions that can be taken, from the point of view of crossing data and in various dimensions of analyze. Calculating the average value of the reserve anticipation range 'from 0 to 15', 'from 21 to 30' and making the difference between the two, as shown in Figure 5, quickly demonstrates that savings of 354.59 thousand of Brazilian reais would be achieved. Thus, if all reservations made in the period presented, were better planned and made in advance, between 21 and 30 days, there would be a reduction of approximately 20% in the gross total, which concludes that an action is needed to minimize this cost. Analyzing the expenses with accommodation, it is suggested, through the feasibility research in the central table Appendix I a.

64

The decision to hire the fixed service of corporate apartments, in the main cities, where the company's commitments take place, demonstrated significant rates of cost reduction. It is also clear that other analyzes are necessary, such as: operational cost in the management of reservations, distance between the place of lodging and the commitments of the employees. However, it is a vision that is necessary for business improvement options.

Figure 5: Simulation of Economy

Mi = R$ 1 Million

Source: (Authors)

Other analyzes that can be obtained are the macro dimension of the use of expenses. Appendix I b shows information of a high degree of relevance to the business. In the table on the upper right, it shows the total of each service for the period, on the upper left it brings an index that each type of service represents on the total expense, enabling the quantitative analysis of which one should take cost optimization action, that is, the most impactful for the business. In addition to the average values of each one of them in the lower left table, which is estimated to be the average tickets that will be used for the projection of the budget, for this type of expenditure, for the next period and finally the total amount, which is quite expressive for the company, since 3.2 million are believed that any manager would focus on planning and engagement actions, in order to reduce this expense as much as possible.

3.4 Experiments Via Time Series

An ARIMA time series experiment was carried out, using historical data, from September of 2017 to September of 2019, of air travel expenses of company X, object of the study, in

order to project a budget forecast of said expense, making inference on the possibility of automating the process of composing the company's entire budget. The experiment starts with importing and reading the data. At this point, some characteristics were observed in the data structure. It was identified that there were several lines with the same date and with the same value, which means, for example, that the same air flight and the same date was booked for more than one employee of company X, as a result of which some treatments were necessary , because the system did not recognize periods in the database, as well as a solution, the grouping was adopted adding the values of each of the duplicate dates.

The data was read again, but even so, the data set was not recognized as a valid period, because there were no travel reservations on all days of the month. Thus, it was opted for a new grouping based in weeks, as this would obtain the necessary number of weeks per year, which is 52, it was indicated, via code, for the program to group the days of the week and add the weekly values of them. A new reading of the data was carried out and it was found that it was recognized with valid periods, in this way the database was plotted again.

It was noticed when observing the graph on the x-axis that the data were random. This is because it was not indicated for the program which year and week the period starts, if the occurrence was identified, the frequency of 52 weeks was indicated, together with the year 2017 and week 35 as the beginning, in the function that transforms the data frame (column and row variables) into time series, which can be seen in the graph plotted in Figure 7. It was possible to observe that the the program carried out the transformation as indicated, dividing the periods by semesters, and between the periods 2018.5 and 2019.0 it can be inferred that the maximum value of approximately 29,000.00 and the minimum value understood after the 2019.5 period, but it is not possible to classify whether the time series shows trend or seasonality.

Thus, it is only possible to visualize that there is no change in the cycle and that it is a random series, but to project a forecast the program needs to decompose and classify this series. The periods are represented and classified on the y-axis in observed (original basis), trend (trend), seasonal (seasonality) and random (randomness), which validates the necessary conditions for the program to carry out the temporal projection. Before proceeding with the forecast, through ARIMA, it is necessary to run a line of code to validate the expected projection, with the objective of determining that it is projected 26 weeks (six months). With all the data and periods validated, in fact one can extract the forecast for the first six months of 2020 from the budget, as shown in Figure 6.

Figure 6: Plot of projected forecast for the first 6 months of 2020
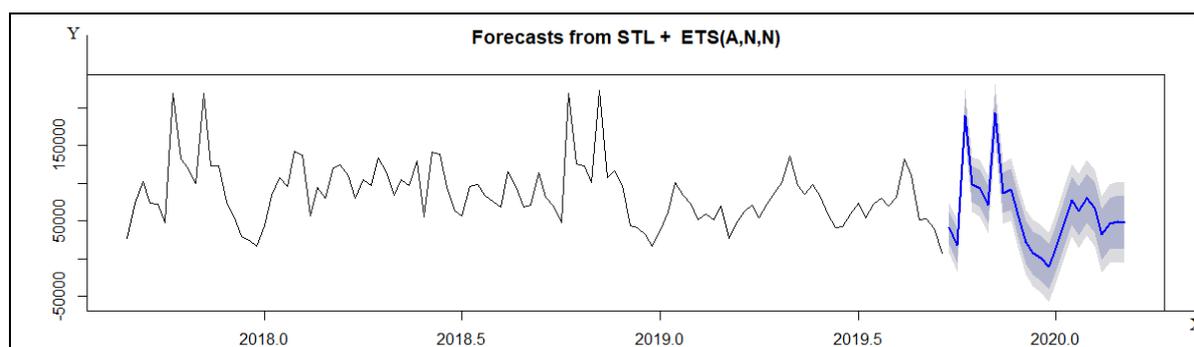
Source: (Authors)

The result shown in the graph in Figure 7 shows that the experiment can be used as a possible budget forecasting tool, as only air travel data was used, considering only the variables days and values, other experiments will be needed to determine the wide variety of variables and situations included in the budget process.

4. Conclusion

It was presented that the concept of BI is based on the transformation of data into information, then decisions and finally actions. In the results of the study, the evolution of this concept was shown, and it is possible to conclude that a set of specific practices and knowledge is necessary to generate information, which makes business management more visual, in terms of results. It is concluded that, with the use of Power BI®, it is possible to optimize time and minimize company costs, by reducing operations with the use of compiled data, producing statistical data with graphical visualizations in several dimensions where it is possible to observe all the evolution of information generating more hits and less expenses.

Finally, the use of ARIMA proved to be effective, with due prior analysis of the base to be analyzed. Analyzing the related works, it can be seen that, as a differential, the study presents the results quantifying the possible financial reductions in the budget and enabled the prediction of expenses by ARIMA.

References

[1] Božič, K., and Dimovski, V. (2019). Business intelligence and analytics for value creation: The role of absorptive capacity. *International journal of information management*, 46, 93-103.

[2] Ain, N., Vaia, G., DeLone, W. H., and Waheed, M. (2019). Two decades of research on business intelligence system adoption, utilization and success–A systematic literature review. *Decision Support Systems*, 125, 113113.

[3] Richards, G., Yeoh, W., Chong, A. Y. L., & Popovič, A. (2019). Business intelligence effectiveness and corporate performance management: an empirical analysis. Journal of Computer Information Systems, 59(2), 188-196.

[4] De Castro, L. M. & Da Silva, M. A. L. (2018) Business Intelligence (BI): Análise comparativa entre as ferramentas líderes no mercado. e-RAC, v. 8, n. 1.

[5] Gartner, I. T. (2018) *Glossary: Business Inteligence*.

[6] Chatfield, C., and Xing, H. (2019). *The analysis of time series: an introduction with R*. CRC press.

[7] Lira, A. C. D. Q.; Gomes, M. D. L. B. & Cavalcanti, V. Y. S. D. L. (2019) Uma reflexão em busca de uma configuração: estratégia empresarial, competitividade, estratégia de produção. *Revista Capital Científico-Eletrônica (*RCCʔe)-ISSN 2177-4153, v. 17, n. 1, p. 124-139.

[8] Pereira, A. F. R.; Sampaio, F. J. C. D. S. and Guilherme, H. F. (2019) Estimativa da arrecadação própria municipal: um estudo da previsão dos impostos de municípios paraibano e potiguar através das séries temporais. *Brazilian Journal of Development*, v. 5, n. 6, p. 5675-5699.

Appendix I – Savings Evaluation

Hosting Analysis from January to June 2019
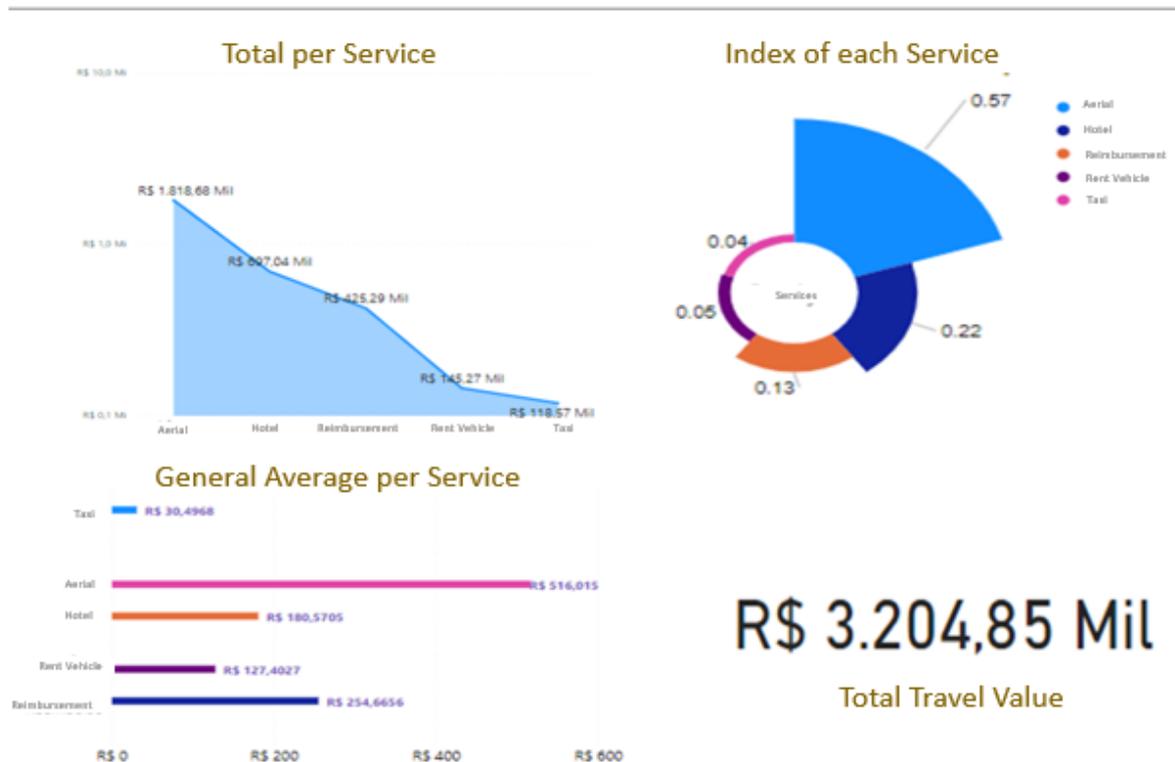#### Number of Hotel Nights by Main Cities

Cities	Total Hotel	Total Corporate	Savings
BH	R$ 74.304,97	R$ 46.946,44	R$ 27.358,53
JOI	R$ 75.227,08	R$ 51.455,23	R$ 23.771,85
SP	R$ 167.043,42	R$ 71.833,89	R$ 95.209,53
Total	**R$ 316.575,47**	**R$ 170.235,56**	**R$ 146.339,91**

a - Viability of corporate apartments

General Analysis of Travel Expenses from January to June 2019

b - Expense Overview
*Mi = R$ 1 Million ;Mil = R$ 1 Thousand

Source: (Authors)