Makting Area Campaign

Hugo Loureiro 14660

*Abstract* – The importance of data studies and filtering of these are vital for better decision-making in a future advertising campaign, where the campaigns to be made will be better restricted, thus restraining the target audience. Using data in different formats, ETL functions were applied where data were filtered based on certain previously defenible rules, so it is possible to have a much clearer perception of the information.

# Introduction

The marketing area is directly connected with the knowledge of customer profiles in order to be able to address in an effecient way the campaigns to be carried out. By knowing the profile of customers most likely to purchase a given product, organizations can restrict their campaigns by reducing the waste associated, for example, by sending promotional posters to customers who, at the outset, do not fit the profile of consumer customers of the product being marketed or the campaign being carried out.

The dataset integrates 500 records, associated with 500 customers who were targeted by a previous campaign. In the sequence of the campaign carried out, a certain set joined it, acquiring the promoted product, while another set of customers did not join the said campaign. By analyzing this data set, it is intended to identify the profile of customers who joined the previous campaign, in order to make them the privileged target of a future campaign. The main focus of this work will be the processing of the data so that they can be inserted in the final multidimensional model.

# Concepts

# DATA WAREHOUSE

The data warehouse is the center of the architecture for information systems. It supports informational processing, providing a solid platform for integrated historical data from which the appropriate analyses were made. Provides ease of integration into a world of unrooted application systems. These stores are achieved over time in an evolutionary process.11

# DATA MART

A Data Mart consists of a portion of an organization's data, and its scope is confined to a specific group or department. Data Mart can be independent or dependent on DW. The purpose of a Data Mart is to suppress the informational requirements of a part of the organization. Based on the bottom-up architecture, in development, data marts are derived from the data warehouse. In the top-down architecture, the data warehouse is derived from data marts. This theme will be further developed in the following sections of this paper.12

# ETL Process

Extract-transformation-loading (ETL) tools are pieces of software responsible for extracting data from various sources, cleaning, impersonation and insertion into a data store. Usually these processes should be completed in a certain time window, and in very cassos it is necessary to optimize their execution time.

# Business Intelligence

Business Intelligence is a data-driven Decision Support System (DSS) that combines data collection, data storage, and knowledge management with analytics to provide information to the decision process. Business intelligence highlights the analysis of large volumes of data about the company and its operations.

# 

# Pratical Work

The work was done based on the structure of Figure 2. Taking into account the topics discussed earlier and taking into account our problem, I decided to start by creating the external data sources, which would be the data sources that we would use for this whole project. I created an excel file, with the most varied fields, from correct to incorrect, and already with the appropriate base columns that would support the course of the work. A file has been.txt where they contain all the main roles to be respected when the ETL functions are performed.

Python functions were created (we use the Syder IDE), where respecting the rules in our.txt file, several wipes were done on the data, from removing null values to checking the range of possible values. These functions end in exporting the data to.xlsx The exported files were properly divided by what would later be the dimmense tables. Subsequently, the exported files were.xlsx to be in the right way to be inserted into our star layout.

A picture containing diagram

Description automatically generated

Fig. 2: Architecture

The highest level of granularity is 3 (Campaign - Client - Genero). We have 1 fact table (Markting Campaign) and 10 dimensions (Car, Customer, Genero, CustomerChildren, CreditoHabitacao, Locality, AccountDeadline, Expiration, ContaOrdem, ClienteFamilia). The star model is represented in fig 3.Diagram

Description automatically generated

Fig. 3: Star Model

For the visualization of the star model was used a very useful and also very simple tool, called Dataedo, where after a simple connection with our bd (postgreeSQL) we can quickly produce the E-R model visualized in fig 3.

For the final visualization of the data, we chose Tableau Desktop. The connection with postgreeSQL was simple, and we used this OLAP tool because in addition to being an easy and essentially quick way to visualize the data. This tool offers many features so that it is easy for users to use it. Most people do not need access to all the content softer from external data sources, but on the contrary they already need access (quickly and simply) to information about certain subjects, departments, topics, depending on their work functions.

Tableau helps you create data that can be understood by professionals at any level in an organization, as well as help simplify raw data in a very easily understandable format. Considering our star scheme it is important to have a graphical view.

Diagram

Description automatically generated

Fig. 4: Relação de tabelas – Tableau

For reasons of time, we could not find any graph that fits the results obtained to show them, so we decided to approach a different tactic, we created an html page, in Laravel, so we can make a simple consultation, as we can see in the following figure.

Graphical user interface, text, application

Description automatically generated

Fig. 4: Web form – Makting Canpain

Fig. 4: Deaths and vaccination evolution

# Conclusion

This whole process from data extraction from external data sources to ETL functions are very important for the continuous evolution of data warehouses, the constant evolution of business requirements, different business perspectives, etc.. there is a continuous development of ETL for all types of operations. It is on the basis of these processes/functions that we can access a small but veridica amount of data that we need for a given problem. Despite the multiple OLAP solutions available on the market, we chose Tableau because it's simple, it has a complete set of tools that we can use, thus providing faster delivery.

### References

1. Inmon, W. H. (2000). *WHAT IS A DATA WAREHOUSE?* William H. Inmon.
2. . Naeem, T. (3 de 2 de 2020). *Conceitos de data warehouse: abordagem Kimball vs. Inmon*. Obtido de Astera: https://www.astera.com/pt/type/blog/data-warehouse-concepts/
3. Negash, S., & Gray, P. (2008). Business Intelligence. Obtido de Springer Link: https://link.springer.com/chapter/10.1007/978-3-540-48716-6\_9
4. Simitsis, A., Vassiliadis, P., & Sellis, T. (18 de April de 2005). Optimizing ETL processes in data warehouses. Obtido de IEEE Xplore: https://ieeexplore.ieee.org/abstract/document/1410172
5. Vassiliadiss, P., Simitsis, A., & Skiadopoulos, S. (November de 2002). Conceptual modeling for ETL processes. Obtido de Association for Computing Machinery: https://dl.acm.org/doi/abs/10.1145/583890.583893
6. What is Tableau? Uses of Tableau Software Tool. (2021). Obtido de Guru99: https://www.guru99.com/what-is-tableau.html