Lisbon School of Economics & Management
Universidade de Lisboa

MASTER

DATA ANALYTICS FOR BUSINESS

MASTER´S FINAL WORK

INTERNSHIP REPORT

BUSINESS INTELLIGENCE IN DECISION-MAKING: IMPACT ON CUSTOMER ENGAGEMENT

PATRÍCIA ALEXANDRA DA SILVA ANTÃO

MARCH – 2023

Lisbon School of Economics & Management
Universidade de Lisboa

MASTER
DATA ANALYTICS FOR BUSINESS

MASTER´S FINAL WORK
INTERNSHIP REPORT

BUSINESS INTELLIGENCE IN DECISION-MAKING: IMPACT ON CUSTOMER ENGAGEMENT

PATRÍCIA ALEXANDRA DA SILVA ANTÃO

SUPERVISION:
PROF. WINNIE NG PICOTO
ANDRÉ SILVA

MARCH – 2023

GLOSSARY

BI - Business Intelligence.

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

CSF - Critical Success Factor.

ETL – Extract, Transform and Load.

HnB – Heat not Burn.

KPI – Key Performance Indicator.

MFW – Master's Final Work.

PMI – Philip Morris International.

ABSTRACT

Business Intelligence has become an essential aspect of modern business management, providing organizations with valuable insights into their operations and performance. This Master's Final Work focuses on leveraging BI to improve communication strategies for a business organization.

The internship described in this Master's Final Work involved working on a Business Intelligence project in a company called Tabaqueira. The primary objective of the project was to optimize and improve communication via email, with a focus on enhancing customer engagement and increasing revenue growth.

This study provides a comprehensive overview of the entire BI project, from the theoretical background to the implementation and reporting phase. It covers the various BI tools and techniques used to analyze the available data, from the data source to all the Extract, Transform and Load (ETL) processes that allowed the data to be ready to be explored. The reporting phase includes the presentation of findings using Power BI and recommendations to the stakeholder and the evaluation of the project's success.

The solution provided valuable insights into customer behavior and preferences, allowing the design of communication campaigns to be more relevant and personalized to individual customers.

KEYWORDS: Business Intelligence; Data; Decision-making; Consumer Behaviour; Campaigns.

RESUMO

Business Intelligence tornou-se um aspeto essencial da gestão de empresas modernas, proporcionando às organizações conhecimentos valiosos sobre as suas operações e desempenho. Este Trabalho Final de Mestrado foca-se no uso da BI para melhorar as estratégias de comunicação de uma empresa.

O estágio descrito neste Trabalho Final de Mestrado envolve o trabalho num projeto de *Business Intelligence* na Tabaqueira. O principal objetivo do projeto era otimizar e melhorar a comunicação via e-mail, com foco no aumento do envolvimento do cliente e no crescimento das receitas.

Este estudo fornece uma visão abrangente de todo o projeto de BI, desde o enquadramento teórico até à fase de implementação e elaboração de um relatório. Abrange as várias ferramentas e técnicas de BI utilizadas para analisar os dados disponíveis, desde a fonte dos dados a todos os processos extração e transformação que permitiram que os dados estivessem aptos a ser explorados. A fase de relatório inclui a apresentação dos resultados recorrendo ao *Power BI* e as recomendações ao *stakeholder* e a avaliação do êxito do projeto.

A solução forneceu valiosos elementos sobre o comportamento e preferências do cliente, permitindo a formulação de campanhas de comunicação que sejam mais relevantes e personalizadas para clientes específicos.

PALAVRAS-CHAVE: *Business Intelligence*; Dados; *Decision-making*; Comportamento do Consumidor; Campanhas.

TABLE OF CONTENTS

TABLE OF FIGURES

ACKNOWLEDGMENTS

This MFW is an important step in my life, as it marks the end of a journey. All the steps that brought me here would not have been possible without a set of very important people who were always by my side unconditionally, and I could not go without showing them my gratitude.

First, to ISEG – Lisbon School of Economics and Management, which has been part of my life for 5 years and has seen (and made) me grow and contributed a lot to the person I am today. In this scope, acknowledge, obviously, my supervisor, Professor Winnie Picoto, who was always there for all my doubts and difficulties with her experience and knowledge and supported me even when everything seemed more difficult and distant. Thanks also to Tabaqueira for allowing me to be part of it and for the opportunity to learn from the best.

Secondly, my parents, to whom I owe everything I am today, and all the opportunities they gave me that made me get this far. To my mother, for all the unconditional love and affection, for always having a word of comfort for me, even on more difficult days. To my father, for the wisdom and education he passed on to me and for always having a comforting look in his eyes even when he was short of words.

Third, to my dear Rita, without whom I can no longer live, for all the friendship and for always believing in me even when I doubted myself.

Lastly, to the one who is always by my side. Bruno, thank you for allowing me to grow and learn from you, for your patience at all times, and for all the love that you give me each and every day.

Carolina, Gonçalo, and Mariana thank you for being my support and for being my safe zone. I am very grateful for the opportunity to building memories with you.

1 INTRODUCTION

As a second-year Master's student in Data Analytics for Business, my decision to undertake an internship as my final project was based on the opportunity to apply the knowledge and skills I had acquired. The internship discussed in this master's final work document was established as a collaboration between ISEG - Lisbon School of Economics and Management and Tabaqueira S.A.

The present study aimed to explore the impact of Business Intelligence on business performance at Tabaqueira. Its main goal was to acknowledge how BI can be helpful to Tabaqueira to gather, analyze, and utilize data to make informed decisions and improve performance (Mallam et al., 2021).

With the increasing availability of large amounts of data, organizations have come to recognize the importance of utilizing BI to gather valuable insights and gain a competitive advantage (Khan et al., 2020). In recent years, many organizations have adopted Business Intelligence Systems with the objective of gathering and analyzing data to monitor operations and evaluate firm performance (Bălăceanu, 2007). European Commission expects that the value of data economy will almost triplicate by 2025, reporting that organizations that create innovations based on data show a productivity growth of 5%-10% (European Commission, 2022).

Several industries have recognized the potential of Business Intelligence to support the accomplishment of business goals, including but not limited to Education, Healthcare, Tourism, among others (Hasan et al., 2016). Business Intelligence has proven to be a significant factor in renovating organizations by transforming business data into actionable insights for the benefit of both the organization and its stakeholders.

2 LITERATURE REVIEW

2.1 Business Intelligence

Due to the increasing volume of data generated by organizations, the field of Business Intelligence (BI) has gained substantial importance (Bălăceanu, 2007).

Although the Business Intelligence (BI) term has been around since 1989, there's still no consensus about it, as it is a wide-ranging category of applications and processes including decision-making, collecting, and analyzing data (Wixom & Watson, 2010). Generally, BI is defined as a set of methodologies, technologies, and tools that are used to collect, access, and analyze data to support informed and data-driven decision-making and make more informed decisions (Turban et al., 2011). These methodologies include data mining, reporting, online analytical processing, and predictive modeling (Bălăceanu, 2007).

Business Intelligence enables companies to gather data from heterogeneous sources in order to support demands from multiple users – businesses, stakeholders, and consumers (Khan et al., 2020). For that reason, companies need to gather insights at a fast pace to support tactical decisions according to the strategy and planning of the company (Mallam et al., 2021). Therefore, the benefits of BI are only achievable if the information can be translated into insights and are relevant to support decisions (Popovič et al., 2012).

The use of BI is becoming increasingly important for organizations as the volume and the complexity of data grows, as also the need for more accurate and timely information (Khan et al., 2020). By providing access to real-time and accurate information, BI can support organizations in making informed decisions that can have a significant impact on their success (Bălăceanu, 2007). Furthermore, BI can help organizations to identify new business opportunities, increase efficiency, and enhance customer satisfaction.

According to a study by Dresner Advisory Services (2017), organizations that implement BI effectively can improve their decision-making accuracy by up to 40% and realize a return on investment within the first year of deployment.

BI includes several components such as data sources, Extraction, Transformation, and Loading (ETL) processes, data warehousing, and reporting (Ong et al., 2011).

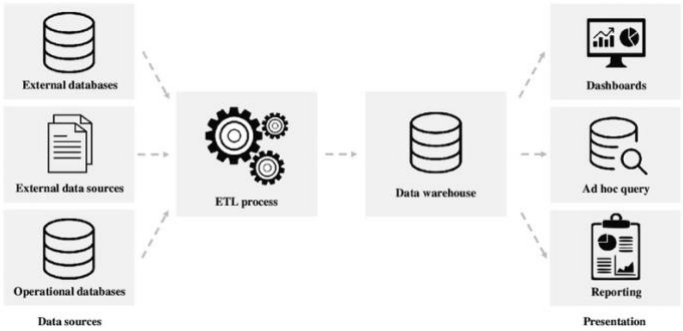

Figure 1 – Business Intelligence Components (Esteves et al., 2019)

2.1.1 BI Components

1. Data Sources

Data sources play a crucial role in BI as they provide the foundation for informed decision-making. They can be operational databases, historical data, external data from market research firms, the internet, or any existing data warehouse environment (Ranjan, 2009). This data can be structured, or unstructured. (Ranjan, 2009).

2. ETL (Extraction, Transformation, and Loading)

Even though the ETL process appears a relatively simple task, these processes can be very manual, expensive and can take a large portion of the total effort of a project (Dayal et al., 2009). These processes are responsible for extracting data from the databases, transform and clean these data, and load them into the data warehouse (Kimball & Caserta, 2004).

3. Data Warehouse

A data warehouse is a large, centralized repository of data that is a core component of BI as it is used to support activities such as reporting, data analysis, and decision-making (Bălăceanu, 2007).

Data Warehouses are designed to allow users to access large volumes of data quickly and easily, and to perform complex calculations and analysis on that data (Ong et al., 2011). This data is used to support decisions and can come from a variety of sources, both internal and external (Al-Debei, 2011).

Most businesses use these data warehouses as their source for all information, and it is defined by 4 concepts (Ong et al., 2011):

Subject-Oriented - All the information that is present in the data warehouse should be organized around subjects, that they describe, or refer to.

Integrated – As the data may come from different sources, with different characteristics, the data warehouse must be unified and aligned.

Non-volatile – Data warehouses should be permanent storage of data, meaning that the past data cannot be altered.

Time -Variant – To gather insights from data, it is important to have historical data, and for that, data should contain a time dimension.

4. Reporting

After populating the data warehouse, data analysis must be delivered to end users. The dashboards allow customization of visualizations, and as the variety of visuals is linked, they allow users to view and share different visuals at the same time. (Gowthami & Pavan Kumar, 2017).

Power BI, Tableau, and Qlik are powerful tools for business intelligence and data visualization, which empower enterprises to analyze, share, and visualize their data interactively and with ease (Elias & Bezerianos, 2011). The primary distinction among them is their cost-effectiveness and user-friendliness.

Although end users can only see the dashboards and their visuals, the correct implementation of ETL processes plays a critical role in ensuring data accuracy and consistency, which is essential for delivering effective and insightful business intelligence reporting (Bălăceanu, 2007).

2.1.2 Implementation in Organizations

The goal companies want to achieve when implementing Business Intelligence is to become a top-level organization when it comes to decision-making, transforming their decisions into more informed ones that impact not just the business but also the customers (Nafeeseh & Al-Mudimigh, 2011).

Starbucks, the American multinational coffeehouse chain, serves as a prime example of a successful BI implementation. The company gathers a significant amount of data from its daily transactions and uses it to improve its business performance. With the help of its mobile app, Starbucks offers personalized promotions and rewards to its customers based on their purchase history. Additionally, BI helps Starbucks to study customer behavior and preferences, which is useful when launching new products or opening new locations. By using BI to analyze customer data, Starbucks can make informed decisions on where to open new stores to best serve its customers (Lee, 2020).

Organizations realized that Business Intelligence encompasses much more than just utilizing technology. According to Miller (2008), the primary reason for a BI implementation failure is the underestimation of both the time required and the amount of data necessary for the implementation to be successful. Along the same line, some authors analyzed the Critical Success Factors (CSF) for the implementation of BI.

Yeoh & Koronios (2010) categorize the Critical Success Factors (CSFs) into three categories: Organizational, Process, and Technology. The Organizational category contains elements such as a well-defined vision, management support, and commitment. The Process category includes user-centered change management, and the Technology category encompasses a business-focused and interactive development methodology. Additionally, maintaining data quality and integrity are considered separate dimensions.

BI implementation can take place across a diverse range of industries, and as a result, the challenges and issues faced can vary based on the specific organization and its objectives for implementation (Hasan et al., 2016). Multinational companies leverage BI for reporting and analysis of both their own activities and those of their competitors, both locally and globally (Saif Ali Shah, 2012). Hospitals use BI to integrate the systems from different units holding patients' information, while hotels use it to identify sales trends and improve performance (Korte et al., 2013).

Despite the differences in objectives, there are some common challenges that are relevant to every industry, such as data management, data authenticity, and quality assurance to ensure the accuracy of the input data for the projects (Olszak & Ziemba, 2012). Another challenge is clearly defining business goals, as BI cannot effectively

support decision-making if the wrong factors are being analyzed. User acceptance is also a key factor that must be taken into consideration (Hasan et al., 2016).

2.2 Data Visualization

To reveal the pattern and trend of variables, the data, and the indicators available must be analyzed. For that, visualizations such as charts and diagrams are used to summarize and explain data (Elias & Bezerianos, 2011).

Dashboards went from just fixed graphics to interactive and customizable ones with the goal of simultaneously interpreting different visuals from distinctive perspectives. For example, when the user clicks a data point, all the visuals present on that page show that data cell from diverse viewpoints (Kitchin et al., 2015).

Dashboards display not only a summary of data, but also enable a drill down into visuals, and allow summary-to-detail analysis (Kitchin et al., 2015). Moreover, dashboards facilitate the exporting of visuals to use in documents.

2.2.1 Business KPIs

Performance measurement is fundamental for businesses to keep track of their efficiency; for that, the critical Key Performance Indicators (KPIs) must be identified (Cai et al., 2009). The choice of the right KPIs plays an important role in any business, as they define if the company is on the right track, for this the SMART rule was created. The SMART rule declares that a KPI must be Specific, Measurable, Attainable, Realistic, and Timely (SMART). It is very common for businesses to choose the wrong KPIs, either for the ease of measure or for the availability of data (Lavy et al., 2010).

Gumbus (2005) defends that the KPIs can be distributed in three categories, 1) the financial implications, where the company evaluates the ones that express costs, being utilized to decision-making in the short and long-term and can be measured per area, person, or product, for example. 2) Operational, which measures the performance of some area or team. 3) Customer Relations, as they cannot be quantified, for example, the opinions of the customers. These responses can detect some aspects of the business or the products that the other KPIs can’t and can be used to understand the demographic characteristics of the customers (age, gender, location, and more).

3 METHODOLOGY

The project was a collaborative effort between two analysts (myself and a senior analyst) and the team responsible for CRM (Customer Relationship Management) and communications, which developed the analysis request. Our main objective as analysts was to investigate the available data and develop new KPIs while updating the current ones. The ultimate goal of the project was to present the current results to the relevant stakeholders and propose possible future actions.

Throughout the project, we worked closely with the team responsible for CRM and communications to ensure that our findings aligned with their goals and objectives. By combining our analytical skills with their expertise in the field, we were able to produce a comprehensive report that provided valuable insights into the organization's performance.

This project followed the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology, which is a widely used framework for data mining and machine learning projects (Shearer, 2000). While we did not strictly adhere to all its phases and activities, we found it to be a useful guide for organizing our work and ensuring that we covered key aspects of the project. The CRISP-DM methodology breaks down a project into six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment (Shearer, 2000).

The first phase was aimed at defining an appropriate business question that could be answered or investigated with the available data. To achieve this, we worked closely with the responsible stakeholder, the CRM team, to understand the context and the relevant business objectives. After several discussions and brainstorming sessions, we identified the key question to be answered: what factors influence the click rate of emails sent? This question was deemed essential to improve the company's marketing strategy and increase customer engagement.

To understand the characteristics and the data we had available, we extracted and treated the data. The data used is stored in a data cloud – Snowflake, a cloud-based data platform. So, through a query in Snowflake, we obtained all the data and performed some data cleaning and transformation actions to prepare it for further analysis.

We then proceeded to create a dashboard in Power BI, which was fed by the Snowflake query mentioned earlier. The dashboard allowed us to explore the data visually and identify any patterns or trends that could be relevant to answering the business question.

Overall, by working closely with the CRM team and leveraging the available data, we were able to define a clear and actionable business question and prepare the data for further analysis.

4 INTERNSHIP DESCRIPTION

4.1 Philip Morris International

Philip Morris International (PMI) is a multinational tobacco company founded in 1847 and headquartered in New York City. The company is known for its popular brands such as Marlboro and Chesterfield and has a presence in over 180 countries worldwide (Philip Morris International, 2021b). As of 2021, PMI had a market capitalization of over $150 billion and reported net revenues of $28.7 billion for the year (Forbes, 2022).

PMI describes building a future of smoke-free products as its mission since 2015 and has made efforts to decrease the risk associated with the products it sells (Philip Morris International, 2020).

To achieve this, PMI has been developing a portfolio of risk-reduced products that are alternatives to conventional *tobacco*.

4.2 Tabaqueira S.A.

Tabaqueira is a leading company in Portugal that has contributed for almost a century to the national economy, being the largest tobacco company in Portugal, and one of the largest in the country (Tabaqueira, 2021).

Since its foundation, Tabaqueira has contributed to the growth of the area of Albarraque, Sintra, having, in addition to the factory, a residential area for its workers and their families (houses, school, canteen, medical post, and a chapel).

Tabaqueira has operated for more than 95 years, and for the last 25 years, it has been a subsidiary of the Philip Morris International group, having reached a market share of more than 60% in Portugal (Jornal de Negócios, 2020).

Since its integration into the PMI group, Tabaqueira tripled its production, exporting 85% of it, totaling €700 million in 2021(ECO Seguros, 2022).

At the end of 2015, Tabaqueira started in Portugal a process already initiated in only 3 other affiliates of the PMI group (Japan, Switzerland, and Italy), the commercialization of an innovative form of tobacco consumption - smokeless and smokeless combustion.

4.3 IQOS

IQOS is a smokeless tobacco product that utilizes HnB (heat-not-burn) technology to release a nicotine-containing vapor instead of smoke (Philip Morris International, 2020). *IQOS* devices consist of two elements - a holder and a charger. This holder contains a metallic blade where the user inserts the tobacco stick (*HEETS*). The blade is then responsible for heating the stick up to 350°C (Adriaens et al., 2018).

In the last months of 2022, PMI has added to its portfolio new blade-less products - *IQOS ILUMA*. These heat the new sticks (*TEREA*) by induction, being an advance in the technology associated with these products, eliminating some pain points for consumers - the blade and the need for regular cleaning of the holder.

PMI claims that this technology is less harmful than traditional smoking and has marketed *IQOS* as a safer alternative (Philip Morris International, 2020). However, the safety of heat-not-burn products has been a point of debate, with some studies suggesting that they may not be as safe as claimed (Hajek et al., 2020). Despite these challenges, *IQOS* has gained popularity and currently is present in more than 50 markets worldwide (Philip Morris International, 2021a).

4.4 Objectives

The internship was developed in the Customer Experience department of Tabaqueira, specifically in the Commercial Intelligence (CI) team. The CI team comprises a diverse set of internal and external stakeholders to the organization, with a broad spectrum of demands such as developing dashboards, conducting data analysis, and providing insights. The primary objective of this internship was to produce valuable insights and visualizations of the data available to expand and support decision-making of several teams and project managers throughout the company.

Email campaigns can provide valuable information to a company in a variety of ways. They can be used to track customer engagement, measure the effectiveness of marketing strategies, and gather data on customer preferences and their demographics. This information can be used to improve customer targeting, personalize marketing messages, and optimize email content to increase open rates and conversion rates. Additionally, data on email campaign performance can be used to inform decisions about budget allocation, staffing, and other business operations.

Having not reached yet the desired goal for communications via email, the present study has as its main objective assessing the performance of current campaigns and identifying areas for improvement in order to reach more consumers, and for that, the subsequent questions were explored:

Question 1: Which age group or gender is engaging more with email campaigns?

Question 2: Are campaigns leading consumers to buy a product?

Question 3: Are consumers who have been in the heated tobacco category for a longer time more open to campaigns?

With these questions, this study aims not only to understand the success of the campaigns but also the loyalty and the different consumer profiles that are present.

4.5 The Internship Activities

In this section, I will provide a detailed description of the activities carried out during the internship, with a focus on the solution's implementation and on the distinct phases described in the methodology section. The goal is to describe the approach taken to develop and implement the solution, as well as the various tools and technologies used to achieve the desired outcome.

4.5.1 Onboarding

During the first two weeks of the internship, I enrolled in an extensive training program that delivered an in-depth overview of Tabaqueira's operations and the tobacco industry. This training included sessions about the company's history, mission, and values, as well as an overview of the various products and services offered by the organization. Additionally, I was provided with an introduction to the tobacco industry, its history, and current trends and challenges.

Following this initial training period, I then spent the next two weeks focused on building my technical skills, particularly in the areas of data analysis and visualization. During that time, I received training on various tools and technologies, including Snowflake, SQL, and Power BI. This training was critical in enabling me to work with the company's vast amounts of data and to gain insights into customer behavior and

preferences, which were essential for developing effective solutions to improve the customer support system.

4.5.2 Data Preparation

The data extraction and processing phase was a crucial step in the project, as it involved obtaining the necessary data to address the project objectives. To accomplish this, we utilized the Snowflake platform, which allowed us to access multiple tables containing information about customer communications, purchases, and profiles. By query, we were able to extract the relevant data and prepare it for analysis.

The first step was to identify and select the relevant columns from the emails table. Once the necessary fields were identified, we then used a left join to combine data from other tables, specifically the consumer and purchase tables. This allowed us to merge information from different sources and obtain a more comprehensive view of customer behavior and preferences. The use of left join was a technical choice that enabled us to retain all the data from the email table while still incorporating relevant information from the other tables. We started by performing a join operation between the emails table and the consumer table, using the USER_ID field as the common identifier.

```
FROM DB_PT_PRD.DATA_PRODUCT.V_MKT_CLOUD_SENT A
LEFT JOIN DB_PT_PRD.DATA_PRODUCT.V_CRM_DB B ON A.USER_ID=B.USER_ID
```

By joining the "DEVICEACTIVITIES" table with the previously merged data, using the same USER_ID field, we were able to incorporate information on the date when a new device was registered in the user's account. To ensure that this information was relevant for our analysis, we filtered the data so that only device registration dates that occurred after the email was sent were included. This allowed us to focus on understanding the influence of email communications on customer purchases, as any device registrations prior to the email could not have been influenced by the email.

```
LEFT JOIN (SELECT OWNERID, ACTIVITYDATE, DEVICEID, ACTIVITYTYPE FROM
DB_PT_PRD.DATA_PRODUCT.V_DCS_DEVICEACTIVITIES ON DEV.OWNERID=A.USER_ID AND
DEV.ACTIVITYDATE >= SENT_DATE
```

To extract and incorporate information on email interactions, we leveraged three key tables in the dataset, each of which contained data on opened, clicked, and unsent emails. To incorporate this information into our analysis, we used binary variables, or flags, associated with each email ID to indicate whether the email was opened, clicked, or had

any delivery problems. By using these flags, we were able to create a more detailed picture of how customers engaged with email communications, including which emails were most successful in generating clicks and conversions.

```
CASE WHEN AUX_OPEN IS NULL THEN 0 ELSE 1 END AS "FLG_OPEN", OPENS.EVENTDATE AS
"OPEN_DATE",

CASE WHEN AUX_CLICK IS NULL THEN 0 ELSE 1 END AS "FLG_CLICK", CLICK.EVENTDATE AS
"CLICK_DATE",
CASE WHEN AUX_BOUNCE IS NULL THEN 0 ELSE 1 END AS "FLG_BOUNCE", BOUNCE.EVENTDATE AS
"BOUNCE_DATE"
```

During the data processing phase, we encountered an issue where some emails exceeded the maximum character limit, causing Snowflake to split them into two separate lines.

To address this, we used an auxiliary field (AUX_JOB_USER) that linked the campaign ID with the user ID and specified that only one instance of each email's sending date should appear in the final table. By using this approach, we were able to effectively combine data on split emails and avoid duplicate entries.

```
ROW_NUMBER() OVER (PARTITION BY AUX_JOB_USER ORDER BY A.EVENTDATE DESC ) AS "RN"
FROM DB_PT_PRD.DATA_PRODUCT.V_MKT_CLOUD_SENT A

WHERE RN = 1
```

Consequently, the final overview of all the variables contained within our dataset was:

Table I - Description of Variables

Variable Name	Variable Description
JOB_ID	Campaign identifier
EMAIL_NAME	Campaign name – the subject of the email
SENT_DATE	Sent date
USER_ID	User identifier
AGE_GROUP	User's age (18-24, 25-29, 30-39, 40-49, 50+)
GENDER	User's gender
REGISTRATION_DATE	User's registration date
AREA	User's living area
ACCEPTANCE_ON_NOTIFICATIONS	1 if the consumer accepted the notifications
FLG_OPEN	1 if the consumer opened the email
OPEN_DATE	Open Date
OPEN_DEVICE	Device where the email was opened
FLG_CLICK	1 if the consumer clicked on the email
ALIAS	Where the consumer clicked
CLICK_DATE	Click Date
FLG_BOUNCE	1 if the email bounced back
BOUNCEREASON	Reason for the email to bounce back
ACTIVITY_DATE	Registration date of a device bought

4.5.3 Reporting

To facilitate our analysis, we used Power BI as our primary data visualization tool. Power BI allowed us to load data from different data sources. From there, we established a connection to Snowflake, importing the data from the query we had written earlier into

Power BI. This allowed us to create a new dataset within Power BI that was based on our Snowflake query. We needed to perform some additional data transformations to prepare the data for analysis. To accomplish this, we used Power Query in Power BI.

We were able to define the type of variables that were not correctly identified. In addition, using Power BI, we were able to define measures for click rate, open rate, and bounce rate, which helped us analyze email engagement.

```
OPEN_RATE = DIVIDE(SUM('E-MAILS'[FLG_OPEN]), COUNTROWS('E-MAILS'))
CLICK_RATE = CALCULATE(SUM('E-MAILS'[FLG_CLICK]), FILTER('E-MAILS','E-MAILS'[ALIAS]<>
"unsubscribe"))/COUNTROWS('E-MAILS')
BOUNCE_RATE = SUM('E-MAILS'[FLG_BOUNCE])/COUNTROWS('E-MAILS')
```

Then, we calculated the number of users who linked a device 15 days after receiving an email. Subsequently, we divided the number of users who linked a device by all the users who received emails, giving us the link rate.

```
BUY NEXT 15 DAYS = if(DATEDIFF('E-MAILS'[SENT_DATE],'E-MAILS'[ACTIVITYDATE],DAY)<=15
&& 'E-MAILS'[ACTIVITYDATE] <> BLANK(),1,0)
LINK_RATE = SUM('E-MAILS'[BUY NEXT 15 DAYS])/COUNTROWS('E-MAILS')
```

The first sheet of our Power BI dashboard was focused on presenting the KPIs related to the email marketing campaign. We used various visualizations to showcase the number of emails sent, the number of consumers reached, as well as the open and click rates. These KPIs were essential for evaluating the success of the campaign and providing insights into consumer behavior. We also included a trend analysis over time, allowing us to identify campaign performance patterns and adjust our strategies accordingly. The visualizations were interactive, enabling us to filter data based on different criteria, such as date range or target audience, providing a deeper understanding of the campaign's effectiveness.

The second sheet of the report was focused on analyzing the click rate per different demographic factors. Specifically, we wanted to understand how the click rate varied across different ages, gender, locations, and maturity (time in the HnB category). For age, gender, and maturity analysis, we used a bar chart to display the average click rate for each group. For the user's location analysis, we used a table displaying the click rate for different geographic regions and the penetration rate of *IQOS* for that region.

On the last sheet, we analyzed the link rate and identified which campaigns led consumers to buy a device in the next 15 days. We created a bar chart that shows the link rate for each campaign, sorted from highest to lowest.

We incorporated slicers into each sheet, providing users with the ability to filter the data presented. These filters included the date, campaign, age group, and type of device associated with the user. By using these slicers, users could easily modify the visuals and better understand the relationship between the different variables.

5 RESULTS

The analysis has been completed and the data has been processed and visualized using Power BI. With the use of various measures and graphs, insights were gained into the email communication of the company, and the factors that influence the click and open rates of emails. This section will present the results of the analysis highlighting the significant findings related to the questions proposed before.

In this context, dividing emails into different groups can help to better target and engage the audience and to better understand the results. In this case, we categorized their emails into three distinct groups: Call to Action, NPS, and Others. The first group, Call to Action, includes emails that are specifically designed to encourage the user to take a specific action, in this case, it includes email campaigns launching new products, promoting existing products, or encouraging customers to participate in user programs. The second group, NPS (Net Promoter Score), is a metric used to measure customer loyalty and satisfaction by asking customers how likely they are to recommend a product or service to others. In this context, the group comprises emails to customers who have recently completed a purchase, asking them to provide feedback on their experience. The third group, Others, included emails that are not related to NPS or Call to Action. These emails include company updates, newsletters, or general marketing content.

It was possible to generate several key findings. First, each consumer received an average of seven emails per year, indicating a significant volume of communications. Second, campaigns promoting new products had a higher click rate than those focused on other messaging, for example, NPS emails. Third, consumers from areas with a higher penetration rate for the heated tobacco category tended to click on emails more frequently. Finally, consumers from regions with higher purchasing power also showed a higher propensity to engage with email campaigns.

5.1 Question 1

After conducting our analysis, we discovered that gender had no significant impact on the reach of email campaigns to consumers. However, age was a crucial factor in determining the effectiveness of these campaigns.

For this analysis, we divided our user base into five distinct age groups: 18-24, 25-29, 30-39, 40-49, and 50+.

Our findings showed that younger individuals were more inclined to open emails than their older equals. On the other hand, older individuals were more likely to click on the email content once they opened it.

These results imply that it is crucial to take age into consideration when designing email campaigns. It may be necessary to use different approaches to effectively target different age groups. For instance, we could create email content that appeals more to the interests of older consumers to increase the likelihood of them opening the email. For younger individuals, we could focus on creating more engaging content that encourages them to click on the email's content.

5.2 Question 2

According to our findings, email advertising had a considerable impact on customers' purchasing decisions. Specifically, we found that consumers who clicked on emails were 50% more likely to acquire a product in the next 15 days than those who did not engage with the content. Furthermore, our research shows that tracking click-through rates can provide useful information about the success of email campaigns. Companies may measure the performance of their email campaigns and adapt their strategy accordingly by analyzing click-through rates.

5.3 Question 3

Based on our findings, consumers who have been in the heated cigarette category for a longer length of time are more engaged with email advertisements. This shows that when customers grow more familiar with a product category, they may become more susceptible to tailored marketing efforts. This research emphasizes the necessity of evaluating the amount of time consumers have been in the category when planning marketing strategy. Furthermore, this research illustrates the potential benefits of developing long-term connections with clients. As customers get more familiar with a product, they may be more likely to interact with marketing activities, resulting in improved brand loyalty and revenue growth.

Upon analyzing the available data, it has been concluded that there are factors that are influencing email performance that are not currently being considered. It has become clear that certain consumers are being excessively contacted, leading to decreased open and click rates. To improve these rates, it is essential to segment campaigns based on factors such as age and maturity (time in the HnB category), ensuring that each group receives a personalized message. This strategy can help increase engagement and prevent burnout among customers who are being overly contacted. Also, it is crucial to prioritize the click rate as a KPI for these campaigns as clicking on an email increases the likelihood of a purchase. By focusing on increasing the click rate, the overall effectiveness of email campaigns can be improved.

After reflecting on the methodology chosen for the project and the steps followed, it can be concluded that the selected method was appropriate. The study required a deep understanding of the available data, and the initial steps of the method allowed for that understanding to be achieved.

Working closely with the CRM team proved to be a significant advantage as it provided valuable insights and opinions. This collaboration enabled the team to approach the data with a more realistic perspective, leading to more informed decisions.

However, some aspects of the process could be improved. The lack of an identifier for each email in the data made it challenging to track individual emails and their performance. This issue was addressed by creating an auxiliary identifier (AUX_JOB_USER), but if the table had a specific identifier for each email, the process would be simplified.

In summary, while the chosen methodology was appropriate for the study, there is always room for improvement. Close collaboration with relevant teams and having a deep understanding of the data available can greatly enhance the accuracy and efficiency of data analysis.

Future projects should consider analyzing all forms of communication with customers, not just emails. This will help ensure that clients are not being excessively contacted and will provide a more comprehensive understanding of customer engagement. Outbound calls, SMS, and NPS calls are additional communication channels that could be evaluated for their impact on customer engagement and overall

effectiveness. Also, analyzing customer feedback through NPS calls can help identify areas for improvement and provide insights for enhancing customer experience.

6 CONCLUSIONS

This internship was a great opportunity to apply all that I learned during my master's program. The theoretical knowledge I gained in the classroom became practical when I had to apply it to real-world problems. For example, I was able to use the statistical methods I learned to analyze data sets and draw insights that informed business decisions. Additionally, I had the chance to work on collaborative projects with people from different backgrounds, which gave me a valuable perspective on teamwork and communication.

After reflection, it can be confidently stated that all objectives of this project were successfully accomplished. The team was able to gather valuable information and insights from the data analysis, which allowed for the development of new communication strategies aimed at improving business performance. The data provided a clear understanding of customer engagement and allowed for the identification of areas for improvement in current marketing campaigns.

In addition, the KPIs defined at the outset of the project proved to be a valuable tool for the team to track the impact of email campaigns. By regularly monitoring these KPIs, the team can make data-driven decisions and adjust their communication strategies accordingly.

Managing a stakeholder such as the CRM team in this project provided an opportunity to gain a better understanding of various aspects of business management. Firstly, it allowed for a deeper understanding of the business. By working with the CRM team, it was possible to gain insights into their objectives and priorities, enabling us to tailor the data analysis to meet their needs. Secondly, the project required effective time management skills to meet the stakeholder's expectations within the given timeframe. By setting realistic deadlines, prioritizing tasks, and efficiently allocating resources, we were able to complete the project on time while delivering quality results.

Overall, the project was a success, with the team achieving all objectives and gaining valuable insights to enhance business performance. By utilizing the information gathered and monitoring the defined KPIs, the team can continue to develop effective communication strategies and improve customer engagement.

REFERENCES

Adriaens, K., van Gucht, D. & Baeyens, F. (2018). IQOS™ vs. e-cigarette vs. tobacco cigarette: A direct comparison of short-term effects after overnight-abstinence. *International Journal of Environmental Research and Public Health*. 15 (12).

Al-Debei, M.M. (2011). Data Warehouse as a Backbone for Business Intelligence: Issues and Challenges. *European Journal of Economics, Finance and Administrative Sciences*. [Online]. (33). Available from: http://www.eurojournals.com. [Accessed: 11 January 2023].

Bălăceanu, D. (2007). Components of a Business Intelligence software solution. *Informatica Economică*. 2. p.pp. 67–73.

Cai, J., Liu, X., Xiao, Z. & Liu, J. (2009). Improving supply chain performance management: A systematic approach to analyzing iterative KPI accomplishment. *Decision Support Systems*. 46 (2). p.pp. 512–521.

Dayal, U., Castellanos, M., Wilkinson, K. & Simitsis, A. (2009). *Data Integration Flows for Business Intelligence*. ACM.

Dresner Advisory Services (2017). *Big Data Analytics Market Study Wisdom of Crowds*. [Online]. Available from: http://www.dresneradvisory.com. [Accessed: 27 January 2023].

ECO Seguros (2022). *Tabaqueira reduz riscos: já há 400 mil utilizadores de IQOS*. [Online]. 24 November 2022. ECO Seguros. Available from: https://eco.sapo.pt/2022/11/24/tabaqueira-reduz-riscos-ja-ha-400-mil-utilizadores-de-iqos/. [Accessed: 8 December 2022].

Elias, M. & Bezerianos, A. (2011). *Exploration Views: Understanding Dashboard Creation and Customization for Visualization Novices*.

Esteves, M., Esteves, M., Abelha, A. & Machado, J. (2019). A proof of concept of a mobile health application to support professionals in a portuguese nursing home. *Sensors (Switzerland)*. 19 (18).

European Commission (2022). *The European Strategy for data at a glance*.

Forbes (2022). *Forbes - Philip Morris International*. [Online]. 2022. Available from: https://www.forbes.com/companies/philip-morris-international/. [Accessed: 27 November 2022].

Gowthami, K. & Pavan Kumar, M.R. (2017). Study on Business Intelligence Tools for Enterprise Dashboard Development. *International Research Journal of Engineering and Technology*. [Online]. 4 (4). p.pp. 2987–2992. Available from: www.irjet.net.

Gumbus, A. (2005). Introducing the balanced scorecard: Creating metrics to measure performance. *Journal of Management Education*. 29 (4). p.pp. 617–630.

Hajek, P., Pittaccio, K., Pesola, F., Myers Smith, K., Phillips-Waller, A. & Przulj, D. (2020). Nicotine delivery and users' reactions to Juul compared with cigarettes and other e-cigarette products. *Addiction*. 115 (6). p.pp. 1141–1148.

Hasan, N.A., Rahman, A.A., Lahad, N.A. & Bahru, J. (2016). Issues And Challenges In Business Intelligence Case Studies. *Jurnal Teknologi*. 8 (2). p.pp. 171–178.

Jornal de Negócios (2020). Menos 250 milhões de cigarros vendidos na pandemia. *Jornal de Negócios*. [Online]. Available from: https://www.jn.pt/nacional/menos-250-milhoes-de-cigarros-vendidos-na-pandemia-13111596.html. [Accessed: 8 November 2022].

Khan, S., Qader, M.R., Ka, T. & Abimannan, S. (2020). Analysis of Business Intelligence Impact on Organizational Performance. In: *International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy*. 26 October 2020, Institute of Electrical and Electronics Engineers Inc.

Kimball, R. & Caserta, J. (2004). *The Data Warehouse ETL Toolkit*. Wiley Publisher.

Kitchin, R., Lauriault, T.P. & McArdle, G. (2015). Knowing and governing cities through urban indicators, city benchmarking and real-time dashboards. *Regional Studies, Regional Science*. 2 (1). p.pp. 6–28.

Korte, D., Frolick, M. & Ariyachandra, T. (2013). Business Intelligence in the Hospitality Industry. *International Journal of Innovation, Management and Technology*. 4 (4).

Lavy, S., Garcia, J.A. & Dixit, M.K. (2010). Establishment of KPIs for facility performance measurement: Review of literature. *Facilities*. 28 (9). p.pp. 440–464.

Lee, S.-Y. (2020). Study on Digital Transformation Strategies: Starbucks Case Study. *Journal of Digital Contents Society*. 21 (10). p.pp. 1809–1816.

Mallam, P., Ashu & Singh, B. (2021). Business Intelligence Techniques Using Data Analytics: An Overview. In: *International Conference on Computing Sciences*. 2021, Institute of Electrical and Electronics Engineers Inc., pp. 265–267.

Miller, G.J. & Queisser, T. (2008). *Does A Business Intelligence Competency Center (BICC) Improve Business Performance Through Better Decision Making?* In: 2008, Florida.

Nafeeseh, R.A. & Al-Mudimigh, A.S. (2011). Justifying ERP Investment: The Role and Impacts of Business Case A Literature Survey. *IJCSNS International Journal of Computer Science and Network Security*. 11 (1).

Olszak, C.M. & Ziemba, E. (2012). Critical Success Factors for Implementing Business Intelligence Systems in Small and Medium Enterprises on the Example of Upper Silesia, Poland. *Interdisciplinary Journal of Information, Knowledge, and Management*. 7.

Ong, I., Siew, P. & Wong, S. (2011). A Five-Layered Business Intelligence Architecture. *Communications of the IBIMA*. p.pp. 1–11.

Philip Morris International (2021a). *Integrated Report*. [Online]. Available from: www.pmi.com. [Accessed: 22 November 2022].

Philip Morris International (2021b). *Philip Morris International - About PMI*. [Online]. 2021. Available from: https://www.pmi.com/who-we-are/about-pmi. [Accessed: 27 December 2022].

Philip Morris International (2020). *PMI'S Statement of Purpose*.

Popovič, A., Hackney, R., Coelho, P.S. & Jaklič, J. (2012). Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. *Decision Support Systems*. 54 (1). p.pp. 729–739.

Ranjan, J. (2009). Business Intelligence: Concepts, Components,Techniques And Benefits. *Journal of Theoretical and Applied Information Technology*. 9 (1). p.pp. 60–70.

Saif Ali Shah, S. (2012). *A case of BI adoption in Pakistan*. University of Skovde.

Shearer, C. (2000). The CRISP-DM Model: The New Blueprint for Data Mining. *Journal of Data Warehousing*. 5 (4). p.pp. 13–22.

Tabaqueira (2021). *Relatório de Sustentabilidade*.

Turban, E., Sharda, R., Dele, D. & King, D. (2011). *Business intelligence : a managerial approach*. Prentice Hall.

Wixom, B. & Watson, H. (2010). The BI-Based Organization. *International Journal of Business Intelligence Research*. 1 (1). p.pp. 13–28.

Yeoh, W. & Koronios, A. (2010). Critical success factors for business intelligence systems. *Journal of computer information systems*. 50 (3). p.pp. 23–32.