Revista Inovação, Projetos e Tecnologias - IPTEC
e-ISSN: 2318-9851
https://doi.org/10.5585/iptec.v11i1.24603
Received: 31 May 2023 – **Approved:** 06 June. 2023
Chief Editor: Profa. Dra. Cristina Dai Prá Martens
Scientific Publisher: Profa. Dra. Cristiane Drebes Pedron

Check for updates

BUILDING A BUSINESS INTELLIGENCE DASHBOARD FOR A LEBANESE COMPANY[1]

CONSTRUINDO UM DASHBOARD DE BUSINESS INTELLIGENCE PARA UMA EMPRESA LIBANESA

Gerard Najem
MSc in Data Analytics for Business
Instituto Superior de Economia e Gestão da Universidade de Lisboa – ISEG-UL.
Lisboa - Portugal
l56775@aln.iseg.ulisboa.pt

Jesualdo Cerqueira Fernandes
PhD in Management
Instituto Superior de Economia e Gestão da Universidade de Lisboa – ISEG-UL.
Lisboa - Portugal
jcf@iseg.ulisnboa.pt

Abstract: In this study, we develop a Business Intelligence dashboard for a Lebanese company to enhance their financial management. Development involved integrating their available data into a physical database using Python and PostgreSQL. From this database, we generated a dashboard with Power BI to visualize and analyze their revenue numbers. The outcome helped the company gain insights into their business performance, optimize their revenue and profit, and to identify data entry issues that may affect their results. The dashboard also provides a foundation for the further integration and analyzation of new data.

Keywords Business Intelligence. Dashboards. Data. ETL. Modelling.

Resumo: Neste estudo, desenvolvemos um *dashboard* de *Business Intelligence* para uma empresa libanesa para melhorar sua gestão financeira. O desenvolvimento envolveu a integração dos dados disponíveis numa base de dados física, recorrendo a *Python* e *PostgreSQL*. A partir dessa base de dados, geramos um *dashboard* com o Power BI para visualizar e analisar os valores de receita. O resultado ajudou a empresa a obter *insights* sobre o desempenho dos seus negócios, otimizar a receita e o lucro e identificar problemas de entrada de dados que podem afetar os resultados. O *dashboard* também fornece uma base para futuras integrações e análises de novos dados.

Palavras-chave: Business Intelligence. Dashboards. Dados. ETL. Modelização

Cite como

American Psychological Association (APA)

Najem, G., & Fernandes, J. C. (2023, jan./jun.). Building a business intelligence dashboard for a Lebanese company. *Revista Inovação, Projetos e Tecnologias - IPTEC,* São Paulo, *11*(1), 1-18, e24603. https://doi.org/10.5585/iptec.v11i1.24603.

[1] This work was supported by FCT, I.P., the Portuguese national funding agency for science, research and technology, under the Project UIDB/04521/2020

1 Introduction

In today's data-driven business environment, companies are confronted with the challenge of managing and analyzing vast volumes of data to extract valuable insights for informed decision-making. The increasing emphasis on collecting and organizing client information, social media analytics, and other potentially valuable data accentuate the importance of being able to efficiently store, manage, and analyze data (Cukier 2013; Li et al. 2018). By strategically tracking these metrics, companies gain a powerful tool that can save them millions of dollars by predicting potential product failures or identifying underperforming departments (Younus and Qureshi 2021).

To delve deeper into this wealth of data and to unlock valuable insights, companies need to analyze data with increasing granularity. This type of analysis allows them to identify challenges, seize opportunities, innovate solutions, and leverage their strengths (Davenport and Harris 2007). In a dynamic marketplace, companies need to understand the target audience and how their operations interact with it in order to maintain competitiveness (Gao and Ren 2020).

The focus of this study is a project carried out with a Lebanese company. The aim of the project was to build a comprehensive interactive dashboard for the company's decision-makers based on a solid database. The objective was to equip key decision-makers with real-time insights and an accurate analysis of their company's operations. The dashboard would enable management to visualize key performance indicators (KPIs) and other essential metrics for monitoring performance (Hernández-Orallo et al. 2020).

To achieve this objective, the project leveraged advanced data analytics to provide an overall view of the company's performance that would empower the management team to make informed decisions. A diverse range of data types, such as sales, customer information, and finances, were used in the project (Wu et al. 2021). Furthermore, the project used modeling techniques on the data, such as entity-relationship diagrams and normalization, to establish a robust and scalable database (García-Sánchez et al. 2018).

The company in question is a small Lebanese start-up that specializes in crafting unique and visually captivating designs inspired by photographs. These designs are transformed into decorative and gift items that are handcrafted from steel. The company is three years old and has no real competitors in the market. It has diversified over time to produce other types of products such as bowties, wine holders, and keychains.

One of the primary objectives of this project was to empower the company by establishing its very own data center. This centralized hub served as a unified platform that

enabled the company to efficiently track and analyze their sales and revenue figures. By consolidating all pertinent data in a single location, the company gained valuable insights into their performance that facilitated its decision-making and strategic planning for future growth and success.

This study provides a comprehensive overview of the problem addressed and the objectives of the project. The subsequent sections present the method used for data analysis and its findings that provide valuable insights for companies seeking to enhance their decision-making through data-driven approaches.

2 Literature review

This section presents a literature review on four essential topics that were considered in this project: the significance of the quality of the data for the analysis, the effect of big data on companies' decision-making, the enhancement of efficiency and effectiveness through the use of data, and the role of data in financial management.

The company in question is a small Lebanese start-up that specializes in crafting unique and visually captivating designs inspired by photographs. These designs are transformed into decorative and gift items that are handcrafted from steel. The company is three years old and has no real competitors in the market. It has diversified over time to produce other types of products such as bowties, wine holders, and keychains.

One of the primary objectives of this project was to empower the company by establishing its very own data center. This centralized hub served as a unified platform that enabled the company to efficiently track and analyze their sales and revenue figures. By consolidating all pertinent data in a single location, the company gained valuable insights into their performance that facilitated its decision-making and strategic planning for future growth and success.

This study provides a comprehensive overview of the problem addressed and the objectives of the project. The subsequent sections present the method used for data analysis and its findings that provide valuable insights for companies seeking to enhance their decision-making through data-driven approaches.

2.1 Quality of data

In today's business world, data originates from diverse sources, including customer interactions, online reviews, social media analytics, market research, and sales data (Gandomi

and Haider 2015). The challenge for companies is to consolidate these disparate data into a unified and meaningful format for informed decision-making (Breslin 2019). Achieving this format requires a profound understanding of the sources and types of data as well as the effective tools for analyzation.

After unifying the data, companies need to identify the specific data points most beneficial for their operations. This identification involves careful consideration of the relevant key performance indicators (KPIs) aligned with their business objectives, such as customer retention, revenue growth, or market share (Kim and Lee 2012). By pinpointing valuable data points, companies can create more accurate predictive models and optimize their decision-making.

Furthermore, the process of unifying and analyzing data is not a one-time event; rather, it is an ongoing process that necessitates regular review and updates (Lenzerini 2002). With the constant evolution of the business landscape, companies must stay up to date with the latest trends and technologies to ensure they are leveraging the most relevant data to drive their business forward.

Wieder et al. (2012) highlight the importance of quality data for effective business intelligence. The authors find that while business intelligence tools can provide valuable insights and improve decision-making, they are only as effective as the quality of the data they are analyzing. The authors emphasize that poor quality data can lead to inaccurate or incomplete analysis that leads to flawed decision-making and poor performance. They also find that issues with the quality of the data can come from a variety of sources, such as inconsistent formats, incompleteness, or incorrect entry.

To address these issues, the authors suggest that companies should prioritize the practices that management can use to ensure the data's quality, such as cleansing and validation, so that their business intelligence tools are working with accurate and reliable data. They also recommend regular monitoring and auditing of the quality to identify and address any issues that may arise (Wieder et al. 2012).

Overall, Wieder et al. (2012) highlight the critical importance of the quality of data for effective business intelligence and underscore the need for companies to prioritize management practices that ensure quality to achieve optimal performance.

The company in question is a small Lebanese start-up that specializes in crafting unique and visually captivating designs inspired by photographs. These designs are transformed into decorative and gift items that are handcrafted from steel. The company is three years old and

has no real competitors in the market. It has diversified over time to produce other types of products such as bowties, wine holders, and keychains.

One of the primary objectives of this project was to empower the company by establishing its very own data center. This centralized hub served as a unified platform that enabled the company to efficiently track and analyze their sales and revenue figures. By consolidating all pertinent data in a single location, the company gained valuable insights into their performance that facilitated its decision-making and strategic planning for future growth and success.

This study provides a comprehensive overview of the problem addressed and the objectives of the project. The subsequent sections present the method used for data analysis and its findings that provide valuable insights for companies seeking to enhance their decision-making through data-driven approaches.

2.2 Role of big data in reporting and decision-making

Fernández (2020) explores the role of big data in reporting and decision-making. With the advent of modern technologies, big data have become a valuable resource for companies looking to improve their processes for reporting and making decisions (Gandomi and Haider 2015). The use of big data can help companies access and analyze large amounts of data from various sources to create more accurate and comprehensive reports. These reports can provide valuable insights into customer behaviour, market trends, and other key metrics that can help companies make more informed decisions (Manyika et al. 2011). By using analytics tools for big data, companies can also streamline their reporting processes by reducing the time and resources needed to produce reports.

When it comes to decision-making, companies can use big data to identify patterns and trends in large datasets that would be difficult or impossible to identify using traditional methods (Gandomi and Haider 2015). These data can be used to identify new market opportunities, optimize pricing strategies, and improve customer engagement. Overall, the role of big data in reporting and decision-making is becoming increasingly important as companies seek to gain a competitive edge in today's rapidly changing business environment (Manyika et al. 2011).

The company in question is a small Lebanese start-up that specializes in crafting unique and visually captivating designs inspired by photographs. These designs are transformed into decorative and gift items that are handcrafted from steel. The company is three years old and

has no real competitors in the market. It has diversified over time to produce other types of products such as bowties, wine holders, and keychains.

One of the primary objectives of this project was to empower the company by establishing its very own data center. This centralized hub served as a unified platform that enabled the company to efficiently track and analyze their sales and revenue figures. By consolidating all pertinent data in a single location, the company gained valuable insights into their performance that facilitated its decision-making and strategic planning for future growth and success.

This study provides a comprehensive overview of the problem addressed and the objectives of the project. The subsequent sections present the method used for data analysis and its findings that provide valuable insights for companies seeking to enhance their decision-making through data-driven approaches.

2.3 Improvement of efficiency and effectiveness

Popescu (2012) explores how business intelligence solutions can help companies improve their efficiency and effectiveness.

According to Popescu (2012), these solutions can improve efficiency by streamlining the collection and analysis of data. By automating these processes, companies can reduce the time and resources needed to collect and analyze data, allowing them to make faster and more accurate decisions. Business intelligence solutions can also help companies identify inefficiencies in their processes and operations that allow them to optimize their workflows and reduce waste.

In terms of effectiveness, these solutions can provide companies with valuable insights into their customers, markets, and competitors (Popescu 2012). This information can be used to develop more effective marketing and sales strategies, improve customer engagement, and identify new business opportunities.

Popescu (2012) also notes that these solutions can help companies improve their decision-making. By providing decision-makers with real-time access to data and insights, companies can make more informed decisions that are based on accurate and up-to-date information. Business intelligence solutions can also provide decision-makers with tools and visualizations that make it easier to interpret complex data and identify trends and patterns.

Overall, the use of these solutions can lead to general improvements in efficiency and effectiveness for companies (Popescu 2012). By leveraging the power of data and analytics,

companies can optimize their processes, improve their decision-making capabilities, and gain a competitive edge in today's rapidly changing business environment.

The company in question is a small Lebanese start-up that specializes in crafting unique and visually captivating designs inspired by photographs. These designs are transformed into decorative and gift items that are handcrafted from steel. The company is three years old and has no real competitors in the market. It has diversified over time to produce other types of products such as bowties, wine holders, and keychains.

One of the primary objectives of this project was to empower the company by establishing its very own data center. This centralized hub served as a unified platform that enabled the company to efficiently track and analyze their sales and revenue figures. By consolidating all pertinent data in a single location, the company gained valuable insights into their performance that facilitated its decision-making and strategic planning for future growth and success.

This study provides a comprehensive overview of the problem addressed and the objectives of the project. The subsequent sections present the method used for data analysis and its findings that provide valuable insights for companies seeking to enhance their decision-making through data-driven approaches.

2.4 Role in financial management

Bray (2011) explores how business intelligence dashboards can help companies improve their financial management.

According to Bray (2011), these dashboards can provide financial managers with a real-time view of their company's financial performance. This perspective allows them to quickly identify potential issues or areas of concern and take action to address them.

In addition to providing real-time insights into financial performance, these dashboards can also improve collaboration and communication within an company. By providing a centralized view of financial data, Bray (2011) explains that dashboards can facilitate collaboration between different departments and teams. This collaboration can lead to more effective communication and decision-making, as all stakeholders have access to the same information.

Another benefit of business intelligence dashboards is that they can be customized to meet the specific needs of a company. Financial managers can choose which KPIs and metrics

to display in the dashboard, based on their unique needs and goals. This flexibility allows companies to tailor their financial management to their specific requirements that improves their efficiency and effectiveness (Bray 2011).

Overall, according to Bray (2011), the use of business intelligence dashboards can play a crucial role in financial management. By providing real-time insights, facilitating collaboration and communication, and offering flexibility and customization, dashboards can help companies improve their financial performance and achieve their goals.

3 Method

This section gives an outline of the project such as data exploration, tool selection, model design, and ETL (Extract, Transform, Load) implementation.

3.1 Datasets

The data were evaluated to determine the key data points, their values, and their relevance for the analysis. No data cleaning was performed prior to this evaluation as all changes were made in later steps.

3.2 Tools

For the ETL process, Python and PostgreSQL were selected as the preferred tools because of their user-friendly capabilities in manipulating and cleaning data. The CSV files were integrated into Python and subsequently connected to PostgreSQL for seamless integration into the tables.

Python is a popular and powerful programming language for integrating and analyzing data due to its large ecosystem of libraries, such as NumPy and Pandas. It is also versatile and can be used for a wide range of other tasks.

PostgreSQL is an open-source relational management system for databases known for its scalability and ability to handle extensive data and high transaction volumes. It offers a wide range of built-in types of data and advanced features for their manipulation that allow for versatility and customization according to specific requirements. With robust data security features, PostgreSQL ensures the protection of sensitive information. The platform benefits from a thriving community of developers who provide active support to users. Its extensibility

is evident through numerous third-party extensions and add-ons that make it a favorite option for projects that manipulate data.

Microsoft Power BI was chosen for building the dashboards due to its user-friendly interface, flexibility, and integration capabilities. It offers robust modeling, analysis, and visualization features for data, along with scalability and automatic updates as a cloud-based platform. The extensive support resources from Microsoft and a large user community make Power BI a popular choice for companies seeking intuitive and powerful visualization and analysis tools for data.

Thus, the whole flow that data usually pass through was covered.

4 Problem context and execution

This section has a description of the project's execution and deliverables. The company previously recorded the sales and other data manually with a cloud-based spreadsheet software called "Airtable." To process the data, files had to be exported manually and be labelled based on their content. Each file was renamed following a specific format for easy integration. The first six digits represented the year and month, while the remaining characters signified the category of data within the file.

4.1 Data model

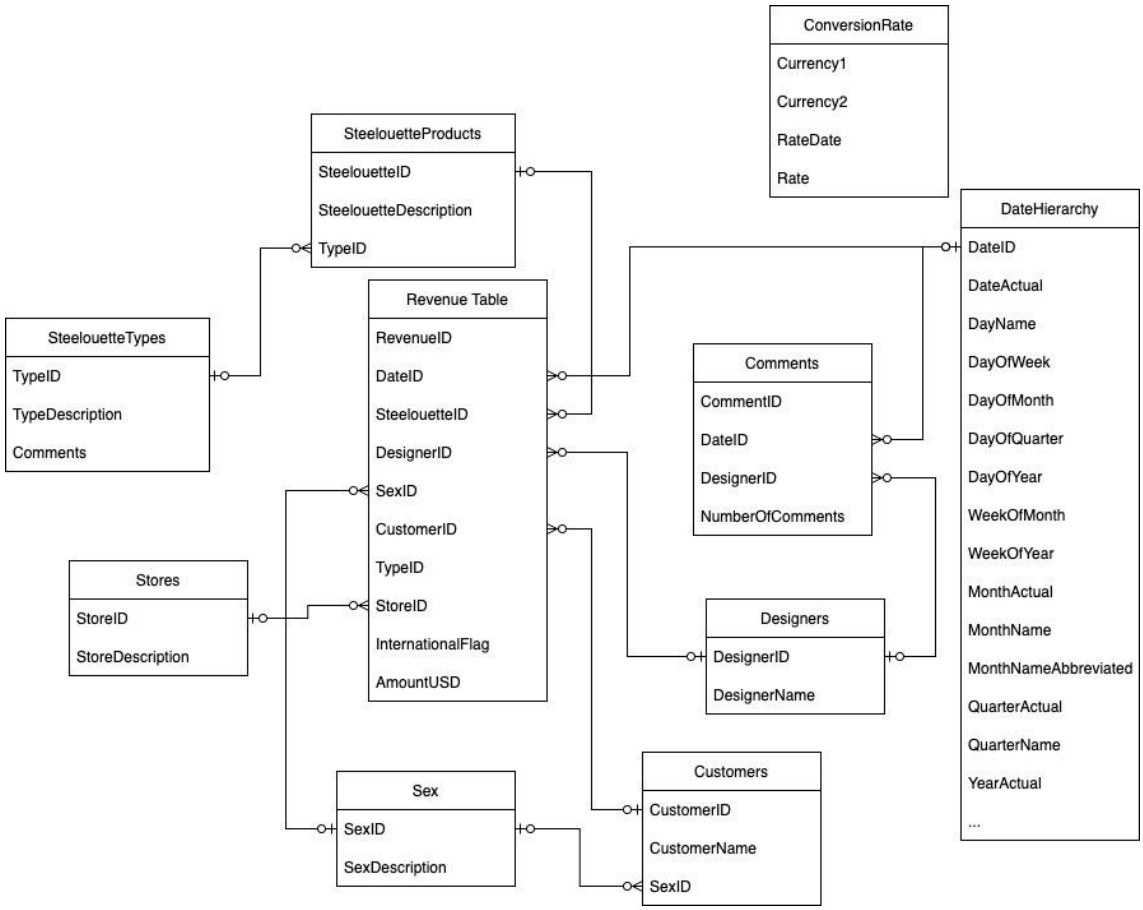
This subsection presents the data model that was built for the company; it contained two main areas: revenue and customer satisfaction.

To build the model, we used a star schema that is a database schema used to warehouse and organize data into a central fact table surrounded by dimension tables. This schema simplifies queries, improves performance, and provides scalability. The design is simple and easy to maintain that make it efficient for large data sets.

The model includes two fact tables: revenue and comments. The revenue table contains information about each product sold, while the comments table holds the number of comments per order per designer that allows them to evaluate customer satisfaction and track employee performance. The main dimension tables are date, products, types, stores, gender, customers, and designers that provide the flexibility to analyze revenue numbers from different perspectives and make better marketing and workflow decisions to maximize performance and profit.

Fig. 1.

Database model built for the project



Source: Created by the authors.

The tables adhere to normalization rules, with each table having both logical and physical primary keys (PK) generated by a sequence with increments for every inserted record. Foreign keys (FK) were implemented to enforce proper linkage between records during insertion to ensure accurate and consistent data.

4.2 ETL process

The ETL process is divided into three main parts. The chronology is essential for two important reasons: the first one ensures that the tables are being filled in an order that would complete the dimensions tables before the fact ones so that there is no violation of the rules of the FK. The second one divides the process into different steps for faster identification of possible errors and integration issues.

The first part is an SQL script that creates the model tables and the staging tables that store the raw data before transformation and processing. This chunk uses DDL queries to create the PK, FK, and indexes.

Fig. 2.

DDL commands to create table, PK, and FK

```
Create Table Steelouette_Revenue
(
    RevenueId           SERIAL Primary Key,
    DateId              int,
    SteelouetteId       int,
    SexId               int,
    DesignerId          int,
    CustomerId          int,
    TypeId              int,
    StoreId             int,
    InternationalFlag   int,
    AmountUSD           float
);

Create Table Comments
(
    CommentId           SERIAL Primary Key,
    DateId              int,
    DesignerId          int,
    NumberOfComments    float
);

ALTER TABLE Steelouette_Revenue
    ADD CONSTRAINT fk_revenue_type FOREIGN KEY (TypeId) REFERENCES Steelouette_type (TypeId);

ALTER TABLE Steelouette_Revenue
    ADD CONSTRAINT fk_revenue_store FOREIGN KEY (StoreId) REFERENCES Steelouette_Stores (StoreId);

```

Source: Created by the authors.

The second step is a Python script with the objective of extracting the data from the CSV files and inserting them into data frames using pandas. First, the conversion rate data are integrated. These data were exported from a website called lirarate.org into a CSV, and then inserted into the rate table in the database using a connection to Postgres.

Fig. 3.

Python Script To Integrate Conversion Rate Data

```
In [2]: df_rates = pd.read_csv("/Users/gerardnajem/Desktop/ISEG/MFW/usd-to-lbp-market-rate.csv")
        df_rates.rename(columns = {'USD to LBP':'Rate'}, inplace = True)
        df_rates.rename(columns = {'DateTime':'RateDate'}, inplace = True)
        df_rates.insert(0, column='Currency2', value='LBP')
        df_rates.insert(0, column='Currency1', value='USD')

In [3]: conn = psycopg2.connect(
            database="Steelouette", user='postgres', password='011095', host='localhost', port='5432'
        )

In [4]: def execute_values(conn, df, table):

            tuples = [tuple(x) for x in df.to_numpy()]

            cols = ','.join(list(df.columns))
            # SQL query to execute
            query = "INSERT INTO %s(%s) VALUES %%s" % (table, cols)
            cursor = conn.cursor()
            try:
                extras.execute_values(cursor, query, tuples)
                conn.commit()
            except (Exception, psycopg2.DatabaseError) as error:
                print("Error: %s" % error)
                conn.rollback()
                cursor.close()
                return 1
            print("the dataframe is inserted")
            cursor.close()

        execute_values(conn, df_rates, 'Conversion_rate')
```

Source: Created by the authors.

The same process has been applied to the CSV files containing the data; however, it was divided into eight different parts, since each type of file needed to be linked to its specific information and labelled to simplify the process later on. The needed columns were selected from the data frames; others were merged and then integrated into their specific staging table.

Fig. 4.

Python Script To Integrate The Revenue And Comments Data

```
for i in csv_files:
    if re.search('R',i):
        if R.empty:
            string = "/Users/gerardnajem/Desktop/ISEG/MFW/Data/" + i
            R=pd.read_csv(string)
            R.insert(0, column='Type', value='Non-Customized')
            R.insert(0, column='Week', value='1')
            R.insert(0, column='Month', value= i[4:6])
            R.insert(0, column='Year', value= i[0:4])
        else:
            string = "/Users/gerardnajem/Desktop/ISEG/MFW/Data/" + i
            R1=pd.read_csv(string)
            R1.insert(0, column='Type', value='Non-Customized')
            R1.insert(0, column='Week', value='1')
            R1.insert(0, column='Month', value= i[4:6])
            R1.insert(0, column='Year', value= i[0:4])
            R=pd.concat([R1,R])
    elif re.search('Bowties',i):
        if KeyChains.empty:
            string = "/Users/gerardnajem/Desktop/ISEG/MFW/Data/" + i
            Bowties=pd.read_csv(string)
            Bowties.insert(0, column='Type', value='Bowties')
            Bowties.insert(0, column='Week', value='1')
            Bowties.insert(0, column='Month', value= i[4:6])
            Bowties.insert(0, column='Year', value= i[0:4])
        else:
            string = "/Users/gerardnajem/Desktop/ISEG/MFW/Data/" + i
            Bowties1=pd.read_csv(string)
            Bowties1.insert(0, column='Type', value='Bowties')
            Bowties1.insert(0, column='Week', value='1')
            Bowties1.insert(0, column='Month', value= i[4:6])
            Bowties1.insert(0, column='Year', value= i[0:4])
            Bowties=pd.concat([Bowties1,Bowties])
```

Source: Created by the authors.

The third part is the actual transformation and manipulation of the data. It starts by filling the dimension tables; some of them contain static data such as the sex, the date hierarchy, the stores, and the types. The others, such as customers and products, are filled by checking the information that was collected from the CSV files. Afterwards, several transformations were performed on the data to ensure consistency and to identify outliers and data points that do not present any additional value to the analysis.

Fig. 5.

Filling Some Dimension Tables With Static Data

```
INSERT INTO Sex (SexDescription) values ('Female');
INSERT INTO Sex (SexDescription) values ('Male');
INSERT INTO Sex (SexDescription) values ('Unknown');

INSERT INTO Steelouette_type (TypeDescription, Comments) values ('Customized','');
INSERT INTO Steelouette_type (TypeDescription, Comments) values ('Bowtie','');
INSERT INTO Steelouette_type (TypeDescription, Comments) values ('KeyChain','');
INSERT INTO Steelouette_type (TypeDescription, Comments) values ('Wine','');
INSERT INTO Steelouette_type (TypeDescription, Comments) values ('Crib','');
INSERT INTO Steelouette_type (TypeDescription, Comments) values ('Non-Customized','');
INSERT INTO Steelouette_type (TypeDescription, Comments) values ('Unknown','');

```

Source: Created by the authors.

Fig. 6.

Filling Some Dimension Tables With Dynamic Data

```
INSERT INTO Steelouette_products (SteelouetteDescription, TypeId)
SELECT distinct INITCAP(TRIM(unnest(string_to_array(r.steelouettename, ','))));,
        (select typeid from steelouette_Type where typedescription = 'Bowtie')
  FROM Bowties_TEMP r
  where r.steelouettename <> 'Unknown'
UNION ALL
SELECT distinct INITCAP(TRIM(unnest(string_to_array(r.steelouettename, ','))));,
        (select typeid from steelouette_Type where typedescription = 'KeyChain')
  FROM Keychains_TEMP r
  where r.steelouettename <> 'Unknown'
UNION ALL
SELECT 'Wine',
        (select typeid from steelouette_Type where typedescription = 'Wine')
UNION ALL
SELECT 'Crib',
        (select typeid from steelouette_Type where typedescription = 'Crib')
```

Source: Created by the authors.

This step is the most important one, since it ensures that the data are clean, correct, and understandable to eventually facilitate the building of the dashboard and to extract the insights and conclusions from the project.

Fig. 3.

Example Of The Manipulation And Transformation Of Data

```
CREATE TABLE R_TEMP AS
SELECT CAST(R.Year      AS INT) AS Year,
       CAST(R.Month     AS INT) AS Month,
       CAST(R.Week      AS INT) AS Week,
       CAST(COALESCE(NULLIF(SUBSTRING(R.Date from 1 for
                 (CASE WHEN POSITION(',' in R.Date) = 0 THEN length(R.Date)
                       ELSE POSITION(',' in R.Date) - 1
                       END
                 )
                ), ''),'1/1/1970') AS DATE)  AS Date,
       CAST(INITCAP(TRIM(R.ClientName)) AS VARCHAR(100))  AS ClientName,
       CAST(COALESCE(NULLIF(R.SteelouetteName,''),'Unknown') AS VARCHAR(100)) AS SteelouetteName,
       CAST(CASE WHEN TRIM(R.Boy) = 'checked' THEN 'Male'
            WHEN TRIM(R.Girl) = 'checked' THEN 'Female'
            ELSE 'Unknown'
            END                     AS VARCHAR(10)) AS Sex,
       CAST(COALESCE(NULLIF(R.Number, ''),'0') AS FLOAT) AS Number,
       CAST(CASE WHEN NULLIF(R.SteelouetteName,'') like '%bowtie%' THEN 'Bowties'
                 WHEN NULLIF(R.SteelouetteName,'') like '%keychain%' THEN 'KeyChain'
            ELSE R.Type END AS VARCHAR(100))        AS Type,
       CAST(CASE WHEN position('DHL' in R.DelOpt) <> '0' THEN 'International' ELSE 'Local' END AS Varchar(20)) DelOpt,
       CAST(NULLIF(REGEXP_REPLACE(REGEXP_REPLACE(NULLIF(R.ValueUSD,''), '[^.0-9]+', '', 'g'), E'\\r?\\n', '', 'g'),'') AS FLOAT)
       CAST(NULLIF(REGEXP_REPLACE(REGEXP_REPLACE(NULLIF(R.Value,''), '[^.0-9]+', '', 'g'), E'\\r?\\n', '', 'g'),'') AS FLOAT) AS
  FROM StagingR R;
```

Source: Created by the authors.

4.3 Dashboards

The project was concluded by creating a dynamic dashboard that displayed the main findings in three parts.

The first part gave an overview of the business, with the key figures and trends. The second part showed the same figures but segmented by categories, to identify the performance and seasonality of different products. The third demonstrated the importance of quality data by highlighting the missing values that could affect the analysis.

Some of the calculations were done before connecting Power BI to the database, while others were done using DAX during the development of the tiles and pages.

The dashboard was designed to allow the user to choose a time period; filter the data by clicking on a specific time; or drill down into the date hierarchy from yearly to quarterly, monthly, or daily levels.

5 Findings

In this section, we present the main results that provide insights into the company's products, types, and collaborating stores. Additionally, the authors offer several tips to improve data and production analysis.

On the overview page, alongside revenue and number of items sold, the percentage of international sales is highlighted. This is crucial due to the unstable economic situation in

Lebanon. International sales contribute significantly, accounting for 3% of items sold but generating almost 12% of the revenue.

The data indicate that women were the primary customers, constituting 48% of total sales. However, around 40% of the items sold did not specify the gender which could affect these numbers.

Sales peaked at the end of the year before Christmas with November and December recording the highest numbers in terms of sales and revenue. There was also a slight increase in sales during the summer that was influenced by expats visiting Lebanon and purchasing gifts for friends and family.

Furthermore, a table was added at the request of the company to present the average comments of customers for each customized item produced. The table was segmented by designers to identify potential sources of customer dissatisfaction.

In the detailed view, several key findings emerged. Approximately 82% of the products sold were customized, and 92% of sales occurred either through social media or in-store. Excluding these two channels, the top 5 products sold and the top 5 stores were showcased in a tree map that highlighted the main income streams.

Another graph illustrates the performance of specific product categories or newly released collections. For example, the "bowtie" collection launched in Spring 2022 had peak sales during the summer wedding season. Similarly, the "crib" collection was exclusively available in December for Christmas. Non-customized products were produced year-round, with a slight increase in March before Mother's Day and at the end of the year.

As mentioned earlier, the purpose of the last page is to highlight data gaps. Notably, 95% of revenue was not linked to any specific product, and 95% of items sold were not assigned to a product. Additionally, 41% of sales were attributed to customers without a specified gender, contributing only 7% of the revenue. Moreover, 13% of sales and 12% of revenues came from uncategorized products, mainly due to issues with the data collection at other stores.

Other miscellaneous findings include 20% of the items designed lacked an identified designer and 5% of products sold lacked a clear customer association. Although these gaps might not have a significant effect on the business, addressing them would improve the quality of the analysis.

Overall, addressing the data gaps would enhance the business and elevate the accuracy and insights derived from the analysis.

6 Conclusion

The results gathered from the analysis of the company's data provided valuable insights into the products, types, and stores the business collaborates with. The percentage of revenue generated from international sales highlighted the importance of diversifying revenue streams and the data also showed that women were the main customers, and the peak in sales occurred at the end of the year and during the summer months. The detailed view shows that customized products made up the majority of sales, and social media and in-store purchases were the main sales channels.

Many of these conclusions were visible in the dashboards that demonstrated the importance of business intelligence in financial management. As discussed in the literature review, this analysis has an effect on both the revenue streams and the effectiveness and efficiency of the workflow in the company; and all these points could be tracked in the dashboards in real time, or when more data points are added.

Moreover, the gaps in the data, such as unidentified products, designers, and customers, may not have an immediate effect, but addressing them will improve the quality of the analysis and ultimately benefit the company. Thus, this step is as essential as adding new information to the model, since the quality of data helps to achieve optimal performance as discussed in the literature review.

Overall, this analysis provides a roadmap for improving the company and increasing revenue in the future by building business intelligence tools that could help track in real time all the information that is needed down to its smallest detail.

References

Bray T (2011) The role of business intelligence dashboards in financial management. Credit Control 32 (5/6): 29–33

Breslin J G (2019) Developing an organizational data strategy. J Bus Manag 24(1): 35-42 https://doi.org/10.3390/app11073186

Cukier K (2013) Data, data everywhere: A special report on managing information. The Economist 25 https://doi.org/10.4324/9781315856469

Davenport T H, and Harris J G (2007) Competing on analytics: The new science of winning. Harvard Business Press, https://doi.org/10.5860/choice.44-6322

Fernández A (2020) El papel del Big Data en el reporting y la toma de decisiones. Contabilidad y Direccion 31, 21–36

Gandomi A, and Haider M (2015) Beyond the hype: Big data concepts, methods, and analytics. Inter J Info Manag 35(2), 137–144 https://doi:10.1016/j.ijinfomgt.2014.10.007

Gao J, and Ren G (2020) Social media data analytics: An overview. IEEE Transac Compu Social Systems 7(6): 1621–1633

García-Sánchez F, Pérez-González D, and Pavón-Mariño P (2018) A survey of data modeling techniques for big data. Info Systems 73: 1–19

Hernández-Orallo J, Ramírez-Quintana M J, and Flach P A (2020) Explainable AI: Interpreting, explaining and visualizing deep learning. Springer https://doi.org/10.1007/978-3-030-28954-6

Kim S H, and LeeJ N (2012) The impact of organizational context and information technology on employee knowledge-sharing capabilities. Public Admin Rev 72(4), 527–536 https://doi:10.1111/j.1540-6210.2006.00595.x

Lenzerini M (2002) Data integration: A theoretical perspective. Proceedings of the ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems 233–246 https://doi:10.1145/543613.543644

Li X, Huang J, Wang L, and Lu X (2018) Big data analytics in supply chain management: A comprehensive overview and future directions. Computers Indus Engineer 116: 299–316 https://doi.org/10.3390/bdcc6010017

Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, and Byers A H (2011) Big data: The next frontier for innovation, competition, and productivity. McKinsey Global Instit: 1–156

Popescu S (2012) Business intelligence solutions - A way of general improvement of efficiency and effectiveness. Rev Inter Compar Manag / Revista de Manag Comparat Inter 13(1): 88–95

Wieder B, Ossimitz M-L, and Chamoni P (2012) The impact of business intelligence tools on performance: A user satisfaction paradox? Inter J Econ Sciences and Applied Res 5(3): 7–32

Wu X, Zhu X, Wu G Q, and Ding W (2021) Data mining with big data. IEEE Trans Know Data Engineer 33(5): 1855–1878 https://doi.org/10.1109/TKDE.2013.109

Younus A, and Qureshi M A (2021) Data analytics for business decision making: A comparative study of regression techniques. In Proceedings of the 4th International Conference on Future Networks and Distributed Systems (pp 229–241)