Final Project: WorldWideIMS

Carolina Bellani - M20170098
Gonçalo Passão - M20170450
Sofia Jerónimo - M20170070

List of Contents

Introduction to the WorldWideIMS

WorldwideIMS (WWIMS) is a 5-years old wholesale novelty goods importer and distributor operating from the San Francisco bay area.

As a wholesaler, WWIMS's customers are mostly companies who resell to individuals. WWIMS sells to retail customers across the United States including speciality stores, supermarkets, computing stores, tourist attraction shops, and some individuals. WWIMS also sells to other wholesalers via a network of agents who promote the products on WWIMS's behalf. While all of WWIMS's customers are currently based in the United States, the company is intending to push for expansion into other countries and possibly other continents.

WWIMS buys goods from suppliers including novelty and toy manufacturers, and other novelty wholesalers. They stock the goods in their WWIMS warehouse and reorder from suppliers as needed to fulfil customer orders. They also purchase large volumes of packaging materials and sell these as a convenience for the customers.

Recently, WWI started to sell a variety of edible novelties such as chilli chocolates.

From January of 2013, WWI started to record all the relevant business information into an Online Transaction Processing (OLTP) database and stored into an integrated Data Warehouse (DW).

Overview of the business scenario

From January 2013 until May 2016, the company had a total of 231,398 sales transactions, 13 suppliers of 8 categories (Clothing, Courier Services, Financial Services, Insurance Services, Marketing Services, Novelty Goods, Packaging and Toy); 18 staff employees and 663 clients of 5 categories (Computer Store, Corporate, Gift Store, Novelty Shop and Supermarket).
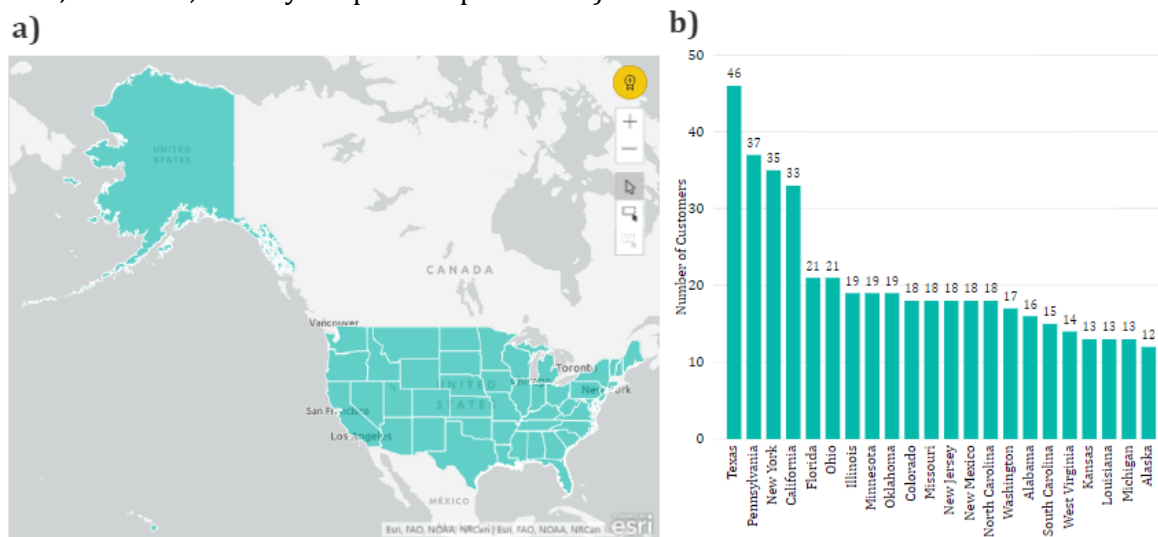
Fig. 1 - *Active customers distribution. a) Customers distribution in the United States and b) Number of customers for 20 states.*

The active clients are from all around the USA (**Fig. 1**), but these clients are predominantly located in three regions: Southern Region (Texas with 46 customers), East Coast (New York with 35 customers, Pennsylvania with 37 customers), West Coast (California with 33 customers). Therefore, 23% of the total amount of clients are located in those 3 States, while most of the northern part of the US, the company have few customers.
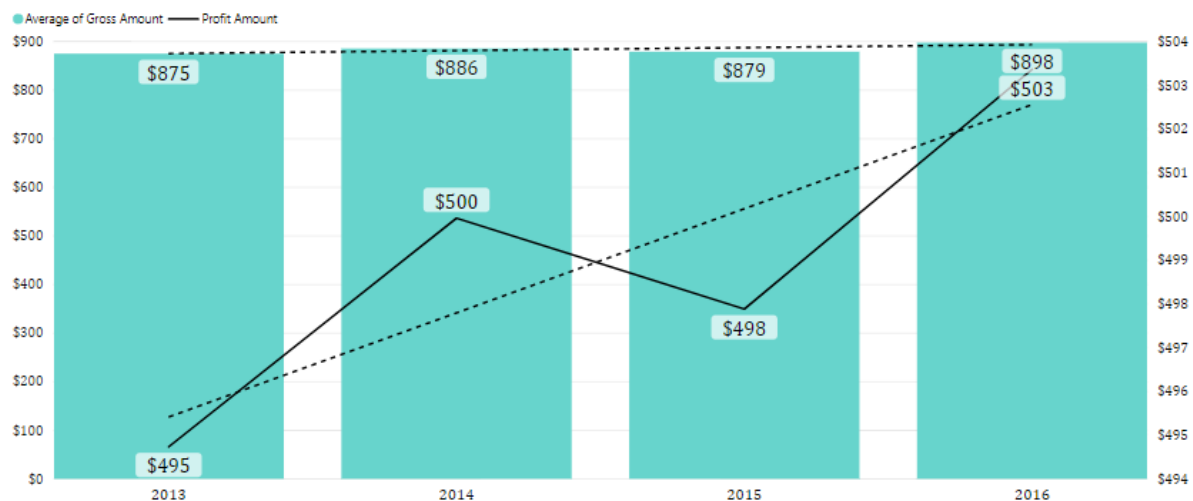
Fig. 2 – *Average gross per transaction and respective profit for each year (2013-2016). The line corresponds to the average profit per transaction and the columns to the average gross amount per transaction. The amount of sales is labelled inside the rectangle for both average gross and profit amount per transaction. The dashed lines correspond to the trend of gross and profit amount.*

The average gross per transaction is gradually increasing throughout the years, being $898 per transaction for 2016 (**Fig. 2**). There is some fluctuation in the average profit per transaction, although for 2016 has drastically increased. The trends show an increase for both average gross amount and profit amount per transaction, being a slight increase of the gross amount and a pronounced increase of the profit amount.

Fig. 3 – *Number of products per each of the categories.*

The company is currently selling 227 different products of 7 different categories (Chocolate, Clothing, Masks, Mug, Packaging, Toys and USB). The company has most variety of products for packaging and clothing (**Fig. 3**).
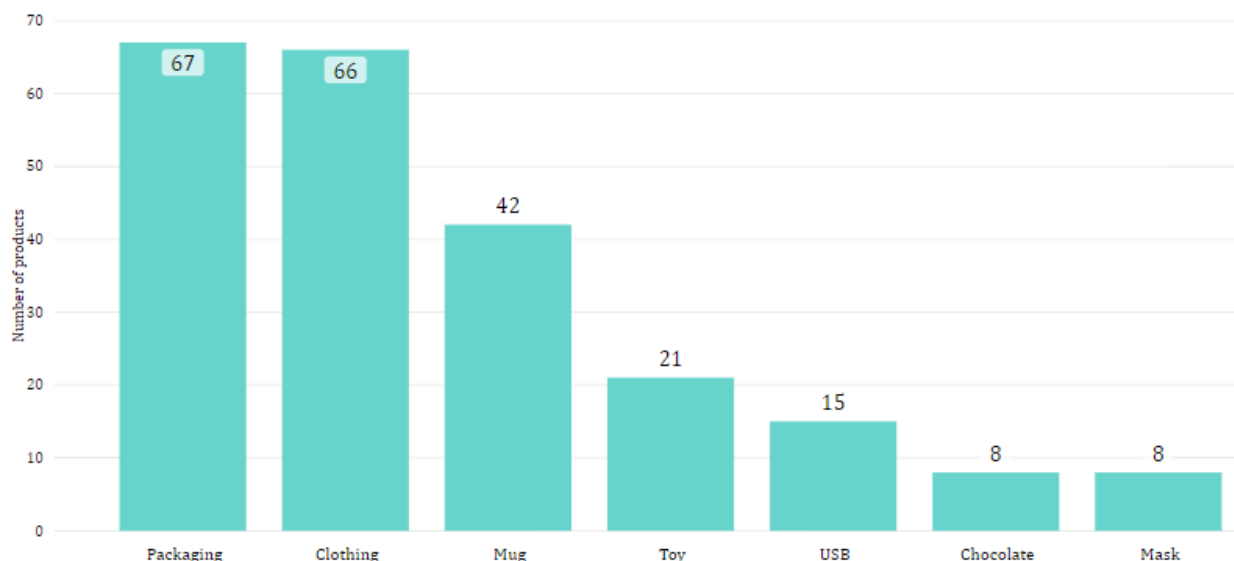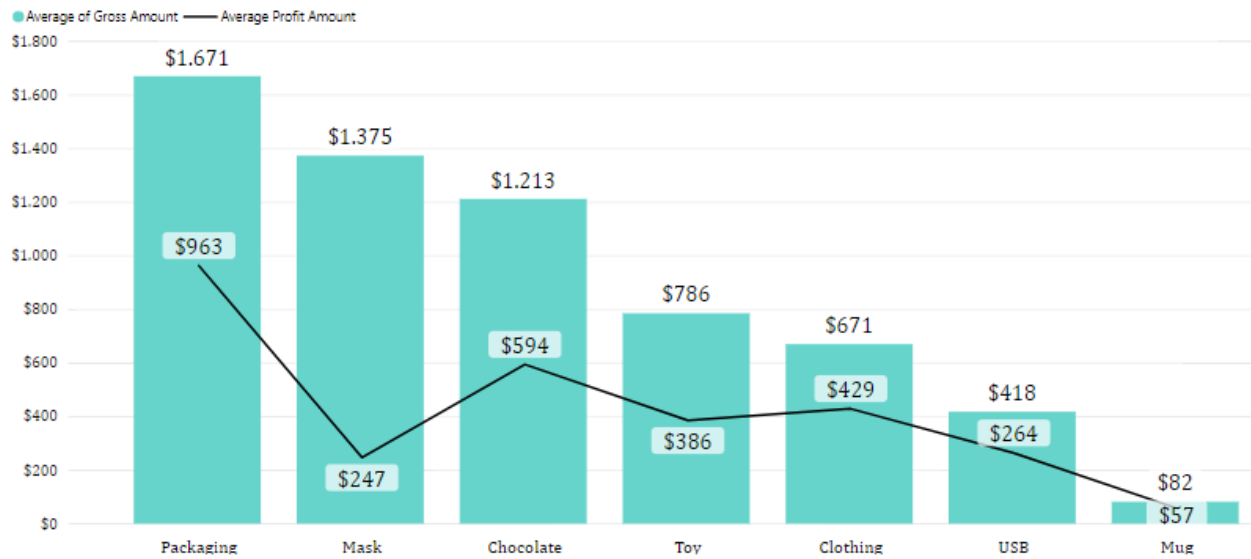
Fig. 4 – *Average gross and respective profit per category. The line corresponds to the average profit per category and the columns to the average gross amount per category.*

Packaging, Fantasy Masks and Chocolates are the most sold products by the company with most average sales and profit (**Fig. 4**). The least sold product is the mug (**Fig. 4**).

Business problem and details of business needs

Since the company has been growing throughout the years and the sales prospective for 2016 are very positive, the business executives are now willing to expand its business internally or even start an internationalization process by maintaining a sustainable growth of the company. In order to do that, we were contacted and requested to build a business solution that will assist to distribute funds to expand in different locations while maintaining the stability of the current organization in the next years. We interviewed the WWIMS's chief executive officer (CEO) and the Sales and Manufactures Departments to find out the types of questions to be answered with our analytics solution and review the reports they used to find clues about important analytical elements to analyse the maximization of profits and company operations optimization.

Main questions to be answered with the analytics solution

Table 1 - Business needs asked from the CEO and correspondent measures.

Business Need	Measures	Report
How many sales and income amounts occurred for past years and how they evolve monthly or throughout years?	Gross Amount, Profit Amount	Sales and CEO views
What were the worst and best locations in terms of profit amount and gross amount by state?	Gross Amount, Profit Amount	Sales view
What were the best 8 salespersons by name?	Gross Amount	CEO view
What was the top seller, profitable and costly products by category?	Total Gross Amount, Total Profit Amount, Profitability, Products sold, Average Gross Amount and Total Cost	Sales and CEO views
What was the top discount and full price selling products by category?	Total Gross Amount, Total Discount	Sales view
How much the clients are paying of taxes?	Client Tax Amount	Sales view
What is the top 5 profitable product characteristics (colour, category), by location, by year?	Profit Amount	Sales and CEO views
A forecast for how much revenue will the company have for the next quarter?	Profit Amount, Gross Amount	Sales and CEO views

5

Which is the category with biggest discount amount?	Discount Amount	CEO view
Which is the relation between product amount and discount amount by manufactures' locations?	Product Amount, Discount Amount	Expansion view

Description of source systems and of source data

- **WWIMS Data Warehouse**

The Data Warehouse has a star schema (***Fig. 5***) with the fact sales containing all the mentioned measures and six dimensions regarding the Date, Location, Client, Product, Supplier and Staff.

The different hierarchies for the different dimensions allow us to study different granularities concerning the sales transactions.

This Date's Dimension is divided into different hierarchies: Date (full date), Month, Trimester, Quarter, Semester, Year. If the interested date to be according to the fiscal Date, the hierarchy is fiscal Month, fiscal Year.

The hierarchy of the Client's Dimension goes from name, category of the client (Computer Store, Corporate, Gift Store, Novelty Shop and Supermarket) and discount group (none, Tailspin Toys and Wingtip Toys).

Concerning the Dimension Supplier, the hierarchy is supplier name, followed by the reference and the category.
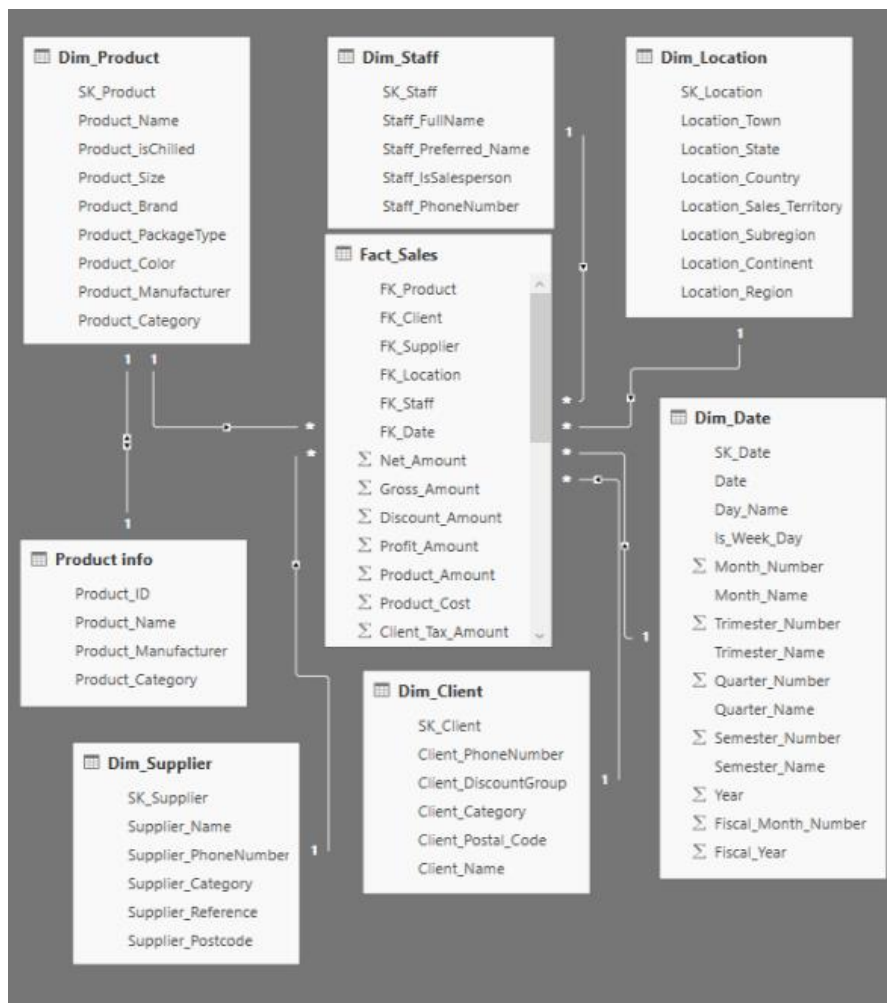

Fig. 5 – *DW Star Schema viewed in PowerBI Relationships.*

The product dimension has also a hierarchy that is the product name, then the size, brand, package type followed by the category, is chilled and manufacturer.

The hierarchy of Staff name, Staff preferred name, after the phone number and finally if it a salesperson is referred to Staff dimension.

The sixth dimension 'Location' has a hierarchy that goes from Town, State, Country, Sales Territory, Sub Region, Region and finally Continent.

- **External Data Sources**
- **Csv file: Product info**

This file corresponds to the information about the product manufacturer and category and was merged with the Product Dimension and therefore it has a relationship with that dimension (***Fig. 5***).

- **Wikipedia: countries general information**

To capture the main information of the old market locations and possible new manufactures, we uploaded the general table of Wikipedia for Canada, United States, China, Japan, Vietnam and Bangladesh countries (***Fig. 5***).

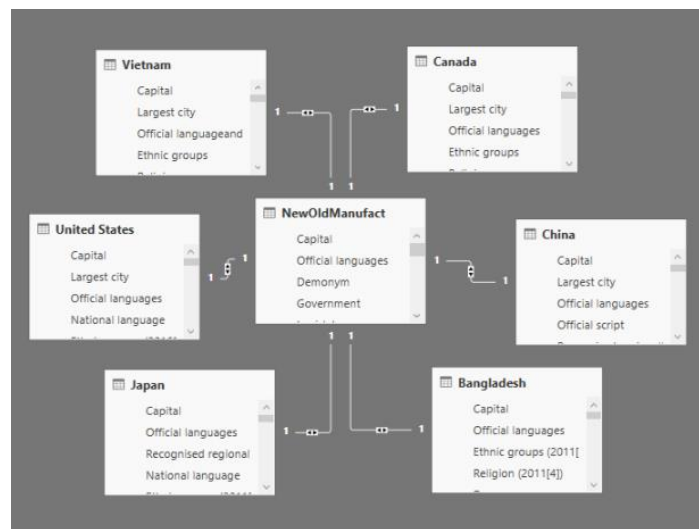***Fig. 6 –*** *External information about the interested countries.*

- **Competitors information**

We created four excel files with ten american company competitors: Wolux, Virsham, ADPIL, Prayag, Hisdustan Corporate, Brunga, JB Exports, Yaro and Kidsvilles. The details are about their locations, the sold products and the sales (***Fig. 7***). We also generated their logo.
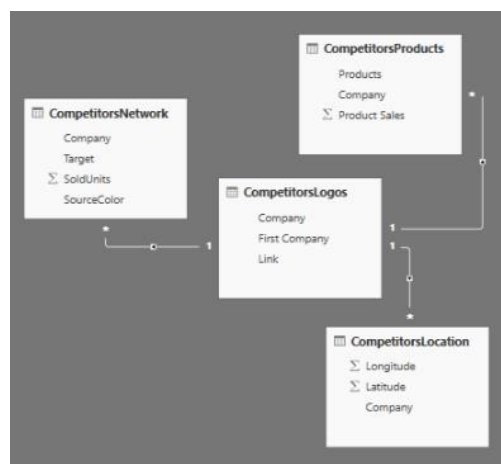

Fig. 7 – *Competitors details.*

- **Social media: WWIMS's Facebook page and Twitter #WWIMS tweets**
We extract the information about the WWIMS's Facebook page and twitter because it is extremely important to know the dynamic of the social media and the clients' feedback.

Description of the multi-dimensional schema
Our multi-dimensional schema consists in the Data Warehouse, with countries, Facebook, Twitter and competitors information.

Analytical solution
To build our model for our analytical solution, we used Power BI Desktop. It is a powerful tool that allows an easy interface to create extremely professional reports and dashboards.

- **Data integration and transformation**
- **Data Warehouse**
We got the data from the 'SQL Server database', we selected the name of the DW 'WWIMS_DW' and selected the fact tables and dimensions described above. We then checked if all the formats were correct, in some cases, some changes were necessary to do and they are described below for each one of the dimension's tables.

Dimension Client
We removed the variables concerning the slowly changing dimensions. The column of the 'Client_Name' was split by the delimiter '[' and removed because the information of the location is already in the dimension of location. The format of the postal code was replaced to whole number.
Dimension Staff
The columns for the slowly changing dimensions were removed, the data conversion row was removed from the dimension Staff. The first name and last Staff's name were separated by a '.' in the column 'Staff_FullName' and replaced by a blank space.
Dimension Supplier
In the Supplier dimension, the supplier name had 'SUP:' prefix that was removed. The 'Supplier_Postcode' format was replaced to whole number.
Dimension Product
The new column for 'Product_Category' and 'Product_Manufacturer' in the Product dimension was obtained by combining an external data 'Product_info.csv' containing the information of the categories and manufacturers of all listed products. This external data is integrated by a left-outer merge of the two queries by 'Product_Name'. There was Toy and Toys category, so we replaced those cases as Toy.
The 'Stk.' Prefix in the product name was removed. In order to have a most complete column of the colors of the products, we duplicated the column and extract the text between the delimiters, in this case, the parenthesis. To have the correct color of the product 'Developer joke mug – (hip, hip, array)', we removed the parenthesis of this cell. The blank places, 'variety', 'female' and 'male' cells were replaced with 'null'. The color 'Light Brown' was replaced to 'Brown'. By doing all this, all the colors were right assembled in the new column 'Product_Color' and replaced the previous one. The blank places, 'variety', 'female' and 'male' cells were replaced with 'null'. The color 'Light Brown' was replaced to 'Brown'.

- **Countries information**
We append the required Wikipedia information for all the countries and created a unique table called "NewOldManufact". In this table, we took in consideration to change the format of GDP pro capita, to numeric and in dollar. For the other variables, we just checked the inconsistencies and integrated with other web searches but we didn't pay attention to the format, since we want to show the information through a "Multi-row card" visualization.

- **Social media**
- o **Extraction of the Facebook data**

  We establish the connection to 'Facebook' data source by authenticating into our Facebook account and finally we selected the Object ID 'WWIMS' and connection 'posts'.

- o **WWIMS's Facebook Posts data transformation**

  We extracted the data from the posts of 'WWIMS' Facebook page, we took out the 'null' messages. To have the date and time in the correct column and format, we replaced the delimiter, in this case 'T' with a blank space. The date and time column was changed to the format 'Date/Time/Timezone'. We filter for the last 500 posts by keeping the first 500 rows.

  To get the number of comments per post, we needed to expand the 'object_link' column and select 'connections'. We repeat the procedure, but this time for 'comments' and we then aggregate the '#Count of message', this new column was named 'Count of message'.

  We removed the column 'story'.

- o **WWIMS's Facebook Comments data transformation**

  Due to the limitation of the Cognitive Services of Microsoft Azure, it is only possible to extract analyze a certain number of rows and for the time spending by the Services to analyze every comment, we decide to extract only the last 1000 comments. We think this is a satisfactory number to have a feeling of the customer satisfaction.

  In order to have the comments, in a different query 'Facebook Comments', we expanded the 'object_link' once again, select 'connections' and after 'comments'. Then we expanded 'created time', 'message' and 'id'. The date and time were handle like described before. We filter the top 1000 rows that correspond to the last comments and changed the format the text of the text by clicking in the format 'clean' and 'lowercase'.

- o **Extraction of the Twitter data**

  We started by creating a Twitter account and afterwards an App in the https://apps.twitter.com/, this will generated consumer key (API Key), consumer secret (API Secret), access Token and access Token Secret that are going to be used in the code.

  To extract the data, we used the R Script and the following code:

```
library(twitteR)
# Change the next four lines based on your own consumer_key, consume_secret, access_token, and access_secret.
consumerKey <- "insert consumer key here"
consumerSecret <- " insert consumer secret here"
accessToken <- "insert access token here"
accessSecret <- "insert access secret here"
options(httr_oauth_cache = TRUE)
setup_twitter_oauth(consumer_key = consumerKey,
          consumer_secret = consumerSecret,
          access_token = accessToken,
          access_secret = accessSecret)
my_name <- twitteR::searchTwitter("#WWIMS", n = 1000, since ='2018-01-01', retryOnRateLimit = 1e3)
my_name_df <- twListToDF(my_name)
```

  Once again, we decided to extract only the last 1000 tweets.

  After, we selected to load the table 'my_name_df' and changed the name of the table to 'Twitter'.

- o **#WWIMS Twitter data transformation**

  The extracted table had some columns that were not relevant for our case and for that reason, they were removed. The tweet's text was cleaned and lowercase. Similarly to the Facebook's comments, we only kept the first 1000 tweets that were afterwards invoked the function for the sentiment and key phrases analysis.

o **Microsoft Cognitive Services**
We used the Microsoft Cognitive Services and got our Text Analytics API key in the Microsoft Azure Portal through our subscription.
In order to have a connection to the services, we had to define the permissions for the sources 'https://westeurope.api.cognitive.microsoft.com/text/analytics/v2.0/sentiment' and 'https://westeurope.api.cognitive.microsoft.com/text/analytics/v2.0/keyPhrases' as 'Public' for the Privacy Level and the 'Credentials' to access the Web Content as 'Anonymous' in the 'Data source settings'.

o **Get Sentiment Results function**
This function was created to be used for both Facebook's comments and Tweets.
After we created a custom function in a new blank query and named it 'Get Sentiment Results'. We created the following code containing the unique API key obtained from the Microsoft Cognitive Services in the Advanced Editor:

```
let
    #"Get Sentiment Results" = (text) => let
    apikey      = "insert the api key here",
    endpoint    = "https://westeurope.api.cognitive.microsoft.com/text/analytics/v2.0/sentiment",
    jsontext    = Text.FromBinary(Json.FromValue(Text.Start(Text.Trim(text), 5000))),
    jsonbody    = "{ documents: [ { language: ""en"",id: ""0"", text: " & jsontext & " } ] }",
    bytesbody   = Text.ToBinary(jsonbody),
    headers     = [#"Ocp-Apim-Subscription-Key" = apikey],
    bytesresp   = Web.Contents(endpoint, [Headers=headers, Content=bytesbody]),
    jsonresp    = Json.Document(bytesresp),
    sentiment   = jsonresp[documents]{0}[score]
in  sentiment,
    Custom1 = #"Get Sentiment Results"
in
    Custom1
```

We then add a column 'Get Sentiment Result' in both the Facebook comments and Twitter tables for the invoke custom function, we selected the function query 'Get Sentiment Results' and for the text we selected 'Comment'(Facebook comments) and 'Tweet' (Twitter). These columns were then formatted to 'Decimal number'.
The Sentiment scores go between 0 and 1, being more than 0.5 positive and less than 0.5 negative.

o **Get Key Phrases Results function**
This function was created to be used for both Facebook's comments and Tweets.
We created another custom function in a new blank query for the 'Get Key Phrases Results' and created a very similar code than the one before:

```
let
    #"Get Key Phrases Results" = (text) => let
    apikey      = "insert the api key here",
    endpoint    = "https://westeurope.api.cognitive.microsoft.com/text/analytics/v2.0/keyPhrases",
    jsontext    = Text.FromBinary(Json.FromValue(Text.Start(Text.Trim(text), 5000))),
    jsonbody    = "{ documents: [ { language: ""en"", id: ""0"", text: " & jsontext & " } ] }",
    bytesbody   = Text.ToBinary(jsonbody),
    headers     = [#"Ocp-Apim-Subscription-Key" = apikey],
    bytesresp   = Web.Contents(endpoint, [Headers=headers, Content=bytesbody]),
    jsonresp    = Json.Document(bytesresp),
    keyphrases  = Text.Lower(Text.Combine(jsonresp[documents]{0}[keyPhrases], ", "))
in  keyphrases,
    Custom1 = #"Get Key Phrases Results"
in
    Custom1
```

We add a column 'Get Key Phrases Results' for both Facebook's comments and Tweets tables for the invoke custom function, we selected the function query 'Get Key Phrases Results' and for the text we selected 'Comment' (Facebook comments) and 'Tweet' (Twitter). The Key Phrases results are going to be used for the Cloud Analysis.

- **Reports and Dashboards**
  Business reporting is an essential part of the business intelligence movement toward improving managerial decision making. We tried to converge into dashboards and visual analytics, some prescriptive, descriptive and predictive views. WWIMS is in need of information to make accurate and timely decisions, and information is essentially the contextualization of data. It needs to ensure that the sales department is functioning properly, provide results, persuade all the parts to act.
  We will take in consideration of outcome-oriented metrics, key performance indicators, which are enterprise-wide agreed target to be tracked against over a period of time.
  To have a general and fast idea, firstly we analyzed the dataset using quick insights and we were inspired by them in the creation of the report.

- **New calculated measures**
  To achieve our business goals, we created the new measures in the Fact_sales table (**Table 2**) using DAX formula language.

Table 2. List of measures and its respective DAX code formulas.

Measures	DAX formulas
Average Gross Amount	AVERAGE(Fact_Sales[Gross_Amount])
Products Sold	COUNTROWS(Fact_Sales)
Total Client Tax Amount	SUM(Fact_Sales[Client_Tax_Amount])
Total Discount Amount	SUM(Fact_Sales[Discount_Amount])
Total Gross Amount	SUM(Fact_Sales[Gross_Amount])
Total Net Amount	SUM(Fact_Sales[Net_Amount])
Total Product Amount	SUM(Fact_Sales[Product_Amount])
Total Product Cost	SUM(Fact_Sales[Product_Cost])
Total Profit Amount	SUM(Fact_Sales[Profit_Amount])
Total Profit Amount YTD	TOTALYTD([Total Profit Amount],Dim_Date[Date])
Total Profit Amount QTD	TOTALQTD([Total Profit Amount], Dim_Date[Date])
Profitability	Fact_Sales[Total Profit Amount]/Fact_Sales[Total Gross Amount]*100
Total Cost	Fact_Sales[Total Gross Amount]-Fact_Sales[Total Profit Amount]
Product Solds YearBef	CALCULATE([Products Sold],SAMEPERIODLASTYEAR('Dim_Date'[Date]))
Total Net Amount YearBef	CALCULATE([Total Net Amount],SAMEPERIODLASTYEAR('Dim_Date'[Date]))
Total Profit Amount YearBef	CALCULATE([Total Profit Amount],SAMEPERIODLASTYEAR('Dim_Date'[Date]))

Total Product Amount Target	CALCULATE([Total Product Amount],SAMEPERIODLASTYEAR('Dim_Date'[Date]))*4

- **CEO view: overview of the company**
  A high-level overview of how the company results are going, regarding its strategy for the CEO of the company.
  It is possible to see all the main measures important for the grow of the company (Total Profit Amount, Products Sold, Average Gross Amount, Total Cost, Total Discount Amount and Profitability), following the desired granularity of the research in the card visualizations: period, region and product category. Considering the period, it is possible to select day, month and year through a slicer: we payed close attention to the 2016 year because the CEO required it, but we thought it was a good idea to have the possibility to select the past years too. Regarding the region, it is possible to select them using the tree map if they have a total profit in the selected period bigger than 7,700$.
  In the line chart, it is noticeable the presence of seasonality for the Profit and the Gross Amount and, analyzing separately, also of Discount Amount.
  A table sorter ranks the categories following the main measures: we can notice there is not totally agreement between them. The new product chilled chocolate is more sold, gives more profit and costs less than the old product mask. Packaging is the most sold category and has the best total profit amount. Two gauges of Net Amount and Profit Amount show the differences with the previous same period. A forecast for the following 150 days, useful to give an idea about what will happen for the July, August and September 2016.
  Moreover, there are the best sales people showed to be considered for an eventual bonus: Hudson is the first one and he almost duplicates the second, Jack.

- **Sales: operational view focusing in the sales**
  In this report, we can have an overall of the sales and customer preferences in relation to the products that WWIMS have available, this information is extremely relevant for the Sales Department.
  In order to allow to visualize the different plots in different granularities we add a HierarchySlicer, this allows to select the date per month, quarter and year.
  If the company want to visualize the graphs concerning the fiscal year, we add the filter called Slicer.
  We add a card visualization for Product Amount, Net Amount, Gross Amount, Profit Amount, Product Cost, Discounts amount and Client Taxes Amount.
  The Pie Chart shows the most profitable products in average by product category, which we can see that until the last upload of the DW, the winners were packaging and chocolate. The other Pie Chart have similar information, but instead of measuring the profitability, it accounts the Total Product Amount. So the top sold product is, by far, packaging related products.
  The Donut Chart informs us about the percentage of product that was sold in a certain color. In general, the most sold products are sold in brown and black (28.14% and 20.74%, respectively). This chart can be interesting when clicking in one of the categories, therefore filtering the most preferred colors for each of the categories, for instance the items of clothing tend to be both white and black (39.43% and 41,54%). These charts provide information about client preference that is important for the decision markers when deciding which product categories should the company invest the most, as well as which colors each product should have in the future.
  The top clients in terms of gross amount spent are showed in the Stacked bar chart and they are Tailspin and Wingtip Toys which are both Novelty Shops. It is important to keep a really good relation to those resellers as they have a high value for WWIMS company.
  We also provide the Funnel chart to highlight the total profit by Sales Territories and that used together with the filtering per category, it can help to better manage which products should have every stores and which stores have a low performance and consequently should close their doors.
  The Line and clustered column plot is meant to show the temporal change of the total gross amount and profit amount to see the performance of the overall company to better decide appropriate business strategies.

- o **Advanced analytics: Sales forecast**
  We use a time-series model to forecast the gross amount for the next four months by using the downloaded add-in 'Forecasting TBATS' from the marketplace that integrates a R script. Beforehand, we had to install in Rstudio or R, the necessary packages to run the code, such as: zoo, scales, reshape2, ggplot2, plotly, forecast, lubridate, htmlwidgets, XML. When this was done, we just needed to select in the forecasting settings, the forecast length and seasonal factor.
  This model helps to predict how well the business will perform in the future and help to prevent possible drops in the sales by changing the sales strategies before it is too late.

- **Expansion view**
- o **Manufactures expansion**
  From January 2016 the company chose to expand the manufactures in the US, decreasing the ones in Japan.
  To achieve the goal of growth, we compare the scenario of old and new interested manufactures:
  - Canada, close to US; it has a strategic position for the company's business
  - Bangladesh and Vietnam; they are known for the low salaries and we can see how different is the GDP per capita comparing to US and Canada.
  The scenario includes the capital, the currency, the total GDP and per capita with the respective world position and the time zone.
  Also in this case, it is possible to select the year period with a slicer.
  An important relation between sold products and discount amount is showed in the clustered bar chart, but it is also showed illustrated that the majority of the sold products are made in China in the stacked area chart and in the map.
  Below each flag, there is a multi-row card with the country scenario: the total GDP for Canada, China, Japan and US is of the order of trillion, besides of the order of billion for Vietnam and Bangladesh. A column char analyzes the GDP pro capita and in the legend, it is possible to see how the country is classified in a world view: US has the highest column and it is 11th, BD has the lowest column and it is 139th.

- o **Products selling growth**
  To maintain the stability of the current selling and to achieve the target, which consists to quadruplicate the sold products, it is possible to see in the map all the profit amount by specific region and the target compared with the actual value in the gauge. Moreover, with the attribute slicer it is possible to see the region and its profit amount and to search for the interested one.
  Texas has the highest profit but looking to the products sold target needs still improvements.

- **Social Media: WWIMS's Facebook and Twitter**
  The information about the WWIMS's Facebook page and twitter is utterly important for the Marketing team to know how dynamic is being the social media contents and the clients' feedback.
  We added the card of the most relevant features that will be very useful when combining with the other plots, such as, number of Facebook comments, number of tweets, total of retweets that the tweets had and finally the number of likes a tweet received.
  One important measure is to know which posts are receiving the most comments because this can bring insights of what people, and possibly customers, want to see in the Facebook's page. For that, we add the stacked bar chart with the number of comments for each of the 500 posts, sorted it by number of comments and showed only the top 10.
  To better understand which words are being more used, we used the 'WordCloud 1.6.0' that was download from the Marketplace. This visualization was created for both the 1000 last Facebook comments and 1000 last Tweets. After downloading the add-in, we selected the 'Get Key Phrases Results' as categories, and the count of 'Get Key Phrases Results' as values to have the total number of times that each Key Word was used in the text and thereafter, have a higher or lower word size in the Word Cloud.
  Another relevant measure is to understand the customer satisfaction about the business and its products, for that a sentiment analysis is useful to score the satisfaction and therefore, identify possible issues that

were not accounted before and need to be solved. This analysis also allows to recognize what makes a customer happy and increase those values inside the company.

This average of this measure can be visualized in the Bullet chart (that was downloaded from the marketplace) for both Facebook and Twitter, once again for the last 1000. The limits of the axis were created by the two measures Bad (=0) and Good (=1). If the average of the sentiment analysis score is higher than 0,5, the person and possibly customer is showing a positive and happy comment, if it is below 0,5, the respective comment should be read carefully in order to identify the reason behind that bad feedback.

Finally, the Timeline add-in was created for the Tweets and there you can see the Twitter username, number of retweets, likes and sentiment analysis for the respective tweet. This can also be a good tool to be used combining with other plots of the report.

- **Competitors View**
  In the words of the military general Sun Tzu "If you know both yourself and your enemy, you can win a hundred battles without jeopardy". For that reason, after we explored the different views in our company, we decided to explore our direct competition.

  First, we gathered information about the top 9 companies that might be involved in the same market as our organization. In this report we focused mostly on giving the option to the business user to explore a particular company or set of companies and the products they offer. For this reason, we created two slicers, one for the product categories and one for the specific companies.

  After the basics are defined, we need to select which business strands we want to fully explore and comprehend. With the expansion view in mind, we created a location map marking the locations of these companies, so we can fully grasp the geographical influence. Another thing to take into consideration is the interactions between all 10 companies, the 9 previously mentioned and our organization. The import/export negotiations and, most importantly, the alliances and dependencies of these companies. For that reason, we created an industry network using the custom visual "Force-Directed Graph" which shows the relationships taking a company(export) and their target(import). Finally, we created two sales graphs, one for the amount of sales by company so we can understand the money flow and the influence of each organization and another one for the sales by product category, to interpret the market share in each product category.

- **Excel analysis: Pivot table and Total sales by product color and category**

  It is possible to connect the data model created in the PowerBI with the Excel. We decided to create a new measure 'Average of cost' in Measure calculated in MDX which is [Measures].[Total Cost]/[Measures].[Products Sold].

  In the pivot table, we selected as axis the categorical variables Product_Manufacturer and in the values the new created measure 'Average of cost', this analysis allowed us to see that Japan is the country offering the best average prices, while USA has the highest production prices.

  This workbook was exported into Power BI in the Workbooks section inside the project WWI.

- **Dashboard**
  Our final dashboard aims to provide an overall view of the story told by each individual report. With this defined goal, we selected the main components of each report, these include the profit amount, sales amount, products sold, etc.

  First, we begin with the overall important values we want to show and the corresponding previsions, followed by the specific characteristics that will influence these values such as the geographical properties. After we need to take into consideration what makes the money flow in our business, the products. So, we follow the same logic as previously mentioned for the products. Finally, we integrated an extra view to our dashboard in two different ways. One was through the social media influence which can't be belittled in now days, while the other was the organizations that offer the main competition to our business and our cash flow.

In conclusion, we aimed to have a dashboard that summarizes the main aspects of the reports, although it must be taken into consideration that the dashboard on its own offers a simplified view of a complex business and must be used altogether with the individual reports.

We created an alert if the profitability was below 50%.

We integrated also market orders data streaming, with a bar chart showing in real time and for each trade type the orders.

In addition to that, a phone view of the dashboard was also done.

Conclusions

The analytical solution for WWIMS started with the examination of the business including financial statistics and the identification of the business need for a business intelligence solution. Based on these business requirements, we were able to derive some questions for building the fundamentals for the further design development. This was accomplished by the description of the source systems and external data to interpret the content and to determine the analysis solutions. We created different reports, trying to include the most relevant facts for the company. We added bookmarks to show them. At the end, we summarized the results in the dashboard.