

Declaration

I hereby declare that this project work entitled “Marketing Campaign insight Analysis has been prepared by me during the Summer Internship from 01/07/2022 to 15/08/2022 under the guidance of Ms. Khouloud Abid the CEO and Cofounder of the host company **Auzy**.

I also declare that this project is the result of my own work and my personal effort, except for extracts and summaries for which the original references are stated herein.

Acknowledgment

I would like to express my deepest and sincerest appreciation to all those who provided me with the possibility to complete this report.

I give special gratitude to my supervisor, Ms. Khouloud Abid, whose contributions in stimulating suggestions and encouragement, helped me to coordinate my project.

I also appreciate the effort that he has invested in guiding me to achieve the goal.

Last but not least, many thanks to the examiner, for agreeing to evaluate my work.

Table of Contents

General Introduction	7
Chapter 1: Internship Framework	8
I. Introduction.....	8
II. Presentation of the host company	8
III. Preliminary study	9
III.1. Introduction	9
III.2. Context of the Project	9
III.3. Study of the existing	9
III.4. Problem Statement.....	9
III.5. Proposed solution.....	9
III.6. Conclusion.....	10
Chapter 2: Technologies and environment	11
I. Introduction.....	11
II. General concepts of Business intelligence	11
II.1. BI Definition	11
II.2. Key Elements of BI Structure	11
II.3. BI Methods.....	13
II.4. BI Applications	14
II.5. Advantages of Business Intelligence	16
II.6. Disadvantages of Business Intelligence.....	17
III. Work tools and Used languages.....	17
III.1. Power BI [1]	17
III.2. DAX [2]	18
III.3. Python [3].....	19
IV. Conclusion.....	19
Chapter 3: Realized Work.....	20
I. Introduction.....	20
II. Environment Preparation	20
III. Import and Clean data in Power BI Desktop	24

IV.	Creating the Dashboard	34
V.	Conclusion	40
	Bibliography	42

Figures Table

Figure 1.1 - Company's logo.....	8
Figure 1.2 - Company's website interface.....	8
Figure 2.1- Key Elements of BI Structure.....	11
Figure 2.2-Sales overview dashboard.....	14
Figure 2.3- detailed report.....	15
Figure 2.4- Microsoft Power BI offers dashboards that give users insight into performance management.....	15
Figure 2.5- Benefits of Business Intelligence for Marketing.....	16
Figure 2.6- Enhanced Customer Experience and Increased Effectiveness.....	16
Figure 2.7– Logo power BI Desktop.....	17
Figure 2.8 - The parts of Power BI.....	18
Figure 2.9 - Logo DAX.....	18
Figure 2.10 - Logo Python.....	19
Figure 3.1- Microsoft store.....	19
Figure 3.2- browser.....	20
Figure 3.3-Download Python from python.org.....	20
Figure 3.4-Python Setup was successful.....	21
Figure 3.5- install NumPy package.....	22
Figure 3.6- upgrade the pip.....	22
Figure 3.7- Install pandas package.....	23
Figure 3.8- Install Matplotlib.....	23
Figure 3.9 – CSV file.....	24
Figure 3.10 –Data in the CSV file.....	24
Figure 3.11 – checking python scripting in Power BI.....	25
Figure 3.12 – Python script.....	25
Figure 3.13 – Import Data Command.....	26
Figure 3.14 – Loading the Data.....	26
Figure 3.15 – Data loaded.....	27

Figure 3.16 – The Data.....	27
Figure 3.17 – Data Model.....	28
Figure 3.18 – Cleaning Data rules.....	28
Figure 3.19 – Data contains fields with the wrong types.....	29
Figure 3.20 – Changing certain fields types.....	29
Figure 3.21 – run a new python script.....	30
Figure 3.22 – Adding the column “is_correct_lang” in the python script.....	30
Figure 3.23 – “is_correct_lang” column is added in the data.....	31
Figure 3.24 – Adding the column “channel_code” in the python script.....	31
Figure 3.25 – “channel_code” column is added in the data.....	32
Figure 3.26 – Adding the column “Dow” in the python script.....	32
Figure 3.27 – “Dow” column is added in the data.....	32
Figure 3.28 – Adding the column “conversion_rate” in the python script.....	33
Figure 3.29 – “conversion_rate” column is added in the data.....	33
Figure 3.30 – Adding the column “conversion_rate” in the python script.....	34
Figure 3.31 – “retention_rate” column is added in the data.....	34
Figure 3.32 – python script for the graph “Marketing channels across age groups” from the page overview.....	35
Figure 3.33 – python script for the graph “Conversion rate by age group” from the page Conversion Rate.....	35
Figure 3.34 – python script for the graph “Retention Rate by Subscribing Channel” from the page Retention Rate.....	36
Figure 3.35 – python script for the graph “Test Allocation for Email” from the page Test Allocation.....	36
Figure 3.36 – Overview page.....	37
Figure 3.37 – Conversion Rate page.....	38
Figure 3.38 – Retention Rate page.....	39
Figure 3.39 – Test Allocation page.....	40

General Introduction

As part of the second year of engineering, each student must complete a summer internship.

The internship is a way for the student to integrate into professional life to have a general idea of the functioning of the company and to enrich his knowledge.

The purpose of this internship is to make a practical intervention that allows us to apply our scientific and technical knowledge related to the field of computer science, and to acquire new ones.

Through this report, I will therefore detail the six weeks of my internship in three chapters. In the first chapter, I will start with a presentation of the host organization, the subject, and the objective of the internship. Then, I will present the technologies and software used.

Finally, I will illustrate all the work done and the results.

Chapter 1: Internship Framework

I. Introduction

In this chapter, First, we present the host organization in which our project took place, then we describe the context of the project and specify the needs and the problem of the existing, finally we propose the ultimate solution.

II. Presentation of the host company

Auzy is a medical-psycho-social support startup, created in 2019 by software and telecom engineer Khouloud Abid.

Auzy presents a mobile and web application that organizes personalized care to inform, detect, diagnose, treat and support parents in their journey to cure or control their children's mental health issues or disorders such as autism and hyperactivity.



Figure 1.1 - Company's logo

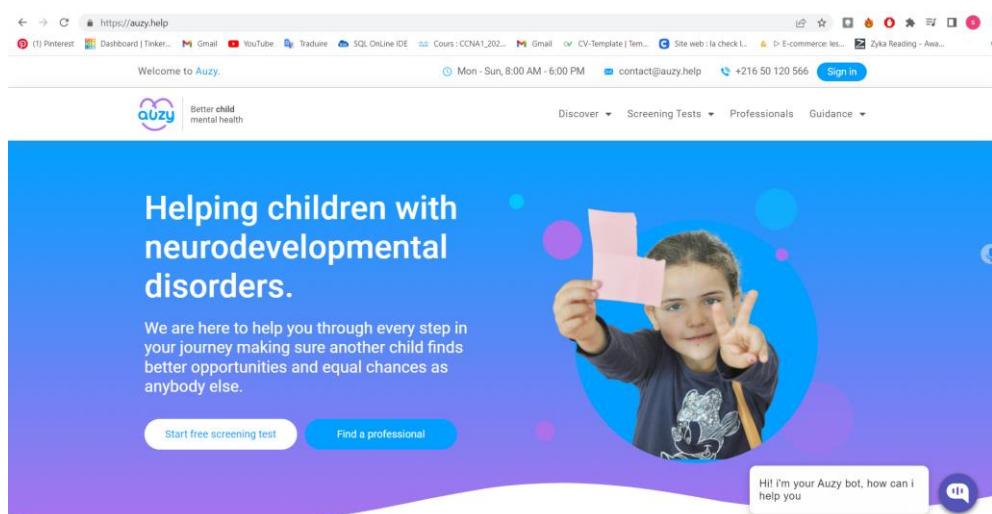


Figure 1.2 - Company's websi

III. Preliminary study

III.1. Introduction

In this chapter we describe the context of the project, then we specify the problem statement and we finally propose the ultimate solution.

III.2. Context of the Project

Whatever the services or the product that a company sells or offers, each company's primary goal is to reach the largest number of users, gain new customers, and maintain the old ones. therefore marketing campaigns become a crucial activity for any company.

III.3. Study of the existing

Auzy has Launched various marketing campaigns on different social media platforms, those marketing campaigns offer a giant amount of data, and based on this data, managers and the marketing team, are going to take decisions concerning adjusting to customer needs, improving sales indications, and optimizing the targeted campaign.

III.4. Problem Statement

The data collected from marketing campaigns is really large, and the human capabilities and skills aren't efficient enough to analyze the data and used it correctly.

In order for the managers to take the right decisions and to optimize a company's services, they need guided detailed information about the situation, the usual data formats make it hard and overwhelming for the decisions makers in a company to understand fully the situation and to be conscious of every possible scenario.

III.5. Proposed solution

In order to create a comprehensive view of the business, and to help stakeholders make better, actionable decisions we are going to use **Business Intelligence** tools to combine data from multiple sources and transform it into a digested format which going to provide the

decisions makers in a company with Data visibility, Accurate reports, and a Streamlined process.

III.6. Conclusion

In this first chapter, we defined the general framework of the Internship followed by the study of its existence which enabled us to prepare an accurate Design for our proposed solution.

In the next chapter, we will present the development and design technologies used in the realization of our solution.

Chapter 2: Technologies and environment

I. Introduction

In this chapter, first, we define the field of **Business intelligence** then we specify the elements of BI Structure and its different methods and applications, following this we indicate its advantages and its limits. Finally, we identify the work tools and the used languages in this project.

II. General concepts of Business intelligence

II.1. BI Definition

Business intelligence (BI) is a set of technology-driven processes that combines business analytics, data mining, data visualization, data tools, infrastructure and reporting, to help extract related and essential information from a large amount of internal and external data. These insights can help various organizational functions like marketing and sales enable better decision-making and calculated actions.

II.2. Key Elements of BI Structure

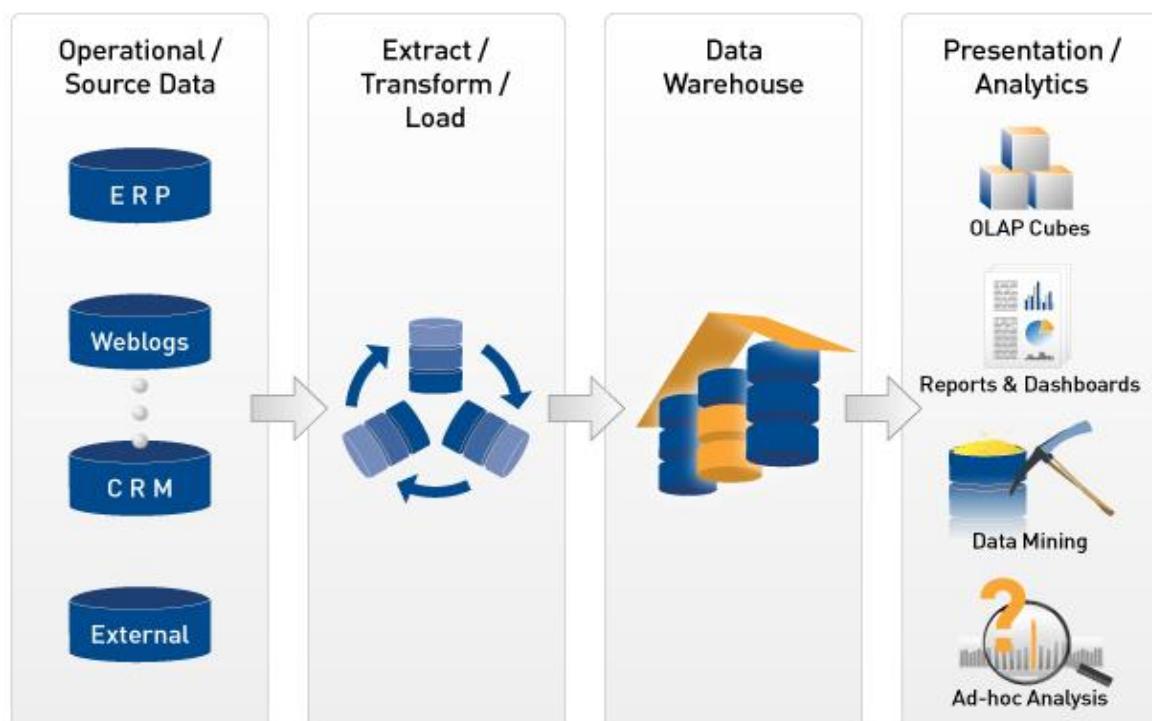


Figure 2.1- Key Elements of BI Structure

II.2.1. Data Warehouse

It's a special database or repository of data that is specially designed for supporting the decision-making process. It varies from an easy query and reporting to a more complex setup. Different methodologies such as data extraction, metadata, transformation, and loading (**ETL**) are performed in the data warehouse.

II.2.2. Business Analytics

It's the combination of different applications, technologies, and operations that are used for regular repetitive investigation and exploration of historical business performances. Business Analytics is used to have a clear view of the various inputs and forces of business planning.

It enables the development of various on-demand reports, and analysis of data and queries, and several software tools are available for the users. It's mainly included in online analytical processing (**OLAP**). Several BI techniques and tools are available for the end users which can be segmented into three categories:

- i. Reporting and Queries
- ii. Advanced Analytics
- iii. Data, Text, and Web Mining

II.2.3. Business Performance Management (BPM)

BPM is a setup of defining, executing, and controlling the organization's business strategy by establishing a relationship between the factual methods with objectives. BPM uses analysis queries and reporting.

The main objective of **BPM** is the optimization of the comprehensive performance of an organization.

II.2.4. User Interface

There are two types of user interface:

- i. **Dashboards:** Various information is organized and presented in a manner that is easy to understand. Various trends, exceptions, and key performance indicators are presented by these dashboards.
- ii. **Visualization Tools:** there are various types of visualization tools including multidimensional cube presentation to virtual reality, and technologies similar to Geographical Information Systems (**GISs**).

II.3. BI Methods

Business Intelligence has various methods of collecting, storing, and analyzing data from business operations or activities to optimize performance such as:

- **Data mining:** using databases, statistics, and machine learning (ML) to uncover trends in large datasets.
- **Reporting:** Sharing data analysis to stakeholders so they can make conclusions and decisions.
- **Performance metrics and benchmarking:** Comparing current performance data to historical data to track performance against goals, usually using customized dashboards.
- **Descriptive analytics:** Using data analysis to find out what happened
- **Querying:** to Ask the data-specific questions, the business intelligence pulls the answers from the data sets.
- **Statistical analysis:** Taking the results from descriptive analytics and exploring the data using statistics such as how this trend happened and why.
- **Data visualization:** Turning data analysis into visual representations such as charts, graphs, histograms, etc, to more easily consume data.
- **Visual analysis:** In order to communicate insights on the fly and to stay in the flow of analysis, we explore data through visual storytelling.
- **Data preparation:** Compiling and identifying the dimensions and measurement of multiple data sources, and preparing them for data analysis.

II.4. BI Applications

II.4.1. Sales Intelligence

Business intelligence gathers data on particular **KPIs** like customer demographics, conversion rates, sales metrics, etc. Then it organizes this data into structured visualizations such as graphs, pie charts, and scattergrams.

Stakeholders can identify trends from this data which provide insights into customer behavior and business operations.

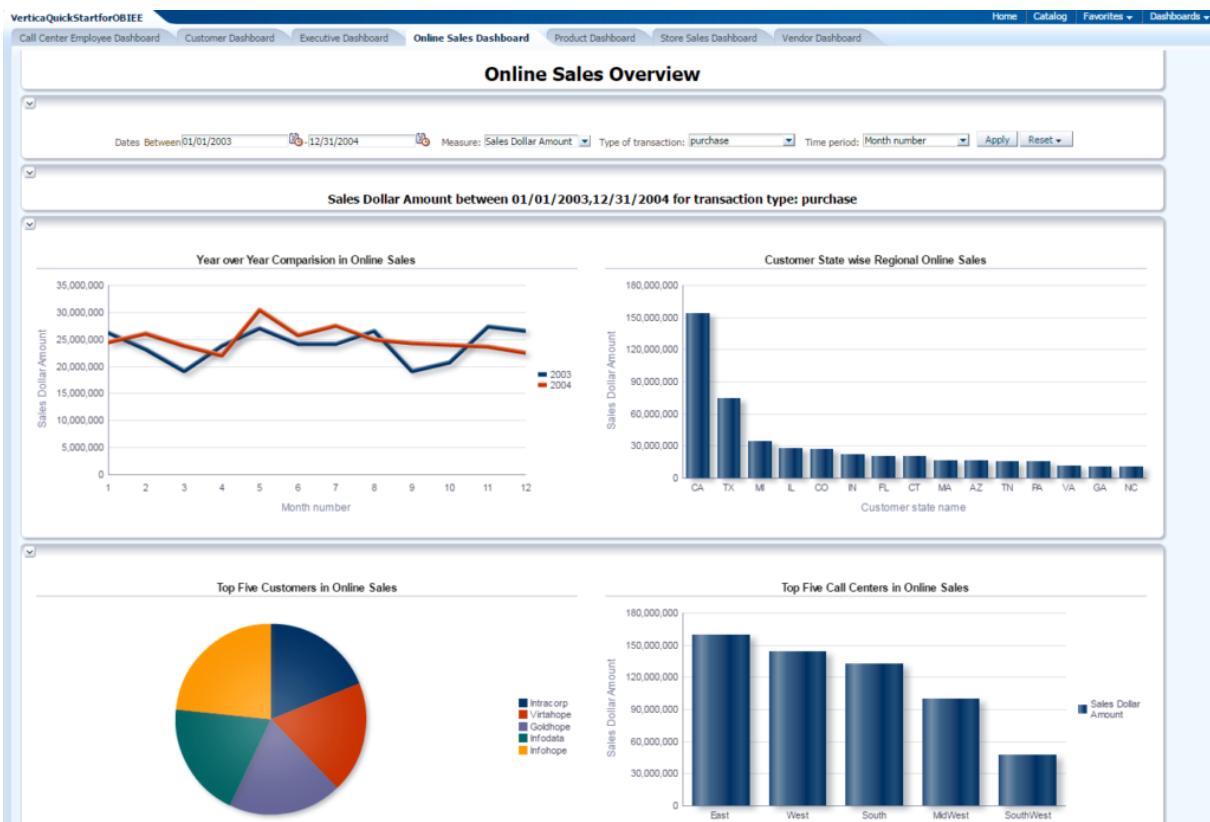


Figure 2.2-Sales overview dashboard

II.4.2. Visualization

Presenting the data in intuitive visuals and easy-to-comprehend formats, business intelligence systems allow even the least experienced employee to draw insights from data.

II.4.3. Reporting

A crucial business application of BI is reporting.

Reporting and data analysis are similar, but they vary significantly in purpose, delivery, tasks, and value. Reporting is the process of organizing data in summaries with the intention of monitoring business performance.

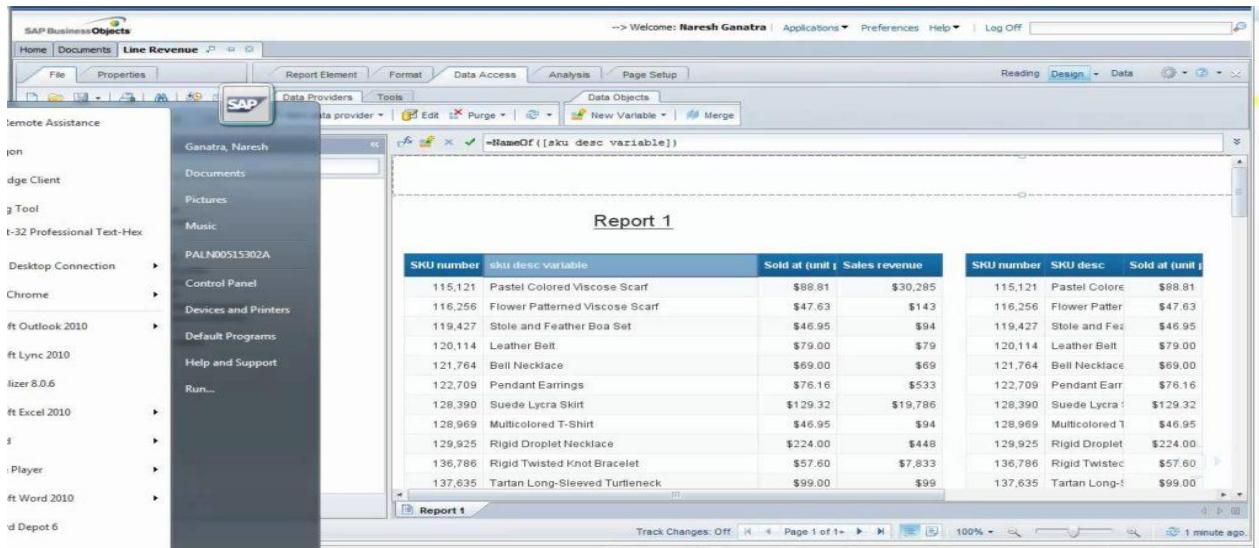


Figure 2.3- detailed report

II.4.4. Performance management

With BI applications, organizations can monitor goal progress based on pre-defined or customizable timeframes. The data-driven goals may include project completion deadlines, target delivery time, or sales goals.

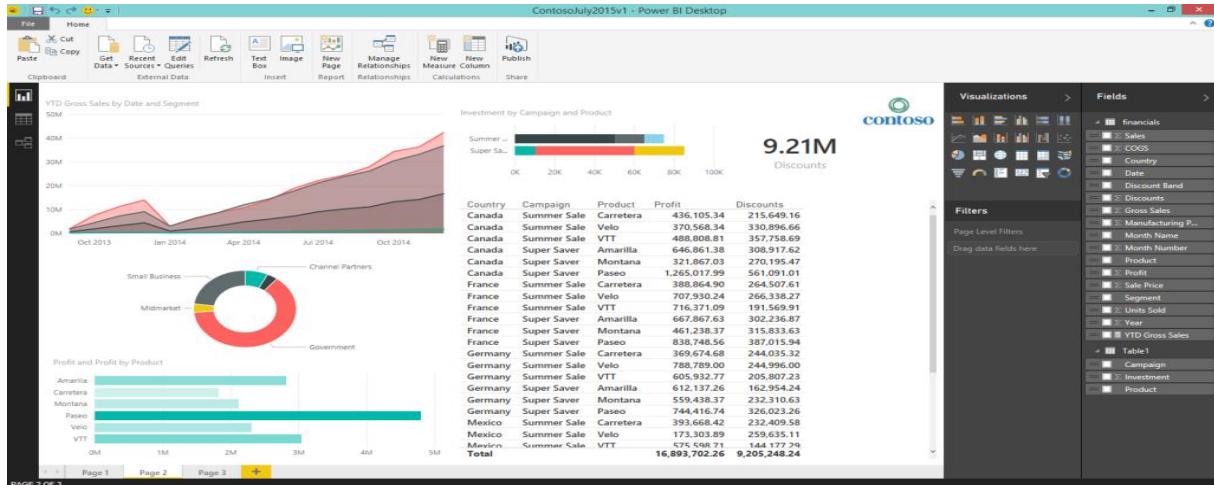


Figure 2.4- Microsoft Power BI offers dashboards that give users insight into performance management

II.5. Advantages of Business Intelligence

- Ability to gain customer insights.
- Improved Performance.
- Improved business operations visibility.
- Real-time data availability.
- Reduction of delays and backlogs.
- Better relationships and deals with customers and suppliers.
- Minimize Wastage.
- Reduced inventory costs.
- Problem detection.
- Accurate reports.



Figure 2.5- Benefits of Business Intelligence for Marketing



Figure 2.6- Enhanced Customer Experience and Increased Effectiveness

II.6. Disadvantages of Business Intelligence

- A lot of Historical Data.
- Cost.
- Difficulty: Data skills gap.
- Disordering of Commercial Set-ups.
- Limited Applications.
- Implementation is a Time-Consuming Event.
- User resistance.

III. Work tools and Used languages

III.1. Power BI [1]



Figure 2.7– Logo power BI Desktop

Power BI is a collection of software services, apps, and connectors that work together to turn your unrelated sources of data into coherent, visually immersive, and interactive insights. Your data may be an Excel spreadsheet, or a collection of cloud-based and on-premises hybrid data warehouses. Power BI lets you easily connect to your data sources, visualize and discover what's important, and share that with anyone or everyone you want.

Power BI consists of several elements that all work together, starting with these three basics:

- ❖ A Windows desktop application called **Power BI Desktop**.
- ❖ An online SaaS (Software as a Service) service called the **Power BI service**.
- ❖ Power BI **mobile apps** for Windows, iOS, and Android devices.

Beyond those three, Power BI also features two other elements:

- **Power BI Report Builder**, for creating paginated reports to share in the Power BI service. Read more about [paginated reports](#) later in this article.
- **Power BI Report Server**, an on-premises report server where you can publish your Power BI reports, after creating them in Power BI Desktop. Read more about [Power BI Report Server](#) later in this article.

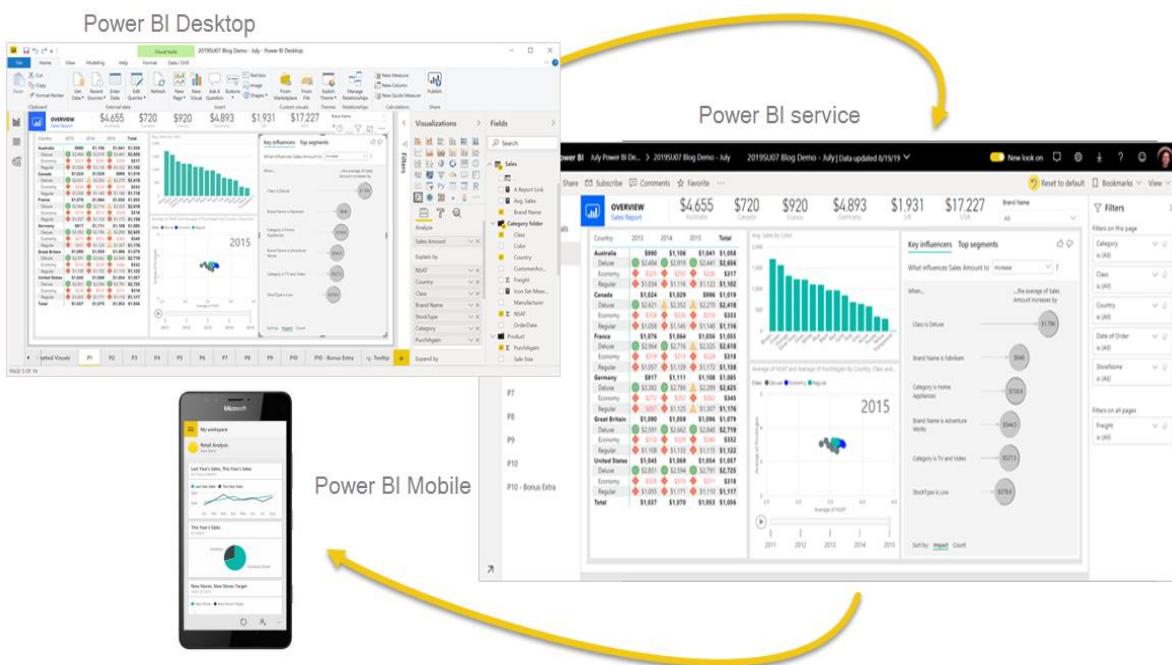


Figure 2.8 - The parts of Power BI

III.2. DAX [2]



Figure 2.9 - Logo DAX

Data Analysis Expressions (DAX) is a formula expression language used in Analysis Services, Power BI, and Power Pivot in Excel. DAX formulas include functions, operators, and values to perform advanced calculations and queries on data in related tables and columns in tabular data models.

This article provides only a basic introduction to the most important concepts in DAX. It describes DAX as it applies to all the products that use it. Some functionality may not apply to certain products or use cases. Refer to your product's documentation describing its particular implementation of DAX.

DAX formulas are used in measures, calculated columns, calculated tables, and row-level security.

III.3. Python [3]



Figure 2.10 - Logo Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance.

Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

IV. Conclusion

In this chapter, we clarified the field of business intelligence and we specified the tools and technologies we used in this work, in the next chapter we are going to demonstrate the realized work.

Chapter 3: Realized Work

I. Introduction

In this chapter, we will go through the steps of creating this detailed project with code, interfaces, and final result.

II. Environment Preparation

First, we install the power BI desktop from the Microsoft store.



Figure 3.1- Microsoft store.

It can also be installed through the browser: <https://powerbi.microsoft.com/en-us/downloads/>

 A screenshot of a web browser displaying the Power BI download page. The header shows the URL "https://powerbi.microsoft.com/en-us/downloads/". The main content features a yellow banner with the text "Download Power BI tools and apps". Below the banner, there are three sections: "Microsoft Power BI Desktop", "Microsoft Power BI Mobile", and "Microsoft on-premises data gateway". Each section includes a brief description and download links for Microsoft, App Store, and Google Play.

Tool	Description	Download Options
Microsoft Power BI Desktop	With the Power BI Desktop you can visually explore your data through a free-form drag-and-drop canvas, a broad range of modern data visualizations, and an easy-to-use report authoring experience.	Download > Advanced download options >
Microsoft Power BI Mobile	Access your data anywhere, anytime. These native apps provide live, interactive, mobile access to your important business information.	Get it from Microsoft Download on the App Store GET IT ON Google Play
Microsoft on-premises data gateway	Keep your dashboards and reports up to date by connecting to your on-premises data sources—without the need to move the data.	Download standard mode > Download personal mode >

Figure 3.2- browser.

To run Python scripts in Power BI Desktop, we need to install Python on our local machine.

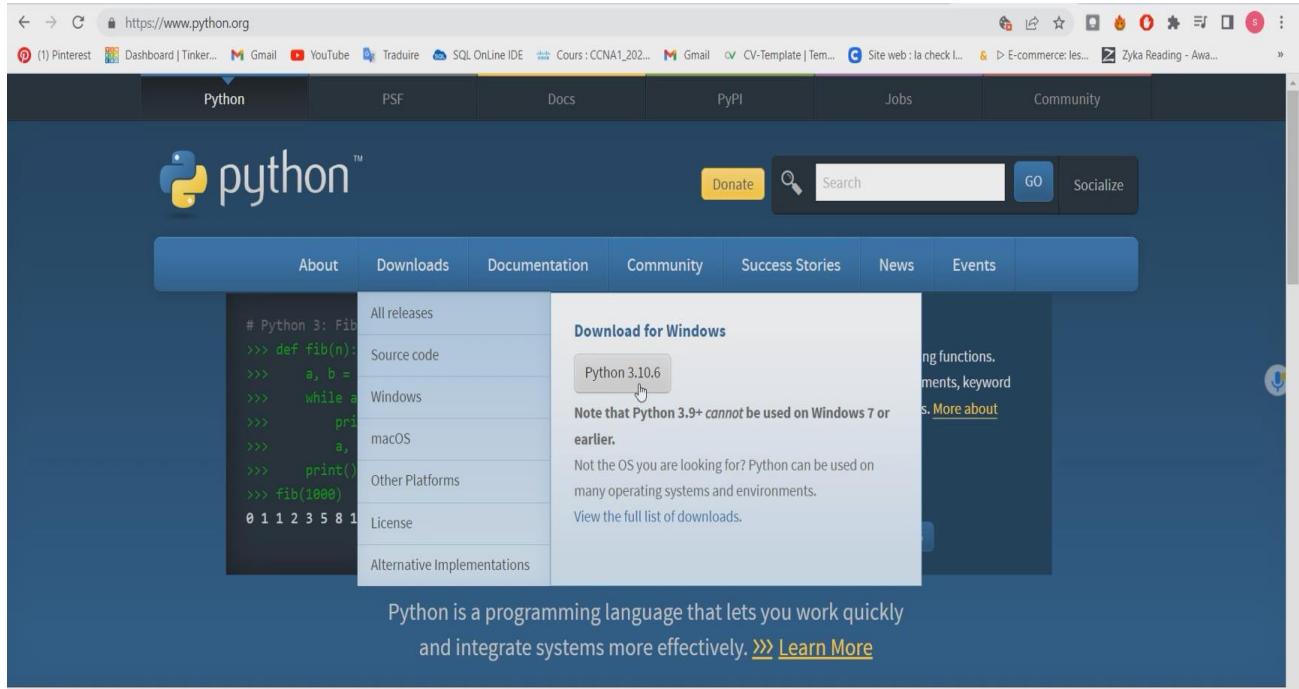


Figure 3.3-Download Python from python.org

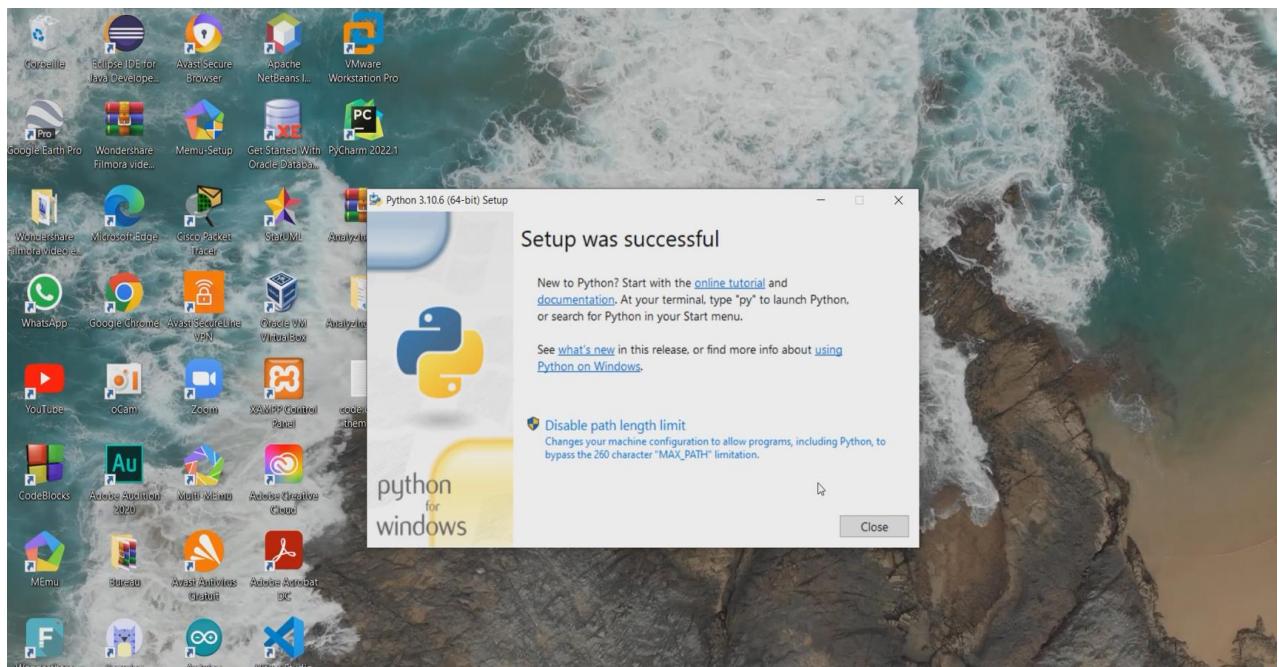


Figure 3.4-Python Setup was successful

The integration of Python into Power BI requires the installation of different Python packages.

Those packages help us create the dashboard with its various charts on the Power BI desktop.

First, we install NumPy package using this command:

Python -m pip install numpy

- ❖ **NumPy [4]:** is the **fundamental package** for scientific computing in Python.

It is a Python library that provides a multidimensional array object, various derived objects (such as masked arrays and matrices), and an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.

```
C:\Users\LENOVO>python -m pip install numpy
Collecting numpy
  Downloading numpy-1.23.2-cp310-cp310-win_amd64.whl (14.6 MB)
    14.6/14.6 MB 2.2 MB/s eta 0:00:00
Installing collected packages: numpy
Successfully installed numpy-1.23.2

[notice] A new release of pip available: 22.2.1 -> 22.2.2
[notice] To update, run: python.exe -m pip install --upgrade pip

C:\Users\LENOVO>python -m pip install --upgrade pip
Requirement already satisfied: pip in c:\users\lenovo\appdata\local\programs\python\python3
Collecting pip
  Downloading pip-22.2.2-py3-none-any.whl (2.0 MB)
    2.0/2.0 MB 2.1 MB/s eta 0:00:00
Installing collected packages: pip
  Attempting uninstall: pip
    Found existing installation: pip 22.2.1
    Uninstalling pip-22.2.1:
      Successfully uninstalled pip-22.2.1
Successfully installed pip-22.2.2

C:\Users\LENOVO>
```

Figure 3.5- install NumPy package

Before installing the second package we need to upgrade the pip package using this command

Python -m pip install –upgrade pip

- ❖ **Pip [5] :** Package Installer for Python (pip) is the de facto and recommended^[5] package-management system written in Python and is used to install and manage software packages.

```
C:\Users\LENOVO>python -m pip install --upgrade pip
Requirement already satisfied: pip in c:\users\lenovo\appdata\local\programs\python\python38-32\script
Collecting pip
  Downloading pip-22.2.2-py3-none-any.whl (2.0 MB)
    2.0/2.0 MB 2.1 MB/s eta 0:00:00
Installing collected packages: pip
  Attempting uninstall: pip
    Found existing installation: pip 22.2.1
    Uninstalling pip-22.2.1:
      Successfully uninstalled pip-22.2.1
Successfully installed pip-22.2.2

C:\Users\LENOVO>
```

Figure 3.6- upgrade the pip

After that we install pandas' package through this command:

Python -m pip install pandas

- ❖ **Pandas [6]:** pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with “relational” or “labeled” data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, **real-world** data analysis in Python.

```
C:\Users\LENOVO>python -m pip install pandas
Collecting pandas
  Downloading pandas-1.4.3-cp310-cp310-win_amd64.whl (10.5 MB)
    10.5/10.5 MB 2.3 MB/s eta 0:00:00
Collecting pytz>=2020.1
  Using cached pytz-2022.2.1-py2.py3-none-any.whl (500 kB)
Collecting python-dateutil>=2.8.1
  Using cached python_dateutil-2.8.2-py2.py3-none-any.whl (247 kB)
Requirement already satisfied: numpy>=1.21.0 in c:\users\lenovo\appdata\local\programs\python\python38-32\script
Collecting six>=1.5
  Using cached six-1.16.0-py2.py3-none-any.whl (11 kB)
Installing collected packages: pytz, six, python-dateutil, pandas
Successfully installed pandas-1.4.3 python-dateutil-2.8.2 pytz-2022.2.1 six-1.16.0

C:\Users\LENOVO>
```

Figure 3.7- Install pandas package

Finally, we install the last package that we need to integrate the python script in Power BI which is Matplotlib through this command:

Python -m pip install Matplotlib

- ❖ **Matplotlib [7]:** Matplotlib is a cross-platform, data visualization and graphical plotting library for Python and its numerical extension NumPy. As such, it offers a viable open-source alternative to MATLAB. Developers can also use matplotlib's APIs (Application Programming Interfaces) to embed plots in GUI applications.

```
C:\Users\LENOVO>python -m pip install matplotlib
Collecting matplotlib
  Downloading matplotlib-3.5.3-cp310-cp310-win_amd64.whl (7.2 MB)
    7.2/7.2 MB 2.2 MB/s eta 0:00:00
Collecting packaging>=20.0
  Using cached packaging-21.3-py3-none-any.whl (40 kB)
Collecting pillow>=6.2.0
  Downloading Pillow-9.2.0-cp310-cp310-win_amd64.whl (3.3 MB)
    3.3/3.3 MB 1.9 MB/s eta 0:00:00
Collecting cycler>=0.10
  Using cached cycler-0.11.0-py3-none-any.whl (6.4 kB)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\lenovo\appdata\local\programs\python\python310\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: numpy>=1.17 in c:\users\lenovo\appdata\local\programs\python\python310\lib\site-packages (from matplotlib) (1.23.2)
Collecting pytz>=2.2.1
  Downloading pytz-2022.1-py3-none-any.whl (98 kB)
    98.3/98.3 kB 1.9 MB/s eta 0:00:00
Collecting kiwisolver>=1.0.1
  Downloading kiwisolver-1.4.4-cp310-cp310-win_amd64.whl (55 kB)
    55.3/55.3 kB 2.8 MB/s eta 0:00:00
Collecting fonttools>=4.22.0
  Downloading fonttools-4.37.1-py3-none-any.whl (957 kB)
    957.2/957.2 kB 2.3 MB/s eta 0:00:00
Requirement already satisfied: six>=1.5 in c:\users\lenovo\appdata\local\programs\python\python310\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
Installing collected packages: pytz, pillow, kiwisolver, fonttools, cycler, packaging, matplotlib
Successfully installed cycler-0.11.0 fonttools-4.37.1 kiwisolver-1.4.4 matplotlib-3.5.3 packaging-21.3 pillow-9.2.0 pytz-2022.1
C:\Users\LENOVO>
C:\Users\LENOVO>
```

Figure 3.8- Install Matplotlib

III. Import and Clean data in Power BI Desktop

The file that contains the data of Auzy's marketing campaign is an **csv file**.

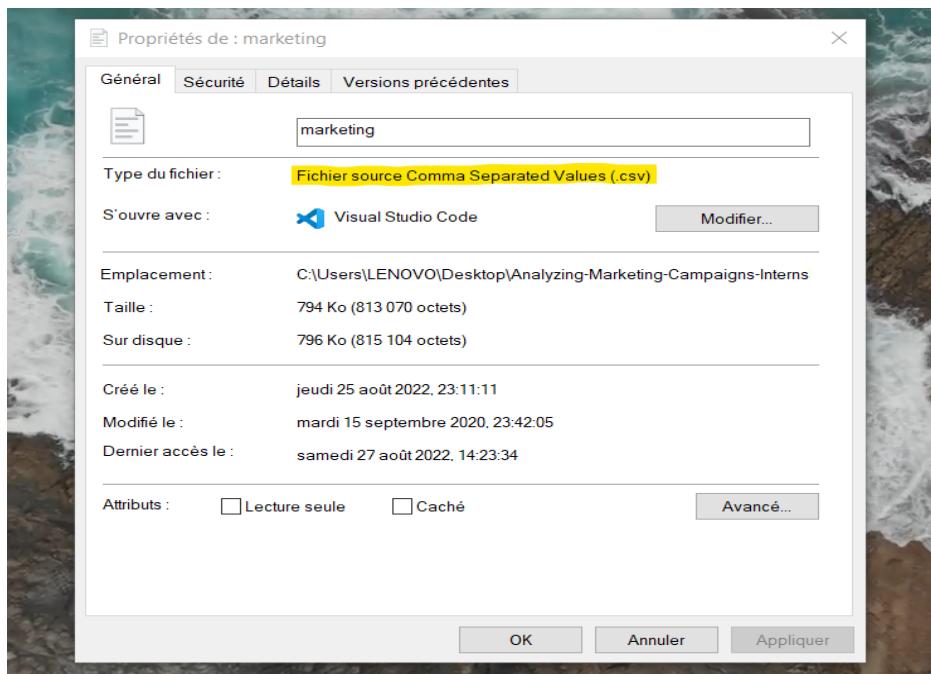


Figure 3.9 – CSV file

Realized Work

```

marketing.csv - Visual Studio Code
marketing.csv
C:\Users\LENOVO\Desktop\Analyzing-Marketing-Campaigns-Interdip> marketing.csv

1 user_id,date_served,marketing_channel,variant,converted,language_displayed,language_preferred,age_group,date_subscribed,date_canceled,subscribing_channel,is_retained
2 a100000029,1/1/18,House Ads,personalization,TRUE,English,English,0-18 years,1/1/18,,House Ads,TRUE
3 a100000030,1/1/18,House Ads,personalization,TRUE,English,English,19-24 years,1/1/18,,House Ads,TRUE
4 a100000031,1/1/18,House Ads,personalization,TRUE,English,English,24-30 years,1/1/18,,House Ads,TRUE
5 a100000032,1/1/18,House Ads,personalization,TRUE,English,English,30-36 years,1/1/18,,House Ads,TRUE
6 a100000033,1/1/18,House Ads,personalization,TRUE,English,English,36-45 years,1/1/18,,House Ads,TRUE
7 a100000034,1/1/18,House Ads,personalization,TRUE,German,German,45-55 years,1/1/18,,House Ads,TRUE
8 a100000035,1/1/18,House Ads,personalization,TRUE,English,English,55+ years,1/1/18,,House Ads,TRUE
9 a100000036,1/1/18,House Ads,personalization,TRUE,English,English,0-18 years,1/1/18,,House Ads,TRUE
10 a100000037,1/1/18,House Ads,personalization,TRUE,English,English,19-24 years,1/1/18,,House Ads,TRUE
11 a100000038,1/1/18,House Ads,personalization,TRUE,English,English,24-30 years,1/1/18,,House Ads,TRUE
12 a100000039,1/1/18,House Ads,personalization,TRUE,English,English,0-18 years,1/1/18,,House Ads,TRUE
13 a100000040,1/1/18,House Ads,personalization,TRUE,English,English,19-24 years,1/1/18,1/18/18,House Ads,FALSE
14 a100000041,1/1/18,House Ads,personalization,TRUE,Arabic,Arabic,24-30 years,1/1/18,2/22/18,House Ads,FALSE
15 a100000042,1/1/18,House Ads,personalization,TRUE,English,English,0-18 years,1/1/18,3/9/18,House Ads,FALSE
16 a100000043,1/1/18,House Ads,personalization,TRUE,English,English,19-24 years,1/1/18,2/18/18,House Ads,FALSE
17 a100000044,1/1/18,House Ads,personalization,TRUE,Arabic,Arabic,24-30 years,1/1/18,3/18/18,House Ads,FALSE
18 a100000045,1/2/18,House Ads,personalization,TRUE,English,English,0-18 years,1/2/18,2/4/18,House Ads,FALSE
19 a100000046,1/2/18,House Ads,personalization,TRUE,English,English,19-24 years,1/2/18,1/9/18,House Ads,FALSE
20 a100000047,1/2/18,House Ads,personalization,TRUE,English,English,24-30 years,1/2/18,2/4/18,House Ads,FALSE
21 a100000048,1/2/18,House Ads,personalization,TRUE,German,German,0-18 years,1/2/18,2/21/18,House Ads,FALSE
22 a100000049,1/2/18,House Ads,personalization,TRUE,German,German,19-24 years,1/2/18,House Ads,TRUE
23 a100000050,1/2/18,House Ads,personalization,TRUE,English,English,24-30 years,1/2/18,House Ads,TRUE
24 a100000050,1/1/18,House Ads,personalization,FALSE,English,English,24-30 years,1/2/18,,House Ads,TRUE
25 a100000051,1/2/18,House Ads,personalization,TRUE,English,English,30-36 years,1/2/18,,House Ads,TRUE
26 a100000051,1/1/18,Push,personalization,FALSE,English,English,30-36 years,1/2/18,,House Ads,TRUE
27 a100000052,1/2/18,House Ads,personalization,TRUE,English,English,36-45 years,1/2/18,,House Ads,TRUE
28 a100000053,1/1/18,House Ads,personalization,FALSE,English,English,36-45 years,1/2/18,,House Ads,TRUE
29 a100000053,1/2/18,House Ads,personalization,FALSE,English,English,45-55 years,1/2/18,,House Ads,TRUE
30 a100000054,1/2/18,House Ads,personalization,TRUE,English,English,45-55 years,1/2/18,,House Ads,TRUE
31 a100000054,1/1/18,House Ads,personalization,TRUE,English,English,55+ years,1/2/18,,House Ads,TRUE
32 a100000054,1/2/18,House Ads,personalization,FALSE,English,English,55+ years,1/2/18,,House Ads,TRUE
33 a100000055,1/2/18,House Ads,personalization,TRUE,English,English,0-18 years,1/2/18,,House Ads,TRUE
34 a100000055,1/1/18,Push,personalization,FALSE,English,English,0-18 years,1/2/18,,House Ads,TRUE
35 a100000056,1/2/18,House Ads,personalization,TRUE,English,English,19-24 years,1/2/18,,House Ads,TRUE
36 a100000057,1/2/18,House Ads,personalization,TRUE,English,English,24-30 years,1/2/18,,House Ads,TRUE
37 a100000057,1/1/18,Facebook,personalization,FALSE,English,English,24-30 years,1/2/18,,House Ads,TRUE

```

Figure 3.10 –Data in the CSV file

To import the CSV file first we need to check that the python scripting is integrated into Power BI Desktop.

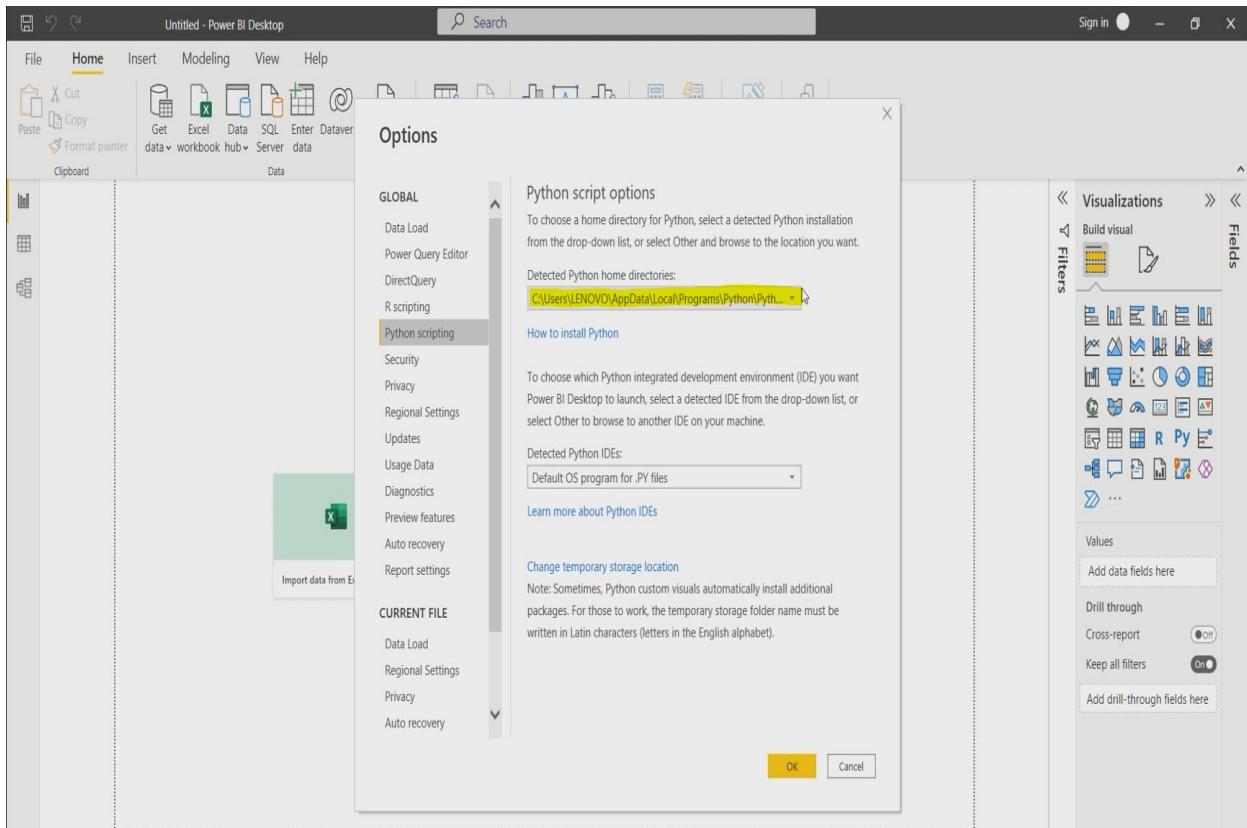


Figure 3.11 – checking python scripting in Power BI

To import data, we open a python script

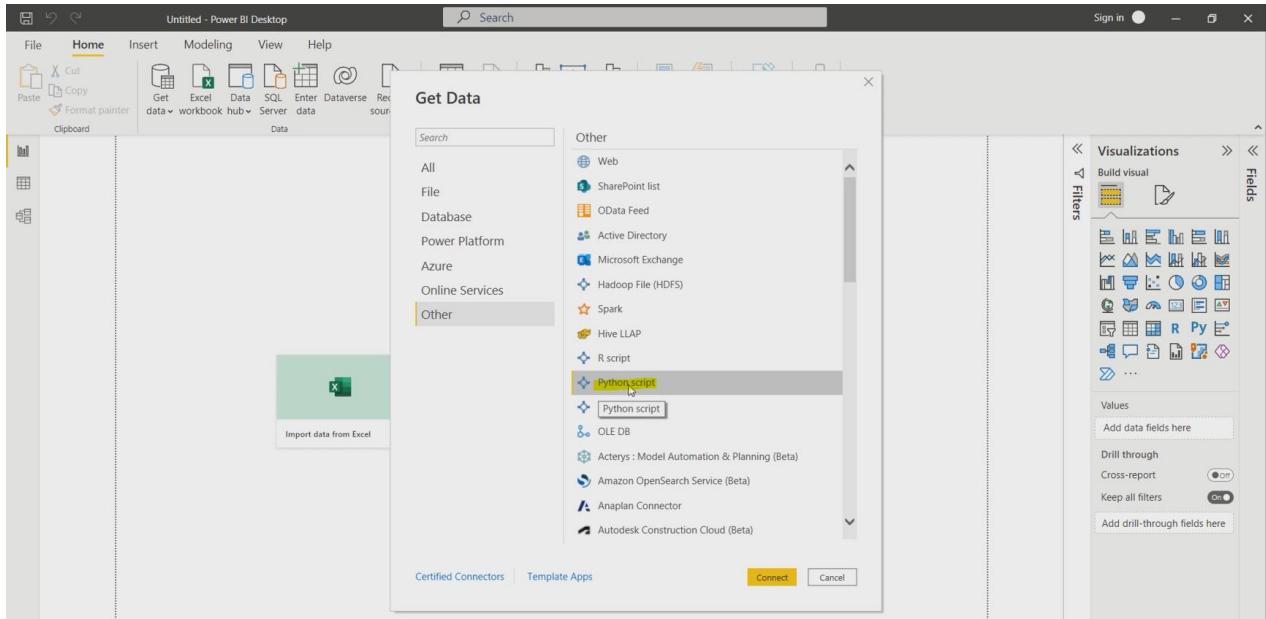


Figure 3.12 – Python script

Then we write this command:

```
marketing=pandas.read_csv('C:/Users/LENOVO/Desktop/Analyzing-Marketing-
Campaigns-Internship/marketing.csv')
```

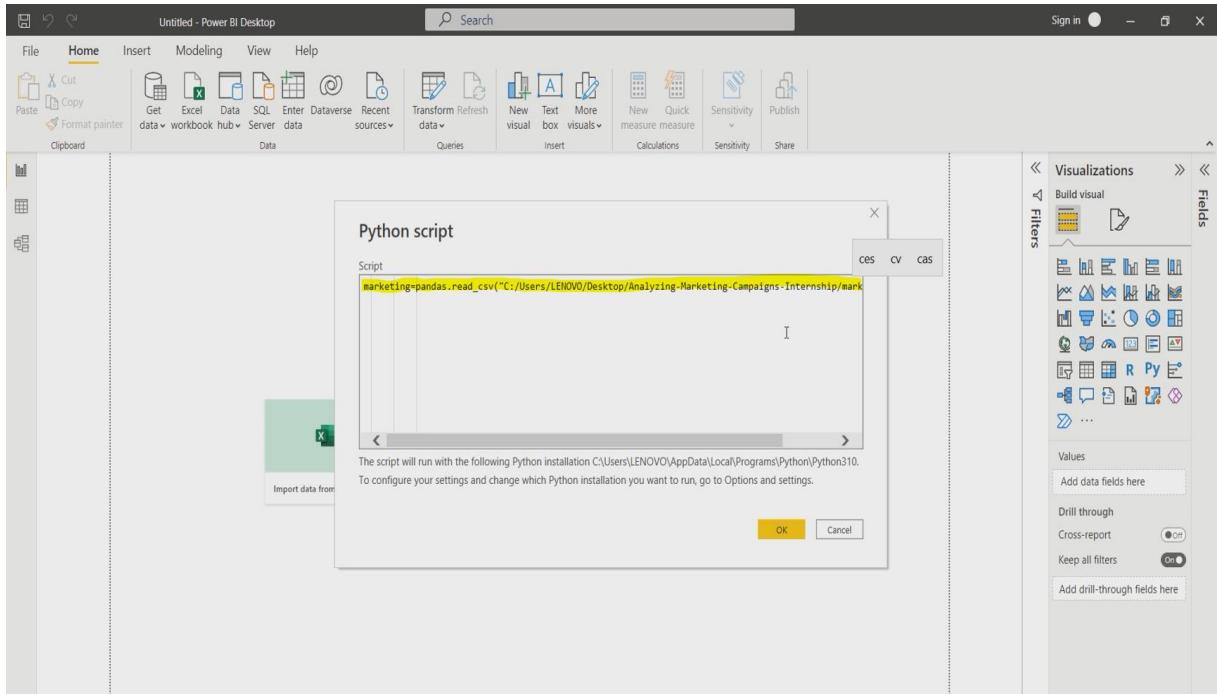


Figure 3.13 – Import Data Command

We select marketing and we click load to import the data.

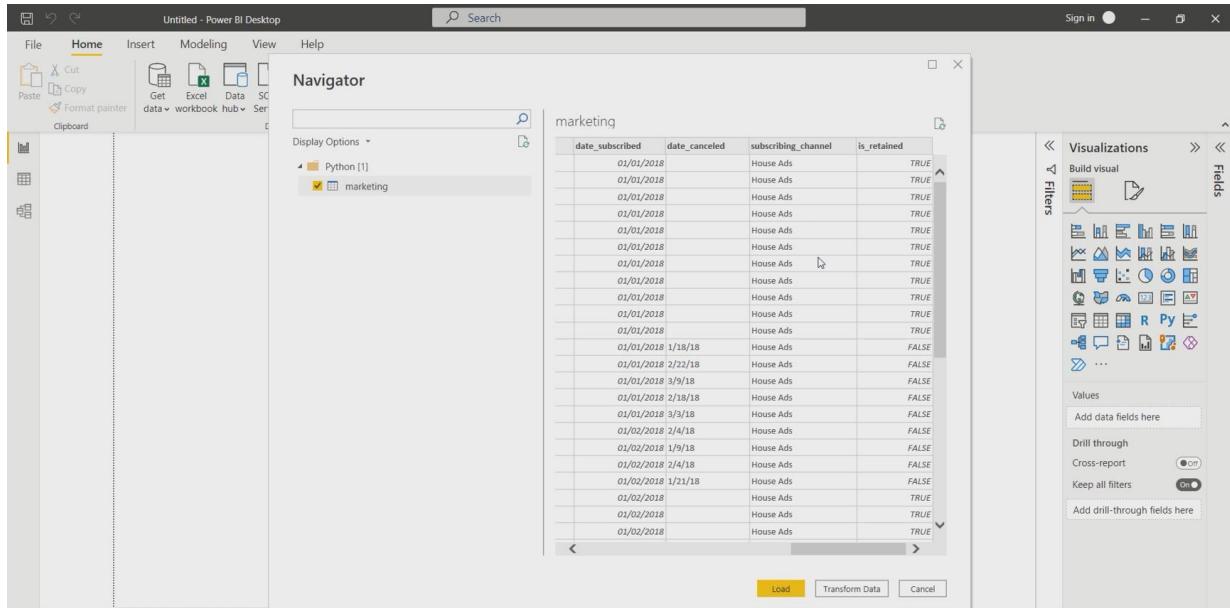


Figure 3.14 – Loading the Data

After the data is loaded, we can see the different fields of the data like age-group, converted, etc.

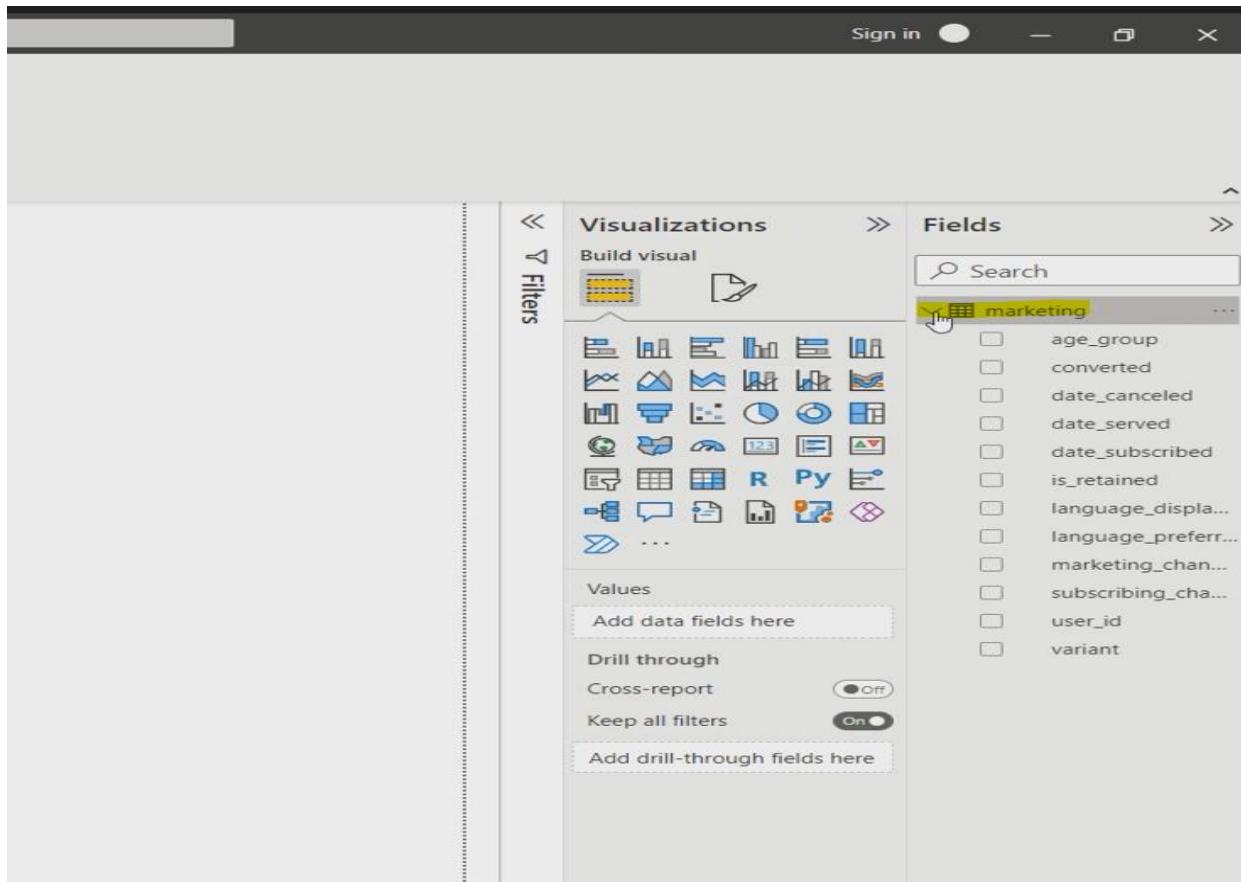
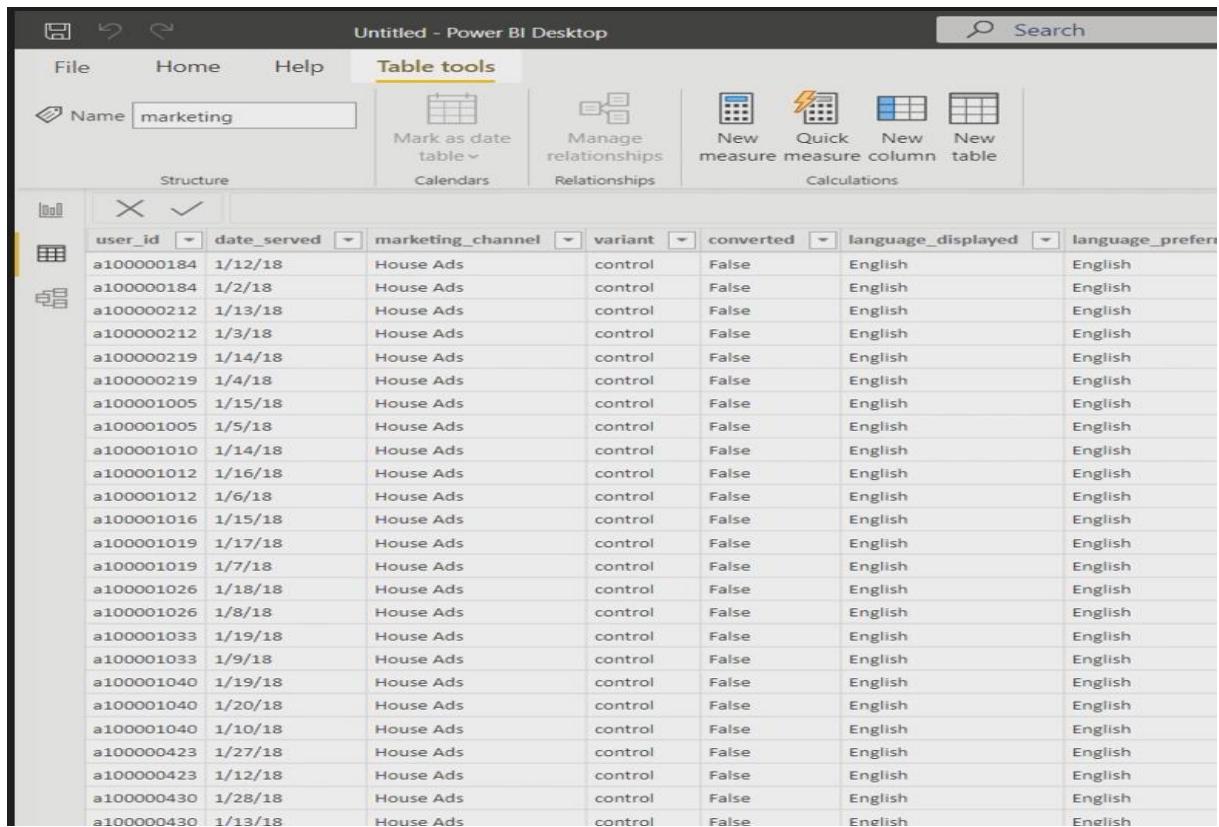


Figure 3.15 – Data loaded

We can view all the data in the form of a tables in Power BI Desktop.

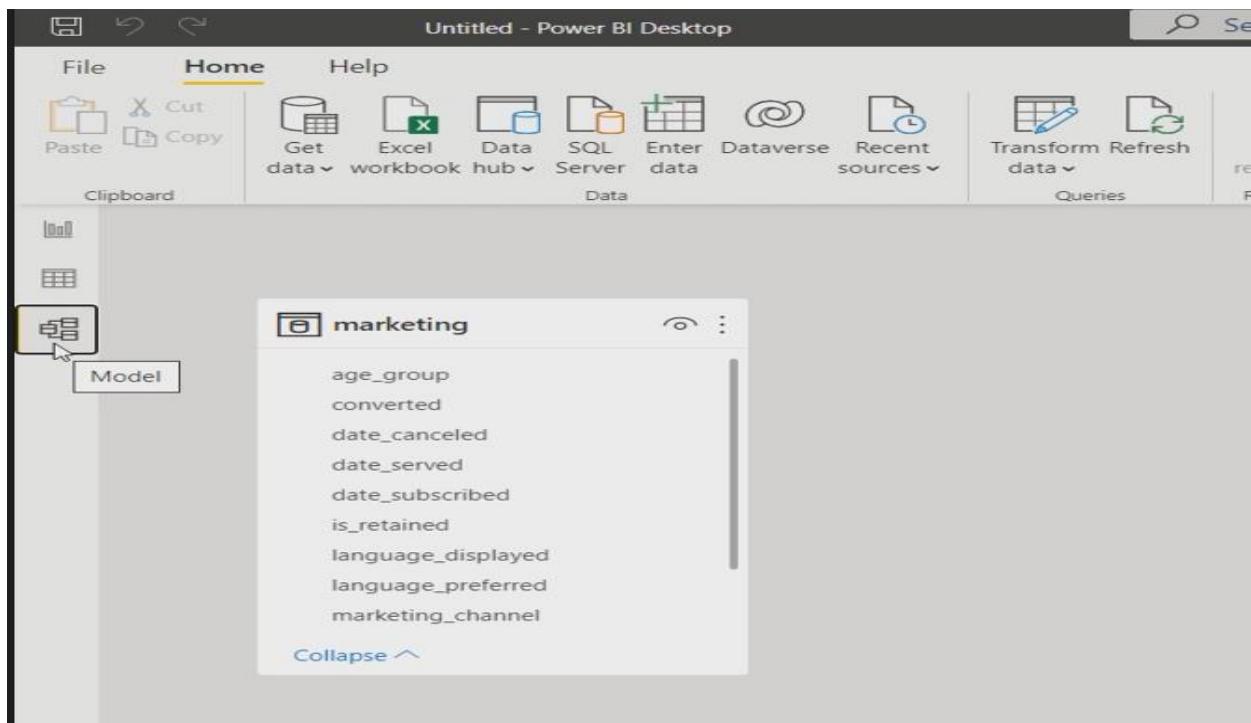


The screenshot shows the Power BI Desktop interface with the title bar "Untitled - Power BI Desktop". The "Table tools" tab is selected in the ribbon. A table named "marketing" is displayed, containing 40 rows of data with columns: user_id, date_served, marketing_channel, variant, converted, language_displayed, and language_preferred. The data shows various user IDs, dates ranging from 1/12/18 to 1/28/18, marketing channels like House Ads, variants like control, and languages like English.

user_id	date_served	marketing_channel	variant	converted	language_displayed	language_preferred
a100000184	1/12/18	House Ads	control	False	English	English
a100000184	1/2/18	House Ads	control	False	English	English
a100000212	1/13/18	House Ads	control	False	English	English
a100000212	1/3/18	House Ads	control	False	English	English
a100000219	1/14/18	House Ads	control	False	English	English
a100000219	1/4/18	House Ads	control	False	English	English
a100001005	1/15/18	House Ads	control	False	English	English
a100001005	1/5/18	House Ads	control	False	English	English
a100001010	1/14/18	House Ads	control	False	English	English
a100001012	1/16/18	House Ads	control	False	English	English
a100001012	1/6/18	House Ads	control	False	English	English
a100001016	1/15/18	House Ads	control	False	English	English
a100001019	1/17/18	House Ads	control	False	English	English
a100001019	1/7/18	House Ads	control	False	English	English
a100001026	1/18/18	House Ads	control	False	English	English
a100001026	1/8/18	House Ads	control	False	English	English
a100001033	1/19/18	House Ads	control	False	English	English
a100001033	1/9/18	House Ads	control	False	English	English
a100001040	1/19/18	House Ads	control	False	English	English
a100001040	1/20/18	House Ads	control	False	English	English
a100001040	1/10/18	House Ads	control	False	English	English
a100000423	1/27/18	House Ads	control	False	English	English
a100000423	1/12/18	House Ads	control	False	English	English
a100000430	1/28/18	House Ads	control	False	English	English
a100000430	1/13/18	House Ads	control	False	Enelish	Enelish

Figure 3.16 – The Data

Power BI show us the fields of the data in a form of model.



The screenshot shows the Power BI Desktop interface with the title bar "Untitled - Power BI Desktop". The "Home" tab is selected in the ribbon. In the center, the "marketing" table is displayed, showing its fields: age_group, converted, date_canceled, date_served, date_subscribed, is_retained, language_displayed, language_preferred, and marketing_channel. The "Model" icon in the left sidebar is highlighted with a yellow box.

Figure 3.17 – Data Model

To create a functioning responsive dashboard in power BI we need to manipulate the data we loaded, yet we can't start manipulating this data unless it's **clean**.

There is a couple of rules that the data has to conform to in order for us to consider it clean or not.

Those rules are:

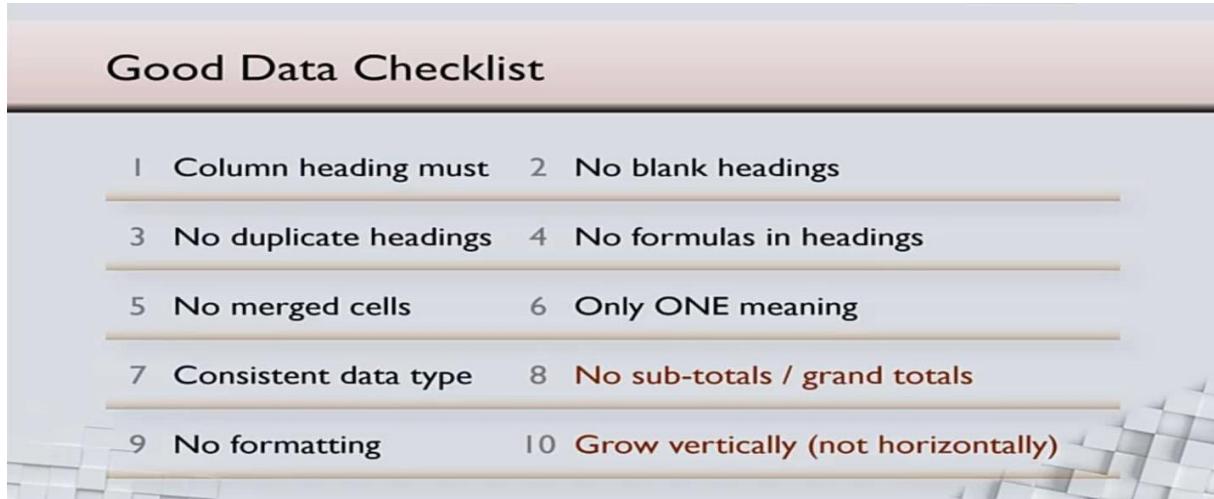


Figure 3.18 – Cleaning Data rules

To manipulate the data and transform it we need to open the **Power Query Editor** in Power BI desktop.

In our scenario, our data sticks to almost all the rules except one rule that our data don't follow which is **the type** of certain fields.

The screenshot shows the Power Query Editor interface with a table named 'marketing'. The table contains the following columns and data:

#	A _{PK} user_id	A _{PK} date_served	A _{PK} marketing_channel	A _{PK} variant	A _{PK} converted	A _{PK} language_displayed	A _{PK} language_pre
1	a100000029	1/1/18	House Ads	personalization	True	English	English
2	a100000030	1/1/18	House Ads	personalization	True	English	English
3	a100000031	1/1/18	House Ads	personalization	True	English	English
4	a100000032	1/1/18	House Ads	personalization	True	English	English
5	a100000033	1/1/18	House Ads	personalization	True	English	English
6	a100000034	1/1/18	House Ads	personalization	True	German	German
7	a100000035	1/1/18	House Ads	personalization	True	English	English
8	a100000036	1/1/18	House Ads	personalization	True	English	English
9	a100000037	1/1/18	House Ads	personalization	True	English	English
10	a100000038	1/1/18	House Ads	personalization	True	English	English
11	a100000039	1/1/18	House Ads	personalization	True	English	English
12	a100000040	1/1/18	House Ads	personalization	True	English	English
13	a100000041	1/1/18	House Ads	personalization	True	Arabic	Arabic
14	a100000042	1/1/18	House Ads	personalization	True	English	English
15	a100000043	1/1/18	House Ads	personalization	True	English	English
16	a100000044	1/1/18	House Ads	personalization	True	Arabic	Arabic
17	a100000045	1/2/18	House Ads	personalization	True	English	English
18	a100000046	1/2/18	House Ads	personalization	True	English	English
19	a100000047	1/2/18	House Ads	personalization	True	English	English
20	a100000048	1/2/18	House Ads	personalization	True	German	German
21	a100000049	1/2/18	House Ads	personalization	True	German	German
22	a100000050	1/2/18	House Ads	personalization	True	English	English
23	a100000050	1/1/18	House Ads	personalization	False	English	English
24	a100000051	1/2/18	House Ads	personalization	True	English	English

The 'Applied Steps' pane on the right shows a step named 'change_type_date'.

Figure 3.19 – Data contains fields with the wrong types

The above figure shows that certain fields which are supposed to be type date, are instead written as a **type text field**.

For better manipulation of the data, we need to transform the type of the fields **"date_served"**, **"date_subscribed"** and **"date_canceled"** to type date instead of text type.

The screenshot shows the Power BI Data Editor interface. The top ribbon has tabs like Fichier, Home, Transform, Add Column, View, Tools, and Help. The Transform tab is selected. Below the ribbon is a toolbar with various icons for data operations like Close & Apply, New Source, Refresh, Properties, Manage, and Sort. The main area shows a query editor with a DAX formula:

```
= Table.TransformColumnTypes(marketing1,{{"user_id", type text}, {"date_served", type date}, {"marketing_channel", type text}, {"variant", type text}, {"converted", type logical}, {"language_displayed", type text}, {"language_preferred", type text}, {"age_group", type text}, {"date_subscribed", type date}, {"date_canceled", type date}, {"subscribing_channel", type text}, {"is_retained", type logical}})
```

Below the formula is a preview of the data in a table format. The columns are: user_id, date_served, marketing_channel, variant, converted, language_displayed, and language. The data shows 16 rows of user information with their respective dates and language preferences.

	user_id	date_served	marketing_channel	variant	converted	language_displayed	language
1	a100000029	01/01/2018	House Ads	personalization	TRUE	English	English
2	a100000030	01/01/2018	House Ads	personalization	TRUE	English	English
3	a100000031	01/01/2018	House Ads	personalization	TRUE	English	English
4	a100000032	01/01/2018	House Ads	personalization	TRUE	English	English
5	a100000033	01/01/2018	House Ads	personalization	TRUE	English	English
6	a100000034	01/01/2018	House Ads	personalization	TRUE	German	German
7	a100000035	01/01/2018	House Ads	personalization	TRUE	English	English
8	a100000036	01/01/2018	House Ads	personalization	TRUE	English	English
9	a100000037	01/01/2018	House Ads	personalization	TRUE	English	English
10	a100000038	01/01/2018	House Ads	personalization	TRUE	English	English
11	a100000039	01/01/2018	House Ads	personalization	TRUE	English	English
12	a100000040	01/01/2018	House Ads	personalization	TRUE	English	English
13	a100000041	01/01/2018	House Ads	personalization	TRUE	Arabic	Arabic
14	a100000042	01/01/2018	House Ads	personalization	TRUE	English	English
15	a100000043	01/01/2018	House Ads	personalization	TRUE	English	English
16	a100000044	01/01/2018	House Ads	personalization	TRUE	Arabic	Arabic

Figure 3.20 – Changing certain fields types

For better data manipulation and because we need those values in certain charts in the dashboard, we need to **create new columns**.

To manipulate the data, we need to run a **new python script** in the Power Query Editor.

The screenshot shows the Power Query Editor interface. The top ribbon has tabs like Fichier, Home, Transform, Add Column, View, Tools, and Help. The Transform tab is selected. Below the ribbon is a toolbar with various icons for data operations like Group By, Use First Row as Headers, Transpose, Reverse Rows, Detect Data Type, Rename, Pivot Column, Fill, Move, Unpivot Columns, Split Column, Format, Extract, Parse, Merge Columns, Statistics, Standard, Scientific, Trigonometry, Date, Time, Duration, and Run Python script. The Run Python script icon is highlighted with a yellow box.

Figure 3.21 – run a new python script

❖ Adding the column named “is_correct_lang”

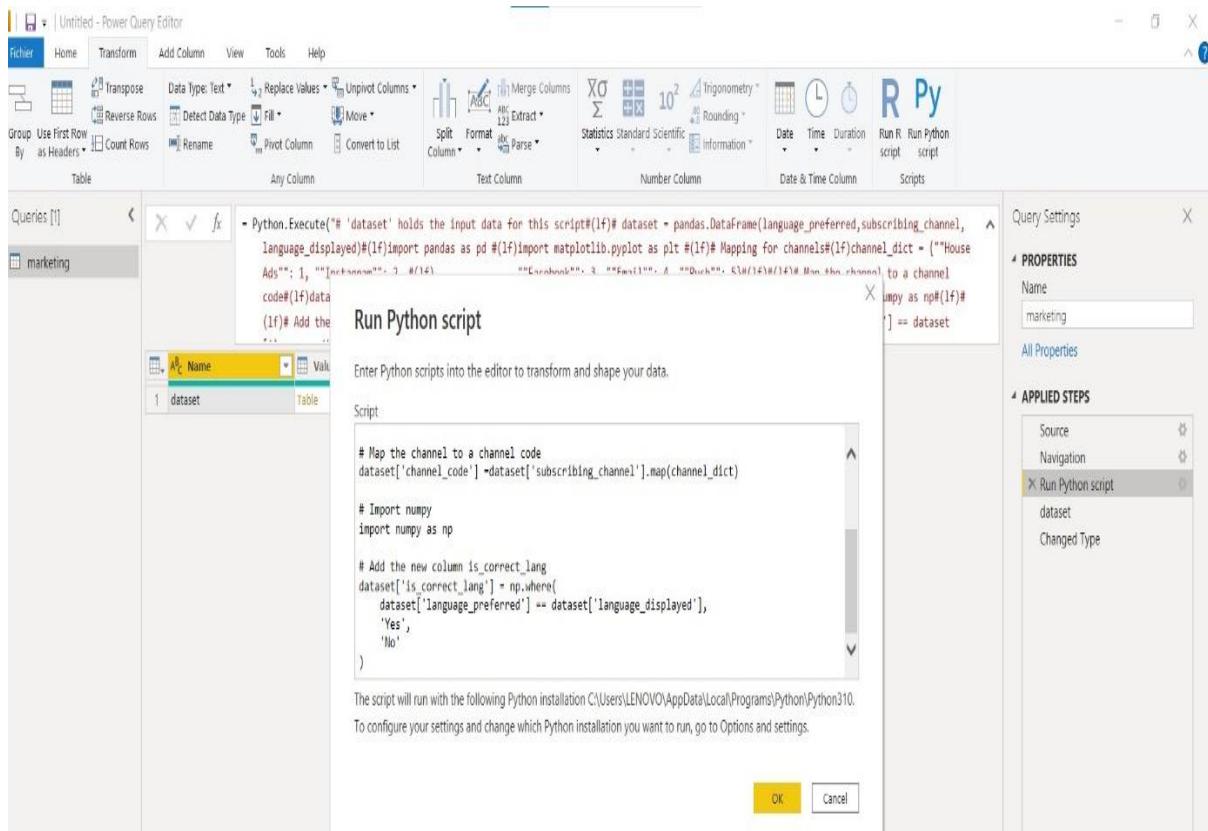


Figure 3.22 – Adding the column “is_correct_lang” in the python script

	date_cancelled	subscribing_channel	is_retained	DoW	channel_code	is_correct_lang
1		House Ads	True	0.0	1.0	Yes
2		House Ads	True	0.0	1.0	Yes
3		House Ads	True	0.0	1.0	Yes
4		House Ads	True	0.0	1.0	Yes
5		House Ads	True	0.0	1.0	Yes
6		House Ads	True	0.0	1.0	Yes
7		House Ads	True	0.0	1.0	Yes
8		House Ads	True	0.0	1.0	Yes
9		House Ads	True	0.0	1.0	Yes
10		House Ads	True	0.0	1.0	Yes
11		House Ads	True	0.0	1.0	Yes
12	2018-01-18	House Ads	False	0.0	1.0	Yes
13	2018-02-22	House Ads	False	0.0	1.0	Yes
14	2018-03-09	House Ads	False	0.0	1.0	Yes

Figure 3.23 – “is_correct_lang” column is added in the data

❖ Adding the column named “channel_code”

The screenshot shows the Power Query Editor interface with the 'Run Python script' dialog open. The script defines a mapping dictionary for channels and then maps the 'subscribing_channel' column to the 'channel_code' column.

```

# Mapping for channels
channel_dict = {"House Ads": 1, "Instagram": 2, "Facebook": 3, "Email": 4, "Push": 5}

# Map the channel to a channel code
marketing['channel_code'] = marketing['subscribing_channel'].map(channel_dict)
    
```

Figure 3.24 – Adding the column “channel_code” in the python script

The screenshot shows the Power Query Editor interface with the 'Applied Steps' pane open. The 'Removed Columns' step is highlighted, indicating that the 'channel_code' column was added by removing the 'subscribing_channel' column.

Figure 3.25 – “channel_code” column is added in the data

❖ Adding the column named “Dow”

The screenshot shows the Power Query Editor interface with the 'Run Python script' dialog open. The script uses pandas to map channels to codes and then adds a 'Dow' column based on the day of the week of the 'date_subscribed' column.

```

# Add the new column is_correct_lang
dataset['is_correct_lang'] = np.where(
    dataset['language_preferred'] == dataset['language_displayed'],
    'Yes',
    'No'
)

dataset['Dow'] = dataset['date_subscribed'].dt.dayofweek
    
```

Figure 3.26 – Adding the column “Dow” in the python script

The screenshot shows the Power Query Editor interface. On the left, there's a 'Queries [1]' pane with a single entry named 'marketing'. The main area displays a table with six columns: 'date_canceled', 'subscribing_channel', 'is_retained', 'DoW', 'channel_code', and 'is_correct_lang'. The 'DoW' column is highlighted with a yellow background. On the right, there's a 'Query Settings' pane with sections for 'PROPERTIES' (Name: marketing) and 'APPLIED STEPS' (Source, Navigation, Removed Columns, Run Python script, dataset).

Figure 3.27 – “Dow” column is added in the data

- ❖ Adding the column named “conversion_rate”

The screenshot shows the Power Query Editor interface. In the center, a 'Run Python script' dialog box is open. It contains a code editor with the following Python script:

```

Script
# 'dataset' holds the input data for this script
# dataset = pandas.DataFrame(user_id, converted)

import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

# calculate the number of people we marketed to
total = dataset['user_id'].nunique()
# Calculate the number of people who subscribed
subscribers = dataset[dataset['converted'] == True]['user_id'].nunique()

# calculate the conversion rate
conversion_rate = subscribers / total
    
```

The dialog also includes a note about the Python installation path and configuration options. At the bottom, there are 'OK' and 'Cancel' buttons.

Figure 3.28 – Adding the column “conversion_rate” in the python script

Conversion rate resembles the percentage of customers who subscribed to our marketing channel, from the total of people we marketed to on those channels in the marketing campaign.

The screenshot shows the Power Query Editor interface. The 'marketing' query now includes an additional column, 'conversion_rate%', which is the result of the Python script. The table has seven columns: 'date_canceled', 'subscribing_channel', 'is_retained', 'DoW', 'channel_code', 'is_correct_lang', and 'conversion_rate%'. The 'conversion_rate%' column contains values like 14.09. The 'APPLIED STEPS' section in the Query Settings pane now includes 'dataset'.

Figure 3.29 – “conversion_rate” column is added in the data

- ❖ Adding the column named “**retention_rate**”

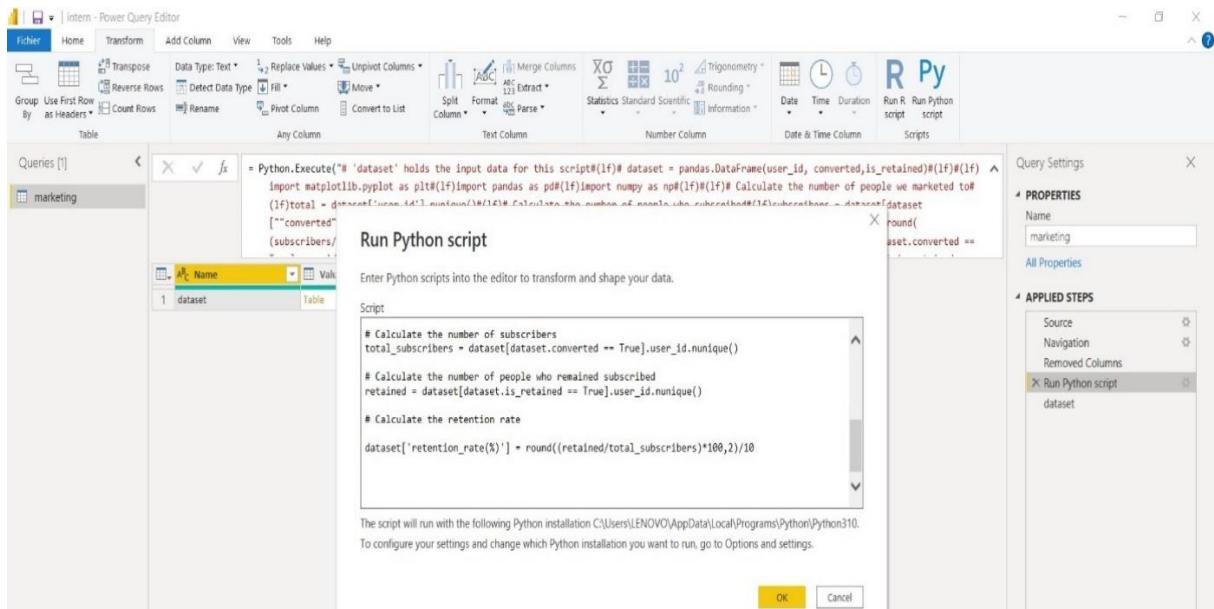


Figure 3.30 – Adding the column “conversion_rate” in the python script

Retention rate resembles the percentage of customers who retained subscribed and didn't cancel their subscription.

channel	is_retained	DoW	channel_code	is_correct_lang	conversion_rate[%]	retention_rate[%]
1	True	0.0	1.0	Yes	14.09	67.64099999999999
2	True	0.0	1.0	Yes	14.09	67.64099999999999
3	True	0.0	1.0	Yes	14.09	67.64099999999999
4	True	0.0	1.0	Yes	14.09	67.64099999999999
5	True	0.0	1.0	Yes	14.09	67.64099999999999
6	True	0.0	1.0	Yes	14.09	67.64099999999999
7	True	0.0	1.0	Yes	14.09	67.64099999999999

Figure 3.31 – “retention_rate” column is added in the data

IV. Creating the Dashboard

We create different charts on every page of the dashboard by writing a python script for each graph.

Let's take the example of Marketing channels across age groups graph, this graph shows us, for each social media platform the number of users for each age group.

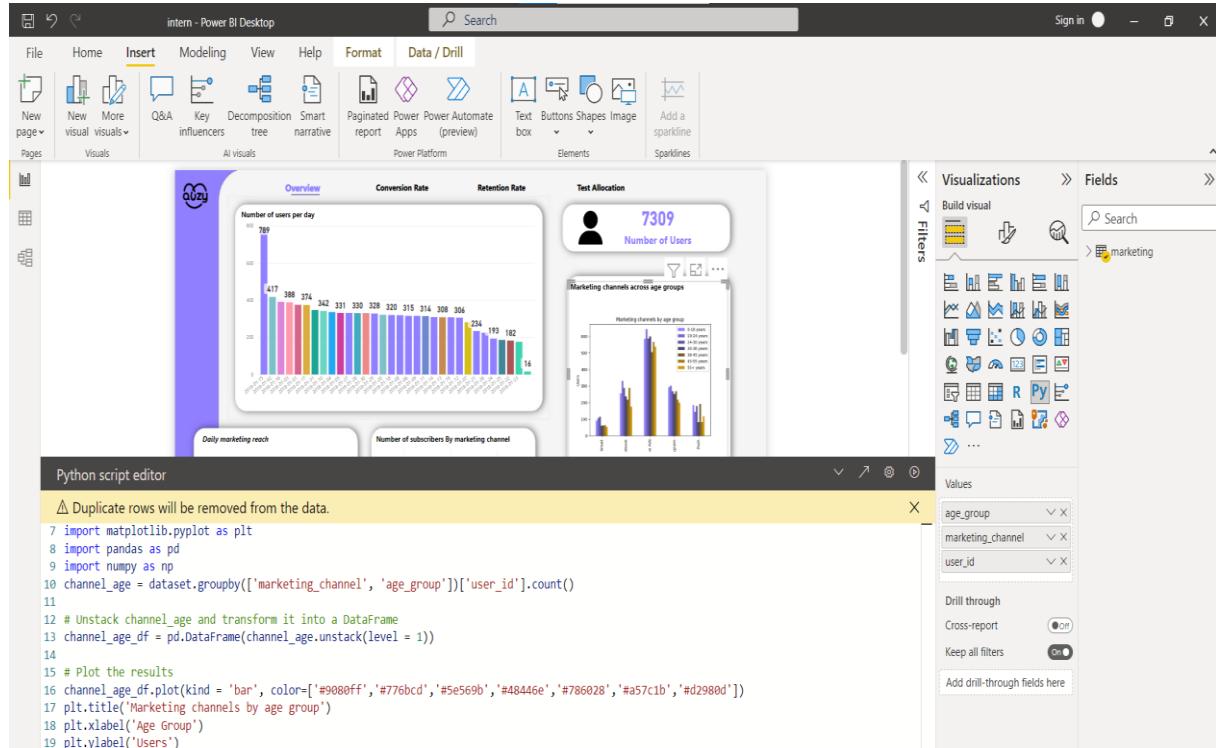


Figure 3.32 – python script for the graph “Marketing channels across age groups” from the page overview

The second example is a graph from the second-page Conversion rate, this graph, show us the number of subscribers from each age group.

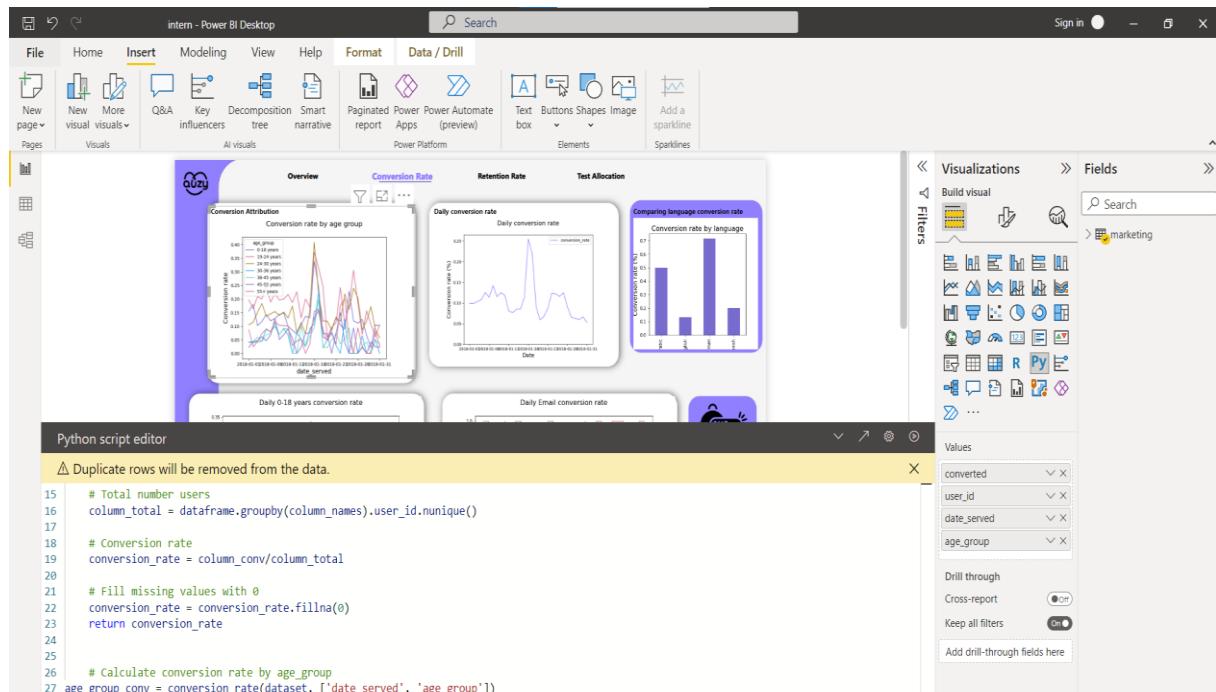


Figure 3.33 – python script for the graph “Conversion rate by age group” from the page Conversion Rate

The third example is the Retention Rate by Subscribing Channel graph, which show us the percentage of retained subscribers for each social media platform per subscription date.

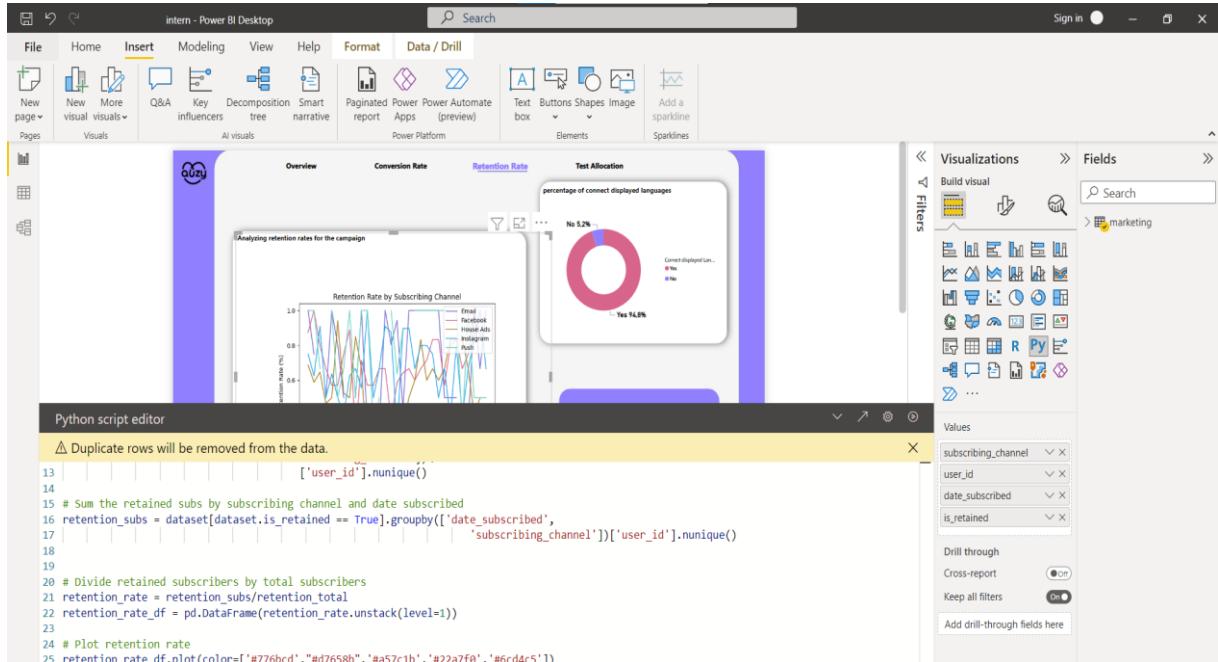


Figure 3.34 – python script for the graph “Retention Rate by Subscribing Channel” from the page Retention Rate

The fourth example is Test Allocation for Email graph, this graph shows us the result of the test allocation for the email.

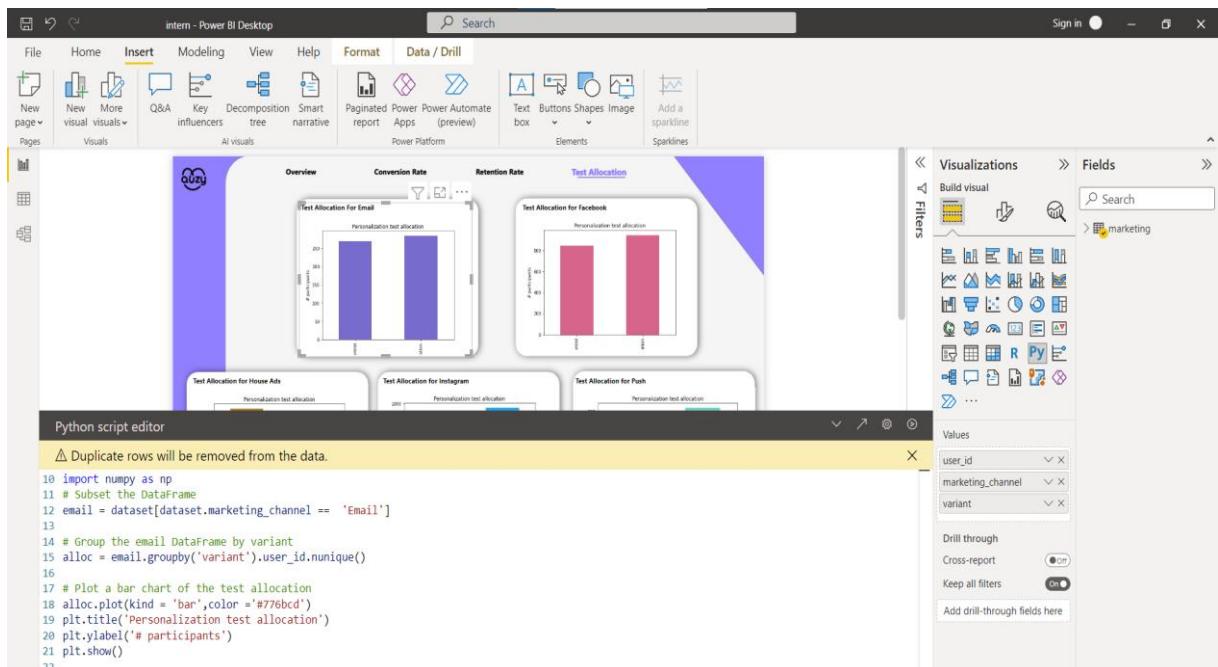


Figure 3.35 – python script for the graph “Test Allocation for Email” from the page Test Allocation

The dashboard consists of **4 different pages**, the first page is the **overview** page.

On this page, we find the total number of users, number of users per day, marketing channels, daily marketing reach, Marketing channels across age groups, and the Number of subscribers by marketing channel.

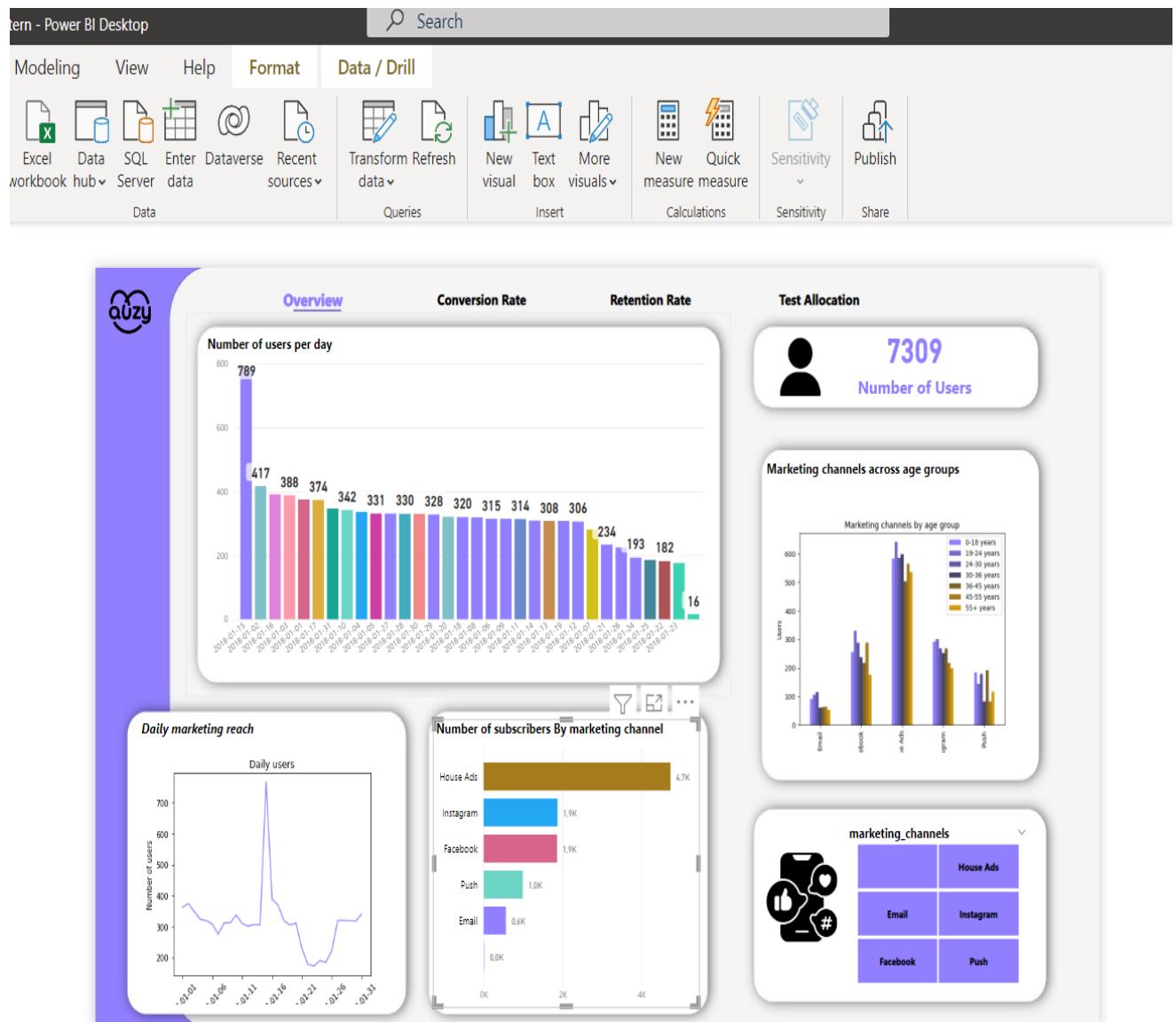


Figure 3.36 – Overview page

The second page is the Conversion Rate page, this page contains the Daily conversion rate, Conversion rate by age group, Comparing language conversion rate, percentage of conversion rate, Daily Email conversion rate, and Daily 0-18 years conversion rate.

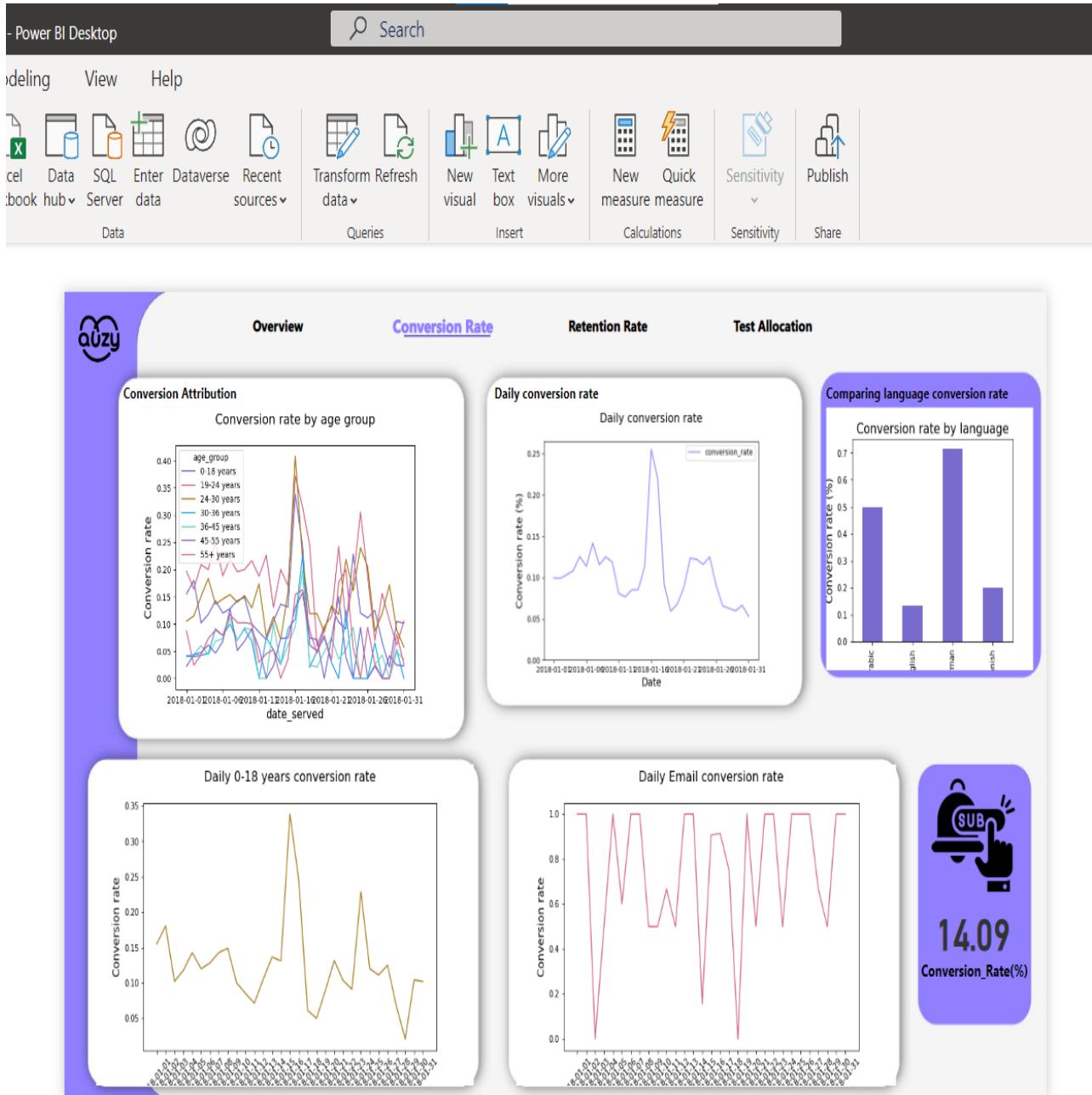


Figure 3.37 – Conversion Rate page

The third page is the Retention Rate page, this page contains the percentage of connected displayed languages, the Retention Rate by Subscribing Channel, and the retention rate percentage.

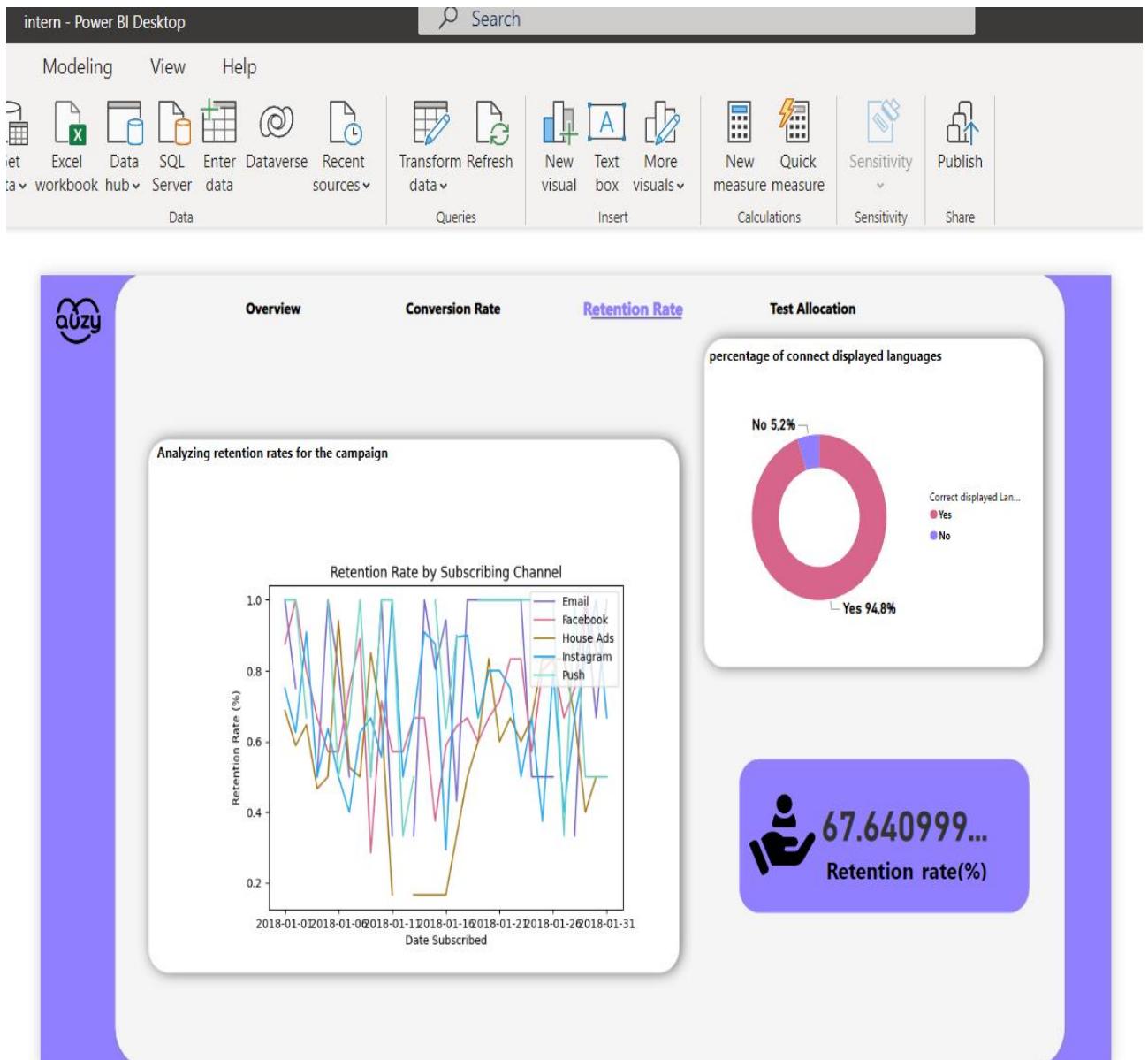


Figure 3.38 – Retention Rate page

The fourth page is the Test Allocation page, this page contains test allocation for different social media platforms (Email, Facebook, House ADS, Instagram, and Push).

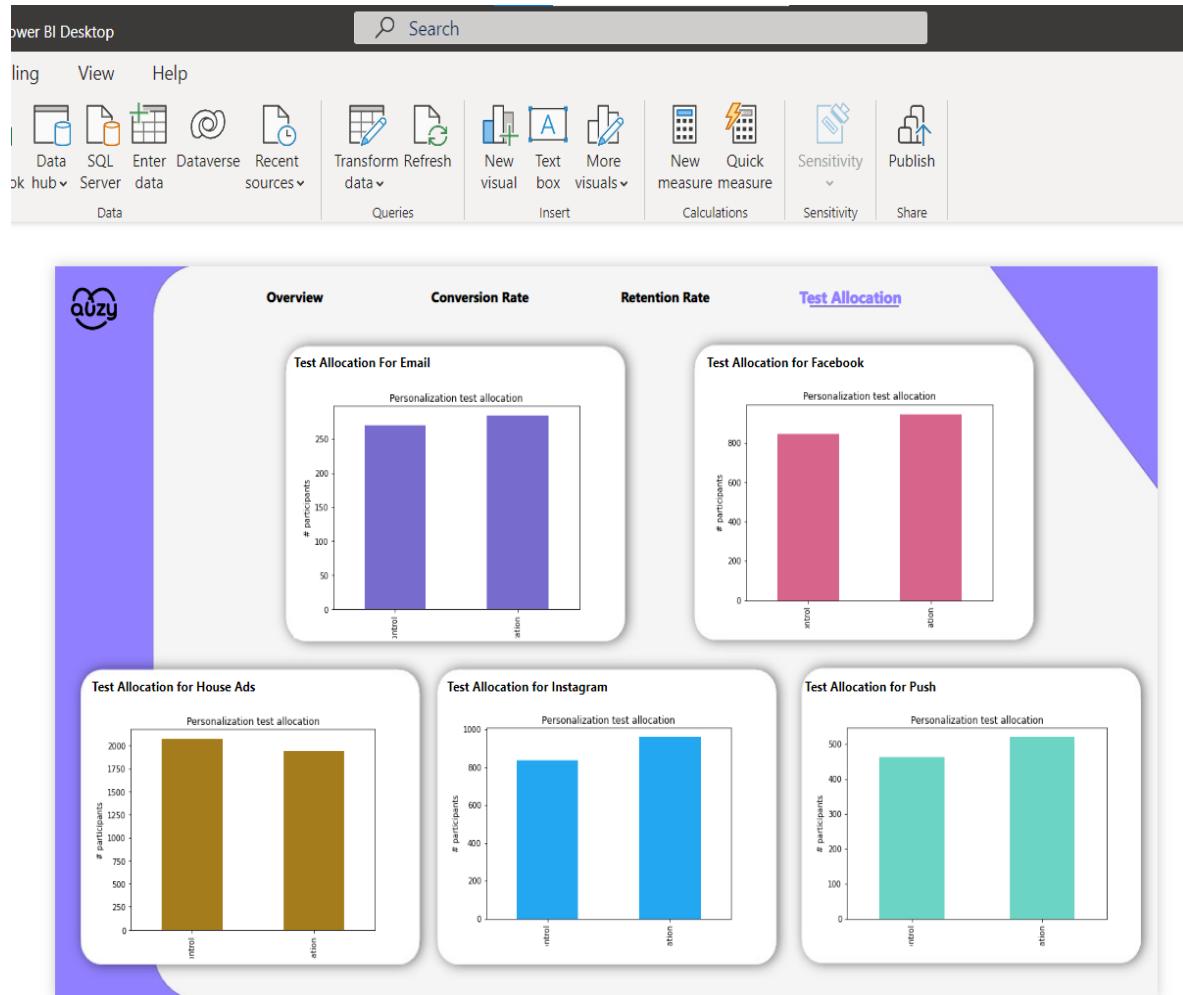


Figure 3.39 – Test Allocation page

V. Conclusion

In this chapter, we first presented our working environment, then we delivered the different pages of the BI dashboard.

General Conclusion

In conclusion, I can say that this Internship was an amazing chance to discover a tremendously new field, which is business intelligence, this field has expanded exponentially in recent years and it's expected to continue growing.

Moreover, I have reinforced my knowledge of different technologies, software, and concepts such as Power BI, Reporting, Descriptive analytics, and data visualization.

finally, this internship introduced me to professional life and teamwork in the host company's headquarters.

Bibliography

- [1] : <https://docs.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>
- [2] : <https://docs.microsoft.com/en-us/dax/dax-overview>
- [3] : <https://www.python.org/doc/essays/blurb/>
- [4] : <https://numpy.org/doc/stable/user/whatisnumpy.html>
- [5] : [https://en.wikipedia.org/wiki/Pip_\(package_manager\)](https://en.wikipedia.org/wiki/Pip_(package_manager))
- [6] : https://pandas.pydata.org/docs/getting_started/overview.html
- [7] : <https://www.activestate.com/resources/quick-reads/what-is-matplotlib-in-python-how-to-use-it-for-plotting/#:~:text=Matplotlib%20is%20a%20cross%2Dplatform,embed%20plots%20in%20GUI%20applications.>