Tech Saksham

Case Study Report

Data Analytics with Power BI

“Real-Time Analysis of Bank Customers”

“BHARATHIYAR ARTS AND SCIENCE COLLEGE FOR WOMEN’S”

NM ID	NAME
0D9BA20C719C3ECA170493FA45DD83C4	PARAMESWARI R

Trainer Name :R UMAMAHESWARI

Master Trainer: R UMAMAHESWARI

ABSTRACT

In the digital age, data has become an invaluable asset for businesses, particularly in the banking sector. The proposed project, “Real-Time Analysis of Bank Customers,” aims to leverage Power BI, a leading business intelligence tool, to analyze and visualize real-time customer data. This project will enable banks to gain deep insights into customer behavior, preferences, and trends, thereby facilitating data-driven decision-making and enhancing customer satisfaction. The real-time analysis will allow banks to respond promptly to changes in customer behavior or preferences, identify opportunities for cross-selling and up-selling, and tailor their products and services to meet customer needs. The project will also contribute to the broader goal of digital transformation in the banking sector, promoting efficiency, innovation, and customer-centricity.

INDEX

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

In today's competitive banking landscape, understanding customer behavior and preferences is crucial for customer retention and revenue generation. However, banks often face challenges in analyzing customer data due to the sheer volume and velocity of data generated. Traditional data analysis methods are time-consuming and often fail to provide real-time insights. This lack of real-time analysis can lead to missed opportunities for customer engagement, cross-selling, and up-selling, impacting the bank's revenue generation and customer satisfaction. Furthermore, the complexity and diversity of customer data, which includes transaction history, customer feedback, and demographic data, pose additional challenges for data analysis.

1.2 Proposed Solution

The proposed solution is to develop a PowerBI dashboard that can analyze and visualize real-time customer data. The dashboard will integrate data from various sources such as transaction history, customer feedback, and demographic data. It will provide a comprehensive view of customer behavior, preferences, and trends, enabling banks to make informed decisions. The dashboard will be interactive, user-friendly, and customizable, allowing banks to tailor it to their specific needs. The real-time analysis capability of the dashboard will enable banks to respond promptly to changes in customer behavior or preferences, identify opportunities for cross-selling and up-selling, and tailor their products and services to meet customer needs.

1.3 Feature

- **Real-Time Analysis**: The dashboard will provide real-time analysis of customer data.

- **Customer Segmentation**: It will segment customers based on various parameters like age, income, transaction behavior, etc.
- **Trend Analysis**: The dashboard will identify and display trends in customer behavior.
- **Predictive Analysis**: It will use historical data to predict future customer behavior.

1.4 Advantages

- **Data-Driven Decisions**: Banks can make informed decisions based on real-time data analysis.
- **Improved Customer Engagement**: Understanding customer behavior and trends can help banks engage with their customers more effectively.
- **Increased Revenue**: By identifying opportunities for cross-selling and up-selling, banks can increase their revenue.

1.5 Scope

The scope of this project extends to all banking institutions that aim to leverage data for decision-making and customer engagement. The project can be further extended to incorporate more data sources and advanced analytics techniques, such as machine learning and artificial intelligence, to provide more sophisticated insights into customer behavior. The project also has the potential to be adapted for other sectors, such as retail, healthcare, and telecommunications, where understanding customer behavior is crucial. Furthermore, the project contributes to the broader goal of digital transformation in the banking sector, promoting efficiency, innovation, and customer-centricity.

CHAPTER 2

SERVICES AND TOOLS REQUIRED

2.1 Services Used

Core Banking Systems

- Core banking systems are the backbone of any financial institution, providing the infrastructure and functionality necessary to manage all aspects of banking operations.

Customer Relationship Management (CRM)

- CRM software is used by banks and financial institutions to manage interactions with customers and improve the overall customer experience.

Anti-Money Laundering (AML)

- AML software helps banks and financial institutions identify and prevent money laundering and other financial crimes. These systems use algorithms and machine learning to detect suspicious activity and flag it for further investigation. Some of the widely used AML software solutions include NICE Actimize, SAS AML, and FICO AML.

2.2 Tools and Software used

Tools:

- **Power BI**: Power BI is a technology-driven business intelligence tool provided by Microsoft for analyzing and visualizing raw data to present actionable information. It combines business analytics, data visualization, and best practices that help an organization to make data-driven decisions.
- **Power Query**: Excel Power Query is a data transformation and preparation tool developed by Microsoft. It allows users to extract, transform, and load data from various sources into Excel or Power BI using a visual interface. It is a powerful tool for data preparation and analysis tasks.

Software Requirements:

- **Power BI Desktop**: Microsoft Power BI Desktop is a data analysis and reporting application that a user can install on a computer to create dashboards and live reports.
- Integrated with Power BI Service, the user can also share these reports with decision-makers and stakeholders to help them understand the current state of the business.

Tableau:

- Tableau is most known for its wide range of data visualization capabilities, and is often used interchangeably with other traditional BI tools. Analysts use it to examine data with SQL and build data solutions for business decision-makers, who in turn use it to analyze data without having to code.

CHAPTER 3

PROJECT ARCHITECTURE

3.1 Architecture

USER **FRONTEND** **BACKEND**

Real-Time Analytics

- Data collection: the process of gathering data of interest from certain data sources that are required to answering research questions.
- Data cleansing – is the process of detecting and correcting (or removing) corrupt or inaccurate records from a record set, table, or database [52];
- Data transformation – denotes preparing data for further analysis via standardization or normalization. It is needed since in many applications of data analysis, the raw data or cleansed data cannot be used directly due to formatting issues. The formatting of values into consistent layouts is performed based on various elements, including industry standards, local standards, business rules and domain knowledge bases.
- Statistical analysis: The goal of this phase is to find interesting patterns so as to identify trends. This phase involves various algorithms and can be broken down into five discrete steps, namely "data nature description", "data relation exploration", "data modelling on how the data relates to the underlying population", "validation of the model", and "predictive analytics". Statistical analysis is often accompanied with or followed by data visualization to interpret the results of the analysis. Note that these phases are akin to the phases in Knowledge Discovery in Databases (KDD), which focuses on data analysis in databases.

There are different type of data processing tools, reflecting different stages of the data analysis activities as introduced above. Examples of such tools are:

- Pre-processing tools, such as Open Refine [53, 54] and Data Wrangler [55]

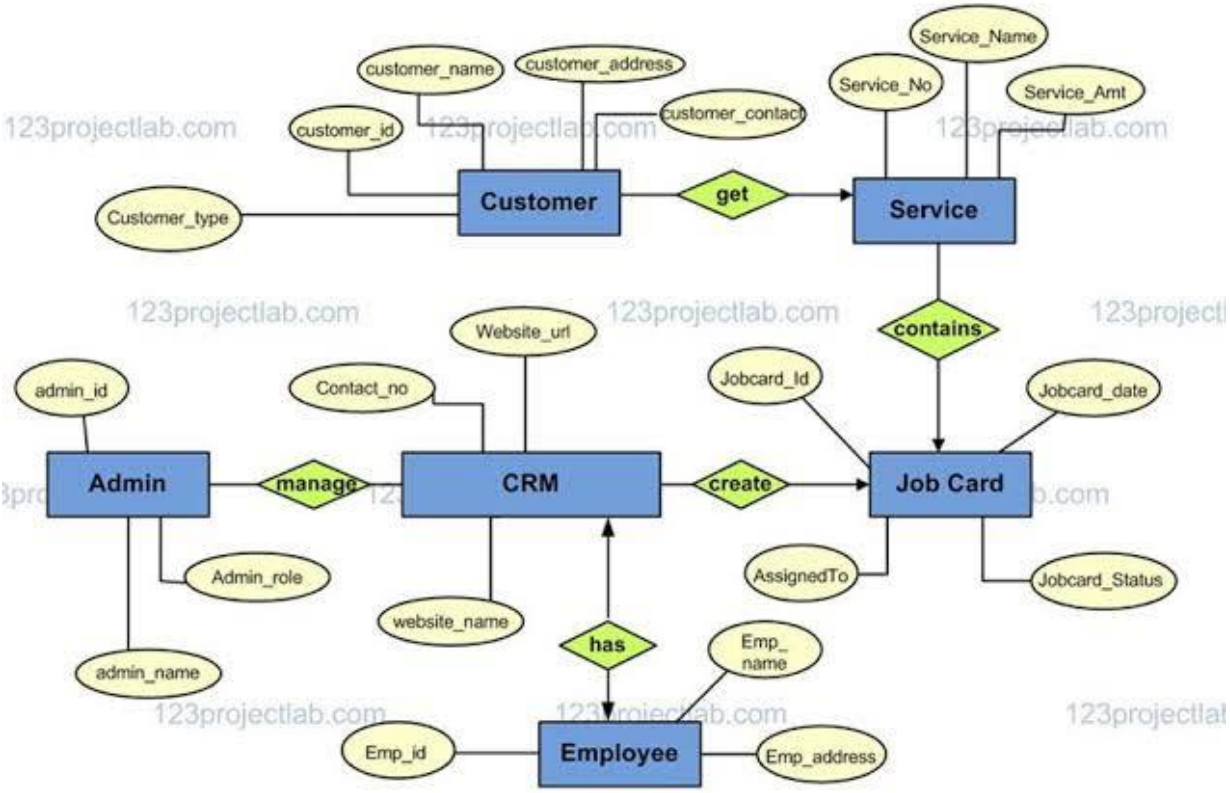
- Database management systems, making use of SQL (e.g. MySQL [56] and PostgreSQL [57]), as well as NoSQL queries to search for data satisfying specific parameters; the NoSQL are more amenable to real-time applications, with some examples of being Cassandra [58] and MongoDB [59]
- Statistical Data analysis tools, such as R, SAS, Stata and SPSS,

CHAPTER 4

MODELING AND RESULT

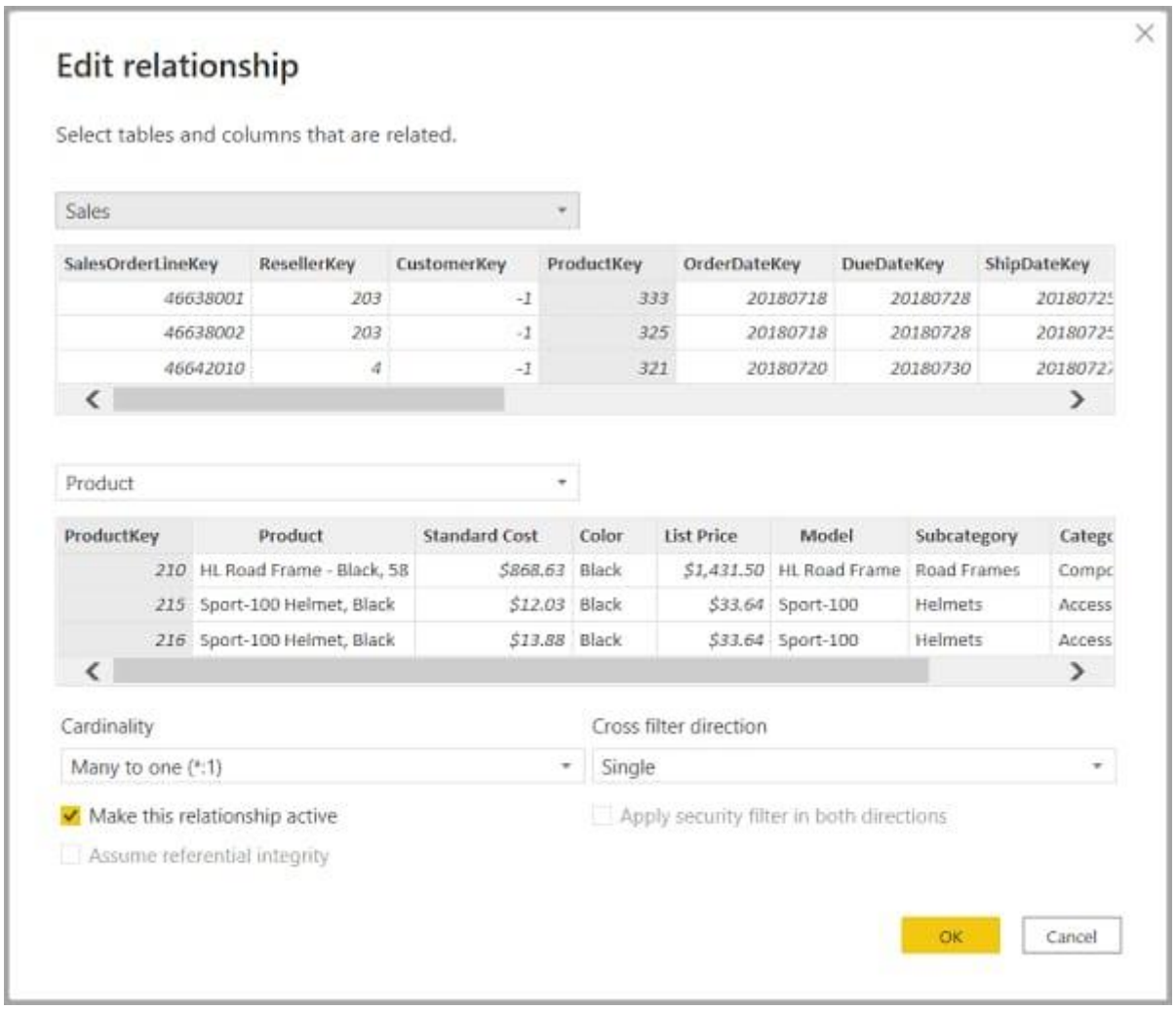
Customer Relationship Management

Customer Relationship Management concept is tendency of banking sector to establish and maintain long-term relationships with customers in order to provide value for customers and banks. This concept allows bank to identify, segment, communicate and build long-term relationships with customers on individual basis.



Manage relationships

Active	From: Table (Column)	To: Table (Column)
☑	card (disp_id)	disp (disp_id)
☑	client (district_id)	district (district_id)
☑	disp (account_id)	account (account_id)
☑	disp (account_id)	loan (account_id)
☑	disp (client_id)	client (client_id)
☑	order (account_id)	account (account_id)
☑	transaction (account_id)	disp (account_id)
☐	account (district_id)	district (district_id)
☐	transaction (account_id)	loan (account_id)



Modelling for Gender and Age data

Notice that the Gender and age of the client are missing from the data. These can be formulated from the birth number YYMMDD where at months (the 3rd and 4th digits) greater than 50 means that client is a Female. We can create a column for Gender.


```
Gender =
VAR stringDate = FORMAT(client[birth_number],"General Number")
VAR month = VALUE(MID(stringDate,3,2))
RETURN IF(month > 50,"F","M")
```

client_id	birth_number	district_id	Gender	Birthday	age
3428	875927	42	F	27/09/1987	13
4354	860813	28	M	13/08/1986	14
3417	855318	35	F	18/03/1985	15
10201	851019	13	M	19/10/1985	15

For birthday, we need to reduce the birth month of the female by 50 and then change the date format to DD/MM/YYYY adding 1900 to the year.

```
Birthday =
VAR stringDate = FORMAT(client[birth_number],"General Number")
VAR stringMonth = VALUE(MID(stringDate,3,2))
VAR mth = IF(stringMonth > 50, stringMonth - 50,stringMonth)
VAR year = VALUE(MID(stringDate,1,2))
VAR day = VALUE(MID(stringDate,5,2))
RETURN FORMAT(DATE(year+1900,mth,day),"DD/MM/YYYY")
```

client_id	birth_number	district_id	Gender	Birthday	age
3428	875927	42	F	27/09/1987	13
4354	860813	28	M	13/08/1986	14
3417	855318	35	F	18/03/1985	15
10201	851019	13	M	19/10/1985	15

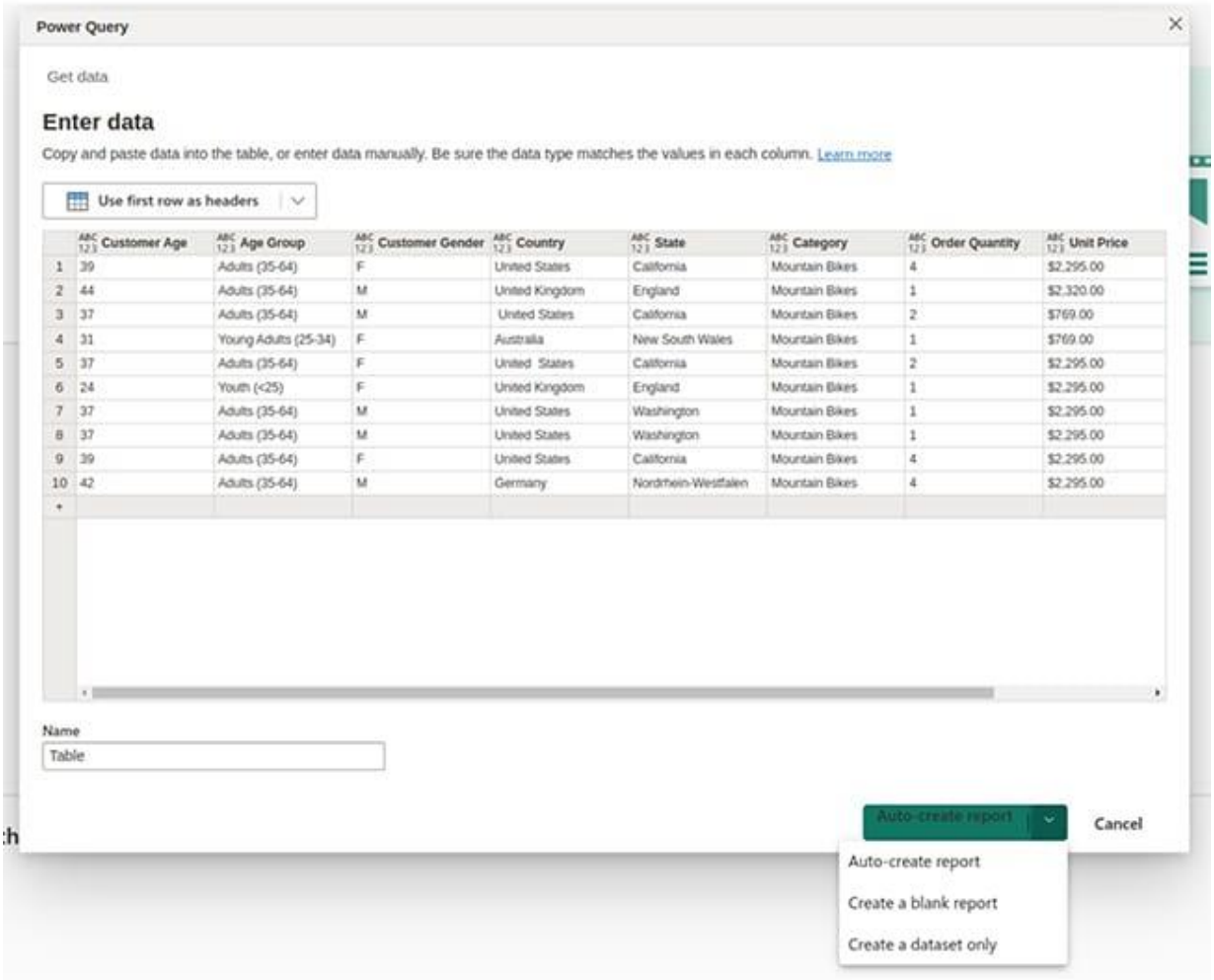
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
age = 1999 -RIGHT(client[Birthday],4)
```

client_id	birth_number	district_id	Gender	Birthday	age	age (groups)
2	450204	1	M	04/02/1945	54	36 -54 Baby Boomers

Entering Data

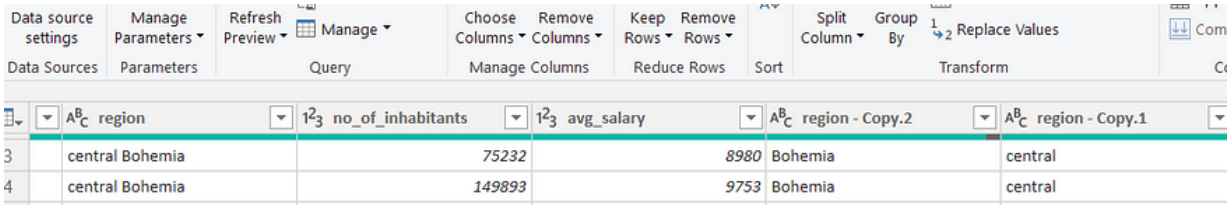
Set some fields to English for easy understanding, we replace values to English with the Power Query Editor.



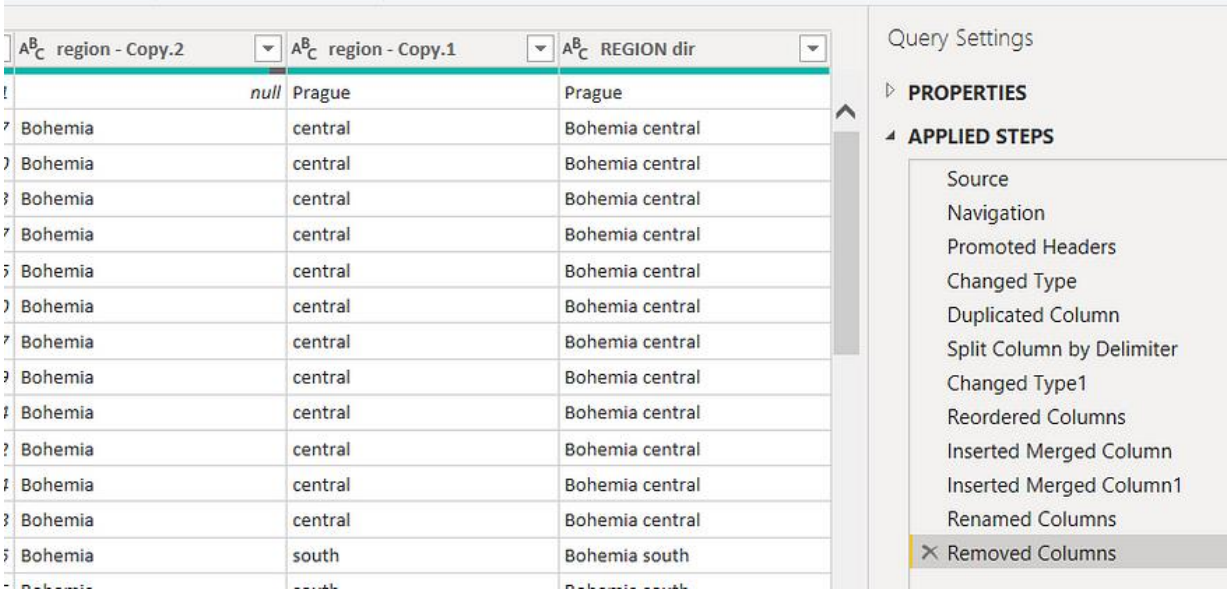
type	+/- transaction	"PRIJEM" stands for credit "VYDAJ" stands for withdrawal
k_symbol	characterization of the transaction	"POJISTNE" stands for insurance payment "SLUZBY" stands for payment for statement "UROK" stands for interest credited "SANKC. UROK" sanction interest if negative balance "SIPO" stands for household "DUCHOD" stands for old-age pension "UVER" stands for loan payment

Changing the order of Region name at Power Query

Duplicate the “district /region” then split column using space as delimiter.



Then merge column by Region and direction. Refer to applied steps for details.



Grouping of age by ranges

- As the customers' age ranges from 12 to 88, we shall group them into different generation age range for easier profiling, we will group the ages into 5 groups.

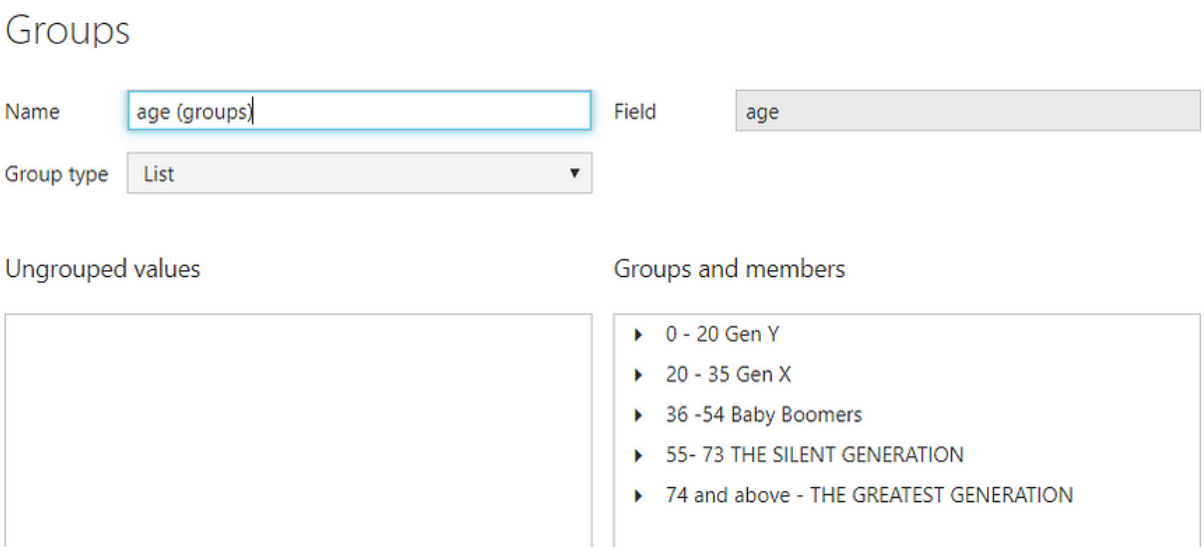
The Gen Y are youths,

- Gen X are young working adults, some starting their families

Baby Boomer are working adults with families.

The silent Generations some are working and retired, living on pensions.

The greatest Generation, retired elderly living on pensions.



- If the sum of ages is X and Y and the ratio of their ages is p:q respectively, then you can determine the age of Y by using the formula shown below: Age of Y = Ratio of Y/ Sum of ratios x sum of ages.

```
=IF(DATEDIF(B2, TODAY(),"y")=0,"",DATEDIF(B2, TODAY(),"y")&" years, ")&
IF(DATEDIF(B2, TODAY(),"ym")=0,"",DATEDIF(B2, TODAY(),"ym")&" months, ")&
IF(DATEDIF(B2, TODAY(),"md")=0,"",DATEDIF(B2, TODAY(),"md")&" days")
```

	A	B	C	D	E
1	**Name**	**Date of birth**	**Age**		
2	Mike	6-Jun-1996	20 years, 3 months, 20 days		
3	Natalie	20-Mar-1983	33 years, 6 months, 6 days		
4	Neal	25-Sep-2000	16 Years, 1 Days		
5	Peter	21-Aug-1984	32 years, 1 months, 5 days		
6	Kate	30-Mar-1990	26 years, 5 months, 27 days		

Values of such as age have also been set as Text.

And name have been categorized as place to be use for the map to show the sum of the inhabitants in each age.

Dashboard

CONCLUSION

The project “Real-Time Analysis of Bank Customers” using Power BI has successfully demonstrated the potential of data analytics in the banking sector. The real-time analysis of customer data has provided valuable insights into customer behavior, preferences, and trends, thereby facilitating informed decision-making. The interactive dashboards and reports have offered a comprehensive view of customer data, enabling the identification of patterns and correlations. This has not only improved the efficiency of data analysis but also enhanced the bank’s ability to provide personalized services to its customers. The project has also highlighted the importance of data visualization in making complex data more understandable and accessible. The use of Power BI has made it possible to present data in a visually appealing and easy-to-understand format, thereby aiding in better decision-making.

FUTURE SCOPE

The future scope of this project is vast. With the advent of advanced analytics and machine learning, Power BI can be leveraged to predict future trends based on historical data. Integrating these predictive analytics into the project could enable the bank to anticipate customer needs and proactively offer solutions. Furthermore, Power BI's capability to integrate with various data sources opens up the possibility of incorporating more diverse datasets for a more holistic view of customers. As data privacy and security become increasingly important, future iterations of this project should focus on implementing robust data governance strategies. This would ensure the secure handling of sensitive customer data while complying with data protection regulations. Additionally, the project could explore the integration of real-time data streams to provide even more timely and relevant insights. This could potentially transform the way banks interact with their customers, leading to improved customer satisfaction and loyalty.

REFERENCES

https://medium.com/analytics-vidhya/analysis-of-bank-customers-using-dashboard-in-power-bi-a366f2b3e563

LINK

https://github.com/githubtraining/hellogitworld.git

