Tech Saksham

Case Study Report

Data Analytics with Power BI

“Real-Time Analysis of Bank Customers”

“A V S College of Arts & Science”

NM ID	NAME
D3F5A98B35DD3D1AD654A50EE7AB	BAJURUDIN J

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 PowerBI, 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

Sr. No.	Table of Contents	Page No.
1	Chapter 1: Introduction	4
2	Chapter 2: Services and Tools Required	7
3	Chapter 3: Project Architecture	9
4	Chapter 4: Modeling and Result	10
5	Conclusion	17
6	Future Scope	18
7	References	19
8	Links	19

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

A major bank wants to improve its customer service and marketing strategies by analyzing the behavior and characteristics of its customers in real-time. Then bank has vast amounts of customer data including transaction history, demographics, account types, and more. The goal is to leverage data analytics using Power BI to gain insights into customer behavior, preferences, and trends. In real time monitoring, we develop a dashboard that provides real-time updates on customer activities such as transactions, account openings, account closures, etc. In customer segmentation, we utilize clustering algorithms to segment customers based on demographics, transaction patterns, account types, and other relevant factors.

1.2 Proposed Solution

For real-time dashboard, we develop a dynamic Power BI dashboard that connects to the bank's data sources and provides real-time updates in key metrics such as transaction volumes, account activities, customer demographics and more. We can utilize Power BI's streaming capabilities to ensure timely updates. In customer segmentation, we implement clustering algorithms within Power BI to segment customers based on various attributes such as age, income, account balances, transaction frequency, and product usage. Visualize these segments to understand the different customer personas and their behaviors.

1.3 Feature

The "Real time analysis of bank customers" feature in Power BI could involve monitoring anfd analyzing various metrics related to bank customers in real-time. This includes

1. **Customer demographics:** Tracking demographics such as age, gender, location, and income level of bank customers.
2. **Account activities:** Monitoring real-time account activities such as deposits, withdrawals, transfers, and account balances.
3. **Customer behavior:** Analyzing customer behavior patterns such as frequency of transactions, preferred banking channels(online, mobile, in-person) and product usage.

1.4 Advantages

The advantage of implementing real-time analysis of bank customers using Power BI include:

i) **Immediate insights:** Real-time analysis allows banks to monitor customer behavior as it happens, enabling timely responses to changing trends, potential issues, or opportunities.

ii) **Enhanced decision-making:** With up-to-the-minute data at their disposal, decision-makes can make informed choices quickly,leading to more effective strategies for customer acquisition,retention, and satisfaction

iii) **Proactive fraud detection:** Real-time analytics can help banks detect and respond to fraudulent activities as they occur ,minimizing financial losses and protecting customers assets

iv) **Improved regulatory compliance**: Real-time monitoring can help banks ensure compliance with regulations by promptly identifying and addressing any discrepancies or violations.

1.5 Scope

The scope of real-time analysis of bank customers using Power BI encompasses various aspects of customer data and interactions within the banking ecosystem. Here are some key components of its scope:

1. **Customer demographics:** Analyzing demographic information such as age, gender, income level, and location to understand the characteristics of different customer segments.

2. **Transaction monitoring:** Tracking real-time transactional data including deposits, withdrawals, transfers, payments, and purchases to identify patterns, trends, and anomalies.
3. **Channel analytics:** Examining customer interactions across different channels such as online banking, mobile apps, ATMs, and branches to optimize channel performance and customer experience.

CHAPTER 2

SERVICES AND TOOLS REQUIRED

2.1 Services Used

- **Data Collection and Storage Services**: Banks need to collect and store customer data in real-time. This could be achieved through services like Azure Data Factory, Azure Event Hubs, or AWS Kinesis for real-time data collection, and Azure SQL Database or AWS RDS for data storage.

- **Data Processing Services**: Services like Azure Stream Analytics or AWS Kinesis Data Analytics can be used to process the real-time data.

- **Machine Learning Services**: Azure Machine Learning or AWS SageMaker can be used to build predictive models based on historical data.

2.2 Tools and Software used

Tools:

- **PowerBI**: The main tool for this project is PowerBI, which will be used to create interactive dashboards for real-time data visualization.

- **Power Query**: This is a data connection technology that enables you to discover, connect, combine, and refine data across a wide variety of sources.

Software Requirements:

- **PowerBI Desktop**: This is a Windows application that you can use to create reports and publish them to PowerBI.

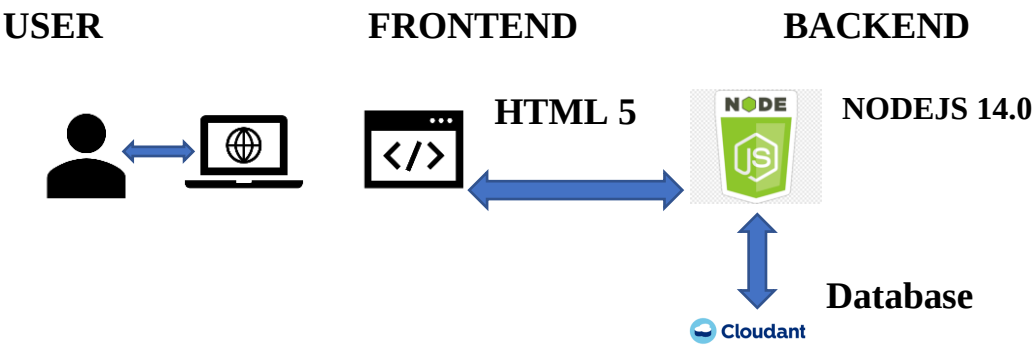
- **PowerBI Service**: This is an online SaaS (Software as a Service) service that you use to publish reports, create new dashboards, and share insights.

- **PowerBI Mobile**: This is a mobile application that you can use to access your reports and dashboards on the go.

CHAPTER 3

PROJECT ARCHITECTURE

3.1 Architecture



Here's a high-level architecture for the project:

1. **Data Collection**: Real-time customer data is collected from various sources like bank transactions, customer interactions, etc. This could be achieved using services like Azure Event Hubs or AWS Kinesis.
2. **Data Storage**: The collected data is stored in a database for processing. Azure SQL Database or AWS RDS can be used for this purpose.
3. **Data Processing**: The stored data is processed in real-time using services like Azure Stream Analytics or AWS Kinesis Data Analytics.
4. **Machine Learning**: Predictive models are built based on processed data using Azure Machine Learning or AWS SageMaker. These models can help in predicting customer behavior, detecting fraud, etc.
5. **Data Visualization**: The processed data and the results from the predictive models are visualized in real-time using PowerBI. PowerBI allows you to create interactive dashboards that can provide valuable insights into the data.
6. **Data Access**: The dashboards created in PowerBI can be accessed through PowerBI Desktop, PowerBI Service (online), and PowerBI Mobile.

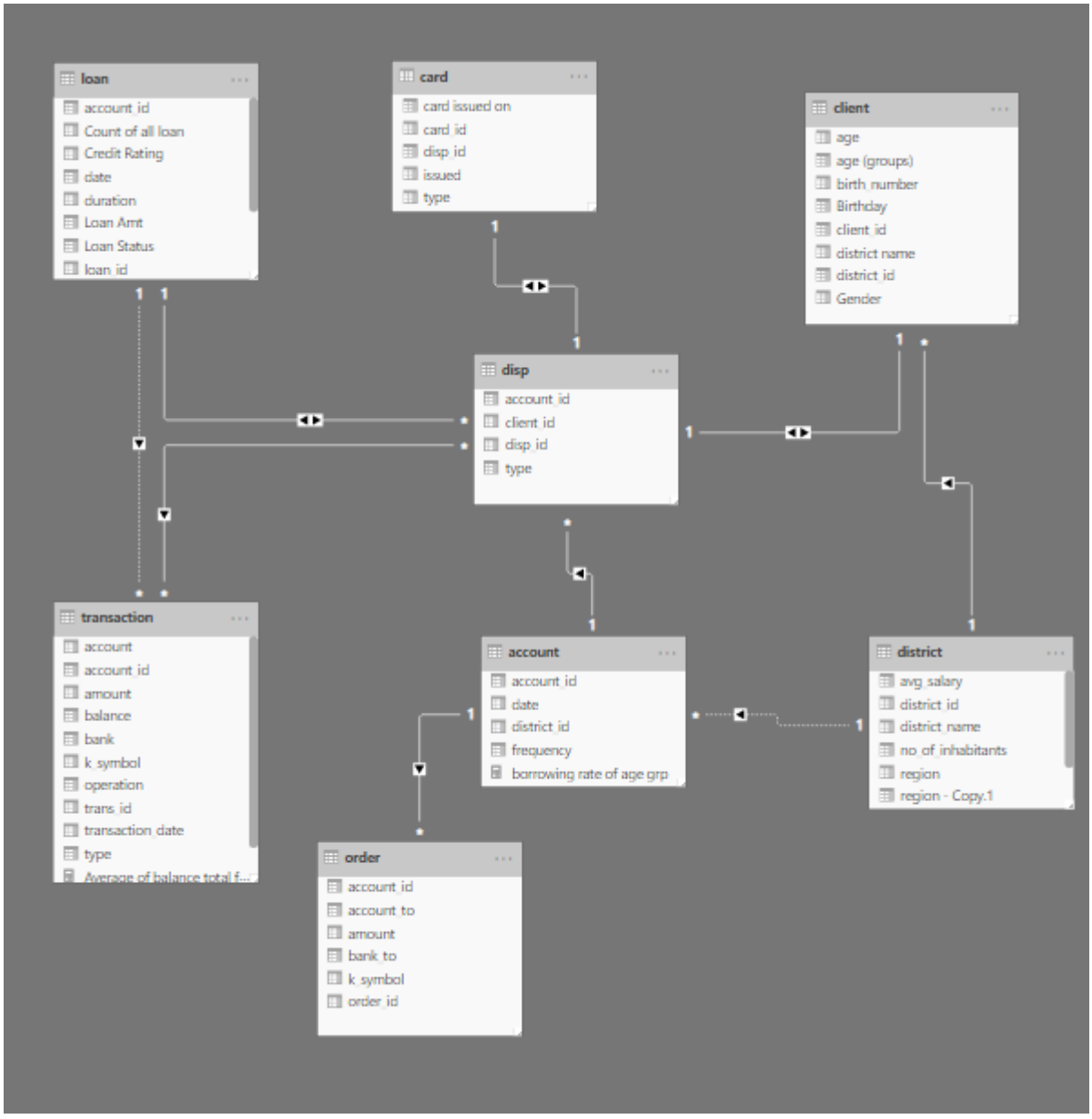
This architecture provides a comprehensive solution for real-time analysis of bank customers. However, it's important to note that the specific architecture may vary depending on the bank's existing infrastructure, specific requirements, and budget. It's also important to ensure that all tools and services comply with relevant data privacy and security regulations.

CHAPTER 4

MODELING AND RESULT

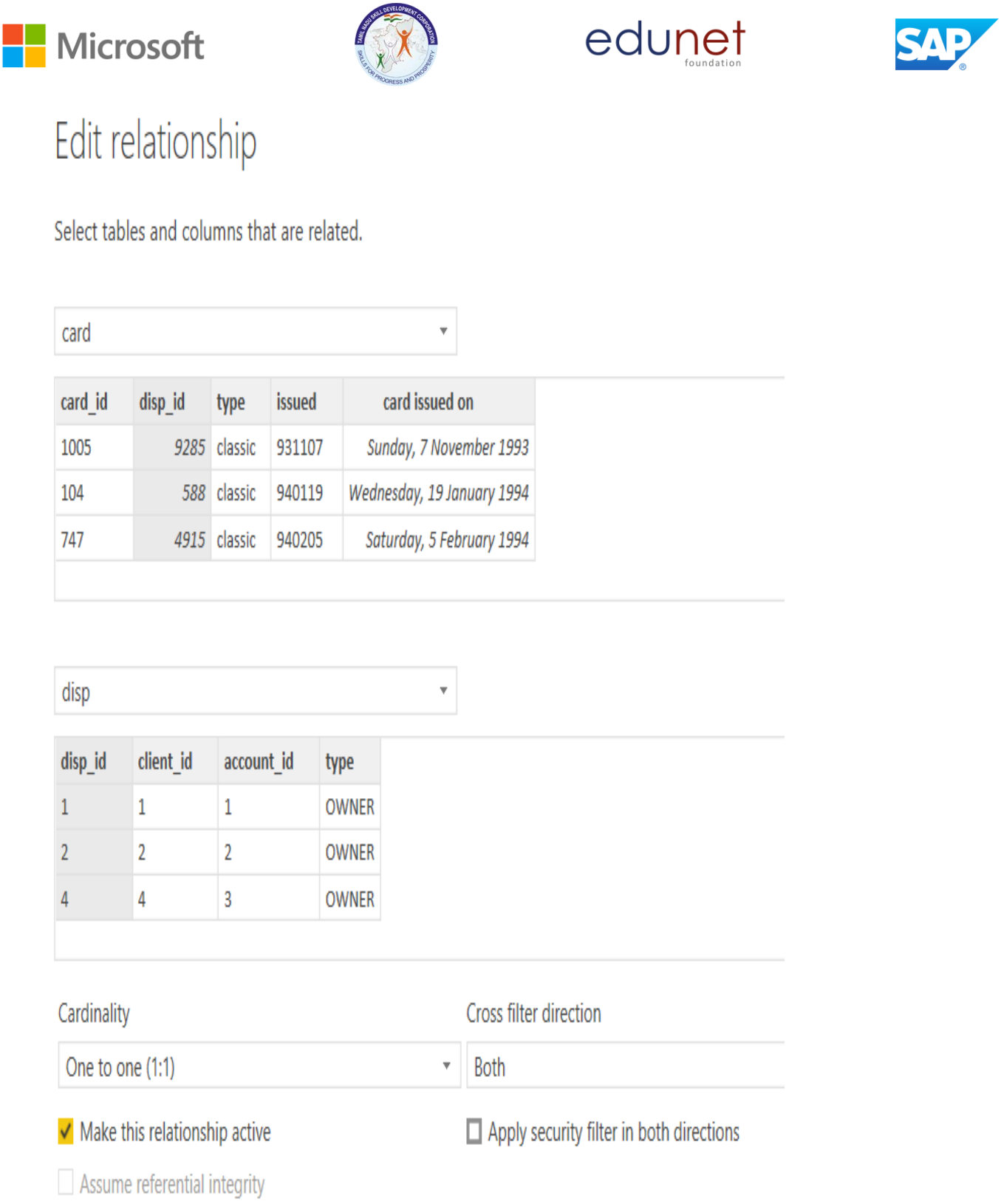
Manage relationship

The “disp” file will be used as the main connector as it contains most key identifier (account id, client id and disp id) which can be use to relates the 8 data files together. The “district” file is use to link the client profile geographically with “district id”



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.


```
1 Gender =
2 VAR stringDate = FORMAT(client[birth_number],"General Number")
3 VAR month = VALUE(MID(stringDate,3,2))
4 RETURN IF(month > 50,"F","M")
5
```

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.

```
1 Birthday =
2 VAR stringDate = FORMAT(client[birth_number],"General Number")
3 VAR stringMonth = VALUE(MID(stringDate,3,2))
4 VAR mth = IF(stringMonth > 50, stringMonth - 50,stringMonth)
5 VAR year = VALUE(MID(stringDate,1,2))
6 VAR day = VALUE(MID(stringDate,5,2))
7 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

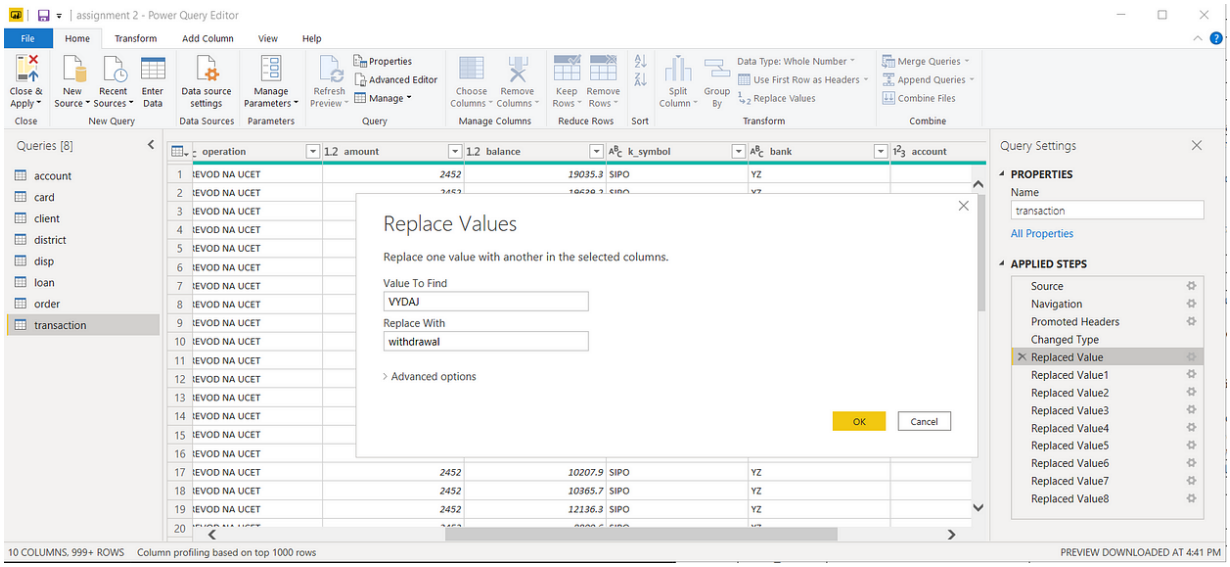
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
1 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

Replacing values

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

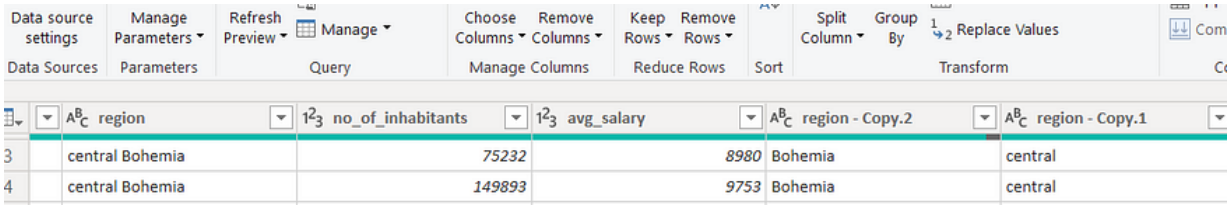


type	+/- transaction	"PRIJEM" stands for credit "VYDAJ" stands for withdrawal

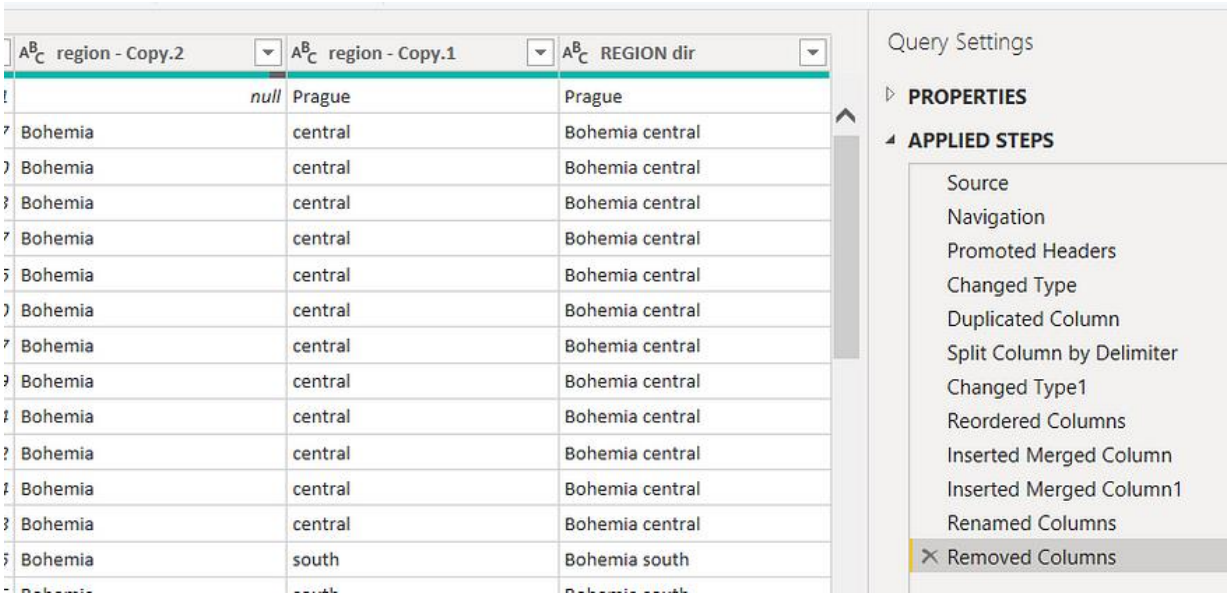
k_symbol	characterization of the transaction	"POJISTNE" stands for insurance payment "SLUZBY" stands for payment for statement "UROK" stands for interest credited "SANKC. UROK" sanction interest if negative balance "SIPO" stands for household "DUCHOD" stands for old-age pension "UVER" stands for loan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



region - Copy.2	region - Copy.1	REGION dir
null	Prague	Prague
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	south	Bohemia south

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.

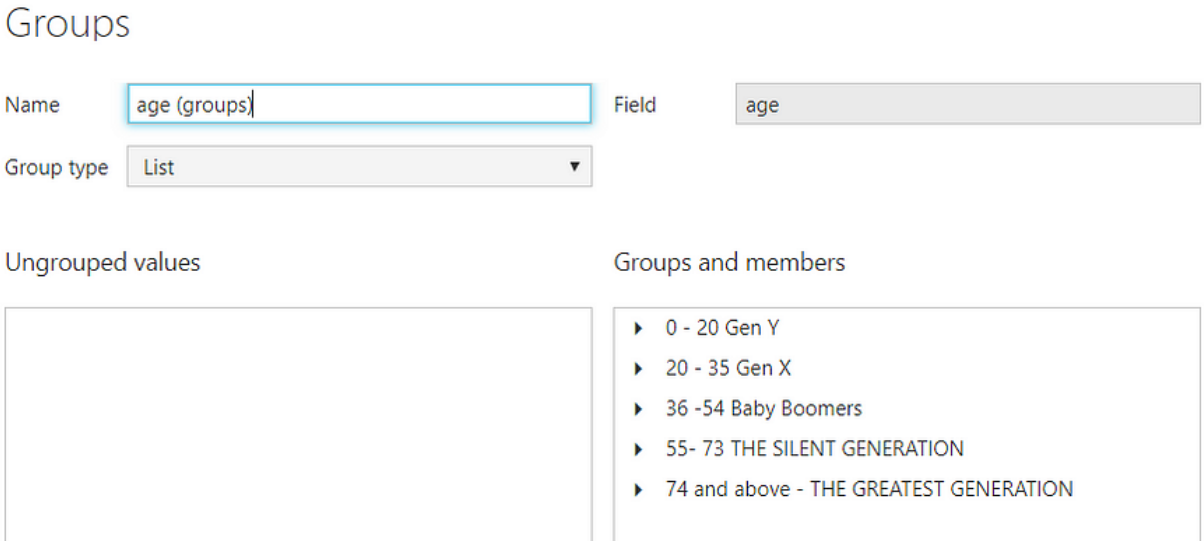
The Gen Y are youths,

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.



Dashboard

CONCLUSION

In conclusion, leveraging real-time data analysis of bank customers through Power BI offers numerous benefits. It allows banks to gain deeper insights into customer behavior, preferences, and trends, enabling them to make informed decisions promptly. By identifying patterns in customer transactions, banks can personalize services, improve customer satisfaction, and drive revenue growth. Additionally, real-time analytics empower banks to detect and respond to potential fraud or risks swiftly, enhancing security measures. Overall, integrating Power BI for real-time analysis enhances operational efficiency, fosters customer trust, and positions banks for success in today's dynamic financial landscape

FUTURE SCOPE

The future scope of this project is vast. With the advent of advanced analytics and machine learning, PowerBI 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, PowerBI'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://www.microteklearning.com/blog/power-bi-for-banking-dashboard-to-analyze-financial-operations/

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

