Tech Saksham

Case Study Report

Data Analytics with Power BI

“Real-Time Analysis of Bank Customers ”

“ Salem Sowdeswari College SALEM-10 ”

NM ID	NAME
5059A48A08EF1B7BA0F436AC5C3265BC	Krishna Prakash.K.G

Trainer Name: R UMAMAHESWARI

Master Trainer: R UMAMAHESWARI

ABSTRACT

In the fast-paced world of banking, understanding customer behaviour in real-time is paramount for ensuring optimal service delivery and maintaining competitive advantage. This abstract presents a methodology for real-time analysis of bank customers utilizing Power BI, a powerful business analytics tool.

The proposed approach involves the integration of real-time data streams from various banking channels such as ATMs, online banking platforms, and mobile applications into Power BI. These data streams encompass diverse customer interactions, including transactions, account inquiries, loan applications, and customer service requests.

INDEX

CHAPTER 1 INTRODUCTION

1.1 Problem Statement:

In the dynamic landscape of banking, understanding and responding to customer behaviour in real-time is essential for maintaining competitiveness and ensuring customer satisfaction. However, traditional methods of analysing customer data often suffer from delays and lack the agility needed to address rapidly evolving trends and customer preferences. Banks face the challenge of efficiently harnessing vast volumes of real-time data from various channels such as ATM transactions, online banking activities, and mobile interactions to derive actionable insights.

Moreover, there is a need to detect anomalies promptly, identify fraudulent activities, and predict customer churn to mitigate risks and preserve customer trust. Traditional analytics approaches may struggle to provide timely alerts and proactive measures, leaving banks vulnerable to financial losses and reputational damage.

1.2 Proposed Solution:

To address these challenges, a real-time analysis solution leveraging Power BI can be implemented in the banking sector. This solution involves the integration of data streams from multiple banking channels into Power BI for continuous monitoring and analysis. Here's how it works:

Real-time data streams from various sources such as ATM networks, online banking platforms, and mobile applications are aggregated and ingested into Power BI. This integration ensures a holistic view of customer interactions across different touchpoints.

Power BI's visualization capabilities enable the creation of interactive dashboards that display key metrics and KPIs in real-time. Bank stakeholders, including executives, managers, and frontline staff, can access these dashboards to monitor customer activities, transaction volumes, account balances, and other relevant indicators.

Power BI's advanced analytical features, including predictive modeling and machine learning algorithms, enable banks to anticipate customer behaviour

and trends. By analysing historical data alongside real-time inputs, banks can identify patterns, detect anomalies, and predict future outcomes such as customer churn or potential fraud.

Power BI can be configured to generate alerts and notifications based on predefined thresholds or anomalies detected in real-time data. This proactive approach enables banks to take immediate action to address issues such as suspicious transactions, account breaches, or unusual spending patterns.

By leveraging Power BI's capabilities for segmentation and customer profiling, banks can deliver personalized insights and recommendations to individual customers. This enhances the overall customer experience and strengthens customer loyalty.

The real-time analysis solution using Power BI facilitates continuous monitoring and iterative refinement of banking operations. By analysing feedback and performance metrics, banks can identify areas for improvement and implement data-driven strategies to enhance service delivery and operational efficiency.

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 behaviour, etc.
- **Trend Analysis**: The dashboard will identify and display trends in customer behaviour.
- **Predictive Analysis**: It will use historical data to predict future customer behaviour.

1.4 Advantages:

- **Data-Driven Decisions:** Banks can make informed decisions based on real-time data analysis.

- **Improved Customer Engagement:** Understanding customer behaviour 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 project's reach encompasses all banking organizations seeking to utilize data for decision-making and engaging with customers. It can expand to integrate additional data origins and advanced analytics methods, like machine learning and artificial intelligence, to offer deeper insights into customer behaviour. Moreover, it holds promise for adaptation across various industries, including retail, healthcare, and telecommunications, where comprehending customer behaviour holds significance. Additionally, it aligns with the overarching objective of digital transformation within the banking sector, fostering efficiency, creativity, and prioritizing customer needs.

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**:

- **Power BI**: The main tool for this project is Power BI, 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.

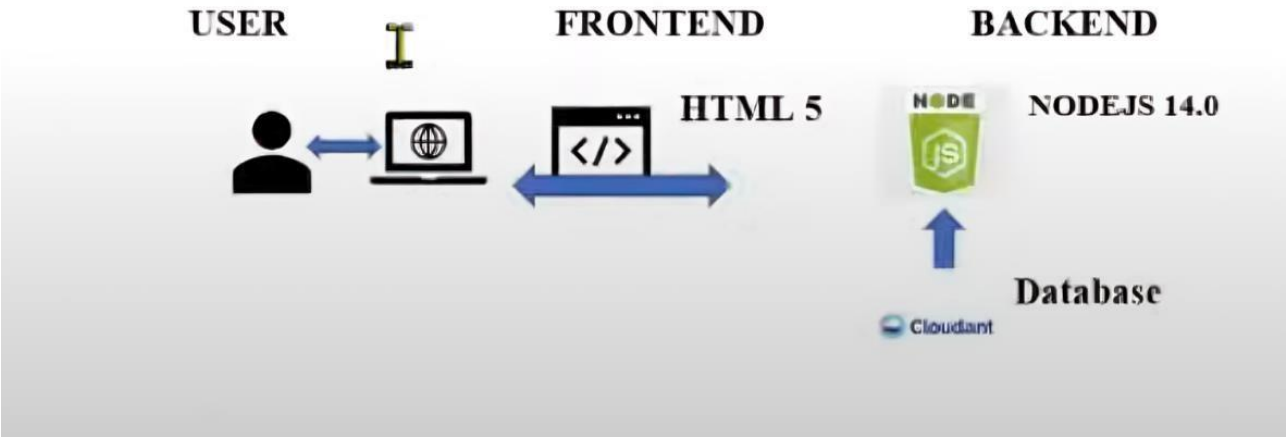
Software Requirements:

- **Power BI Desktop**: This is a Windows application that you can use to create reports and publish them to Power BI.
- **Power BI Service**: This is an online SaaS (Software as a Service) service that you use to publish reports, create new dashboards, and share insights.
- **Power BI 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:

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.

Data Storage: The collected data is stored in a database for processing. Azure SQL Database or AWS RDS can be used for this purpose.

Data Processing: The stored data is processed in real-time using services like Azure Stream Analytics or AWS Kinesis Data Analytics.

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.

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.

Data Access: The dashboards created in PowerBI can be accessed through PowerBI Desktop, PowerBI Service (online), and PowerBI Mobile.

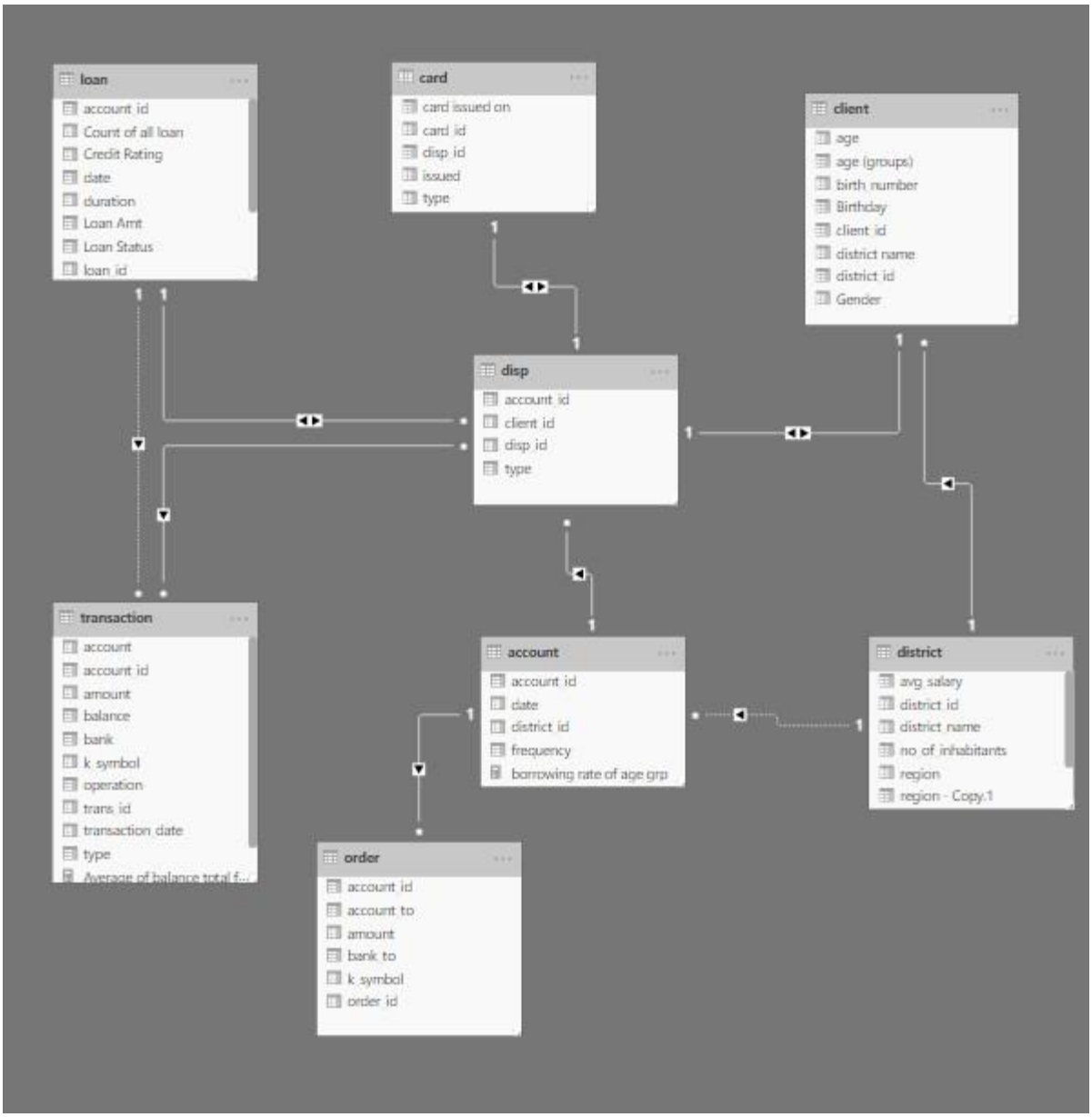
This architecture provides a comprehensive solution for real-time analysis of food delivery apps. However, it's important to note that the specific architecture may vary depending on the food delivery connection, specific requirements, and budget. It's also important to ensure that all tools and services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

CHAPTER 4

MODELING AND RESULT

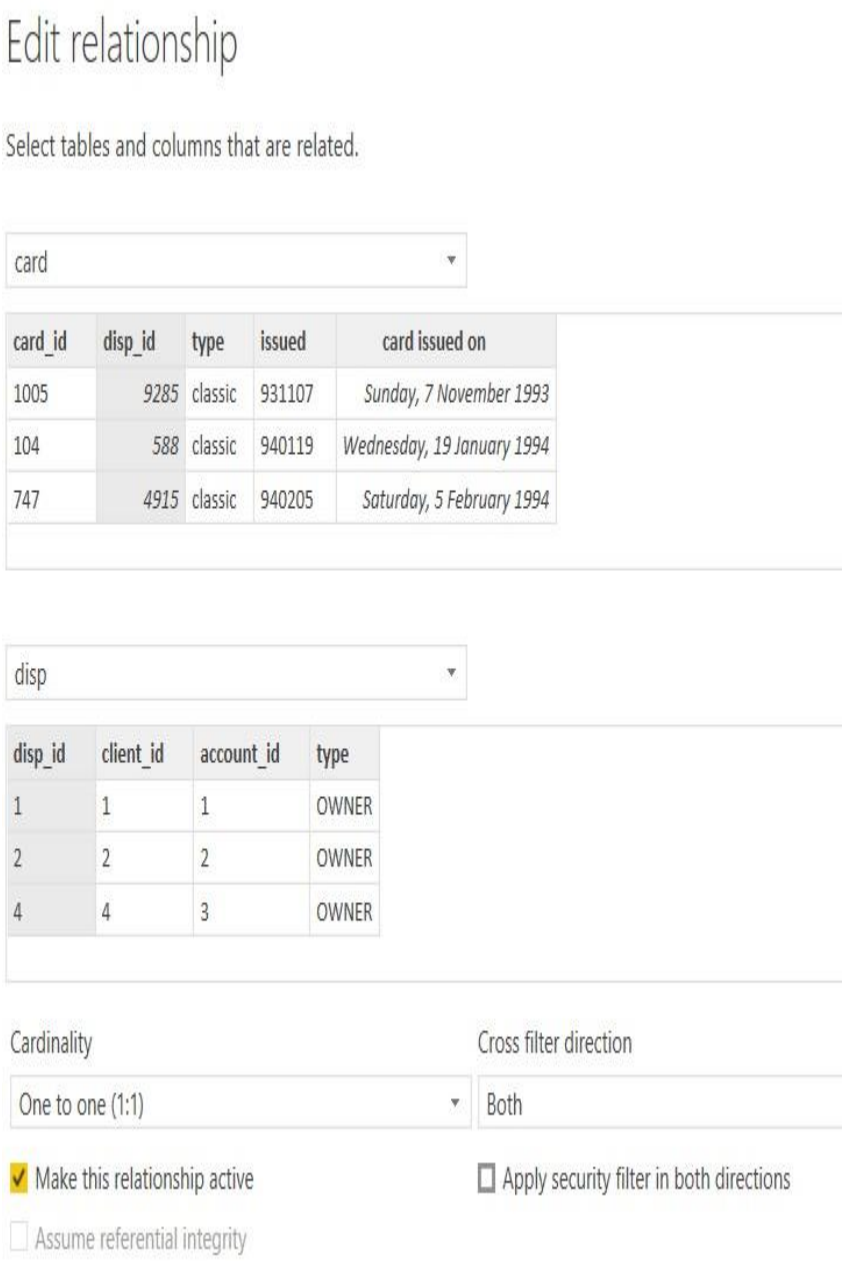
Manage relationship

The "disp" file will be use 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"


loan
account id
Count of all loan
Credit Rating
date
duration
Loan Amt
Loan Status
loan id
card
card issued on
card id
disp id
issued
type
client
age
age (groups)
birth number
Birthday
client id
district name
district id
Gender
disp
account id
client id
disp id
type
transaction
account
account id
amount
balance
bank
k symbol
operation
trans id
transaction date
type
account
account id
date
district id
frequency
borrowing rate of age grp
district
avg salary
district id
district name
no of inhabitants
region
region - Copy.1
order
account id
account to
amount
bank to
k symbol
order 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

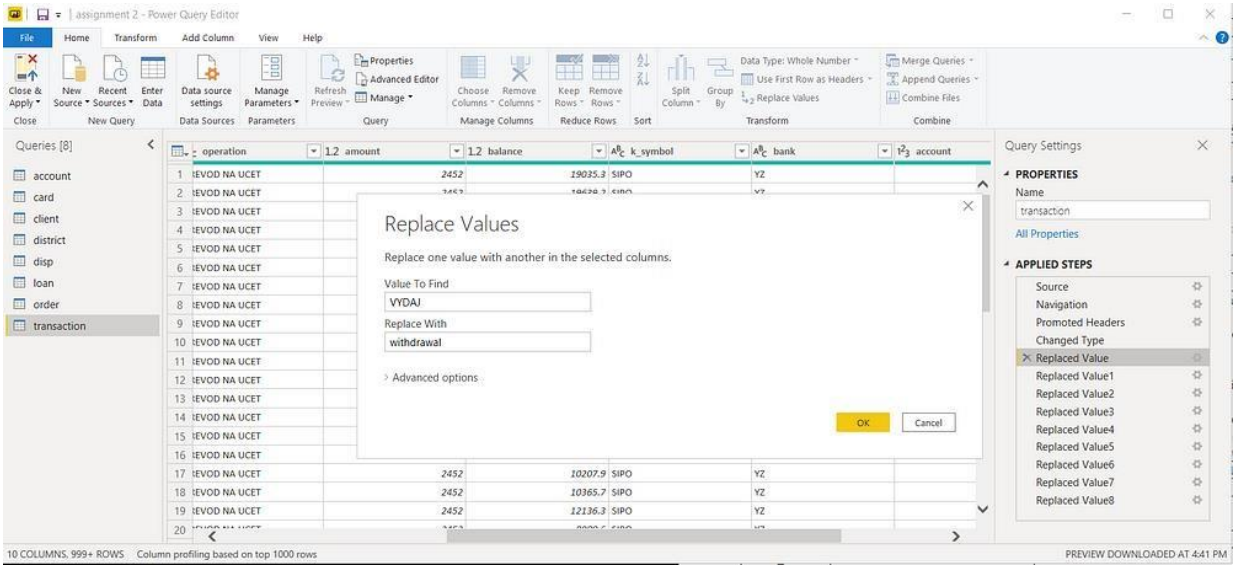
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
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.

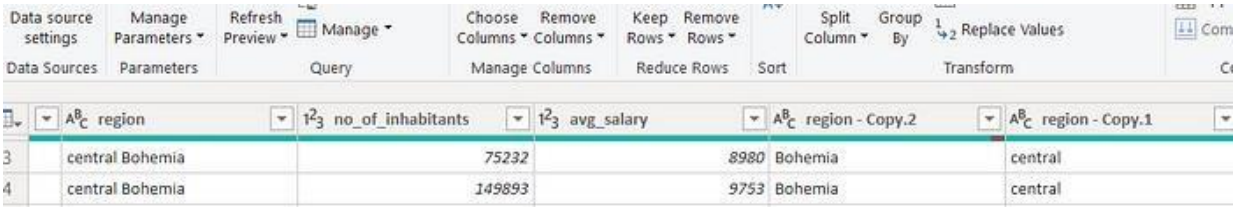


type	+/- transaction	"PRIJEM" stands for credit "VYDAJ" stands for withdrawal

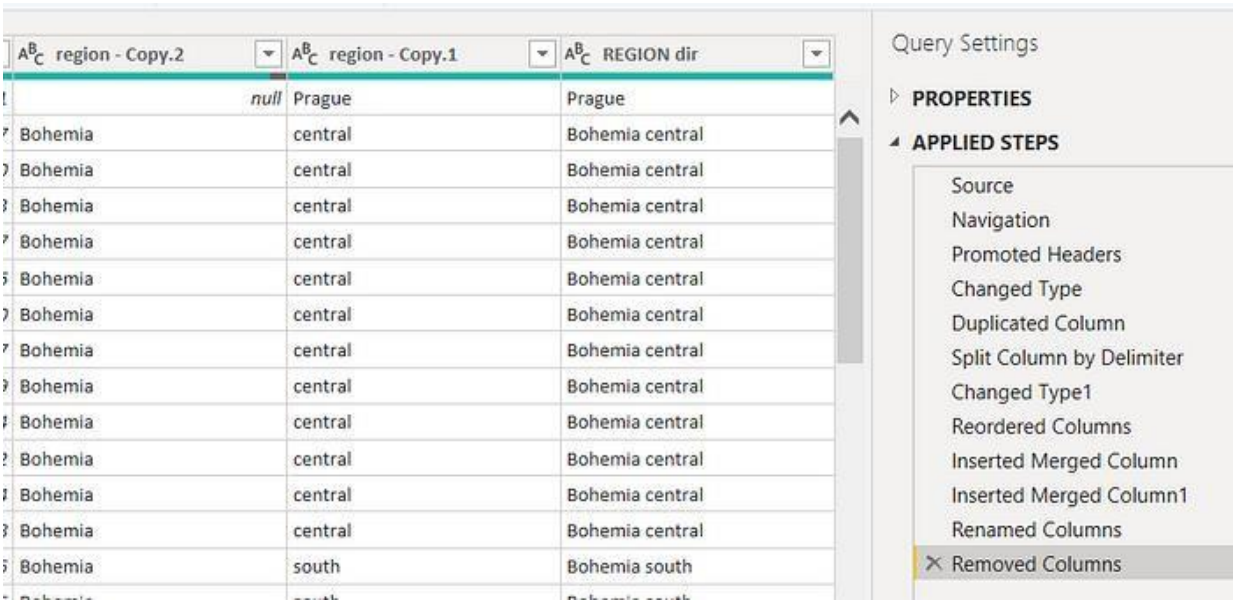
k_symbol	characterization of the transaction	"POJISTNE" stands for insurance payment "SLUZBY" stands for payment for statement "UROK" stands for interest credited "SANKC. UROK" sanction interest if negative balance "SIPO" stands for household "DUCHOD" stands for old-age pension "UVER" stands for loan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

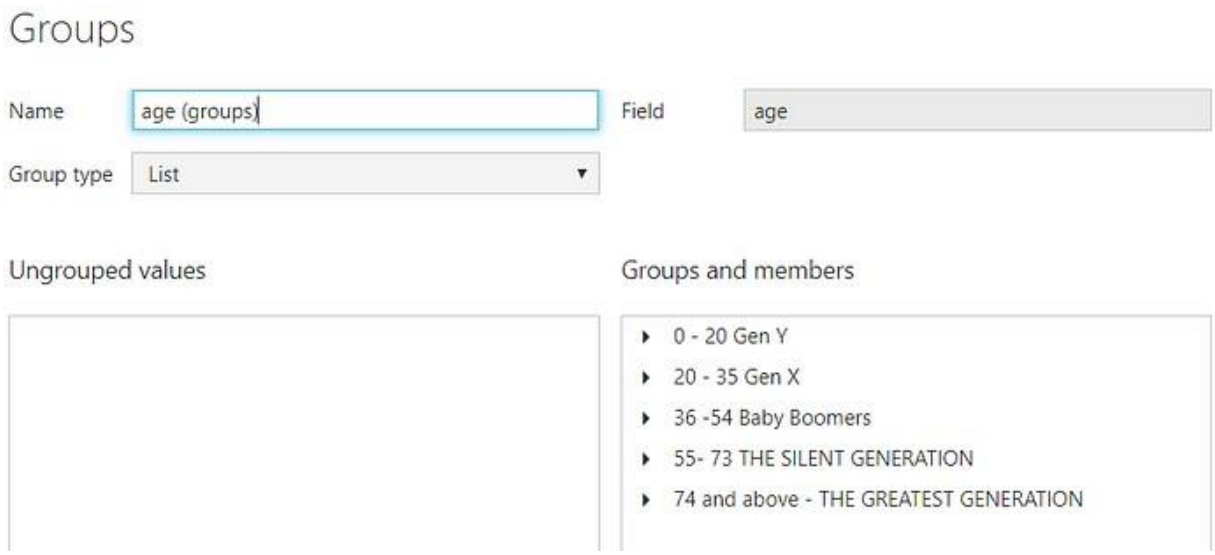
The Gen Y are youths,

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.



Credit Rating and Loan Status

As the Loan status uses A, B, C, D which are not reader friendly. We can add a column to represent what it stands for, we also simplify the classification of those with late or default on payment as bad credit, refer to the table below for details on the new columns added.

Status in "loan" data	New column "loan status"	New column "credit rating"
'A' stands for contract finished no problems	Fully Repaid	Good
'B' stands for contract finished loan not payed	Default	Bad
'C' stands for running contract OK so far	Timely Payment	Good
'D' stands for running contract client in debt	Late payment	Bad

```
1 Loan Status =
2 IF(loan[status]="A","Repaid Full",
3 IF(loan[status]="B","Default",IF (loan[status]="c","Timely payment","Late payment" )))
```

loan_id	account_id	date	Loan Amt	duration	payments	status	Credit Rating	Loan Status
6059	5196	971228	79,824 Kč	12	6652	A	GOOD	Repaid Full
6727	8505	971210	42,840 Kč	12	3570	A	GOOD	Repaid Full

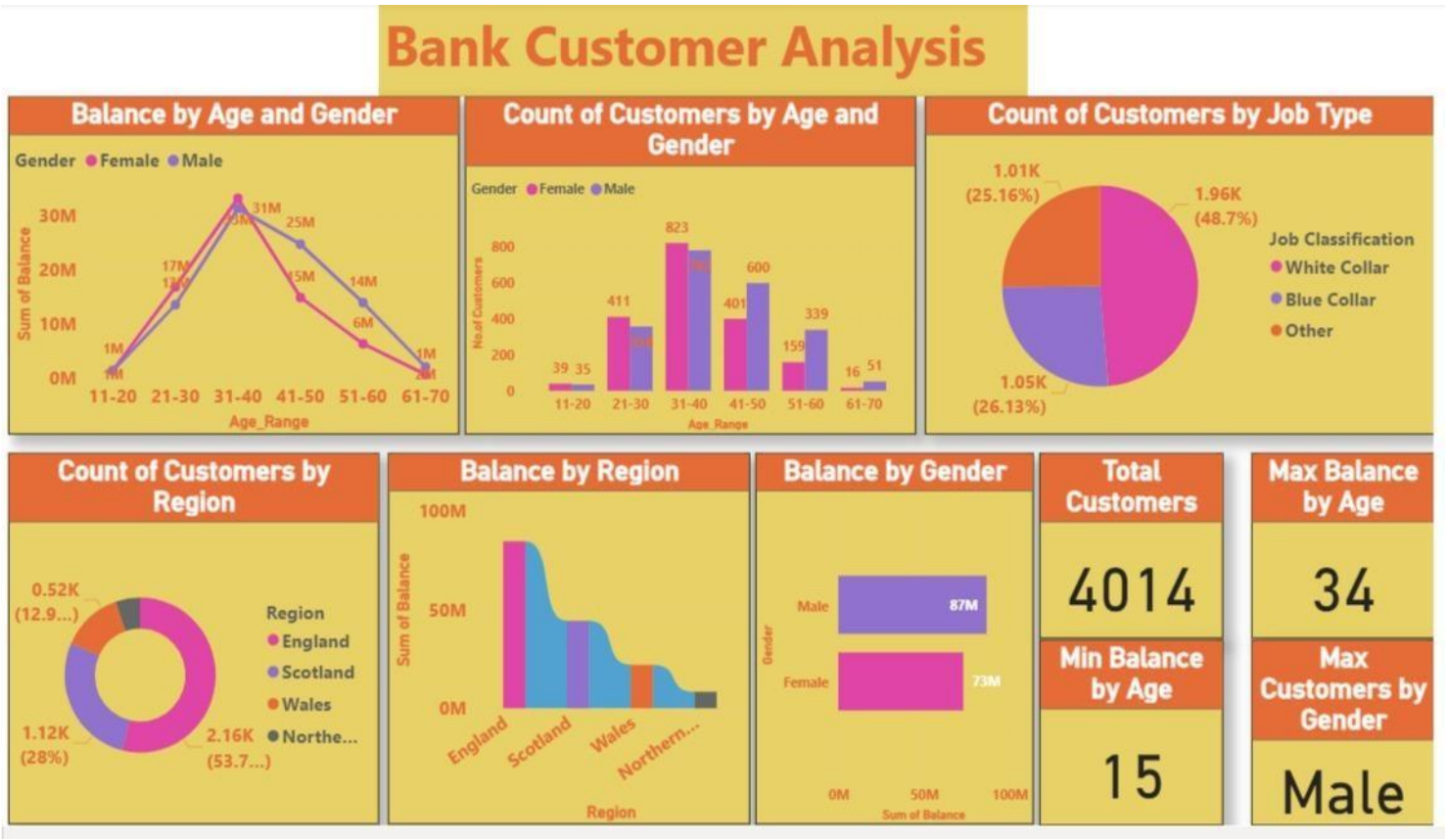
```
1 Credit Rating =
2 IF(loan[status]="A","GOOD",
3 IF(loan[status]="B","BAD",IF (loan[status]="c","GOOD","BAD" )))
```

loan_id	account_id	date	Loan Amt	duration	payments	status	Credit Rating	Loan Status
5221	1284	981205	52,512 Kč	12	4376	C	GOOD	Timely payment
5841	4268	981104	41,988 Kč	12	3499	C	GOOD	Timely payment

Values of such as "account Id" have also been set as Text.

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

Dashboard


Bank Customer Analysis
Balance by Age and Gender
Gender Female Male
Sum of Balance
30M
20M
10M
0M
11-20 21-30 31-40 41-50 51-60 61-70
Age_Range
Count of Customers by Age and Gender
Gender Female Male
No.of Customers
800
600
400
200
0
39 35
411
823
401
600
159
339
16 51
11-20 21-30 31-40 41-50 51-60 61-70
Age_Range
Count of Customers by Job Type
1.01K (25.16%)
1.96K (48.7%)
1.05K (26.13%)
Job Classification
White Collar
Blue Collar
Other
Count of Customers by Region
0.52K (12.9...)
1.12K (28%)
2.16K (53.7...)
Region
England
Scotland
Wales
Northe...
Balance by Region
100M
50M
0M
Sum of Balance
England
Scotland
Wales
Northern...
Region
Balance by Gender
Male
87M
Female
73M
Gender
0M
50M
100M
Sum of Balance
Total Customers
4014
Max Balance by Age
34
Min Balance by Age
15
Max Customers by Gender
Male

CONCLUSION

In conclusion, the implementation of real-time analysis of bank customers using Power BI offers a transformative solution for banks aiming to enhance their operational efficiency, customer engagement, and overall performance. By leveraging Power BI's robust capabilities in data visualization, predictive analytics, and integration with various data sources, banks can gain valuable insights into customer behavior in real-time.The project extends beyond traditional methods of customer analysis, allowing banks to anticipate needs, detect anomalies, and mitigate risks promptly. Moreover, the integration of advanced analytics techniques such as machine learning enables banks to predict future trends and personalize offerings, thereby improving customer satisfaction and loyalty.Furthermore, the scalability of the project enables its adaptation across different sectors, emphasizing its broader relevance beyond banking. As data privacy and security remain paramount, the project underscores the importance of implementing robust data governance strategies to ensure compliance with regulations and safeguard customer information.In essence, the real-time analysis of bank customers using Power BI represents a significant step towards digital transformation in the banking sector, promoting innovation, efficiency, and customer-centricity. As banks continue to evolve in response to changing market dynamics and customer expectations, the insights derived from this project will play a crucial role in driving strategic decision-making and fostering long-term success.

FUTURE SCOPE

The future potential of this project is extensive. With the advancements in analytics and machine learning, PowerBI can be utilized to forecast future trends by analyzing historical data. By incorporating these predictive analytics, the bank could foresee customer requirements and offer proactive solutions. Moreover, PowerBI's ability to integrate with diverse data sources allows for the inclusion of a wider range of datasets, providing a more comprehensive understanding of customers. As data privacy and security gain increasing importance, future phases of this project should prioritize the implementation of robust data governance measures. This ensures the safe management of sensitive customer information while adhering to data protection regulations. Additionally, the project could explore incorporating real-time data streams to deliver even more timely and pertinent insights. This has the potential to revolutionize how banks engage with their customers, ultimately enhancing customer satisfaction and loyalty.

REFERENCES

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

LINK

https://github.com/KP0112/Real-Time-Analysis-of-Bank-Customers-Data-Analytics-with-Power-BI-