Tech Saksham

Case Study Report

Date Analysis with Power BI

“Real-Time Analysis of bank customers”

“VIDHYAA ARTS AND SCIENCE COLLEGE KONGANAPURAM”

NM ID	NAME
3BA1220083A8FB89AAC4FCCC10D0828D	KARTHI A

Trainer Name: R UMAMAHESWARI

Master Trainer: R UMAMAHESWARI

ABSTRACT

In the digital age, date 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 busienesses intelligence too, to analyze and visualize real-time customer date. This project will enable banks to gain deep insights into customer behavior, preferences, and trends, thereby facillitaing date-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 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

1.1Problem Statement

In today's competitive banking landscaps, understanding customer behavior and preferences is crucil for customer retentionand revenue generation. However , banks often face challenges in analyzing customer date due to the sheer volumeand velocity of date generated. Traditional date 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 date,which includes transaction histtry, customer feedback, and demographic date, pose additional challenges for date analysis.

1.2 Proposed solution

The proposed solution is to develop a Power BI dashboard that can anaiyze and visualize real-time customer date. The dashboard will integrate date form various 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 provided real-time analysis of customer data.
- **Customer segmentation:** It will segment customer based on various parameters like age, income, transaction behavior, etc.
- **Trend Analysis:** The dashboard will identify and displa trends in customer behavior.
- **Predictive Analysis:** It will use historical date to predict future customer behavior.

1.4 Advantages

- **Date-Driven Decisions:** Banks can make informed decisions based on real-time date 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 aimto leverage date for decision- making and customer engagement. The project can be further exended to incorporate more date sources and advanced analysis techniques, such as machine learing and artificial intelligence, to provide more sophisticated insights into customer behavior. The project also has potential to be adapted fo other sectors, such as retail, healthcare, and telecommunications, where underatandin 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 Servlces Used

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

- **Date Processing Services:**
  Services like Azure Stream Analysis or AWS Kinesis Date Analysis can be used to process the real- time date.

- **Machine Leaming Sevices:**
  Azure Machine Learning or AWS Sagemaker can be used to build predictive models based on historical date.

2.2 Tools and Software used

Tools:

- **Power BI:** The main tool for this project is Power BI , which will to create interactive dashboards for real-time date visualization.

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

Software Requiremenets:

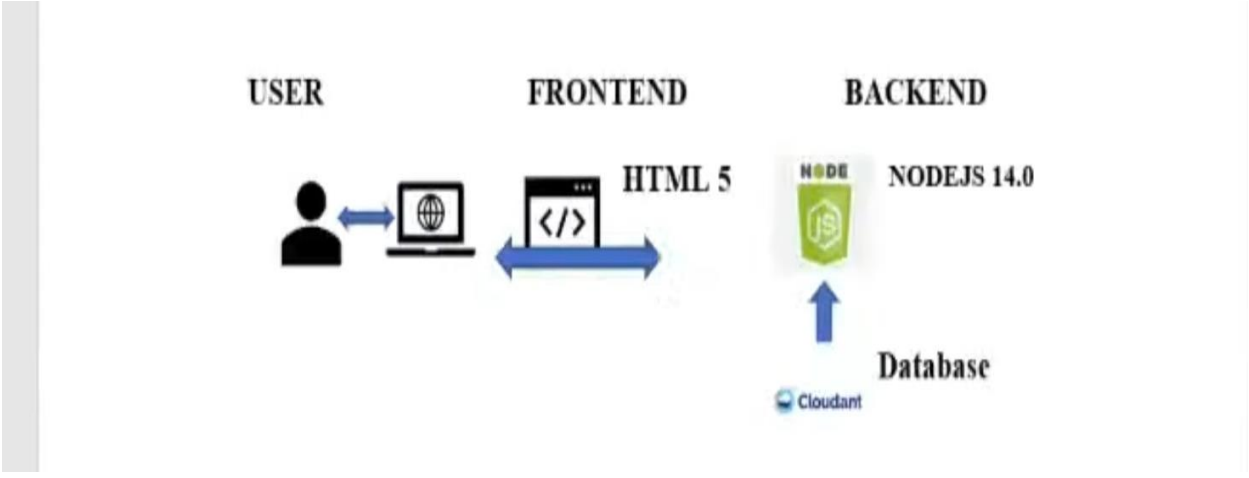
- **PowerBI Desktop:** This is a windows application that you canto create’” reports and publish them to PowerBI .

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

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

CHAPTER 3
PROJECT ARCHITECTURE

3.1 Architecture



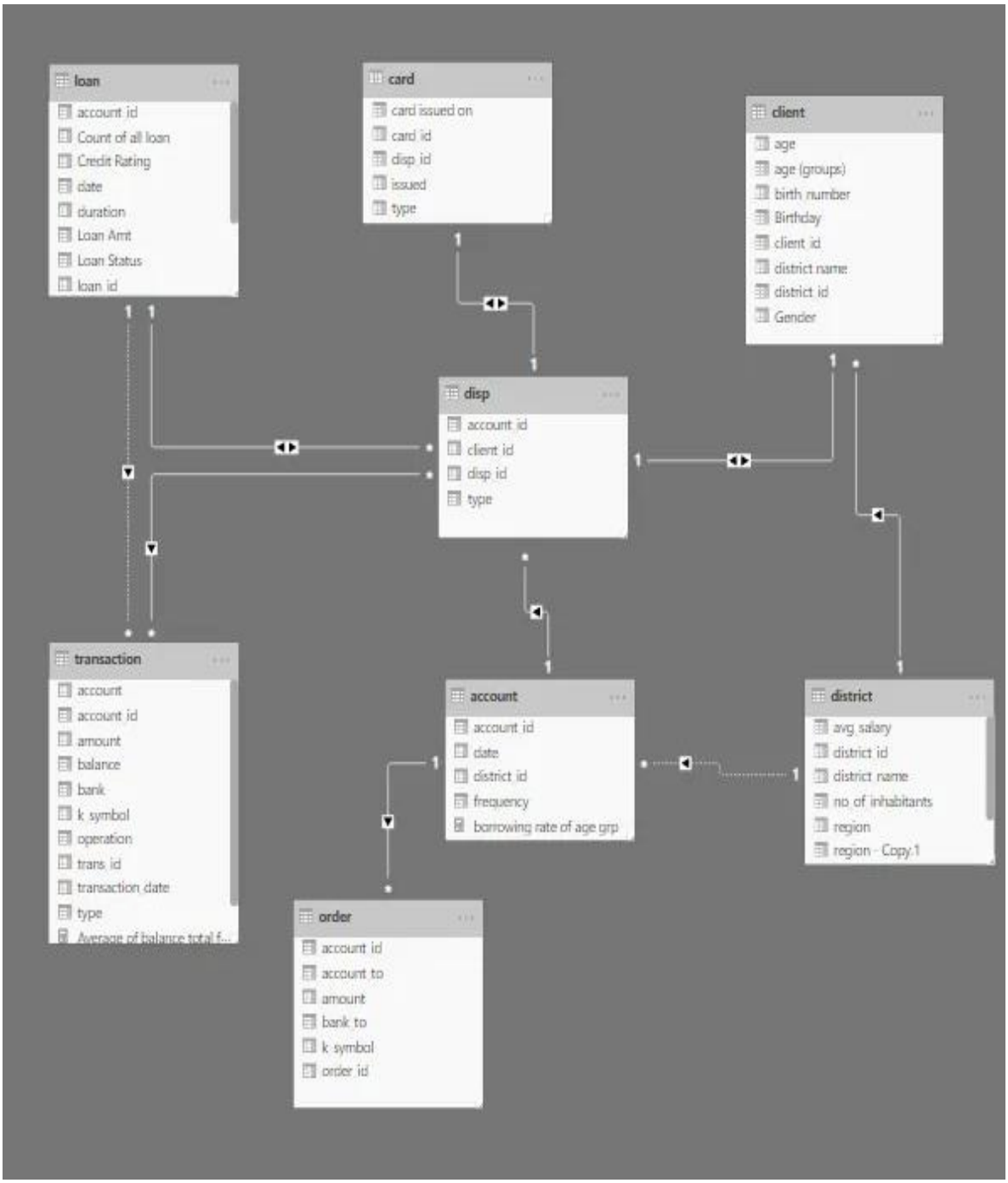
1. **Date collection:**
   Real -time customer date is collected from various sources like bank transactions, customer interactions, etc. This could be achieved using services like Azure Event Hubs or AWS Kinesis.
2. **Date Storage:**
   The collected date is processed in real-time using services like Azure Stream Analytics or AWS RDS Kinesis Date Analysis.
3. **Machine Learning :**
   Predictive models are built based on processed date using Azure machine learning or AWS Kinesis Date Analysis.
4. **Date Proessing:**
   The stored date is processed in real-time using services like Azure Stream Analytics or AWS Kinesis Date Analysis.
5. **Date Access:** The dashboards created in Power BI can be accessed through Power BI Desktop, Power BI Service(online), and Power BI 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 uses 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

Edit relationship

Select tables and columns that are related.

card

card_id	disp_id	type	issued	card issued on
1005	*9285*	classic	931107	*Sunday, 7 November 1993*
104	*588*	classic	940119	*Wednesday, 19 January 1994*
747	*4915*	classic	940205	*Saturday, 5 February 1994*

disp

disp_id	client_id	account_id	type
1	1	1	OWNER
2	2	2	OWNER
4	4	3	OWNER

Cardinality

One to one (1:1)

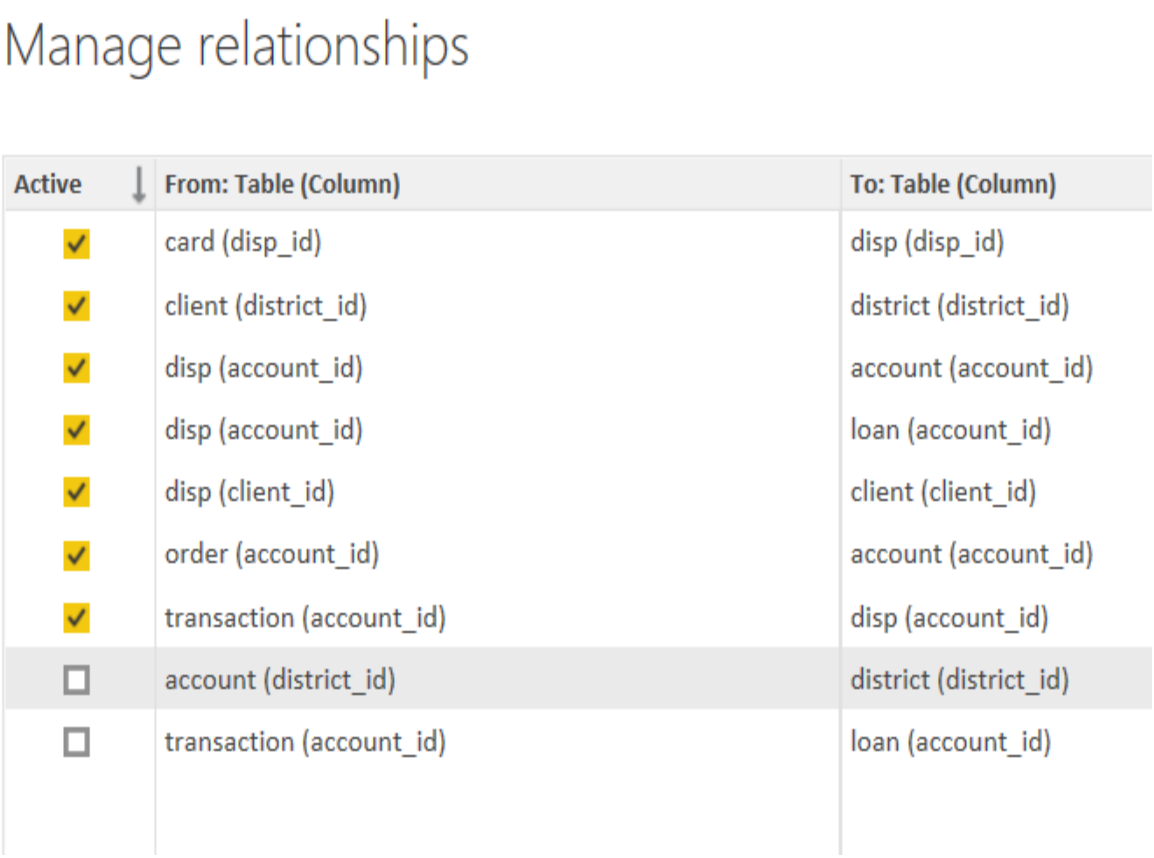
Cross filter direction

Both

☑ Make this relationship active

☐ Apply security filter in both directions

☐ Assume referential integrity



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

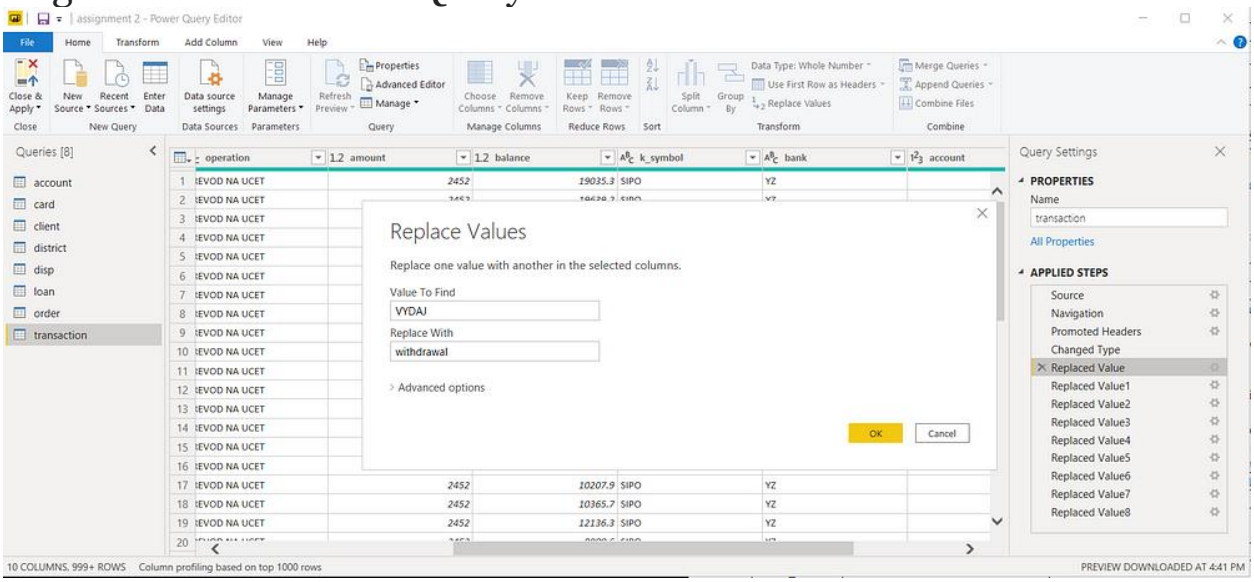
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
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

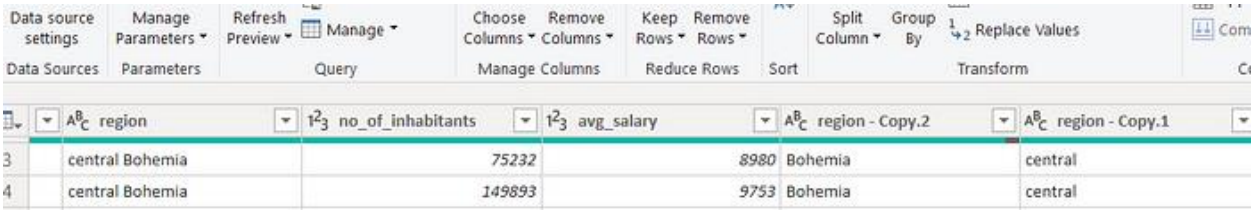
Replacing values

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

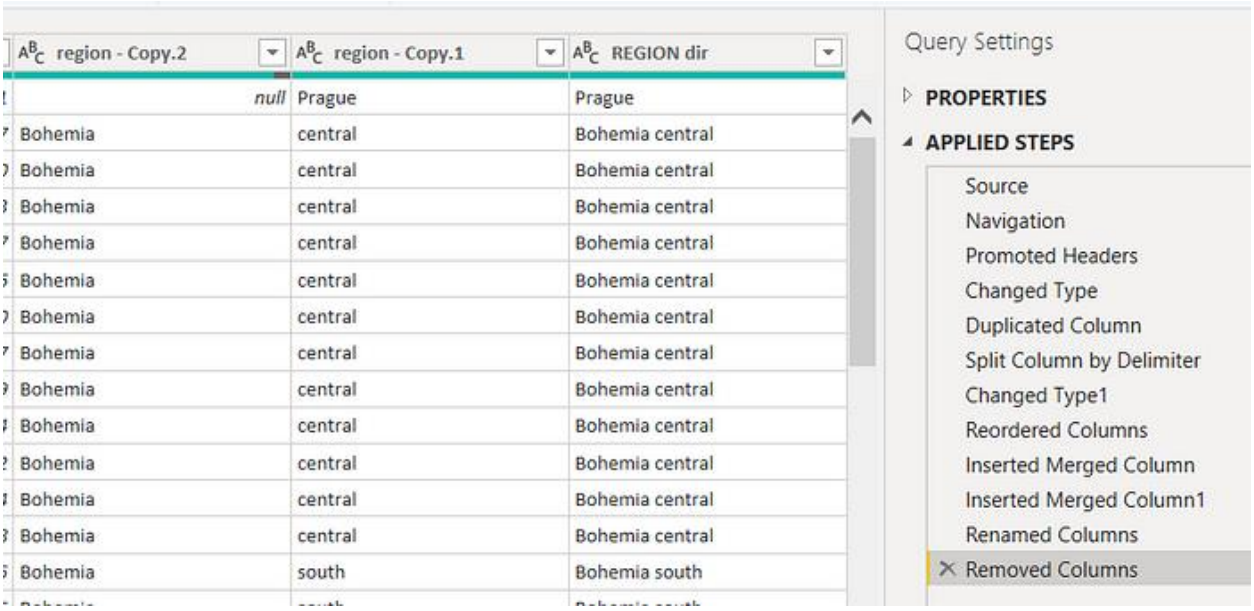


Changing the order of Region name at Power Query

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. Referring to https://marketingartfully.com/customer-demographics-age-ranges-numbers/ (Jacobsen, 2019), we will group the ages into 5 groups.

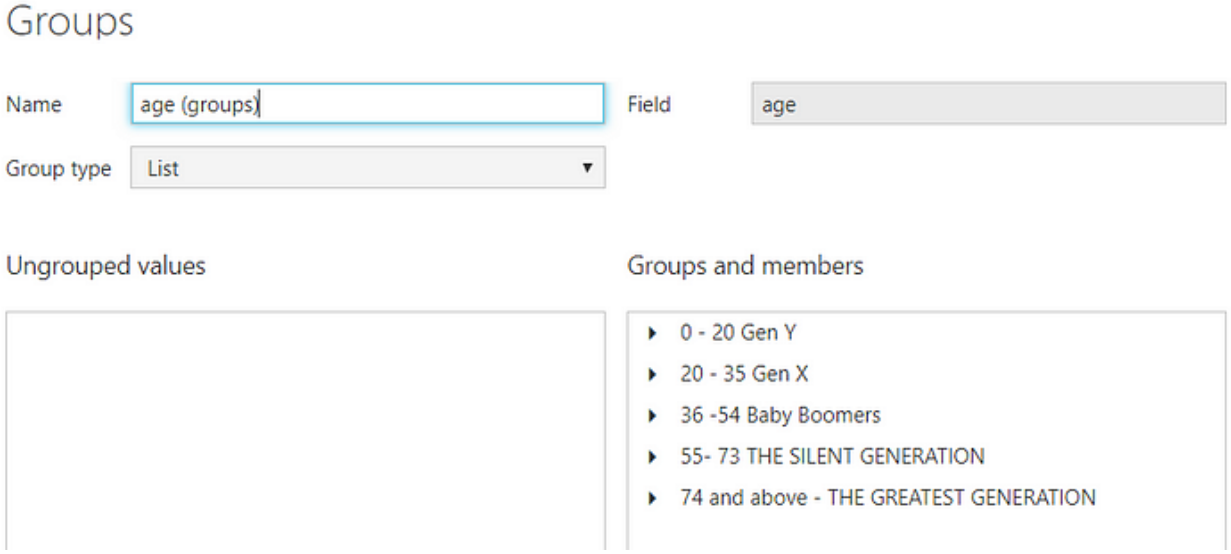
The Gen Y are youths,

Gen X are young working adults, some starting their families

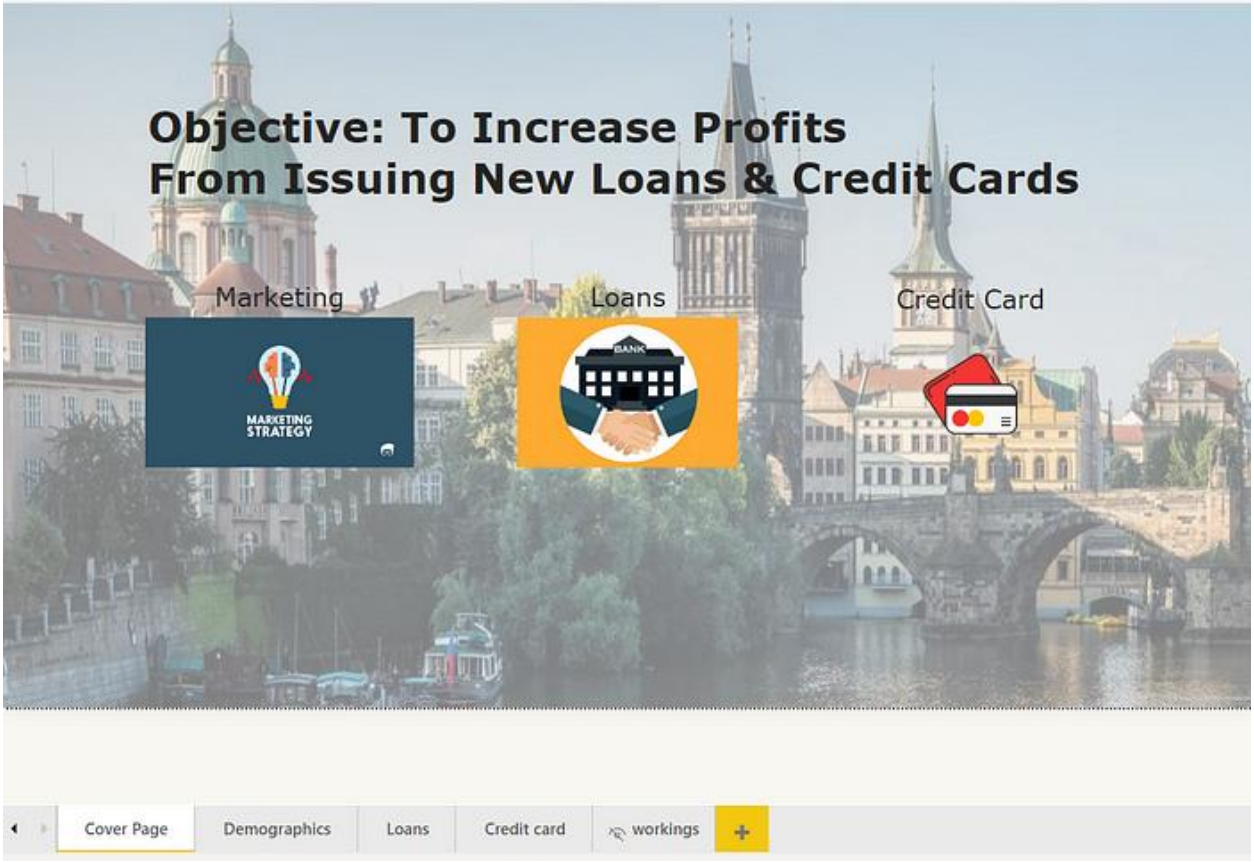
Baby Boomer are working adults with families.

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

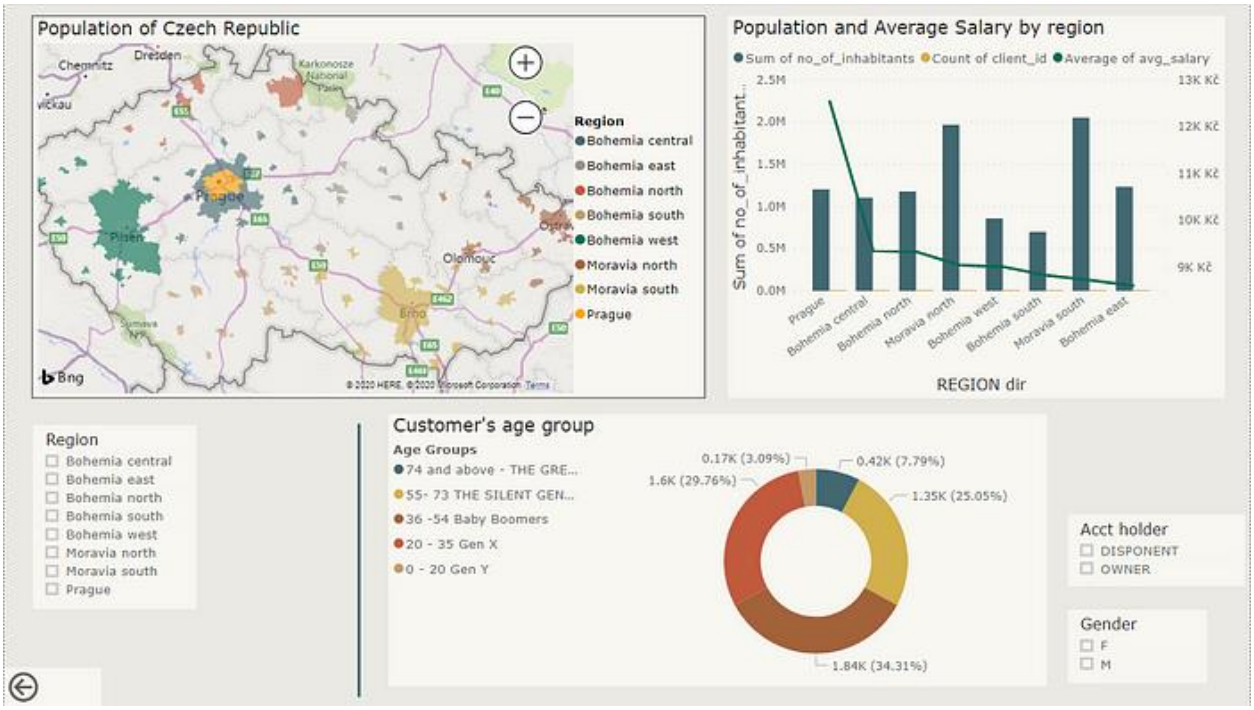
The greatest Generation, retired elderly living on pensions.



The cover shows our objectives with 3 buttons gg to each respective page, demographics analysis for marketing, Loan analysis and Credit card analysis.



The demographics page displays the population distribution of Czech Republic and their average salary rank from highest. User can filter by region to see where are the people located. There is also a donut chart which shows the age group distribution which can be drilldown to gender. User can also filter the gender and the type of account holder.



The side of the loan page show the filters that can be used such as loan status. The top shows a table of the total no. of loan and the sum of the loan amount and also a bar chart of the loan amt. Lower half of the page shows traits of bad credit that we have found out which will be explained later.

Credit Rating
- (Blank)
- BAD
- GOOD

Loan Status
- (Blank)
- Default
- Late payment
- Repaid Full
- Timely payment

Age (groups)
- 0 - 20 Gen Y
- 20 - 35 Gen X
- 36 -54 Baby Boom...
- 55- 73 THE SILENT...
- 74 and above - TH...

Gender
- F
- M

age (groups)	No. of Loan	Sum of Loan Amt	borrowing rate of age grp
55- 73 THE SILENT GENERATION	139	19,488,552 Kč	20.38%
36 -54 Baby Boomers	307	45,346,416 Kč	45.01%
20 - 35 Gen X	253	40,281,432 Kč	37.10%
0 - 20 Gen Y	17	2,794,320 Kč	2.49%
Total	**682**	**103,261,740 Kč**	**100.00%**

682
Total Loan

Total Loan Amount:
103 Mil Kc

Sum of Loan Amt by Credit Rating and Loan Status

Bad Credit

Credit Card ownership

type	Default	Late payment	Total
classic	1	2	3
gold		1	1
junior	1		1
Total	**2**	**3**	**5**

Amount of Bad Debts

REGION dir	Default	Late payment	Total
Bohemia central	508,572 Kč	1,214,652 Kč	**1,723,224 Kč**
Bohemia east		1,532,124 Kč	**1,532,124 Kč**
Bohemia north	215,388 Kč		**215,388 Kč**
Bohemia south	400,476 Kč	1,189,068 Kč	**1,589,544 Kč**
Bohemia west	556,920 Kč	755,040 Kč	**1,311,960 Kč**
Moravia north	1,252,668 Kč	3,043,080 Kč	**4,295,748 Kč**
Moravia south	801,468 Kč	2,363,784 Kč	**3,165,252 Kč**
Prague	626,856 Kč	1,120,056 Kč	**1,746,912 Kč**
Total	**4,362,348 Kč**	**11,217,804 Kč**	**15,580,152 Kč**

Average account balance

age (groups)		Default	Late payment	Repaid Full	Timely payment	Total
74 and above - THE GREATEST GENERATION	25,166.24					**25,166.24**
55- 73 THE SILENT GENERATION	29,160.94	30,149.49	36,072.12	37,522.70	44,825.41	**30,512.77**
36 -54 Baby Boomers	36,980.25	36,403.88	29,206.56	41,022.13	42,033.75	**38,083.00**
20 - 35 Gen X	37,397.51	26,515.27	30,345.15	43,992.02	42,782.65	**38,606.65**
0 - 20 Gen Y	36,462.94		18,931.78	35,040.01	45,121.63	**37,492.99**
Total	**32,430.09**	**31,469.73**	**30,391.87**	**41,539.49**	**42,638.09**	**34,060.66**

CONCLUSION

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

1. https://marketingartfully.com/customer-demographics-age-ranges-numbers/
2. https://www.feedough.com/what-is-marketing-strategy-examples-components-planning/

LINK

1. https://marketingartfully.com/customer-demographics-age-ranges-numbers