Tech Saksham

Case Study Report

Data Analytics with Power BI

“REAL TIME ANALYSIS OF BANK CUSTOMERS”

“Dr.ZAKIR HUSAIN COLLEGE OF ART’S AND SCIENCE”

NM ID	NAME
1CF78E5F01CEA878C5147EB66748219C	MOHAMED KAHLEEF.P

Trainer Name: S.KRISHNAMOORTHY

Master Name:S.KRISHNAMOORTHY

ABSTRACT

In the dynamically evolving landscape of banking, understanding customer behavior in real-time is paramount for providing personalized services, detecting fraud, and optimizing operational efficiency. This paper proposes a comprehensive framework for real-time analysis of bank customer behavior leveraging advanced data analytics techniques. The framework encompasses data collection, processing, analysis, and action, enabling banks to adapt swiftly to changing customer needs and market trends.

INDEX

CHAPTER 1 INTRODUCTION

In the rapidly evolving landscape of banking and finance, the ability to understand and respond to customer needs in real-time has become a strategic imperative for financial institutions. With the advent of digital technologies and the proliferation of data, banks are increasingly leveraging real-time analysis of customer behavior to enhance service delivery, detect fraudulent activities, and optimize operational efficiency. This introduction sets the stage for exploring the significance and implications of real-time analysis of bank customers, outlining its key objectives, challenges, and potential benefits.

Today, customers expect seamless, personalized experiences from their banks, whether they're conducting transactions, seeking financial advice, or managing their accounts online. Traditional banking models, with batch processing and static customer segmentation, are no longer sufficient to meet these expectations. Instead, banks are turning to real-time analysis to gain actionable insights into customer behavior, preferences, and needs as they occur.

Real-time analysis of bank customers involves the continuous monitoring and analysis of various data sources, including transactional data, social media interactions, website browsing behavior, and mobile app usage. By harnessing advanced analytics techniques such as machine learning, predictive modeling, and natural language processing, banks can distill vast amounts of data into meaningful insights in real-time.

The primary objectives of real-time analysis in banking include:

.

Enhanced Customer Experience: By understanding individual customer preferences and behavior in real-time, banks can offer personalized recommendations, targeted promotions, and proactive support, thereby improving overall customer satisfaction and loyalty.

.
.

Fraud Detection and Prevention: Real-time analysis enables banks to detect fraudulent activities as they occur, rather than after the fact. By monitoring transaction patterns, user behavior, and other indicators in real-time, banks can swiftly identify activities, minimizing financial losses and protecting customers' assets.

.
.

Operational Efficiency: Real-time insights empower banks to optimize internal processes, resource and decision-making. By automating routine tasks, streamlining workflows, and identifying bottlenecks in real-time, banks can improve efficiency and reduce cost

.

Despite its potential benefits, real-time analysis of bank customers also poses several challenges. These include managing large volumes of data in real-time, ensuring data privacy and security, integrating disparate data sources, and deploying sophisticated analytics models in a rapidly changing environment.

In conclusion, real-time analysis of bank customers represents a transformative opportunity for financial institutions to stay competitive, drive innovation, and deliver superior value to

their customers. By embracing real-time analytics, banks can unlock new insights, mitigate risks, and capitalize on emerging opportunities in the digital age of banking.

CHAPTER 2

Real-Time Data Collection and Processing Techniques

Real-time data collection and processing are fundamental components of any real-time analysis system. This chapter delves into the methodologies, technologies, and best practices involved in collecting and processing customer data streams in real-time within the banking environment.

.

Data Collection Methods:

.

- Online Banking Platforms: Banks collect real-time data from online banking platforms where customers conduct various transactions, such as fund transfers, bill payments, and account inquiries.
- Mobile Apps: With the widespread use of mobile banking apps, banks capture real-time data from customer interactions on mobile devices, including transactions, balance inquiries, and location data.
- ATMs: Data generated from ATM transactions, including withdrawals, deposits, and balance inquiries, provides valuable insights into customer behavior and preferences.
- Branch Interactions: Customer interactions at bank branches, such as account openings, loan applications, and consultations, generate real-time data that can be captured for analysis.

.

Data Collection Technologies:

- Application Programming Interfaces (APIs): Banks use APIs to integrate with various data sources and systems, enabling real-time data access and retrieval.

- Sensors and IoT Devices: Sensors and Internet of Things (IoT) devices installed in ATMs, branches, and other banking infrastructure capture real-time data on customer interactions, foot traffic, and environmental conditions.
- Transaction Logs: Transaction logs record details of customer transactions, including timestamps, transaction types, and amounts, providing a rich source of real-time data for analysis.

Data Processing Infrastructure:

- Stream Processing Frameworks: Stream processing frameworks such as Apache Kafka, Apache Flink, and Apache Spark Streaming enable the ingestion, processing, and analysis of data streams in real-time.
- Distributed Computing: Banks deploy distributed computing technologies to handle the high volume and velocity of real-time data streams efficiently, ensuring scalability and fault tolerance.
- In-Memory Computing: In-memory computing platforms leverage fast memory storage to process real-time data streams with low latency, enabling rapid analysis and decision-making.

Data Quality and Governance:

- Data Quality Assurance: Banks implement data quality checks and validation rules to ensure the accuracy, completeness, and consistency of real-time data.

- Data Governance Frameworks: Robust data governance frameworks define policies, procedures, and standards for managing and securing real-time data, ensuring compliance with regulatory requirements and industry best practices.

CHAPTER 3

Applications of Real-Time Analytics for Customer Insights

Chapter 3: Applications of Real-Time Analytics for Customer Insights

Introduction:

Real-time analytics empowers banks to gain immediate insights into customer behavior, preferences, and needs. This chapter explores the practical applications of real-time analytics for deriving actionable customer insights in the banking sector.

1. Customer Segmentation:

- Definition: Customer segmentation involves categorizing customers into distinct groups based on shared characteristics, behaviors, or preferences.
- Real-Time Analysis: Real-time analytics enables banks to dynamically segment customers in real-time based on transactional data, demographic information, browsing behavior, and other contextual factors.
- Applications: Banks can use real-time customer segmentation to personalize marketing messages, tailor product offerings, and optimize service delivery for different customer segments.

2. Personalized Marketing and Recommendations:

- Definition: Personalized marketing involves delivering targeted messages and offers to individual customers based on their preferences, interests, and past interactions.
- Real-Time Analysis: Real-time analytics allows banks to analyze customer data streams in real-time and generate personalized recommendations and offers instantaneously.

- Applications: Banks can leverage real-time analytics to deliver personalized marketing emails, SMS messages, and in-app notifications, promoting relevant products, services, and promotions to customers based on their transaction history, browsing behavior, and preferences.

3. Fraud Detection and Prevention:

- Definition: Fraud detection and prevention involve identifying and mitigating fraudulent activities in real-time to protect customers and minimize financial losses.
- Real-Time Analysis: Real-time analytics enables banks to monitor transactional data streams in real-time and detect suspicious patterns, anomalies, and deviations from normal behavior.
- Applications: Banks can deploy real-time fraud detection models that use machine learning algorithms to analyze transaction data in real-time and flag potentially fraudulent transactions for further investigation or intervention.

4. Customer Journey Analysis:

- Definition: Customer journey analysis involves tracking and analyzing customer interactions across multiple touchpoints and channels to understand their end-to-end journey and identify areas for improvement.
- Real-Time Analysis: Real-time analytics enables banks to capture and analyze customer interactions in real-time, providing insights into the effectiveness of marketing campaigns, website usability, and customer service interactions.
- Applications: Banks can use real-time customer journey analysis to optimize the customer experience, identify pain points, and implement targeted interventions to improve satisfaction and loyalty.

5. Operational Optimization:

- Definition: Operational optimization involves streamlining processes, allocating resources, and improving efficiency to deliver superior servic/e and reduce costs.

- Real-Time Analysis: Real-time analytics enables banks to monitor operational data streams in real-time and identify opportunities for optimization, such as reducing wait times, improving resource utilization, and automating manual processes.

CHAPTER 4 MODELING

Modeling for real-time analysis of bank customers involves leveraging data analytics and machine learning techniques to gain insights into customer behavior, preferences, and patterns in real-time. Here's a general outline of steps you might take:

.

Data Collection: Gather data from various sources within the bank's ecosystem. This could include transaction data, customer demographics, account information, website interactions, and even data from external sources like social media or economic indicators.

.
.

Data Preprocessing: Cleanse and preprocess the data to handle missing values, outliers, and inconsistencies. This step is crucial for ensuring the quality of the data before analysis.

.
.

Feature Engineering: Extract relevant features from the data that can be used to behavior. characterize customer These features could include transaction frequency, amount, account balances, demographics, etc.

.
.

Real-time Data Streaming: Implement mechanisms for ingesting real-time data streams from various sources. Technologies like Apache Kafka or Apache Flink could be used for this purpose.

.

loan_id	account_id	date	Loan Amt	duration	payments	status	Credit Rating	Loan Status	Count of all loan
6059	5196	971228	79,824 Kč	12	6652	A	GOOD	Repaid Full	682
6727	8505	971210	42,840 Kč	12	3570	A	GOOD	Repaid Full	682

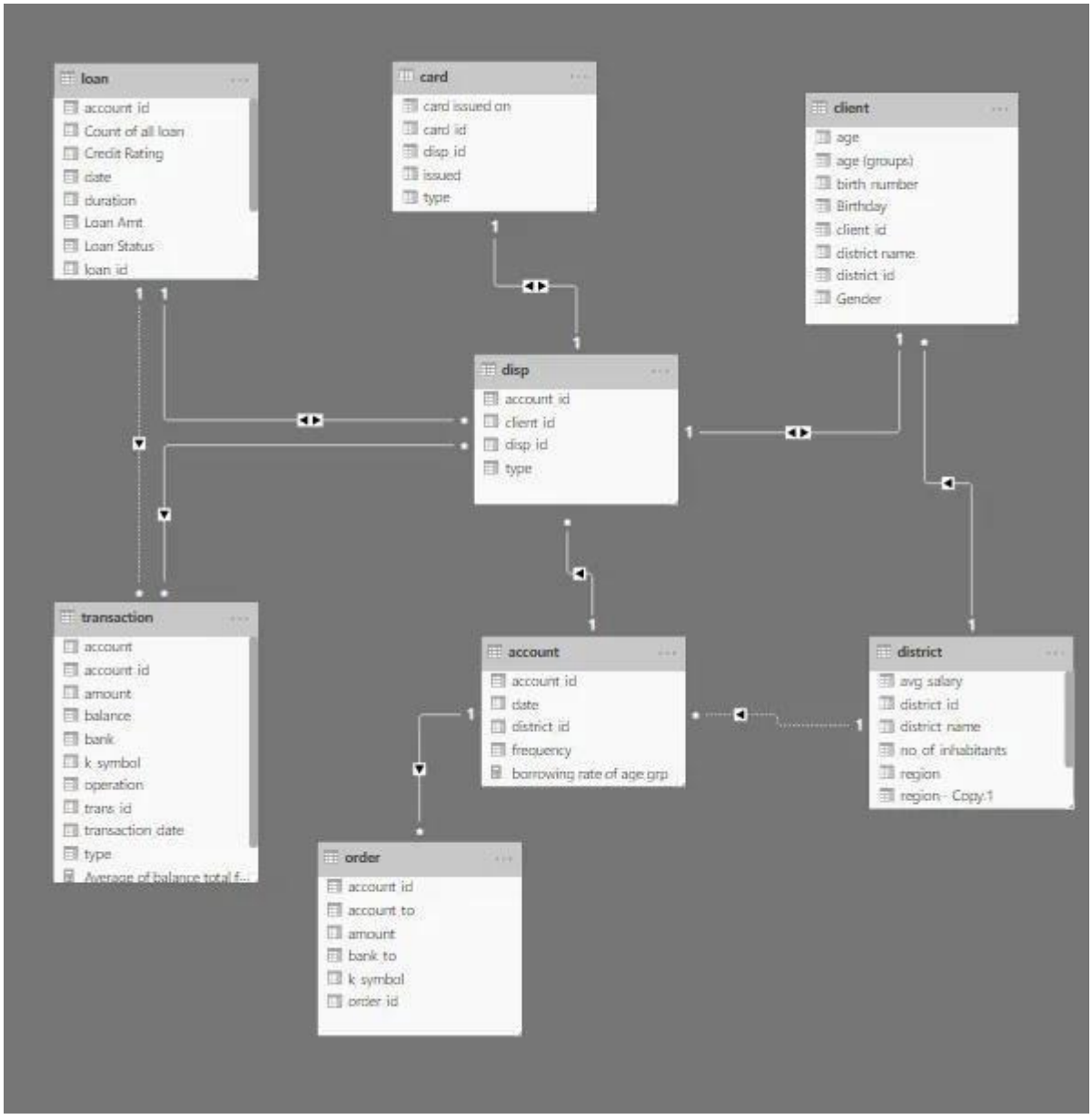
And a 12 months loan issued in Dec 1997 has been repaid in full.

loan_id	account_id	date	Loan Amt	duration	payments	status	Credit Rating	Loan Status	Count of all loan
6059	5196	971228	79,824 Kč	12	6652	A	GOOD	Repaid Full	682
6727	8505	971210	42,840 Kč	12	3570	A	GOOD	Repaid Full	682

Data Modelling

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”



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)

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

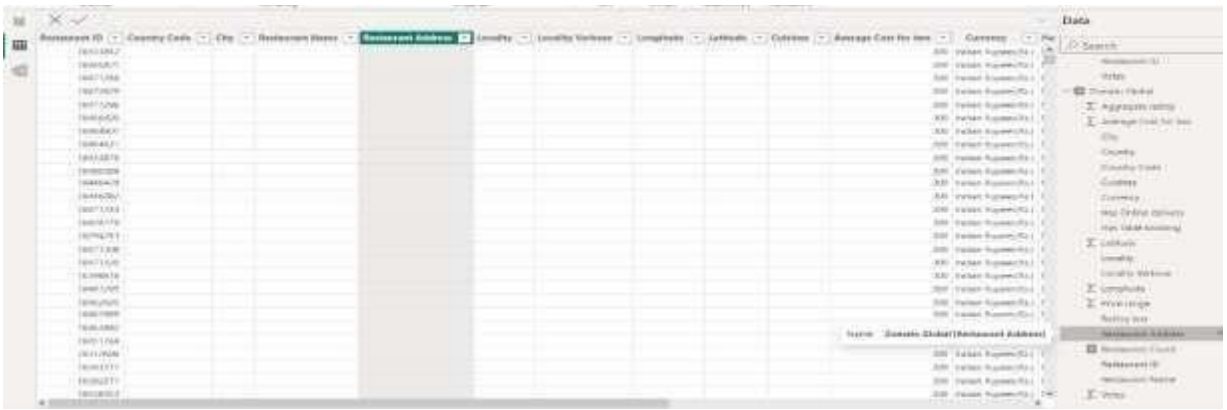
Cardinality: One to one (1:1)

Cross filter direction: Both

☑ Make this relationship active

☐ Apply security filter in both directions

☐ Assume referential integrity



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.

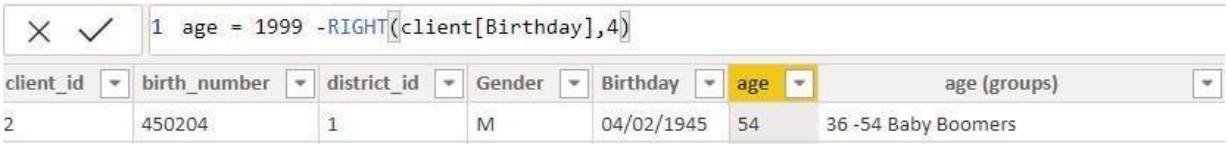
And District name have been categorized as place to be use for the map to show the sum of the inhabitants in each region.

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.

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.

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

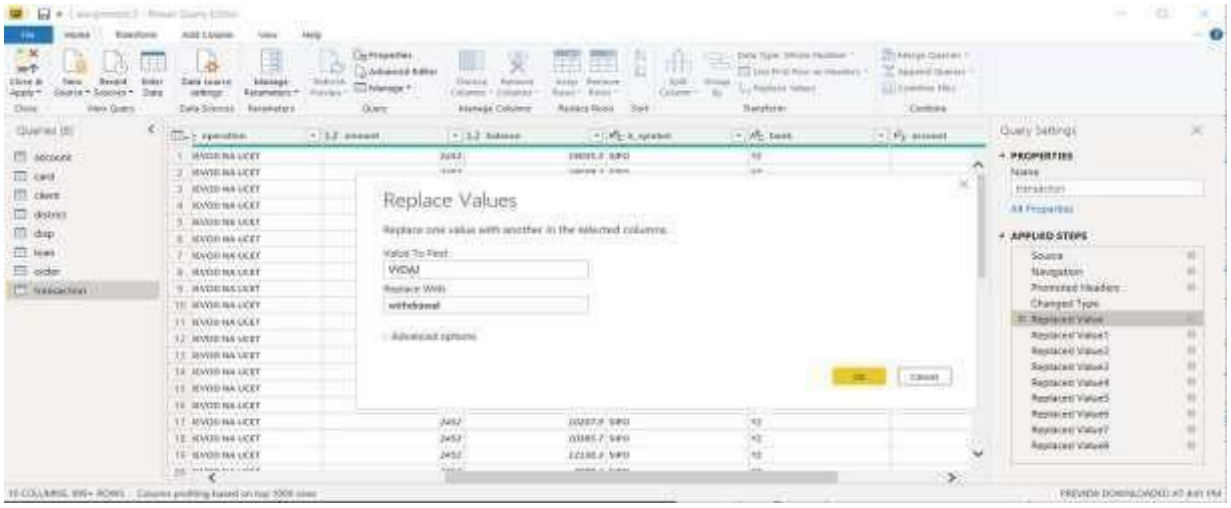


```
1  age = 1999 -RIGHT(client[Birthday],4)
```

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

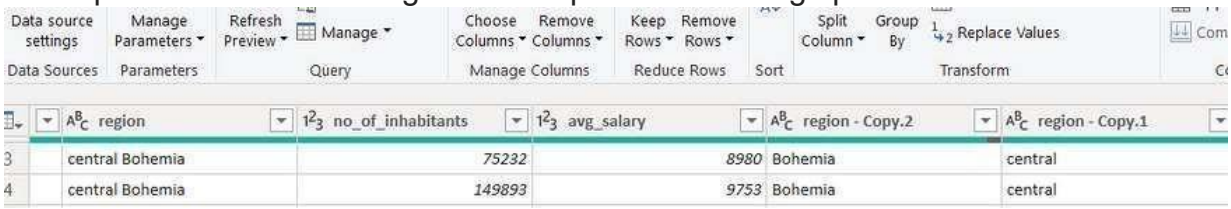
Replacing values

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



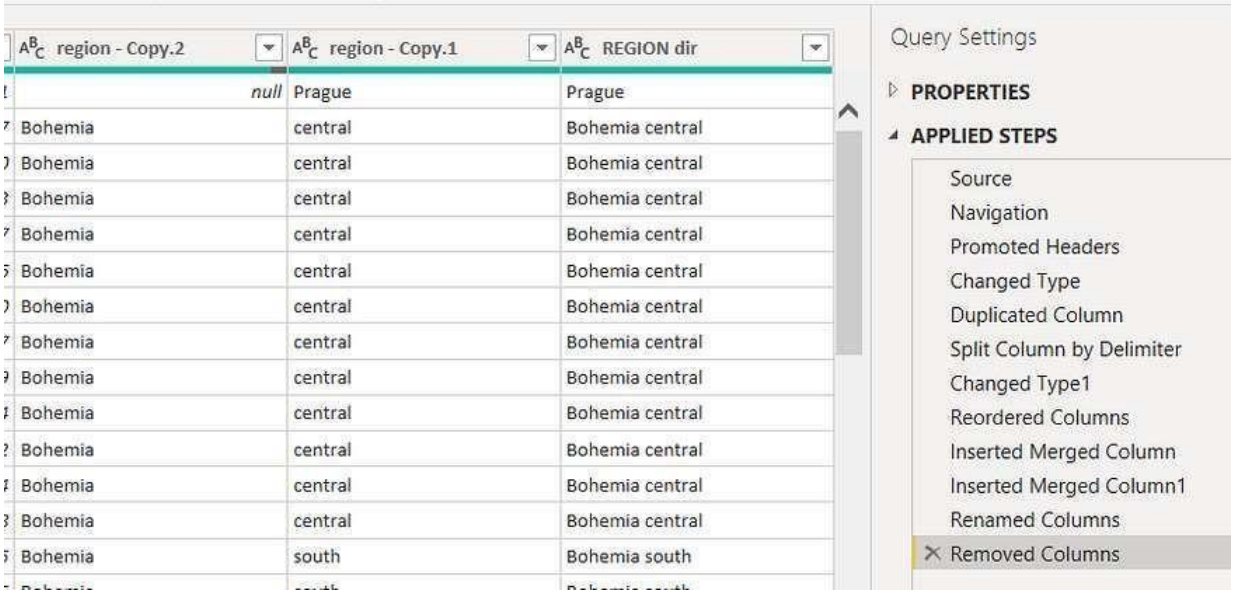
Changing the order of Region name at Power Query

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



	region	no_of_inhabitants	avg_salary	region - Copy.2	region - Copy.1
3	central Bohemia	75232	8980	Bohemia	central
4	central Bohemia	149893	9753	Bohemia	central

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.

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.

CONCLUSION

Conducting real-time analysis of bank customers provides invaluable insights that can enhance customer experiences, optimize operations, and mitigate risks. By harnessing advanced analytics and cutting-edge technology, banks can promptly respond to customer needs, detect fraudulent activities swiftly, and personalize services effectively. The ability to analyze customer data in real-time enables banks to offer tailored products and services, anticipate market trends, and enhance competitive advantage. Moreover, real- time analysis facilitates proactive decision-making, ensuring timely interventions and improving overall business performance. As technology continues to evolve, investing in real-time analysis capabilities remains imperative for banks to stay agile, relevant, and resilient in an ever-changing financial landscape.

By leveraging real-time data insights, the bank has achieved competitive advantages, driving growth and sustainability in the ever-evolving banking landscape. This case study underscores the transformative potential of real-time analysis in the banking sector and highlights the importance of leveraging advanced analytics to stay ahead in today's digital era.

FUTURE SCOPE

The future scope of real-time analysis in the banking sector is vast, with ongoing advancements in technology and data analytics opening up new possibilities for enhancing customer experiences, optimizing operations, and managing risks. Here are some potential areas of future development:

1. **AI-Powered Personalization**: As artificial intelligence (AI) and machine learning algorithms continue to evolve, banks can leverage real-time analysis to offer highly personalized banking experiences. By analyzing vast amounts of customer data in real-time, banks can anticipate individual needs, preferences, and financial goals, allowing for tailored product recommendations, customized offers, and proactive financial advice.

2. **Predictive Analytics for Fraud Detection**: Real-time analysis combined with predictive analytics holds great promise for improving fraud detection and prevention in banking. By continuously monitoring transactional data and applying advanced anomaly detection techniques, banks can identify suspicious patterns and potential fraud in real-time, enabling swift intervention to mitigate losses and protect customer assets.

3. **Real-Time Credit Scoring and Risk Management**: Traditional credit scoring models are often based on historical data and may not capture real-time changes in customer creditworthiness. Real-time analysis offers the opportunity to enhance credit scoring models by incorporating dynamic data streams such as transaction history, social media activity, and even real-time income verification. This enables banks to make more accurate and timely lending decisions while effectively managing credit risk.

4. **Instant Financial Insights and Alerts**: With the rise of mobile banking and digital channels, there is growing demand for instant access to financial insights and alerts. Real-time analysis can power mobile banking apps and digital platforms to provide

customers with real-time updates on account balances, spending patterns, and budgeting tips. Additionally, real-time alerts can notify customers of unusual account activity, payment due dates, or potential overdrafts, empowering them to make informed financial decisions in real-time.

5. **Continuous Compliance Monitoring**: Regulatory compliance is a critical aspect of banking operations, requiring banks to monitor transactions and activities for compliance with anti-money laundering (AML) and know your customer (KYC) regulations. Real-time analysis can automate compliance monitoring processes by flagging suspicious transactions, monitoring customer profiles for changes, and generating real-time reports for regulatory authorities. This helps banks streamline compliance efforts, reduce compliance risks, and avoid costly penalties.

Overall, the future of real-time analysis in banking holds immense potential for transforming customer experiences, enhancing operational efficiency, and managing risks in an increasingly digital and data-driven environment. By embracing emerging technologies and leveraging real-time data analytics, banks can stay ahead of the curve and deliver value-added services that meet the evolving needs of customers and regulators alike.

REFERENCES

1.https://latinia.com/en/resources/real-time-banking-analytics

2.https://www.datatobiz.com/blog/customer-analytics-in-banking/

LINK

https://github.com/kahleef/kahleefcasestudy

