ANALYSIS OF UNEMPLOYMENT

IN REBUBLIC OF INDIA USING

DATA ANALYTICS WITH POWER BI

College Name

GOVERNMENT ARTS&SCIENCE COLLEGE

VEDARANYAM

NM ID	NAME
asbdu369cb21s206510	M.prabu

Trainer name:

Uma   maheshwari

ABSTRACT

Today, enterprises are looking for innovative ways to digitally transform their businesses - a crucial step forward to remain competitive and enhance profitability. There are key technology enablers that support an enterprise's digital transformation efforts, including analytics. Real-time insights and data in motion via analytics helps organizations to gain the business intelligence they need for digital transformation. From a business perspective, the potential benefits it can offer an organization are many - you can use location and contextual data to create better customer experiences; create radically new databased products for your business; make more informed decisions in complex scenarios; carry out effective monitoring and analysis;detect even the smallest change and trigger immediate action; and extend your solutions to analyze thepast, present, and the future. While these benefits are applicable to most organizations across diverse industries, a key advantage of analytics is that it can be customized to create solutions to meet the specific requirements of a particular industry.This white paper will focus on the business benefits extended to the banking & finance industry and discuss some common use cases within this domain.

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, crossselling, 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 realtime 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

Algorithms such as Clustering help a computer program to model 'normal' behavior by looking at past transaction trends. Therefore, this helps banks to identify new types of fraud by looking for transactions that differ from the normal behaviour that the machine learning algorithm has modelled.

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

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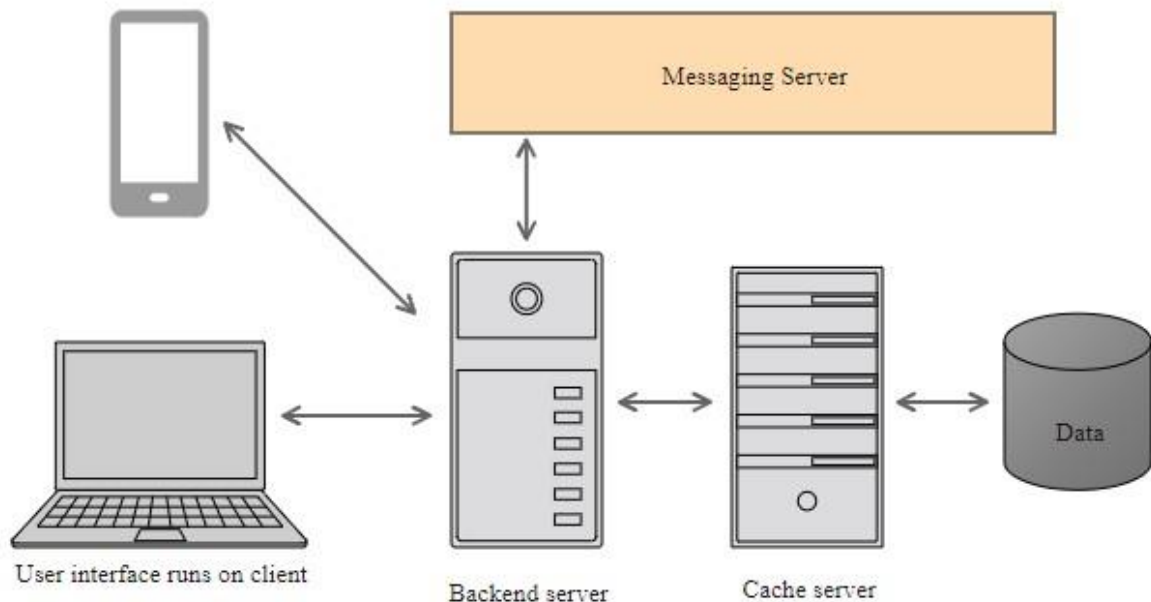
Here's a high-level architecture for the project:

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"
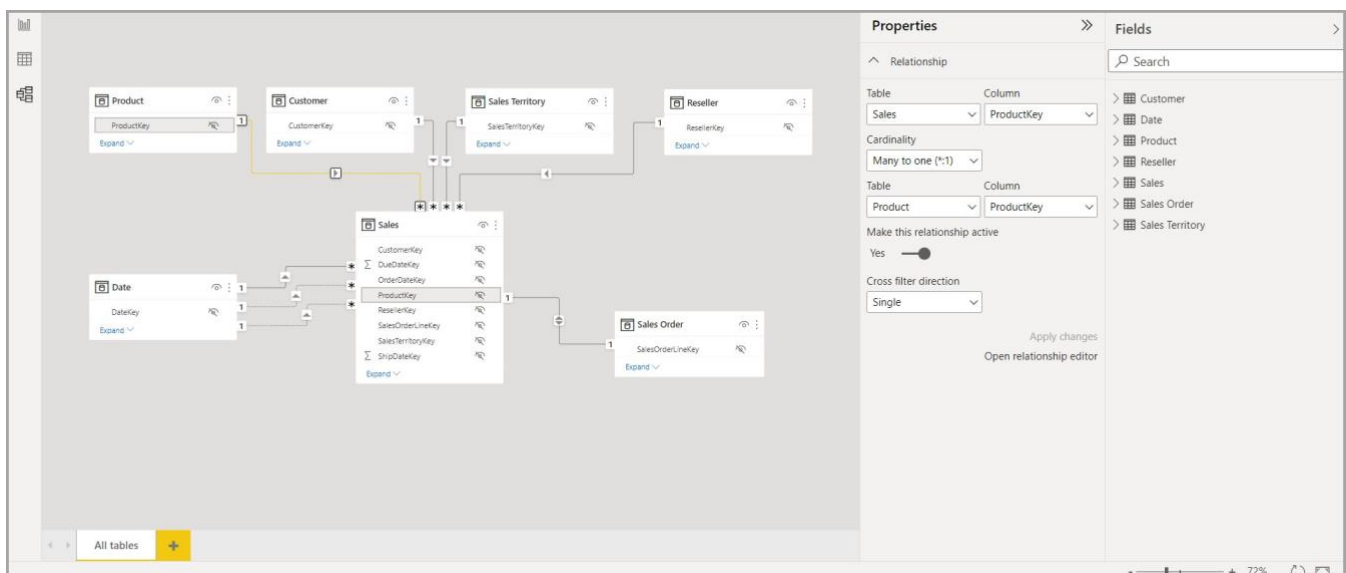
Manage relationships

Active	From: Table (Column)	To: Table (Column)
☑	EmployeeRole (Employee)	ProjectTickets (SubmittedBy)
☐	ProjectTickets (OpenedBy)	EmployeeRole (Employee)

New...   Autodetect...   Edit...   Delete

Close

Edit relationship

Select tables and columns that are related.

Sales ▾

SalesOrderLineKey	ResellerKey	CustomerKey	ProductKey	OrderDateKey	DueDateKey	ShipDateKey
46638001	203	-1	333	20180718	20180728	20180725
46638002	203	-1	325	20180718	20180728	20180725
46642010	4	-1	321	20180720	20180730	20180727

‹ ▬▬▬▬▬▬▬ ›

Product ▾

ProductKey	Product	Standard Cost	Color	List Price	Model	Subcategory	Catego
210	HL Road Frame - Black, 58	$868.63	Black	$1,431.50	HL Road Frame	Road Frames	Compo
215	Sport-100 Helmet, Black	$12.03	Black	$33.64	Sport-100	Helmets	Access
216	Sport-100 Helmet, Black	$13.88	Black	$33.64	Sport-100	Helmets	Access

‹ ▬▬▬▬▬▬▬▬▬▬▬▬▬ ›

Cardinality

Many to one (*:1) ▾

☑ Make this relationship active

☐ Assume referential integrity

Cross filter direction

Single ▾

☐ Apply security filter in both directions

[OK] [Cancel]

Create relationship

Select tables and columns that are related.

ProjectHours

Ticket	SubmittedBy	Hours	Project	DateSubmit
1001	Brewer, Alan	22	Blue	Tuesday, January 1, 2013
1002	Brewer, Alan	26	Red	Friday, February 1, 2013
1003	Ito, Shu	34	Yellow	Tuesday, December 4, 2012

CompanyProject

ProjName	Priority
Blue	A
Red	B
Green	C

Cardinality

Many to one (*:1)

Cross filter direction

Single

☑ Make this relationship active

☐ Apply security filter in both directions

☐ Assume referential integrity

OK   Cancel

n this report, we have performed an analysis of the Unemployment Rate in European Union over the last 10 years using features such as advanced conditional formatting - Segmentation and Heat Maps, Sorting, Variance Analysis and many more

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

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.

Priority	Hours
A	256
B	256
C	256
Total	**256**

Visualizations

Fields

Search

∧ ⊞✓ CompanyProject

  ✓ Priority

  ☐ ProjName

∧ ⊞✓ ProjectHours

  ∨ ☐ 📅 DateSubmit

  ✓ Σ Hours

  ☐ Project

  ☐ SubmittedBy

  ☐ Σ Ticket

Create relationship

Select tables and columns that are related.

ProjectHours ▾

Ticket	SubmittedBy	Hours	Project	DateSubmit
1001	Brewer, Alan	22	Blue	Tuesday, January 1, 2013
1002	Brewer, Alan	26	Red	Friday, February 1, 2013
1003	Ito, Shu	34	Yellow	Tuesday, December 4, 2012

CompanyProject ▾

ProjName	Priority
Blue	A
Red	B
Green	C

Cardinality   Cross filter direction

Many to one (*:1) ▾   Single ▾

✓ Make this relationship active   ☐ Apply security filter in both directions

☐ Assume referential integrity

OK   Cancel

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.

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

(a) (b) (c) (d)

(e) (f) (g) (h)

Retail Banking

⋮ ✕

Date Filter
Jul 1 2020 : Jul 31 2021 ⌄

Deposits vs Goal

167,279

+378% ▲
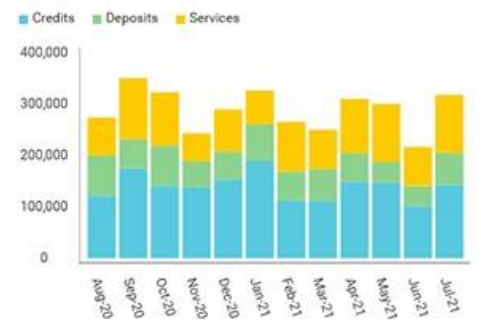vs deposits goal / Month to Date

Deposits vs Goal

■ deposits — deposits goal

(bar chart: Aug-20 through Jul-21, left axis 0–4,000,000, right axis 0–40,000)

Deposits vs Goal

— accounts opened — accounts closed

(line chart: Aug-20 through Jul-21, axis 0–400,000)

Opened Account by Referral

● Credits ● Deposits ● Services

(donut chart: 48.6%, 31%, 20.3%)

Opened Account by Referral

■ Credits ■ Deposits ■ Services

(stacked bar chart: Aug-20 through Jul-21, axis 0–400,000)

↗
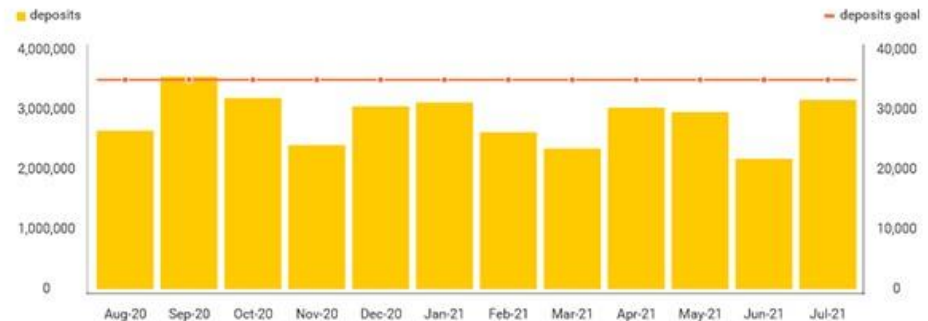
Credit Rating
☐ (Blank)
☐ BAD
☐ GOOD

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%**

Sum of Loan Amt by Credit Rating and Loan Status

Loan Status
● Default
● Late payment
● Repaid Full
● Timely payment

(bar chart: GOOD 69M Kč / 19M, BAD 11M Kč / 4M, Credit Rating)

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

682
Total Loan

Total Loan Amount:
103 Mil Kc

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**

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

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

Today, enterprises are looking for innovative ways to digitally transform their businesses - a crucial step forward to remain competitive and enhance profitability. There are key technology enablers that support an enterprise's digital transformation efforts, including analytics. Real-time insights and data in motion via analytics helps organizations to gain the business intelligence they need for digital transformation. From a business perspective, the potential benefits it can offer an organization are many - you can use location and contextual data to create better customer experiences; create radically new databased products for your business; make more informed decisions in complex scenarios; carry out effective monitoring and analysis;detect even the smallest change and trigger immediate action; and extend your solutions to analyze thepast, present, and the future. While these benefits are applicable to most organizations across diverse industries, a key advantage of analytics is that it can be customized to create solutions to meet the specific requirements of a particular industry.This white paper will focus on the business benefits extended to the banking & finance industry and discuss some common use cases within this domain.

FUTURE SCOPE

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, crossselling, 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.

REFERENCES

https://github.com/Prabu-2005/Prabu-2005