

ANALYSIS OF UNEMPLOYMENT IN REBUBLIC OF INDIA USING DATA ANALYTICS WITH POWER BI

College Name

GOVERNMENT ARTS AND SCIENCE COLLEGE

NM ID	NAME
asbdu369cb21s206518	C.Suganya

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 the past, 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

Sr. No.	Table of Contents	Page No.
1	Chapter 1: Introduction	4
2	Chapter 2: Services and Tools Required	6
3	Chapter 3: Project Architecture	8
4	Chapter 4: Modeling and Result	10
5	Conclusion	19
6	Future Scope	20
7	References	21
8	Links	22

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, cross-selling, 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 real-time 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.

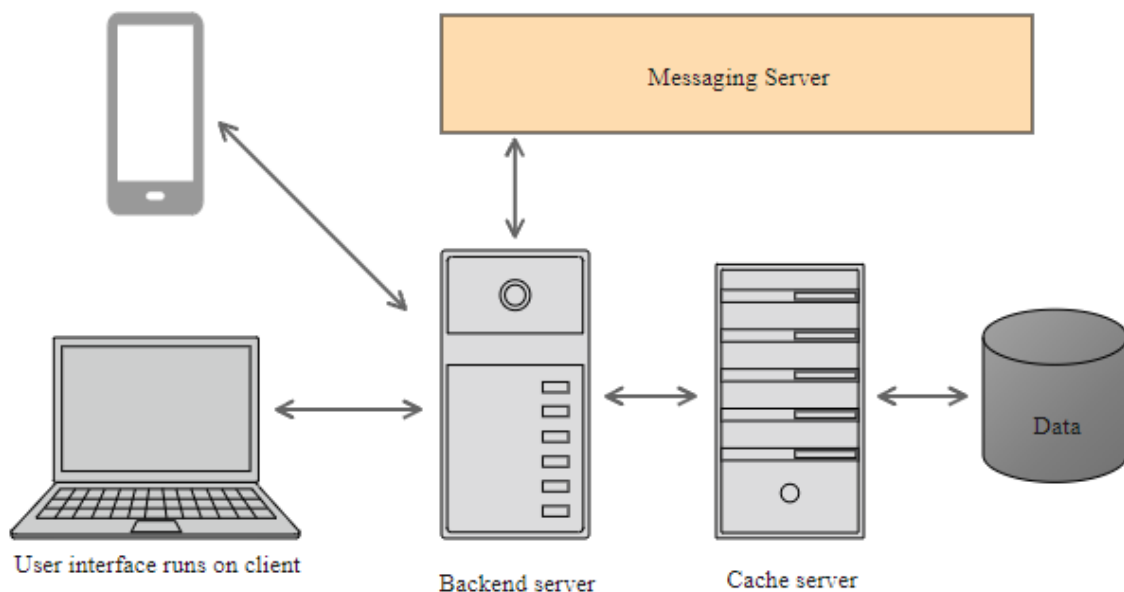
- **PowerBI Service:** This is an online SaaS (Software as a Service) service that you use to publish reports, create new dashboards, and share insights.
- **PowerBI Mobile:** This is a mobile application that you can use to access your reports and dashboards on the go.

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.

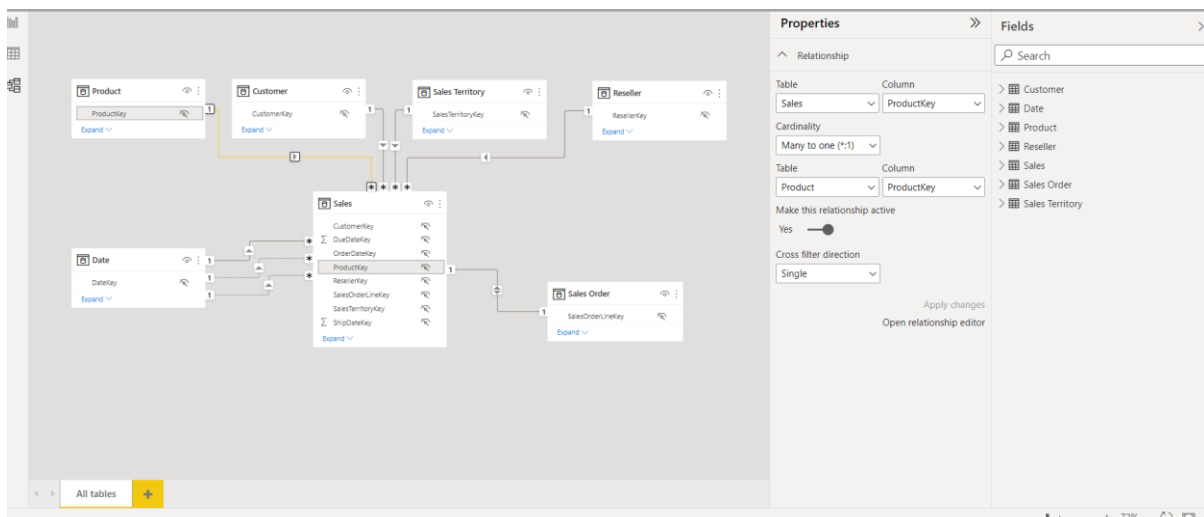
This architecture provides a comprehensive solution for real-time analysis of bank customers. However, it's important to note that the specific architecture may vary depending on the bank's existing infrastructure, specific requirements, and budget. It's also important to ensure that all tools and services 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)
<input checked="" type="checkbox"/>	EmployeeRole (Employee)	ProjectTickets (SubmittedBy)
<input type="checkbox"/>	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	Category
210	HL Road Frame - Black, 58	\$868.63	Black	\$1,431.50	HL Road Frame	Road Frames	Components
215	Sport-100 Helmet, Black	\$12.03	Black	\$33.64	Sport-100	Helmets	Accessories
216	Sport-100 Helmet, Black	\$13.88	Black	\$33.64	Sport-100	Helmets	Accessories

Cardinality

Many to one (*:1)

Cross filter direction

Single

☒ Make this relationship active

☐ Apply security filter in both directions

☐ Assume referential integrity

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



2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2019

Home Advanced

Manage Columns



Unemployment rates by educational attainment level and NUTS 2 regions (%)

Level	All Levels		Levels 0-2		Levels 3 - 4		Levels 5 - 8	
	Unemployment	Variance Analysis (Unemployment)	Unemployment	Variance Analysis (Unemployment)	Unemployment	Variance Analysis (Unemployment)	Unemployment	Variance Analysis (Unemployment)
European Union - 27 countries (from 2020)	101.5	0.0%	103.5	0.0%	91.4	0.0%	61.5	0.0%
Madrid, Bataan, Simak, Silit	251.4	147.7%	262.8	43.2%	245.2	168.3%	215.9	247.0%
Var, Mus, Bitlis, Makkari	150.1	56.2%	151.6	17.3%	202.2	121.2%	157.1	155.4%
Ciudad de Melilla	181.9	197.4%	400.7	151.3%	245.2	168.3%	119.2	93.8%
Dytiki Makedonia	209.9	185.6%	267.7	40.0%	320.6	259.5%	249.0	384.9%
Dytiki Ellada	266.6	162.7%	248.3	35.5%	307.0	235.9%	238.0	274.0%
Sarıırmak, Diyarbakir	167.8	65.3%	167.6	8.0%	184.2	101.5%	149.0	142.3%
Ciudad de Ceuta	500.6	205.0%	244.0	142.3%	204.0	123.2%	151.7	114.1%
Mayotte	217.3	114.1%	252.5	54.0%	100.2	97.2%	100.0	100.0%
Güneydogu Anadolu	173.4	70.8%	175.0	4.5%	180.2	97.2%	154.8	150.4%
La Réunion	273.6	169.6%	370.0	100.2%	286.4	213.3%	97.0	59.0%
Extremadura	294.9	190.5%	369.1	101.4%	259.8	184.2%	179.4	191.7%
Andalucia	315.8	211.1%	408.1	122.6%	290.2	217.5%	189.5	208.1%
Canarias	296.5	194.1%	374.2	104.1%	285.0	209.0%	261.3	227.3%
Guadeloupe	250.0	140.3%	339.6	65.3%	259.3	183.5%	109.5	70.0%
Calabria	220.3	117.0%	259.7	41.7%	215.8	156.1%	154.2	150.7%
Kentriki Makedonia	250.7	147.0%	281.0	53.3%	276.7	204.9%	192.2	212.5%
RUP FR - Régions ultrapériphériques françaises	251.0	149.1%	355.6	94.0%	256.9	181.1%	93.4	51.9%
Suz (ES)	304.4	199.9%	393.1	114.5%	282.7	209.3%	100.7	193.0%
Attiki	233.2	129.8%	337.7	84.2%	265.4	190.4%	158.3	157.4%
Vorota Ellada	142.9	74.5%	250.1	40.0%	200.1	100.0%	100.0	100.0%

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.

Priority	Hours
A	256
B	256
C	256
Total	256

Filters

Visualizations

Fields

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

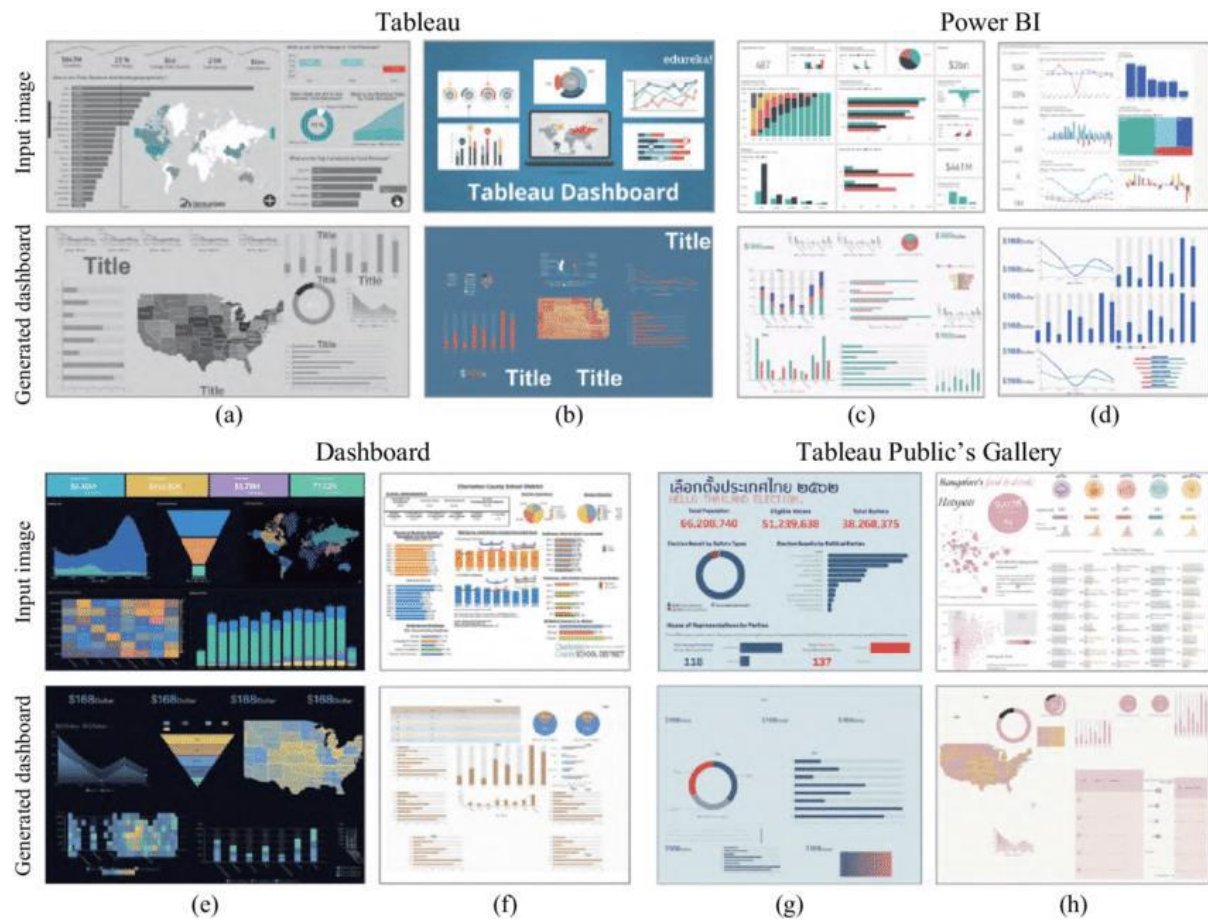
Credit Rating and Loan Status

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





Retail Banking



Date Filter

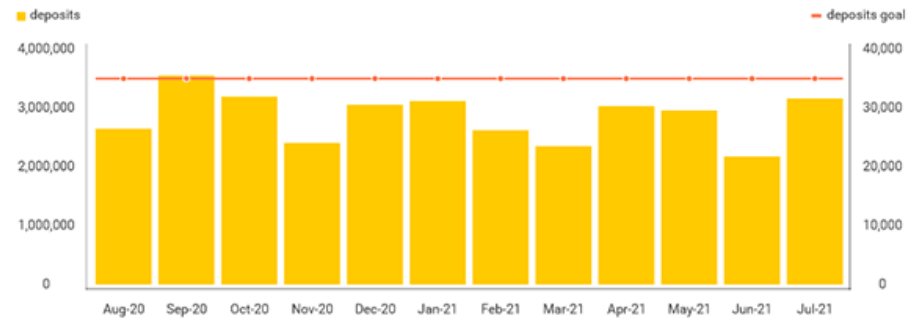
Jul 1 2020 : Jul 31 2021 ▾

Deposits vs Goal

167,279

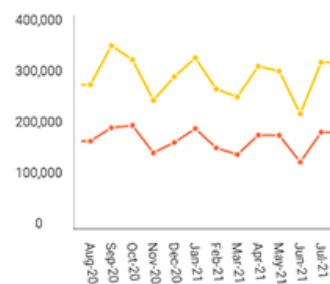
+378%▲
vs deposits goal / Month to Date

Deposits vs Goal



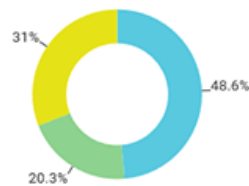
Deposits vs Goal

accounts opened accounts closed



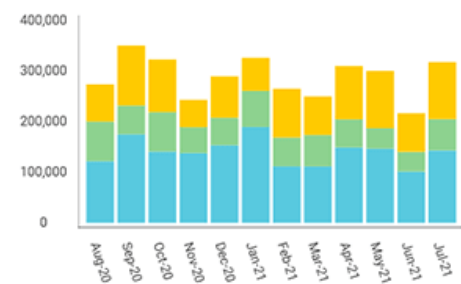
Opened Account by Referral

Credits Deposits Services



Opened Account by Referral

Credits Deposits Services



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 the past, 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, cross-selling, 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://www.softwebsolutions.com/data-analytics-banking-dashboard.html>

LINK