Tech Saksham

Case Study Report

Data Analysis with Power BI

Analysis of Commercial Electricity Consumption in Indian States

A.P.C Mahalaxmi College for Women

NM ID	NAME
9D2D0AE906DB686BC2848978D23275A3	B.SUBHA

R.UMAMAHESWARI	**Trainer Name**
R.UMAMAHESWARI	**Master Trainer**

ABSTRACT

Energy consumption during food processing varies with the product, the degree of processing, the processes involved, and the form of energy used – thermal, electricity, or both. Of the greenhouse gases (GHGs), carbon dioxide, methane, and nitrogen oxides are primarily a consequence of burning fossil fuels and, in some cases because of the intrinsic process, for example, fermentations. Other important emissions are intended or unintended release of hydrofluorocarbons (HFCs) from refrigeration equipment. Because most of the food processing industries are specific to the types of products

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

A problem statement is a concise summary of the user's needs and specifications that must be met. A problem statement brings the organization together around the user problem, providing everyone a clear goal to work toward. A powerful problem statement is one that is focused on people. We have two personas for the qualitative participants' group. So two problem statements are created based on the user's characteristics and the user's insights.

1.2 Proposed Solution

Use an advanced power strip to reduce "vampire loads"--electricity that is wasted when electronics are not in use. Lighting -- Purchase energy-efficient lighting, operate them efficiently, and incorporate more daylighting into your home using energy-efficient windows and skylights.

1.3 Feature

- **Electricity Consumption**: The dashboard will provide the analysis of electricity consumption 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

Electric power has many advantages domestically and industrially, as most of the equipment run by electric power. Brightness in the night is only possible by the use of electricity. Almost all the factories and industries are running due to electric power. The advantage of electric power is its reliable and uninterrupted supply runs the equipment efficiently and continuously. The transportation of electricity is easy once the transmission lines are functional. They work for years and need no or very less maintenance. The invention of electric power is one of the best inventions which have changed human life drastically. It allows people to do more leisure activities.

1.5 Scope

Aggregate electricity demand could grow from 949 TWh in 2015 to between 2074 TWh (low GDP, high efficiency) and 2785 TWh (high GDP, low efficiency), with a mid-value of 2338 TWh (6.2 percent CAGR) by 2030. The big changes in sectoral shares (and therefore growth rates) occur in the commercial and agriculture sectors—commercial likely surpassing agricultural (irrigation pumping) demand in 2030 when it was less than half of the former in 2015.

Industrial and domestic remain the largest consumers, with greater uncertainty (range of possible outcomes) around the later.

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:

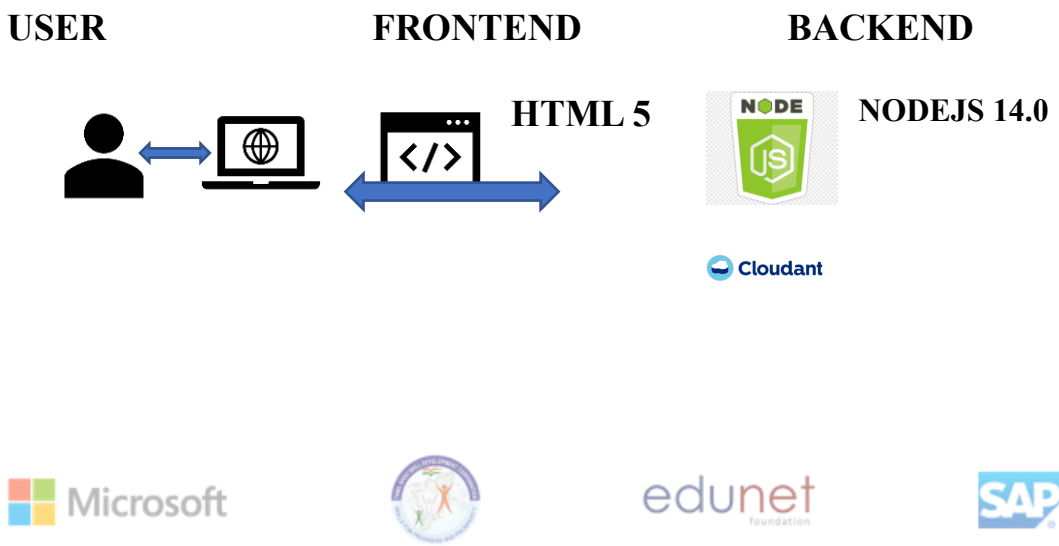
- **PowerBI Desktop**: This is a Windows application that you can use to create reports and publish them to PowerBI.

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

CHAPTER 3

PROJECT ARCHITECTURE

3.1 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.

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

3.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.

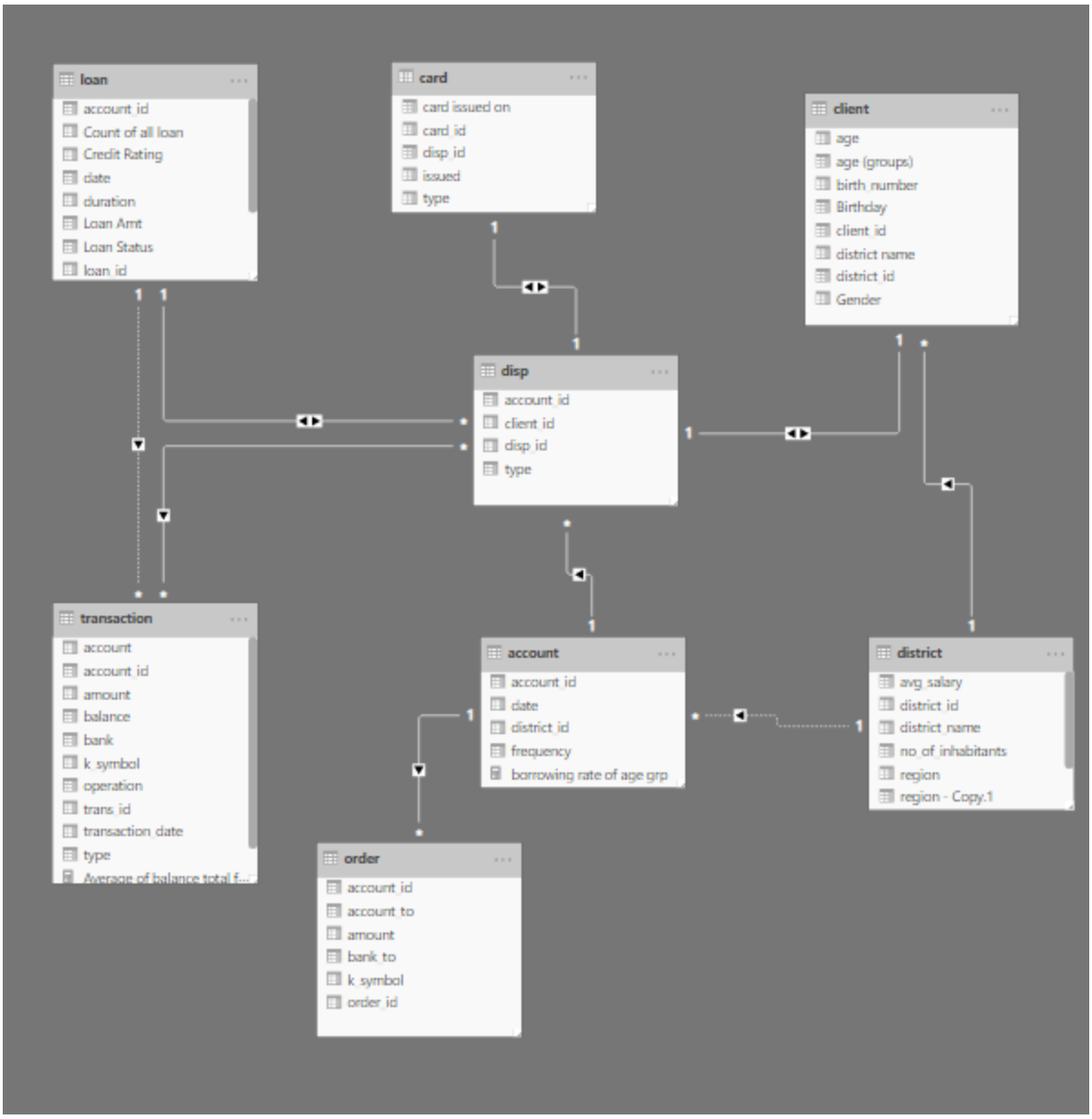
4.Data Visualization: The processed data and the results from the predictive models are visualized in real-time using PowerBI. PowerBI allows you to create interactive dashboards that can provide valuable insights into the data.

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”

Microsoft edunet foundation SAP

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

```
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
```

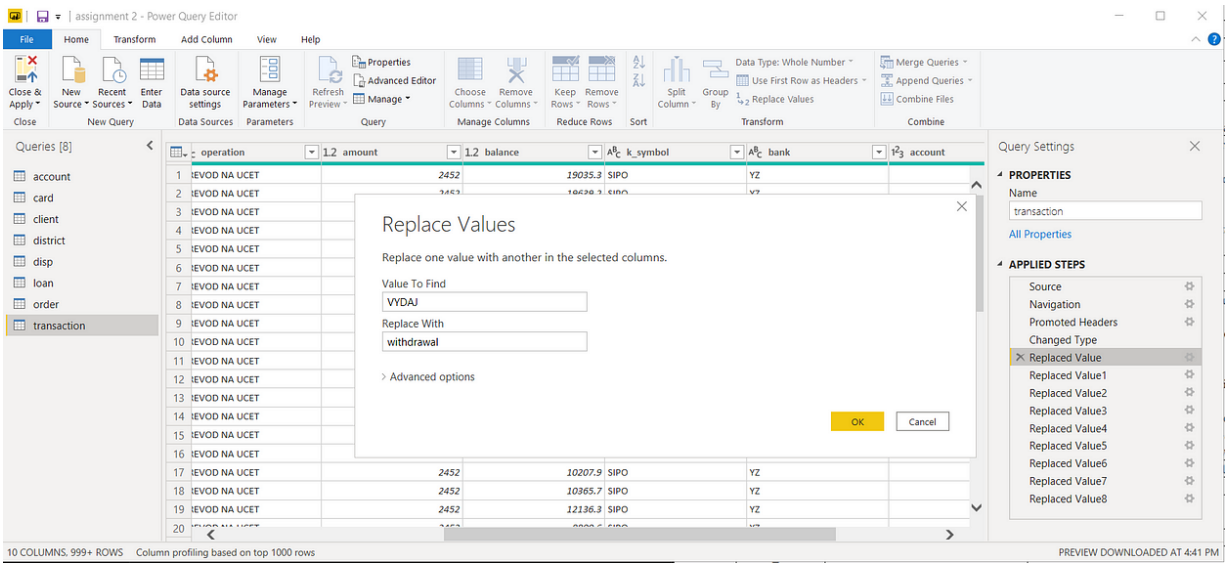
client_id	birth_number	district_id	Gender	Birthday	age
3428	875927	42	F	27/09/1987	13
4354	860813	28	M	13/08/1986	14
3417	855318	35	F	18/03/1985	15
10201	851019	13	M	19/10/1985	15

```
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

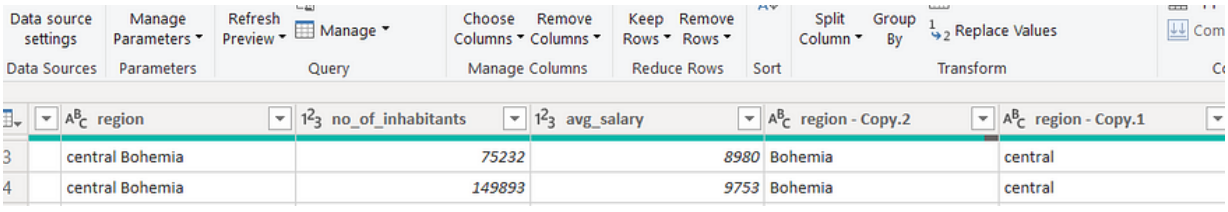
```
1  age = 1999 -RIGHT(client[Birthday],4)
```

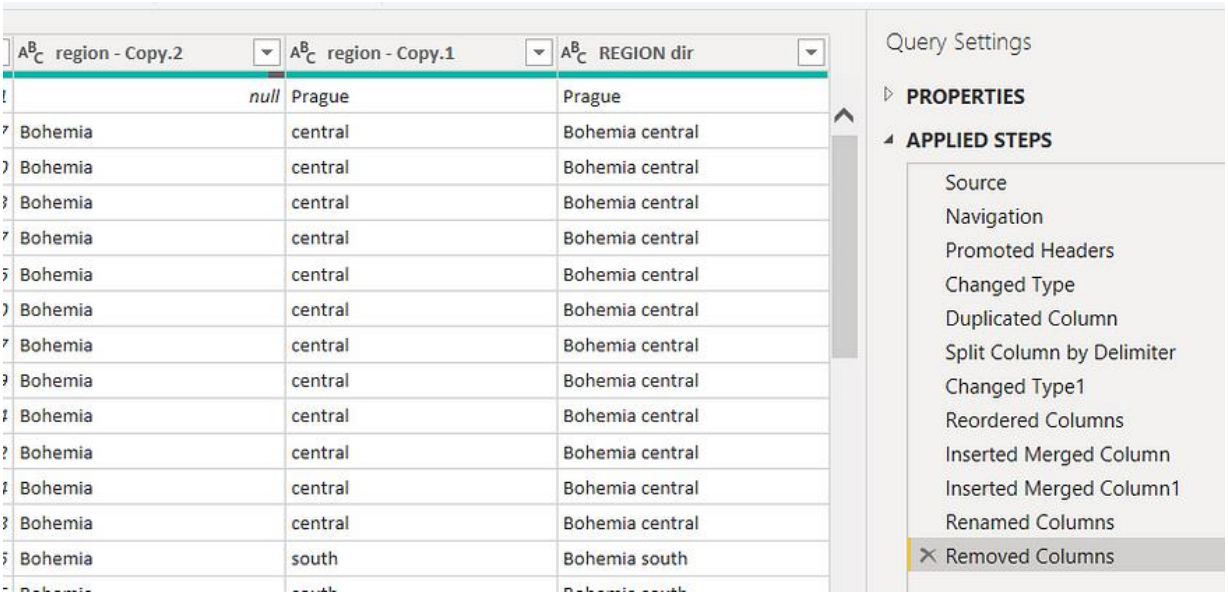
client_id	birth_number	district_id	Gender	Birthday	age	age (groups)
2	450204	1	M	04/02/1945	54	36 -54 Baby Boomers



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





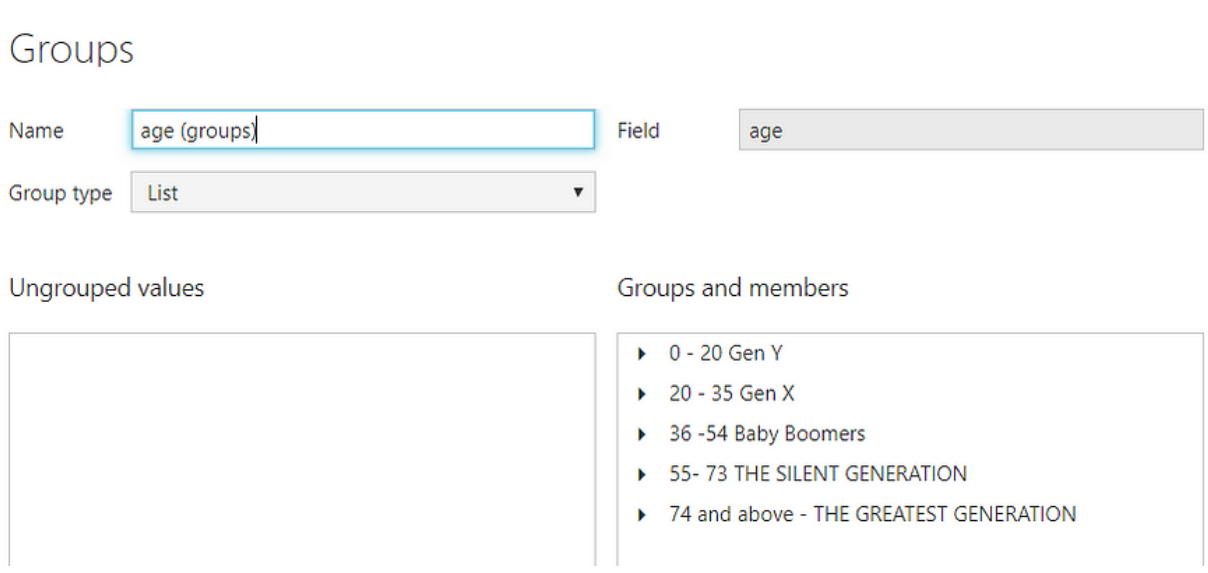
region - Copy.2	region - Copy.1	REGION dir
null	Prague	Prague
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	central	Bohemia central
Bohemia	south	Bohemia south

Query Settings

PROPERTIES

APPLIED STEPS

- Source
- Navigation
- Promoted Headers
- Changed Type
- Duplicated Column
- Split Column by Delimiter
- Changed Type1
- Reordered Columns
- Inserted Merged Column
- Inserted Merged Column1
- Renamed Columns
- Removed Columns



Groups

Name: age (groups)

Field: age

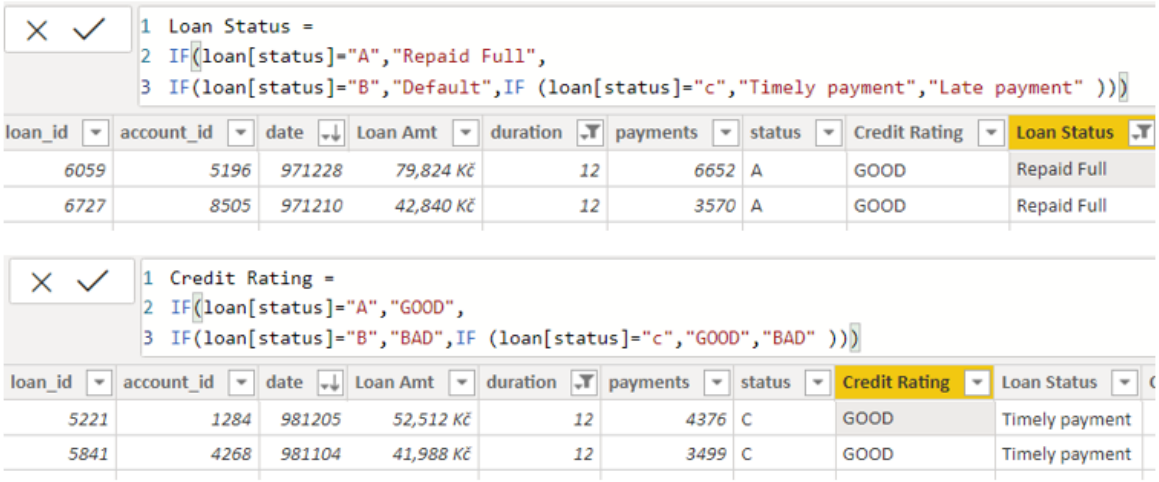
Group type: List

Ungrouped values

Groups and members

- 0 - 20 Gen Y
- 20 - 35 Gen X
- 36 -54 Baby Boomers
- 55- 73 THE SILENT GENERATION
- 74 and above - THE GREATEST GENERATION

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



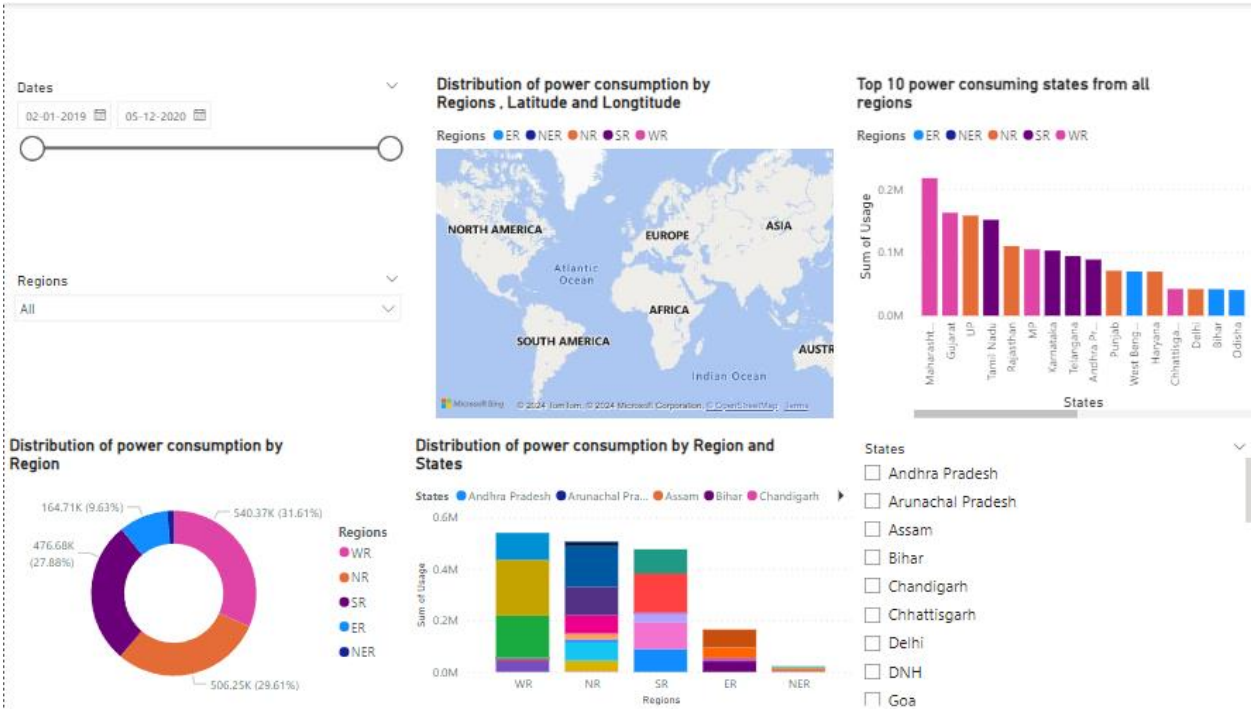
```
1 Loan Status =
2 IF(loan[status]="A","Repaid Full",
3 IF(loan[status]="B","Default",IF (loan[status]="c","Timely payment","Late payment" )))
```

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

Dashboard


Dates
02-01-2019
05-12-2020
Regions
All
Distribution of power consumption by Regions , Latitude and Longtitude
Regions ER NER NR SR WR
NORTH AMERICA
EUROPE
ASIA
Atlantic Ocean
AFRICA
SOUTH AMERICA
Indian Ocean
Top 10 power consuming states from all regions
Regions ER NER NR SR WR
Sum of Usage
0.2M
0.1M
0.0M
States
Distribution of power consumption by Region
164.71K (9.63%)
540.37K (31.61%)
476.68K (27.88%)
506.25K (29.61%)
Regions
WR
NR
SR
ER
NER
Distribution of power consumption by Region and States
States Andhra Pradesh Arunachal Pra... Assam Bihar Chandigarh
0.6M
0.4M
0.2M
0.0M
Sum of Usage
WR
NR
SR
ER
NER
Regions
States
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Delhi
DNH
Goa

CONCLUSION

The project “Analysis of commercial electricity consumption” using PowerBI has successfully demonstrated the potential of data analytics in the electricity sector. The real-time analysis of customer data has provided valuable insights into customer behavior, preferences, and trends, thereby facilitating informed decision-making. The interactive dashboards and reports have offered a comprehensive view of customer data, enabling the identification of patterns and correlations. This has not only improved the efficiency of data analysis but also enhanced the bank’s ability to provide personalized services to its customers. The project has also highlighted the importance of data visualization in making complex data more understandable and accessible. The use of PowerBI has made it possible to present data in a visually appealing and easy-to-understand format, thereby aiding in better decision-making.

FUTURE SCOPES

As Indian economy is continuing to make a remarkable progress, it is estimated that the country is set to experience the largest increase in energy demand over a period of 20 years. The ever-expanding industrialization and urbanization will primarily drive the energy demand that is forecasted to reach 405 Gigawatts of renewable energy capacity by 2030. Having said that, several concerns are being raised on how it is going to meet the soaring demand without creating any exacerbating impact on the economy and its citizens.

REFERENCES

https://www.sciencedirect.com/topics/earth-and-planetary-sciences/energy-consumption

https://timesofindia.indiatimes.com/blogs/voices/the-future-of-energy-sector-in-india-trends/

LINK