Tech Saksham

Case study Report

Data Analytics with power BI

“Analysis of commercial Electricity Consumption in Indian State”

“A.P.C. Mahalaxmi college for women”

NM ID	NAME
33AC7ACC08914D7AEDD41FF748A1EAC5	V.SUJITHA

R.UMAMAHESHWARI	Trainer Name
R.UMAMAHESHWARI	Master Trainer

ABSTRACT

This work explains how to analyze the commercial electricity consumption of Indian consumers, and extract key components such as heating, ventilation and air conditioning (HVAC), residential lighting, and street lighting consumption from the total consumption. To avoid explicit modeling of dependencies on time of day and on working versus non-working days, least-squares fitting for outside temperature and natural illumination dependency proceeds independently for each hour of the day. Cubic polynomials model dependencies on Steadman apparent temperature and on log-scale illumination, but spline surfaces are best when considering these variables jointly. The primary focus is on residential consumption, but the same techniques can be used for studying street lighting, commercial and industrial consumption.

INDEX

Chapter 1

INTRODUCTION

1.1Problem statement

In Today ’s world Our Indian population has been increasing day by day. Amidst of this, Government has to provide a lot of basic amenities to its civilians .One of the important facility is Commercial Electricity Consumption in Indian state .

1.2 Proposed Solution

The proposed solution is to develop a Power BI 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 electricity Boards to make informed decisions. The dashboard will be interactive, user-friendly, and customizable, allowing banks to tailor it to their specific 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

In addition to reducing energy expenses and protecting against rising energy costs, an energy consumption analysis can also result in operational benefits. Through building system upgrades, organizations can experience:

- Increased operational efficiency
- Improved equipment reliability and longevity
- Reduced downtime due to equipment malfunctions
- Self-regulation of the building's smart systems, providing future savings

1.5 Scope

- Growth in commercial floor-space area and air-conditioning coverage and use will be the key influencing factors for commercial electricity demand growth. Whereas commercial heating, ventilation and air-conditioning (HVAC) already consume over 50 percent of its total demand, this could likely grow up to 75 percent as

total commercial demand grows to 247-348 TWh (7.6-10.1 percent CAGR) by 2030. This points to the big role of new technologies in centralised and inverter based air-conditioning, as well as building shell optimisation including insulation, roofing, window glazing, window design

• As the Indian government plans to increase electrification of rail-route kilometers from 40 percent presently to 77 percent by 2022, the level of electricity consumption achieved by 2030 could be 35-43 TWh, growing at 5.0-6.3 percent CAGR from 17 TWh in 2015. Whereas the impact of electrification is minor on electricity demand, significant savings on account of efficiency gains and oil substitution accrue.

• The potential electricity demand from transition to electric vehicles (EV) across private and fleet categories, even with 100 percent EV sales by 2030 turns out to be less than 100 TWh, which is only 10 percent of aggregate consumption in 2015. However, the aggregate load potential ofsuch a transition can lead to significant volatility in instantaneous demand, with EVs potentially contributing 50 percent to peak load by 2030 in case of unmanaged charging. • Increasing urbanisation facilitated by shrinking household sizes and concentration of economic opportunities in urban areas will increase pressure on municipal services such as public lighting and water-pumping. These plus the

‘miscellaneous’ demand asreported by CEA are likely to more than double from 2015 levels of 49 TWh to 104- 115 TWh (5.1-5.8 percent CAGR) in 2030. Here again, while the quantum of electricity demanded is manageable (representing about 5 percent of total demand) enhancing the supporting infrastructure to keep pace with urban life

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:

- **Power BI**: 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:

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

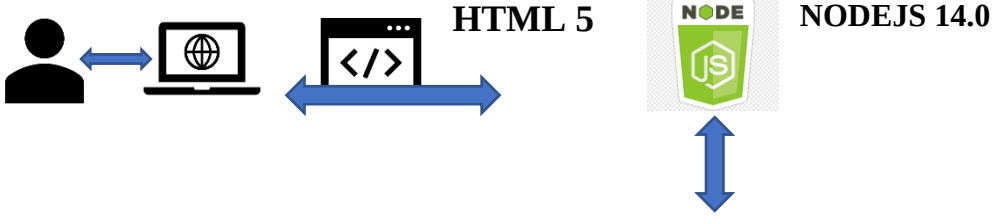
- **Power BI Service**: This is an online SaaS (Software as a Service) service that you use to publish reports, create new dashboards, and share insights.

- **Power BI 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



Database

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 Sage Maker. 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, Power BI Service (online), and PowerBI Mobile.

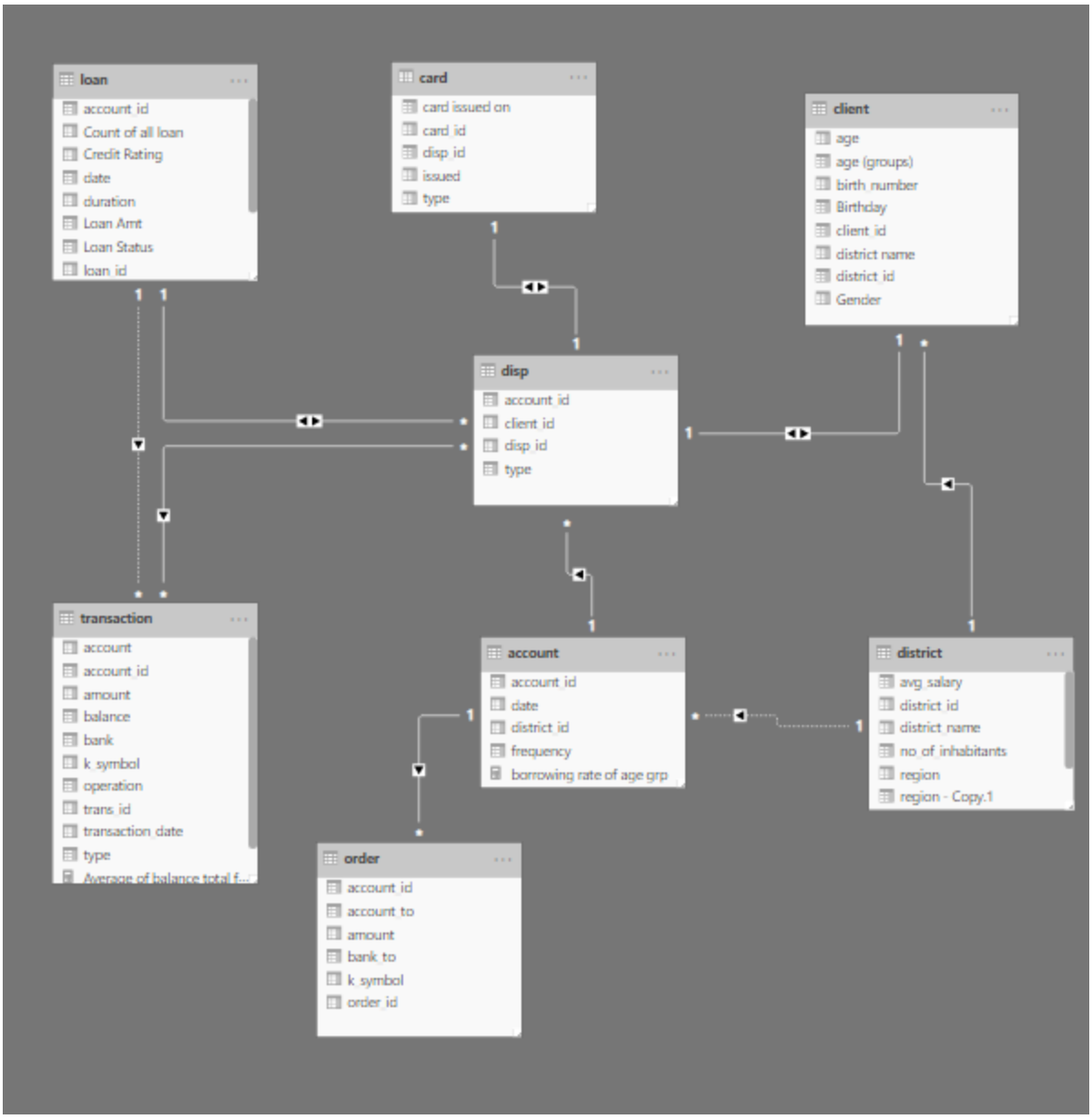
This architecture provides a comprehensive solution for real-time life

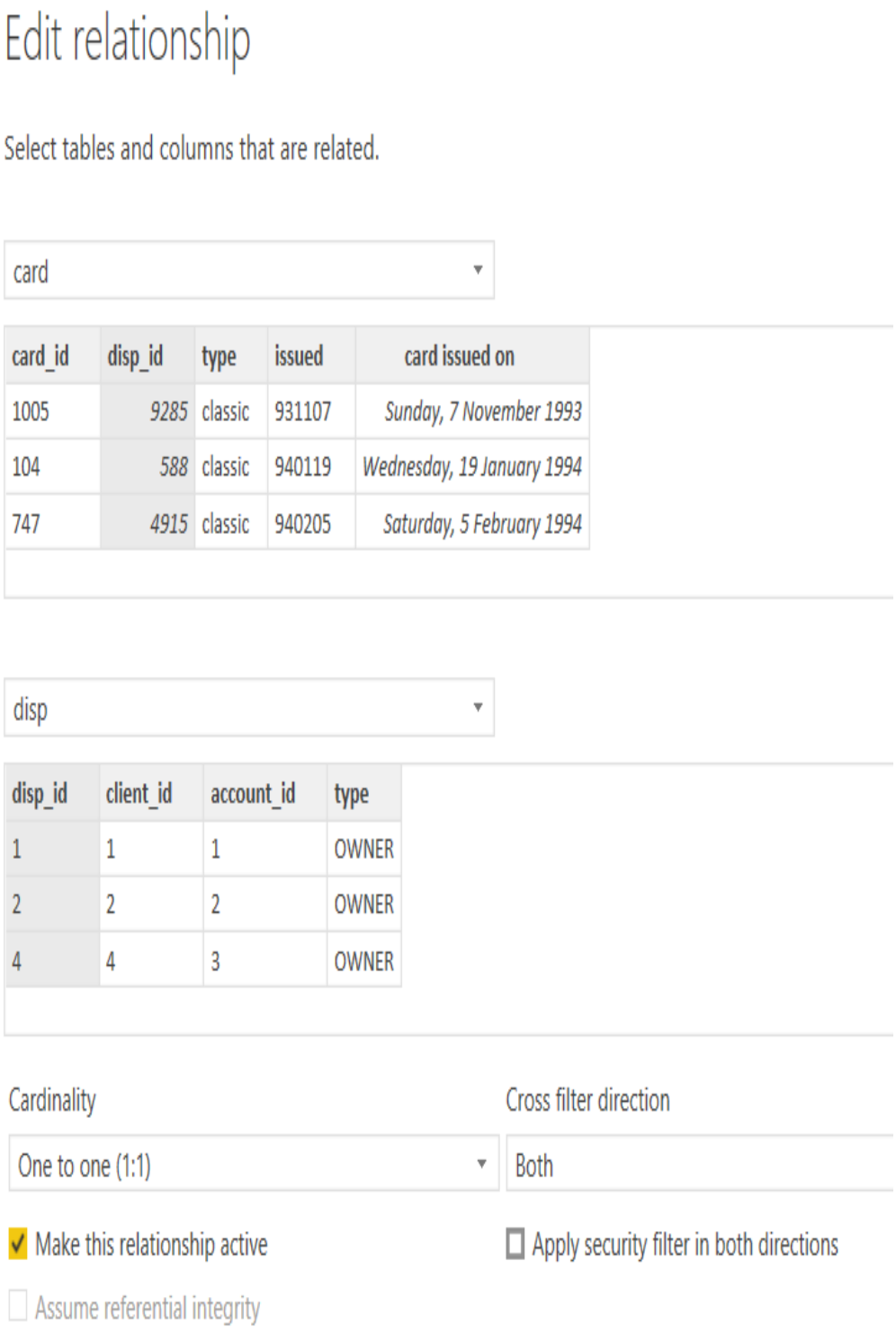
CHAPTER 4

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”


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



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

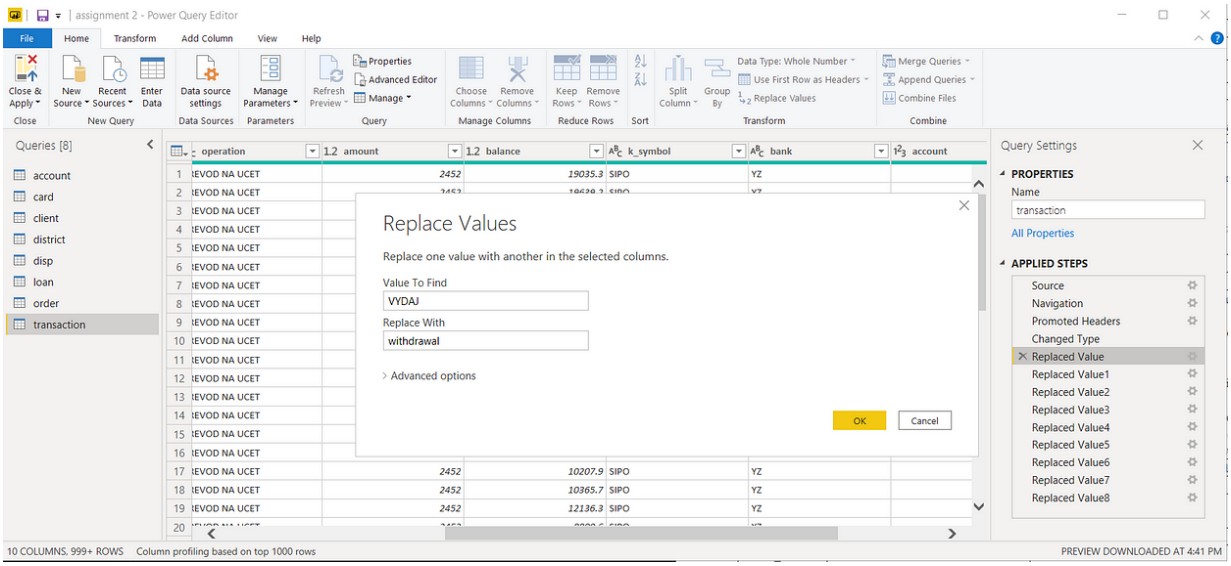
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
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

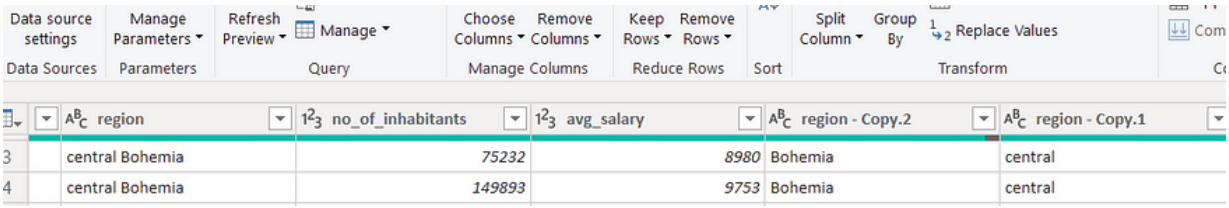
Set some fields to English for easy understanding, we replace values to English with the Power Query Editor.



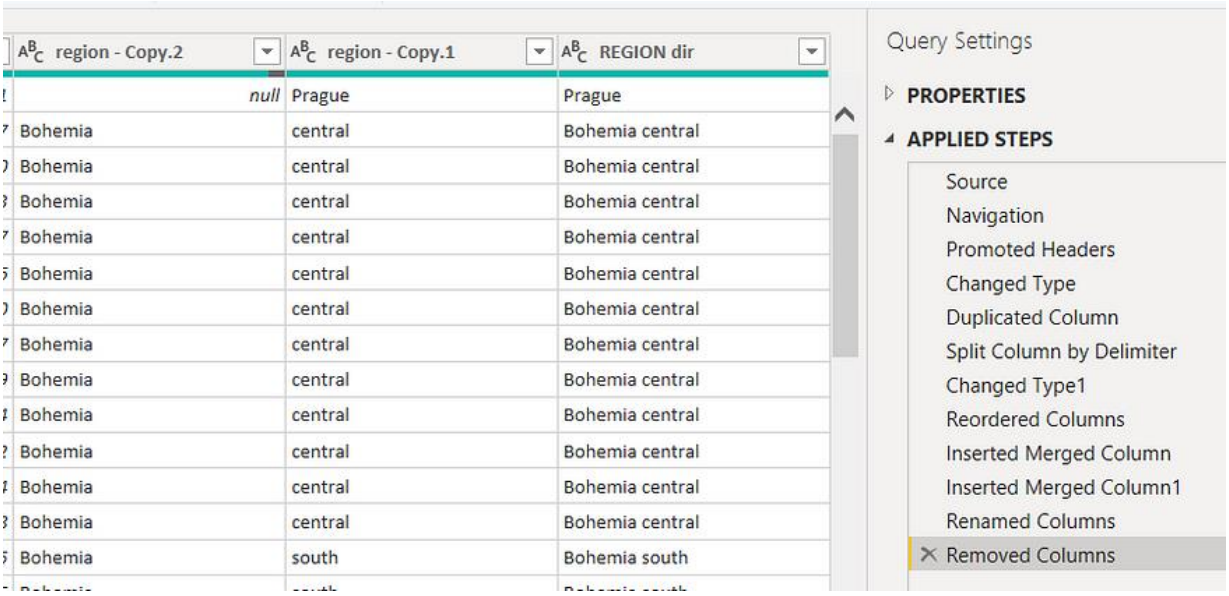
type	+/- transaction	"PRIJEM" stands for credit "VYDAJ" stands for withdrawal
k_symbol	characterization of the transaction	"POJISTNE" stands for insurance payment "SLUZBY" stands for payment for statement "UROK" stands for interest credited "SANKC. UROK" sanction interest if negative balance "SIPO" stands for household "DUCHOD" stands for old-age pension "UVER" stands for loan payment

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, we will group the ages into 5 groups.

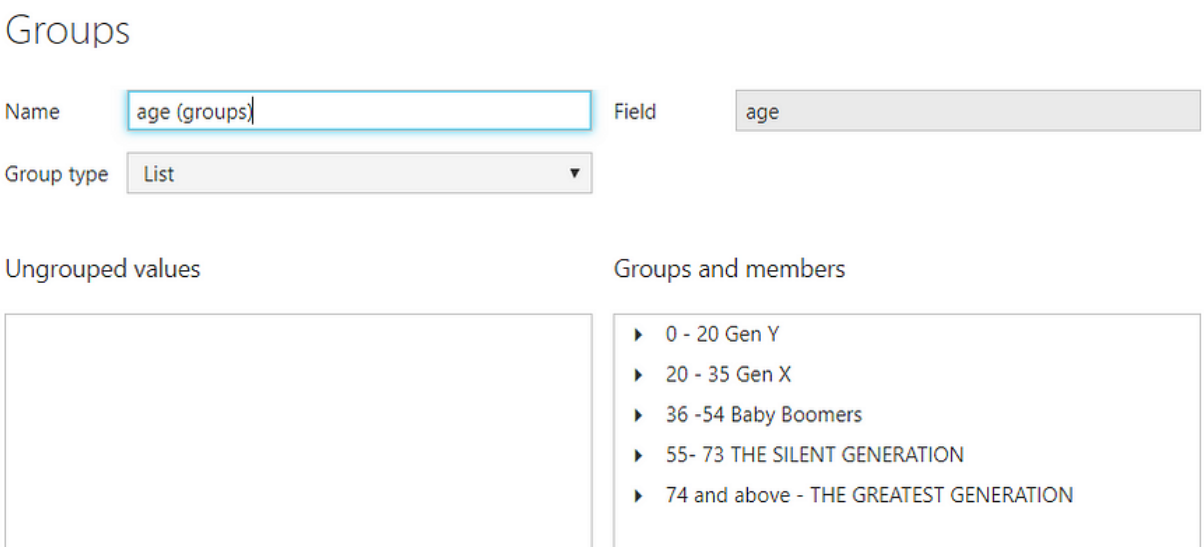
The Gen Y are youths,

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.

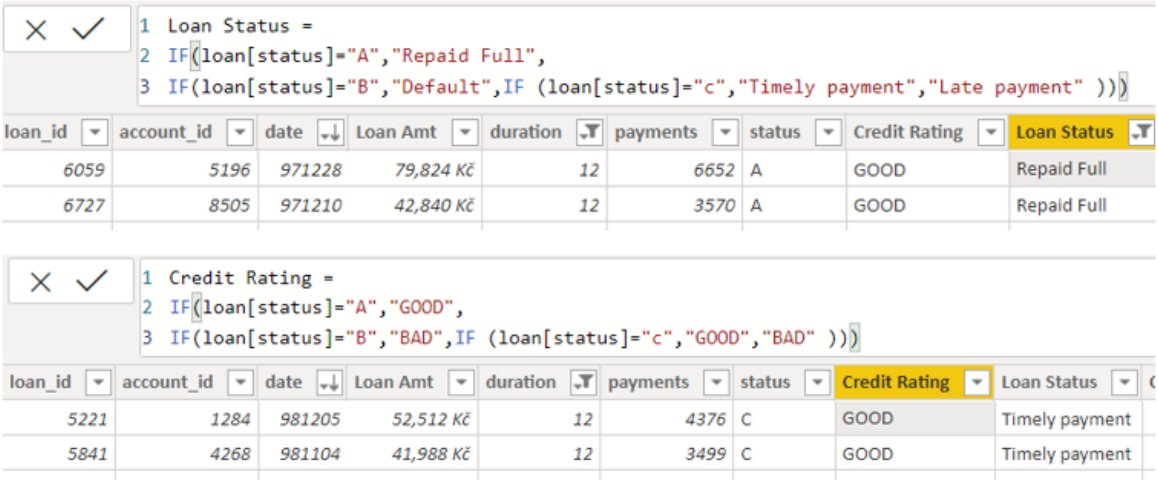
The greatest Generation, retired elderly living on pensions.



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

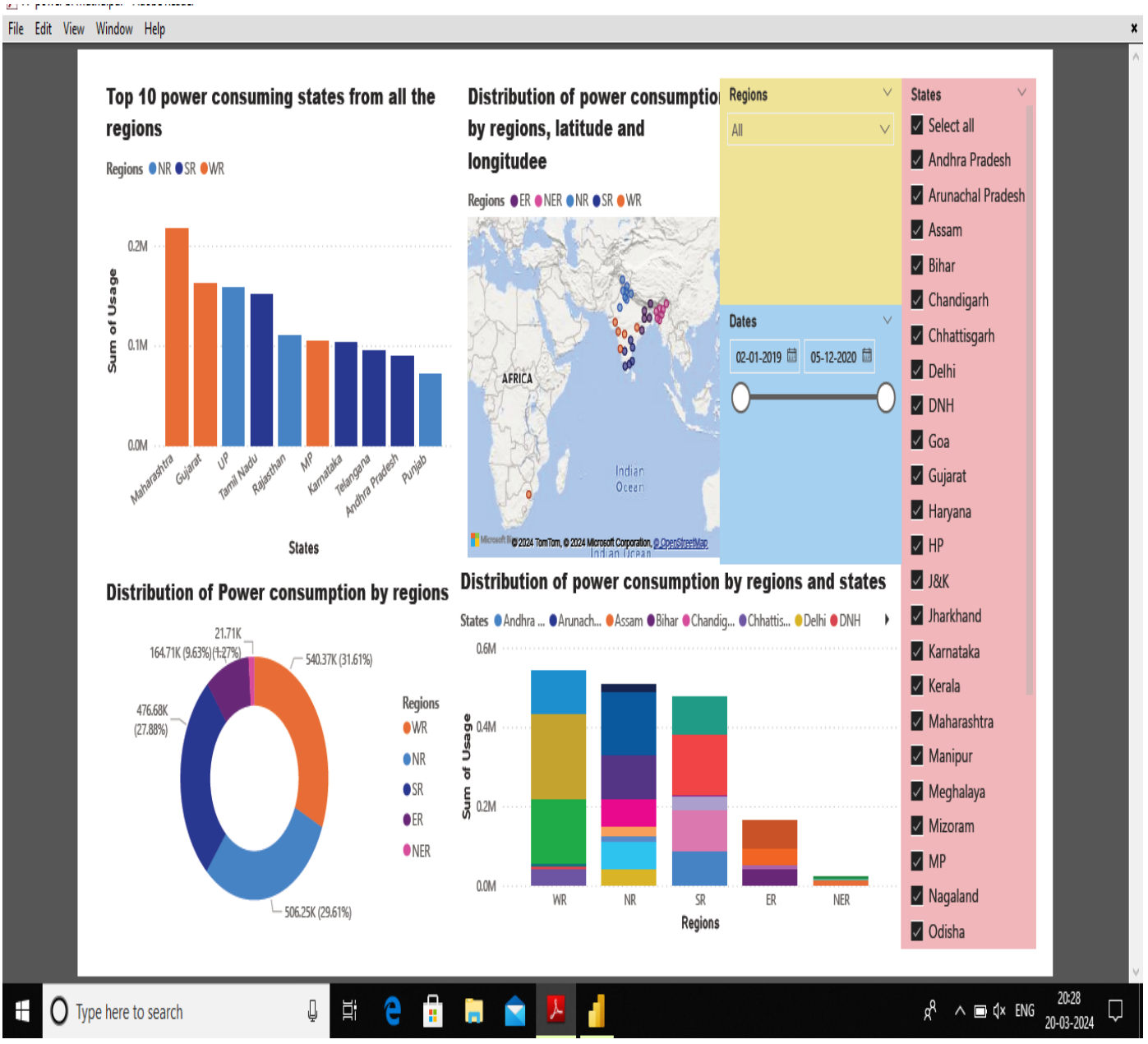
```
1 Credit Rating =
2 IF(loan[status]="A","GOOD",
3 IF(loan[status]="B","BAD",IF (loan[status]="c","GOOD","BAD" )))
```

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.

Dashboard


File Edit View Window Help
Top 10 power consuming states from all the regions
Regions NR SR WR
Sum of Usage
0.2M
0.1M
0.0M
Maharashtra
Gujarat
UP
Tamil Nadu
Rajasthan
MP
Karnataka
Telangana
Andhra Pradesh
Punjab
States
Distribution of power consumption by regions, latitude and longitudee
Regions ER NER NR SR WR
AFRICA
Indian Ocean
Regions
All
Dates
02-01-2019
05-12-2020
States
Select all
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Delhi
DNH
Goa
Gujarat
Haryana
HP
J&K
Jharkhand
Karnataka
Kerala
Maharashtra
Manipur
Meghalaya
Mizoram
MP
Nagaland
Odisha
Distribution of Power consumption by regions
21.71K
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 regions and states
States Andhra ... Arunach... Assam Bihar Chandig... Chhattis... Delhi DNH
0.6M
0.4M
0.2M
0.0M
Sum of Usage
WR
NR
SR
ER
NER
Regions
Type here to search
ENG
20:28
20-03-2024

CONCLUSION

The project “Analysis of commercial electricity consumption in Indian state” using Power BI has successfully demonstrated the potential of data analytics . The real-time analysis of consumer data has provided valuable insights into consumer behavior, preferences, and trends, thereby facilitating informed decision-making. The interactive dashboards and reports have offered a comprehensive view of consumer data, enabling the identification of patterns and correlations. This has not only improved the efficiency of data analysis but also enhanced the electricity consumption ’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 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, PowerBI'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

- https://www.brookings.edu/wp-content/uploads/2018/10/The-future-of-Indian-electricity-demand.pdf
- https://www.sciencedirect.com/topics/engineering/commercial-energy-consumption
- https://www.mdpi.com/2071-1050/13/21/11586


Microsoft
edunet
foundation
SAP
TAMIL NADU SKILL DEVELOPMENT CORPORATION
SKILLS FOR PROGRESS AND PROSPERITY

LINK