Tech Saksham

Case Study Report

Data Analytics with Power BI

“Analysis of Commercial Electricity Consumption in Indian States”

“Ambai Arts College”

NM ID	NAME
4468406C5240E3D8A2D40AB610E56695	Subisha U

Trainer Name: K. Umamaheshwari

Master Trainer: K. Umamaheshwari

Master Trainer

ABSTRACT

The project begins by collecting electricity consumption data from various sources, including smart meters, energy monitoring devices, and utility bills. The data is then processed, cleaned, and transformed into a format suitable for analysis using Power BI. Using Power BI's interactive dashboards, reports, and visualizations, users can explore key metrics such as total energy consumption, peak demand periods, and distribution across different time periods and locations. Advanced analytics capabilities, including trend analysis, forecasting, and anomaly detection, provide deeper insights into consumption patterns and potential efficiency gains. Moreover, the project leverages Power BI's integration capabilities to combine electricity consumption data with other relevant datasets, such as weather data, occupancy patterns, and renewable energy generation, to provide a comprehensive view of factors influencing electricity usage.

INDEX

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

"Increasing electricity consumption in urban areas is leading to higher demand on the grid, resulting in strain during peak hours and potential power outages. This poses challenges for sustainable energy management, infrastructure planning, and environmental sustainability. Addressing these issues requires innovative solutions to optimize consumption, reduce wastage, and promote renewable energy sources."

1.2 Proposed Solution

Smart grids utilize advanced metering, real-time monitoring, and automated control systems to optimize energy distribution, reduce transmission losses, and balance supply and demand more efficiently. Additionally, promoting energy conservation measures, such as incentivizing energy-efficient appliances, implementing demand response programs, and encouraging renewable energy generation at the local level, can help alleviate strain on the grid and promote sustainability. Additionally, investing in energy storage solutions, such as batteries and pumped hydro storage, can help store excess energy during off-peak hours for use during periods of high demand, further enhancing grid stability and reliability.

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

Residential Sector: Encouraging energy-efficient practices among households, promoting the use of energy-efficient appliances, and implementing home energy management systems.

Commercial Sector: Implementing energy management strategies in commercial buildings, such as lighting controls, HVAC optimization, and occupancy sensors, to reduce electricity consumption.

Industrial Sector: Introducing energy-efficient technologies and processes in industries, conducting energy audits, and implementing demand-side management programs to optimize electricity usage in manufacturing facilities.

Transportation: Promoting the adoption of electric vehicles (EVs) and developing charging infrastructure to reduce reliance on fossil fuels for transportation, thus indirectly impacting electricity consumption.

CHAPTER 2

SERVICES AND TOOLS REQUIRED

2.1 Services Used

- **Data Collection and Storage Services**: Electricity consumption of Indian states is collected from various industrial sectors consuming the power and the energy production resources are also collected and stored in cloud.

- **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 Power BI, 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.

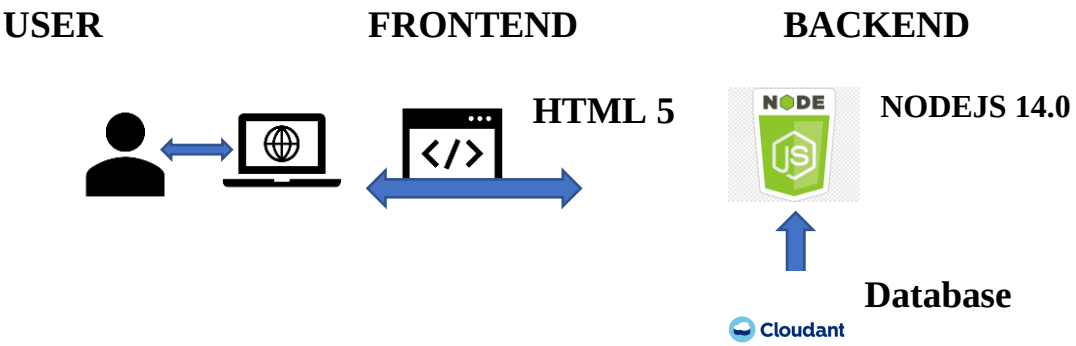
Software Requirements:

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

CHAPTER 3

PROJECT ARCHITECTURE

3.1 Architecture



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

Data Collection: Gather energy consumption data from various sources, such as smart meters, sensors, or utility bills. Ensure that the data is accurate and covers a suitable timeframe for analysis.

1. **Data Storage**: The collected data is stored in a database for processing. Azure SQL Database or AWS RDS can be used for this purpose.
2. **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 Sage Maker. These models can help in predicting customer behavior, detecting fraud, etc.
4. **Data Visualization**: Power BI offers a wide range of interactive visualizations, including charts, graphs, maps, tables, and custom visuals. These visualizations help in presenting data in a visually appealing and meaningful way, making it easier to understand and analyze.

5. **Data Access**: The dashboards created in Power BI can be accessed through Power BI Desktop, Power BI Service (online), and Power BI Mobile.

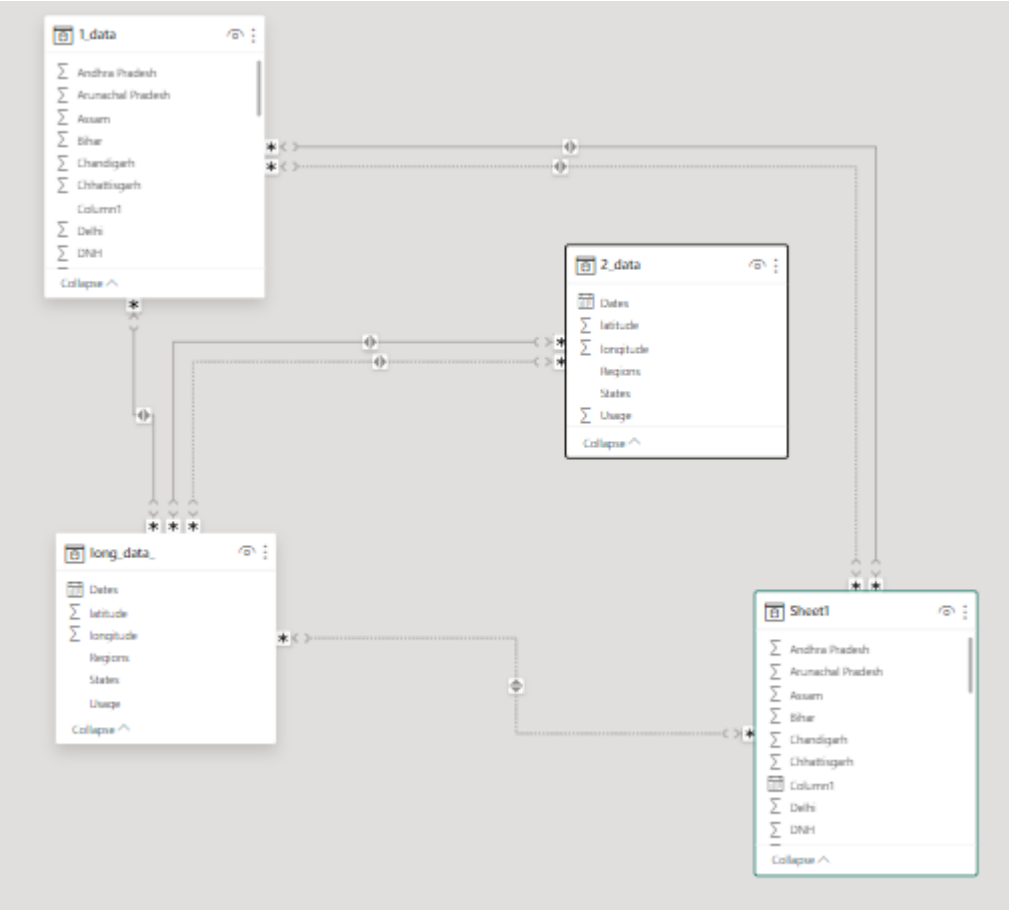
This architecture provides a comprehensive solution for real-time analysis of power consumption. However, it's important to note that the specific architecture may vary depending on the electricity's existing infrastructure, specific requirements, and budget. It's also important to ensure that all tools and services comply with relevant data privacy and security regulations.

CHAPTER 4

MODELING AND RESULT

Manage relationship

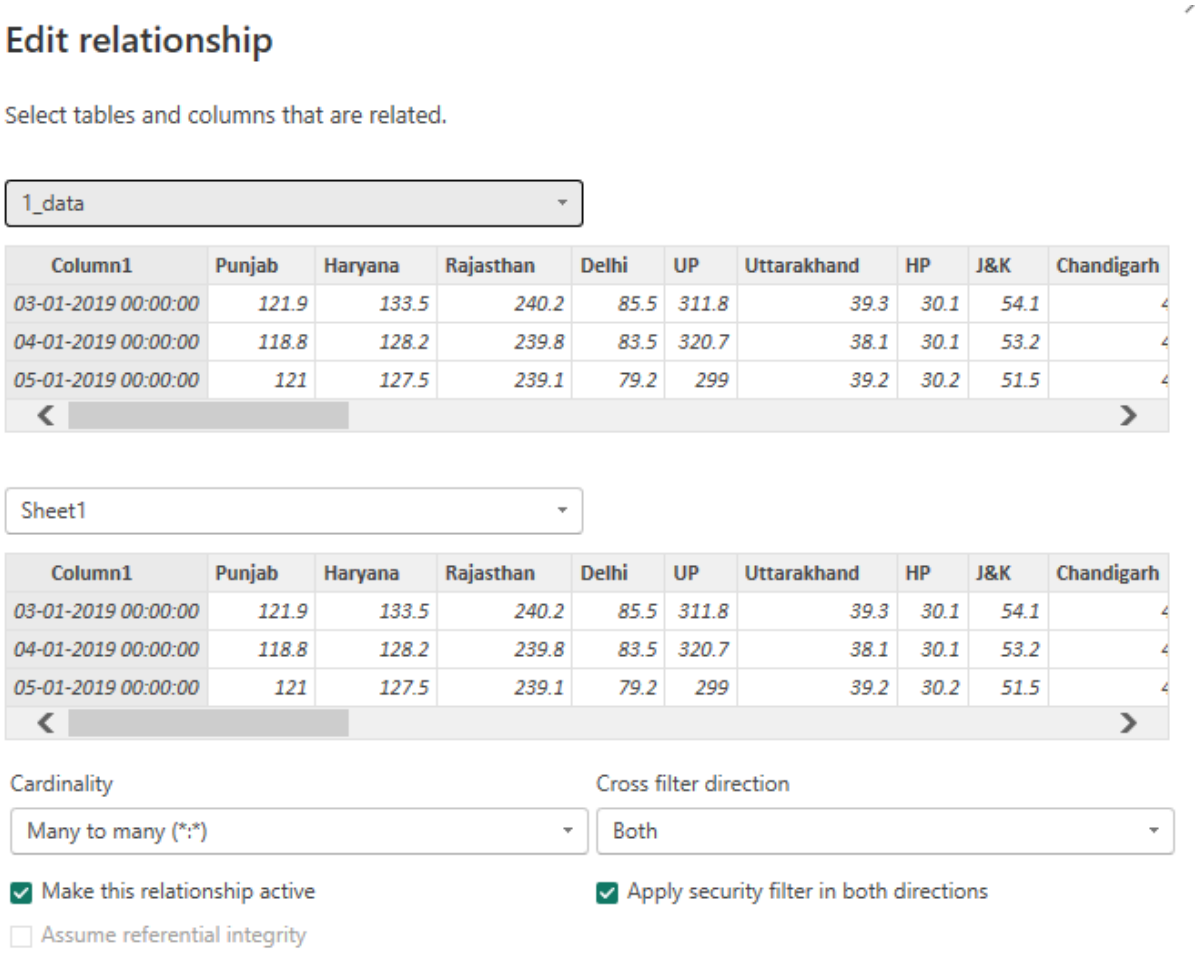
The “disp” file will be used as the main connector as it contains most key identifier (account id, client id and disp id) which can be use to relates the 8 data files together. The “district” file is use to link the client profile geographically with “district id”



Manage relationships

Active	From: Table (Column)	To: Table (Column)
☑	1_data (Column1)	Sheet1 (Column1)
☐	1_data (Punjab)	Sheet1 (Punjab)
☑	1_data (Tripura)	long_data_ (Usage)
☐	2_data (latitude)	long_data_ (latitude)
☑	2_data (States)	long_data_ (States)
☐	long_data_ (longitude)	Sheet1 (Delhi)

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



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

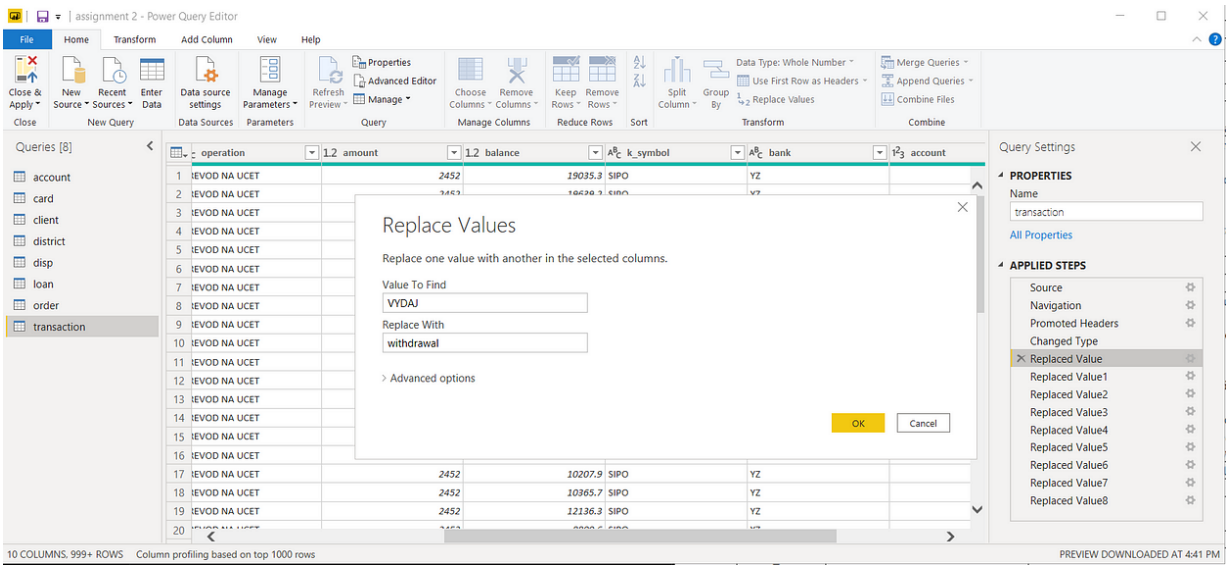
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
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

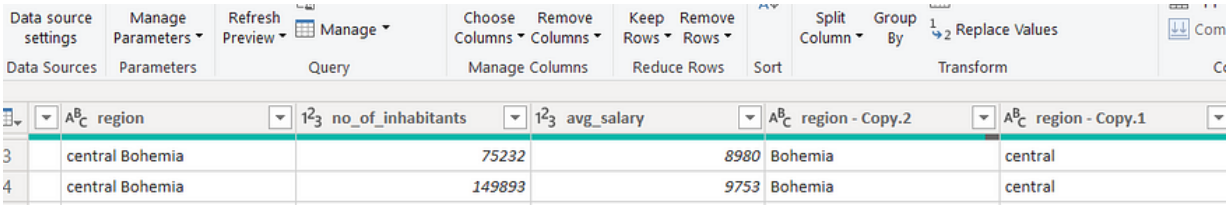
Set some fields to English for easy understanding, we replace values to English with the Power Query Editor.



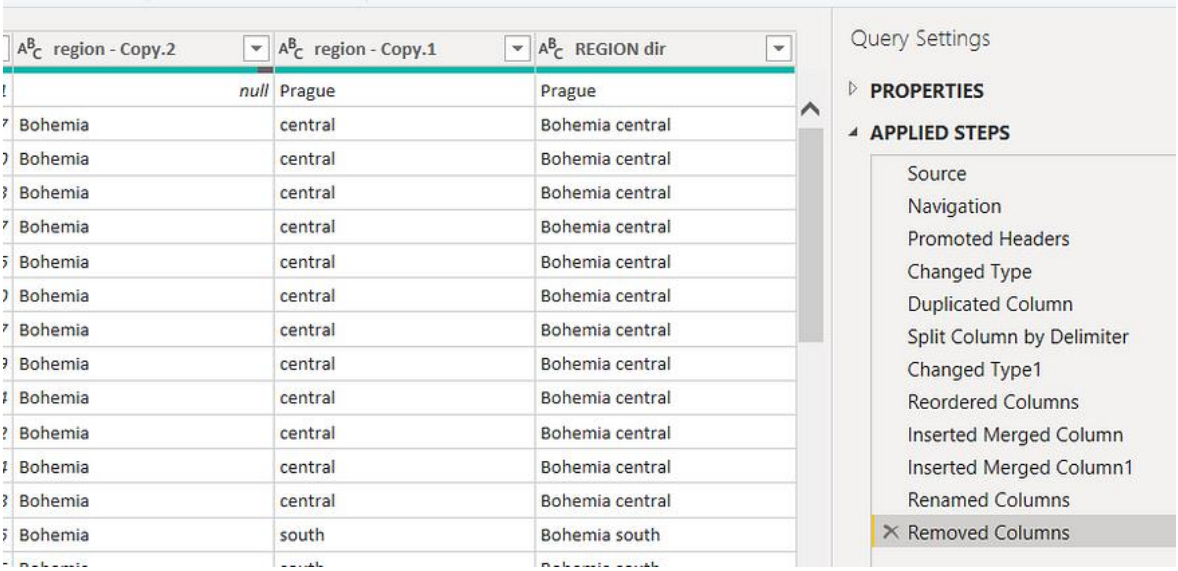
type	+/- transaction	"PRIJEM" stands for credit "VYDAJ" stands for withdrawal
k_symbol	characterization of the transaction	"POJISTNE" stands for insurance payment "SLUZBY" stands for payment for statement "UROK" stands for interest credited "SANKC. UROK" sanction interest if negative balance "SIPO" stands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



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.

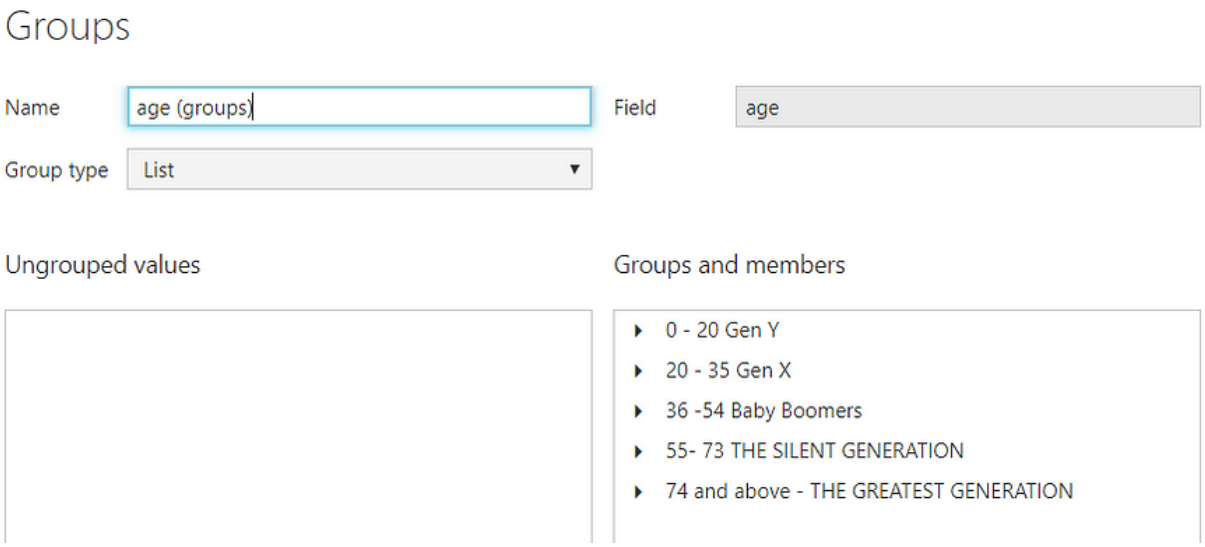
The Gen Y are youths,

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.



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

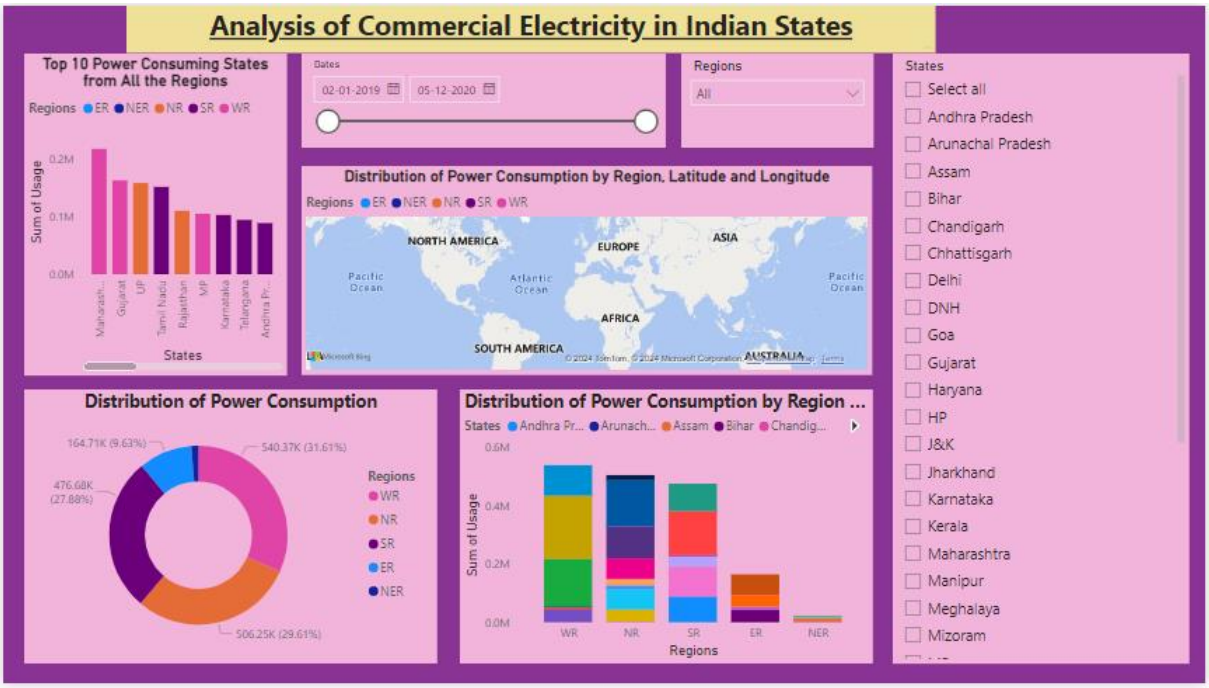
```
1 Credit Rating =
2 IF(loan[status]="A","GOOD",
3 IF(loan[status]="B","BAD",IF (loan[status]="c","GOOD","BAD" )))
```

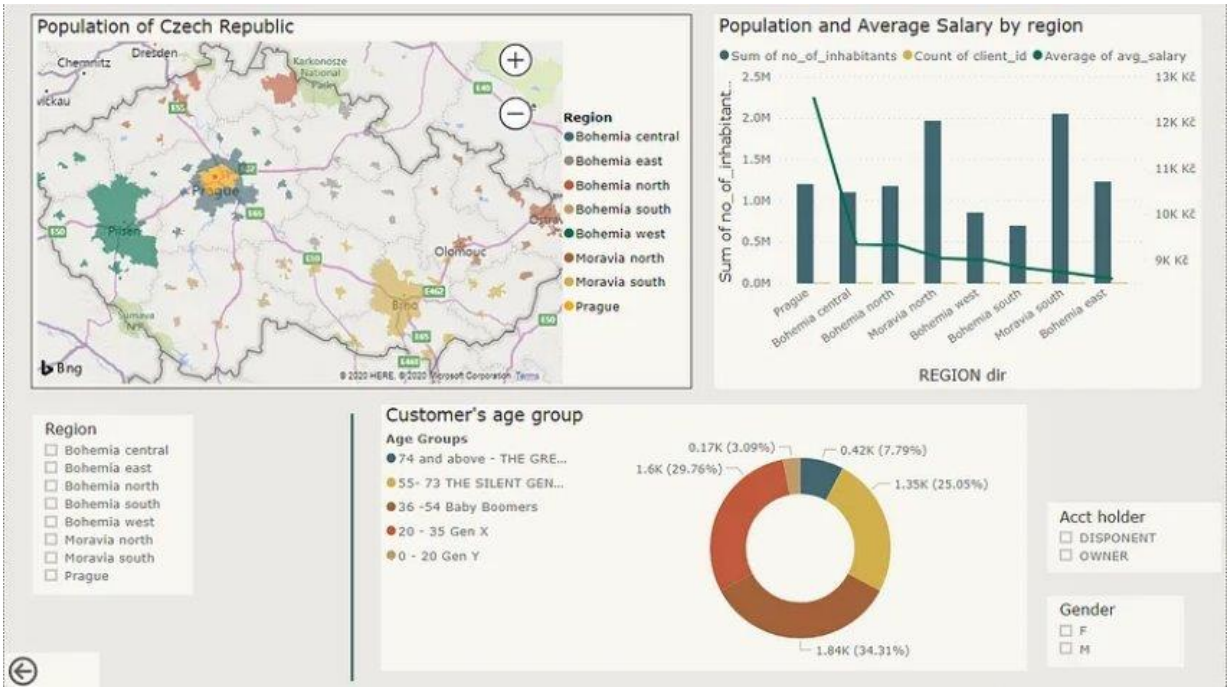
loan_id	account_id	date	Loan Amt	duration	payments	status	Credit Rating	Loan Status
5221	1284	981205	52,512 Kč	12	4376	C	GOOD	Timely payment
5841	4268	981104	41,988 Kč	12	3499	C	GOOD	Timely payment

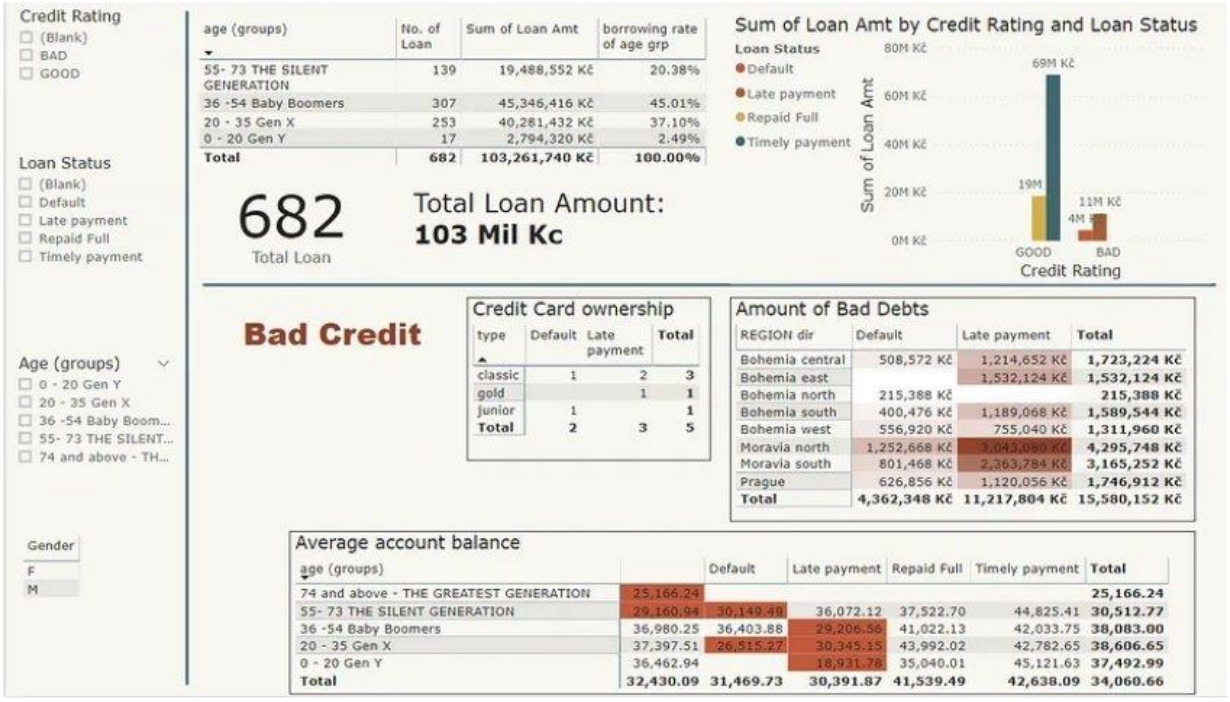
Values of such as "account Id" have also been set as Text.

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







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%**

type	Default	Late payment	Total
classic	1	2	3
gold		1	1
Junior	1		1
Total	**2**	**3**	**5**

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)		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

In conclusion, effectively managing electricity consumption is crucial for ensuring sustainability, reducing costs, and mitigating environmental impact. By implementing smart metering, energy management systems, and demand response programs, individuals, businesses, and utility companies can monitor, analyze, and optimize electricity usage in real-time. Furthermore, integrating renewable energy sources, deploying energy storage solutions, and promoting energy efficiency measures contribute to a more resilient and sustainable energy infrastructure. Public awareness and education play a vital role in fostering a culture of energy conservation and responsible consumption practices. Overall, by leveraging technology, data-driven insights, and collaborative efforts, we can achieve a more efficient, reliable, and environmentally friendly electricity consumption ecosystem for the future.

.

FUTURE SCOPE

Smart Grid Evolution: Continued development and deployment of smart grid technologies will enable more efficient electricity distribution, real-time monitoring, and dynamic pricing mechanisms to optimize consumption and grid stability.

Internet of Things (IoT) Integration: Integration of IoT sensors and devices into electrical infrastructure and appliances will enable more granular data collection, predictive maintenance, and automated energy management strategies.

Artificial Intelligence (AI) and Machine Learning: AI and machine learning algorithms will play a crucial role in analyzing complex data sets, predicting electricity demand patterns, optimizing energy usage, and identifying opportunities for efficiency improvements.

Energy Storage Innovations: Advancements in energy storage technologies, such as batteries, hydrogen storage, and flywheels, will enable greater integration of renewable energy sources, grid flexibility, and resilience against supply disruptions.

REFERENCES

https://ijarsct.co.in/Paper9099.pdf

https://wecindia.in/indias-yearly-energy-consumption-analysis/

LINK

