Tech Saksham

Case Study Report

Data Analytics with Power BI

"Analysis of Commercial Electricity Consumption in Indian States"

"Ambai Arts College"

NM ID	NAME
73127941779E25A8F934E4B55C7C2AF9	Muthulakshmi E

Trainer Name K.Umamaheswari

Master Name K.Umamaheswari

ABSTRACT

This study employs Power BI (Business Intelligence) to analyze commercial electricity consumption patterns in India. With the increasing demand for electricity in the commercial sector due to economic growth and urbanization, understanding consumption trends becomes paramount for efficient resource management and policy formulation. The dataset used comprises historical electricity consumption data from various commercial establishments across different regions of India. Through interactive visualizations and data analysis tools offered by Power BI, this study aims to identify key factors influencing commercial electricity consumption, such as industry type, geographical location, seasonal variations, and economic indicators. By leveraging Power BI's capabilities, including data modeling, visualization, and predictive analytics, this research provides insights into patterns, trends, and correlations within the commercial electricity consumption landscape. The findings of this study can aid policymakers, energy regulators, and commercial stakeholders in making informed decisions regarding energy efficiency measures, infrastructure planning, and sustainable development initiatives.

INDEX

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

The consumption of energy plays a vital role to determine the economic growth of a country. Our society is energy based one. Today the demand for energy has been growing rapidly due to increase infrastructure development, rise in per capita income, modernization, rapid population growth, industrialization, urbanization. Further this leads to change in the consumption of energy from traditional aspect to new modern technological aspect in the country. The supply of energy has not been increasing in the same proportion of the demand for energy. The country's economic condition influences the quantity of energy to be consumed by the various sectors from the availability of energy resources in India. Even though to fulfill the present energy need from our own resources still we are in the position to import energy resources from another country. In the present study, an attempt is made is to understand the changing commercial energy production and consumption pattern in India.

1.2 Proposed Solution

To achieve this goal, big data analytics has emerged as a promising tool for analyzing electricity consumption patterns and identifying opportunities for energy savings. By analyzing large volumes of data generated by smart meters and other monitoring devices, utility companies can gain insights into the behavior of individual consumers, as well as the overall trends in consumption. This information can be used to develop targeted strategies for reducing energy use, promoting energy efficiency, and managing peak demand. The paper aims to present a study on the electricity consumption patterns of 500 residential consumers in a particular area over 24 months, using big data analytics. The data was obtained from the MSEB and analyzed to identify trends in consumer behavior.

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

India is the third largest producer of electricity in the world. During the fiscal year (FY) 2022–23, the total electricity generation in the country was 1,844 TWh, of which 1,618 TWh was generated by utilities. The gross electricity consumption per capita in FY2023 was 1,327 kWh. In FY2015, electric energy consumption in agriculture was recorded as being the highest (17.89%) worldwide. The per capita electricity consumption is low compared to most other countries despite India having a low electricity traffic. The Indian national electric grid has an installed capacity of 416.0 GW as of 31 March 2023. Renewable energy plants, which also include large hydroelectric power plants, constitute 40.7% of the total installed capacity. India's electricity sector is dominated by fossil fuels, in particular coal, which produced about three-quarters of the country's electricity. The government declared its efforts to increase investment in renewable energy. Under the government's 2023-2027 National Electricity Plan, India will not build any new fossil fuel power plants in the utility sector, aside from those currently under construction. It is expected that non-fossil fuel generation contribution is likely to reach around 44.7% of the total gross electricity generation by 2029–30.

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:

- **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.

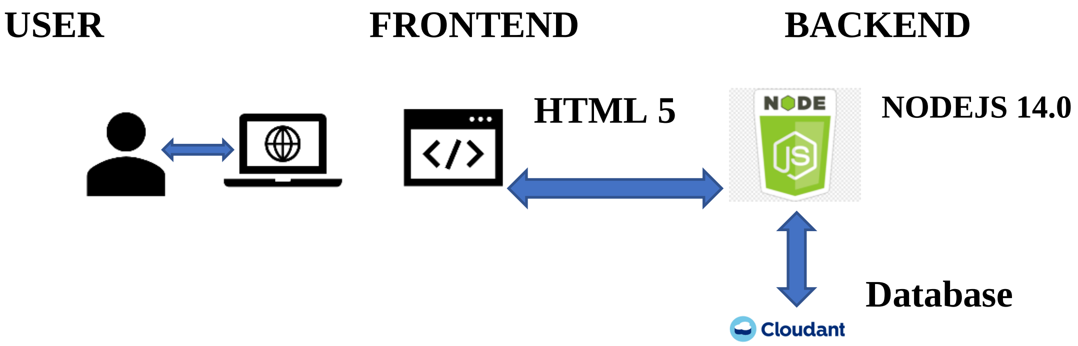
Software Requirements:

- **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



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 SageMaker. 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 PowerBI can be accessed through PowerBI Desktop, PowerBI Service (online), and PowerBI Mobile.

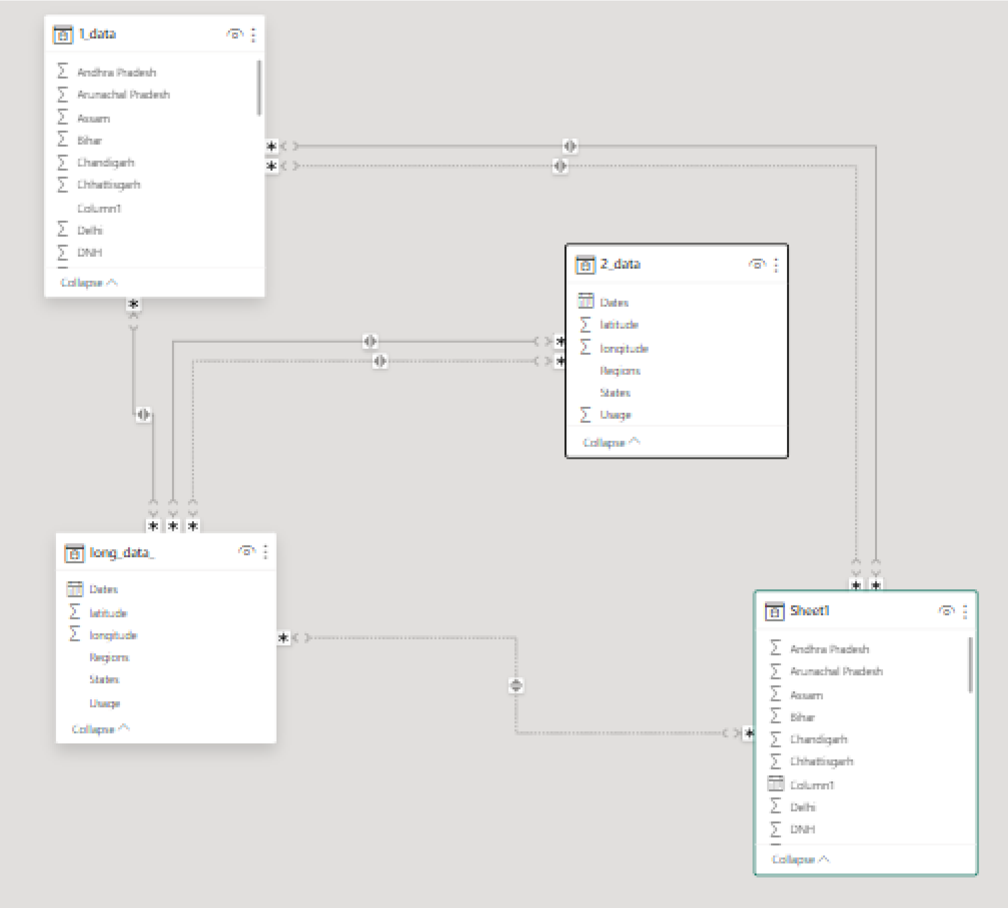
This architecture provides a comprehensive solution for real-time analysis of power consumption. However, it’s important to note that the specific architecture may vary depending on the electricity’s existing infrastructure, specific requirements, and budget. It’s also important to ensure that all tools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

Edit relationship

Select tables and columns that are related.

1_data

Column1	Punjab	Haryana	Rajasthan	Delhi	UP	Uttarakhand	HP	J&K	Chandigarh
03-01-2019 00:00:00	121.9	133.5	240.2	85.5	311.8	39.3	30.1	54.1	4
04-01-2019 00:00:00	118.8	128.2	239.8	83.5	320.7	38.1	30.1	53.2	4
05-01-2019 00:00:00	121	127.5	239.1	79.2	299	39.2	30.2	51.5	4

Sheet1

Column1	Punjab	Haryana	Rajasthan	Delhi	UP	Uttarakhand	HP	J&K	Chandigarh
03-01-2019 00:00:00	121.9	133.5	240.2	85.5	311.8	39.3	30.1	54.1	4
04-01-2019 00:00:00	118.8	128.2	239.8	83.5	320.7	38.1	30.1	53.2	4
05-01-2019 00:00:00	121	127.5	239.1	79.2	299	39.2	30.2	51.5	4

Cardinality: Many to many (*:*)

Cross filter direction: Both

- [x] Make this relationship active
- [x] Apply security filter in both directions
- [ ] Assume referential integrity

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.


```
Gender =
VAR stringDate = FORMAT(client[birth_number],"General Number")
VAR month = VALUE(MID(stringDate,3,2))
RETURN IF(month > 50,"F","M")
```

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
724	855114	46	F	14/01/1985	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.

```
Birthday =
VAR stringDate = FORMAT(client[birth_number],"General Number")
VAR stringMonth = VALUE(MID(stringDate,3,2))
VAR mth = IF(stringMonth > 50, stringMonth - 50,stringMonth)
VAR year = VALUE(MID(stringDate,1,2))
VAR day = VALUE(MID(stringDate,5,2))
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

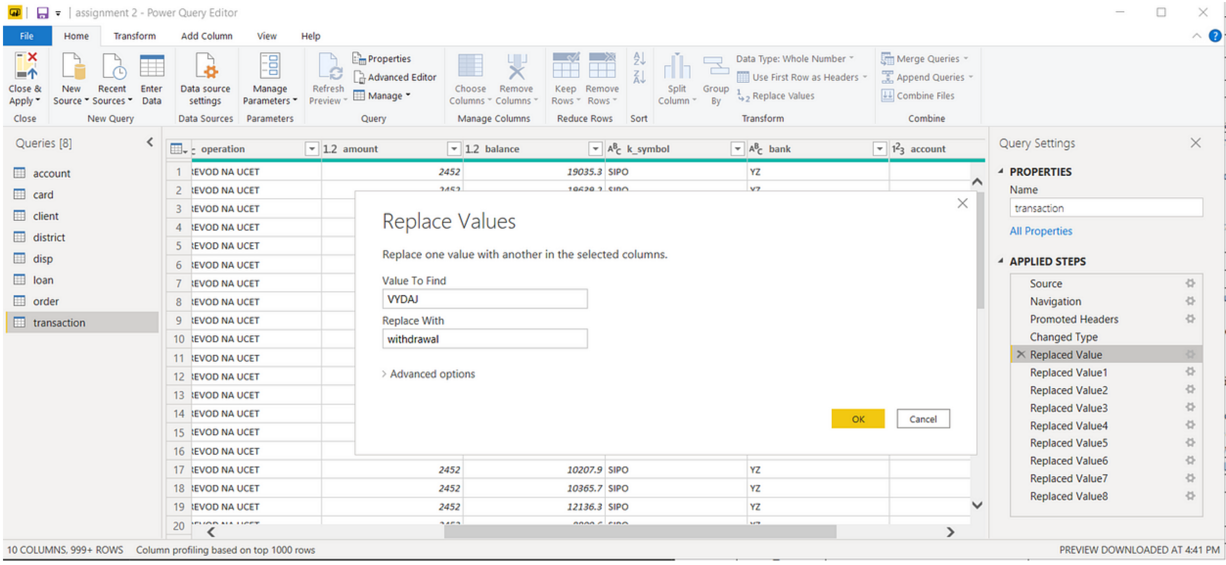
For Age, we shall assume it is year 1999 as explain previously and use it to minus from the birth year.

```
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.

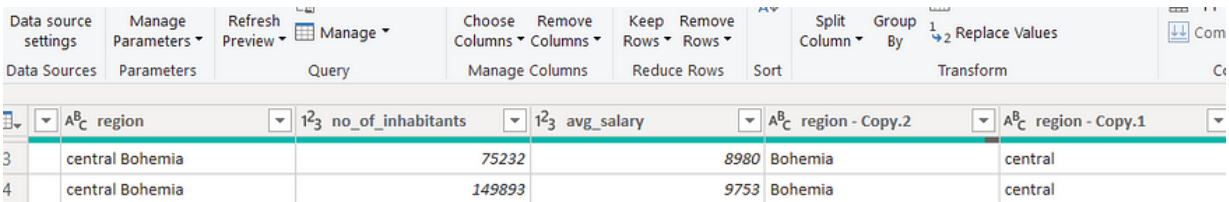


type	+/- transaction	"PRIJEM" stands for credit "VYDAJ" stands for withdrawal

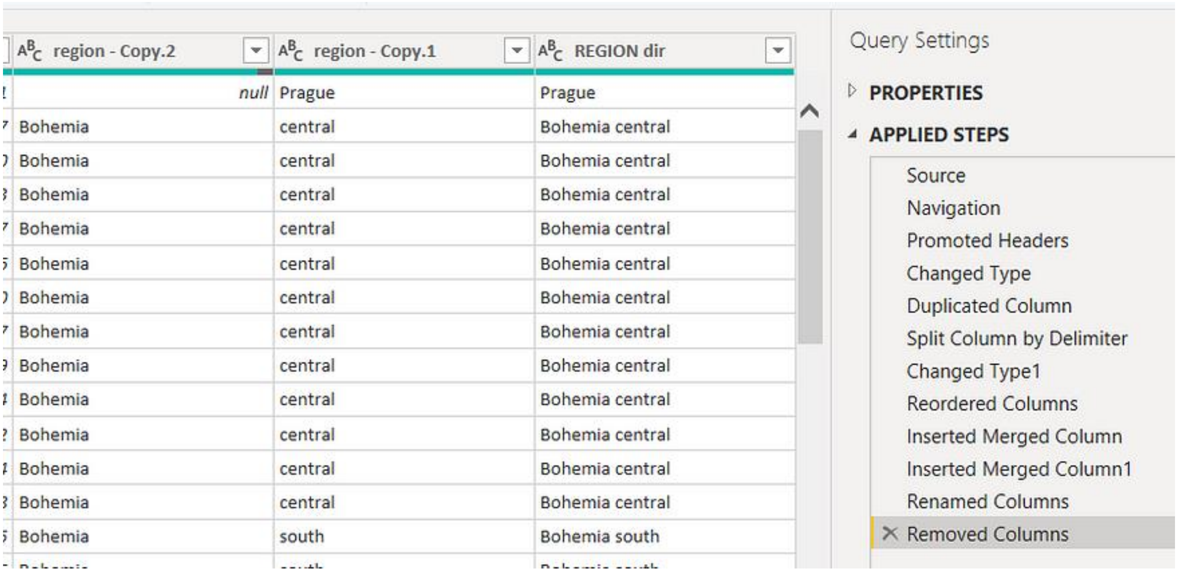
k_symbol	characterization of the transaction	"POJISTNE" stands for insurance payment "SLUZBY" stands for payment for statement "UROK" stands for interest credited "SANKC. UROK" sanction interest if negative balance "SIPO" stands for household "DUCHOD" stands for old-age pension "UVER" stands for loan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

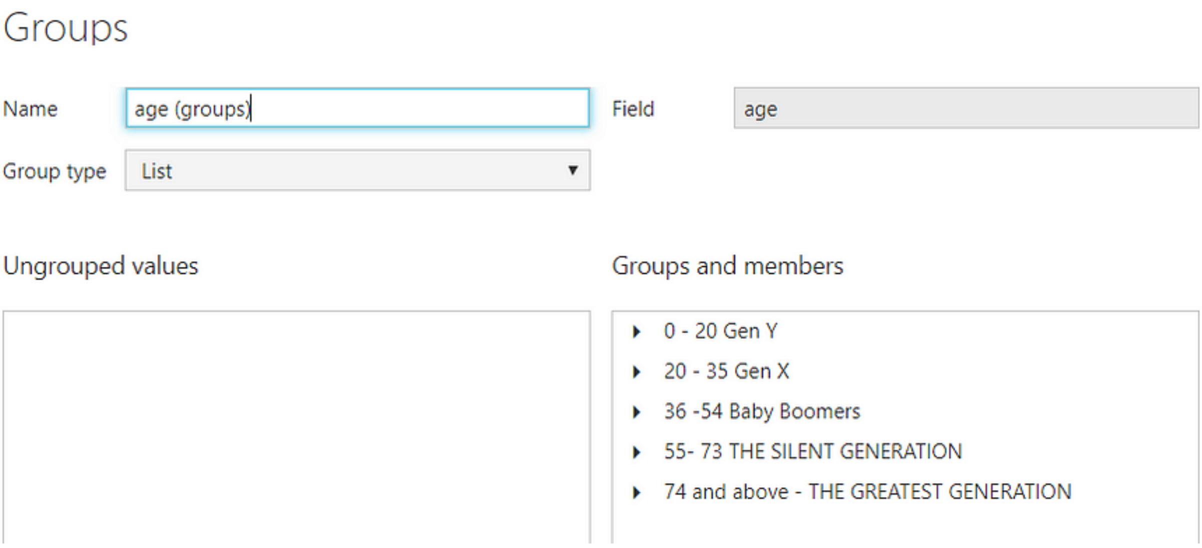
The Gen Y are youths,

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

```
Loan Status =
IF(loan[status]="A","Repaid Full",
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

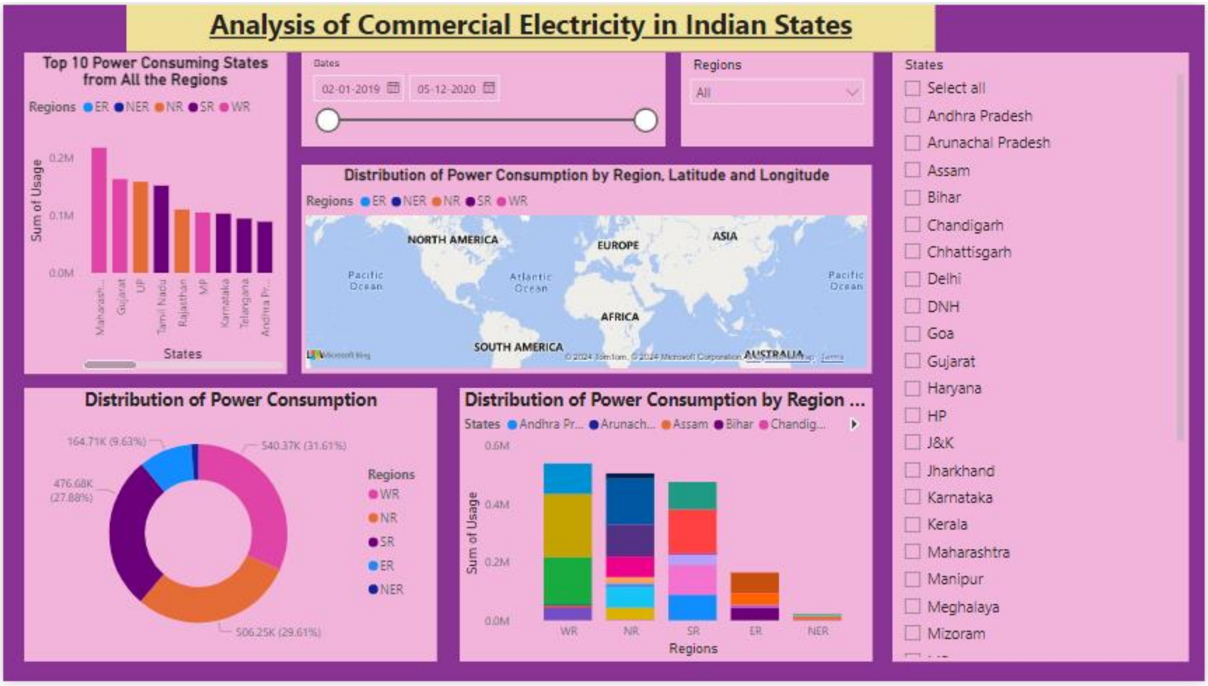
```
Credit Rating =
IF(loan[status]="A","GOOD",
IF(loan[status]="B","BAD",IF (loan[status]="c","GOOD","BAD" )))
```

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

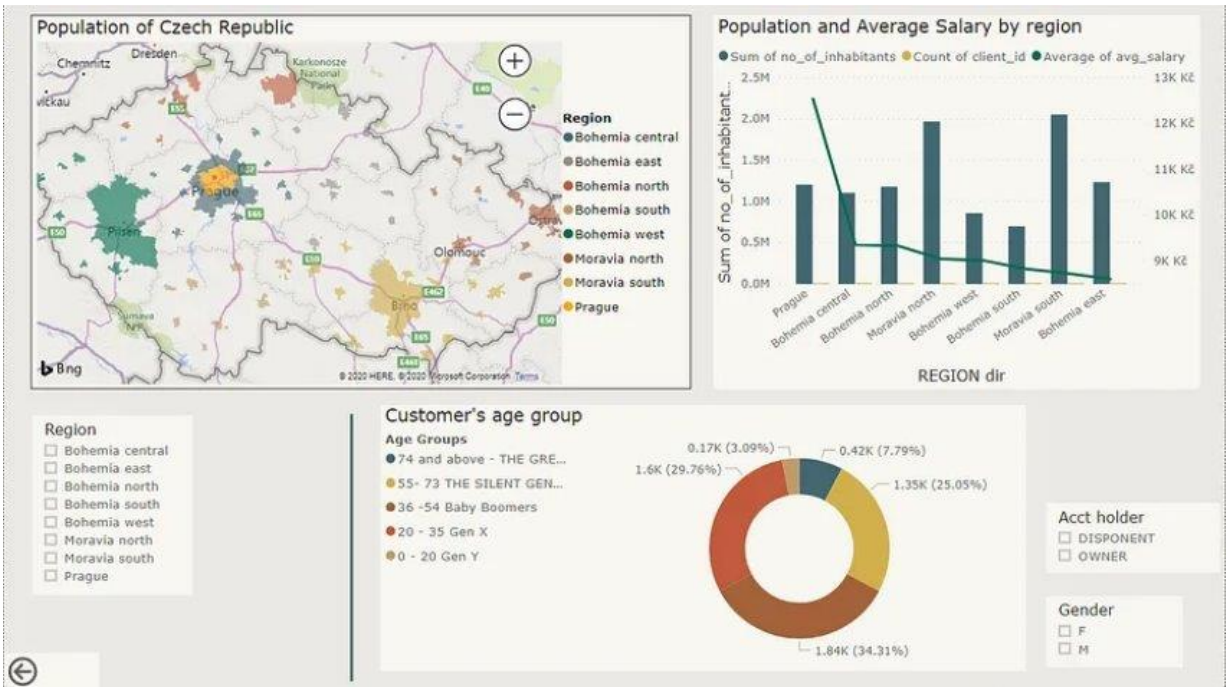
Values of such as "account Id" have also been set as Text.

And District name have been categorized as place to be use for the map to show the sum of the inhabitants in each region.

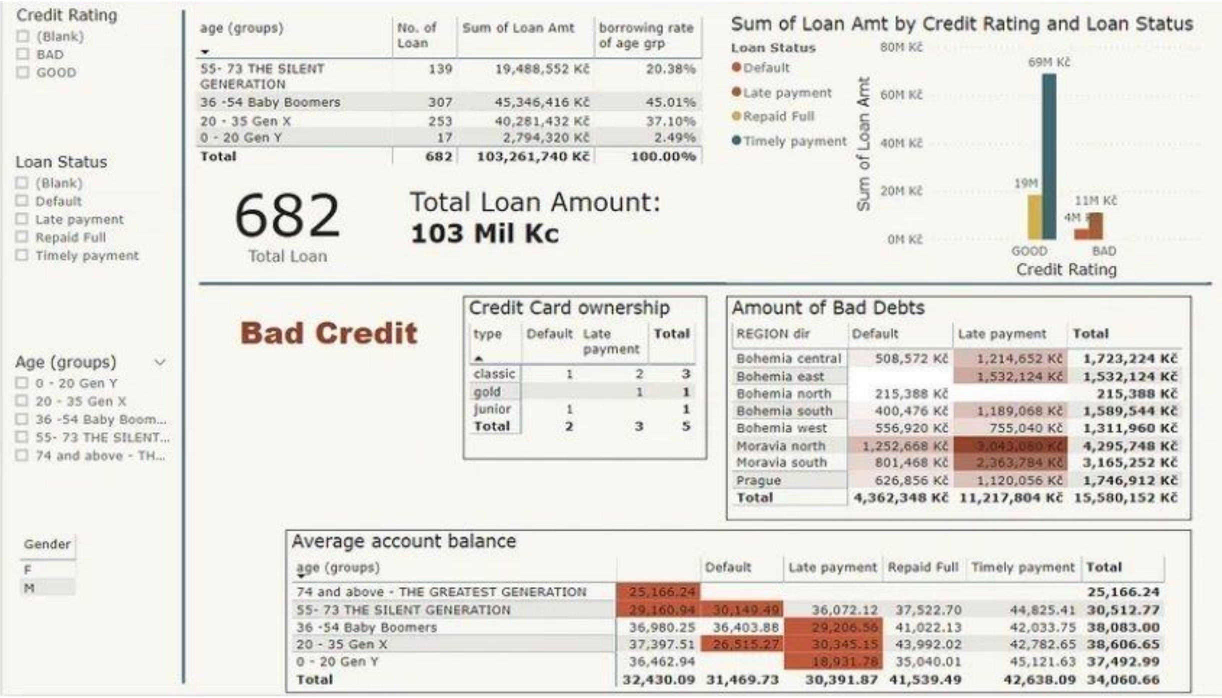
Dashboard



Analysis of Commercial Electricity in Indian States
Top 10 Power Consuming States from All the Regions
Distribution of Power Consumption by Region, Latitude and Longitude
Distribution of Power Consumption
Distribution of Power Consumption by Region ...
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



Population of Czech Republic
Population and Average Salary by region
REGION dir
Region
Customer's age group
Acct holder
DISPONENT
OWNER
Gender
F
M


Credit Rating
(Blank)
BAD
GOOD
Loan Status
(Blank)
Default
Late payment
Repaid Full
Timely payment
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
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%
682
Total Loan
Total Loan Amount:
103 Mil Kc
Sum of Loan Amt by Credit Rating and Loan Status
Loan Status
Default
Late payment
Repaid Full
Timely payment
Sum of Loan Amt
80M Kč
60M Kč
40M Kč
20M Kč
0M Kč
69M Kč
19M
11M Kč
4M
GOOD
BAD
Credit Rating
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
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

In conclusion, our research paper has successfully developed a consumption and prediction model that can be utilized to forecast electricity consumption for the next few months. The model, developed using the Prophet algorithm, has shown reasonable forecasting accuracy, which can be used to draw inferences about the demand-side increase and to help bolster the supply side of the equation. We have also developed a dashboard using the Power BI tool that can provide insights into the electricity consumption patterns, such as the top 10 consumers, the top 5 values and counts of months by year and users, and the average values by user. These insights can be leveraged to strategize policies and optimize energy usage in homes. Although the current model has been developed using a limited dataset, we have shown that it can perform even better with more data. Overall, our research provides a strong foundation for future studies on electricity consumption prediction and optimization.

FUTURE SCOPE

Future scenarios of sectoral value-added and overall and per capita GDP, are used to determine service demands in certain sectors, that in turn determine the employment of various appliances and equipment to convert electricity into end-use services (such as water pumped, steel produced, lighting and air conditioning, etc.). The government over successive years has prioritised the manufacturing sector to draw surplus labor from primary activities and enhance growth, productivity and meaningful employment. This has reflected in the National Manufacturing Policy (NMP) as well as the more recently launched Make in India (MII) campaign. Among key priorities has been increasing the share of manufacturing from 16-17 percent at present to 25 percent of GDP (Bhattacharjee, 2015). If this is achieved by 2030, then industry will need to grow at 1.5 percentage points higher than overall GDP, assuming services retains its current 60 percent share.

REFERENCES

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

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

