OPTIMISING ENERGY DEMAND AND CONSUMPTION THROUGH DATA-DRIVEN STRATEGIES PHASE-V

COLLEGE CODE: 8100

COLLEGE NAME: UNIVERSITY COLLEGE OF ENGINEERING(BITCAMPUS), TRICHY-620024

TECHNOLOGY: DATA ANALYTICS

TEAM NAME:

STUDENT DETAILS :

- HARIHARAN V
- ROHITH JOHANAN D
- **MOHAN R**
- PRAVEEN KRISHNA R
- ARUN T

SUBMITTED BY

MOHAN R

NM ID: au810022104002

Table of Contents

HYPER LINK FOR ACCESSING ALL PDF'S AND TO DOWNLOAD ---> click here

Abstract :

Due to factors like urbanization, technological advancements, and environmental concerns, the current landscape of energy consumption and demand is undergoing significant changes. In this regard, maximizing energy use and achieving sustainability objectives require leveraging the power of data-driven strategies. The objective of this project is to investigate how patterns of consumption, energy demand, and the use of data-driven approaches for efficient management are interdependent. The project aims to provide insights into energy usage trends, pinpoint possible areas for efficiency gains, and create workable demand-side management strategies by utilizing extensive data analysis, predictive modeling, and algorithmic optimization techniques.

Introduction:

The dynamics of energy demand and consumption in today's quickly changing world offer opportunities as well as challenges to stakeholders in a variety of sectors. Efficient energy use and sustainable consumption practices are more important than ever as urbanization picks up speed and technology advancements continue to transform industries.The project sums up a thorough investigation into the complex relationships between energy demand, consumption trends, and the game-changing possibilities of data-driven approaches. By utilizing algorithmic learning, predictive modeling, and data analytics.

Methodologies:

Data Collecting: Gather relevant datasets encompassing energy consumption records, weather patterns, demographic information, and other pertinent variables.

Data cleansing: Cleanse and preprocess the data to handle missing values, outliers, and inconsistencies.Normalize or scale the data to ensure compatibility across different features.

Exploratory Data Analysis (EDA): Conduct EDA to gain insights into the distribution, trends, and correlations within the energy consumption data.Visualize key

metrics such as energy usage patterns over time, peak demand periods, and geographical variations.Identify potential factors influencing energy demand, such as seasonal trends, economic activities, and population dynamics.

Predictive Modeling: Develop predictive models to forecast future energy demand based on historical consumption data and relevant predictors.Utilize time series forecasting techniques such as ARIMA (Auto Regressive Integrated Moving Average) or machine learning algorithms like LSTM (Long Short-Term Memory) for capturing temporal dependencies.

By integrating these methodologies, the project aims to develop data-driven insights and actionable strategies for optimizing energy demand and consumption, thereby fostering sustainability, resilience, and efficiency in energy management practices.

Existing works:

GridWatch: GridWatch is a web-based platform that provides real-time monitoring and visualization of energy consumption, generation, and grid conditions. It aggregates data from smart meters, IoT sensors, and grid infrastructure to provide stakeholders with insights into energy demand patterns, renewable energy integration, and grid stability.

Energy Management Dashboard: Energy management dashboards offer web-based interfaces for building owners, facility managers, and energy professionals to monitor and analyze energy usage in real-time. These platforms provide interactive visualizations, customizable reports, and actionable insights to optimize energy consumption, identify inefficiencies, and implement energy-saving measures.

Demand Response Platforms: Demand response platforms enable utilities, energy aggregators, and consumers to participate in demand response programs and virtual power plant (VPP) initiatives through web-based interfaces. These platforms facilitate real-time communication, event notification, and energy curtailment strategies, allowing

participants to adjust their energy usage in response to grid conditions and market signals.

Energy Monitoring and Feedback Tools: Web-based energy monitoring and feedback tools provide consumers with access to personalized energy usage data, recommendations, and comparative benchmarks. These platforms offer interactive dashboards, mobile applications, and social engagement features to empower consumers to track their energy consumption, set energy-saving goals, and receive real-time feedback on their progress.

Renewable Energy Marketplaces: Renewable energy marketplaces leverage web-based platforms to facilitate peer-to-peer trading of renewable energy certificates (RECs), carbon credits, and energy attributes. These platforms connect renewable energy producers with consumers, enabling transparent, secure, and decentralized transactions for renewable energy procurement and carbon offsetting.

Energy Analytics Platforms: Energy analytics platforms offer web-based tools for data analysis, predictive modeling, and optimization of energy consumption across various sectors. These platforms integrate historical data, real-time sensor feeds, and machine learning algorithms to identify energy-saving opportunities, optimize resource allocation, and enhance operational efficiency in buildings, industrial processes, and transportation systems.

Proposed works:

Data-Driven Demand Response Programs: Several studies have explored the implementation of demand response programs leveraging data analytics and machine learning techniques to optimize energy usage and reduce peak demand. Research in this area focuses on developing predictive models for demand forecasting, identifying demand response opportunities, and designing incentive mechanisms to encourage consumer participation.

Smart Grid Technologies: The deployment of smart grid technologies, including advanced metering infrastructure (AMI), smart meters, and IoT-enabled sensors, has revolutionized energy management practices. Existing works emphasize the role of data analytics in enabling real-time monitoring, grid optimization, and demand-side management, thereby enhancing grid reliability and efficiency.

Predictive Analytics for Energy Forecasting: Research efforts in predictive analytics have led to the development of sophisticated models for energy forecasting, encompassing techniques such as time series analysis, machine learning, and deep learning. These models provide valuable insights into future energy demand patterns, enabling utilities and grid operators to optimize resource planning, mitigate risks, and make informed decisions.

Optimization of Energy Consumption Patterns: Optimization algorithms and mathematical modeling techniques play a pivotal role in optimizing energy consumption patterns across various sectors, including residential, commercial, and industrial domains. Existing works focus on developing optimization frameworks for energy-efficient scheduling, load balancing, and resource allocation, considering factors such as cost constraints, environmental objectives, and demand-response dynamics.

Integration of Renewable Energy Sources: With the increasing penetration of renewable energy sources such as solar and wind power, there is a growing need to integrate these intermittent resources into the energy grid efficiently. Research in this area emphasizes the use of data-driven strategies for forecasting renewable energy generation, optimizing grid operation, and managing energy storage systems to ensure reliability and stability.

System Requirements for Energy demand and consumption:

Python: Python is a versatile programming language widely used for data analysis, machine learning, and scientific computing. Libraries such as **Pandas**, **NumPy**,

and **SciPy** provide powerful tools for data manipulation, statistical analysis, and visualization

IBM Cognos Analytics is a comprehensive business intelligence and analytics platform . it may be used for energy demand and consumption optimization, it can still play a valuable role in the broader context of data-driven decision-making and analysis within the energy sector.

Visual Studio is a popular integrated development environment (IDE) developed by Microsoft. It provides a comprehensive set of tools for software development, including code editing, debugging, version control, and collaboration features.

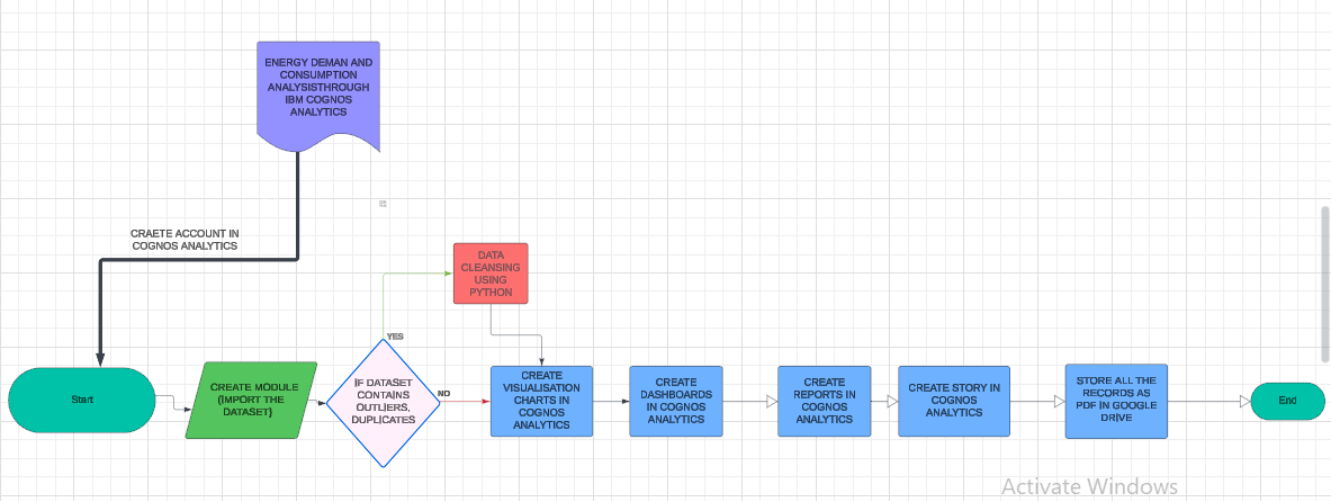
Excel: It is a powerful spreadsheet program developed by Microsoft, widely used for various purposes including data analysis, calculations, and visualization

Google Image Seeking: A modern web browser such as Google Chrome, Mozilla Firefox, Safari, or Microsoft Edge.

Internet Connection: Stable and high-speed internet connection for quick image loading and searches.

Visual Studio Code for HTML Web Page Creation: Optional Visual Studio Code extensions for enhanced HTML, CSS, and JavaScript development.

Work Flow Diagram:



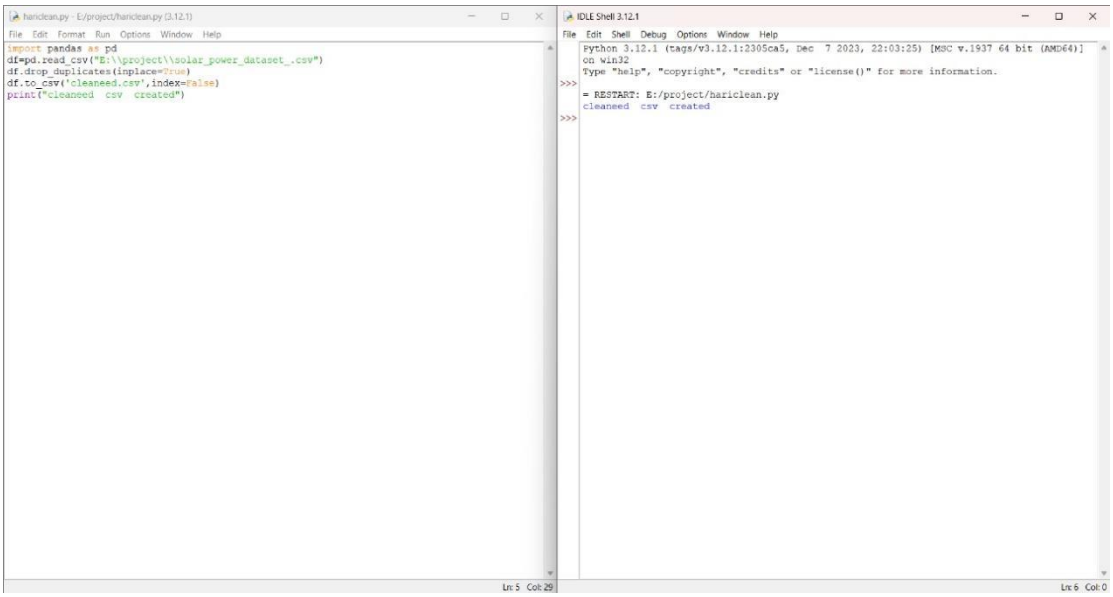
8.DATA COLLECTION AND PRE-PROCESSING

DATA SET FILE LINK:

https://drive.google.com/drive/folders/10rHmJfI9S-lUPhXA1gvpVMNrWnx-t24V?usp=sharing

8(a) Data cleaning:

Data cleaning, often referred to as data cleansing or data scrubbing involves spotting and rectifying errors, discrepancies and inaccuracies, in datasets to enhance their quality and credibility for analysis purposes. This step is vital, in the data preparation process as it guarantees that the data is precise thorough and correctly structured before any analysis or modeling procedures are undertaken. The screenshot **1(a),(b),(c)** shows the **Data cleaning processes**

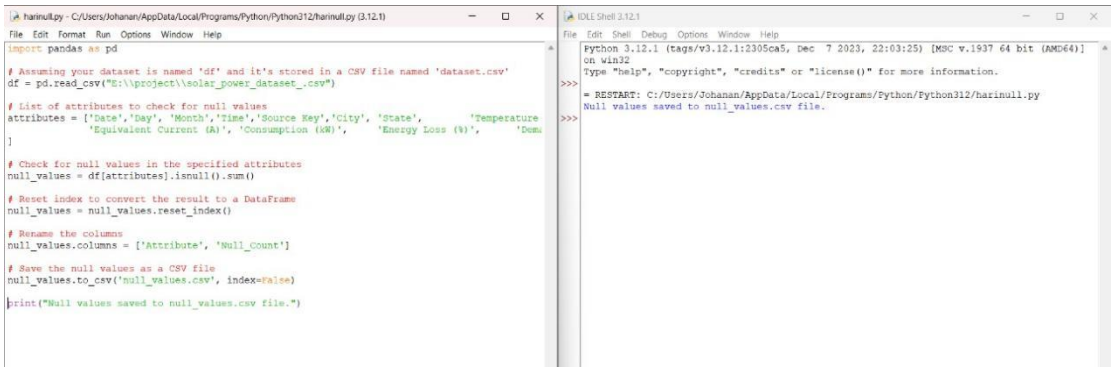


1(a) Data cleaning

8(b) Missing values:

Null values, also known as missing values, are data

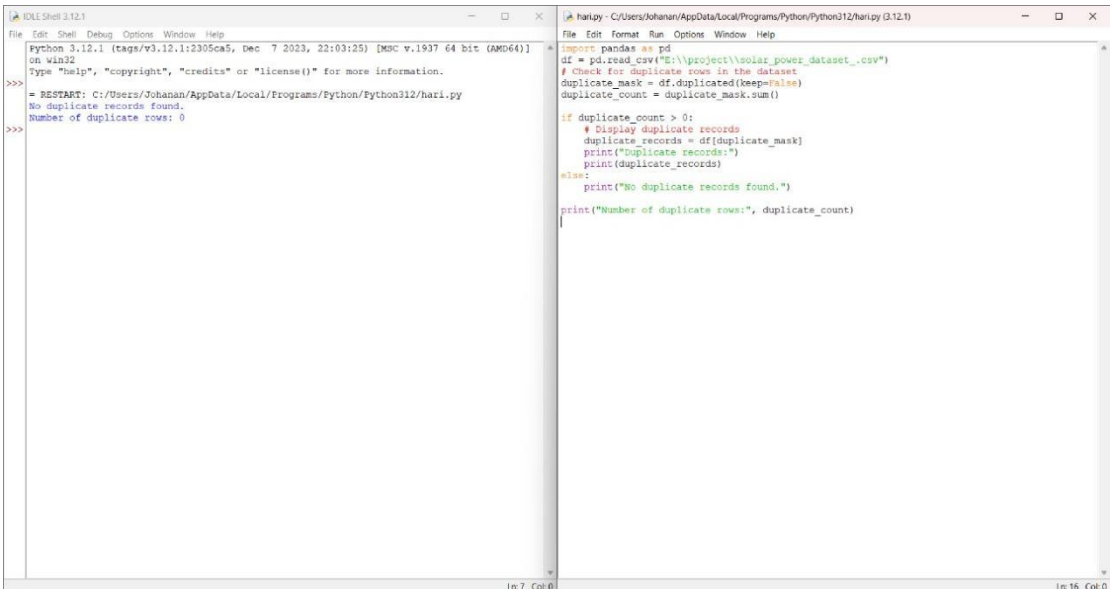
Points that are not present or are undefined in a data set. Dealing with null values is a critical part of the data cleaning process because they can lead to inaccurate analyses or models if not handled properly.



1(b) Missing values

8 C) Outliers:

Outliers are data points that significantly differ from other observations in a dataset. Identifying and appropriately handling outliers is essential in data analysis and modeling to ensure that they do not unduly influence results or conclusions.



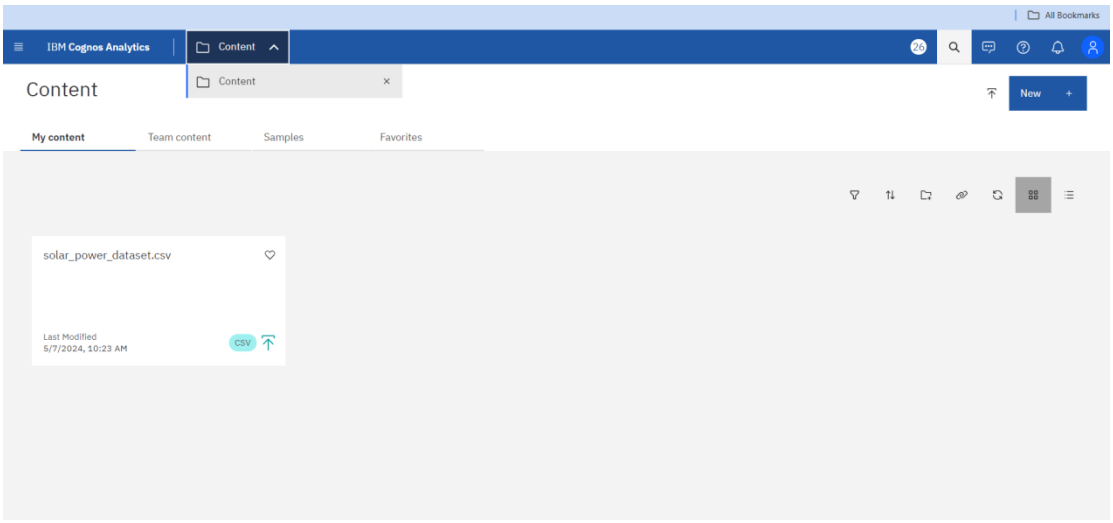
1(c) Outliers

8 D) Data Module creation:

The Data module is a tool that enables users to combine data from sources into one view, for reporting and analysis. In Cognos Analytics converting a data module usually involves adjusting its structure and content to better align with the requirements of reporting or analysis tasks.

Data Module Creation:

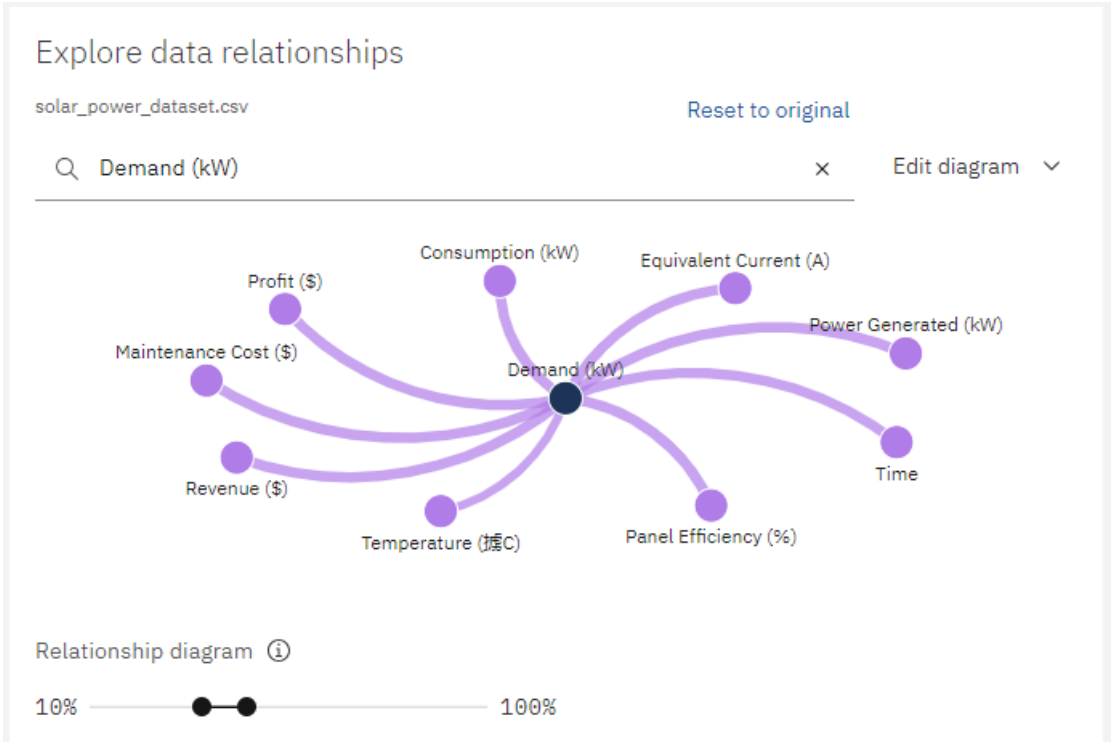
Creating a Data Module in Cognos Analytics allows you to combine and model data from various sources into a single coherent dataset for reporting and analysis



Data Uploading in Cognos Analytics

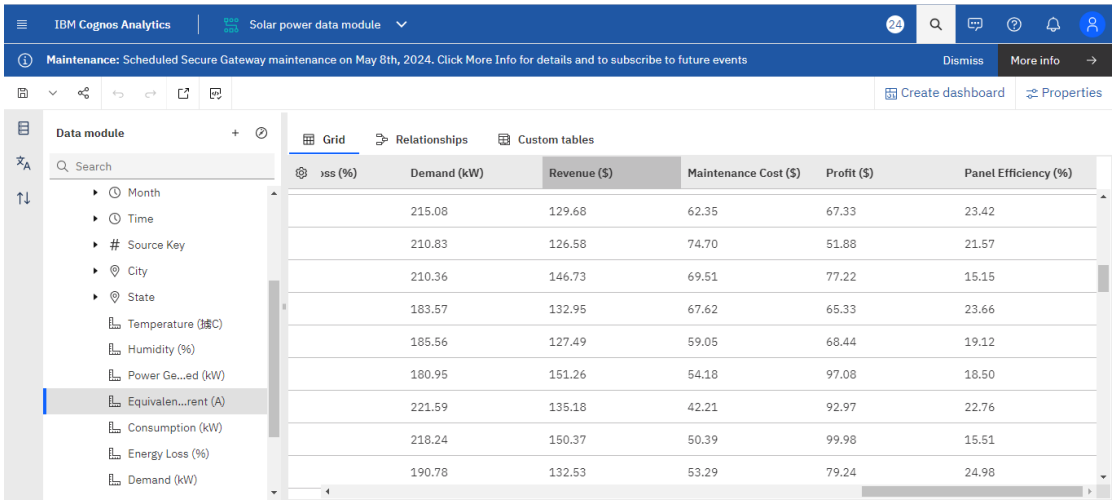
8 E) Data Relationships:

Data relationships are established to define how data from different sources or tables are related to each other. These relationships are crucial for building accurate reports, dashboards, and analyses that integrate data from multiple sources.



8 F) Data Formatting:

Data formatting is the process of defining how data is displayed in reports, dashboards, and visualizations. It allows you to control the appearance of data values, such as numbers, dates, and text, to make them more readable and meaningful to users.



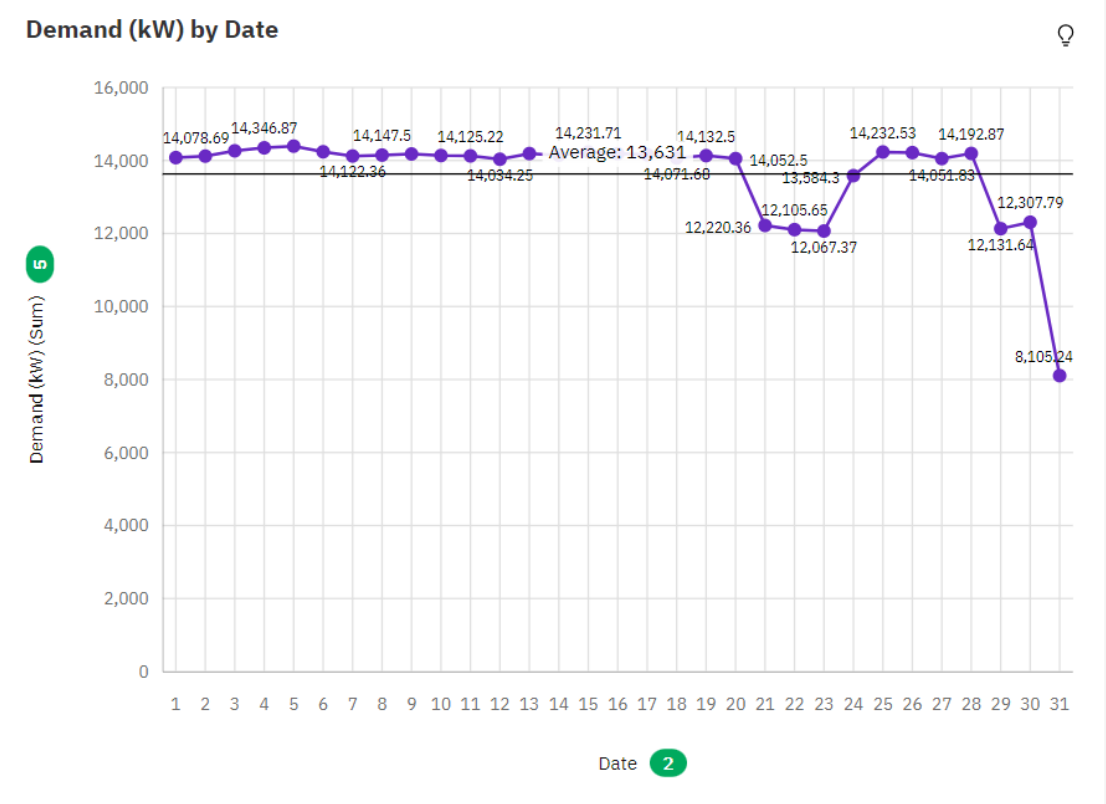
Data Formatting

9.Data Visualization:

DRIVELINK:

https://drive.google.com/drive/folders/1Cr_9Pj4qJepMWlqvNLLOXAL0LVbACDzA?usp=sharing

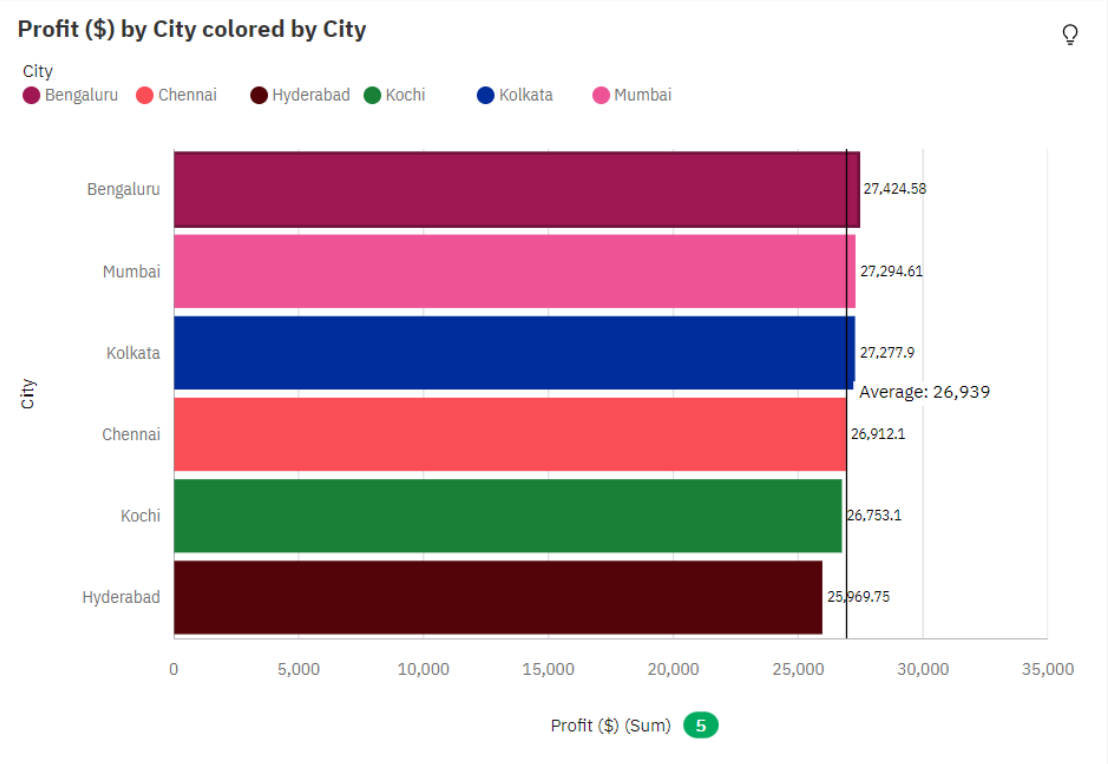
1. Line chart: Demand (kW) by Date



Insights:

- The average value of Demand(kw) is 13,631
- Over all **dates**, the sum of **Demand (kW)** is nearly 423 thousand.
- **Demand (kW)** ranges from over eight thousand, when **Date** is 31, to over fourteen thousand, when **Date** is 5.
- For **Demand (kW)**, the most significant values of **Date** are 5, 4, 3, 6, and 25, whose respective **Demand (kW)** values add up to over 71 thousand, or 16.9 % of the total.

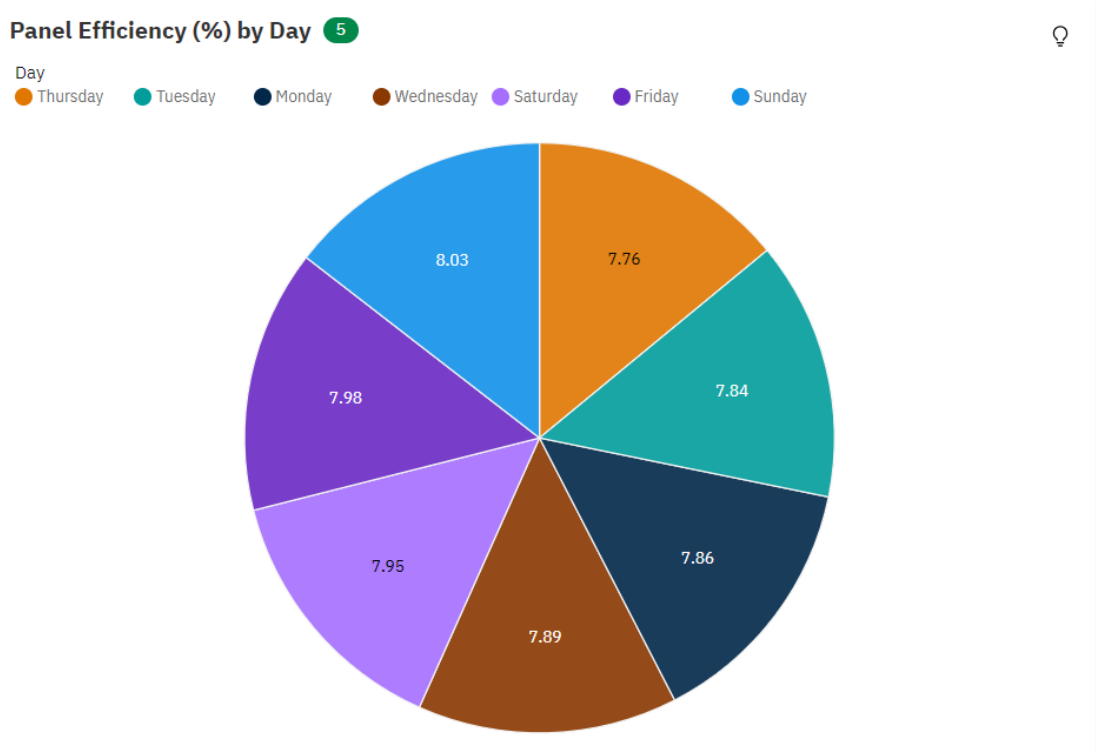
2. Bar chart: Profit ($) by City coloured by City



Insights:

- The average value of profit($) is 26,939
- Over all **cities** and **cities**, the sum of **Profit ($)** is almost 162 thousand.
- The summed values of **Profit ($)** range from nearly 26 thousand to over 27 thousand.
- For **Profit ($)**, the most significant values of **City** are Bengaluru, Mumbai, Kolkata, Chennai, and Kochi, whose respective **Profit ($)** values add up to almost 136 thousand, or 83.9 % of the total.

3. Pie chart: Panel Efficiency (%) by Day



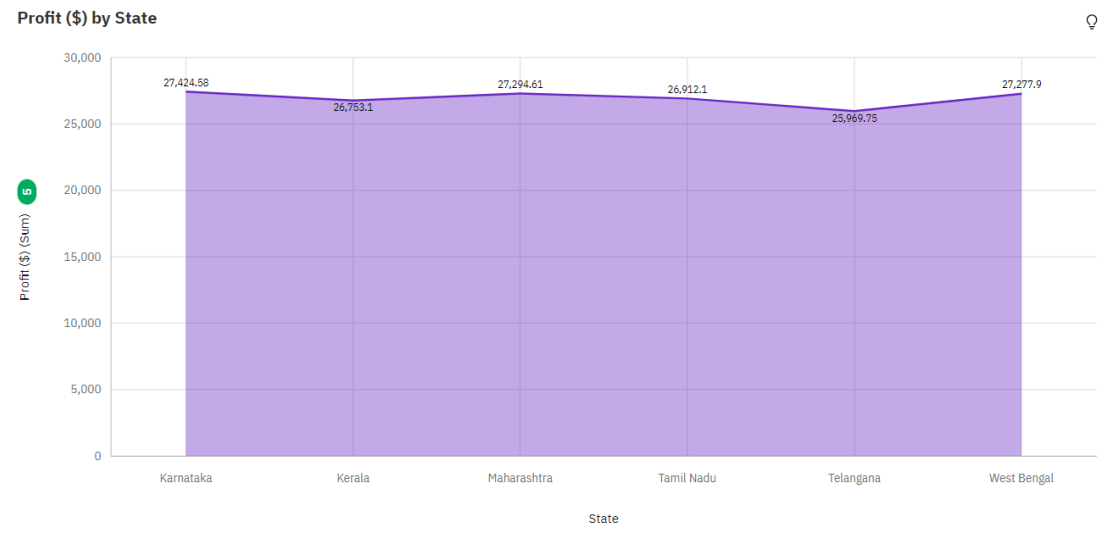
Insights:

- The average value of panel Efficiency(%) is 7.904.
- There is no reliable predictive relationship between Day and panel Efficiency (%).
- Over all **days**, the average of **Panel Efficiency (%)** is 7.904.
- The average values of **Panel Efficiency (%)** range from 7.762, occurring when **Day** is Thursday, to 8.034, when **Day** is Sunday.
- Monday (14.4 %), Sunday (14.4 %), Tuesday (14.4 %), Saturday (14.4 %), and Wednesday (14.3 %) are the most frequently occurring categories of **Day** with a

combined count of 3,596 items with **Panel Efficiency (%)** values (71.9 % of the total)

.

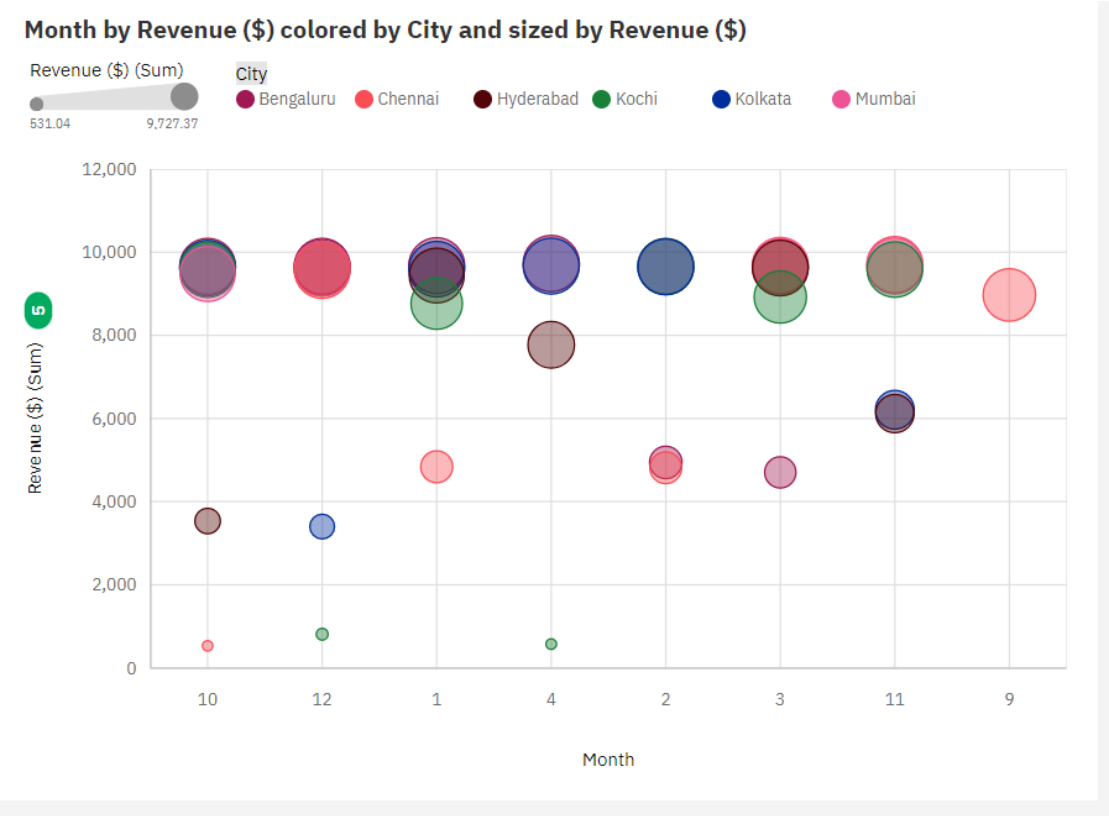
4. Area chart: Profit ($) by State



Insights:

- Over all **states**, the sum of **Profit ($)** is almost 162 thousand.
- **Profit ($)** ranges from nearly 26 thousand, in Telangana, to over 27 thousand, in Karnataka.
- For **Profit ($)**, the most significant values of **State** are Karnataka, Maharashtra, West Bengal, Tamil Nadu, and Kerala, whose respective **Profit ($)** values add up to almost 136 thousand, or 83.9 % of the total.

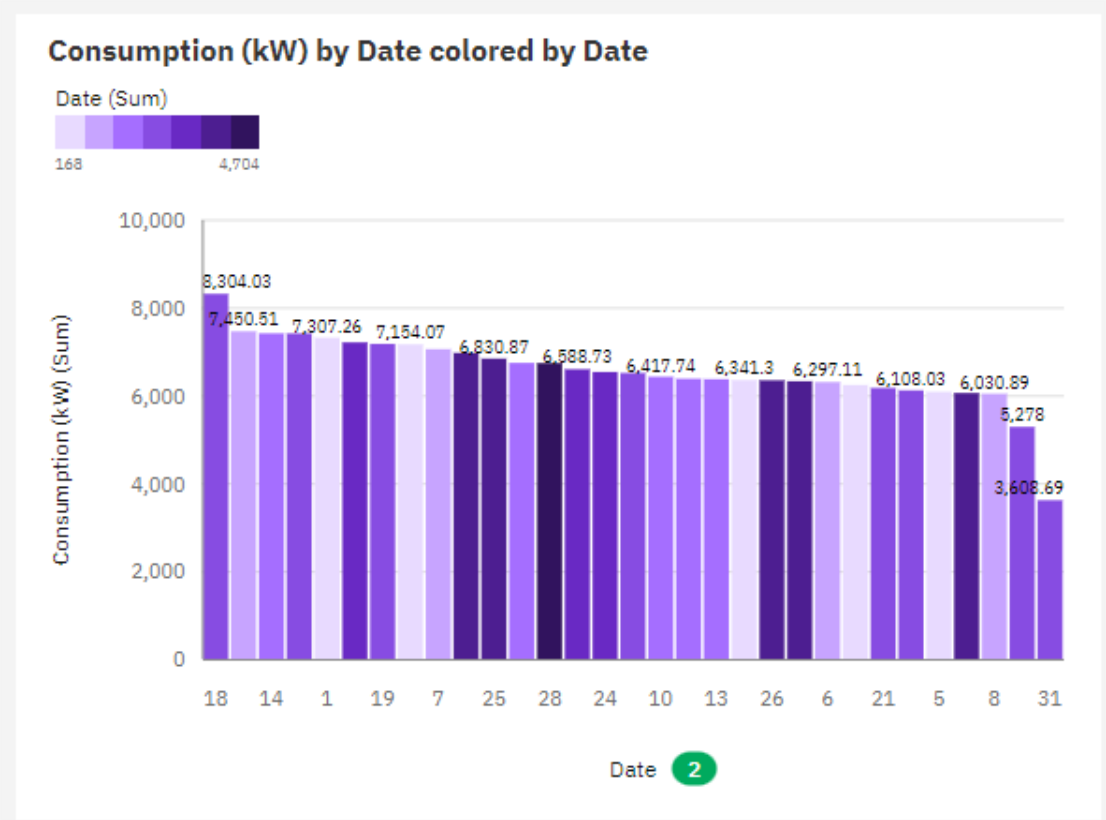
5. Bubble chart: Month by Revenue ($) colored by City and sized by Revenue ($)



Insights:

- Over all **months** and **cities**, the sum of **Revenue ($)** is nearly 287 thousand.
- The summed values of **Revenue ($)** range from 531 to over 9500.
- For **Revenue ($)**, the most significant values of **Month** are 12, 3, 10, 1, and 11, whose respective **Revenue ($)** values add up to over 211 thousand, or 73.7 % of the total.
- For **Revenue ($)**, the most significant values of **City** are Bengaluru, Kolkata, Mumbai, Chennai, and Kochi, whose respective **Revenue ($)** values add up to almost 241 thousand, or 83.9 % of the total.

6. COLUMN CHART: CONSUMPTION (KW) BY DATE COLORED BY DATE



INSIGHTS:

- Across all **dates**, the sum of **Consumption (kW)** is over 203 thousand.**Consumption (kW)** ranges from over 3500, when **Date** is 31, to almost 8500, when Date is 18.

9. DASHBOARD SCREENSHOTS & ANALYSIS:

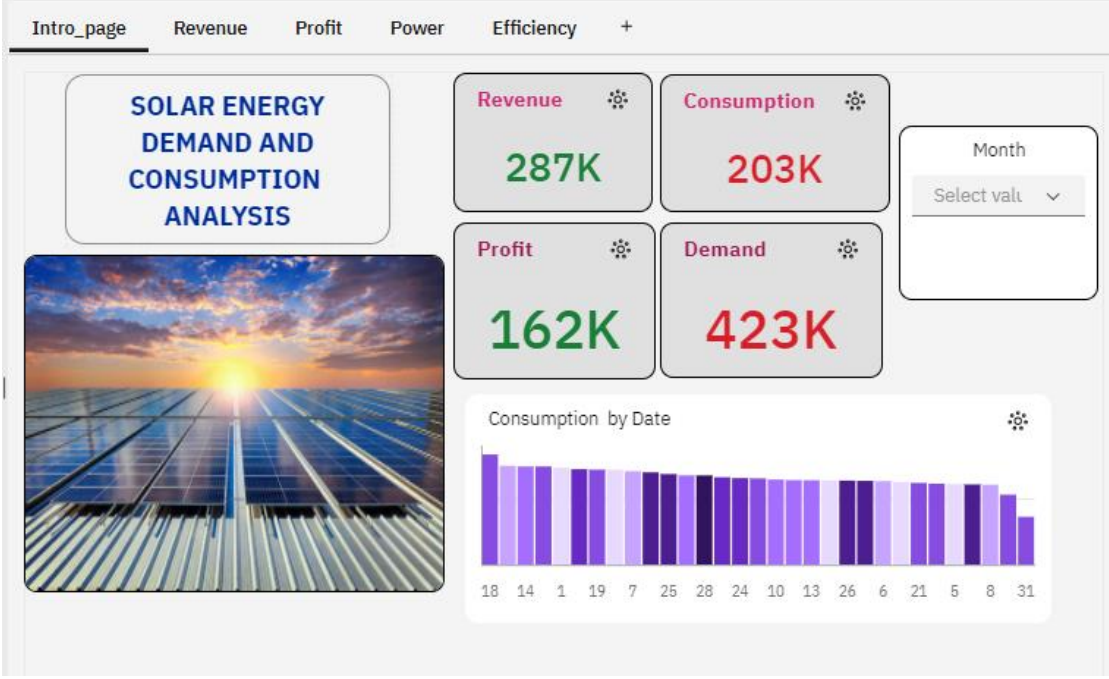
DASHBOARD WEBSITE LINK:

https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FSolar%2BPower%2Bdashboard&action=view&mode=dashboard&subView=model0000018f8ab00012_00000000

DASHBOARD DRIVE LINK:

https://drive.google.com/drive/folders/1ibr0z6cP6RrBJXP96TgHYvtvP9jX8GPN?usp=sharing

1. INTRO_PAGE SCREENSHOT:



REVENUE INSIGHTS:

- Month 9 has the lowest total Revenue ($) at nearly 9 thousand, followed by 4 at almost 28 thousand.
- Month 12 has the highest total Revenue ($) at 42656.42, followed by 3 at 42573.25.

- Add insight to favorites
- The overall number of results for Revenue ($) is five thousand

PROFIT INSIGHTS:

- Hyderabad has the lowest total profit at nearly $26,000, followed closely by Kochi at almost $27,000.
- Bengaluru leads with the highest total profit at $27,424.58, followed by Mumbai at $27,294.61.
- There are a total of 5,000 results for profit**.**

CONSUMPTION INSIGHTS:

- Hyderabad has the lowest total consumption of kilowatts (kW) at over 31,000, followed by Bengaluru at over 33,000.
- Chennai leads with the highest total consumption of kW at over 35,000, followed by Mumbai at almost 35,000.
- There are a total of 5,000 results for consumption**.**

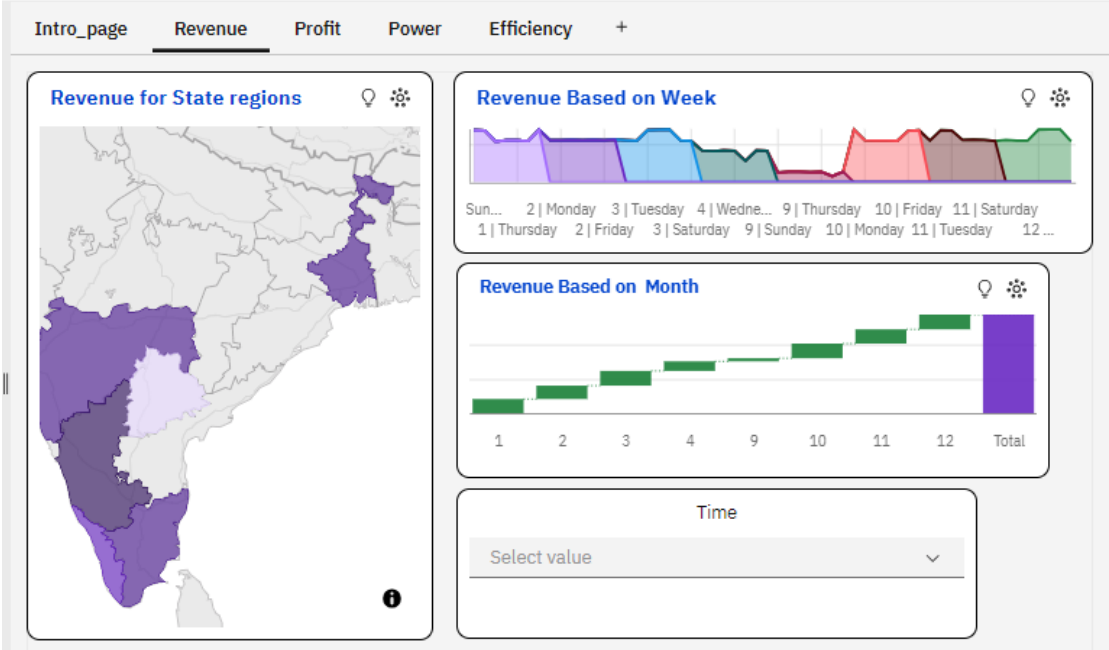
DEMAND INSIGHTS:

- September (Month 9) has the lowest total demand of kW at over 13,000, followed by April (Month 4) at over 40,000.
- March (Month 3) has the highest total demand of kW at over 63,000, followed by January (Month 1) at almost 63,000.
- There are a total of 5,000 results for demand.

CONSUMPTION INSIGHTS:

- Consumption (kW) ranges from over 3,500 when the date is 31, to almost 8,500 when the date is 18.
- Across all dates, the sum of Consumption (kW) is over 203,000.

2. REVENUE DASHBOARD SCREENSHOT:



REVENUE FOR STATE INSIGHTS:`

- Month 12 has the highest revenue at almost $43,000, with the state of Karnataka contributing the most at over $9,500.
- In Month 4, State Karnataka contributed $9,727, whereas West Bengal contributed $9,666.
- State Karnataka has the highest total revenue due to Month 4.
- Revenue ranges from over $46,000 in Telangana to over $48,000 in Karnataka.
- The most significant contributors to revenue are Karnataka, West Bengal, Maharashtra, Tamil Nadu, and Kerala, with their respective revenue values adding up to almost $241,000 or 83.9% of the total.
- Across all states, the sum of revenue is nearly $287,000.

REVENUE FOR WEEK INSIGHTS:`

- Day Friday has the highest Average Humidity (%) but is ranked #6 in Total Revenue ($).

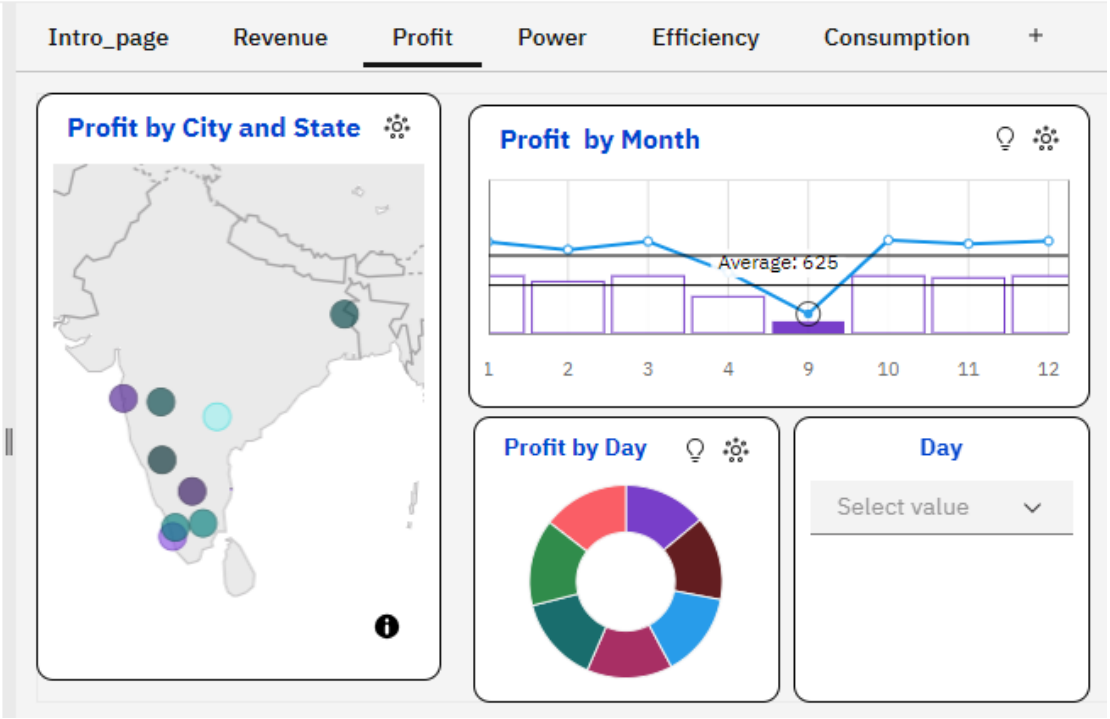
- Month 12 has the highest Revenue ($) at almost 43 thousand, out of which City Bengaluru contributed the most at over 9500.
- Day Wednesday has the highest Total Revenue ($) but is ranked #4 in Average Humidity (%).

- City Bengaluru Revenue ($) from Month 4 is 9,727, whereas Kolkata is only 9,666.
- City Bengaluru has the highest total Revenue ($) due to Month 4.
- The summed values of Revenue ($) range from 809.5 to nearly seven thousand.
- For Revenue ($), the most significant values of Month - Day are 3|Thursday, 12|Thursday, 10|Sunday, 12|Friday, and 12|Wednesday, whose respective Revenue ($) values add up to almost 35 thousand, or 12.1 % of the total.
- For Revenue ($), the most significant values of Month are 12, 3, 10, 1, and 11, whose respective Revenue ($) values add up to over 211 thousand, or 73.7 % of the total.
- Over all month - days and months, the sum of Revenue ($) is nearly 287 thousand.

REVENUE FOR MONTH INSIGHTS:`

- Revenue ($) is unusually low when Month is 9.
- Month 12 has the highest Revenue ($) at almost 43 thousand, out of which City Bengaluru contributed the most at over 9500.
- City Bengaluru Revenue ($) from Month 4 is 9,727, whereas Kolkata is only 9,666.
- City Bengaluru has the highest total Revenue ($) due to Month 4.
- Revenue ($) ranges from nearly nine thousand, when Month is 9, to almost 43 thousand, when Month is 12.
- Over all months, the sum of Revenue ($) is nearly 287 thousand.

3. PROFIT DASHBOARD SCREENSHOT:



PROFIT BY CITY INSIGHTS:

- City Bengaluru has the highest total profit ($) due to Month 4.
- It also has the highest values of both profit ($) and state.
- Month 10 has the highest profit ($) at over $24,000, with City Kolkata contributing the most at over $5,500.
- Kolkata has a profit ($) of over $5,500 for Month 10.
- Profit ($) ranges from nearly $26,000 in Hyderabad to over $27,000 in Bengaluru.
- The most significant values of profit ($) by city are Bengaluru, Mumbai, Kolkata, Chennai, and Kochi, whose respective profit ($) values add up to almost $136,000, or 83.9% of the total.
- Across all cities, the sum of profit ($) is almost $162,000.

PROFIT BY STATE INSIGHTS:

- State Karnataka has the highest total profit ($) due to Month 4.
- Month 10 has the highest profit ($) at over $24,000, with State West Bengal contributing the most at over $5,500.
- West Bengal has a profit ($) of over $5,500 for Month 10.

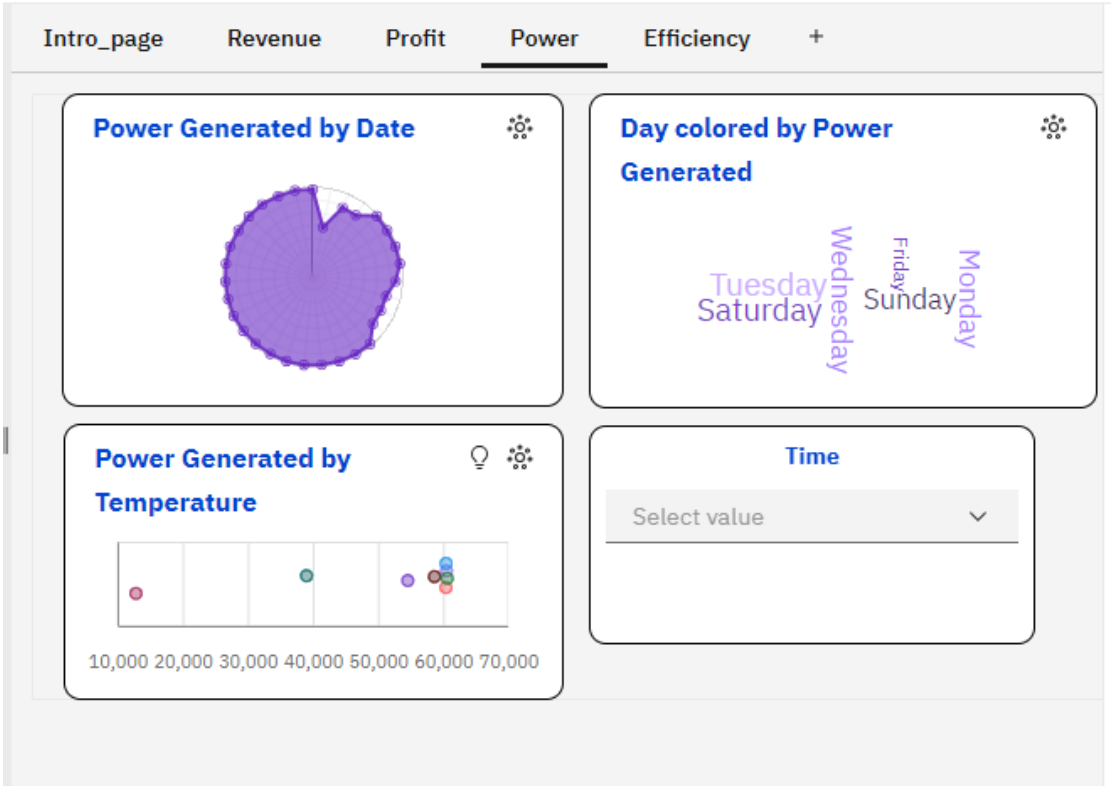
PROFIT BY MONTH INSIGHTS:

- The total number of results for Month, across all months, is five thousand.
- City Bengaluru has the highest total Profit ($) due to Month 4.
- Profit ($) is unusually low when Month is 9.
- Month 10 has the highest Profit ($) at over 24 thousand, out of which City Kolkata contributed the most at over 5500.
- Kolkata has a Profit ($) of over 5500 for Month 10.
- Profit ($) ranges from over five thousand, when Month is 9, to over 24 thousand, when Month is 10.
- For Profit ($), the most significant values of Month are 10, 12, 3, 1, and 11, whose respective Profit ($) values add up to over 119 thousand, or 73.6 % of the total.
- Over all months, the sum of Profit ($) is almost 162 thousand.

PROFIT BY DAY INSIGHTS:

- Based on the current forecasting, Profit ($) may reach over 22 thousand by Day Monday+1.
- Profit ($) ranges from almost 23 thousand, when Day is Thursday, to over 23 thousand, when Day is Saturday.
- For Profit ($), the most significant values of Day are Saturday, Sunday, Monday, Wednesday, and Tuesday, whose respective Profit ($) values add up to over 116 thousand, or 72 % of the total.
- Over all days, the sum of Profit ($) is almost 162 thousand.

4. POWER GENERATED DASHBOARD SCREENSHOT:



POWER GENERATED BY DATE INSIGHTS:

- It is projected that by Monday+1, 16 will exceed 2 in Power Generated (kW) by 24.

POWER GENERATED BY TEMPERATURE INSIGHTS:

- Month 2 has the highest Average Temperature but is ranked #6 in Total Power Generated (kW).
- 2 Month accounted for 20% of Kolkata Power Generated (kW) compared to 10% for Chennai.
- 2 Month accounted for 20% of Kolkata Power Generated (kW) compared to 10% for Chennai.
- Month 12 has the highest Power Generated (kW) at almost 61 thousand, out of which City Chennai contributed the most at nearly 14 thousand.
- Month 11 has the highest Total Power Generated (kW) but is ranked #2 in Average Temperature .
- City Kolkata has the highest total Power Generated (kW) due to Month 2.
- There is no reliable relationship between Temperature and Power Generated (kW).

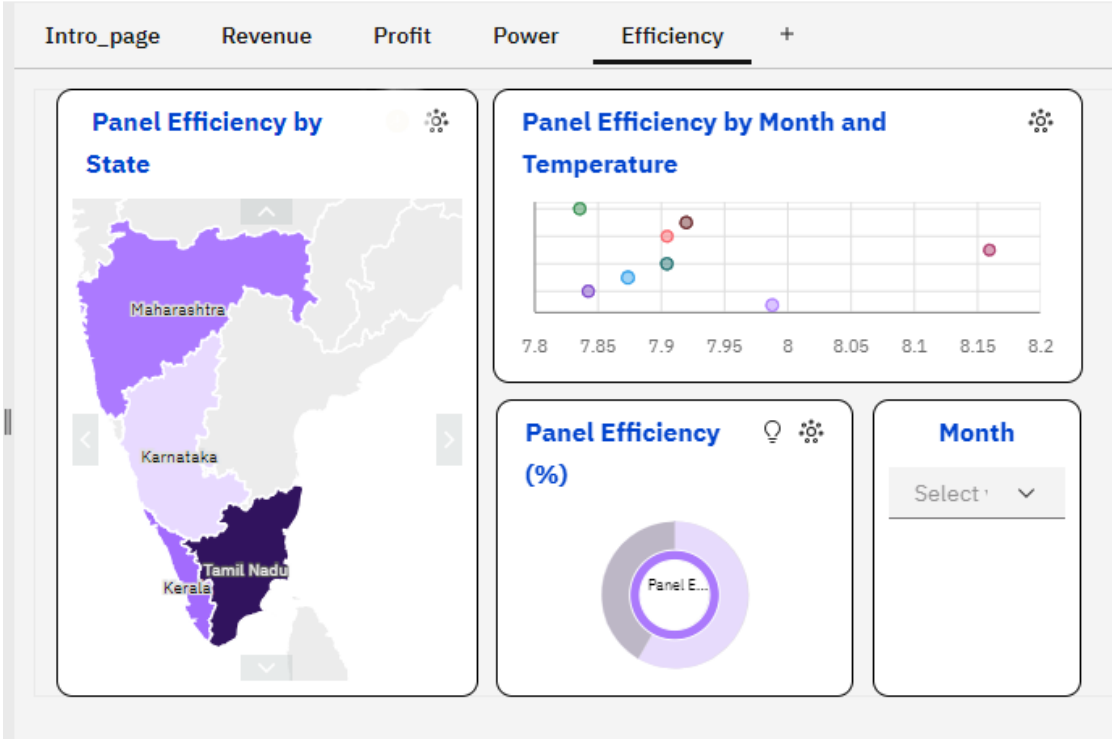
- Chennai has a Power Generated (kW) of nearly 14 thousand for Month 12.
- 11 (16.7 %), 3 (16.7 %), 1 (13.3 %), 12 (13.3 %), and 2 (13.3 %) are the most frequently occurring categories of Month with a combined count of 528 items with Temperature values (73.3 % of the total) .
- Over all months, the average of Temperature is 32.61.
- The average values of Temperature range from 31.81, occurring when Month is 10, to 33.01, when Month is 2.

POWER GENERATED BY DAY INSIGHTS:

- Based on the current forecasting, Power Generated (kW) may reach almost 58 thousand by Day Monday+1.
- It is projected that by Monday+1, Hyderabad will exceed Bengaluru in Date by 113.
- Monday (14.4 %), Sunday (14.4 %), Tuesday (14.4 %), Saturday (14.4 %), and Wednesday (14.3 %) are the most frequently occurring categories of Day with a combined count of 3,596 items with Date values (71.9 % of the total) .
- Monday (14.4 %), Sunday (14.4 %), Tuesday (14.4 %), Saturday (14.4 %), and Wednesday (14.3 %) are the most frequently occurring categories of Day with a combined count of 3,596 items with Power Generated (kW) values (71.9 % of the total) .
- Over all days, the average of Date is 15.55.
- Over all days, the average of Power Generated (kW) is 81.44.
- The total number of results for Date, across all days, is five thousand.
- The total number of results for Power Generated (kW), across all days, is five thousand.
- The overall number of results for Panel Efficiency (%) is five thousand.
- Date ranges from nearly 11 thousand, when Day is Friday, to nearly 12 thousand, when Day is Sunday.

- Power Generated (kW) ranges from almost 57 thousand, when Day is Thursday, to almost 59 thousand, when Day is Tuesday.

5. PANEL EFFICIENCY DASHBOARD SCREENSHOT:



PANEL EFFICIENCY BY STATE INSIGHTS:

- Month 12 has the highest Panel Efficiency (%) at 48.73, out of which State Kerala contributed the most at 9.672**.**
- State Kerala Panel Efficiency (%) from Month 12 is 9.672, whereas Tamil Nadu is only 7.654.
- State Kerala has the highest average Panel Efficiency (%) due to Month 12.
- Karnataka (16.8 %), Tamil Nadu (16.8 %), West Bengal (16.8 %), Maharashtra (16.8 %), and Kerala (16.8 %) are the most frequently occurring categories of State with a combined count of 4,200 items with Panel Efficiency (%) values (84 % of the total) .
- Over all states, the average of Panel Efficiency (%) is 7.904.
- The average values of Panel Efficiency (%) range from 7.781, occurring in Karnataka, to 8.018, in Tamil Nadu.

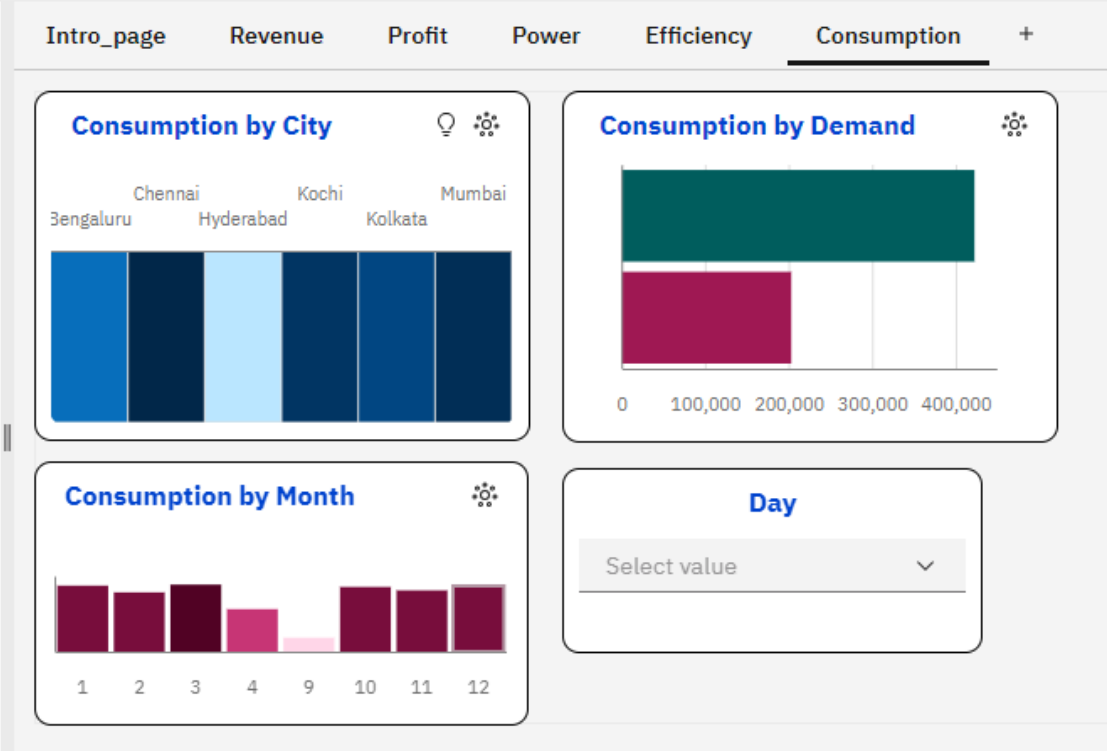
PANEL EFFICIENCY BY MONTH AND TEMPERATURE INSIGHTS:

- Month 9 has the highest Average Panel Efficiency (%) but is ranked #8 in Average Temperature .
- Month 3 has the highest Average Temperature but is ranked #6 in Average Panel Efficiency (%).
- Month 12 has the highest Panel Efficiency (%) at 48.73, out of which City Kochi contributed the most at 9.672.
- City Kochi Panel Efficiency (%) from Month 12 is 9.672, whereas Chennai is only 7.654.
- City Kochi has the highest average Panel Efficiency (%) due to Month 12.
- The overall number of results for Month is 8.
- The overall number of results for Panel Efficiency (%) is five thousand.
- The overall number of results for Temperature is five thousand.

PANEL EFFICIENCY INSIGHTS:

- Day Thursday has the lowest average Panel Efficiency (%) at 7.76, followed by Tuesday at 7.84.
- Day Sunday has the highest average Panel Efficiency (%) at 8.03, followed by Friday at 7.98.
- Based on the current forecasting, Panel Efficiency (%) may reach 7.859 by Day Monday+1.

6. CONSUMPTION DASHBOARD SCREENSHOT :



CONSUMPTION BY CITY INSIGHTS:

- 12 Month accounted for 22% of Chennai Consumption (kW) compared to 20% for Mumbai.
- Month 3 has the highest Consumption (kW) at nearly 31 thousand, out of which City Mumbai contributed the most at nearly 8 thousand.
- City Chennai has the highest total Consumption (kW) due to Month 12.
- Mumbai has a Consumption (kW) of nearly eight thousand for Month 3.
- Consumption (kW) ranges from over 31 thousand, in Hyderabad, to over 35 thousand, in Chennai.
- For Consumption (kW), the most significant values of City are Chennai, Mumbai, Kochi, Kolkata, and Bengaluru, whose respective Consumption (kW) values add up to almost 172 thousand, or 84.6 % of the total.
- Across all cities, the sum of Consumption (kW) is over 203 thousand.

CONSUMPTION BY DEMAND INSIGHTS:

- Month 3 has the highest values of both Demand (kW) and Consumption (kW).
- The overall number of results for Consumption (kW) is five thousand.
- The overall number of results for Demand (kW) is five thousand.

CONSUMPTION BY MONTH INSIGHTS:

- 12 Month accounted for 22% of Chennai Consumption (kW) compared to 20% for Mumbai.
- Month 3 has the highest Consumption (kW) at nearly 31 thousand, out of which City Mumbai contributed the most at nearly 8 thousand.
- City Chennai has the highest total Consumption (kW) due to Month 12.
- Mumbai has a Consumption (kW) of nearly eight thousand for Month 3.
- Consumption (kW) ranges from nearly 6500, when Month is 9, to nearly 31 thousand, when Month is 3.
- For Consumption (kW), the most significant values of Month are 3, 1, 12, 10, and 11, whose respective Consumption (kW) values add up to nearly 150 thousand, or 73.7 % of the total.
- Across all months, the sum of Consumption (kW) is over 203 thousand.

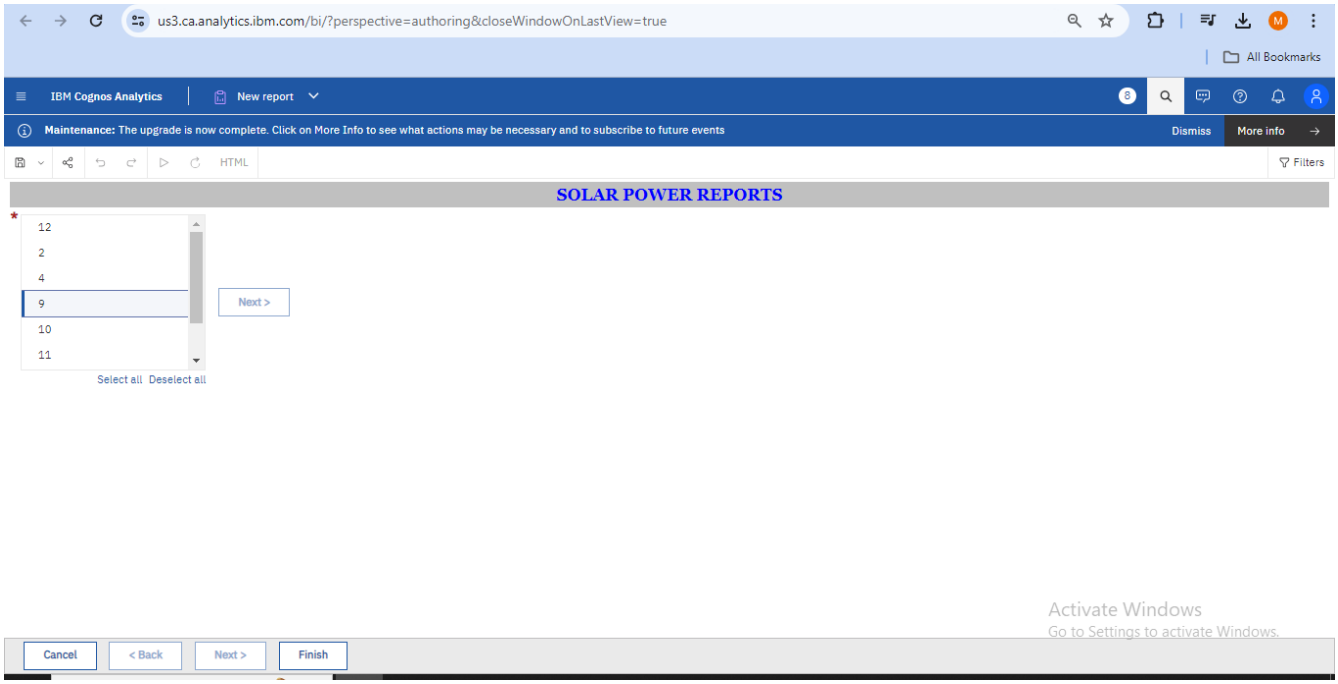
10. REPORTS LINK:

https://us3.ca.analytics.ibm.com/bi/?pathRef=.my_folders%2Fsolar%2Bpower%2Breport&action=edit

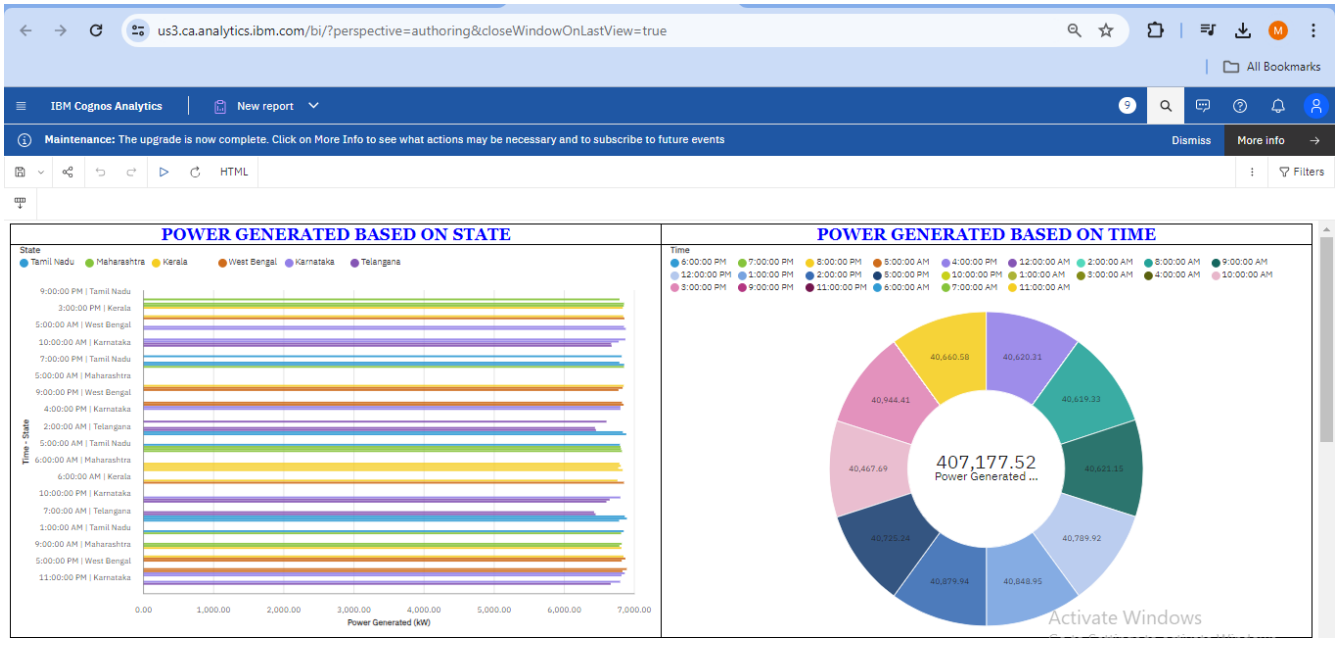
REPORT DRIVE PDF LINK:

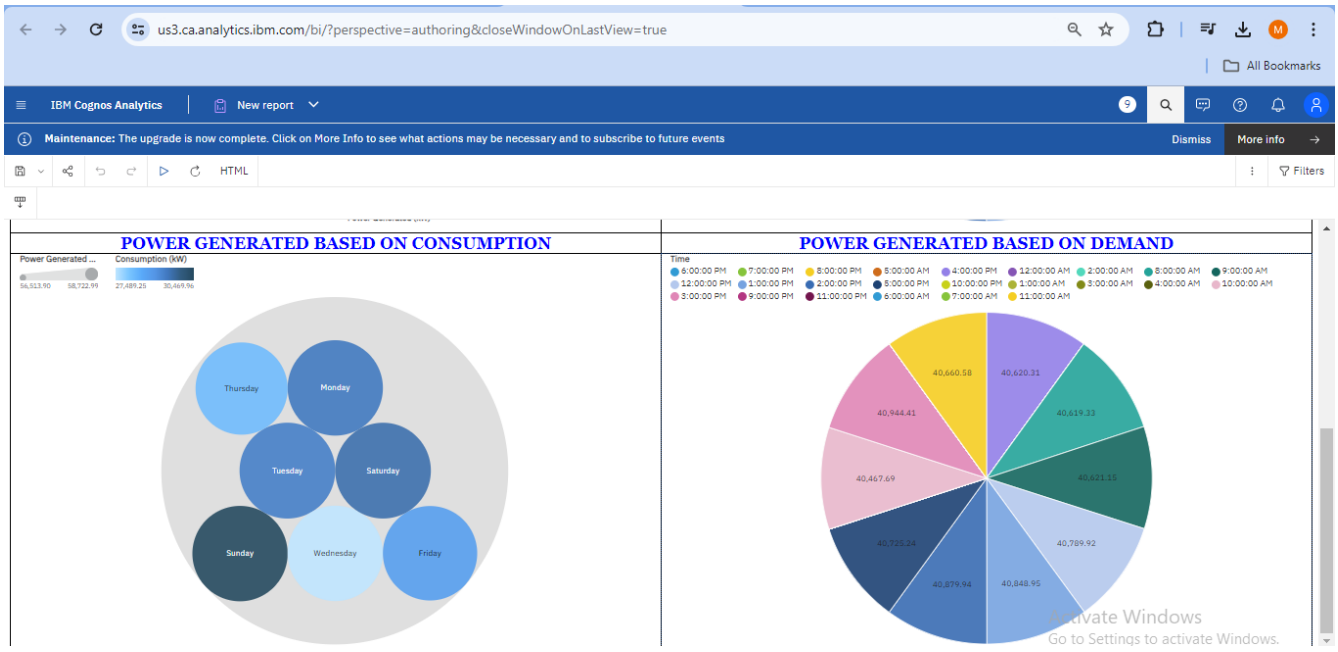
https://drive.google.com/drive/folders/1KCDODl0QB_8mRxd2S2ape2OPYHDD7hU7?usp=sharing

PROMPT PAGE REPORT BY MONTHS:

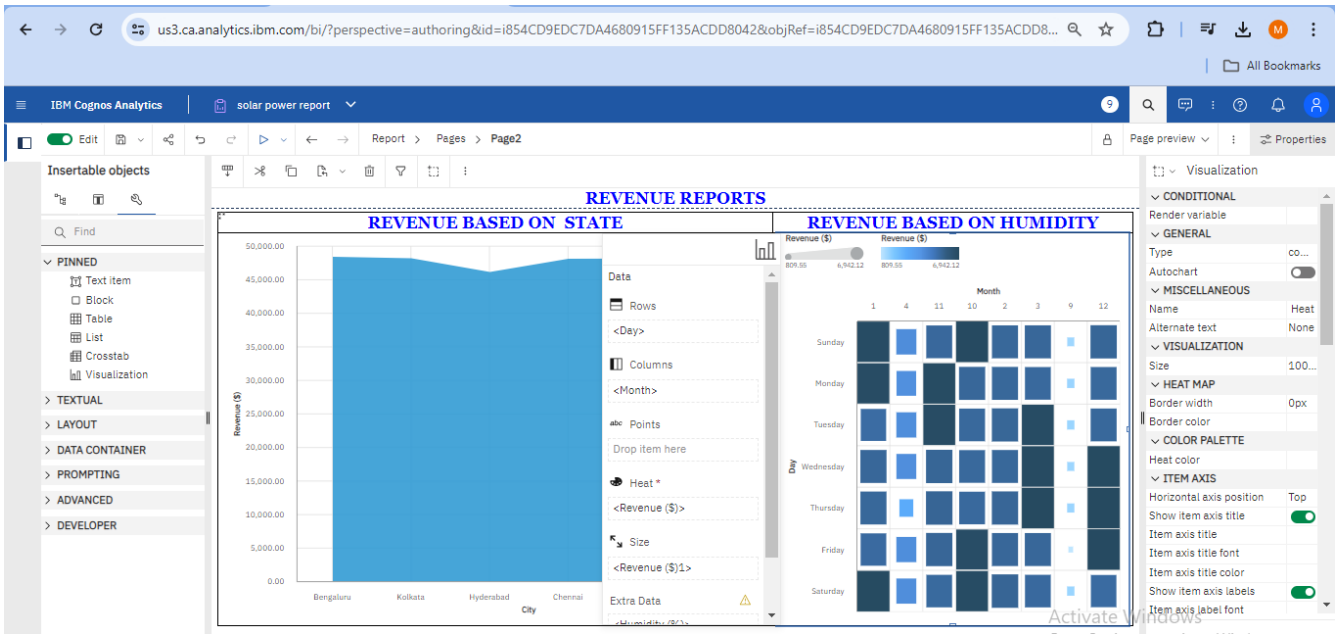


POWER GENERATED REPORTS:

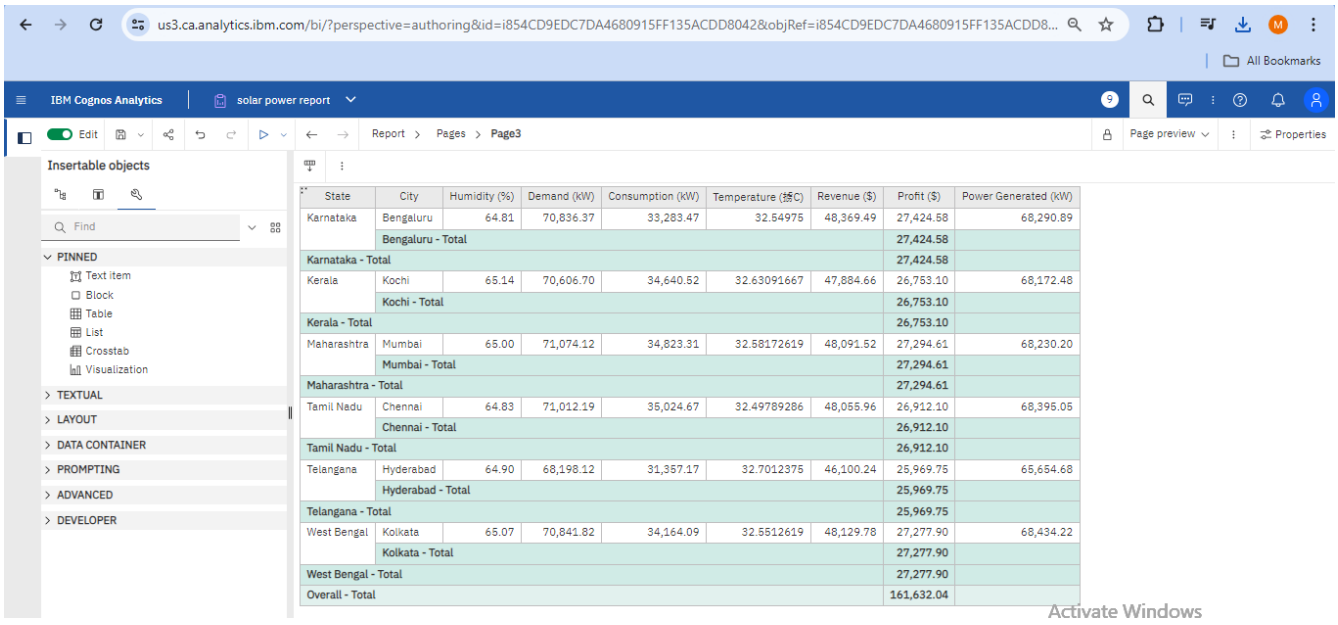




REVENUE REPORTS:



STATE, CITY AND HUMIDITY GROUPED REPORTS:



State	City	Humidity (%)	Demand (kW)	Consumption (kW)	Temperature (℃)	Revenue ($)	Profit ($)	Power Generated (kW)
Karnataka	Bengaluru	64.81	70,836.37	33,283.47	32.54975	48,369.49	27,424.58	68,290.89
	Bengaluru - Total						**27,424.58**	
Karnataka - Total							**27,424.58**	
Kerala	Kochi	65.14	70,606.70	34,640.52	32.63091667	47,884.66	26,753.10	68,172.48
	Kochi - Total						**26,753.10**	
Kerala - Total							**26,753.10**	
Maharashtra	Mumbai	65.00	71,074.12	34,823.31	32.58172619	48,091.52	27,294.61	68,230.20
	Mumbai - Total						**27,294.61**	
Maharashtra - Total							**27,294.61**	
Tamil Nadu	Chennai	64.83	71,012.19	35,024.67	32.49789286	48,055.96	26,912.10	68,395.05
	Chennai - Total						**26,912.10**	
Tamil Nadu - Total							**26,912.10**	
Telangana	Hyderabad	64.90	68,198.12	31,357.17	32.7012375	46,100.24	25,969.75	65,654.68
	Hyderabad - Total						**25,969.75**	
Telangana - Total							**25,969.75**	
West Bengal	Kolkata	65.07	70,841.82	34,164.09	32.5512619	48,129.78	27,277.90	68,434.22
	Kolkata - Total						**27,277.90**	
West Bengal - Total							**27,277.90**	
Overall - Total							**161,632.04**	

11. STORY IN COGNOS ANALYTICS:

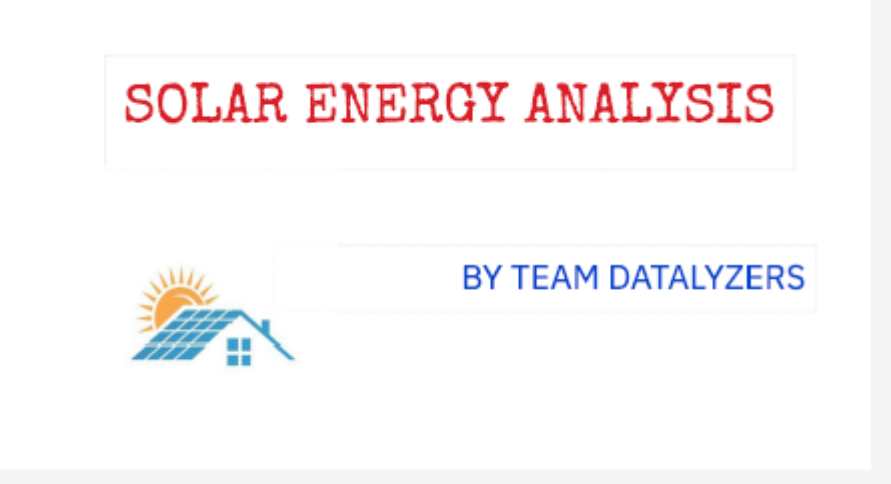
STORY WEBSITE LINK:

https://us3.ca.analytics.ibm.com/bi/?perspective=story&pathRef=.my_folders%2FSolar%2BEnergy%2Bstory&action=view&sceneId=-1&sceneTime=0

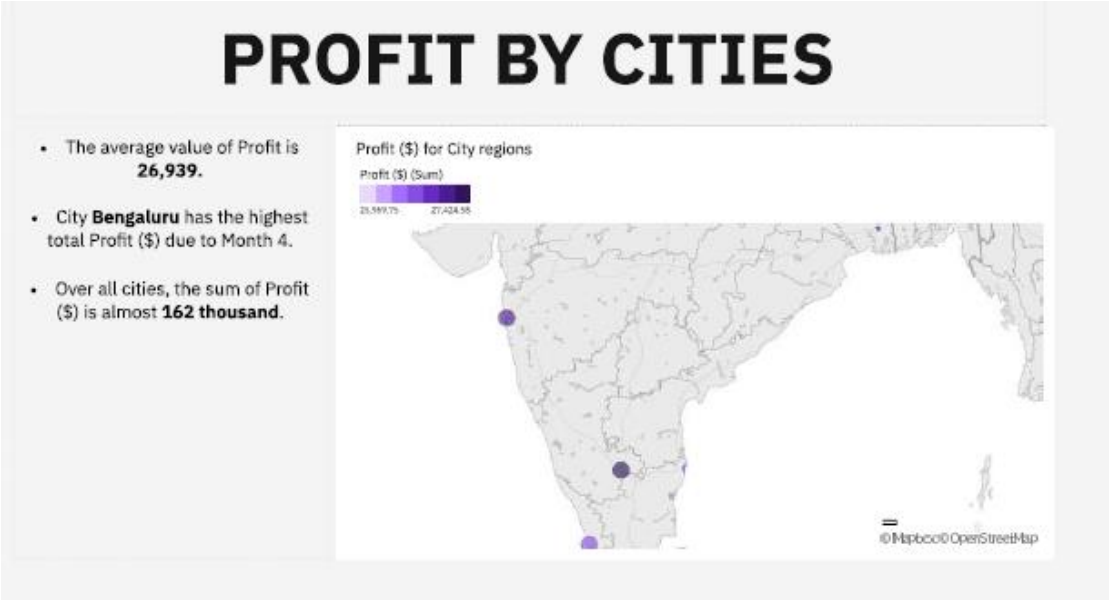
STORY IN DRIVE PDF LINK:

https://drive.google.com/drive/folders/1Jp92CvroNozxwyxwQdVyZFjnsQMQEoZg?usp=sharing

INTRO SCENE:

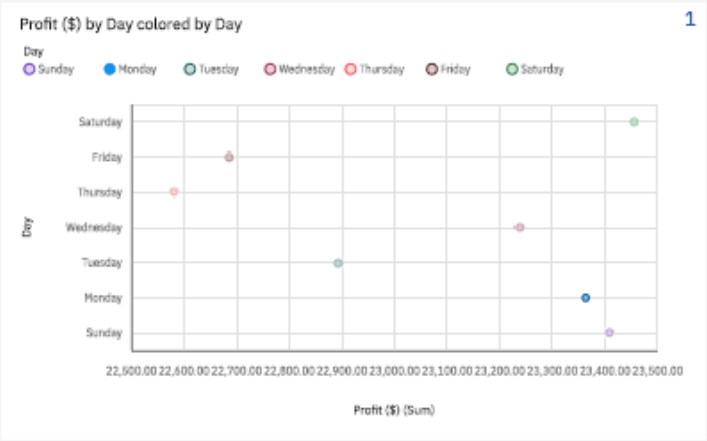


PROFIT SCENES:



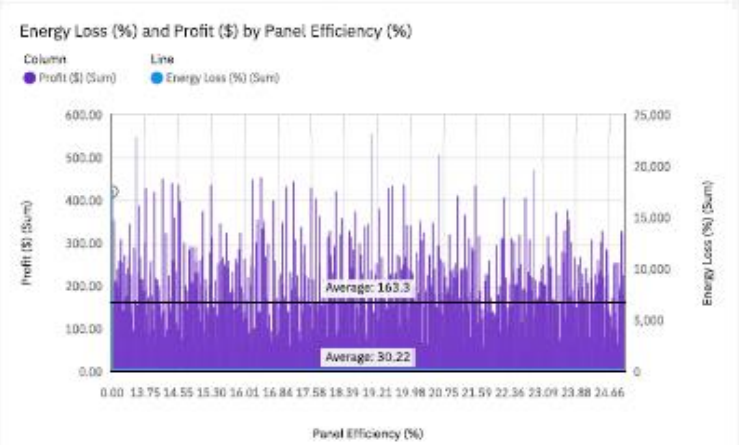
PROFIT BASED ON DAYS

- Based on the current forecasting, Profit ($) may reach over 22 thousand by **Day Monday+1**.
- The overall number of results for **Day** is **7**.

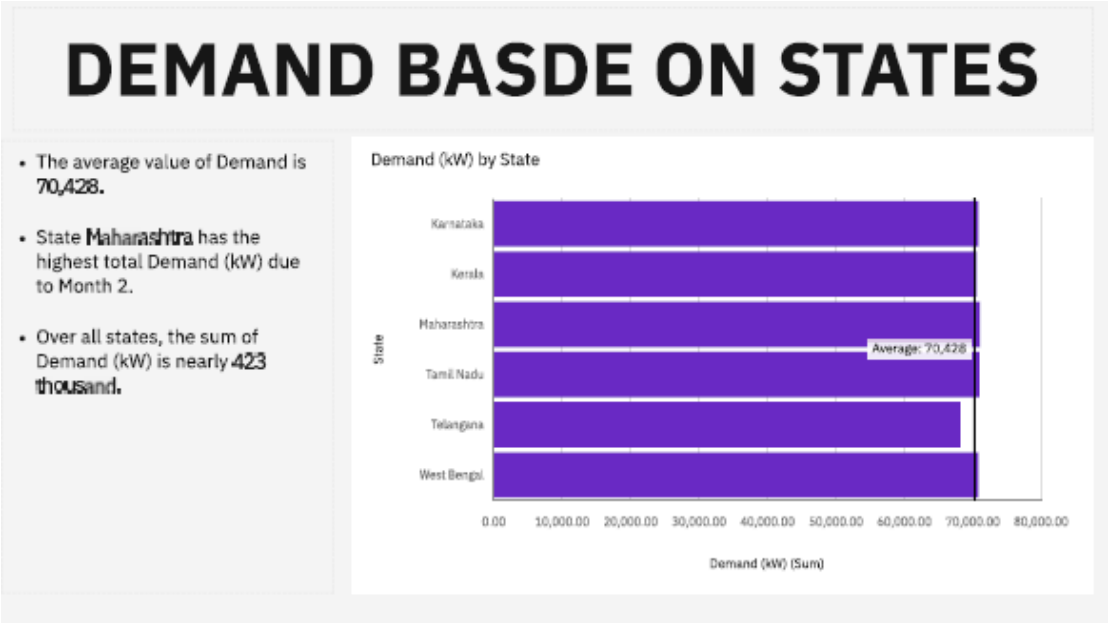


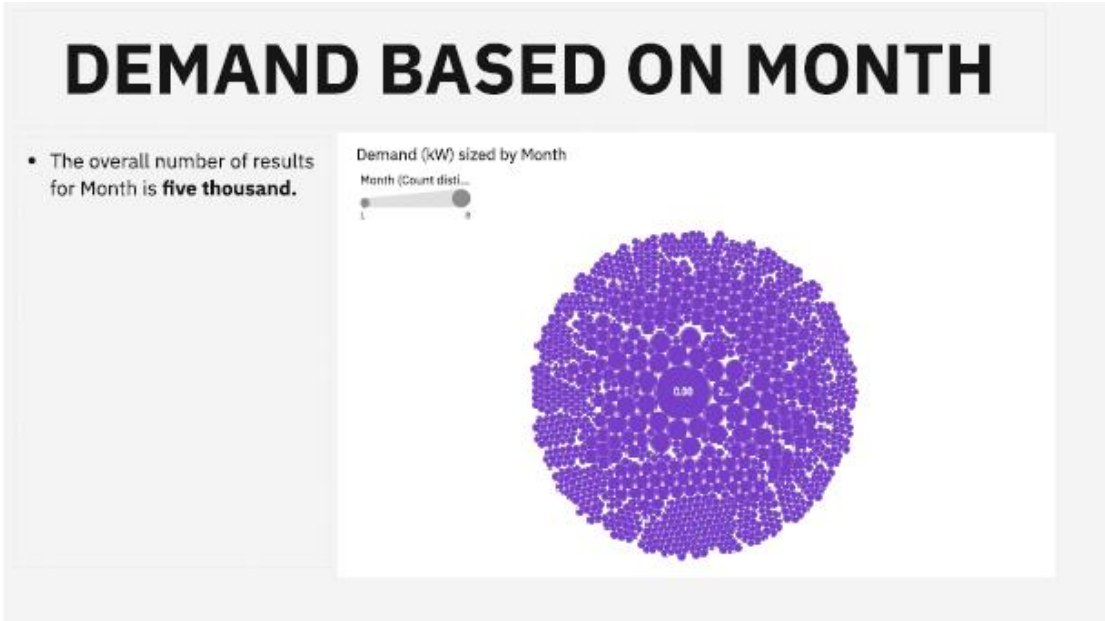
PROFIT BASED ON EFFICIENCY

- Profit ($) is unusually high when Panel Efficiency (%) is **19.06, 13.57, 20.66** and **22.9.**
- Over all values of Panel Efficiency (%), the sum of Profit ($) is almost **162 thousand.**
- Over all values of Panel Efficiency (%), the sum of Energy Loss (%) is almost **30 thousand.**



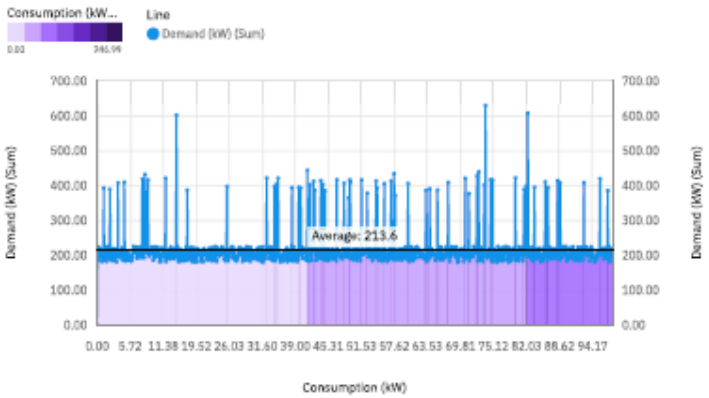
DEMAND SCENES:





DEMAND BASED ON CONSUMPTION

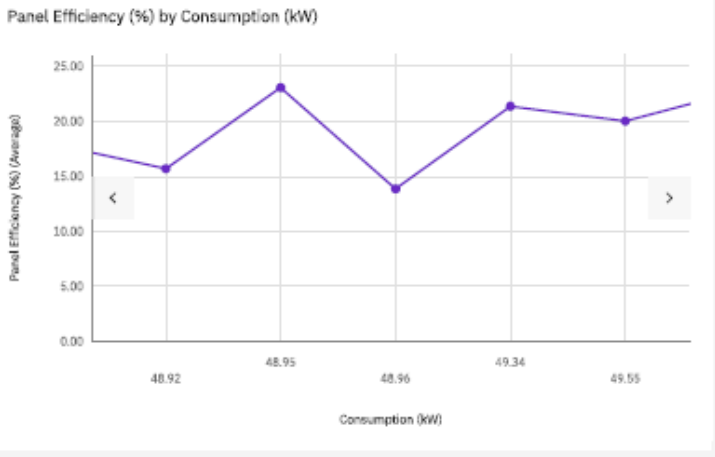
- The average value of Demand is 213.6
- Demand (kW) ranges from 0, when Consumption (kW) is 0, to 628.7, when Consumption (kW) is 73.88.
- 127.25 has a Demand (kW) of 443.2 for Month 12.



CONSUMPTION SCENES:

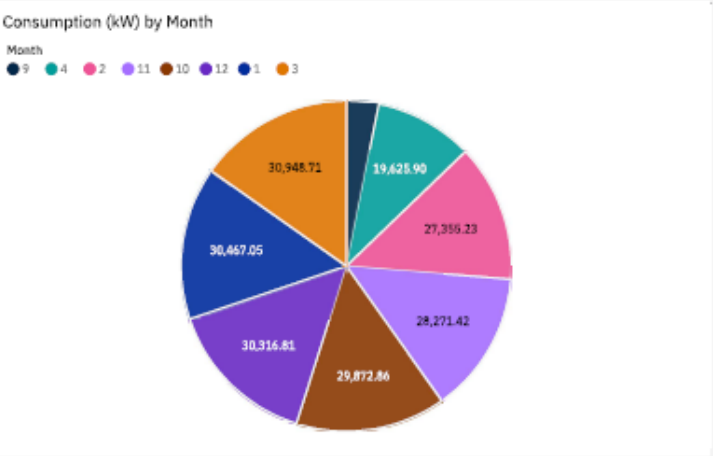
CONSUMPTION BASED ON EFFICIENCY

- Consumption (kW) 82.33 has the highest average Panel Efficiency (%) due to Day Tuesday.
- 43.17 has a Panel Efficiency (%) of 25 for Day Wednesday.
- The average values of Panel Efficiency (%) range from 0, occurring when Consumption (kW) is 0, to 25, when Consumption (kW) is 18.76.



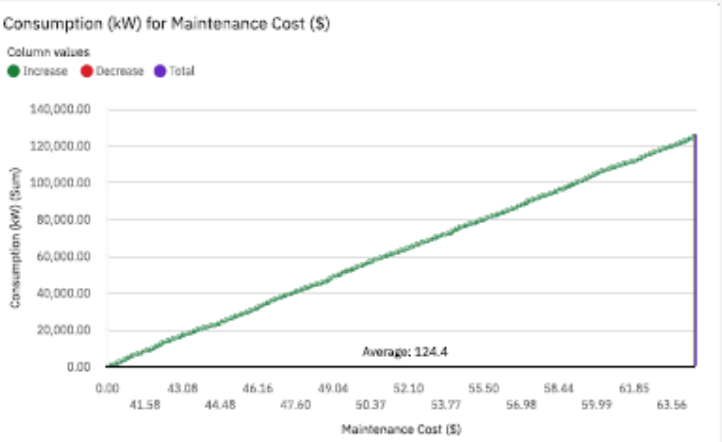
CONSUMPTION BASED ON MONTH

- Consumption (kW) is unusually low when Month is **9.**
- City **Chennai** has the highest total Consumption (kW) due to Month **12**.
- **Mumbai** has a Consumption (kW) of nearly eight thousand for Month **3**.
- Across all months, the sum of **Consumption** (kW) is over **203 thousand.**

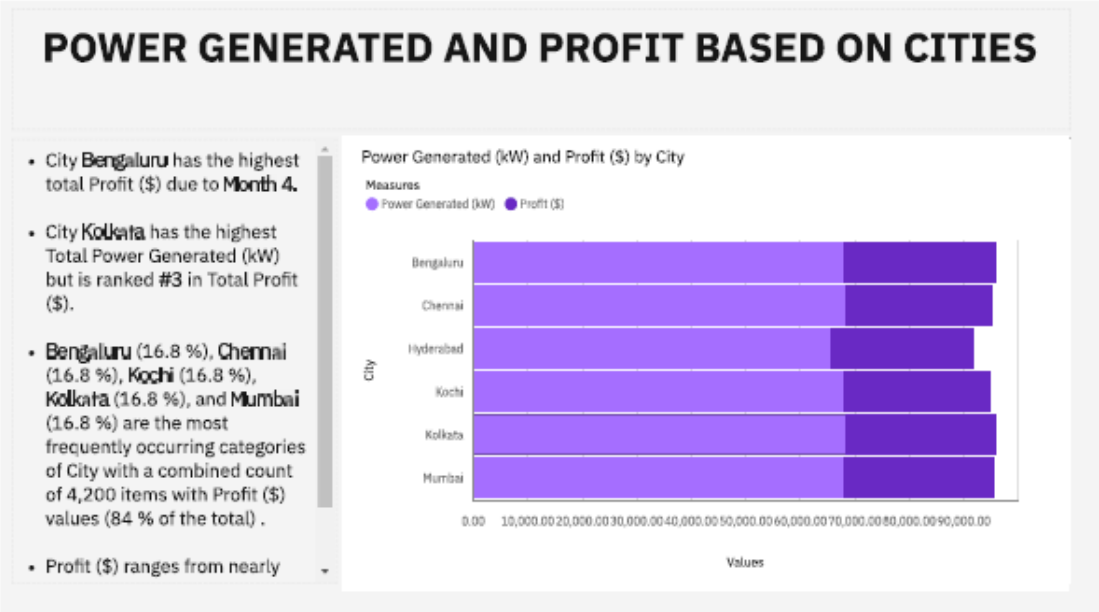


CONSUMPTION BASED ON MAINTANENANCE COST

- Consumption (kW) is unusually high when the values of **Maintenance Cost ($)** are 66.22, 42.21, 58.72, 73.15, 40.97 and more.
- Consumption (kW) ranges from 0, when Maintenance Cost ($) is 0, to 624.2, when Maintenance Cost ($) is 66.22.
- Across all values of **Maintenance Cost ($)**, the sum of Consumption (kW) is over **203 thousand.**



POWER GENERATED SCENE:



CONCLUSION SCENE:

12. WEBPAGE CREATION FOR DASHBOARD :

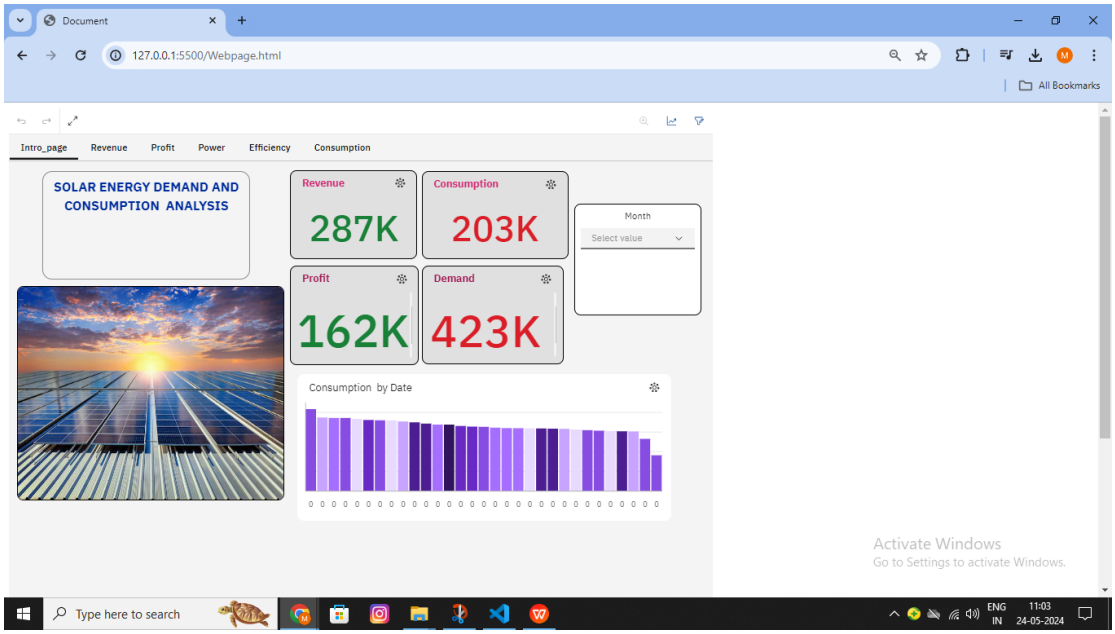
WEB PAGE CODE:

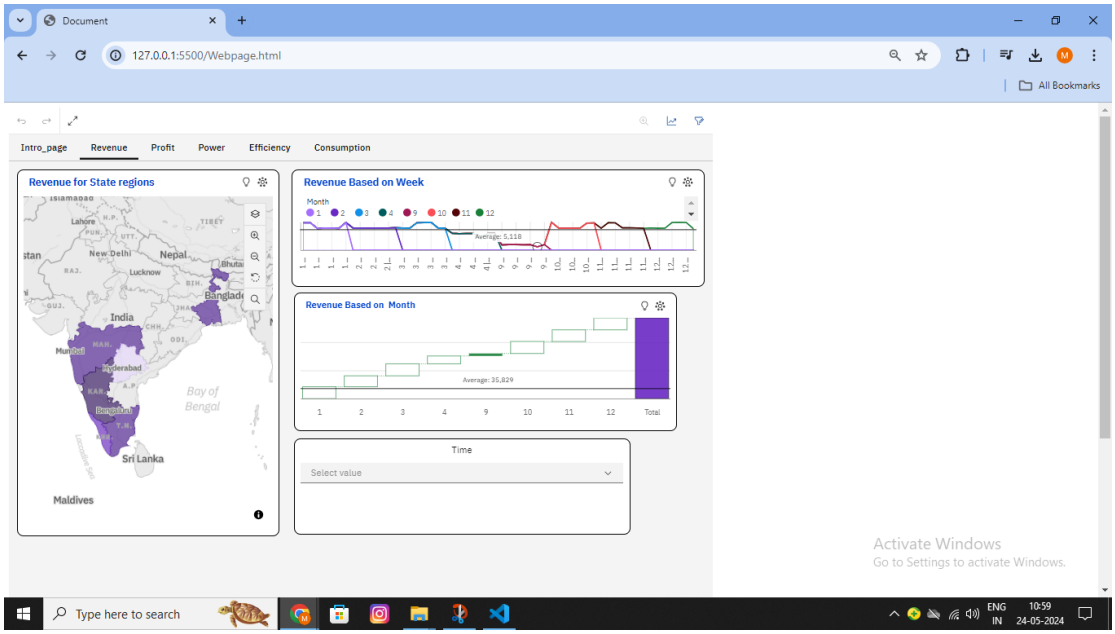
```
<!DOCTYPE html>
<html lang="en">
<head>
    <meta charset="UTF-8">
    <meta name="viewport" content="width=device-width, initial-
scale=1.0">
    <title>Document</title>
</head>
<body>
    <iframe
src="https://us3.ca.analytics.ibm.com/bi/?perspective=dashboard&amp;amp;p
athRef=.my_folders%2FSolar%2BPower%2Bdashboard&amp;amp;closeWin
dowOnLastView=true&amp;amp;ui_appbar=false&amp;amp;ui_navbar=false&am
p;shareMode=embedded&amp;action=view&amp;mode=dashboard&am
p;subView=model0000018f84a3fb89_00000000" width="1080"
height="1080" frameborder="0" gesture="media" allow="encrypted-
media" allowfullscreen=""></iframe>
</body>
</html>
```

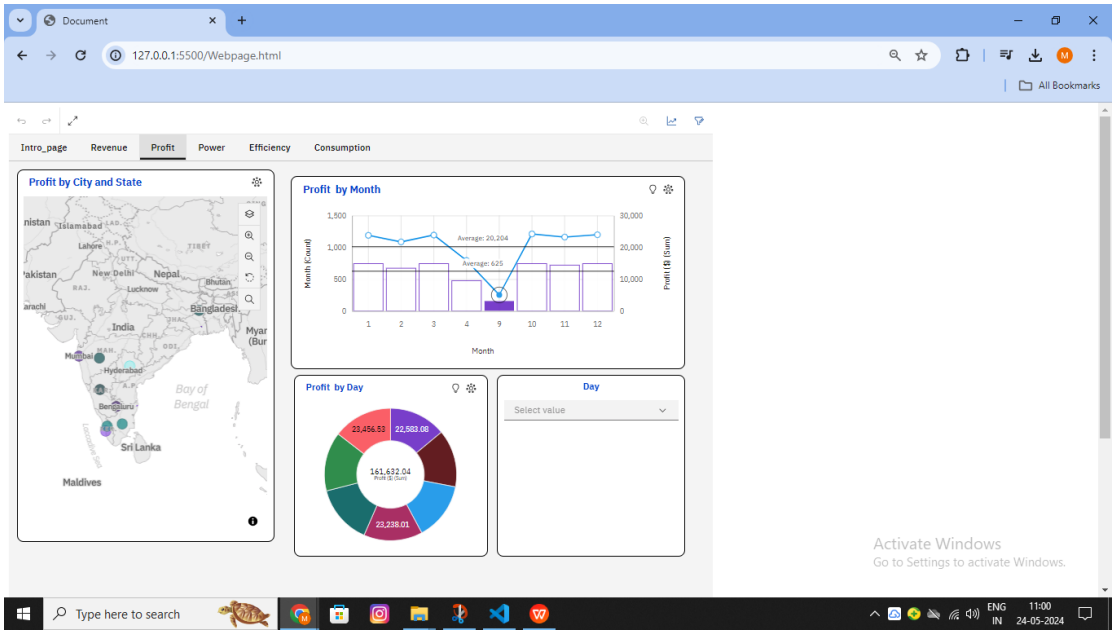
WEEB PAGE LINK:

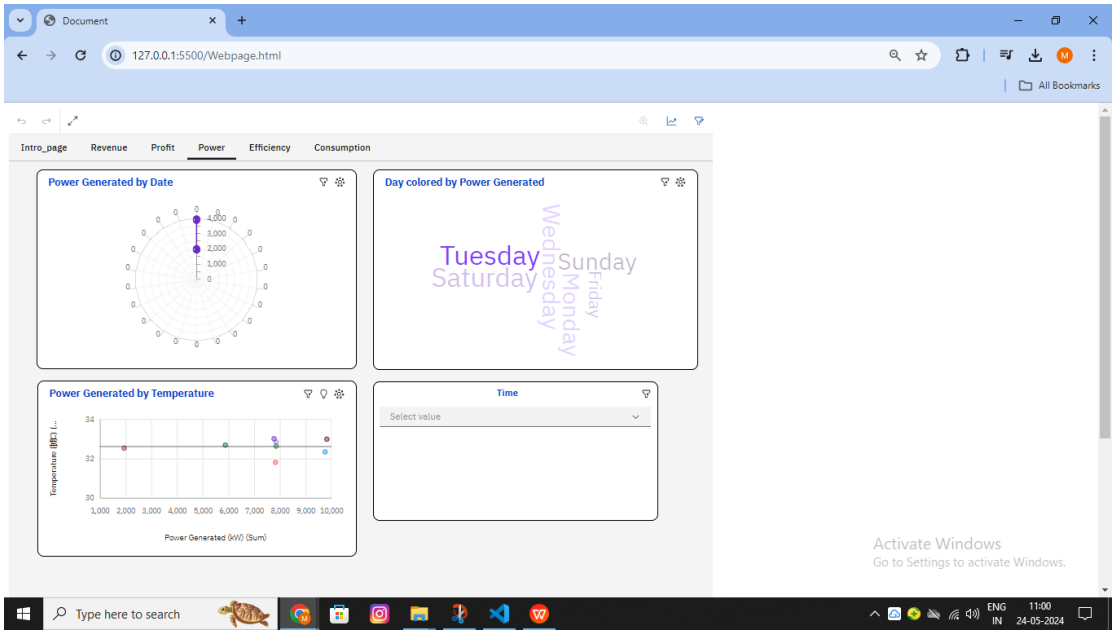
http://127.0.0.1:5500/Webpage.html

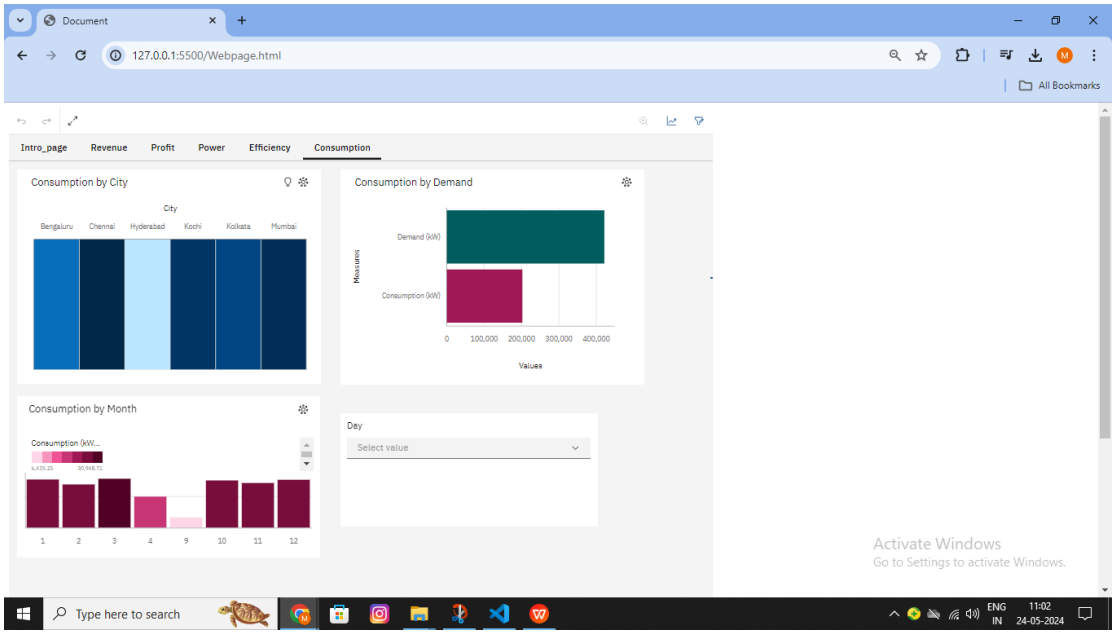
WEB PAGE SCREENSHOTS:






Document
127.0.0.1:5500/Webpage.html
All Bookmarks
Intro_page
Revenue
Profit
Power
Efficiency
Consumption
Profit by City and State
Profit by Month
Average: 20,204
Average: 625
Month (Count)
Profit ($) (Sum)
Month
Profit by Day
23,456.53
22,583.08
161,632.04
23,238.01
Day
Select value
Activate Windows
Go to Settings to activate Windows.
Type here to search
ENG
IN
11:00
24-05-2024


Document
127.0.0.1:5500/Webpage.html
All Bookmarks
Intro_page
Revenue
Profit
Power
Efficiency
Consumption
Power Generated by Date
Day colored by Power Generated
Tuesday
Saturday
Sunday
Wednesday
Monday
Friday
Power Generated by Temperature
Power Generated (kW) (Sum)
Time
Select value
Activate Windows
Go to Settings to activate Windows.
Type here to search
ENG
IN
11:00
24-05-2024



13. DESCRIPTIVE ANALYSIS:

Descriptive analysis in the context of datasets involves summarizing and organizing the characteristics of a data set. When you collect data, the first step in statistical analysis is to describe the features of the responses or observations.

Measures of Variability: These describe how spread out the values are.

Key measures include:

Range: The difference between the maximum and minimum values.

Variance: The average of the squared differences from the mean.

Standard Deviation: The square root of the variance.

Descriptive analysis helps you understand your data before diving into more complex statistical techniques. It provides a foundation for further exploration and hypothesis testing

The Descriptive Analysis is done with the help of python which is as given below:

desc_ana.py - D:/python/desc_ana.py (3.11.1)

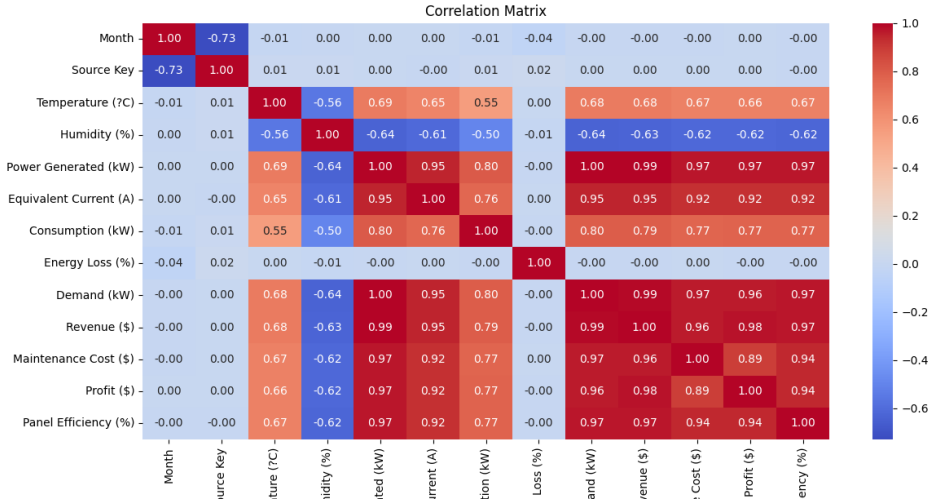
File Edit Format Run Options Window Help

```
import pandas as pd

# Load the dataset
data = pd.read_csv('C:\\Users\\MOHAN\\Documents\\Python\\solar_power_dataset_.csv')

# Generate summary statistics
summary_statistics = data.describe()
print(summary_statistics)
```

IDLE Shell 3.11.1

File Edit Shell Debug Options Window Help

```
Python 3.11.1 (tags/v3.11.1:a7a450f, Dec  6 2022, 19:58:39) [MSC v.1934 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
========================= RESTART: D:/python/desc_ana.py =========================
              Month   Source Key  ...   Profit ($)  Panel Efficiency (%)
count  5000.000000  5000.000000  ...  5000.000000           5000.000000
mean      6.383200  2500.500000  ...    32.326408              7.904420
std       4.366437  1443.520003  ...    39.503042              9.608823
min       1.000000     1.000000  ...     0.000000              0.000000
25%       2.000000  1250.750000  ...     0.000000              0.000000
50%       4.000000  2500.500000  ...     0.000000              0.000000
75%      11.000000  3750.250000  ...    72.937500             17.740000
max      12.000000  5000.000000  ...   114.180000             25.000000

[8 rows x 13 columns]
>>>
```



14. PREDICTIVE ANALYSIS :

Predictive analytics is the process of using data to forecast future outcomes. It combines historical data with statistical modeling, data mining techniques, and machine learning to identify patterns that might predict future behavior.

The Descriptive Analysis is done with the help of python which is as given below:

File Edit Format Run Options Window Help

```
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

# Load the dataset
data = pd.read_csv('C:\\Users\\MOHAN\\Documents\\Python\\solar_power_dataset_.csv')

# Preprocessing
# Dropping non-numeric columns and handling missing values
data = data.select_dtypes(include=[float, int]).dropna()

# Feature selection
features = ['Temperature (?C)', 'Humidity (%)', 'Equivalent Current (A)', 'Consumption (kW)',
            'Energy Loss (%)', 'Demand (kW)', 'Revenue ($)', 'Maintenance Cost ($)', 'Panel Efficiency (%)']
target = 'Power Generated (kW)'

X = data[features]
y = data[target]

# Splitting the data
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

# Building and training the model
model = LinearRegression()
model.fit(X_train, y_train)

# Making predictions
y_pred = model.predict(X_test)

# Evaluating the model
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print(f'Mean Squared Error: {mse}')
print(f'R^2 Score: {r2}')
```

IDLE Shell 3.11.1

File Edit Shell Debug Options Window Help

```
Python 3.11.1 (tags/v3.11.1:a7a450f, Dec  6 2022, 19:58:39) [MSC v.1934 64 bit (AMD64)] on win32
Type "help", "copyright", "credits" or "license()" for more information.
>>>
========== RESTART: D:/python/desc_ana.py ==========
Mean Squared Error: 49.55767295533946
R^2 Score: 0.9946764446800183
>>>
```

15. ACTION PLAN:

1. Understanding the Dataset

Goal: Identify the objectives of your analysis and the key questions you want to answer.

Action: Review the dataset and understand the meaning of each column.

2. Data Preprocessing

Cleaning the Data:

Handle missing values (e.g., imputation or removal).

Correct any inconsistencies or errors in the data.

Transforming Data:

Convert categorical data into numerical form (if necessary).

Normalize/standardize numerical features.

Action: Write a script to clean and preprocess the data.

3. Exploratory Data Analysis (EDA)

Data Visualization:

Use histograms, box plots, scatter plots, and correlation heatmaps to understand data distributions and relationships using **IBM cognos Analytics.**

Action: Conduct EDA and document insights.

4. Model Evaluation

Performance Metrics:

Evaluate models using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), R^2 Score, etc.

Model Comparison:

Compare the performance of different models.

Action: Evaluate and compare models, select the best-performing one.

5. Model Deployment

Saving the Model:

Save the trained model using joblib or pickle.

Creating an API:

Develop a simple API (e.g., using Flask or FastAPI) to serve the model.

Action: Implement model deployment strategy.

8. Project Documentation

Report Findings:

Document the entire process, from data preprocessing to model evaluation.

Stories:

Create visualizations and slides to present the project findings using story scenes in **IBM cognos Analytics.**

Action: Compile documentation and prepare a presentation.

17. CONCLUSION :

The project successfully leveraged the capabilities of Cognos Analytics to collect, preprocess, explore, and visualize solar power data. The interactive dashboards, detailed reports, engaging data story, and comprehensive webpage collectively provided a robust platform for understanding and analyzing solar power generation patterns. This holistic approach enabled stakeholders to make informed decisions, identify areas for improvement, and optimize solar power generation processes. The project demonstrated the power of data analytics in transforming raw data into actionable insights and showcased the effectiveness of Cognos Analytics in facilitating data-driven decision-making.