**SOUTHERN ALBERTA INSTITUTE OF TECHNOLOGY**

**SCHOOL FOR ADVANCED DIGITAL TECHNOLOGY (SADT)**

**An analysis of energy consumption in Alberta and strategies for sustainable energy practices**

**By**

**GROUP 5**

**Linta Jose George**

**George Agbakpe**

**Kamalpreet Singh**

**PRAJWAL NAGARAJ**

**APRIL 2024**

# **PART 5**

## **BUILDING THE ANALYTICAL SOLUTION**

**Risk and Reward**

In analysing AESO energy consumption data comes with both risks and rewards. Thorough understanding of the risks and rewards are crucial for effective decision-making. Below are an outline of the potential risks and rewards that may arise, along with how to effectively manage same.

**Table 1: Risks**

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Risk** | **Description** | **Management** |
| 1 | Market volatility | The energy industry can be highly volatile due to conditions such as natural disasters or events, weather condition, economic conditions and geopolitical. | Diversifying the energy mix and using major tools for risk management such as hedging, staying up to date on external factors which can eventually has an impact on the market and having systems and tools to forecast and predict to some degree of certainty natural event and some disasters. |
| 2 | Regulatory Changes | Regulatory policies and federal government can be changed, these will impact the energy sector's operations specifically investment in the sector and the direction of those investments. | There is the need to stay updated on regulatory developments, and all parties should be engaged in policy development and so that change in government will not lead to change in policy and investment priorities. |
| 3 | Technological advancement | Rapid technological changes can have a disruptive effect or impact on the energy landscape. For push for green technology such as electric vehicles (EVs) will require more generation capacity to meet demand. | Proper investment in research and development, staying up to date on technological trends, and having systems to adapt quickly to those changes will go a long way to help. This will lead to making use of advanced data analytics and predictive modelling. |
| 4 | Clean environmental push | The ever-growing environmental concerns and awareness will lead to stringent regulations, affecting traditional energy sources in the energy generation mix such as coal and oil. | Adequate investment in clean energy alternatives and sustainable, wide diversification of energy sources, and above all, proactively address environmental concerns will be the better approach to use. |

**Table 2: Rewards**

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Demands** | **Description** | **Management** |
| 1 | Innovation opportunities | With the ever-growing push for clean energy alternative sources and sustainable energy mis. The energy industry is poise for growth and offers a lot of opportunities for new technologies and innovation. | The era of working in silos are far gone, energy regulators and producers are to collaborate with technology partners, invest in research and development and promote a culture of innovation. |
| 2 | Demand for energy | The increasing population, high demand for EVs and industrialization will lead to high demand for energy to sustain the industry. | Proper assessment and dentification of high profitable areas with high growth potential to invest in, and adapt more agile strategies to meet such a increasing demand. |
| 3 | Energy transition effort | The shift globally towards sustainable and cleaner energy sources also offer an avenue for investment in renewable energy projects which can offer more generation capacity. | Proper evaluation and diversification into the renewable sources and exploring collaboration with existing companies in the green energy industry and adequately aligning strategies with global sustainability targets |

**Data Collection**

The dataset for the project is on Alberta Electric System Operator (AESO) public data source portal(<https://public.tableau.com/app/profile/market.analytics/viz/AnnualStatistics_16161854228350/Introduction>). Hence, the dataset is available, and access is not restricted. The downloaded data covers a period of 8 years (1st January 2015 to 31st December 2023).

A screenshot of a computer

Description automatically generated

Source: <https://public.tableau.com/app/profile/market.analytics/viz/AnnualStatistics_16161854228350/Introduction>

Below is the structure of the downloaded data. Though the data is cleaned, the format is not suitable for analysis. Hence the need to move to the next stage our architecture overview which is the data preprocessing stage.

There is the need to process the data and convert it to a format suitable for analysis. The essence is to clean and preprocess the data to ensure accuracy. Handle missing values, outliers, and any inconsistencies.

**Table 3: Raw demand data structure**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Date | Hourly Profile | Season | Date (MST) | Calgary | Central | Edmonton | Losses | Northeast | Northwest | South |
| 01/01/2015 | OFF PEAK | WINTER | 01/01/2015 00:00:00 | 1028.338459 | 1326.524094 | 1375.897642 | 279.0501962 | 907.5675665 | 989.6784688 | 922.8939968 |

**Table 3: Raw generation data structure**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Fuel Type** | **Date - MST** | **Date** | **Date (MPT)** | **Date (MST)** | **Maximum Capacity** | **System Available** | **System Capacity** | **System Generation** | **Total Generation** |
| Coal | 08/10/2023 00:00:00 | 08/10/2023 | 08/10/2023 03:00:00 | 08/10/2023 02:00:00 | 820 | 195 | 820 | 273.1354 | 272.0294955 |
| Coal | 08/10/2023 00:00:00 | 08/10/2023 | 08/10/2023 06:00:00 | 08/10/2023 05:00:00 | 820 | 380 | 820 | 377.808 | 377.5864382 |
| Coal | 08/10/2023 00:00:00 | 08/10/2023 | 08/10/2023 23:00:00 | 08/10/2023 22:00:00 | 820 | 380 | 820 | 376.112 | 375.9049385 |

**Data Preprocessing**

The data was transformed and converted into a format suitable for analysis. The data was therefore loaded into Microsoft Power BI.

A screenshot of a computer

Description automatically generated

Power Query was then launched to transform the data by deleting unnecessary columns such as totals (system load), dates in different regions and unpivoting the data in a format suitable for analysis. Below is the figure of the transformation process.

A screenshot of a computer

Description automatically generated

After unpivoting, the two columns which names are not descriptive enough were renamed to Region and Load. This will therefore make them descriptive enough for analysis to be performed on them.

A screenshot of a computer

Description automatically generated

The table below is the output from cleaning and preprocessing, which is ready and are in the format for analysis. We then move to the next stage, which is how to save the clean data.

***Unpivoted data structure***

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Date** | **Hourly Profile** | **Season** | **Date (MST)** | **Region** | **Load** |
| 01/01/2015 | OFF PEAK | WINTER | 01/01/2015 00:00:00 | Calgary | 1028.338459 |
| 01/01/2015 | OFF PEAK | WINTER | 01/01/2015 00:00:00 | Central | 1326.524094 |

**Data Storage**

To make the data available to all team members, a folder was created in OneDrive where the data in CSV was saved and shared for ease of collaboration with version control enabled.

**A screenshot of a computer

Description automatically generated**

**Data Analysis Tools**

As indicated in the proposal, Power BI and Python were the data analysis tools used in our analysis. Power BI was basically used for the initial data cleaning and preprocessing into a format suitable for analysis.

The cleaned data in a csv format was loaded into Jupyter Notebook using Python. Below is the loading process (python code to load the data) and print the first five (5) rows of data from the data frame.

**A screenshot of a computer

Description automatically generated**

To be sure the imported data was of correct format and datatype and suitable for analysis, we again converted the data type to appropriate format to avoid errors during the analysis phase. Below is the code and the output from the datatype conversion.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

**Exploratory Data Analysis**

Before we perform exploratory data analysis on the data, we first checked for missing values though we know the data is clean. This is to be sure we are working with data with no missing values of significance.

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

With the results above we are sure and confident that there are no missing values, and we are confident that we will be working with a very clean data when we proceed. Exploratory Data Analysis involves visualizing and understanding the data to gain insights. Here are some techniques we applied to the energy consumption data.

We used Histogram of Load to visualize the distribution of the 'Load' variable.

A graph of a number of blue and white lines

Description automatically generated with medium confidence

Box Plots by Region was explored to determine the load distribution in different regions including system or transmission loses.

**A screenshot of a computer screen

Description automatically generated**

Violin plot of Load by Season was utilized to visualize the distribution of load by season.

**A screenshot of a computer screen

Description automatically generated**

**Statistical Analysis**

We then proceeded to calculate summary statistics such as mean, median, standard deviation, minimum, and maximum for the 'Load' column to ascertain the spread of the data or values in there. Below is the code and output from that.

A screenshot of a data analysis

Description automatically generated

Descriptive analysis for the power generation dataset

A screenshot of a computer

Description automatically generated

A plot of time series of load for each region to see an overview of demand across Alberta.

**A graph of different colored lines

Description automatically generated**

Grouped Analysis was conducted to group the data by 'Hourly Profile', 'Season', and 'Region' and calculate mean load for each group.

**A screenshot of a computer program

Description automatically generated**

**A graph of a line graph

Description automatically generated with medium confidence**

**Presentation of Solution and Results**

#### **General objective**

The main objective of our project is to analyse energy consumption datasets from Alberta Open Data, identify trends and patterns in energy usage, and develop actionable insights to promote sustainable energy practices.

#### **Specific objectives**

The various specific objectives of our project are to:

1. analyze historical data can reveal long-term trends, such as increasing or decreasing energy consumption patterns, which can influence future planning.
2. develop predictive models based on historical data can assist in forecasting future energy consumption, supporting proactive planning and decision-making.
3. investigate the correlation between weather patterns (temperature, humidity, etc.) and energy consumption can assist in forecasting and adapting to seasonal variations.
4. understand when energy demand is highest can help in optimizing resource allocation and planning for increased capacity during peak times.
5. identify regional variations can help tailor energy policies and initiatives to address specific needs and challenges in different areas.
6. based on the analysis, propose actionable strategies and recommendations to encourage sustainable energy practices among consumers, businesses, and industries.

#### **Business questions**

The business questions of our project are as follows:

1. Are there noticeable trends in energy consumption over the years?
2. Can predictive modelling be used to anticipate future energy consumption trends?
3. How do weather conditions impact energy consumption?
4. What are the peak hours and seasons of energy consumption?
5. Are there specific regions with notable differences in energy consumption patterns?
6. What strategies can be implemented for promoting sustainable energy practices?

***Solution***

Based on the specific objectives above and the business questions, we hereby present our solution below.

**A graph showing a line of blue and orange lines

Description automatically generated**

The above graph confirms there are variations or seasonality in both energy consumption and demand within Alberta over the years. This could be attributed to weather conditions and seasons within the year.

**A graph of a graph with numbers and a line

Description automatically generated with medium confidence**

The monthly energy consumption and generation was plotted. This again collaborates with the finding in the above graph. It also proves the weather conditions affecting generation and demand have been influenced by the seasons.

**A graph of energy generation

Description automatically generated with medium confidence**

This plotted for a specific year just to be sure it runs through the years. And it was confirmed, no matter the year those conditions play a role. Only 2020 shows a different trend and this can be attributed to the covid-19 lockdown. Below is the plot for 2020.

A graph of energy generation

Description automatically generated with medium confidence

**A graph of different colored lines

Description automatically generated**

**A graph of different colored lines

Description automatically generated**

**A graph of a graph showing the number of fuel types

Description automatically generated with medium confidence**

**A graph showing a number of energy

Description automatically generated with medium confidence**

**A graph showing a number of energy generation

Description automatically generated**

**A graph of a line

Description automatically generated**

**A graph with a line going up

Description automatically generated**

**A graph of different seasons

Description automatically generated with medium confidence**