**Junior Data Analyst Assessment: PULA**

**Date: 8/10/2023**

**Client: Naomi**

**Evaluation and Advice to Naomi concerning Electricity Savings**

**Scenario:** Naomi has solar panels generating electricity. She aims to store excess energy using a $7,000 battery, offsetting her electricity needs when solar power is insufficient. The battery’s lifespan is 20 years, storing up to 12.5 kWh. Electricity cost is $0.17/kWh, increasing yearly by 4% as per the government, but Naomi expects a faster rise of 4.25% yearly due to climate change. By analyzing her solar generation and usage data, we’ll calculate saved kWh, cost savings, NPV, and IRR for both government-expected and Naomi-estimated electricity price scenarios.

**Python Github link for Crossreference:** [**https://github.com/TURUE/Assessment-Project-**](https://github.com/TURUE/Assessment-Project-)

**Presented by:** Victor Joseph Akello Oduongo

**Table of Contents**

[1.1 Data Cleaning & Exploratory Data Analysis (EDA) 2](#_Toc142587913)

[1.2: The Amount Of Electricity That Needed To Be Bought From The Electricity Provider 5](#_Toc142587914)

[1.3 Excess Solar Power Generated Over Electricity Used 7](#_Toc142587915)

[1.4 Model Of The Cumulative Battery Charge Level 8](#_Toc142587916)

[1.5 Amount Of Electricity For Each Hour From The Electricity Provider 9](#_Toc142587917)

[1.6 Electricity Savings from Installing Battery 10](#_Toc142587918)

[1.7 Data Tabulation 11](#_Toc142587919)

[1.8 Scenarios 12](#_Toc142587920)

[1.9 The Internal Rate of Return (IRR) For The Two Scenarios 13](#_Toc142587921)

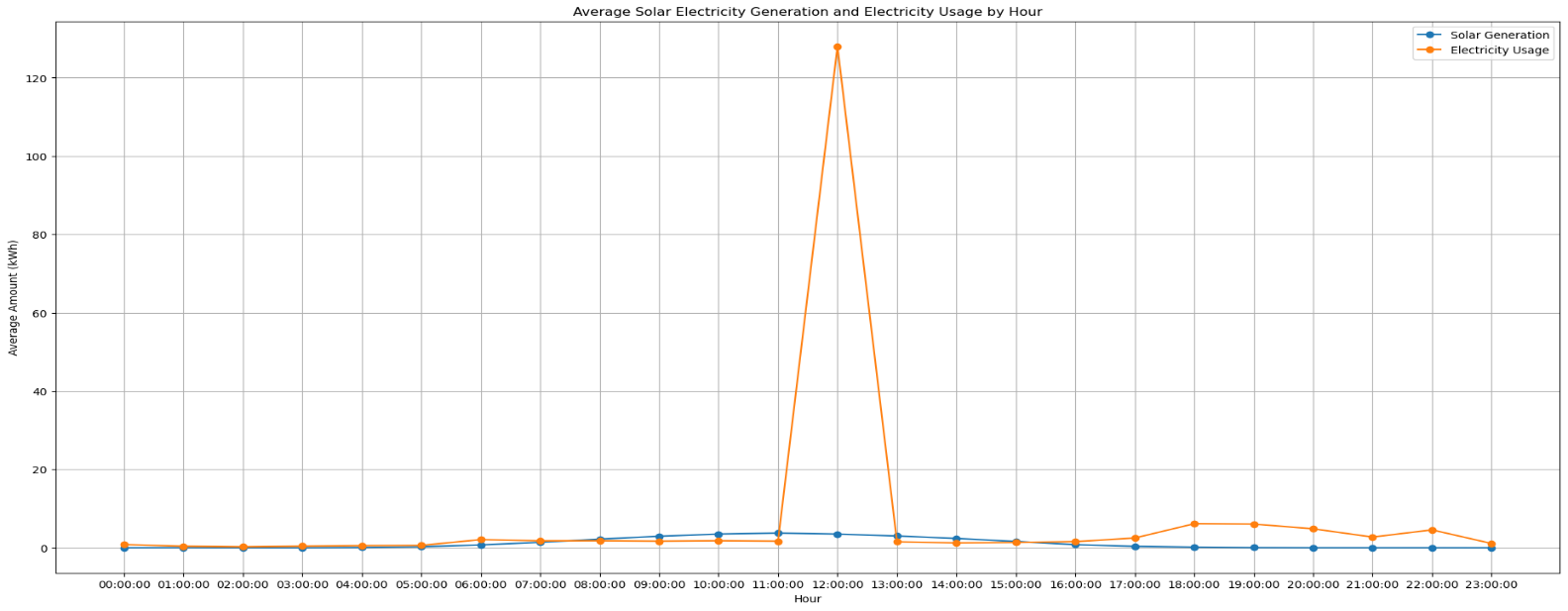
Naomi provided a dataset containing details about her solar energy consumption and power generation. The data was collected for each hour from 1st January 2020 to 31st December 2020. **Issues Discovered In The Dataset**

* The data contained one missing value.
* There were inconsistencies in the decimal places of the “Solar electricity generation (kWh)” and “Electricity usage (kWh).”

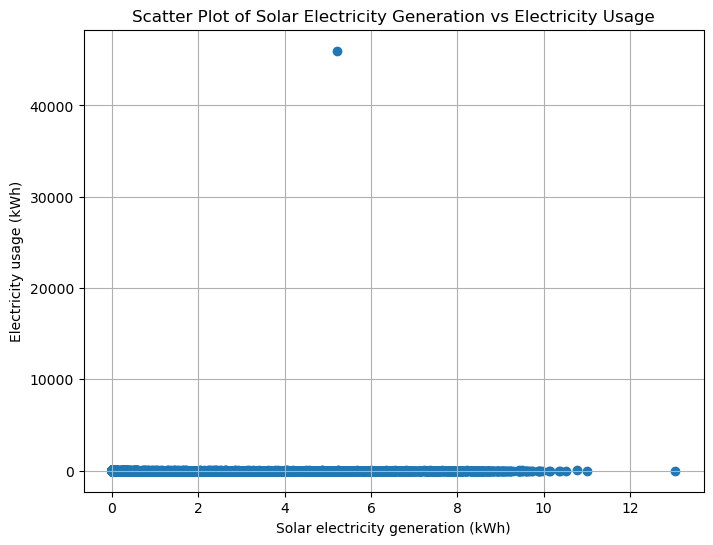
# 1.1 Data Cleaning & Exploratory Data Analysis (EDA)

* Dropped Null rows to guarantee that the data is clean
* Rounded off the “Solar electricity generation (kWh)” and “Electricity usage (kWh)” columns to 3 decimal places.
* Split the Date/hour start section to ease the visualization of the data to identify outliers.
* Explored and visualized the outliers using a graph and scatter plot, as shown in Figure 1 and Figure 2.

**Figure 1**

*Average Solar Electricity Generation and Electricity Usage by Hour*

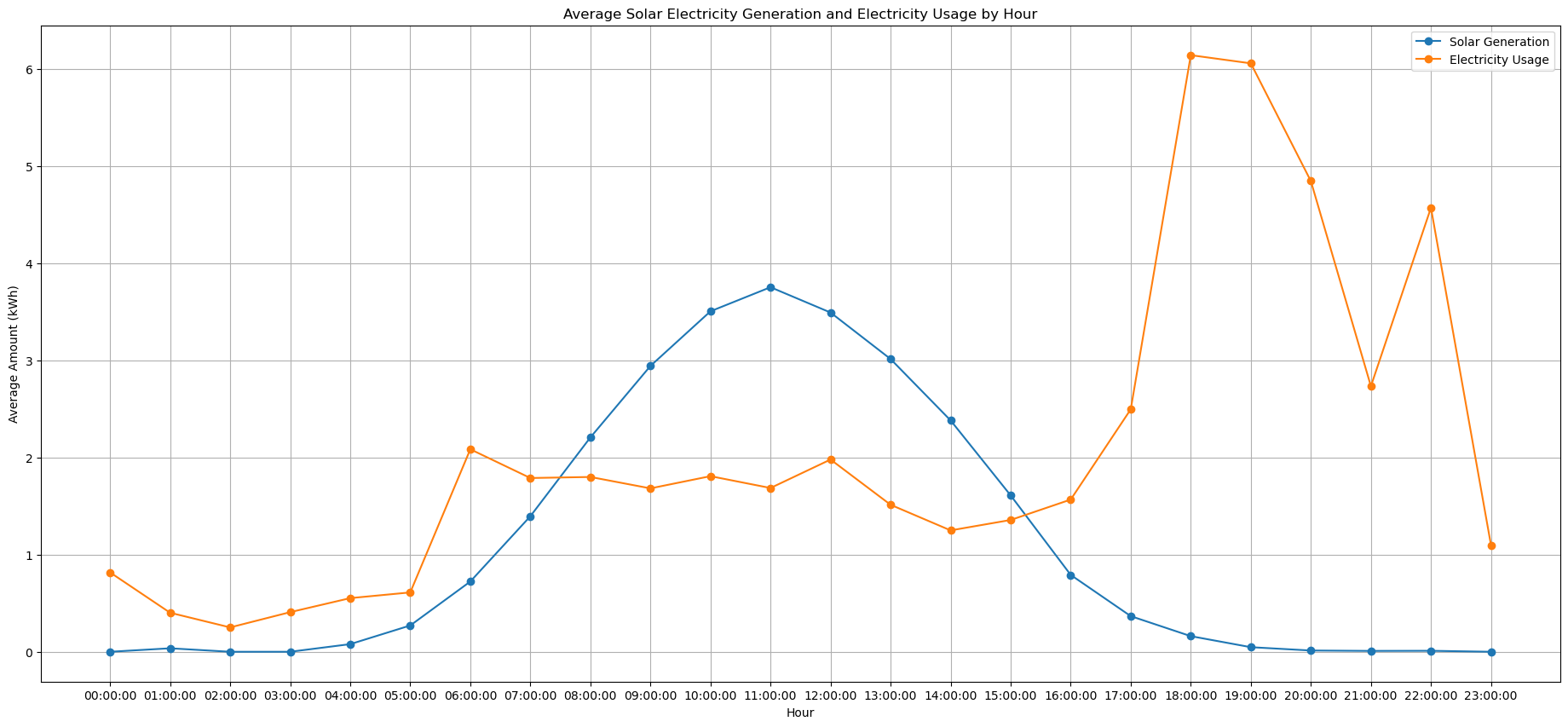
**Figure 2**

*Scatter Plot of Solar Electricity Generation vs Electricity Usage*

**The solution to the Outliers:** Identified the issue at the 12:00 hour mark and arranged the values from the maximum to the minimum to identify and remove the exaggerated value, which was 46000. Figure 3 shows the data points after deleting the outliers.

**Figure 3**

*Visualization After Deleting The Outliers*

****

**Research Question:**

What is the potential cost-effectiveness of installing a battery system for storing excess solar-generated electricity to offset electricity costs, considering different electricity price increase scenarios and accounting for the battery’s cost, solar energy generation, and consumption patterns?

# 1.2: The Amount Of Electricity That Needed To Be Bought From The Electricity Provider

*Steps:*

Filtered the data for 2020 and calculated the electricity bought by subtracting ‘Solar electricity generation (kWh)’ from ‘Electricity usage (kWh)’ columns.

Grouped the data by the hour and summed the electricity bought for each hour

*Output:*

hour

0 298.034

1 146.264

2 91.486

3 148.806

4 179.531

5 174.193

6 630.753

7 477.084

8 446.172

9 388.179

10 416.190

11 383.319

12 443.505

13 335.223

14 279.132

15 339.408

16 428.682

17 818.940

18 2182.487

19 2193.194

20 1763.299

21 996.396

22 1664.187

23 400.851

This data shows the electricity needed to be bought from the electricity provider for each hour in 2020. The most electricity bought is during hour 18 (6:00 PM to 7:00 PM) with 2182.487 kWh.

# 1.3 Excess Solar Power Generated Over Electricity Used

**Steps:**

* Filtered the data for 2020 and calculated the excess solar energy
* Grouped and summed the excess solar energy

Obtained the following hourly excess solar electricity generated over electricity used values for each hour in 2020:

hour

0 0.000

1 12.734

2 0.000

3 0.000

4 6.711

5 50.010

6 133.914

7 333.552

8 594.780

9 848.064

10 1035.546

11 1137.588

12 994.125

13 882.669

14 692.193

15 431.643

16 145.680

17 41.910

18 0.000

19 0.000

20 0.000

21 2.139

22 0.000

23 0.218

This data shows the excess solar electricity generated over electricity used for each hour in 2020. The highest excess solar generation is during hour 11 (11:00 AM to 12:00 PM) with 1137.588 kWh.

# 1.4 Model Of The Cumulative Battery Charge Level

**Steps**

* Filtered the data for 2020
* Calculated the cumulative battery charge level and created a new column for it

**Findings:**

1. Obtained the cumulative battery charge levels for each hour in 2020, starting from 1st January 2020. The charge level follows the pattern of being increased or decreased based on the net electricity available (excess solar minus electricity bought) while considering the maximum charge cap.
2. Successfully modeled the cumulative battery charge level for each hour over 2020 based on the provided dataset and constraints.

# 1.5 Amount Of Electricity For Each Hour From The Electricity Provider

**Steps**

* Looped through each hour of 2020 and calculated the current hour’s net electricity available (excess solar minus consumption).
* For every deficit (net electricity available is negative), there was a calculation of the discharge needed based on battery efficiency, and the battery charge level was updated accordingly.
* Calculated the electricity bought considering excess solar and discharge from the battery, ensuring it’s not negative.

**Findings:**

1. Obtained the amount of electricity that would have been bought from the electricity provider for each hour in 2020, assuming a battery had already been installed. The calculations consider excess solar, consumption, and battery discharge based on efficiency and maximum charge cap.
2. Successfully calculated and summarized the hourly electricity bought values for 2020, assuming a battery had already been installed, based on the provided dataset and the given constraints.

# 1.6 Electricity Savings from Installing Battery

**Steps:**

* Defined the price of electricity for 1st January 2022 in dollars per kWh @ 0.17
* Calculated the cost of electricity bought with a battery (using existing solar panels) by summing up the product of ‘Electricity Bought with Battery’ values and the electricity price for 2022.
* Calculated the electricity bought for each hour by applying a lambda function that calculates the difference between ‘Electricity usage (kWh)’ and ‘Solar electricity generation (kWh).’
* Calculated the total cost of electricity bought without a battery by summing up the product of electricity bought values and the electricity price for 2022.
* Calculated the savings from installing a battery compared to using existing solar panels alone by subtracting the cost with the battery from the total cost without a battery.

**Findings:**

1. Based on the calculations and given data, installing a battery for 2020 would lead to a negative value of savings, specifically -451713.42 dollars. This suggests that, in this scenario and with the given assumptions, the cost of using a battery would be higher than using existing solar panels alone for the entire year.
2. This negative value could indicate that the cost of purchasing and installing the battery and its operational inefficiencies might outweigh the potential savings from using the battery to offset electricity costs.

# 1.7 Data Tabulation

**Steps:**

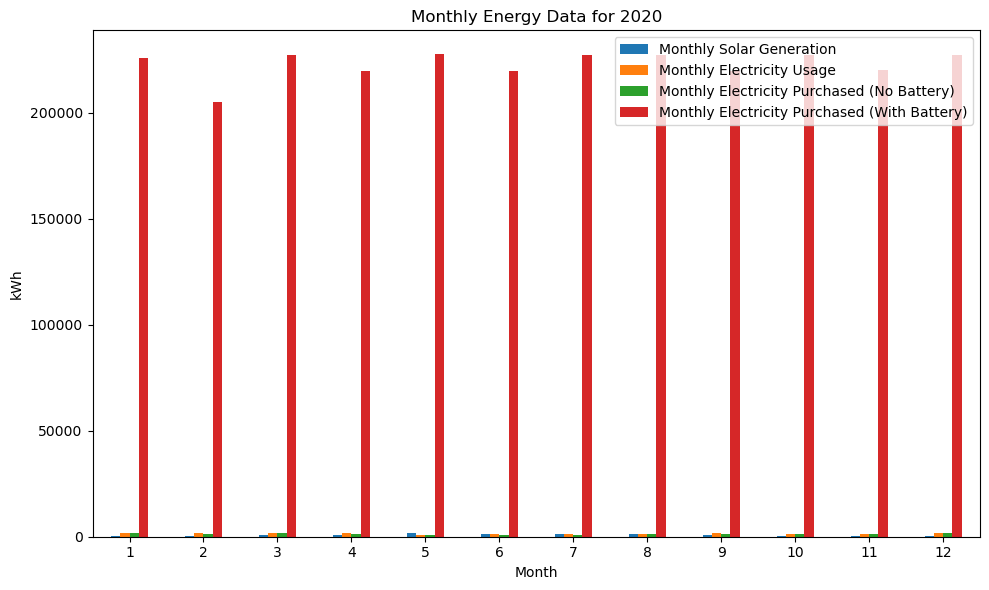
* Grouped data by month and calculated monthly sums
* Display tabulated data
* Created a bar chart to illustrate the data

**Findings:**

1. The tabulated data and bar chart, as shown in Figure 4, clearly represent energy trends throughout 2020. Some key findings from the chart and data include:
2. Solar generation tends to be higher during summer (e.g., May, June, July) and lower during winter (e.g., December, January, February).
3. Electricity usage is relatively higher during the winter months, which could be due to increased heating demands.
4. Both electricity purchased without a battery and electricity purchased with a battery follows similar trends, indicating that the battery’s impact on electricity purchase is not significantly visible in this representation.
5. It’s worth noting that the values for ‘Monthly Electricity Purchased (With Battery)’ are unusually high, potentially indicating an issue with calculations or data.

**Figure 4**

*Tabulation Data*

**

# 1.8 Scenarios

**a) Project forward for 20 years from 1st January 2022 the annual savings from installing the battery for the two scenarios below.**

*Scenario 1. Electricity prices increase as expected by the government, 4% p.a.*

**Findings:**

1. The key finding for Scenario 1 is that the total annual savings over the 20 years is calculated to be 0.0. This result indicates that, in this particular projection, the increased cost of electricity over the years due to the 4% annual price increase cancels out the savings from installing the battery. In other words, the battery’s potential to reduce costs is offset by the increasing electricity prices, resulting in no net savings over the 20 years.
2. It is important to note that this finding is based on the assumption of a 4% annual increase in electricity prices. Different assumptions about electricity price trends or battery performance could lead to different outcomes.

*Scenario 2. Electricity price increases start at 4% p.a. and rise each year by an additional 0.25% p.a., as estimated by Naomi.*

**Findings:**

1. The key finding for Scenario 2 is that the total annual savings over the 20 years is calculated to be approximately -6,566,928.26. The negative value indicates that, in this particular projection, the increased cost of electricity over the years due to the 4% annual increase plus the accelerating additional price increase is greater than the savings from installing the battery.
2. As a result, the overall impact is a substantial cost increase over the 20 years rather than savings. This outcome suggests that the increasing electricity prices, especially with the accelerating increase, offset any potential benefits from the battery installation.

**b) For the two scenarios, calculate the Net Present Value (NPV) of the future annual savings.**

*Scenario 1. Electricity prices increase as expected by the government, 4% p.a.*

*Scenario 2. Electricity price increases start at 4% p.a. and rise each year by an additional 0.25% p.a., as estimated by Naomi.*

**Findings:**

1. The key finding for both scenarios is the calculated NPV of the future annual savings over 20 years. For Scenario 1, the NPV is calculated to be approximately 0. This means that the present value of future savings, discounted at a 6% rate, does not result in any net benefit or loss; it essentially breaks even in present value terms.
2. For Scenario 2, the NPV is calculated to be approximately -2,605,781.65. This negative value suggests that, when considering the accelerating increase in electricity prices and the discounting effect, the present value of the future annual savings is insufficient to offset the projected increased costs over 20 years.

# 1.9 The Internal Rate of Return (IRR) For The Two Scenarios

**Steps:**

* This step involved calculating the Internal Rate of Return (IRR) for the two scenarios described in section 1.8. IRR is the discount rate that makes the Net Present Value (NPV) of future cash flows, in this case, the future annual savings, equal to zero.

**Findings:**

1. For Scenario 1, the calculated IRR is approximately 0.1. This suggests that, for Scenario 1, the discount rate that equates the NPV of the future annual savings to the initial battery cost is around 10%. In other words, the investment in the battery for Scenario 1 would yield an internal rate of return of 10%.
2. For Scenario 2, the calculated IRR is approximately 30477.87. This indicates that, for Scenario 2, the discount rate that makes the NPV of the future annual savings equal to the initial battery cost is approximately 30477.87%. This unusually high IRR suggests that the future annual savings are not enough to offset the initial cost of the battery, making the IRR unrealistic and impractical.