

Watt- Seer Personalized Energy Coach - GenAI Approach

Team Members: Ashwini Apte, Rao Parasa, Suresh Srinivas

● Summary

In a world of growing environmental consciousness and rising energy costs, **WattSeer** leverages GenAI to offer personalized coaching for energy consumption. This project aims to guide individuals in reducing electricity usage based on their behavioral patterns, historical consumption, and actionable insights—powered by AI.

This blog demonstrates innovative approaches in applying the GenAI / ML capabilities and techniques to real world problems leveraging energy efficient solutions. We have implemented robust prompt engineering techniques such as few shot prompts, Chain Of Thought and ensured structured relevant response generation through evaluation techniques learnt. We applied function calling to ascertain the top five days of peak usage and generate key insights. Asserted the use case by providing necessary context and captivating data visualizations such as line charts, bar charts and tables.

● Use Case

As a user, I want to improve my daily energy usage and save energy costs by understanding the billing aspects of energy consumption.

● Problem Statement

Energy consumption patterns vary widely, and most users receive generalized advice that may not apply to their specific needs. The lack of personalized insights hampers efforts to conserve energy effectively. This project addresses the challenge:

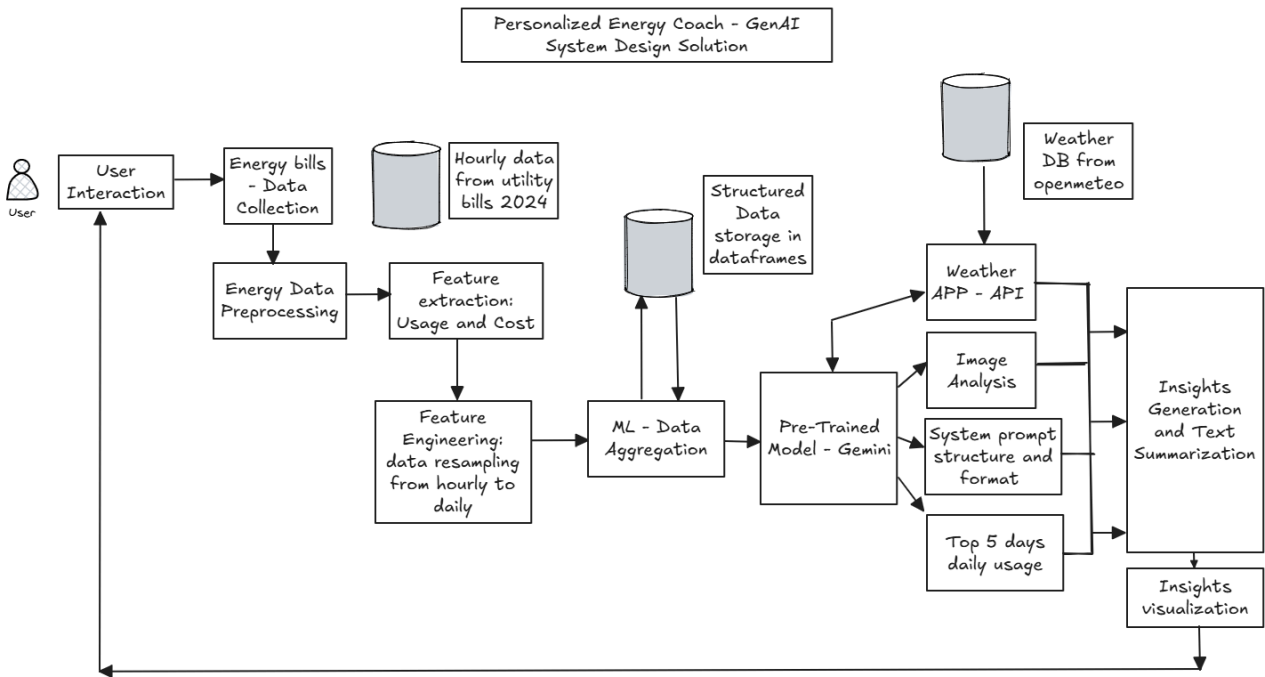
“How can we build a personalized energy coach that understands user behavior and recommends energy-saving actions using generative AI and time-series data?”

● GenAI Solution - The Approach

The solution reinforces a three fold approach in terms of data resampling, aggregation to find the top5 and the RAG based response generation using Gemini evolving through effective prompting techniques. Traditional data visualizations and tables don't always help users make sense of their consumption. GenAI adds a layer of interpretation, turning raw data into clear narratives, recommendations, and even conversations. Here's how we approached the solution:

- Raw **data is gathered** from monthly energy bills from the past one year. **Preprocessed** to focus on variables such as hourly usage and cost.
- This data is then **resampled** to reflect daily energy consumption against the cost to capture any spikes and identify usage patterns for the year on a dual axis chart.
- Next, the daily aggregated data is processed through a **pre-trained LLM** - Gemini 2.5 to **summarize using GenAI Capabilities** and provide recommendations using **function calls** to find top usage days and **prompt engineering** techniques.

- **Interfaced with Weather API** for the LLM to analyse and summarize the weather influence on energy consumption.
- **Image processing** to analyse and summarize neighbors energy usage from his paper bills.
- Following through a **few shots to the chain of thought** providing necessary **context**, the prompt responses evolve delivering more relevant recommendations. The prompt response is **evaluated** using structured rubric **applying LLM as a judge** technique to rate the relevance and correctness of the prompt. Logical workflow architecture:



High Level System Design Architecture Diagram

The WattSeer project combines time-series forecasting, behavioral segmentation, and natural language recommendations to coach users on reducing energy usage. Its GenAI capabilities include:

- Personalized advice via LLM-generated responses in form of insights and summarization
- Forecasting future energy consumption
- Function calling and Weather API calling
- Supporting tailored tips with contextual data insights

• Dataset Overview

The dataset comprises:

- Hourly energy consumption data
- External variables (e.g., temperature, device types)
- User metadata (location, home size, appliance info)

These data points are used to uncover hidden consumption patterns and support personalized recommendations.

```

import pandas as pd

# Load with tab separator and column names
df = pd.read_csv(
    "/kaggle/input/energy-data-meter-2/energy_hourly_data_2024_meter2.dat",
    sep="\t",
    names=["start_time", "end_time", "consumption", "provided_cost",
          "start_minus_prev_end", "end_minus_prev_end"],
    header=None,
    on_bad_lines='skip'
)

# Define exact datetime format (includes time zone offset)
datetime_format = "%Y-%m-%d %H:%M:%S%z"

# Explicit parse
df['start_time'] = pd.to_datetime(df['start_time'], utc=True, format=datetime_format, errors='coerce')
df['end_time'] = pd.to_datetime(df['end_time'], utc=True, format=datetime_format, errors='coerce')

# Drop bad rows
df.dropna(subset=['start_time', 'end_time'], inplace=True)

# Set index and select key columns
df.set_index('start_time', inplace=True)
df = df[['consumption', 'provided_cost']]

```

● Preprocessing, Analysis and Visualizations

The data undergoes preprocessing, including missing value treatment, scaling, and time-based grouping. Below is a key snippet for resampling and feature extraction:

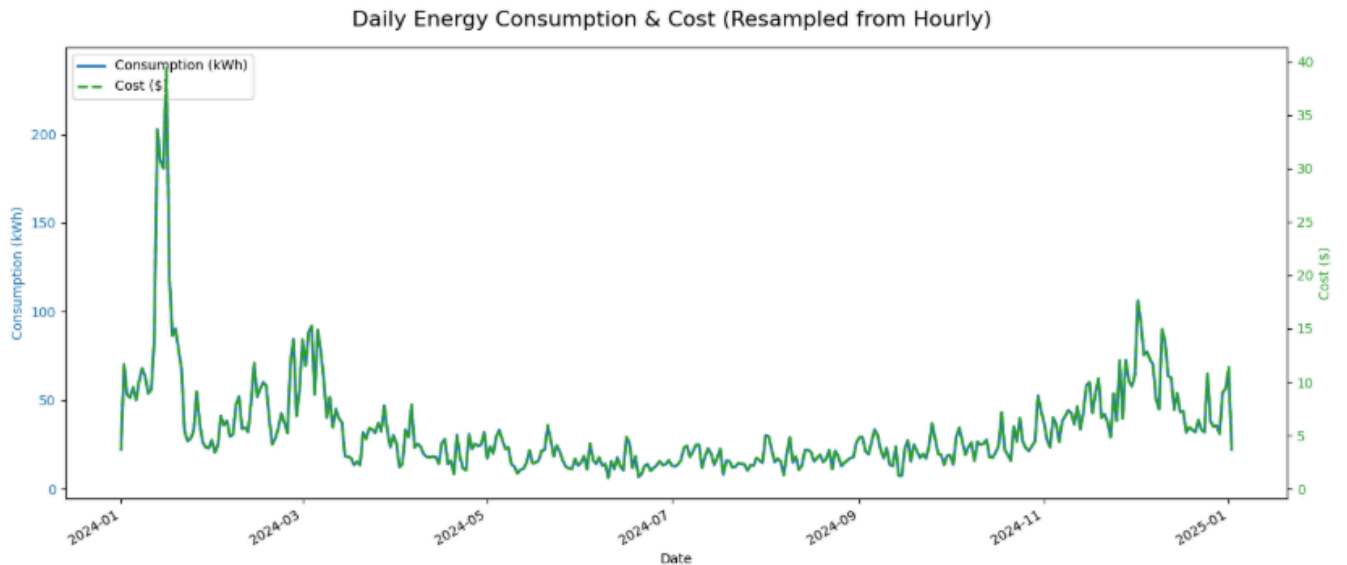
Code: data cleanup, resampling to daily, summarizing usage & cost

```

# Ensure correct types
df['consumption'] = pd.to_numeric(df['consumption'], errors='coerce')
df['provided_cost'] = pd.to_numeric(df['provided_cost'], errors='coerce')
# Resample hourly data to daily totals (IMPORTANT)
daily_data = df.resample('D').agg({
    'consumption': 'sum',
    'provided_cost': 'sum'
})
print(daily_data.head())

```

	consumption	provided_cost
start_time		
2024-01-01 00:00:00+00:00	22.064	3.661962
2024-01-02 00:00:00+00:00	70.308	11.669019
2024-01-03 00:00:00+00:00	53.101	8.813173
2024-01-04 00:00:00+00:00	51.451	8.539322
2024-01-05 00:00:00+00:00	57.207	9.494646



Dual-axis line chart that visualizes daily energy consumption (kWh) and cost (\$), using matplotlib. Here we understand the spike and identify the root cause.

Next, we analyse and summarize the top 5 energy usage days that lead to the spikes. By focusing on reducing the massive energy demand during cold weather and trimming down the constant baseload usage, the household can significantly lower its overall energy consumption and costs.

```
MODEL_ID="gemini-2.5-pro-exp-03-25"
system_prompt="You are a skilled Energy Analyst. You will help with variety of tasks about Energy Usage. You will create graphs, help them analyze the data and more."
chat = client.chats.create(
    model=MODEL_ID,
    config=types.GenerateContentConfig(
        system_instruction=system_prompt
    )
)
prompt = f"""The following table summarizes a household's daily energy usage (kWh) and daily cost (in $) over a fixed-rate plan:

{summary_stats}. Summarize this data"""
response = chat.send_message(prompt)
Markdown(response.text)
```

Okay, here's a summary of the provided household energy usage and cost data:

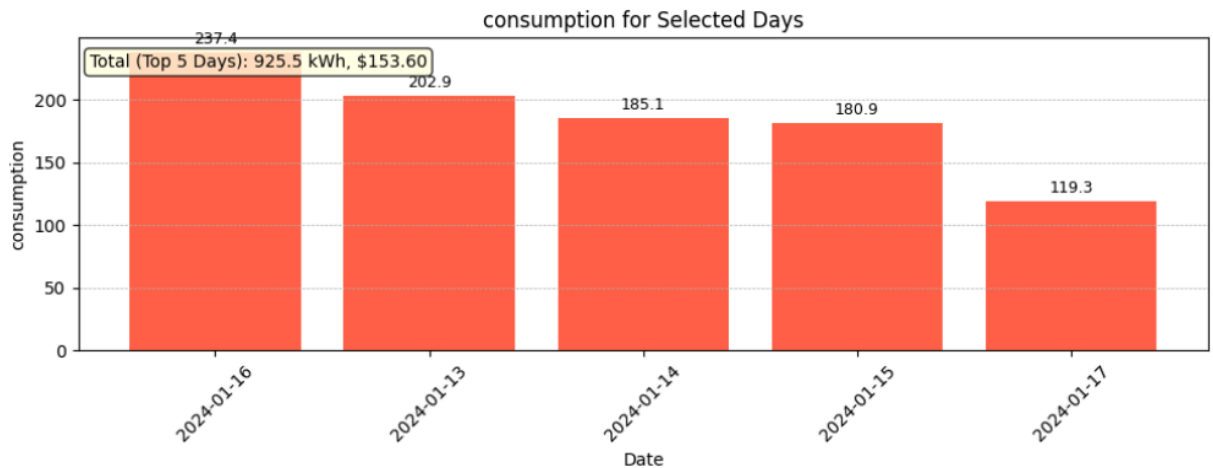
Summary of Household Daily Energy Usage & Cost (368 Days)

- Observation Period:** The data covers 368 days of energy usage and cost for the household.
- Average Usage & Cost:** On average (`mean`), the household consumed **33.11 kWh** per day, costing **\$5.50**.
- Typical Usage & Cost (Median):** A more typical day (the median or `50%` percentile) saw **25.38 kWh** of consumption, costing **\$4.21**. The fact that the average is higher than the median indicates that there are some days with significantly higher usage pulling the average up.
- Variability:** There's substantial variation in daily usage, as shown by the standard deviation (`std`) of **26.62 kWh** (and **\$4.42** for cost). This means daily consumption often deviates quite a bit from the average.
- Range:**
 - Lowest:** The lowest recorded daily consumption was **6.29 kWh** (costing **\$1.04**).
 - Highest:** The highest recorded daily consumption was extremely high at **237.43 kWh** (costing **\$39.41**). This maximum value is a significant outlier compared to the average and median.
- Middle Range (Interquartile Range):** Half of the days (`25%` to `75%`), the household's consumption fell between **17.18 kWh** and **39.65 kWh**, with costs ranging from **\$2.85** to **\$6.58**.
- Fixed Rate Confirmation:** As expected with a fixed-rate plan, the cost statistics directly mirror the consumption statistics. Dividing the mean cost by the mean consumption ($5.50/33.11kWh$) suggests a fixed rate of approximately $\$0.166$ per kWh*. This rate holds consistently across the min, max, and percentile values as well.

In short: The household's daily energy use varies considerably, with a typical day around 25 kWh (`4.21`), but occasional days see extremely high usage (upto `237kWh/39.41`). The cost is directly proportional to consumption at a fixed rate of about $\$0.166/kWh$.

Top 5 Energy Usage Days

	consumption	provided_cost
start_time		
2024-01-16 00:00:00+00:00	237.428000	39.405925
2024-01-13 00:00:00+00:00	202.870000	33.670334
2024-01-14 00:00:00+00:00	185.065000	30.715238
2024-01-15 00:00:00+00:00	180.857000	30.016836
2024-01-17 00:00:00+00:00	119.278000	19.796570



From the insights and analysis, we summarize the top energy usage days:

```
top_days = daily_data.sort_values('consumption', ascending=False).head(5)
low_days = daily_data.sort_values('consumption').head(5)

summary = f"""
Top energy usage days:\n{top_days.round(2).to_string(index=True)}

Lowest energy usage days:\n{low_days.round(2).to_string(index=True)}

Now provide insights and suggestions.
"""

response = chat.send_message(summary)
```

📊 Extreme Winter Energy Peaks

Total for Top 5 Days:

- 🔋 Energy Consumed: **925.50 kWh**
- 💰 Cost Incurred: **\$153.60**

Okay, let's analyze these high and low usage days:

Insights:

- Strong Seasonality:** There's a clear seasonal pattern. The highest usage days are clustered tightly in **mid-January**, a peak winter period in the Northern Hemisphere. Conversely, the lowest usage days occur in **June, August, and September**, typically milder months where heating and often cooling demands are minimal.
- Heating is Likely the Primary Driver of Peaks:** The extremely high consumption in January (reaching over 7 times the average daily use on Jan 16th!) strongly suggests **electric heating** (like a heat pump, electric furnace, or extensive use of space heaters) running heavily, likely during a cold snap. Heating systems are typically the largest energy consumers in households that use electric heat.
- Baseload Consumption:** The lowest usage days (around 6-8 kWh) likely represent the household's **baseload energy**

Suggestions for the Household:

1. Focus on Winter Heating Efficiency:

- **Thermostat Management:** If using electric heat, even a small reduction in the thermostat setting (e.g., 1-2 degrees Fahrenheit or Celsius) during winter can lead to significant savings. Consider a programmable or smart thermostat to automatically lower temperatures when away or asleep.
- **Weatherization:** Check for air leaks around windows, doors, electrical outlets, and ductwork. Sealing these leaks can significantly reduce how hard the heating system has to work. Ensure adequate attic and wall insulation.
- **Heating System Check-up:** Ensure the heating system (furnace, heat pump) is clean (especially filters) and functioning efficiently. Inefficient systems use more energy.
- **Space Heater Use:** If electric space heaters are used, ensure they are used strategically to heat only occupied areas and are not left running unnecessarily. They can consume a lot of power.

2. Reduce Baseload Consumption: While lower, even the 6-8 kWh on minimum days can potentially be reduced:

- **Phantom Loads:** Unplug electronics, chargers, and appliances when not in use, or use smart power strips.
- **Appliance Efficiency:** When replacing appliances, choose ENERGY STAR rated models, especially for refrigerators/freezers which run 24/7.
- **Lighting:** Switch remaining incandescent or CFL bulbs to LEDs.

● Interface with Weather App:

For further analysis and capturing the influence of weather conditions on energy usage, we added an interface with an external weather app to get the weather data for the year and have the LLM summarize the usage. This helps the user better understand the impact of external weather on usage behavior patterns that could cause those spikes.

```
top_day_weather_prompt = f"""
I am analyzing my top 5 highest energy usage days:

Energy usage summary:
{top_days_str}

Weather conditions on those days (hourly temperature in Fahrenheit):
{weather_summary}

Please help me understand:
- How might the temperature patterns explain my high energy usage?
- Are there behavioral or system-level improvements I can consider?
"""

response = chat.send_message(top_day_weather_prompt)
Markdown(response.text)
```

Here's the summarization response from Gemini pretrained LLM:

Okay, let's break down this data. The combination of your highest energy usage days and the corresponding temperatures provides very strong clues.

How Temperature Explains High Usage:

1. **Direct Correlation with Extreme Cold:** The four highest usage days (Jan 13-16) correspond directly to periods of intense cold. Average daily temperatures were only in the mid-teens Fahrenheit, with minimums dropping significantly lower (down to 6.6°F on Jan 14th). In such conditions, any heating system has to work incredibly hard and run almost continuously to maintain a comfortable indoor temperature, leading to massive energy consumption.
2. **Sustained Cold Stress:** The fact that these days are consecutive (Jan 13-16) is significant. After the first day or two of deep cold, the thermal mass of the house itself cools down, meaning the heating system has to overcome not just air infiltration but also heat loss *through* the building materials. This sustained stress likely contributes to the extremely high usage, particularly on Jan 16th, even though Jan 14th had a slightly lower average temperature.
3. **Heating System Thresholds (Especially for Heat Pumps):** If you have a heat pump, extremely low temperatures (often below 20-25°F, depending on the model) force it to rely heavily on less efficient auxiliary or emergency heat strips (usually electric resistance heat). This dramatically increases electricity consumption compared to the heat pump's normal operation mode. These temperatures clearly fall into that range.
4. **Lingering Effect (Jan 17th):** While Jan 17th was significantly warmer (average 29.6°F), the energy usage (119 kWh) was still very high compared to your overall average. This could be due to:
 - **Thermal Lag:** The house structure was still recovering from the preceding deep freeze and required substantial heating to maintain the setpoint.
 - **Setpoint Recovery:** You might have turned the thermostat up on the 17th after setting it lower during the coldest days, causing the system to run hard to reach the new target.
 - **Baseline:** Even an average near freezing still requires significant heating effort compared to milder weather.

Potential Improvements (Behavioral & System):

Behavioral:

1. **Thermostat Strategy:**
 - **Setbacks:** During extreme cold snaps, lowering the thermostat temperature significantly (e.g., by 5-10°F) when asleep or away can yield substantial savings, even if the system runs longer to recover.
 - **Avoid Overshooting:** Resist the urge to crank the thermostat way up when feeling cold. Set it to your desired comfortable temperature and let the system reach it gradually. Sudden large increases can sometimes trigger less efficient heating modes (like aux heat on a heat pump).
2. **Minimize Heat Loss:**
 - Be extra vigilant about keeping doors and windows closed. Minimize unnecessary opening/closing.
 - Ensure fireplace dampers are closed when not in use.
 - Use kitchen and bathroom exhaust fans sparingly, as they vent heated air outside.
3. **Auxiliary Heat Awareness:** If you used supplemental electric space heaters during this period, be aware they consume a *lot* of electricity (typically 1.5 kWh per hour each). Their use could explain spikes on specific days (like Jan 16th being higher than the colder Jan 14th). Use them strategically only in occupied rooms and for limited durations.

System-Level:

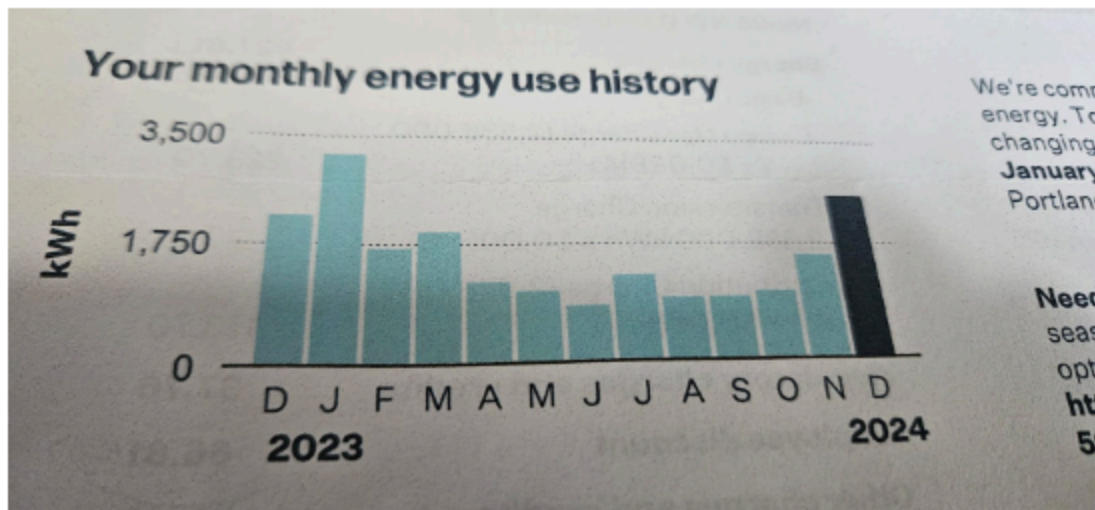
1. **Insulation (Crucial):** This is likely the biggest factor. Insufficient insulation (especially in the attic and potentially walls or basement/crawlspace) means heat escapes quickly, forcing the heating system to run constantly in cold weather. Assess your insulation levels (R-value) and consider upgrades.
2. **Air Sealing (Equally Crucial):** Small cracks and gaps around windows, doors, electrical outlets, plumbing penetrations, and recessed lighting can add up to significant heat loss (like having a window open). Sealing these leaks (caulking, weatherstripping) is often a very cost-effective improvement. A professional energy audit with a blower door test can pinpoint major leaks.
3. **Heating System Performance:**
 - **Maintenance:** Ensure your furnace or heat pump is serviced annually for peak efficiency. Clean/replace filters regularly (monthly during heavy use).
 - **Type/Efficiency:** Understand your system. Older systems are less efficient. If you have a heat pump, investigate its performance in deep cold and the settings for auxiliary heat activation. Sometimes these can be adjusted (though often limited by the equipment's capability at low temps).
4. **Windows & Doors:** Older, single-pane, or poorly sealed windows/doors are major sources of heat loss. Consider storm windows/doors, window insulation film, heavy curtains, or eventual replacement with energy-efficient units.

In summary, the data clearly shows your highest usage is driven by intense cold stressing your heating system. The most impactful long-term solutions involve improving your home's thermal envelope (insulation and air sealing) and ensuring your heating system is functioning optimally. Smart thermostat use provides additional savings on top of these fundamental improvements.

It has returned a very well structured explanation on the weather impact, usage behavior and key factors influencing the bills. It has provided concise behavioral and system level improvements for the user to understand how the usage can be streamlined and potentially save the bills!

Additionally we have explored the GenAI Image processing capabilities to reflect the usage parameters:


```
import PIL
from IPython.display import display, Image, Markdown, Audio
Image('/kaggle/input/neighborbill/Neighbor-Bill.jpeg', width=500)
```



```
from google import genai
from google.genai import types
from IPython.display import display, Image, Markdown, Audio

genai.__version__
```

```
'1.7.0'
```

	consumption	provided_cost
start_time		
January 2024	2252.38	373.83
February 2024	1259.13	208.98
March 2024	1323.14	219.60
April 2024	663.37	110.10
May 2024	588.09	97.61
June 2024	440.75	73.15
July 2024	509.68	84.59
August 2024	569.72	94.56
September 2024	643.60	106.82
October 2024	813.59	135.03
November 2024	1340.26	222.44
December 2024	1692.75	280.95
January 2025	89.19	15.25

Now that we have analysed both the user and his neighbor’s energy usage, let’s **compare** and see the results:

```
compare_prompt = f"""
Here is my neighbor's extracted monthly energy usage:

{neighbor_usage_summary}

Here is my own energy usage during the same months:

{monthly_text}

Please compare our energy usage and suggest why there might be differences. We both are in Portland, OR and are a two person household.
"""

response_compare = client.models.generate_content(
    model='gemini-2.0-flash',
    contents=compare_prompt
)

Markdown(response_compare.text)
```

Okay, let’s break down the energy usage comparison between you and your neighbor and explore some potential reasons for the differences.

1. Calculating Usage Ratios:

I'll calculate the ratio of your energy consumption to your neighbor's for each overlapping month. This gives us a sense of how different your consumption patterns are.

Month	Neighbor's Usage (kWh)	Your Usage (kWh)	Ratio (Your Usage / Neighbor's Usage)
January 2024	3200	2252.38	0.704
February 2024	1635	1259.13	0.770
March 2024	1938	1323.14	0.683
April 2024	1172	663.37	0.566
May 2024	1030	588.09	0.571
June 2024	800	440.75	0.551
July 2024	1258	509.68	0.405
August 2024	913	569.72	0.624
September 2024	921	643.60	0.699
October 2024	1005	813.59	0.809
November 2024	1560	1340.26	0.859
December 2024	2500	1692.75	0.677

The potential reasons for the differences in energy usage collated by the LLM:

2. Potential Reasons for Differences in Energy Usage:

Since you both live in Portland, OR, are a two-person household, and are neighbors, some factors can be ruled out (like extreme climate differences). However, several other variables could contribute:

- **Heating and Cooling Systems:**
 - **Type of Heating:** What type of heating do you each have (e.g., gas furnace, electric heat pump, electric resistance heating)? Electric resistance heating (like baseboard heaters) is generally the largest energy bill in the winter. If your neighbor has electric resistance heating and you have a heat pump or gas furnace, that could explain a significant difference.
 - **Thermostat Settings:** Do you keep your thermostat set at different temperatures? Even a few degrees difference can add up over time. Your neighbor might prefer a warmer house in the winter.
 - **Programmable Thermostats:** Do you or your neighbor use a programmable thermostat to automatically adjust the temperature when you are away or asleep?
- **Insulation and Home Efficiency:**
 - **Insulation Levels:** Are your homes similarly insulated (walls, attic, windows)? Poor insulation leads to greater heat loss in the winter and heat gain in the summer, requiring more energy to maintain.
 - **Window Efficiency:** Are your windows single-pane, double-pane, or low-E? Window efficiency significantly affects heating and cooling costs.
 - **Air Leaks:** Are there noticeable drafts around windows, doors, or other areas? Air leaks contribute to heat loss and gain.
- **Appliances:**
 - **Appliance Efficiency:** What is the age and energy efficiency of your appliances (refrigerator, washer, dryer, dishwasher)? Older, less efficient appliances use more energy. A very old refrigerator, for example.
 - **Appliance Usage:** How often do you use appliances like the washer/dryer or dishwasher? Your neighbor might do laundry or run the dishwasher more frequently.
 - **Electric Vehicle:** Does either home have an electric vehicle that is charged using home electricity?
- **Lighting:**
 - **Type of Lighting:** Do you use mostly LED bulbs, or are some incandescent or older fluorescent bulbs still in use? LEDs are far more energy-efficient.
 - **Lighting Habits:** Do you or your neighbor tend to leave lights on in rooms that are not in use?
- **Water Heating:**
 - **Type of Water Heater:** Is it a gas or electric water heater? An electric water heater can be a significant energy user.
 - **Water Heating Habits:** How long are your showers? Do you take baths frequently? Using more hot water increases energy consumption.
- **Other Electrical Devices:**
 - **Computers/Gaming Consoles:** Do you have a lot of computers, gaming consoles, or other electronics that are left on or in standby mode? These can draw power even when not actively used.
 - **Pool/Hot Tub:** Does either home have a pool or hot tub that requires electricity to heat and maintain?
- **Home Size and Layout:**
 - If your homes are significantly different in size (square footage), this could contribute to different energy usage, but if they are next door to each other, this is less likely.

3. Are the Usage Levels Typical for Similar Homes in Portland, OR?

It's hard to say definitively without more information. However, we can make some general assessments:

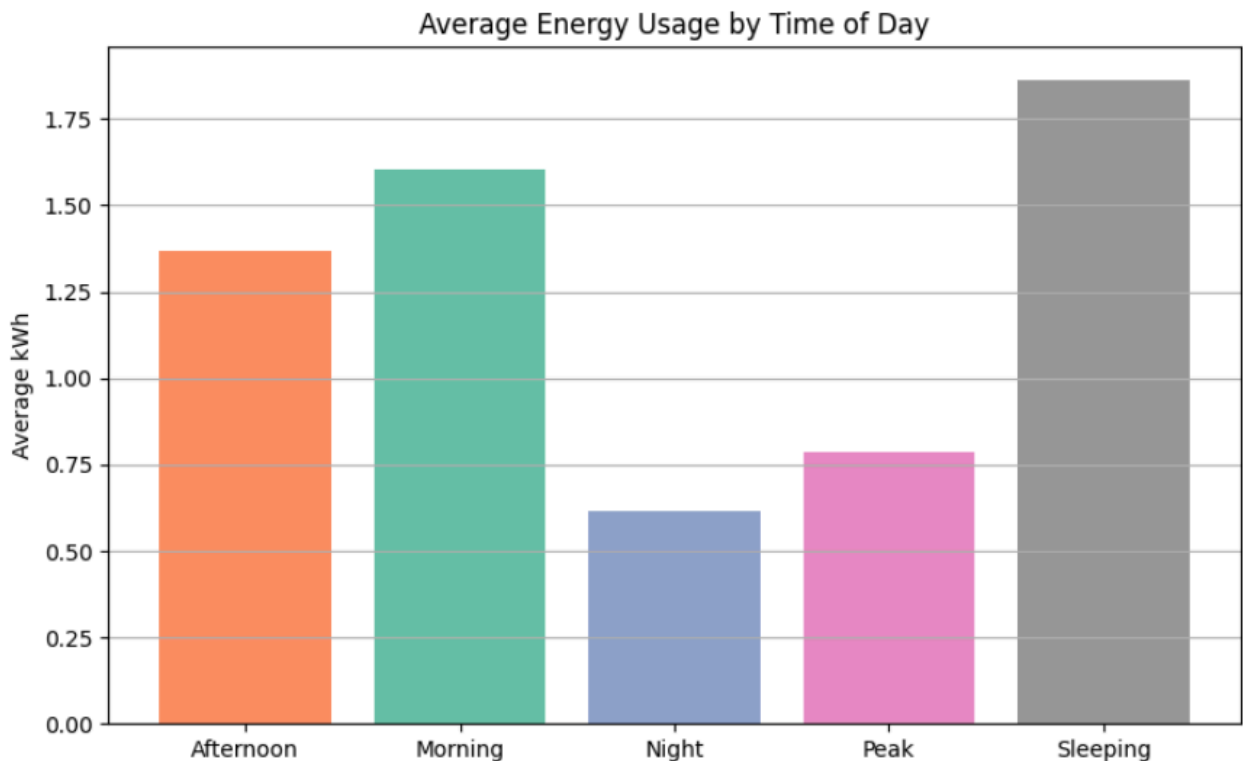
- **Your Usage:** Your consumption is on the lower end, especially in the summer months. It is likely that you are using less than an average household.
- **Neighbor's Usage:** The neighbor's January and December usage (~3200 kWh and ~2500 kWh) seem *high*, especially for a two-person household in Portland. This suggests potential inefficiencies or different usage patterns as outlined above. The lower values in Spring through Fall (900-1200 kWh) are more in line with typical usage for the area.

Recommendations:

1. **Talk to Your Neighbor (If You're Comfortable):** A casual conversation about energy bills might reveal some obvious differences in habits or equipment.
2. **Energy Audit:** Consider getting an energy audit of your home. Many utility companies offer them, sometimes for free or at a reduced cost. They can identify specific areas where you can improve your home's energy efficiency.
3. **Analyze Your Own Habits:** Track your energy usage more closely and see if you can identify any areas where you can cut back.
4. **Compare to Averages:** Look up average energy usage for homes of similar size and type in Portland. This can give you a better benchmark for comparison.

By carefully considering these factors, you can start to understand why your energy usage differs from your neighbor's and identify ways to potentially reduce your own consumption if desired.

Finally capturing the **average and zero energy usage by time of the day** to detect the lower spikes and usage pattern for the day:



```
usage_summary = grouped_df.to_string(index=False)

summary = f"""
Here is my average energy usage during different times of day:

{grouped_df.to_string(index=False)}

Please analyze which period uses the most energy and suggest 1 or 2 behavioral changes to reduce peak usage.
"""

response = chat.send_message(summary)
Markdown(response.text)
```

Okay, let's analyze this time-of-day usage:

Analysis:

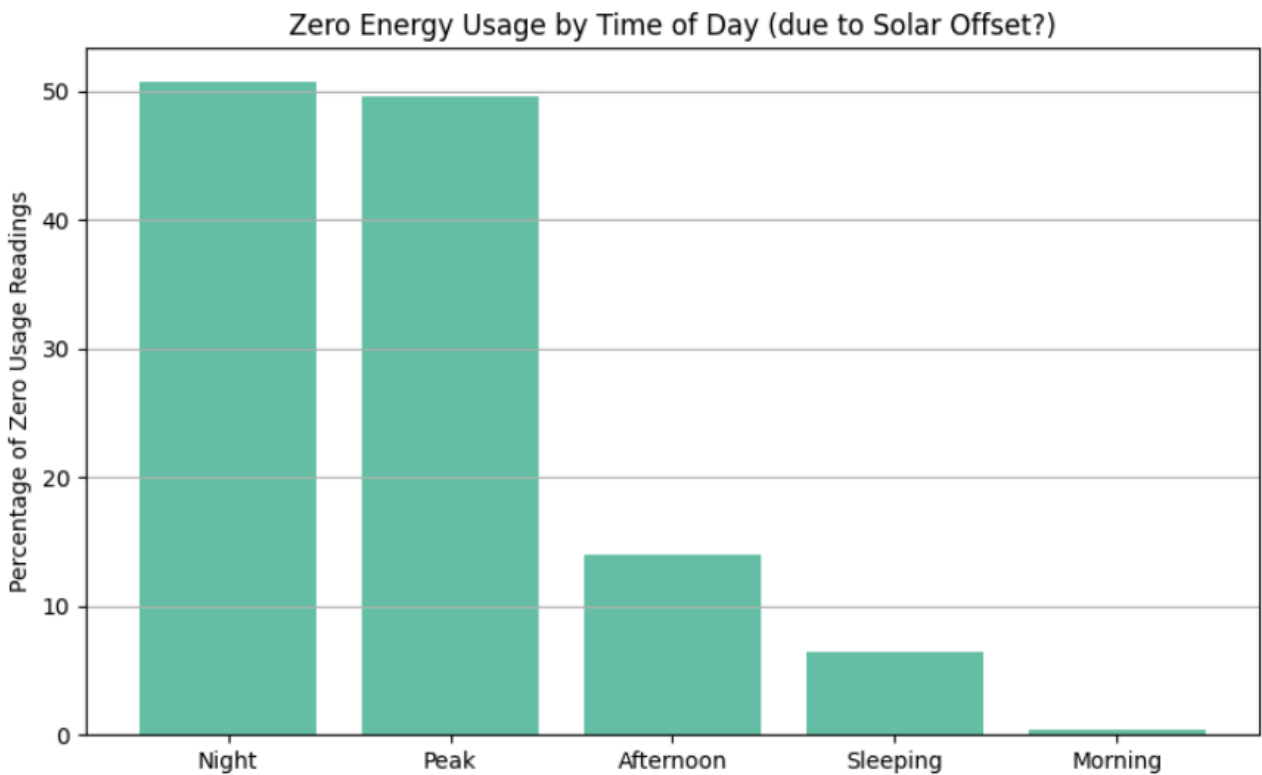
- Based on the average kWh provided, the **"Sleeping" period (1.86 kWh average)** shows the highest energy consumption. This is somewhat counter-intuitive, as household activity is usually lowest the continuous load running overnight, most likely **heating or cooling (HVAC)** working to maintain the temperature setpoint, especially if the house is not well-insulated or if there's a large temperature change. Refrigeration also runs continuously, but HVAC is usually the dominant factor for such high overnight usage, particularly if you have electric heating running in winter.
- The **"Morning" period (1.60 kWh average)** is the next highest, likely due to waking activities, lights, appliances for breakfast, and the HVAC system potentially adjusting the temperature after an overnight period.
- The period labeled **"Peak" (0.78 kWh average)** shows lower usage compared to "Sleeping", "Morning", and "Afternoon". This might indicate the label "Peak" refers to a utility's peak pricing hours rather than consumption period on average according to this data slice.

Behavioral Suggestions to Reduce Peak Usage:

Given the highest average usage is during the "Sleeping" period, but typical "peak reduction" strategies target late afternoon/early evening (which might fall into your "Afternoon" or labeled "Peak" times):

- Adjust Overnight Thermostat Settings:** Since the "Sleeping" period shows the highest average use (likely HVAC), a key behavioral change is to adjust your thermostat significantly during these hours or warmer in summer while you sleep. Even a few degrees can reduce how often the HVAC system needs to run overnight, directly lowering consumption during this high-usage period identified in your data.
- Shift Large Appliance Use:** While your "Sleeping" average is highest, many utilities focus on reducing strain during afternoon/evening peaks. Get in the habit of running high-consumption appliances **dryer outside of the late afternoon/early evening hours** (e.g., run them later at night if that doesn't conflict with the high "Sleeping" usage cause, or during the middle of the day if possible). This targets lower usage during your "Afternoon" or labeled "Peak" periods.

By managing your thermostat overnight and shifting appliance loads away from common peak times, you can target both the highest average usage shown in your data and the periods typically most impacted by time-of-use pricing.



	Time of Day	Zero Usage Count	Total Count	Zero Usage %
0	Night	559	1101	50.772025
1	Peak	546	1101	49.591281
2	Afternoon	309	2202	14.032698
3	Sleeping	143	2202	6.494096
4	Morning	9	2203	0.408534

● Evaluation Metrics

To assess the AI model's effectiveness - GenAI components, evaluation metrics used:

- **Prompt response relevance rating**
- **Accuracy score**
- **Groundedness**

using **LLM as a judge** evaluation technique.

Human in the loop approach - Verification with actual utility bills

User Satisfaction (qualitative): A future direction for deployment feedback loop.

```

import enum

# Define the evaluation prompt
SUMMARY_PROMPT = """\
# Instruction
You are an expert evaluator. Your task is to evaluate the quality of the responses generated by AI models.
We will provide you with the user input and an AI-generated responses.
You should first read the user input carefully for analyzing the task, and then evaluate the quality of the responses based on the Criteria provided.
You will assign the response a rating following the Rating Rubric and Evaluation Steps. Give step-by-step explanations for your rating, and only provide the rating at the end.

# Evaluation
## Metric Definition
You will be assessing summarization quality, which measures the overall ability to summarize text. Pay special attention to length constraints, summarization quality, and the ability to capture the main points.

## Criteria
Instruction following: The response demonstrates a clear understanding of the summarization task instructions, satisfying all of the instructions.
Groundedness: The response contains information included only in the context. The response does not reference any outside information.
Conciseness: The response summarizes the relevant details in the original text without a significant loss in key information without being too verbose.
Fluency: The response is well-organized and easy to read.

## Rating Rubric
5: (Very good). The summary follows instructions, is grounded, is concise, and fluent.
4: (Good). The summary follows instructions, is grounded, concise, and fluent.
3: (Ok). The summary mostly follows instructions, is grounded, but is not very concise and is not fluent.
2: (Bad). The summary is grounded, but does not follow the instructions.
1: (Very bad). The summary is not grounded.

## Evaluation Steps
STEP 1: Assess the response in aspects of instruction following, groundedness, conciseness, and verbosity according to the criteria.
STEP 2: Score based on the rubric.

```

```

# Define a structured enum class to capture the result.
class SummaryRating(enum.Enum):
    VERY_GOOD = '5'
    GOOD = '4'
    OK = '3'
    BAD = '2'
    VERY_BAD = '1'

def eval_summary(prompt, ai_response):
    """Evaluate the generated analysis against the prompt used."""

    chat = client.chats.create(model='gemini-2.5-pro-exp-03-25')

    # Generate the full text response.
    response = chat.send_message(
        message=SUMMARY_PROMPT.format(prompt=prompt, response=ai_response)
    )
    verbose_eval = response.text

    # Coerce into the desired structure.
    structured_output_config = types.GenerateContentConfig(
        response_mime_type="text/x.enum",
        response_schema=SummaryRating,
    )
    response = chat.send_message(
        message="Convert the final score.",
        config=structured_output_config,
    )
    structured_eval = response.parsed

    return verbose_eval, structured_eval

text_eval, struct_eval = eval_summary(prompt=top_day_weather_prompt, ai_response=weather_energy)
Markdown(text_eval)

```

Evaluation of the prompt and its response from the image - bill submitted to the LLM:

```
prompt = [
    f"This is my neighbor's electric bill. Please extract the monthly energy usage (in kWh) for all the months. EXAMPLE Output December 2023:
    PIL.Image.open("/kaggle/input/neighborbill/Neighbor-Bill.jpeg")
]

response = client.models.generate_content(
    model='gemini-2.5-pro-exp-03-25',
    contents=prompt
)
neighbors_energy = response.text
Markdown(response.text)
```

```
text_eval, struct_eval = eval_summary(prompt=[prompt, neighbor_usage_summary], ai_response=neighbors_energy)
Markdown(text_eval)
```

- **STEP 1: Assess the response in aspects of instruction following, groundedness, conciseness, and verbosity according to the criteria.**

- **Instruction following:** The AI successfully followed the instructions. It extracted the monthly energy usage in kWh for all the months presented in the kWh).
- **Groundedness:** The response is grounded in the provided image (the bar chart). The extracted values are reasonable visual approximations of the bars they appear accurate or appropriately rounded (e.g., Dec 2023 is just above 2000, Jan 2024 is clearly ~3200, Feb 2024 is around 1700, etc.). The AI's value matches the prompt text.
- **Conciseness:** The response is concise, providing only the requested information without unnecessary text.
- **Fluency:** The response is well-formatted as a list and easy to read.

- **STEP 2: Score based on the rubric.**

- The summary follows instructions perfectly.
- The summary is fully grounded in the provided image.
- The summary is concise.
- The summary is fluent.
- Therefore, the rating is 5.

Rating: 5

● Limitations

- Limited real-time user feedback loop for adaptive learning
- Dependence on high-quality time-series data
- Data is looking back - The problem has occurred and energy has been used. Past consumption.
- LLM does not yet have all the data such as thermostat data.
- Generalization to different geographies or household types may require model fine-tuning
- Dataset - from one or two consumers from the same neighborhood and similar household type.
- Some energy usage is in gas and not electricity alone.

● Future Enhancements

- Integrate smart meter API data for real-time coaching
- Introduce reinforcement learning for dynamic tip generation
- Deploy as a web-based AI assistant dashboard
- Connect with home automation systems for action-oriented insights
- Additional energy sources and data will improve the results.

- Conclusion

WattSeer personalized Energy coach exemplifies how GenAI can be a game changer in sustainability. By integrating explainable AI with personalized insights, it goes beyond dashboards—coaching individuals toward more energy-efficient lifestyles. It demonstrates how GenAI can bridge the gap between complex energy data and end-user understanding. It reflects on impact, next steps, and potential to scale this to households or community energy planning.

The result is a highly interpretable, intuitive, and empowering energy experience.

- References

Day 1 and Day 2 Prompt templates, example input/output

Structured output turned into visual/summary form