Nafas

Air Purifier Usage Patterns and PM2.5 Trends

A Case Study by Sai Sreekar Vankayalapati

Contents

Problem Statement: Optimizing Purifier Usage Patterns

The Challenge

My data analysis reveals significant inconsistency in how customers use their air purifiers.
Many users alternate between two problematic behaviors:

Device Underutilization (not using the device when they should) and

Unnecessary Operation (using the device even when unnecessary.)

Business Impact

This usage pattern inconsistency creates a critical business challenge:

- **Diminished Health Benefits:** Inconsistent usage prevents customers from experiencing the full health benefits of clean indoor air.
- **Brand Perception Risk:** When customers don't see improved air quality despite owning our products, they may question device effectiveness rather than usage habits.
- **Customer Dissatisfaction:** Higher energy bills and premature device wear may negatively impact customer satisfaction and lifetime value.

Dataset at a Glance

Overview

- 7 users, 3 device types (Pure40, Pure60, AirTest)
- 30-day period, 1-minute intervals
- Key fields: Timestamp, Power, Mode, Speed, PM2.5, Room
- Derived features: is_active, hour, weekday/weekend

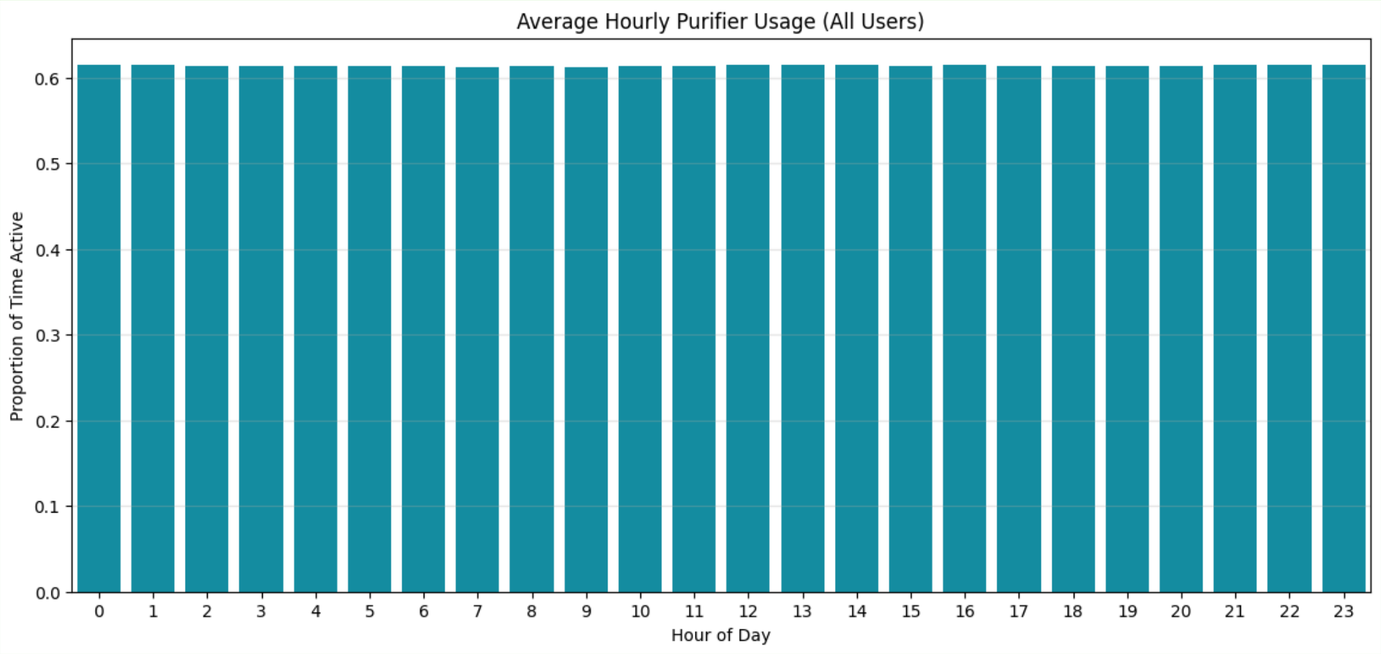
User-Device Mapping

Member UUID	Devices Owned
---	---
1f8a0dd4-72c2-4ff4-bfb9-8b8ded756ec5	AirTest x1; Pure40 x4
2e13b5a2-99eb-4277-a51b-d37c765a0a2e	AirTest x4; Pure40 x1
2f04498d-3408-4bd9-a545-653364db22fa	Pure40 x1
4c5f4c12-a85f-46e8-b075-b58e63b7bf71	Pure40 x5
5ab22bb7-c837-4703-acf6-b9137820d670	Pure40 x1
2dd89330-253f-4659-90cc-55ee33bcdb38	Pure60 x1
5d5c95ae-47d0-4194-b0d0-d2a29b6cd3f8	AirTest x1; Pure60 x1

Assumptions

- Outdoor PM2.5 levels <12 are considered safe and >35 are considered unsafe. Data is analyzed using the provided indoor pm2.5 levels (as there is no outdoor data given), assuming that the indoor values also follow similar trends to outdoor PM2.5
- Due to the low amount of data points, there is no significant variance in different days of the month.
- Higher speed consumes higher electricity.
- As there is no geographical data provided and the number of users is small, individual analysis for each user and device type was done.
- One user (5d...) contained multiple mode values for the same timestamps. The device was assumed to be off.

1. Are We Purifying Smartly or Just Constantly?



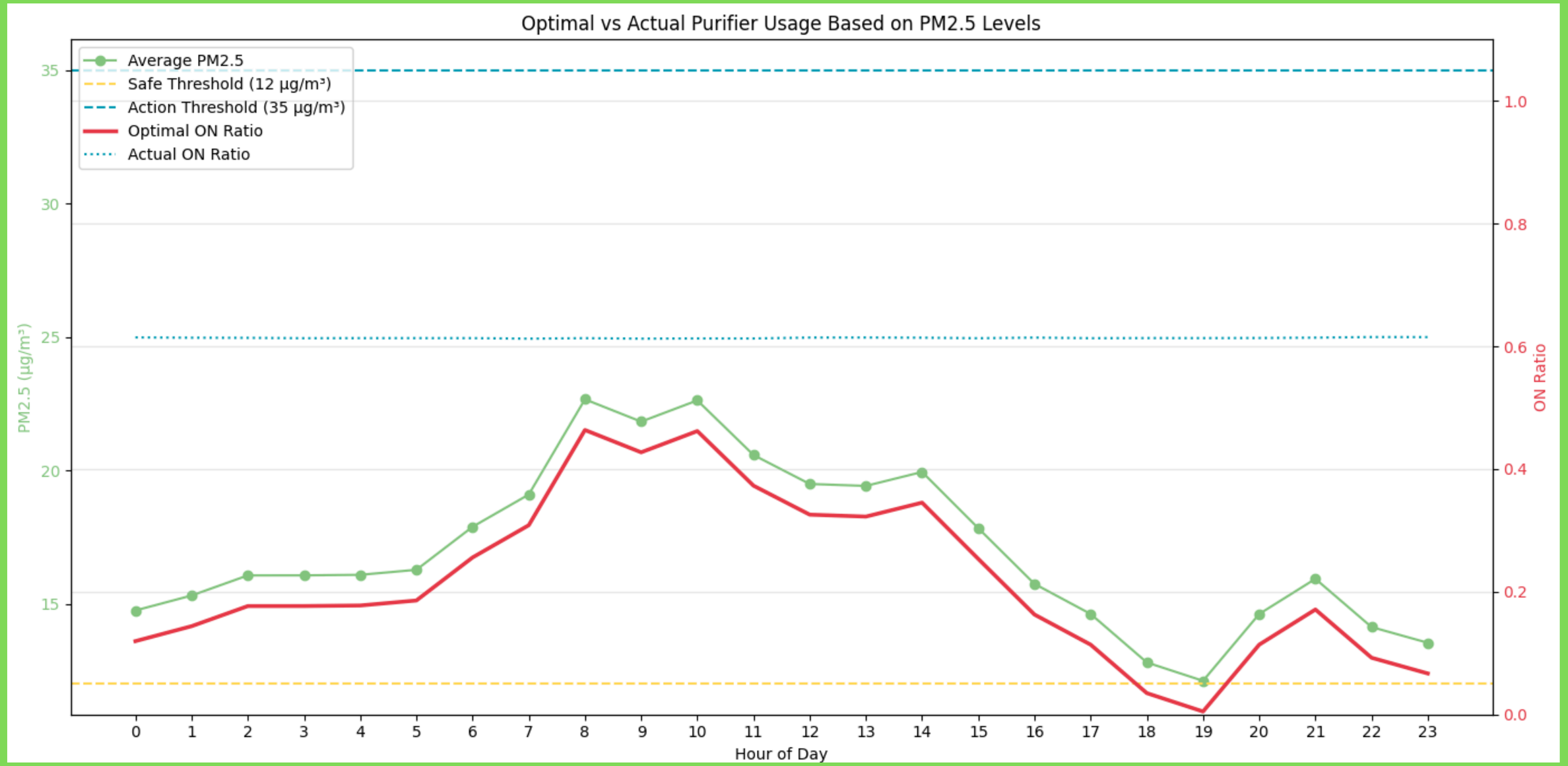
(ON rate = number of rows with power 1 / total rows

Observation

The average ON-rate is ~0.63, meaning that on average, a purifier is active around 63% of the times at any given hour— with no change across the day.

Interpretation

Usage doesn’t correlate to pollution spikes, which may lead to wasted energy or ineffective purification.



Note: The optimal PM2.5 values are for outdoor PM2.5. Indoor PM2.5 is typically much lower than outdoor PM2.5 (as is evident in the graph)

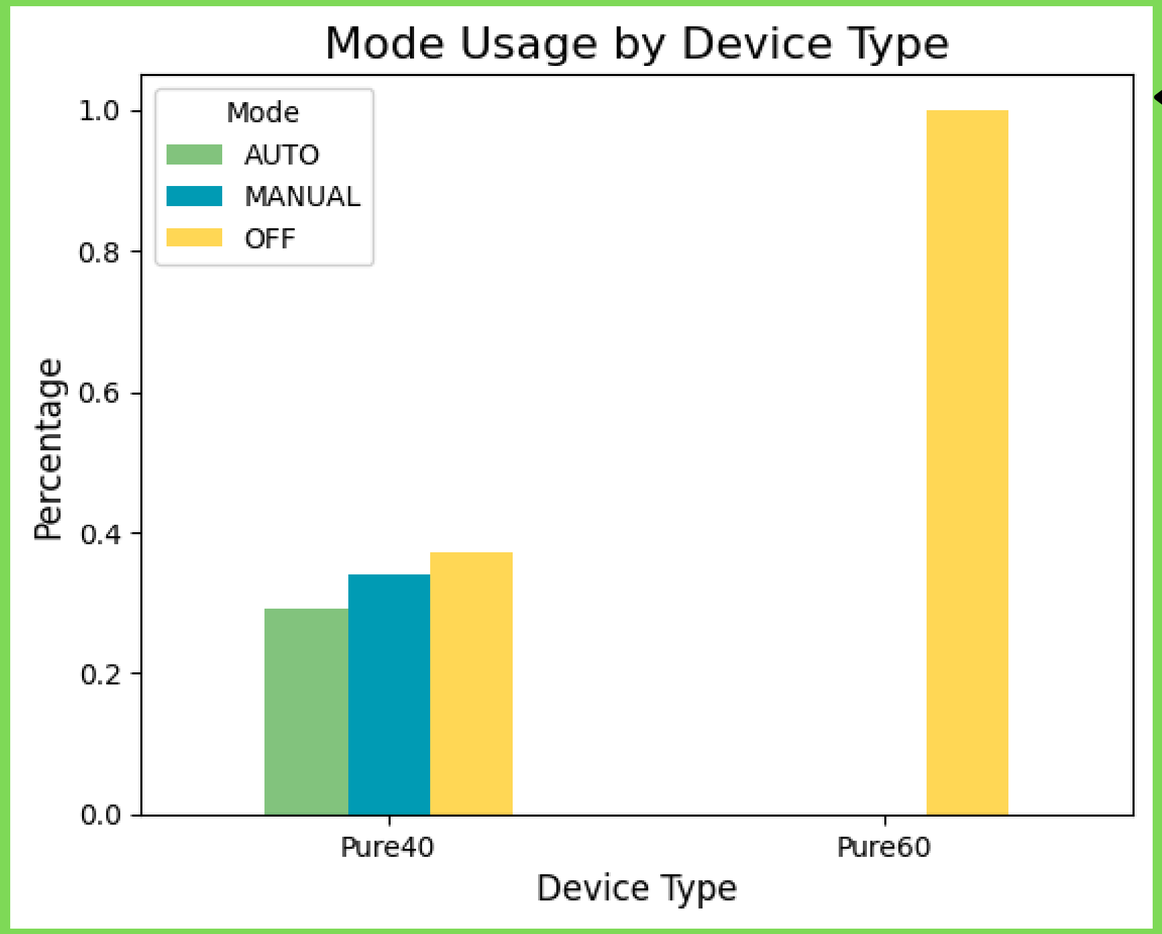
<-- What Can We Observe from that Plot?

- It is clear that the usage does not correlate to PM2.5 levels or spikes.
- The optimal usage should follow the PM2.5 levels rising and following, either exactly or pre-emptively (starting the purifier before the pollution spikes) but that is not the case.
- PM2.5 value reaches it's highest between 8 AM - 10 AM, most likely due to people commuting to work/school.
- Another similar spike in PM2.5 is visible during the night (8PM - 10PM)
- One interesting observation in the indoor PM2.5 levels is another spike at around 1-2 PM. This could possibly be due to cooking or cleaning around those times which are known to increase particulate matter levels in the air.

Insights:

During peak hours, it is better to use the purifiers in high speeds, whereas during the nights or other stable times, low speed auto mode would suffice.

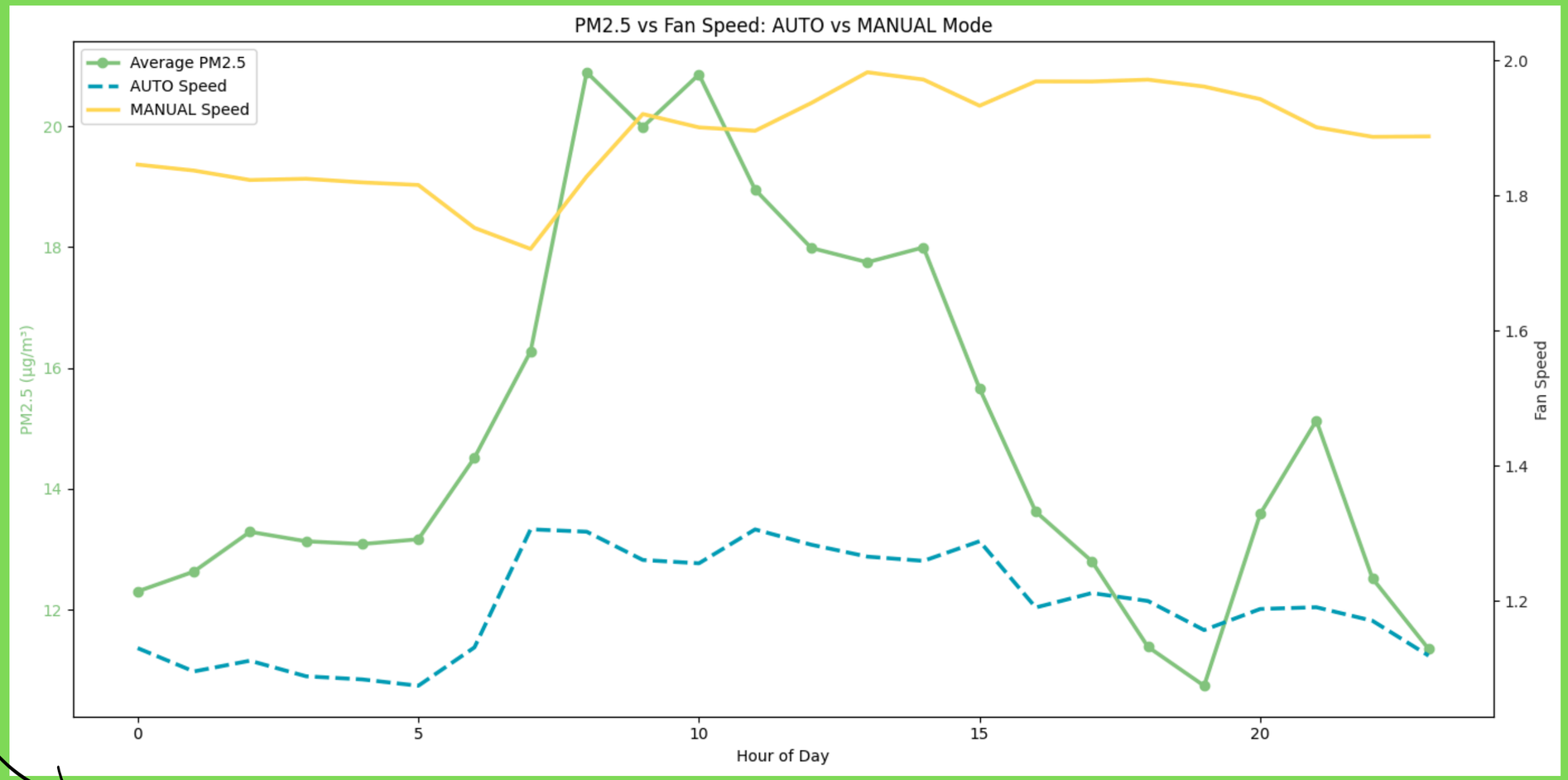
2. How are Different Device Types and Modes Used?



- **Pure60 device is not used at all.**
- **Pure40 device is also turned off more than it is on.**
- **Pure40 is not used much in Auto mode as well.**

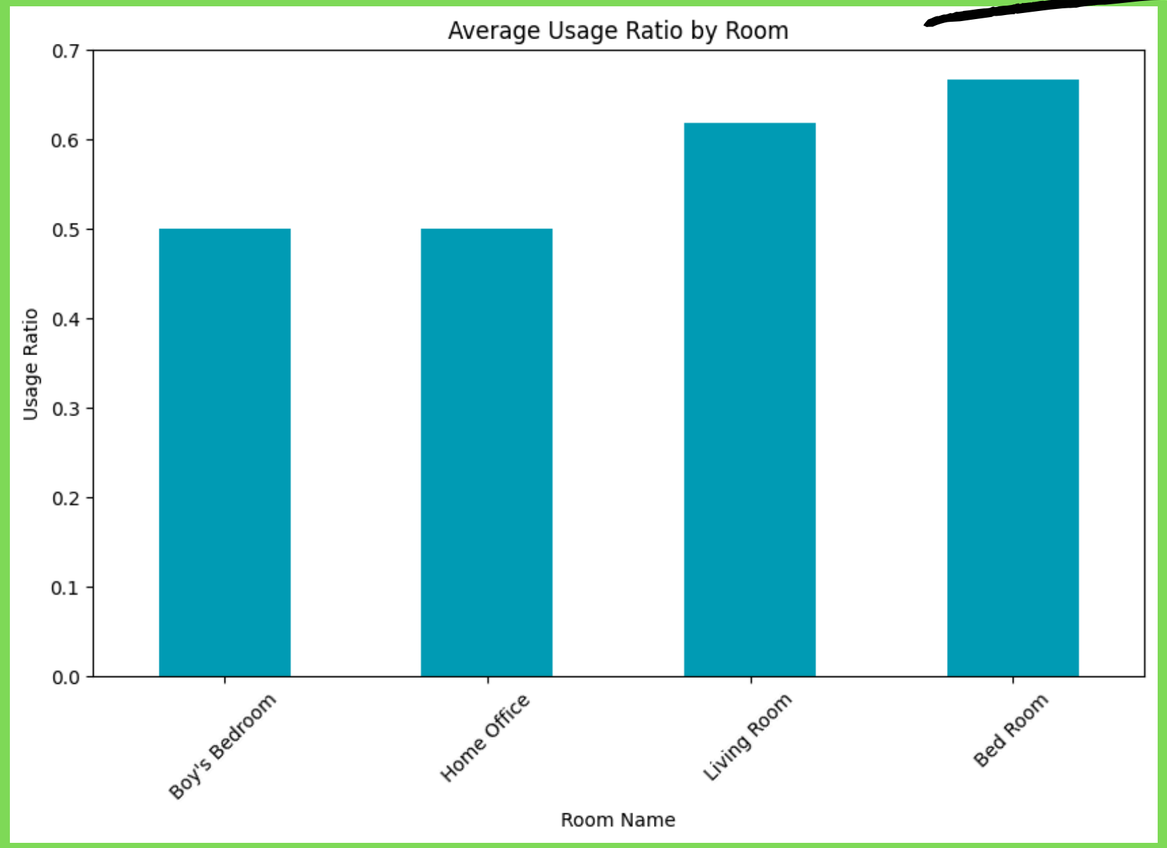
Interpretation:

- **The above results shows that most of the given users are not proactive**
- **For such users, Auto mode would benefit them a lot, however, there seems to be very little usage of Auto mode. This could potentially indicate a lack of trust in the Auto mode.**



- In order to understand Purifier usage in more detail, it is important to look at Manual Mode vs Auto Mode usage.
- From the above plot, it can be observed that Auto mode works almost perfectly in sync with PM2.5 levels, with only minute differences. It could be potentially made better by leveraging other features such as external PM2.5, user patterns/preferences, etc.
- Manual mode is relatively less accurate with more error, which could be a problem.

3. But WHY is the Usage Like This?

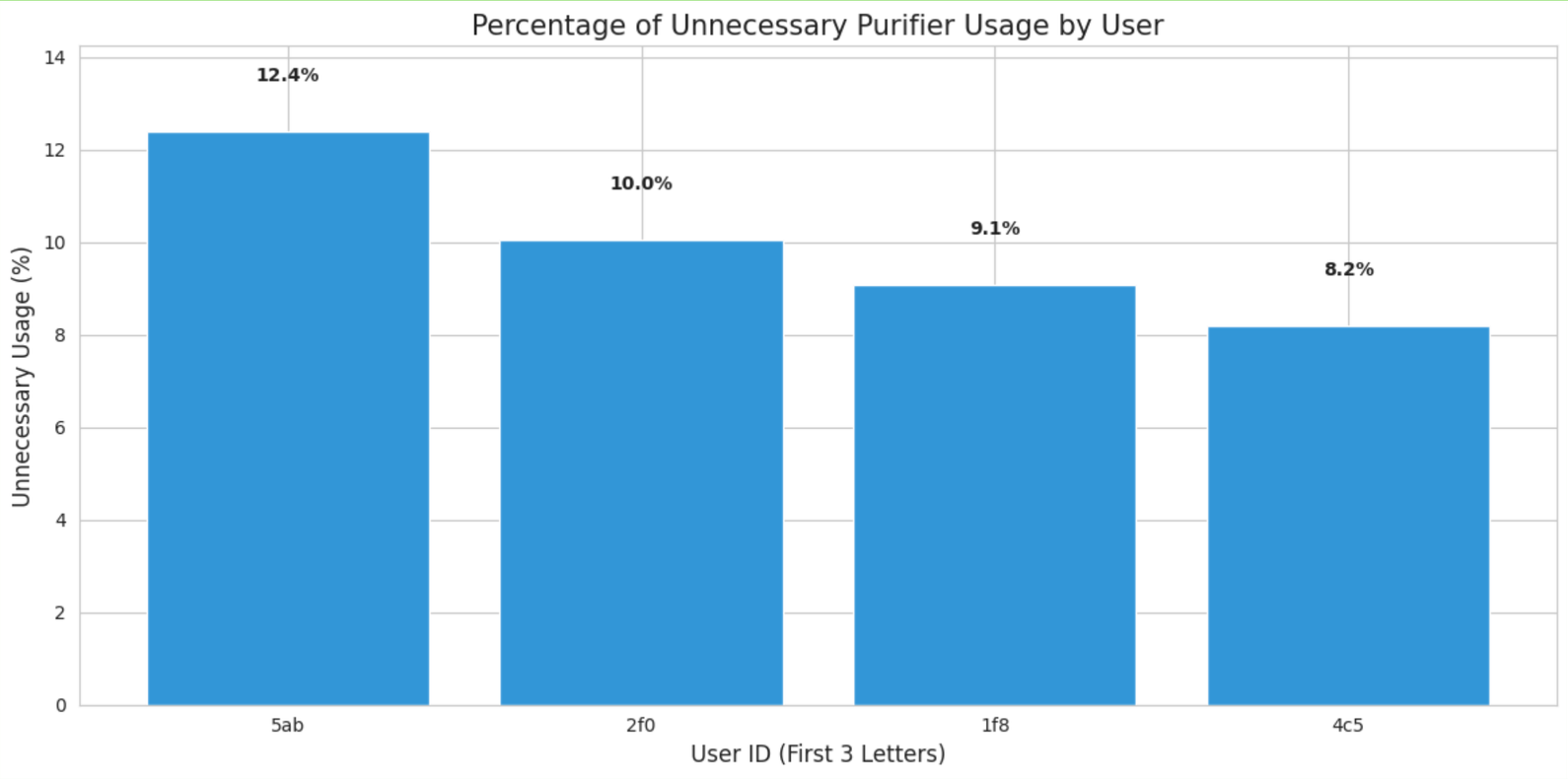


- It can be seen that Home Office and Baby's Bedroom have the least purifier usage.
- This could be due to adults not using those rooms much, hence forgetting to turn on the devices there.
- Similarly, Bedroom and Living room have higher usage, possibly due to the users being in those rooms more.

Interpretation:

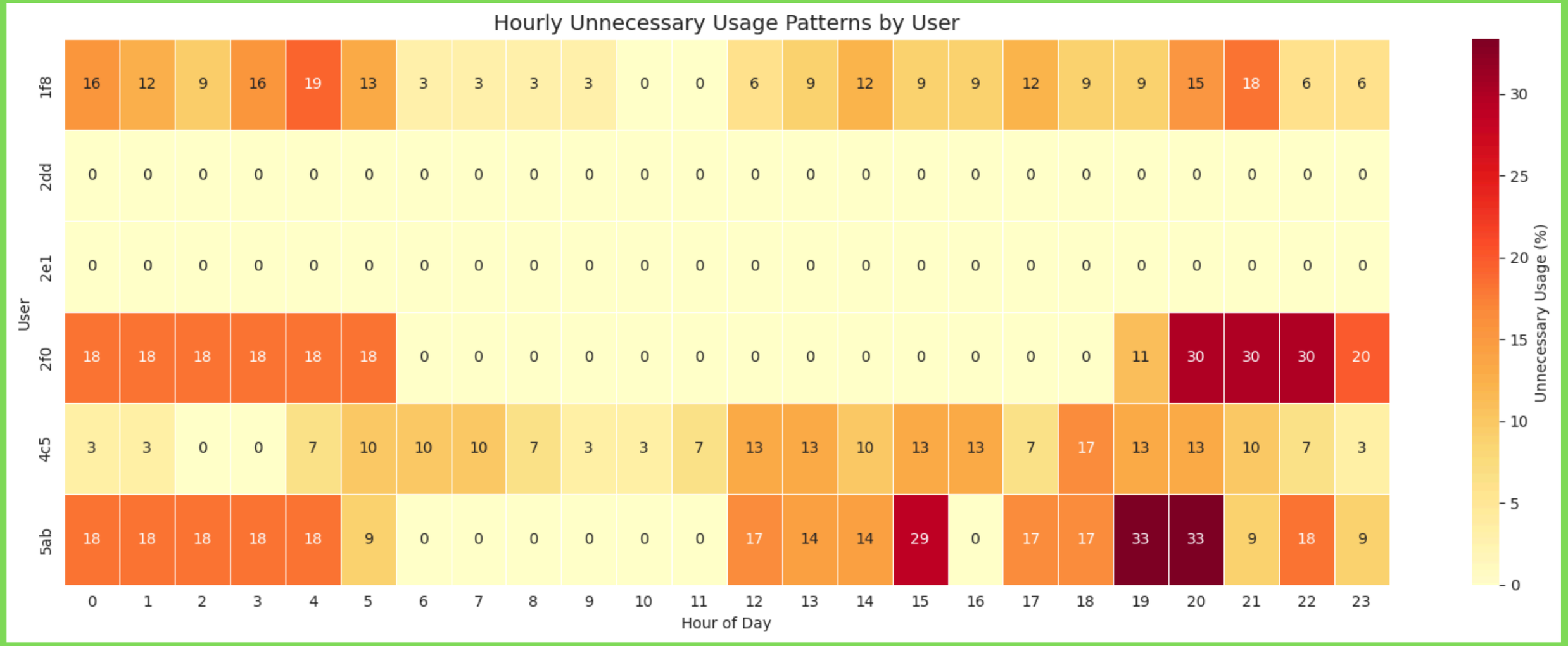
- Having polluted air in baby's room is potentially harmful and must be addressed.
- As we saw, night-time pollution levels are low and stable and the users could be benefitted by changing their devices to Auto mode or low speed settings

4. Can We Save Our Customers Some Money?



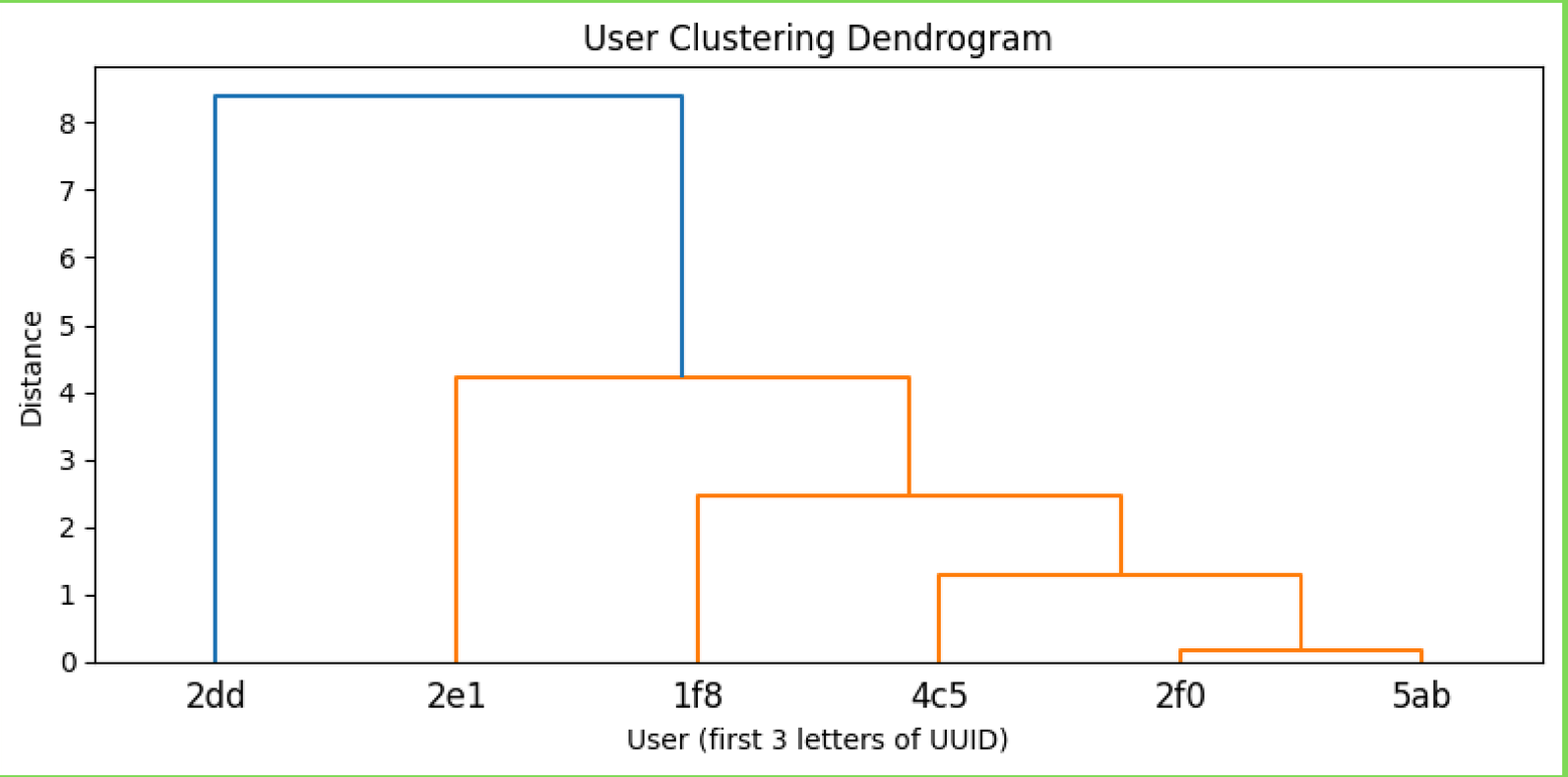
- Even when PM2.5 inside is less than 2, **some users are still keeping the purifiers on,** which is a **waste of electricity.**
- We could save these unnecessary usage if the device is turned off

Assumed Thresholds: Lower = 2 and Upper = 20.



- **The heatmap above shows during which hours of the day purifiers are unnecessarily being used (given the assumed thresholds)**
- **If these are removed, the user could save quite some money, improving their experience with our device.**

5. Do Any Users Have Similar Behaviours?



- This diagram can help us visualize the “distance” between the usage patterns of all users.
- The smaller the distance, the more similar the user behaviours are.
- It can be observed that users 5ab, 2f0 and 4c5 are relatively similar while user 2dd is an outlier having totally different usage.

Interpretation:

- Similar behaviours could help us give them similar solutions.

Summary

- Purifier usage is roughly constant for each hour of the day and for each day of the month.
- Purifier usage is different for each room of the house.
- Pure60 device is not used at all.
- Pure40 is mostly used in manual mode, but it is also off for most of the time.
- Very little utilization of Auto mode by these users.
- Auto mode perfectly matches the indoor PM2.5 variations
- Actual usage is very different from Optimal usage, mainly due to users using in manual mode.
- Some users are using the purifiers when they need not be, which wastes electricity.
- Users overall don’t seem to be proactive and could be benefitted with reminders or awareness about how to use the devices better

Recommendations to Solve Underutilization

Increase Auto Mode Adoption:

- Add in-app explanations and real-time feedback about why the purifier is changing speed.
- Auto Mode can also take user preferences into account (Do they want to be more economical vs faster purification, etc.)
- can motivate people to be more proactive with their usage. For instance, we can create a "Purifier Usage Score" -- similar to Sleep Score -- that conveys to the user how well they are using the devices. Using a simple numeric metric makes it easy for the user to track and improve their usage. Explore possible rewards/discounts for users having a good score and neighbourhood leaderboard systems.

Improve User Education and Engagement:

- Send informative articles and real-time news to help users understand when and why to run their purifiers.
- Send personalized notifications to turn on/off, to open/close windows, educational tidbits, etc. (For reference: Similar approach to Duolingo)

Improving Auto Mode Functioning by Building a Better Forecasting Model

Recommendations to Solve Overutilization

Reduce Unnecessary Usage:

- Implement smart scheduling and auto-shutdown during low PM2.5 periods (Auto Mode Improvements).
- Send alerts when users are running purifiers at high speed unnecessarily for Manual users.
- Introduce an energy-saving mode that caps fan speed when air is already clean.

Targeted Alerts:

- Notify users if their purifier is running for long periods with low PM2.5, suggesting they turn it off or lower the speed.

General Recommendations

- **Optimize Device Placement and Hardware:**
  - Use room-specific presets and motion sensors to ensure purifiers are active in occupied or critical rooms (like bedrooms at night).
  - For future products, consider hardware upgrades (e.g., quieter motors, better sensors).

- **Leverage Data for Continuous Improvement:**
  - Use analytics dashboards and A/B testing to refine algorithms and user engagement strategies.
  - Regularly review usage and air quality data to identify new opportunities for improvement.

- **Filter and Maintenance Reminders:**
  - Remind users to replace filters based on actual usage for sustained performance.

Future Scope

- Leveraging outdoor data such as meteorological features to do more in-depth analysis.
- Geographical analysis using coordinates and other location-based features for the users.
- Improving the Forecasting model for PM2.5, usage, etc. using more features such as user preferences, meteorological variables, etc. Transformer models can be explored as they are proven to work best for such Time series forecasting problems. This would significantly boost the Auto Mode experience.
- Graph Neural Networks can be studied based on location data to learn interactions between locations and other spatial patterns.
- Econometrics of the device usage can be studied by analyzing costs of electricity in said locations and estimates can be drawn on how much electricity can be saved.
- Extensive Exploratory data analysis using features such as age, income, etc. to understand the customer demographics better