Project | Sustainability Impact Analysis for Intel

INTRODUCTION: As you learned listening in on the strategy meeting with Dr. Alvarez and Intel's Sustainability Team, Intel is committed to reducing its carbon footprint and improving the sustainability of its devices – not just during manufacturing, but throughout the entire lifecycle.

A key part of this effort is their repurposing programs, which play a central role in achieving these sustainability goals. Repurposing and recycling programs aim to reduce e-waste, energy consumption, and CO_2 emissions by extending the life of existing devices, and thus reducing the need for new device manufacturing. Like Michael Campbell said: the average household in the US has anywhere from 3–5 PCs devices, tablets, notebooks, desktops that are perfectly functional, but not being used!

One challenge Intel faces is determining which devices in its repurposing program should be prioritized for the maximum environmental benefit. That's where data analysis comes in! To help with this, Intel gathered data on each device repurposed or recycled in 2024.

Your task is to evaluate the effectiveness of Intel's current repurposing strategy and provide a data-driven recommendation to help guide the program's direction and optimize sustainability efforts.

HOW IT WORKS: Follow the prompts in the questions below to investigate the data. Post your answers in the provided boxes: the **yellow boxes** for the queries you write and **blue boxes** for your text-based analysis. Once you're done, you'll submit your ***completed*** .pdf file to HQ for feedback from The Accelerator Team.

SQL App: Here's the link to our specialized SQL app, where you'll write your SQL queries and interact with the data.

NOTE: The dataset you are working with is designed for The Global Career Accelerator to reflect the key characteristics and structure of Intel's real data, while protecting their confidentiality and proprietary information. Be aware that any conclusions or results derived from this dataset should be viewed as hypothetical and for illustrative purposes only.

— Data Set **Descriptions**

In this project you'll query 2 different datasets, `intel.device_data` and `intel.impact_data`, that you will join together for your analysis. Here you'll find the data dictionary for each dataset.

`intel.device_data`

- `device_id`: Unique identifier for each repurposed device
- `device_type`: Type of device, values are either "Laptop" or "Desktop"
- `model_year`: The year the device was manufactured (e.g., 2018, 2019, etc.)

`intel.impact_data`

- `impact_id`: Unique identifier for the repurposed device's impact record (e.g., "LP20NA141592")
- `device_id`: Unique identifier linking the impact record to a specific device in the `intel.device_data` table
- `usage_purpose`: The specific purpose for which the device is being repurposed, values are Education & Digital Literacy, Corporate & Enterprise, Government & Public Sector, Environmental Sustainability Programs, and Social Impact & Non-Profit
- `power_consumption`: Power consumption of the device in watts (W) when in use (e.g., 50W, 75W)
- `energy_savings_yr`: Estimated energy savings per device per year when repurposed compared to a new device, measured in kilowatt-hours (kWh)
- `co2_saved_kg_yr`: Estimated CO2 emissions saved per device per year from manufacturing a new device, measured in kilograms (kg).
- `recycling_rate`: The percentage of the device that is recyclable (e.g., 80%, 90%).
- `region`: The geographical region where the device was repurposed, values are "North America", "Europe", and "Asia"

— **Task 1:** Organizing and Understanding the Data

We'll start by **joining** the device data with the impact data, allowing for a comprehensive analysis of device types, model years, repurpose regions, and energy savings in one dataset.

A. Simply write a query that returns all of the columns from both tables, joining the two on the `device_id` column. Be sure to choose the appropriate join so that all relevant

data is included in your result. **Note:** your query will have more than 150,000 rows (the max display for SQLPad!)

(paste your query below 👇)

```
SELECT
  *
FROM
  intel.device_data AS a
  INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
```

B. To your joined dataset, add a new column called `device_age` calculated by subtracting the model_year from 2024. Paste your query below and double check that the values in your new column make sense. For example, a 2019 device should be 5 years old.

(paste your query below 👇)

```
SELECT
  a.device_id,
  a.device_type,
  a.model_year,
  2024 - a.model_year AS device_age,
  b.*
FROM
  intel.device_data AS a
  INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
```

C. Order your joined data by model_year (oldest to newest). Do you notice more older (5+ years) or newer (under 5 years) devices being repurposed? What might that indicate?

(write your **answer** below 👇)

By using OFFSET by 150000 with the addition of ORDER BY both DESC and ASC, it is clear that there are more newer devices being repurposed than older devices. When you order the query from oldest to newest, the first 150,000 outputs listed show that the devices' age are between 5–8 years old,. But when you order the query from newest to oldest, the first 150,000 outputs listed shows that the devices' ages are only between 1–2 years old, meaning there's a possibility of more devices between 2–4 years old that have been repurposed but can't be seen in the SQL interface without OFFSET. By approximation with the OFFSET function, about 450,000 newer devices (device age between 1–4) have been repurposed compared to the approximate 200,000 older devices (device age between 5–8) .

The indication that newer devices are being repurposed more suggests that newer devices may contain systems and programs that are currently being used in the latest devices that are on the market. Choosing to repurpose newer devices over older devices could save time and energy as they may not need as many changes to be considered in the standard state for usage. Newer devices may also be easier to access as it was mentioned that many households have devices that are brand new but go untouched for a long period of time. Thus, giving Intel the opportunity to repurpose these devices for another customer.

D. Bucketing the `device_age` will allow us to analyze trends and patterns in energy savings and CO_2 reductions more effectively than using individual ages. Use a `CASE WHEN` clause to add one more column, called `device_age_bucket`, to your data, that is based on the device_age:

- `WHEN` the `device_age` is less than or equal to 3, `device_age_bucket` should be "newer"
- `WHEN` the `device_age` is greater than 3 but less than or equal to 6, `device_age_bucket` should be "mid-age"
- `WHEN` the `device_age` is greater than 6, `device_age_bucket` should be "older"

HINT: Instead of using e.g. `device_age <= 3`, you need to reference the calculation directly: `2024 - d.model_year <= 3`.

Double check that the values in your new column make sense! For example, a 2019 device should be characterized as "mid-age".

(paste your query below 👇)

```
SELECT
  a.device_id,
  a.device_type,
  a.model_year,
  2024 - a.model_year AS device_age,
  CASE
    WHEN 2024 - a.model_year <= 3 THEN 'newer'
    WHEN 2024 - a.model_year > 3
    AND 2024 - a.model_year <= 6 THEN 'mid-age'
    ELSE 'older'
  END AS device_age_bucket,
  b.*
FROM
  intel.device_data AS a
  INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
```

— **Task 2:** Key Insights

Now it's time to analyze the overall impact of Intel's repurposing program. You will use your final query from **Task 1** together with the `WITH` keyword for the remainder of this Project as you aggregate and analyze the data you've organized and prepped. For a refresher, rewatch "🍿The `WITH` Keyword" in SkillBuilder 6.

A. What is the total number of devices Intel repurposed in 2024?

HINT: The dataset **is** representing all devices repurposed in 2024! You just need to COUNT all the rows in your joined data from Task 1!

(write your **answer** below 👇)

> According to the joined data table from Task 1, the total number of devices Intel repurposed in 2024 was 601,740 devices.

B. Write a query that returns the total number of devices repurposed, the average age of repurposed devices in 2024, the average estimated energy savings (kWh) from repurposed devices per year, and the total CO_2 emissions saved (in tons) from repurposed devices.

Note: CO_2 emissions are typically measured in tons. Since `CO₂_saved_kg_yr` is measured in kg, divide the `SUM(CO₂_saved_kg_yr)` by 1000 to report the total CO_2 emissions saved in tons.

(paste your query below 👇)

```
WITH intel_repurpose_table AS (
  SELECT
    a.device_id,
    a.device_type,
    a.model_year,
    2024 - a.model_year AS device_age,
    CASE
      WHEN 2024 - a.model_year <= 3 THEN 'newer'
      WHEN 2024 - a.model_year > 3
      AND 2024 - a.model_year <= 6 THEN 'mid-age'
      ELSE 'older'
    END AS device_age_bucket,
    b.*
  FROM
    intel.device_data AS a
    INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
)
SELECT
  COUNT(*) AS total_devices_repurposed,
```

```
  ROUND(AVG(device_age), 2) AS
avg_repurposed_device_age,
  ROUND(AVG(energy_savings_yr), 2) AS
avg_estimated_energy_savings,
  ROUND(SUM(co2_saved_kg_yr) / 100, 2) AS
total_CO2_emissions
FROM
  intel_repurpose_table
```

C. Now that you have calculated the average estimated energy savings (kWh) and CO_2 emissions saved (tons), use ChatGPT to help put these numbers into perspective.

Try this prompt: I found that each repurposed device saves approximately of XXX kWh of energy per year and Intel's repurposing program saved XXX tons of CO_2 emissions in one year. Help me understand the significance of these numbers. How would this compare to the energy consumption of a small city or the amount of CO_2 produced by cars? What is the environmental impact of these savings?

What comparisons did you find most impactful in terms of scale? Summarize how much energy and CO_2 emissions were saved and how it compares to something familiar, like powering households or reducing car emissions.

(write your **answer** below 👇)

The most impactful comparisons I found were that the amount of energy emissions saved was equivalent to 15,000 homes for a year and the amount of CO2 emissions conserved was equivalent to removed 14,700 gas powered cars from the road per year. Based on ChatGPT's analysis, the average gas car emits around 4.6 metric tons of CO2 per year. When divided by the total amount of CO2 emissions Intel saved by repurposing devices in 2024 (67684.21 tons), the amount was comparable to removing roughly 14,715 cars off the road. Additionally, ChatGPT stated that the average U.S. household uses about 10,600

kWh of energy per year. When divided by the total amount of energy Intel saved (161.9 million kWh, which was calculated by dividing the total CO2 emissions saved by the amount of CO2 emissions per kWh), the amount of energy saved was equivalent to the energy of 15,270 homes per year.

— **Task 3:** Identifying Trends & Maximizing Sustainability

By grouping our data in different ways, we can uncover patterns in energy savings and CO_2 reductions. These insights will help us determine which categories of devices contribute the most to sustainability efforts and where Intel should focus its repurposing strategy for maximum impact.

A. Write a query that returns the total number of devices, the average energy savings, and the average CO_2 emissions saved (in tons), grouped by device_type.

Note (again): You'll need to divide `AVG(CO₂_saved_kg_yr)` by 1000 to report the average CO_2 emissions saved in tons.

(paste your query below 👇)

```
WITH intel_repurpose_table AS (
  SELECT
    a.device_id,
    a.device_type,
    a.model_year,
    2024 - a.model_year AS device_age,
    CASE
      WHEN 2024 - a.model_year <= 3 THEN 'newer'
      WHEN 2024 - a.model_year > 3
      AND 2024 - a.model_year <= 6 THEN 'mid-age'
      ELSE 'older'
    END AS device_age_bucket,
    b.*
  FROM
    intel.device_data AS a
```

```
    INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
)
SELECT
  device_type,
  COUNT(*) AS total_devices_repurposed,
  ROUND(AVG(energy_savings_yr), 2) AS
avg_estimated_energy_savings,
  ROUND(SUM(co2_saved_kg_yr) / 100, 2) AS
total_CO2_emissions
FROM
  intel_repurpose_table
GROUP BY
  device_type
```

B. Based on the results, which device type contributes the most to energy savings and CO_2 reduction? Why might that be the case?

Hint: Don't forget you can use ChatGPT as your Teammate to help think through your response!

(write your **answer** below 👇)

From a general standpoint, it appears that laptops contribute the most to energy savings by slightly more as laptops were found to have an average energy savings of 25.80 kWh per device while desktops were found to save an average of 25.62 kWh per device. The average amount of CO2 emissions saved was equivalent for both devices as both laptops and desktops saved an average of 0.11 tons of CO2 emissions per year. According to ChatGPT's analysis, a potential reason as to why laptops have higher energy savings compared to desktops could be because laptops are generally built to be more energy efficient with their power-saving features like their CPU, batteries, or screen. Additionally, desktops typically use more power than laptops when active as their features require more energy to run efficiently.

C. Write a query that returns the total number of devices, the average energy savings, and the average CO_2 emissions saved (in tons), now grouped by `device_age_bucket`.

(paste your query below 👇)

```
WITH intel_repurpose_table AS (
  SELECT
    a.device_id,
    a.device_type,
    a.model_year,
    2024 - a.model_year AS device_age,
    CASE
      WHEN 2024 - a.model_year <= 3 THEN 'newer'
      WHEN 2024 - a.model_year > 3
      AND 2024 - a.model_year <= 6 THEN 'mid-age'
      ELSE 'older'
    END AS device_age_bucket,
    b.*
  FROM
    intel.device_data AS a
    INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
)
SELECT
  device_age_bucket,
  COUNT(*) AS total_devices_repurposed,
  ROUND(AVG(energy_savings_yr), 2) AS
avg_estimated_energy_savings,
  ROUND(AVG(co2_saved_kg_yr) / 100, 2) AS
avg_CO2_emissions
FROM
  intel_repurpose_table
GROUP BY
  device_age_bucket
ORDER BY
  total_devices_repurposed ASC
```

D. Based on the result of your query, what do you notice about the relationship between device age and the number of devices repurposed versus the average energy saved?

(write your **answer** below 👇)

The relationship between device age and the number of devices repurposed seems to be that the younger a device is, the more it was repurposed. The device age with the most amount of devices repurposed was "newer" which means the device age was 3 years old and under. The total number of "newer" devices repurposed was 317,191. The device age with the least amount of devices repurposed was "older", which means the device age was 6 years old and up. THe total number of "older" devices repurposed was 20,239.
On the other hand, the relationship between device age and average energy saved was that the older the device was, the more the average amount of energy was saved. For the "older" devices, the average amount of energy saved in kWh was 48.02 kWh, while the average amount of energy saved in kWh for "newer" devices was 19.07 kWh. This relationship could be due to the fact that older devices might not have the same standard energy saving features and when repurposed, can match the energy saving standards of a newer, current device.

E. Finally, write a query that returns the total number of devices, the average energy savings, and the average CO_2 emissions saved (in tons), now grouped by region.

(paste your query below 👇)

```
WITH intel_repurpose_table AS (
  SELECT
    a.device_id,
    a.device_type,
    a.model_year,
    2024 - a.model_year AS device_age,
    CASE
```

```
      WHEN 2024 - a.model_year <= 3 THEN 'newer'
      WHEN 2024 - a.model_year > 3
      AND 2024 - a.model_year <= 6 THEN 'mid-age'
      ELSE 'older'
    END AS device_age_bucket,
    b.*
  FROM
    intel.device_data AS a
    INNER JOIN intel.impact_data AS b ON a.device_id =
b.device_id
)
SELECT
  region,
  COUNT(*) AS total_devices_repurposed,
  ROUND(AVG(energy_savings_yr), 2) AS
avg_estimated_energy_savings,
  ROUND(AVG(co2_saved_kg_yr) / 100, 2) AS
avg_CO2_emissions
FROM
  intel_repurpose_table
GROUP BY
  region
ORDER BY
  avg_CO2_emissions ASC
```

F. How does the carbon intensity of electricity in each region impact the total CO_2 savings from repurposed devices? Are there regions where repurposing leads to significantly higher environmental benefits? Why might that be?

(write your **answer** below 👇)

Based on the output from SQL, Asia seems to be the region with the highest amount of average CO2 emissions saved at 0.15 tons per device per year. North America follows second at 0.10 tons of CO2 per device per year while Europe saves an average of 0.06 tons. While the amount doesn't seem very much in comparison to Asia's, the average

saved CO2 emissions in both North America and Europe still play a good environmental impact as it's still reducing CO2 emissions overall. But when we put it into the perspective of the region's carbon intensity of electricity, according to ChatGPT's analysis, Asia's carbon intensity of electricity is the highest at approximately 0.7–0.9 kg CO2 per kWh while Europe's is the lowest at approximately 0.2–0.3 kg CO2 per kWh. This is due to the fact that Asia mostly relies on non-renewable energy sources like coal while Europe relies mainly on cleaner sources of energy like renewable energy sources and gas. Thus, when put into this perspective, repurposing devices in Asia leads to significantly higher environmental benefits because it breaks through the waste of the nonrenewable energy sources that are typically used to generate electricity for the region.

— **Task 4:** Data-Driven Recommendations

Using the findings from this analysis, we need to summarize key takeaways and develop actionable recommendations for Intel. Remember: the goal is to refine Intel's repurposing strategy to maximize energy savings and CO_2 reductions while ensuring the most effective use of resources.

A. Based on your analysis of the repurposed devices (including energy savings, CO_2 emissions, and device age), write **four** key takeaways in succinct sentences/bullets that summarize the most important patterns and insights from the data. These should be specific, concise, and focused on the implications of repurposing newer versus older devices.

(write your **answer** below 👇)

-The most amount of devices repurposed were newer devices around the age of 1–3 years old , encompassing a percentage of (317,191/601,740)*100% = 52.72%.

-Repurposed newer devices appear to contribute the least to the average energy savings per year (19.07 kWh per device per year) and the least to the average CO2 emissions per year (0.08 tons per year)

-The least amount of devices repurposed were older devices past the age of 6 years old, encompassing a percentage of (20,239/601,740) * 100% = 3.36%.

-Repurposed older devices appear to contribute the most to the average energy savings per year (48.02 kWh per device per year) and the most to the average CO2 emissions per year (0.21 tons per year).

B. Based on your four key takeaways and ChatGPT as your teammate, write a recommendation for Intel on how to improve the repurposing program. Your recommendation should include a clear action or strategy for Intel based on the data and a data-driven justification for why this approach would maximize energy savings and CO_2 reductions.

(write your **answer** below 👇)

A recommendation to improve Intel's repurposing program to increase energy savings and CO2 reductions is to increase awareness about the environmental value of older devices and create more opportunities for personal and institutional avenues to donate older devices for repurpose. Based on the data Intel had provided, analysis has shown that repurposing of older devices has a higher significant impact on average energy savings and CO2 emissions than repurposing newer devices. In fact, the metrics are nearly tripled as repurposed older devices save on average 48.02 kWh of energy and 0.21 tons of CO2 emissions while repurposed newer devices save on average 19.07 kWh of energy and 0.08 tons of CO2 emissions.
Thus, an environmental educational campaign could raise awareness of how repurposing older devices can save triple the amount of energy as newer devices can to incentivize people to donate their older devices for repurpose rather than disposing them as e-waste. This campaign could be pushed towards educational or corporate institutions as a method to connect Intel with areas of high e-waste that could be repurposed into devices that can still be in standard good use.

C. Briefly reflect on how ChatGPT's suggestions influenced your recommendation. Did it help you see something you hadn't considered? What parts of your recommendation were improved based on its response?

(write your **answer** below 👇)

At first, I was only considering a recommendation that took immediate action such as reaching out to companies about older devices rather than new devices or implementing programs that preferred older devices rather than newer devices. But ChatGPT's analysis made me consider long-term recommendations that left impacts on both the public and corporate/high-education spaces by allowing people to understand the impact of repurposing older devices more often than newer devices. Thus, I added to my recommendation the prospect of an environmental education campaign to raise awareness about the higher impact of older devices vs newer devices with the addition of opening connections to companies and schools for donation of older devices to increase repurposement of such devices. This ultimately reduces CO2 emissions and increases energy savings significantly more because it doesn't halt the current efforts of Intel to continue device repurpose but allows them to focus on the specific benefit of seeking older devices for repurpose to generate more devices that are sustainable and meet standard computing systems in the current technology.

— **LevelUp:** Optimizing Repurposing Strategy for Maximum Impact

Now that you've gained insights into the energy savings and CO_2 reductions across different device types and regions, let's use this data to optimize Intel's repurposing strategy for maximum environmental benefit.

A. Add to your final query of Task 3 that returns the total number of devices, the average energy savings, and the average CO_2 emissions saved (in tons), grouped by region, **the percentage** of the total energy savings and CO_2 reductions contributed by each device type within each region.

HINT: To calculate the percentage of the total energy savings, use this formula:
`Total energy savings for the device type / Total energy savings for the region) * 100`
You'll use a similar one for the percentage of the total CO_2 reductions.

Try this prompt: What's the best way to calculate the percentage of CO_2 reductions contributed by each device type in each region?

(paste your query below 👇)

```
Paste your query here.
```

B. Based on the results of your query, analyze the data to answer:

- Which device types in which regions contribute the most energy savings and CO_2 reductions relative to their numbers?
- How can this analysis help Intel prioritize specific device types in certain regions to maximize environmental benefits?

(write your **answer** below 👇)

Write your answer here.

C. In addition to focusing on sustainability, imagine Intel needs to optimize for cost-effectiveness in their repurposing program. How might you adjust your query to incorporate cost data (e.g., cost per repurposed device)? What strategies could Intel use to balance sustainability goals with cost constraints?

(write your **answer** below 👇)

Write your answer here.