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 d
INNER JOIN intel.impact_data i
 ON d.device_id = i.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
  d.*,
  i.*,
  (2024 - d.model_year) AS device_age
FROM
  intel.device_data d
INNER JOIN intel.impact_data i
  ON d.device_id = i.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 👇)

After ordering by model year, I noticed significantly more newer devices (under 5 years) being repurposed, with 437,017 compared to just 164,723 older ones. This makes sense because companies tend to upgrade their tech every few years and pass down still working devices instead of scrapping them.

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
  d.*,
  i.*,
  (2024 - d.model_year) AS device_age,
  CASE
    WHEN (2024 - d.model_year) <= 3 THEN 'newer'
    WHEN (2024 - d.model_year) > 3 AND (2024 -
d.model_year) <= 6 THEN 'mid-age'
    WHEN (2024 - d.model_year) > 6 THEN 'older'
```

```
  END AS device_age_bucket
FROM
  intel.device_data d
INNER JOIN intel.impact_data i
  ON d.device_id = i.device_id
ORDER BY d.model_year ASC
```

— **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 👇)

601,740

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 prepped_data AS (
```

```
    SELECT
        d.*,
        i.*,
        (2024 - d.model_year) AS device_age,
        CASE
            WHEN (2024 - d.model_year) <= 3 THEN
'newer'
            WHEN (2024 - d.model_year) > 3 AND (2024 -
d.model_year) <= 6 THEN 'mid-age'
            WHEN (2024 - d.model_year) > 6 THEN 'older'
        END AS device_age_bucket
    FROM intel.device_data d
    INNER JOIN intel.impact_data i
        ON d.device_id = i.device_id
)
SELECT
  COUNT(*) AS total_repurposed_devices,
  AVG(device_age) AS avg_device_age,
  AVG(energy_savings_yr) AS avg_energy_savings_kwh,
  SUM(co2_saved_kg_yr) / 1000 AS total_co2_saved_tons
FROM prepped_data
```

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 CO2 savings comparison was the most impactful for me. Saving 6,678 tons of CO2 in a single year is about the same as taking about 1,500 cars off the road, which is pretty crazy to think about. On the energy side, each device saves about 25.7 kWh per year, and across all 601,740 devices that adds up to about 15.5 million kWh, which is enough to power around 1,400 homes for an entire 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 prepped_data AS (
  SELECT
    d.*,
    i.*,
    (2024 - d.model_year) AS device_age,
    CASE
      WHEN (2024 - d.model_year) <= 3 THEN 'newer'
      WHEN (2024 - d.model_year) > 3 AND (2024 -
d.model_year) <= 6 THEN 'mid-age'
      WHEN (2024 - d.model_year) > 6 THEN 'older'
    END AS device_age_bucket
  FROM intel.device_data d
  INNER JOIN intel.impact_data i
```

```
    ON d.device_id = i.device_id
)
SELECT
  device_type,
  COUNT(*) AS total_devices,
  AVG(energy_savings_yr) AS avg_energy_savings_kwh,
  AVG(co2_saved_kg_yr) / 1000 AS avg_co2_saved_tons
FROM prepped_data
GROUP BY 1
```

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 👇)

Laptops contribute the most to the overall savings. The average savings per device are pretty much the same for both but there are more than twice as many laptops being repurposed (408,064 compared to 193,676 desktops), which makes them more impactful. This is probably because laptops are more common in corporate upgrade cycles, so the higher volume is really what drives the difference.

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 prepped_data AS (
  SELECT
```

```
    d.*,
    i.*,
    (2024 - d.model_year) AS device_age,
    CASE
      WHEN (2024 - d.model_year) <= 3 THEN 'newer'
      WHEN (2024 - d.model_year) > 3 AND (2024 -
d.model_year) <= 6 THEN 'mid-age'
      WHEN (2024 - d.model_year) > 6 THEN 'older'
    END AS device_age_bucket
  FROM intel.device_data d
  INNER JOIN intel.impact_data i
    ON d.device_id = i.device_id
)
SELECT
  device_age_bucket,
  COUNT(*) AS total_devices,
  AVG(energy_savings_yr) AS avg_energy_savings_kwh,
  AVG(co2_saved_kg_yr) / 1000 AS avg_co2_saved_tons
FROM prepped_data
GROUP BY 1
```

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 older the devices get, the higher the average energy savings are per device. Older devices save about 48kWh on average compared to just 19 kWh for newer ones, which is more than double. However, way more newer devices are being repurposed (317,191) than older devices (20,239), so the program is prioritizing volume over per unit impact.

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 prepped_data AS (
  SELECT
    d.*,
    i.*,
    (2024 - d.model_year) AS device_age,
    CASE
      WHEN (2024 - d.model_year) <= 3 THEN 'newer'
      WHEN (2024 - d.model_year) > 3 AND (2024 -
d.model_year) <= 6 THEN 'mid-age'
      WHEN (2024 - d.model_year) > 6 THEN 'older'
    END AS device_age_bucket
  FROM intel.device_data d
  INNER JOIN intel.impact_data i
    ON d.device_id = i.device_id
)
SELECT
  region,
  COUNT(*) AS total_devices,
  AVG(energy_savings_yr) AS avg_energy_savings_kwh,
  AVG(co2_saved_kg_yr) / 1000 AS avg_co2_saved_tons
FROM prepped_data
GROUP BY 1
```

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 👇)

Even though the energy saved per device is almost the same across all three regions (around 25.7kWh), Asia saves significantly more CO2 per device at 0.0155 tons compared to just 0.0064 tons in Europe. This is probably because Asia's power grid relies heavily on fossil fuels,

meaning every kWh saved prevents a lot more emissions than it would in Europe where the grid is cleaner.

— **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 👇)

1. Even though Intel mostly repurpose newer devices (317,191), older devices save more than twice as much energy per unit (48 kWh vs 19 kWh), meaning the program is currently prioritizing volume over per unit impact.
2. Laptops create the biggest total savings because more than twice as many of them are repurposed compared to desktops (408,064 vs 193,676), even though the average savings per device are nearly identical.
3. Saving electricity in Asia prevents more than double the CO2 per device compared to Europe (0.0155 tons vs. 0.0064 tons) because their power grid relies more heavily on fossil fuels.
4. While each device saves about 25.7 kWh per year on average, doing this across 601,740 devices adds up to roughly 15.5 million kWh total and 6,768 tons of CO2 saved in a single 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 👇)

> Intel should try and focus on sending their oldest laptops to regions like Asia. Older devices save about 48 kWh per unit compared to 19 kWh for newer ones, and laptops already make up the majority of repurposed devices at 408,064 units. On top of that, Asia saves 0.0155 tons of CO2 per device compared to just 0.0064 tons in Europe, meaning the combination of older devices and a fossil fuel heavy grid creates the biggest possible environmental return. The one challenge is that older devices are harder to come by since 20,239 were repurposed in 2024, so Intel may also need to expand their collection partnerships to source more of them.

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 👇)

> ChatGPT was very helpful in helping form my recommendation. When I asked about Asia's power grid, it explained that the region relies heavily on coal specifically, which was the missing piece that connected my two findings together into one clear recommendation. It also suggested factoring in transportation costs since shipping devices to Asia could offset some of the environmental gains, but since we didn't have that data available I decided to stick with what the data I had was showing me.

— **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 👇)

```
WITH prepped_data AS (
  SELECT
    d.*,
    i.*,
    (2024 - d.model_year) AS device_age,
    CASE
      WHEN (2024 - d.model_year) <= 3 THEN 'newer'
      WHEN (2024 - d.model_year) > 3 AND (2024 -
d.model_year) <= 6 THEN 'mid-age'
      WHEN (2024 - d.model_year) > 6 THEN 'older'
    END AS device_age_bucket
  FROM intel.device_data d
  INNER JOIN intel.impact_data i
    ON d.device_id = i.device_id
),
region_totals AS (
  SELECT
```

```
    region,
    SUM(energy_savings_yr) AS total_region_energy,
    SUM(co2_saved_kg_yr) AS total_region_co2
  FROM prepped_data
  GROUP BY region
),
device_totals AS (
  SELECT
    region,
    device_type,
    COUNT(*) AS total_devices,
    AVG(energy_savings_yr) AS avg_energy_savings_kwh,
    AVG(co2_saved_kg_yr) / 1000 AS avg_co2_saved_tons,
    SUM(energy_savings_yr) AS total_device_energy,
    SUM(co2_saved_kg_yr) AS total_device_co2
  FROM prepped_data
  GROUP BY region, device_type
)
SELECT
  dt.region,
  dt.device_type,
  dt.total_devices,
  dt.avg_energy_savings_kwh,
  dt.avg_co2_saved_tons,
  (dt.total_device_energy / rt.total_region_energy) *
100 AS pct_energy_savings,
  (dt.total_device_co2 / rt.total_region_co2) * 100 AS
pct_co2_reductions
FROM device_totals dt
INNER JOIN region_totals rt
  ON dt.region = rt.region
ORDER BY dt.region, dt.device_type
```

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 👇)

> Laptops dominate the savings in every region, accounting for about 68% of both energy savings and CO2 reductions across Asia, Europe, and North America. However, Asia stands out because laptops there save 0.0155 tons of CO2 per device compared to just 0.0065 tons in Europe, making it the highest impact combination by far. Intel should prioritize routing laptops specifically to Asia since putting the highest volume device in the region with the best per device CO2 return will maximize the overall environmental impact of the program.

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 👇)

> To include costs, I would add a cost_per_device column to the query and calculate a cost per ton of CO2 saved by dividing the total cost by the total CO2 saved for each region and device type combination. This would let Intel see which combinations give the best environmental return for their money. Based on what we already know, laptops in Asia would likely still come out on top, but if shipping costs to Asia are significantly higher, Intel could use this metric to decide whether the extra CO2 savings actually justify the added expense.

— Evaluation Rubric

Unlike your Milestones that were evaluated largely based on your effort, the evaluation of your Portfolio Project will follow traditional evaluation methods, with tasks assessed for correctness and assigned point values accordingly.

Partial credit will be given where parts of this task are correct, even if other parts are incorrect or incomplete.

Task title	Max points
Task 1: Organizing and Understanding the Data	40
Task 2: Key Insights	25
Task 3: Identifying Trends & Maximizing Sustainability	60
Task 4: Data-Driven Recommendations	75
TOTAL POINTS:	200
LevelUp	
Optimizing Repurposing Strategy for Maximum Impact	20