



Photo by Maxim Tolchinskiy on Unsplash

When factoring heat generation required for the manufacturing and transportation of products, *Greenhouse gas emissions attributable to products, from food to sneakers to appliances, make up more than 75% of global emissions.*

( Source: The Carbon Catalogue <https://www.nature.com/articles/s41597-022-01178-9> )

Our data, which is publicly available on nature.com, contains product carbon footprints (PCFs) for various companies. PCFs are the greenhouse gas emissions attributable to a given product, measured in CO<sub>2</sub> (carbon dioxide equivalent).

This data is stored in a PostgreSQL database containing one table, `product_emissions`, which looks at PCFs by product as well as the stage of production that these emissions occurred. Here's a snapshot of what `product_emissions` contains in each column:

`product_emissions`

field	data type
<code>id</code>	<code>VARCHAR</code>
<code>year</code>	<code>INT</code>
<code>product_name</code>	<code>VARCHAR</code>
<code>company</code>	<code>VARCHAR</code>
<code>country</code>	<code>VARCHAR</code>
<code>industry_group</code>	<code>VARCHAR</code>
<code>weight_kg</code>	<code>NUMERIC</code>
<code>carbon_footprint_pcf</code>	<code>NUMERIC</code>

field	data type
upstream_percent_total_pcf	VARCHAR
operations_percent_total_pcf	VARCHAR
downstream_percent_total_pcf	VARCHAR

You'll use this data to examine the carbon footprint of each industry in the dataset!

 Projects Data    DataFrame as carbon\_emissions\_by\_industry

-- Using the product\_emissions table, find the number of unique companies and their total carbon footprint PCF for each industry group, filtering for the most recent year in the database. The query should return three columns: industry\_group, num\_companies, and total\_industry\_footprint, with the last column being rounded to one decimal place. The results should be sorted by total\_industry\_footprint from highest to lowest values.

```
SELECT industry_group,
       COUNT(DISTINCT company) AS num_companies,
       ROUND(SUM(carbon_footprint_pcf), 1) AS total_industry_footprint
FROM product_emissions
WHERE year IN (SELECT MAX(year) FROM product_emissions)
GROUP BY industry_group
ORDER BY total_industry_footprint DESC;
```

index	...	↑↓	industry_group	...	↑↓
		0	Materials		
		1	Capital Goods		
		2	Technology Hardware & Equipment		
		3	Food, Beverage & Tobacco		
		4	Commercial & Professional Services		
		5	Software & Services		

Rows: 6

 Expand Table