

## Tesis Final de Maestría

# ***“La digitalización como catalizador de la Sustentabilidad Ambiental empresarial para abordar estrategias de mitigación y reducción de emisiones de dióxido de carbono”***

2022

*Alumno:* Nicolás Alejandro Canusso

*Tutora:* Paulina Becerra

---

## Temario

0) <a href="#">Resumen de la Tesis</a>	
1) <a href="#">Introducción</a>	
1.1) <a href="#">Resumen</a>	4
1.2) <a href="#">Objetivo</a>	5
1.3) <a href="#">Datos</a>	6
1.4) <a href="#">Metodología</a>	6
<hr/>	
2) <a href="#">Marco Teórico</a>	8
2.1) <a href="#">La Sustentabilidad y la Huella de Carbono</a>	
2.2) <a href="#">La Sustentabilidad en las Empresas</a>	
2.3) <a href="#">El Cambio Climático y el Escenario de 1.5°C</a>	
2.4) <a href="#">Gases de Efecto Invernadero y su Medición</a>	
2.5) <a href="#">Medición de la Huella de carbono y otros Indicadores Ambientales</a>	
3) <a href="#">Revisión de la Literatura: “La digitalización y la Sustentabilidad”</a>	14
4) <a href="#">Desafíos de Data Management en la gestión ambiental</a>	16
4.1) <a href="#">End-Game definido para cada etapa</a>	
5) <a href="#">Análisis de Casos</a>	19
5.1) <a href="#">Herramientas digitales en el proceso de Data Gathering</a>	
5.1.1) <a href="#">Cálculo de Emisiones de CO2 para Flota Logística</a>	
5.1.2) <a href="#">API para medición de consumo de Combustible y medición de Ralentí</a>	
5.1.3) <a href="#">Smart Metering - Medición de consumo de Energía en línea</a>	
5.2) <a href="#">Herramientas digitales en el proceso de Data Warehousing</a>	
5.2.1) <a href="#">API/Web Scraping para la adquisición de factores de emisión</a>	
5.2.2) <a href="#">API/Web Scraping para la adquisición de costo diario de combustible</a>	
5.2.3) <a href="#">Bases de datos integradas con APIs para KPIs de flota logística</a>	
5.2.4) <a href="#">Scripts de SQL para cálculo de KPIs Ambientales</a>	

5.3) <a href="#">Herramientas digitales en el proceso de Data Visualization</a>	
5.3.1) <a href="#">Visualización de Principales Indicadores</a>	
5.4) <a href="#">Herramientas digitales en el proceso de Predictive Analytics</a>	
5.4.1) <a href="#">Sentar las bases para predicción de consumo de Energía</a>	
6) <a href="#">Recomendaciones</a>	52
6.1) <a href="#">Para el proceso de Data Gathering</a>	
6.2) <a href="#">Para el proceso de Data Warehousing</a>	
6.3) <a href="#">Para el proceso de Data Visualization</a>	
6.4) <a href="#">Para el proceso de Predictive Analytics</a>	
7) <a href="#">Conclusiones y trabajo a futuro</a>	56
8) <a href="#">Referencias</a>	58

## **0) Resumen de la Tesis**

*En este trabajo de tesis exploramos cómo la digitalización puede actuar como catalizador de la Sustentabilidad Ambiental empresarial contribuyendo con el abordaje de estrategias de mitigación y reducción de emisiones de dióxido de carbono. Mediante el análisis de dos casos de estudio, el primero el análisis de reducción de métricas ambientales en una Operación de Flota Logística y el segundo caso el estudio de Consumos Energéticos de una compañía, analizamos cómo las herramientas de análisis de datos nos permiten contribuir con los desafíos anteriormente mencionados con respecto a los distintos pilares de la gestión de datos.*

*En primera instancia, la tesis se compone de una introducción detallando los aspectos a estudiar de ambos estudios, indicando las herramientas tecnológicas cuya aplicación se estudiará. Se detallan los aspectos referidos a los datos a utilizar, la metodología a seguir, junto con una introducción teórica sobre la importancia de las estrategias de ESG en el mundo empresarial actual y sus necesidades con posible contribución desde el área de digitalización.*

*Seguidamente en la sección de Marco Teórico, la tesis se enfoca en contar los sucesos actuales en términos de cambio climático, necesidades y oportunidades del mundo actual en términos de ESG, las principales métricas ambientales que necesita abordar el mundo empresarial hoy en día junto los desafíos actuales en relación con el desarrollo de una estrategia de sustentabilidad que pueda acompañar el crecimiento económico de una compañía.*

*Posteriormente, se realiza una revisión de la literatura acerca de los estudios realizados referentes a la implementación y oportunidades de aplicación de herramientas digitales en estrategias de sustentabilidad. A partir de este estudio se visibiliza la necesidad de contribuir con la literatura desde el aspecto empírico y práctico, buscando incorporar análisis ambientales-económicos para los procesos operativos de las empresas, junto con el abordaje de sus necesidades de reporting y compliance, entendiendo también los desafíos y oportunidades de las herramientas de datos, con casos de uso que permitan la fácil aplicación para una compañía. Este estudio busca aportar a la literatura desde ese punto de vista.*

*A partir de esta revisión de la literatura, se expone sobre los desafíos del ciclo de Data Management en una gestión ambiental corporativa para seguidamente abordar de lleno ambos casos de estudio. Dicho ciclo compuesto por el proceso de Data Gathering, Data Warehousing, Data Visualization y Predictive Analytics, y se recorren los desafíos en cada una de estas etapas para el desarrollo de una estrategia de sustentabilidad abordando los casos de estudio antes mencionados.*

*Con foco en la optimización de métricas ambientales como consumo de combustible, consumo energético y principalmente confluyendo todo en el indicador de Huella de Carbono, junto con el aspecto de métricas económicas, se aborda la aplicación de herramientas digitales para las etapas del ciclo de Data Management mencionado con el fin de entender la factibilidad técnica-económica-ambiental junto con los beneficios de diferentes herramientas digitales como APIs, Bases de Datos en SQL, Sistemas de Smart Metering, Sistemas de Telemetría, Dashboards digitales y herramientas de Web Scraping.*

*Finalmente, la tesis cierra con un apartado de conclusiones y recomendaciones prácticas para que los líderes de hoy tengan las herramientas adecuadas para entender el poder de las herramientas digitales para contribuir en sus necesidades de sustentabilidad actual.*

## 1) Introducción

### 1.1) Resumen

#### **“La digitalización como catalizador de la Sustentabilidad Ambiental empresarial para abordar estrategias de mitigación y reducción de emisiones de dióxido de carbono”**

*El mundo y la sociedad actual, especialmente las nuevas generaciones, son cada vez más sensibles a las temáticas ambientales y cada vez más demandantes con respecto a los gobiernos, empresas y productos que consumen. Las empresas, buscando dar una respuesta a esta demanda, cumplir con nuevas normativas ambientales, exigencias de inversores y mismo buscando seguir creciendo en el área por ser un valor propio de la compañía, incorporan la sustentabilidad como estandarte.*

*Todo esto conlleva a que las empresas se vean en la necesidad de entender cuáles son los impactos ambientales que generan sus operaciones para tomar decisiones estratégicas que ayuden a mitigarlos. En muchos casos además, las empresas deben reportar su impacto ambiental. Hoy en día es común leer reportes donde industrias, empresas de tecnología, entidades bancarias, y prácticamente empresas de todos los rubros, divulgan sus principales indicadores y resultados ambientales como huella de carbono, consumo de agua, consumo de energía, cantidad de vehículos eléctricos, porcentaje de energía renovable, entre otros. Según el Governance & Accountability Institute, en el año 2012 más del 50% de las empresas que componen el S&P 500 publicaron reportes de Sustentabilidad; confirmaron además que en el año 2019 esa cifra subió a un 90% del total de compañías que componen dicho índice.*

*El cálculo de estas métricas y seguimiento de estos indicadores, suele tener una complejidad relevante. Primeramente debido a que la cantidad de datos necesarios para los mismos no se suele encontrar de manera ordenada y sencilla. El cálculo de la huella de carbono de una compañía, puede considerar datos relativamente sencillos de obtener, como los consumos energéticos anuales de sus sitios, hasta algo mucho más complejo como la cantidad de combustible consumido por los taxis que utilizan diariamente sus colaboradores para trasladarse a las oficinas. En segundo lugar, podemos observar que el volumen de datos que una empresa tiene que manejar para realizar estos cálculos, pueden tomar magnitudes exponenciales.*

*A medida que se quiere afinar el cálculo de estos indicadores ambientales y que los reportes se vuelven más exigentes, la necesidad de las empresas de mantener una buena gestión de datos, que les permita robustecer su cálculo, seguimiento, toma de decisiones y facilitar su reporting, toma una mayor importancia.*

*En este trabajo exploramos, mediante el análisis de dos casos de estudio, cómo las herramientas de análisis de datos nos permiten contribuir con los desafíos anteriormente mencionados con respecto a los distintos pilares de la gestión de datos.*

*En primer lugar, estudiamos un caso de Operación de Flota Logística entendiendo sus impactos y métricas ambientales relevantes. Buscamos entender el ciclo de vida de los datos y cómo las empresas pueden utilizar diversas herramientas digitales como APIs, Bases de datos, sistemas de Internet of Things de Telematics, entre otras, para optimizar sus estrategias operativas en pos de obtener eficiencias ambientales, económicas y mejorar la calidad de datos que requiere la gestión de sustentabilidad ambiental. Entenderemos cómo la no visibilidad de ciertos datos, como las emisiones generadas por Ralentí, solo se pueden detectar y mitigar utilizando ciertas herramientas.*

*En segundo lugar, estudiaremos un caso de Consumos Energéticos de una empresa, dónde buscaremos entender cómo podemos medir sus impactos ambientales, desarrollar una estrategia acorde y encontrar a su vez eficiencias ambientales y económicas, mediante la implementación de herramientas digitales. Trabajaremos además en la visualización de indicadores principales y la gestión del almacenamiento de datos en pos de optimizar los flujos de reporting necesarios de las empresas hoy en día.*

## **1.2) Objetivo de la tesis**

El objetivo de la tesis es poder detectar e implementar las mejores herramientas tecnológicas aprendidas en la maestría sobre Data Management, en las siguientes etapas de la gestión de datos del área de Medio Ambiente: Data ingestion (APIs, Web scrappers, queries de SQL, etc), Data Warehouse (DBMS, ETL, OLAP, etc), Data Visualization (Storytelling, dashboards, effective charts), y Predictive Analytics (algoritmos de clustering como K-means), que permita:

- Robustecer la gestión de datos mediante la optimización de ingestión, almacenamiento y tratamiento de datos.
- Facilitar la adquisición de datos de múltiples fuentes y encauzarlos para su gestión.
- Facilitar el cálculo de indicadores, métricas ambientales relevantes y datos a reportar.
- Dar visibilidad y seguimiento a los principales indicadores ambientales, mediante dashboards que permitan tomar decisiones estratégicas en cuanto a la mitigación del impacto ambiental considerando distintas variables relevantes.
- Sentar las bases para permitir la predicción de escenarios futuros en cuanto a las principales métricas relevantes como huella de carbono, consumo energético, entre otros.

Esto nos permitirá decidir sobre cómo podemos optimizar la gestión de datos de la compañía y en esa línea cuáles son las medidas más efectivas, contemplando distintos criterios, para reducir la huella de carbono de la compañía.

## **1.3) Datos**

Utilizaremos principalmente dos casos de estudio que nos permitirán abordar este estudio.

Para el primer caso de estudio de la Operación de Flota Logística, utilizaremos un extracto de la Bases de Datos llamada “Heavy-Duty Vehicle Activity” del proyecto MOVES (Motor Vehicle Emission Simulator) recopilada conjuntamente entre la Agencia de Protección Ambiental de los Estados Unidos (EPA), el Laboratorio Nacional de Energías Renovables de los Estados Unidos (NREL) y la Universidad de California, compuesta por un Dataset que incluye 564 vehículos comerciales y más de 23.000 días operativos de dichos vehículos, cubriendo siete tipos de vehículos de Heavy-Duty (del tipo que conversamos anteriormente para los casos de Line-Haul o Media milla logística). El cuál contiene algunos de los siguientes datos y a la cual complementaremos con el análisis de otros datos adicionales:

- Cantidad de Vehículos (N°)
- Cantidad de combustible consumido por vehículos de la compañía (Galones)
- Distancia recorrida por vehículos a diesel (Millas)
- Tiempo de Operación (s)
- Tiempo de Ralentí (s)
- Emisiones de CO2 generadas por la huella logística (TonCO2)
- Costo de combustibles (USD)

Por otro lado, para el segundo caso de estudio, correspondiente al consumo de energía de una compañía en un año, utilizaremos la base de datos “Great energy predictor” de la mundialmente reconocida

sociedad de ASHRAE, especializados en gestión energética, donde podremos entrar en contacto y analizar algunos de los siguientes datos:

- Energía consumida de red (Kwh)
- Intensidad Energética (Mwh/m2)
- Toneladas de CO2 evitadas por consumo de energía renovable (Ton CO2)
- Toneladas de CO2 generadas por consumo de energía (Ton CO2)
- Superficie del site (m2)
- Cantidad de personal en el site (N°)

El espacio temporal, será el de un periodo anual de operación empresarial. Por lo cual, utilizaremos los millones de datos (mencionados del listado anterior, entre otros) generados en un año de operación para implementar la gestión de datos en sus tres etapas.

Los datos serán obtenidos tienen una gran multiplicidad de fuentes y formas. Algunas de estas métricas, como las relacionadas con energía, son representadas por datos más bien sencillos, pero de los cuales el desafío está en su obtención, seguimiento y visualización. En otros casos, como la medición de la huella de carbono, intervienen una multiplicidad de datos. Nosotros utilizaremos los sugeridos por el protocolo GHG de medición de huella de carbono.

## 1.4) Metodología

Como mencionamos anteriormente, segmentaremos la gestión del proyecto en tres etapas mencionadas dentro de Data Management , donde implementaremos algunas de las siguientes herramientas aprendidas en la maestría (y otras por investigar), para optimizar y robustecer la gestión de datos:

### Data Ingestion

Dentro de data ingestion, tenemos el desafío de obtener una cantidad de datos exponenciales de una multiplicidad de fuentes de distintas características.

Para ello, utilizaremos distintas herramientas como Web Scrapping, APIs de ingestión, queries de SQL, entre otras, para asegurar una adquisición de datos robusta que nos permita optimizar el tiempo de toma de datos, la calidad de los datos obtenidos y poder encauzar la big data que maneja la compañía para el cálculo del principal KPI Ambiental: Huella de Carbono.

Primeramente comenzaremos el análisis explorando la tecnología de Telematics en una flota de vehículos. Existen dispositivos de IoT (Internet of things) que tienen la capacidad de conectarse con los vehículos y por los cuales podemos recibir información y datos relevantes de su operación en tiempo real. Dichos dispositivos de "Telemática", son de nuestro interés para analizar si pueden colaborar con nuestro desafío de Data Management para la flota de vehículos y dado que el sistema está conectado al vehículo, podemos obtener en tiempo real los datos de consumo de combustible del vehículo lo cual nos permitirá calcular la huella de carbono.

Y para poder conectarnos con los datos recopilados de la misma, utilizaremos una herramienta aprendida en la maestría llamada APIs, la cual es una interfaz de comunicaciones que permite que dos aplicaciones se puedan comunicar entre sí. Para ello, como mencionamos anteriormente, utilizaremos el extracto de la Bases de Datos llamada "Heavy-Duty Vehicle Activity" del proyecto MOVES (Motor Vehicle Emission Simulator)

Para la recopilación de dichos datos se usaron dos dispositivos de Telemática que permiten obtener información fidedigna de la operación del vehículo, dado que se obtuvieron datos directos de sus

motores, mediante el conexionado tipo CAN Bus (Controller Area Network). De esta manera es posible identificar los periodos de tiempo en el cual los camiones estaban en Ralentí (vehículo detenido con el motor prendido) dado que se mide tanto la velocidad de las ruedas como la velocidad del motor (Se define Ralentí cuando la velocidad de la rueda es menor a 1 Mph y la velocidad del motor es mayor a 0 Mph).

El segundo caso de Data Ingestion que analizaremos, será la utilización de sensores de **Smart Metering** para medición de consumo eléctrico. Esta tecnología implica la utilización de sensores de energía, que miden el consumo eléctrico en tiempo real, lo almacenan en la nube y tienen la posibilidad de descargar los datos para análisis vía API u otras metodologías.

Algunos de estos sistemas incluso tienen la capacidad de medir otras variables, como consumo de agua, temperatura ambiente, entre otras, que quedarán fuera del alcance del análisis de este estudio, pero que también contribuyen con la generación de estrategias de sustentabilidad empresarial.

Analizaremos entonces el Dataset *"Energy Consumption"*, donde tenemos un registro de datos de Consumo Energético (en Kwh) de distintas áreas de una empresa (Logística, Oficinas, Fábrica, etc), donde además tenemos registro de datos interesantes como los horarios de los consumos energéticos y los datos de áreas de cada sitio, entre otros.

### **Data Warehouse**

El siguiente desafío es poder trabajar en la implementación de herramientas de Bases de Datos relacionales y algunos métodos aprendidos como OLAP, ETL, Data Lakes, etc; para poder gestionar el almacenamiento de datos de manera óptima y permitiendo que faciliten el cálculo de métricas, el reporte de métricas, la visualización de las mismas y sentar las bases para poder implementar a futuro algoritmos de analytics predictivo, que nos permitan tomar decisiones sobre los métodos más efectivos para reducir la huella de carbono. Desarrollaremos todas las bases de datos en SQL según el formato aprendido durante la maestría.

Primeramente utilizaremos SQL y formato de Bases de datos para generar tablas para el cálculo de huella de logística: una tabla con los datos obtenidos via telemática sobre la flota logística y otra tabla con los factores de emisión que se actualizará con determinada frecuencia.

En segundo lugar, trabajaremos con la utilización de APIs con los valores de precios de combustibles de NASDAQ para alimentar nuestras tablas generadas, de manera de entender las posibilidades y eficiencias posibles a generar con la alimentación de dichas bases de datos con estos inputs de datos.

Por último, utilizaremos scripts de SQL para obtener KPIs principales de flota logística que las empresas necesitan para sus reportes y su gestión ambiental operativa. Para ello desarrollaremos estos scripts y buscaremos encontrar eficiencias con el armado de tablas, de manera de entender cómo esta tecnología de bases de datos nos permite solucionar algunos de los desafíos mencionados anteriormente en la rama de almacenamiento de datos.

### **Data visualization**

La tercera etapa del sistema de gestión de datos que implementaremos, es la visualización de los mismos. En esta etapa desarrollaremos Dashboards, según lo aprendido en las materias de visualización y algoritmos, que nos permitan tomar decisiones claves de negocio en cuanto a la reducción de impacto



ambiental. Definiremos distintos pipelines de datos para abastecer los dashboards, y programaremos scripts que nos permitan incorporar y visualizar los datos de la manera más eficiente posible.

Primeramente visualizaremos los KPIs principales del caso de flota eléctrica, entendiendo los porcentajes de tiempo de Ralentí en función de la flota operativa para entender patrones de consumo e ineficiencias ambientales (y económicas) dentro de dicho proceso logístico. Por otro lado, trabajaremos en visualizaciones del caso de Smart Metering para medición de consumo energético mensual, para entender y detectar posibles causas de consumo excesivo que nos contribuyan a definir estrategias de eficiencia energética. Finalmente, buscaremos visualizar valores de emisiones de CO2 de manera anual, de modo de poder trabajar en proyecciones y reportes que contribuyan a las empresas en visualizar y reportar los impactos ambientales principales según los estándares internacionales.

### **Predictive Analytics**

Por último, sentaremos las bases para que podamos elaborar distintos modelos de machine learning y predictive analytics, de manera de poder predecir consumos de energía, consumos de agua, ahorros de huella, teniendo en consideración distintos parámetros de las operaciones. El objetivo es poder dejar sentadas las bases que nos permitan a futuro entender cómo evolucionarán las operaciones, para poder definir las mejores estrategias de mitigación de impacto ambiental.

Para ello, buscaremos entender las necesidades de predicciones de datos que contienen los procesos de migración de consumos de energía hacia fuentes renovables, y los desafíos que conllevan poder cerrar dichos contratos con generadores de manera off-site cuyo principal input es una proyección de energía a un largo plazo de tiempo.

## **Marco Teórico**

### **2.1) La Sustentabilidad y la Huella de Carbono**

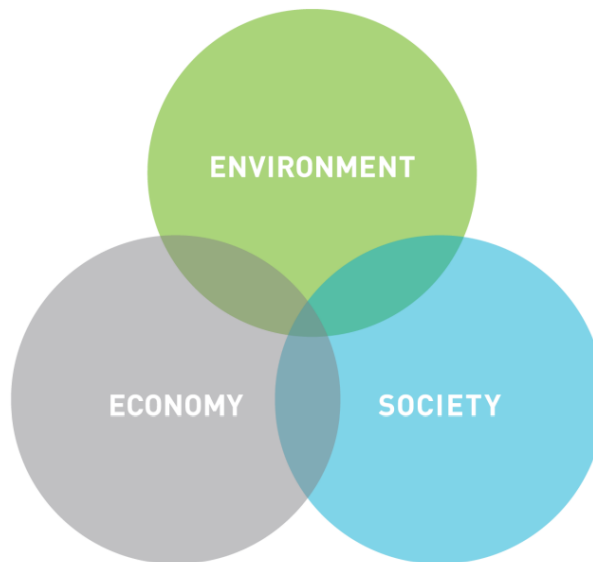
Sin ninguna duda, si buscamos definiciones de Sustentabilidad, vamos a poder encontrar distintos matices y definiciones, según el contexto en la cual nos encontremos. Sin embargo, una de las definiciones que veo con mayor agrado es la generada a partir de las Naciones Unidas llamada "Comisión Brundtland" (anteriormente conocida como *World Commission on Environment and Development*), la cual define a la Sustentabilidad como "la capacidad de cubrir nuestras necesidades actuales sin comprometer la capacidad de que futuras generaciones puedan cubrir sus necesidades a futuro".

Naturalmente por esta definición, se podría pensar que tiene un enfoque fuertemente orientado al aspecto ambiental, por el cual se contemplaría a capacidad de no comprometer recursos naturales, la ecología y el medio ambiente para futuras generaciones; sin embargo también la definición de Sustentabilidad involucra los pilares de la equidad social y el desarrollo económico.

Esta idea de Sustentabilidad es relativamente novedosa, a fines del siglo XX se adoptó la idea de "desarrollo sustentable" incorporando estos tres pilares:

**Figura 1**

*Los pilares del desarrollo sustentable*



Definiendo la sustentabilidad de cada uno de estos pilares como:

#### *Sustentabilidad Ambiental*

Cuando la integridad ecológica se conserva, todos los ecosistemas de la Tierra se mantienen en equilibrio a la vez que los humanos consumen recursos naturales en un ritmo que permite su regeneración en menos de una generación.

#### *Sustentabilidad Económica*

Las comunidades a lo largo del planeta son capaces de mantener su independencia y tener acceso a los recursos que ellos requieren, tanto financieros como otros, para cubrir sus necesidades. Los sistemas económicos permanecen intactos y todos tienen la posibilidad de interactuar con ellos como medios seguros de sustento económico.

#### *Sustentabilidad Social*

Los derechos humanos universales y el poder cubrir sus necesidades básicas alcanzan a todas las personas, que además tienen acceso a suficientes recursos de manera de poder mantener a sus familias y comunidades de manera saludable y segura. Las comunidades fructíferas tienen líderes que aseguran los derechos humanos, de trabajo, culturales y que sean respetados por todos.

## **2.2) La Sustentabilidad en las Empresas**

El concepto de Sustentabilidad se ha extendido hacia el mundo empresarial hace ya tiempo, principalmente debido a las crecientes normativas, exigencias ambientales y demandas por parte de los gobiernos, sociedad e inversores, por lo que las empresas han adoptado con mayor énfasis la importancia de contar con una correcta gestión sustentable.

Es por eso, que en el mundo empresarial nace el concepto de ESG (*Environmental, Social and Governance*), que es un marco para medir y entender cuán sustentable es una cierta organización.

Este marco extiende el concepto de Sustentabilidad, incluyendo la parte de *Governance* que promueve la transparencia y responsabilidad de dicha organización.

Hoy en día, es común que las empresas encuentren un fuerte interés en temáticas de ESG por parte de sus clientes, gobiernos, inversionistas, ONGs, entre otros. Por lo cuál sus esfuerzos se encuentran cada vez más en transparentar sus impactos y contribuciones en estas tres temáticas.

### 2.3) El Cambio climático y el escenario de 1.5°C

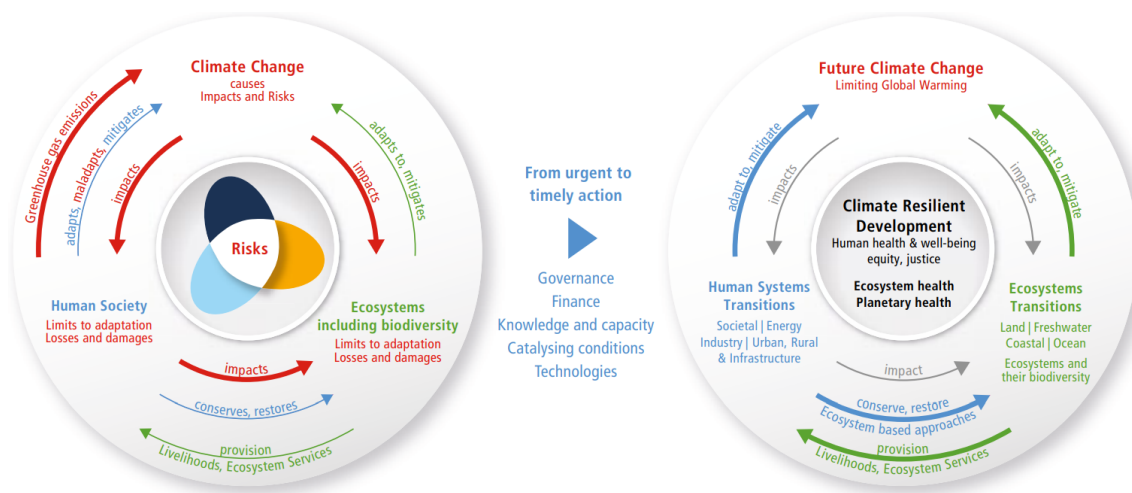
Dentro del mundo ESG de las compañías, nosotros haremos foco en el pilar de *Environment*.

Las Naciones Unidas tienen un cuerpo dedicado a asesorar a la ciencia mundial sobre cambio climático, llamado IPCC (Panel Intergubernamental de Cambio Climático). Dicho ente, se encarga de emitir un reporte anual con las bases científicas sobre el cambio climático, en el cual se describen los potenciales riesgos, impactos y posibilidades de mitigación y adaptación.

En su informe de 2022, el IPCC reconoce la interdependencia entre ecosistemas, clima, biodiversidad y sociedades humanas, definiendo además los riesgos de cambios potencialmente irreversibles en dichos sistemas de mantener el status quo, y propone las acciones necesarias para transicionar hacia un estado resiliente con respecto al cambio climático, en el cual como observamos en el gráfico, la Tecnología aparece como uno de los principales catalizadores para que lo anterior ocurra:

**Figura 2**

*Riesgos e impactos posibles por el cambio climático*



Nota: <sup>24</sup>Climate Change 2022, Impacts, Adaptation and Vulnerability (IPCC, 2022)

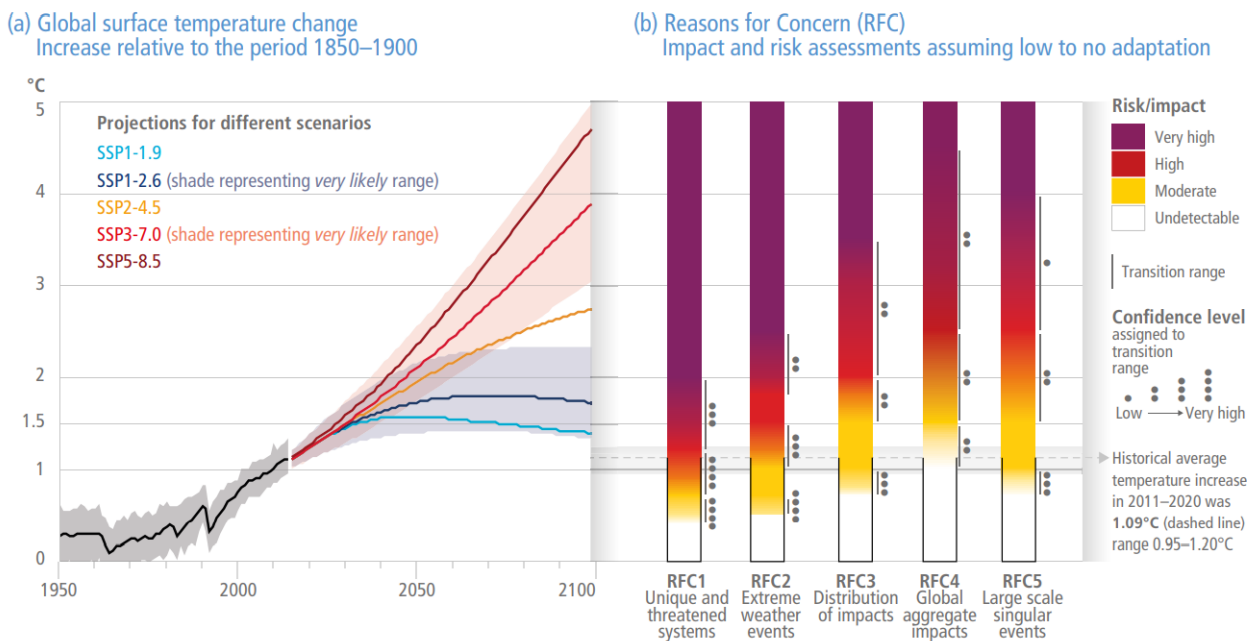
Dentro de dicho reporte, se detallan los posibles riesgos e impactos del cambio climático abarcando desde la escasez de agua, escasez de la producción de alimentos, enfermedades, hasta sus riesgos

económicos debido a los mismos como daños por inundaciones, daños de infraestructura y daños a sectores económicos claves, entre otros.

El IPCC relaciona fuertemente estos posibles riesgos, debido al aumento de de los niveles de calentamiento global (medidos por las Temperatura global de la superficie terrestre), en el cual plantea cinco posibles escenarios climáticos junto con sus proyecciones de temperatura, y las consecuencias climáticas que tendrían:

**Figura 3**

Análisis de riesgo e impacto para los escenarios proyectados por IPCC



Fuente: <sup>24</sup>Climate Change 2022, Impacts, Adaptation and Vulnerability (IPCC, 2022)

De acuerdo al estudio de IPCC, muchas de estas catastróficas consecuencias mencionadas anteriormente, podrían tornarse irreversibles y sus consecuencias podrían ser aún peores de no mantener la temperatura media de la superficie terrestre por debajo de 1.5°C, como se conservaba en periodos pre-industriales.

Según el IPCC, y en consenso con las principales comunidades científicas mundiales, uno de los principales causantes de generar dicho aumento de temperatura, que podría generar consecuencias irreversibles y catástrofes ambientales, sociales y económicas, es la emisión de gases de efecto invernadero.

## 2.4) Gases de efecto invernadero y su medición

Los gases de efecto invernadero, son aquellos gases que absorben y emiten energía radiante en un rango infrarrojo. Estos gases son capaces de absorber calor, generando el conocido “efecto invernadero”. Algunos de ellos son el Dióxido de Carbono (CO<sub>2</sub>), Metano (CH<sub>4</sub>), Óxido Nitroso (N<sub>2</sub>O), entre otros.

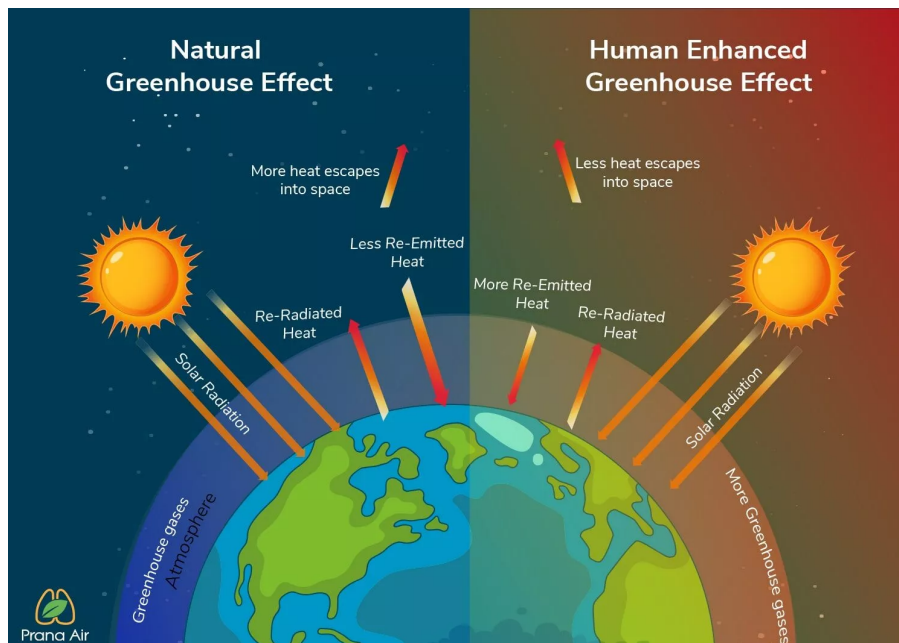
Mantener una cierta temperatura en la superficie terrestre, es fundamental para que la vida como la conocemos en la Tierra perdure. Es por ello, que la radiación solar permite la vida en la Tierra entre otras razones por ello. Esta radiación solar ingresa a la atmósfera terrestre, la cual parte es absorbida y

otra parte es reflejada de vuelta hacia el espacio. En un estado de atmósfera “normal”, con la cantidad necesaria de gases que contenía la atmósfera en épocas pre-industriales, dichos gases absorben calor de los rayos reflejados desde la Tierra al espacio, manteniendo una temperatura que permita la vida en la Tierra. Pero debido al accionar del Humano, se han incrementado los niveles de gases de efecto invernadero en la atmósfera.

Este excedente de gases, están absorbiendo mayor cantidad de calor reflejado por la radiación solar, generando el aumento de Temperatura de la superficie terrestre mencionado anteriormente en el estudio de IPCC, que puede tener consecuencias irreversibles:

**Figura 4**

*Efecto invernadero natural y reforzado por actividad humana*



Fuente: Prana Air

Como mencionamos anteriormente, dentro de las políticas de ESG las empresas deben trabajar para reducir sus impactos y mostrar su compromiso con la Sustentabilidad.

Pero para generar estrategias de mitigación, primero es necesario medir y conocer sus impactos ambientales, como la cantidad de gases de efectos invernaderos generados por su organización.

La medición de los gases de efecto invernadero, es el primer paso para definir una estrategia de mitigación, reportar transparentemente los impactos ambientales de la compañía y contribuir con reducir la temperatura mundial para evitar consecuencias irreversibles.

## **2.5) Medición de la Huella de carbono y otros Indicadores Ambientales**

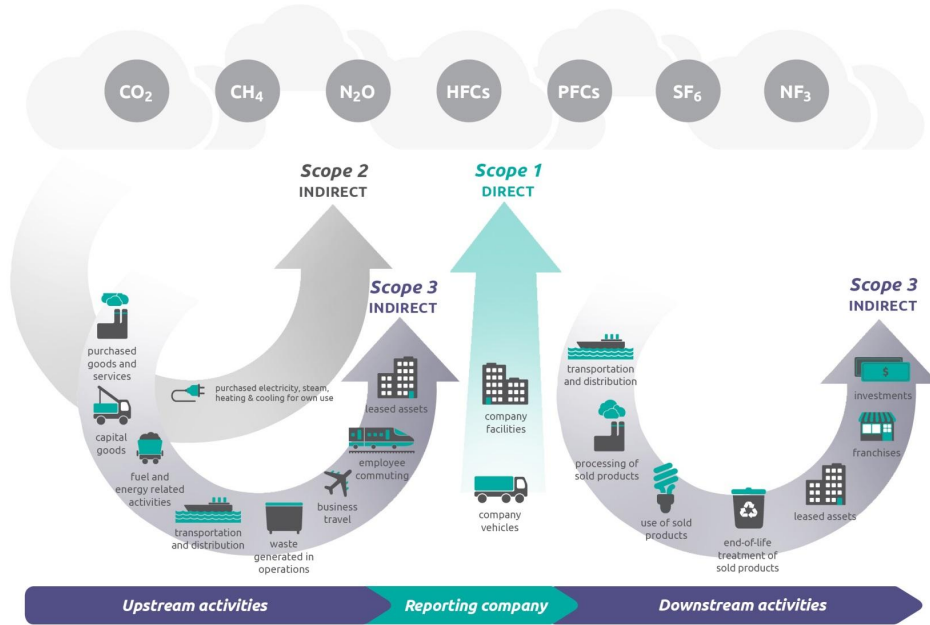
La Huella de Carbono (CO2) es uno de los principales Indicadores Ambientales de cualquier compañía que quiera seguir estándares ESG, así como medir, entender y comunicar sus impactos ambientales, para delinear finalmente una estrategia de mitigación de su impacto.

Para que las empresas puedan calcular y reportar su Huella de Carbono de manera estandarizada, existen diversos protocolos como el GHG Protocol desarrollado por el World Resource Institute, que permiten realizarlo de manera estandarizada.

Dicho protocolo, define tres tipos de Scope para el cálculo del “Inventario de Carbono” a lo largo de toda la cadena de valor de una compañía, como se puede observar en la siguiente imagen proporcionada por el World Resource Institute:

**Figura 5**

*Inventario de Carbono: Tipos de alcance*



Fuente: \*WRI/WBCSD Corporate Value Chain (Scope 3) Accounting and Reporting Standard

Como se puede observar, una empresa debe considerar múltiples fuentes de impacto ambiental a lo largo de la cadena. Se define como Scope 1 a aquellas emisiones que son controladas o pertenecientes a la organización (Ej: emisiones relacionadas con vehículos de la compañía, calderas propias, etc). Se define como Scope 2 a aquellas emisiones que físicamente ocurren en la empresa, pero sin embargo son las consecuentes con el uso energético de la empresa (Ej: compra de: energía, vapor, calefacción o refrigeración). Se define además como Scope 3 a aquellas emisiones resultantes de actividades que no son directamente controladas por la empresa u organización, pero con las cuales la organización impacta indirectamente en su cadena de valor.

El valor de los gases de efecto invernadero, se obtiene de cálculos que normalmente dependen de los llamados “factores de emisión”, que nos permiten convertir todos los tipos de consumo a la misma unidad de gas de efecto invernadero.

Por ejemplo, si conocemos el consumo eléctrico en Kwh de las oficinas, fábricas y centros de distribución de una compañía, podemos convertir la unidad de energía eléctrica en emisiones de CO<sub>2</sub> mediante factores de emisión internacionalmente validados.

Dependiendo del tipo de fuente con la cuál se haya generado la energía, lo cual es normalmente altamente dependiente de la matriz energética del país en el cuál se encuentra dicha empresa, se observan que los factores de emisión pueden tener distinto valor. Una matriz energética mayormente carbónica tendrá valores altos de emisión y una matriz energética con alto porcentaje de Energías Renovables tendrá un factor de emisión más bajo, debido a que la cantidad de emisiones asociadas a dichas energías renovables son menores que las emisiones generadas por fuentes carbónicas.

En la siguiente tabla, proporcionada por el Ministerio de Energía del Gobierno de España podemos ver algunos valores de factores de emisión de distintas fuentes y como el valor del carbón o gas son mayores a los de una fuente renovable como la Biomasa:

**Figura 6**

*Factores de emisión de CO2 para la matriz energética española*

Factores de emisiones de CO2			
	Fuente	Valores aprobados	Valores previos (****)
		kg CO2 /kWh E. final	kg CO2 /kWh E. final
Electricidad convencional Nacional	(*)	0,357	
Electricidad convencional peninsular	(**)	0,331	0,649
Electricidad convencional extrapeninsular	(**)	0,833	0,981
Electricidad convencional Baleares	(**)	0,932	
Electricidad convencional Canarias	(**)	0,776	
Electricidad convencional Ceuta y Melilla	(**)	0,721	
Gasóleo calefacción	(***)	0,311	0,287
GLP	(***)	0,254	0,244
Gas natural	(***)	0,252	0,204
Carbón	(***)	0,472	0,347
Biomasa no densificada	(***)	0,018	neutro
Biomasa densificada (pelets)	(***)	0,018	neutro

Fuente: \*\*FACTORES DE EMISIÓN DE CO2 y COEFICIENTES DE PASO A ENERGÍA PRIMARIA DE DIFERENTES FUENTES DE ENERGÍA FINAL CONSUMIDAS EN EL SECTOR DE EDIFICIOS EN ESPAÑA (RITE)

A partir de esto podemos observar cómo el cálculo de los gases de efecto invernadero de una compañía tiene a priori dos tipos de desafíos:

- La recolección de múltiples tipos de datos de distintas fuentes.
- El cálculo y procesamiento posterior de todos esos datos a unidades equivalentes de gases de efecto invernadero.

Teniendo en cuenta además que para ambos desafíos, los tipos de datos son completamente dinámicos conforme evolucione la compañía y en el segundo desafío son dinámicos conforme a la evolución de los factores de emisión.

Siguiendo el ejemplo anterior, si España decide invertir fuertemente en energías renovables, al entrar en operación todos esos parques eólicos y solares, el factor de emisión cambiará, por tanto la compañía debe seguir la actualización de todos los factores de emisión que los países reportan anualmente de manera pública.

De aquí se desprende además otro matiz, dado que si bien en este punto anterior hicimos foco en la medición y cálculo de los indicadores de Huella de Carbono y otros gases de efecto invernadero generados por una compañía; intrínsecamente se desprende de una estrategia ambiental que la compañía debe poder generar otros indicadores ambientales, como por ejemplo el Consumo de Energía Mensual (Kwh/mes), que permitan tomar estrategias de seguimiento y mitigación.

Por tanto, la medición de la huella desprende múltiples indicadores ambientales de interés para una empresa con los cuales la misma debe medir, hacer seguimiento y proponer estrategias de reducción y/o mitigación.

Para el estudio de esta tesis, nos enfocaremos principalmente en los indicadores ambientales mencionados en el apartado de [Datos](#).

El interés de la realización de esta Tesis recae en la realización de que existe un enorme desafío en términos de Data Management e implementación de tecnologías digitales para la medición, gestión y mitigación de impactos ambientales de una compañía. Principalmente dado que podemos observar que la medición y cálculo de los impactos ambientales de una compañía requieren fuertes herramientas y estructuras de datos para su gestión.

Para ello, primero haremos una revisión de la literatura existente sobre la relación entre la digitalización y la sustentabilidad, para entender cómo podemos contribuir con la gestión de datos del área de sustentabilidad utilizando herramientas digitales.

## **Revisión de la Literatura: “La digitalización y la sustentabilidad”**

Hoy en día, la Sustentabilidad se encuentra en la agenda de todas las organizaciones (Tata, 2021).

La literatura ha detectado a lo largo del tiempo, que existen diferentes factores internos y externos (Rodriguez et al, 2018), por los cuales una organización decide incorporar objetivos y prácticas de sustentabilidad en sus compañías. Algunas de las cuales hemos mencionado en la introducción de esta Tesis de Maestría.

Entre los factores internos más destacados se encuentran: los valores y principios de los dueños, la presión de inversores y el deseo de reducir costos (Rodriguez et al, 2018). Por el lado de factores externos se encuentran: la competencia (Carter & Rogers, 2008), regulaciones y normativas (Carter & Rogers, 2008), presión de la sociedad (Govindan, 2013) y la demanda de los clientes (Hendricks, K.B. & Singhal, V.R., 2005; Hussain, M., 2011). Esto genera por un lado, la necesidad de las empresas en mostrar que son una marca responsable (Tata, 2021), y por otro lado mejorar el reporting y la transparencia, junto con desarrollar buenas perspectivas que las ayuden a mantenerse responsables de sus acciones y decisiones (BNP Paribas, 2018).

Esta creciente importancia de la sustentabilidad ambiental, ha generado que las organizaciones se encuentren constantemente en la búsqueda de impulsar soluciones sistemáticas a lo largo de su cadena de valor que les permita desarrollar capacidades de sustentabilidad (BNP Paribas, 2018). Pero uno de los desafíos más grandes se encuentra en la búsqueda y análisis de datos para llevar a cabo la misma. (Tata, 2021).

Es por esta razón que se han generado diversas investigaciones, proyectos y análisis de organismos privados, públicos y educativos a fin de entender cómo la digitalización, los elementos de la Industria 4.0 y la ciencia de datos pueden contribuir con la sustentabilidad empresarial.

Esto ha llevado incluso a la definición de una nueva rama de Data Analytics, llamada Sustainable Supply Chain Analytics (Deloitte, 2013), también conocido como SSCA, la cual se define como el uso de las herramientas de business analytics para la recolección, análisis y circulación de datos relaciones con la sustentabilidad, con el objetivo de usar esta información para tomar decisiones eficientes y efectivas en desafíos de sustentabilidad (Deloitte, 2013).

Sin embargo, la literatura coincide en que todavía es necesario entender qué tecnología y herramientas son las que habilitarán la adopción de la SSCA en las empresas (Rodriguez et al, 2018). A medida que las empresas buscan nuevas maneras de integrar la sustentabilidad en sus estrategias y modelos de negocio, las soluciones data-driven están evolucionando para cubrir dichas necesidades y allí es donde las nuevas tecnologías relacionadas con datos asumirán un rol central en la estrategia de Sustentabilidad corporativa (BNP Paribas, 2018).

La literatura sobre cómo las tecnologías de Big Data Analytics (BDA) e Innovación puede catalizar la performance de las compañías generando que tengan ventajas competitivas, es extensa (Ramadan et al,



2020). Quienes toman las decisiones en las compañías, se valen del uso de BDA para definir estrategias que se traduzcan en mejoras de performance mediante innovación, competitividad y creación de valor (de Camargo Fiorini et al., 2018; Fosso-Wamba et al., 2017; Tan et al., 2015)

Sin embargo, en la literatura no se encuentra suficiente investigación empírica sobre el valor extraído de herramientas como Big Data para la innovación en el área de Sustentabilidad (Hao et al, 2019).

Las organizaciones necesitan herramientas de datos y de data analytics como catalizadores claves para hacer la transición de gestionar los impactos ambientales hacia tener una perspectiva de toda la cadena de valor y su impacto total (BNP Paribas, 2018). Para lograr el liderazgo de cadenas de suministro sustentables, las empresas deben utilizar sus capacidades de Big Data Analytics para reconocer y analizar diferentes fuentes de datos de manera de proveerse de perspectivas valiosas para la toma de decisión (Hu et al, 2018)

Esta capacidad de integrar recursos intrínsecos de una manera efectiva para lograr objetivos de Sustentabilidad, entregar valor sostenible a sus principales stakeholders y ganar ventaja competitiva sustentable, constituyen las capacidades de sustentabilidad de la empresa (Amui et al., 2017; Dao et al., 2011). Es por esto que la integración de Green Supply Chain Management junto con Big Data management permiten a las empresas desarrollar dichas capacidades que resultan en mejor performance ambiental y social. (Amui et al., 2017; Dao et al., 2011).

La implementación de estas prácticas requiere la aceptación, asimilación y rutinización de la Big Data (Wang et al., 2018; Gunasekaran et al., 2017; Hazen et al., 2012). Así como también la detección y comprensión de herramientas tecnológicas que permitan resolver estos desafíos de recolección y manejo de datos en la gestión sustentable.

Se han estudiado cuáles fueron las tecnologías más utilizadas y con mayor potencial catalizador para la Sustentabilidad en las empresas, y la literatura muestra que en la Industria 4.0 se encuentran grandes posibilidades dentro de las herramientas de Tracking en Tiempo Real, Gestión de Datos, adopción de Energías Renovables, Internet of Things, Big Data entre otros (Jamwal et al, 2021).

Dentro de la gestión de datos para la Sustentabilidad de las empresas, podemos encontrar diferentes desafíos en las distintas etapas de la gestión de datos: Data Ingestion, Data Warehouse, Data visualization y Predictive Analytics.

Podemos observar estudios que nombran herramientas con posibilidades de catalizar la sustentabilidad de una compañía como el trabajo ya mencionado de Jamwal et al, 2021 y existen otros como los trabajos de Camargo Fiorini et al., 2018, Fosso-Wamba et al., 2017 y Tan et al., 2015, donde se comentan desde los aspectos teóricos las implicancias positivas de su implementación y abordajes teóricos.

Sin embargo, en la literatura no existen estudios que complementen los análisis ambientales-económicos de los procesos operativos de las empresas, junto con sus necesidades de reporting y compliance, entendiendo también los desafíos y oportunidades de las herramientas de datos, desde un abordaje completamente práctico, con casos de uso que permitan la fácil aplicación para una compañía.

La literatura, como se observa en los apartados anteriores, se centra en las oportunidades de las herramientas digitales en pos de la contribución de mejorar las capacidades de sustentabilidad de la empresa como analizaron Amui et al., 2017 y Dao et al., 2011, pero no trabaja en cuantificaciones prácticas ni aplicativas de dichas soluciones: *¿Es eficiente respecto a costo e impacto ambiental aplicar cierta tecnología digital? ¿en qué aspectos de reporting, compliance y estrategia de sustentabilidad contribuye? ¿qué desafíos de la gestión de datos de la compañía se puede optimizar con dicha herramienta?*

Con este trabajo pretendemos colaborar con la literatura en que los líderes de sustentabilidad de las empresas de hoy en día, puedan encontrar soluciones y ejemplificaciones prácticas a estas preguntas.

## Desafíos de Data Management en la gestión ambiental

Como mencionamos anteriormente, a partir de las políticas de ESG adoptadas por las empresas, se han creado sectores de Medio Ambiente encargados de medir métricas importantes del negocio como la Huella de Carbono [TonCO<sub>2</sub>e], el Consumo Energético [Kwh], entre otros (Ver detalle en apartado de ["Datos"](#)), los cuales son variados, provienen de múltiples fuentes y en diversos formatos.

Siguiendo el ciclo de vida de los datos, comenzando por su adquisición, nos encontramos primeramente con el desafío de ir a buscar una enorme cantidad de datos para poder generar el cálculo de métricas relevantes de Sustentabilidad para el Negocio. Vemos fundamentalmente que se deben buscar datos de múltiples fuentes, ya sean, bases de datos cuyos datos son generados imputados automáticamente, y bases de datos cuyo input es manual, donde un colaborador se encarga de completar dichos datos.

En esta línea, podemos observar que para ambos tipos de bases de datos utilizadas, existen diferentes desafíos en cuanto al **Data Gathering**.

En primer lugar, para las bases de datos automáticas, los datos no necesariamente se encuentran en el formato o característica necesaria para toma de decisión ágil o cálculo de métricas del área. En muchos casos vemos que se realiza una etapa de procesamiento manual posterior, para adecuar el dato a la necesidad del área. Encausaremos este proceso de "Transform" más adelante, siguiendo el flujo de Extract-Transform-Load conocido como "ETL", pero seguiremos enfocándonos en todos los desafíos del proceso de Adquisición de Datos en primer lugar. Por otro lado, notamos la dificultad de que dichas bases de datos a abordar son múltiples.

Por el lado de las bases de datos manuales existe una mayor cantidad de desafíos. En principio, teniendo en cuenta que la carga es manual existe una mayor cantidad de errores por parte de los usuarios, se pueden colocar valores en formatos que no sirven para el posterior cálculo de las métricas de Sustentabilidad mencionadas. Si quisiéramos conocer el consumo eléctrico en "Kwh" de la compañía, pero tenemos bases cuyo input manual se realizó en otras unidades como "Mwh" o "J", además del valor propio del consumo eléctrico, deberemos adquirir y guardar también el valor de la unidad, lo que traspolado a los miles de valores que debemos obtener no solo ocupa espacio y tiempo de proceso adicional, sino que además obliga a incluir un paso en el procesamiento de dichos valores antes de guardarlos en la misma tabla. Por otro lado, el solicitar estos datos requiere un trabajo doble por parte de las áreas de Sustentabilidad y Medio Ambiente, en primer lugar el de detectar la personas que es responsable de completar dichos datos (que puede modificar debido a rotación o partida) y además hacer seguimiento unívoco a cada persona por la carga de los mismos.

Tenemos a su vez, datos que se necesitan sobre la flota logística que por su inherencia existen dificultades en su adquisición. Pensemos que si una empresa tiene una flota logística y tiene que pedirle a cada Driver que haga seguimiento del Odómetro de su vehículo para anotar la cantidad de kilometraje recorrido, el desafío es inescalable y con muchas posibilidades de errores y faltantes. Sin embargo, ya lo describiremos en detalle en párrafos más adelante, pero para calcular valores de Huella de Carbono [TonCO<sub>2</sub>e] de la logística, se necesitan detalles adicionales de los cuales estos datos a priori no son suficientes, dado que calcular la huella en base a la distancia entre dos puntos no va a ser tan preciso como hacerlo en función del detalle del recorrido exacto en "Kms".

Dentro del siguiente proceso del ciclo de vida de datos, **Data Warehousing**, las empresas se ven en la necesidad de encausar esta multiplicidad de datos en Bases de datos que permita el almacenamiento y

procesamiento de los datos obtenidos inicialmente de la etapa anterior mencionada, para el cálculo de las métricas ambientales relevantes. La falta de un sistema de bases de datos, es una de las causantes en que el cálculo de todas las métricas relevantes a partir de los datos imputados demore tiempos altos. Por otro lado, pensemos que debido a la diversidad de fuentes y tipos de datos, si el proceso de Data Gathering no es sólido la probabilidad de que podamos automatizar procesos de almacenamiento y procesamiento en bases de dato es baja. Esto también es debido a que los datos están descentralizados y constantemente se deben incluir nuevos datos, tablas o valores, las cuales generan un re-procesamiento constante. Otro punto que observamos, es que al momento de querer tomar decisiones de negocio sobre impactos ambientales el tener la información descentralizada, no colabora con la agilidad de tomas de decisiones.

Esto último queda atado al siguiente proceso de **Visualización**, donde el tener múltiples bases de datos también complejiza la posibilidad de generar visualizaciones de datos interesantes que permitan la toma de decisiones. La mayoría de métricas relevantes conllevan un pre-procesamiento y pre-cálculo de diversos parámetros, por lo cual tener una base de datos con un proceso de ETL facilitaría que la visualización sea obtenida directamente por las métricas últimas a medir. Es vital el desarrollo de Dashboards que permitan interpretar fácil y rápidamente impactos ambientales de la compañía y que contribuyan con tomas de decisiones estratégicas con agilidad, pero para ello, necesitamos que los dos procesos anteriores (Data Gathering y Warehousing) sean completamente robustos para asegurar que lo que estamos visualizando nos permite tomar el rumbo hacia la dirección correcta.

El último punto, **Predictive Analytics**, es el estadio al cual queremos desarrollar el terreno para poder potenciar a futuro la posibilidad de predecir ciertas métricas ambientales. Poder determinar predicciones de métricas a futuro, es muy importante para poder definir estrategias eficaces de Sustentabilidad, entender cuales son los proyectos que van a contribuir de mejor manera para reducir los impactos ambientales más importantes, entre otros. Este es el caso por ejemplo del Consumo Eléctrico (KWh), cuyo valor no solo es de gran importancia para calcular indicadores claves dentro de la estrategia ambiental de una compañía como la Huella de Carbono, donde el mismo tiene un peso interesante, sino además entender que la posibilidad de predecir el consumo eléctrico de las operaciones en base a ciertos parámetros actuales es un proceso clave para poder migrar hacia energías renovables y encontrar eficiencias energéticas.

El consumo eléctrico está íntimamente relacionado con la contribución de la Huella de Carbono de una compañía. El cálculo del mismo se realiza en función de indicadores conocidos como “factores de emisión”, los cuales dependen intrínsecamente de la Matriz Energética del país en dónde se calcula. Países con mayor contribución de Energías Renovables en su matriz energética (como Uruguay) tendrán un menor valor de Huella de Carbono emitida en relación con países cuyas matrices energéticas son mayoritariamente carbónicas como Argentina. Es por esto que la migración del consumo eléctrico propio de la empresa a Energías Renovables, es una de las principales estrategias para reducir la Huella de Carbono y conseguir eficiencias económicas. Una de las posibilidades para migrar a Energías Renovables es armar un acuerdo de compra venta de energía también conocido como “Power-Purchase Agreement” (PPA), por el cual una empresa negocia la compra de un volumen de energía a un precio fijo por un periodo normalmente largo de tiempo de a partir de 5 años. Debido a que la empresa se compromete la compra de energía directa a un generador por estos plazos largos, normalmente se consiguen beneficios económicos en comparación con las tarifas regulares del mercado eléctrico.

Sin embargo, conocer cuál será el consumo energético de una empresa en los próximos años, no es algo sencillo de predecir. Especialmente para multiplicidad de sitio en diferentes países. Para ello, las herramientas de Predictive Analytics pueden contribuir mucho en esta línea, pero sin duda que para llegar a este punto es necesario definir el end-game de las etapas antes mencionadas para permitir llegar al estadio de explorar herramientas predictivas.

#### **4.1) End-game definido para cada etapa**

Habiendo analizado los desafíos de Data Management dentro de la gestión ambiental y de sustentabilidad corporativa, donde pudimos barrer los distintos desafíos de todo el ciclo de vida de los datos, estamos en condiciones de definir el End-Game deseado para cada etapa proponiendo soluciones innovadoras con las herramientas de Analytics que contribuyan con los desafíos encontrados.

Para la etapa de **Data Ingestion**, necesitamos tener definido un framework de adquisición de datos que nos asegure:

- Calidad de los datos
- Formato deseado
- Frecuencia deseada
- Accesibilidad a distintas fuentes
- Escalabilidad ante el exponencial crecimiento de los datos a utilizar y de las operaciones de la empresa
- Reducir el tiempo dedicado a la toma de datos

Para ello se analizarán esquemas de Web Scrapping, APIs de ingestión, tecnologías de IoT y queries de SQL que nos faciliten la ingestión de datos para encauzarlas en la siguiente etapa del proceso: Data Warehousing.

Para la etapa de Data Warehousing, es necesario definir un sistema de Bases de Datos (DBMS), donde se puedan volcar los miles de datos obtenidos de la primer etapa y con la cuál se puedan realizar procesamientos ágiles, permitiendo la fácil escalabilidad, asegurando el formato deseado y reduciendo los tiempos de procesamiento. Este sistema de bases de datos nos tendrá que permitir:

- Facilidad en la gestión de datos desde la imputación hasta su procesamiento
- Agilidad en el procesamiento de datos
- Integrabilidad con sistemas de visualización para permitir tomas de negocio
- Escalabilidad ante nuevos escenarios de datos, nuevas métricas ambientales y diversidad de nuevas operaciones.

Para ello se analizarán sistemas de Bases de Datos en SQL, su integración con la multiplicidad de fuentes de datos de la etapa de Data Gathering, junto con definición de procesos que nos permitan la gestión deseada del Data Warehousing.

En la siguiente etapa del proceso, Data Visualization, buscaremos desarrollar Dashboards y métricas cuantitativas, que nos permitan:

- Visualización de las principales métricas ambientales y de sustentabilidad
- Contribuir con métricas visibles que faciliten la toma de decisiones
- Detectar nuevos “insights” claves.
- Contribuir con la definición de nuevas estrategias de sustentabilidad

Para ello, una vez desarrolladas las bases de datos, desarrollaremos Dashboards y effective charts que contribuyan con la toma de decisiones y el story-telling sobre los impactos ambientales. Exploraremos softwares comerciales para el desarrollo de los mismos.

Finalmente para la etapa de Predictive Analytics, una vez sentadas las bases de todas las etapas anteriormente mencionadas, dejaremos un caso base establecido para segmentar el consumo energético de distintos sitios, mediante la utilización de un algoritmo de clustering, que nos contribuya con la proyección de consumos a futuro para la elaboración de Power-Purchase Agreements con Energías Renovables.

## Experimentación y análisis de resultados

### 5.1) Herramientas digitales en el proceso de Data Gathering

#### 5.1.1) Cálculo de Emisiones de CO2 para Flota Logística

Como ya mencionamos anteriormente, uno de los desafíos más grandes dentro del proceso de Data Gathering, es la fiabilidad de la información y la escalabilidad junto con el rastreo de datos en cada periodo.

La logística de vehículos propios y de terceros, son parte de los Scopes de emisiones de gases de efecto invernadero y Huella de carbono de una compañía, como ya comentamos anteriormente. Según el protocolo GHG, el cálculo de emisiones para logística propia y de terceros relacionados, también conocido como “Mobile Combustion”, puede calcularse de dos maneras diferentes:

La primera opción, que según el protocolo es la recomendada para medir las emisiones, es calcular las emisiones de CO2 mediante el consumo de combustibles:

**Figura 7**

*Ecuación para el cálculo de emisiones de CO2 según IPCC*

**EQUATION 3.2.1**  
**CO<sub>2</sub> FROM ROAD TRANSPORT**

$$Emission = \sum_a [Fuel_a \cdot EF_a]$$

Fuente: <sup>24</sup>2006 IPCC Guidelines for National Greenhouse Gas Inventories

Siendo:

- Emission: Emisiones de CO2 [Kgs]
- Fuel (a): Combustible consumido [T]
- EF (a): Factor de emisión [Kg/T]. Equivalente al contenido carbónico del combustible multiplicado por 44/12
- a: Tipo de combustible (Ej: nafta, diesel, gas natural, LPG, etc)

La otra metodología para calcular las emisiones de la logística tanto de CO2 como otros gases (CH4, N2O, etc), son en base a la distancia recorrida (Kms), donde la ecuación es similar a la anterior, con la diferencia de que el factor de emisión está medido en kgCO2/Kms.

El protocolo GHG recomienda fuertemente utilizar la primera metodología de medición basada en el tipo de combustible, dado que es más precisa, y dado que el CO2 representa más del 95% de la contribución de los gases de efecto invernadero, es preferible que las empresas busquen reforzar sus mediciones basadas en consumo de combustible.

Esto representa un enorme desafío en términos de Data Gathering. Pensemos que una empresa al querer medir la Huella de Carbono de su flota logística (ya sea Scope 1 o Scope 3), debe tener una gran cantidad de datos de su flota para poder medir. Analicemos el desafío en detalle:

Supongamos que la empresa decide hacerlo conforme a la segunda metodología (basada en la distancia recorrida (Kms)), dado que los datos de distancia recorrida por su propia flota serían más fáciles de obtener que el consumo de combustible. La empresa necesitaría un detalle de todos los recorridos exactos realizados por cada uno de los vehículos de su flota, tengamos en cuenta además la gigante variabilidad que esto representa. Dependiendo de la demanda de entregas, se podrían incorporar más o menos vehículos, si bien la empresa podría estandarizar las rutas que dichos vehículos realizan conociendo los puntos de inicio y fin, sabemos que cualquier corte de tránsito, decisión del driver, entre otras, podría generar que dichos vehículos opten por rutas diferentes lo cual generaría variabilidad en los Kms.

Si bien se tratan solo de dos simples datos (Cantidad de vehículos y Kms Recorridos), podemos observar que su variabilidad es enorme, y por otro lado en términos de la adquisición de los datos en sí, podemos notar la complejidad para tomar estos datos... los drivers tendrían que completar una planilla manual, podrían olvidarse de completarla o hacerlo de manera incorrecta (puntos de fallas de carga manual que mencionamos en los apartados anteriores), en caso de que un driver sea nuevo tendría que estar capacitado sobre lo mismo, entre otros múltiples desafíos.

Por otro lado, desde el punto de vista Ambiental, cualquier vehículo que recorre una ruta desde un punto A hacia otro B realiza paradas intermitentes, sea por semáforos, manifestaciones, paradas técnicas, entre otras. Si el vehículo se encuentra parado en un semáforo o en una manifestación, pero con el motor encendido durante un cierto tiempo, por más que la distancia recorrida haya sido nula, dado que el motor del vehículo sigue encendido el consumo de combustible ha continuado y por tanto también se han generado emisiones. Este consumo del combustible cuando el vehículo está parado pero con el motor encendido se conoce como **"Ralenti"**. Aquí es donde cobra sentido la recomendación del protocolo GHG sobre medir las emisiones de Mobile Combustion en función del combustible consumido y no sobre la distancia recorrida. Como podemos observar, la medición de Huella de Carbono en función a distancia recorrida jamás podría capturar los valores de Ralenti de una manifestación donde el vehículo se encontró demorado unas horas.

Según el Foto de Defensa Ambiental (Environmental Defense Fund), un camión a diesel puede consumir aproximadamente 1 Gallon de combustible por cada hora de Ralenti (idling). Si consideramos las Tablas 12 y 14 de la WRI GHG Emissions Factors Compilation, podemos observar que para un camión a Diesel los valores de emisiones son:

$$CO2 \text{ Emissions per Mile}_{\text{Heavy Duty Vehicle (Diesel)}} = 1,71 \frac{kg_{CO2}}{\text{Mile}}$$

Si consideramos una ruta de 100 millas, que podríamos obtener del Odómetro del vehículo de manera manual, las emisiones serían:

$$CO2 \text{ Emissions per Hs}_{\text{Heavy Duty Vehicle (Diesel)}} = 1,71 \frac{kg_{CO2}}{\text{Mile}} * 100 \text{ Miles} = 171 kg_{CO2}$$

Pero si pudiéramos conocer el consumo del combustible del vehículo, podríamos utilizar el factor de emisión para Diesel de la misma tabla de WRI GHG Emissions Factors Compilation que sería más próximo a la realidad. Tomando el consumo de combustible para ese trayecto de un promedio de Fuel Efficiency de Heavy Duty Vehicles de 5,5 Miles/Gallon:

$$Avg. \text{ Fuel Efficiency}_{\text{Heavy Duty Vehicle (Diesel)}} \approx 5,5 \frac{\text{Mile}}{\text{Gallon}}$$

$$CO_2 \text{ Emissions per Gallon}_{\text{Heavy Duty Vehicle (Diesel)}} = \frac{100 \text{ Miles}}{5,5 \frac{\text{Mile}}{\text{Gallon}}} * 10,13 \frac{\text{Kg CO}_2}{\text{Gallon}} = 184 \text{ kg}_{CO_2}$$

Si tomamos un vehículo moderno como el Volvo FH I-Save con una sustancial mejora de eficiencia llegando a 8,65 Miles/Gallon:

$$Avg. \text{ Fuel Efficiency}_{\text{Heavy Duty Vehicle (Volvo FH Diesel)}} \approx 8,65 \frac{\text{Mile}}{\text{Gallon}}$$

$$CO_2 \text{ Emissions per Gallon}_{\text{Heavy Duty Vehicle (Diesel)}} = \frac{100 \text{ Miles}}{8,65 \frac{\text{Mile}}{\text{Gallon}}} * 10,13 \frac{\text{Kg CO}_2}{\text{Gallon}} = 117 \text{ kg}_{CO_2}$$

Podemos observar notoriamente, como podemos en dos simples ejemplos mejorar la calidad de nuestras mediciones de Huella de Carbono en logística trabajando con datos reales de consumo de combustible, en vez de distancias. Por tanto, los equipos de Telematics que permiten conexasión con el vehículo y obtener datos en real-time de consumo de combustible nos permitirán no solo facilitar el Data Gathering, (realizándolo via API o descarga de datos masivos desde un data-logger), sino también afinar el valor de Huella de nuestras mediciones, entendiendo también que esto nos permitirá trabajar en la economía de nuestra flota, detectando aquellos vehículos que tienen un bajo rendimiento de Fuel Efficiency y en los cuales se podría reducir el gasto en combustible.

Supongamos que en ese trayecto de 100 millas, el vehículo se encuentra parado un total de 1hs (con paradas intermitentes) sin apagar el motor (en Ralentí). Como mencionamos anteriormente, según EPA:

$$Ralentí \text{ Fuel Consumption}_{\text{Heavy Duty Vehicle (Diesel)}} = 1 \frac{\text{Gallon}}{\text{hs}}$$

Utilizaremos el valor de emisiones para Diesel de la misma tabla de WRI GHG Emissions Factors Compilation, para calcular la incidencia de las emisiones de ralentí aproximada por milla:

$$CO_2 \text{ Emissions per Gallon}_{\text{On-Road Diesel Fuel}} = 10,13 \frac{\text{Kg CO}_2}{\text{Gallon}}$$

$$CO_2 \text{ Emissions per Mile}_{\text{On-Road Diesel Fuel}} = 10,13 \frac{\text{Kg CO}_2}{\text{Gallon}} * 1 \frac{\text{Gallon}}{\text{Hs}} * 1\text{Hs} = 10,13\text{Kg CO}_2$$

$$\% \text{ Emissions}_{\text{Idling/Fuel}} = \frac{10,13 \text{ Kg CO}_2}{117 \text{ Kg CO}_2} \approx 8\%$$

$$\% \text{ Emissions}_{\text{Idling/Fuel}} = \frac{10,13 \text{ Kg CO}_2}{184 \text{ Kg CO}_2} \approx 5\%$$

Por tanto, con este sencillo cálculo podemos ver que el peso de las emisiones generadas por el Ralentí de un motor podría estar entre un 5% y 8%. Por tanto, el contemplar el Ralentí es de suma importancia para el cálculo efectivo de emisiones de carbono, debido a su peso.

Sin ninguna duda todas estas estrategias y decisiones, tendrán su correlato económico, pero como podemos identificar, el caso del Ralentí es un caso donde al poder medirlo no solo podemos detectar eficiencias ambientales (por reducción de emisiones), sino también económicas, teniendo en cuenta que es combustible gastado en realizar una cantidad de kilómetros nula.

Desde el punto de vista Ambiental, tener una clara y precisa valoración de las emisiones, como vimos en este ejemplo de las emisiones de flota, es de suma importancia dado que permite tomar estrategias y decisiones correctas: ¿hay que trabajar sobre la optimización de rutas? ¿sobre la capacitación a los drivers? ¿sobre el tipo de combustible por uno menos contaminante? podrían ser algunas de las preguntas disparadoras para buscar la eficiencia en emisiones de flota, pero solo podremos definir una correcta estrategia Ambiental si medimos los indicadores correctos.

Si nos enfocamos en medir distancias y cantidad de vehículos, no podremos elaborar mucho, pero si podemos lograr medir el consumo de combustible de manera precisa, entendiendo tiempos de Ralentí, podremos buscar optimizar el cálculo de las emisiones de CO2 así como también generar ahorros económicos para la logística de la empresa.

Volviendo al punto inicial de este apartado, como podemos imaginar, el poder medir el consumo de combustible de toda nuestra flota logística es realmente un gran desafío de Data Gathering.

Como mencionamos anteriormente, buscaremos evitar procesos de carga manual de datos ya que conllevan muchas complejidades, son permeables al error humano y no son escalables en el tiempo.

### **5.1.2) API para medición de consumo de Combustible y medición de Ralentí**

Cómo se mencionó anteriormente en el segmento de “Metodología”, utilizaremos el siguiente extracto de la Bases de Datos llamada “Heavy-Duty Vehicle Activity” del proyecto MOVES (Motor Vehicle Emission Simulator) recopilada conjuntamente entre la Agencia de Protección Ambiental de los Estados Unidos (EPA), el Laboratorio Nacional de Energías Renovables de los Estados Unidos (NREL) y la Universidad de California, compuesta por un Dataset que incluye 564 vehículos comerciales y más de 23.000 días operativos de dichos vehículos, cubriendo siete tipos de vehículos de Heavy-Duty (del tipo que conversamos anteriormente para los casos de Line-Haul o Media milla logística).

Aquí podemos ver un extracto de la tabla, con los datos de vehículos junto con sus valores de operación, Ralentí y distancias recorridas:

**Tabla 1**

*Actividad para vehículos pesados (Heavy-Duty)*

```
> head(Idle_Operation)
  VehicleID DayID HourID InstrumentedHours OperationTime IdleTime
1         1     5       1             270         562090  277788
2         1     5       2             270         556965  326278
3         1     5       3             270         543262  355250
4         1     5       4             270         571474  411965
5         1     5       5             270         569150  430249
6         1     5       6             270         561242  445466
  WorkDayIdleTime ExtendedIdleTime MilesDriven IdleFraction
1         44809         232979         3016.217  0.4942056
2         34946         291332         2389.005  0.5858142
3         25721         329529         1897.549  0.6539202
4         30736         381229         1627.872  0.7208814
5         15691         414558         1468.641  0.7559501
6         11738         433728         1236.605  0.7937147
  WorkDayIdleFraction ExtendedIdleFraction
1         0.07971855         0.4144870
2         0.06274362         0.5230706
3         0.04734548         0.6065747
4         0.05378372         0.6670977
5         0.02756918         0.7283809
6         0.02091433         0.7728003
```



Fuente: Análisis basado en la base de datos de <sup>2014</sup>Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Debajo dejamos un detalle del significado de cada uno de las variables medidas:

**Tabla 2**

*Descripción de los campos de la Tabla 1*

Nombre del Campo	Descripción
VehicleID	ID Number of vehicles
Date	Date
DayID	Date ID (weekday = 5, weekend = 2)
HourID	Hour of the day (1-24)
OperationTime	Time with engine operation in the specific hour (sec)
IdleTime	Total engine idle time in the specific hour (sec)
WorkDayIdleTime	Total workday idle time in the specific hour (sec)
ExtendedIdleTime	Total extended idle time in the specific hour (sec)
MilesDriven	Miles driven along this operation time

Fuente: Análisis basado en la base de datos de <sup>2014</sup>Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Siguiendo el proceso de análisis de datos aprendido en la Maestría, comenzaremos analizando dicha bases de datos mediante el siguiente proceso:

1) Eliminación de datos faltantes y "nulls"

Como toda base de datos, es propensa a que contenga "nulls" y datos faltantes. Pensemos que los sistemas interconectados también son proclives a que en el proceso de captura de datos generen faltas por distintos motivos como puede ser la falla del equipo, desconexión temporal o algún otro inconveniente.

En el caso de esta base de datos, vemos que contiene solamente 362 filas (eventos de medición) con errores o "nulls":

**Tabla 3**
*Resumen en R sobre los valores de la Tabla 1*

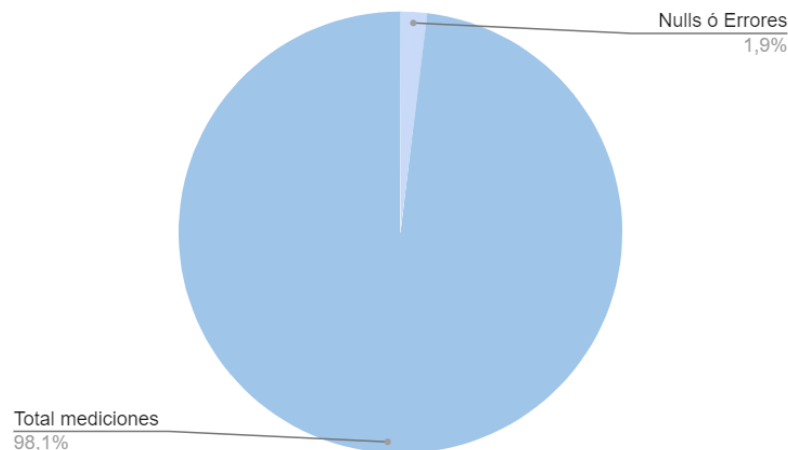
IdleFraction	workDayIdleFraction	ExtendedIdleFraction
Min. :0.0000	Min. :0.0000	Min. :0.0000
1st Qu.:0.1353	1st Qu.:0.0895	1st Qu.:0.0000
Median :0.3127	Median :0.2265	Median :0.0000
Mean :0.3566	Mean :0.2838	Mean :0.0728
3rd Qu.:0.5074	3rd Qu.:0.4212	3rd Qu.:0.0000
Max. :1.0000	Max. :1.0000	Max. :1.0000
NA's :362	NA's :362	NA's :362

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021), Summary en R

El Dataset contiene 18.303 filas, en otras palabras, eventos de medición por hora del día, por tanto es un valor muy bajo representando solamente menos de un 2% de las filas aproximadamente las cuales vamos a eliminar.

**Figura 7**
*Porcentaje de datos faltantes vs total de DataSet*

```
> sum(is.na(Idle_operation$IdleFraction))/nrow(Idle_operation)
[1] 0.01977818
```



Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021), Summary en R

Es interesante poder entender las características de estos faltantes (nulls), para definir si vamos a avanzar eliminándolos del dataset, o si son información válida que nos permite obtener algún insight interesante.

Para ello vamos a revisar algunas características de las filas que tienen datos faltantes, filtrando por aquellas mediciones que contengan faltantes:

**Tabla 4**

*Extracto de filas con NA sobre la actividad de vehículos pesados*

```
> head(Idle_Operation[is.na(Idle_Operation$IdleFraction), ] )
  VehicleID DayID HourID InstrumentedHours OperationTime
5700      134     2     19                6             0
5859      140     2     18                6             0
6233      150     2     1                6             0
6238      150     2     6                6             0
6328      153     5     2                24             0
6329      153     5     3                24             0
  IdleTime WorkDayIdleTime ExtendedIdleTime MilesDriven
5700      0                0                0             0
5859      0                0                0             0
6233      0                0                0             0
6238      0                0                0             0
6328      0                0                0             0
6329      0                0                0             0
  IdleFraction WorkDayIdleFraction ExtendedIdleFraction
5700          NA                NA                NA
5859          NA                NA                NA
6233          NA                NA                NA
6238          NA                NA                NA
6328          NA                NA                NA
6329          NA                NA                NA
```

Fuente: Análisis basado en la base de datos de <sup>2014</sup>Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Como podemos observar, los NAs se generan en los campos fraccionales, en otras palabras, aquellos que dividen dos campos anteriores para obtener porcentajes. El valor NA se genera debido que por ejemplo el campo "Idle Fraction" que nos informa el porcentaje de tiempo que el vehículo estuvo en Ralentí sobre el tiempo operativo total, está dividiendo dos campos que tienen valor cero: el tiempo Operativo (OperationTime) y el tiempo de Ralentí (IdleTime).

A su vez, podemos observar que en todos los casos tanto los Tiempos Operativos como los Tiempos de Ralentí tienen valor 0 (incluso podemos ver esto sumando el valor de todas esas columnas:

**Tabla 5**

*Suma de Tiempos Operativos y Tiempos de Ralentí para los valores con NA*

```
> a <- Idle_Operation[is.na(Idle_Operation$IdleFraction), ]
> sum(a$IdleTime)
[1] 0
> sum(a$MilesDriven)
[1] 1.171851
> sum(a$OperationTime)
[1] 0
```

Fuente: Análisis basado en la base de datos de <sup>2014</sup>Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Podemos observar que si bien el valor de Millas recorridas no es estrictamente cero, es un valor representativamente muy bajo.

Todo parecería indicar que estas mediciones fueron realizadas con el motor de cada vehículo apagado, que por algún motivo el sistema de Telematics siguió almacenando dichos datos. Dado que para el análisis que buscamos realizar entendiendo eficiencias ambientales y económicas, nos interesa entender la operación con motor encendido, vamos a decidir generar un nuevo extracto eliminando estas filas (que representan menos del 2% del total del Dataset).

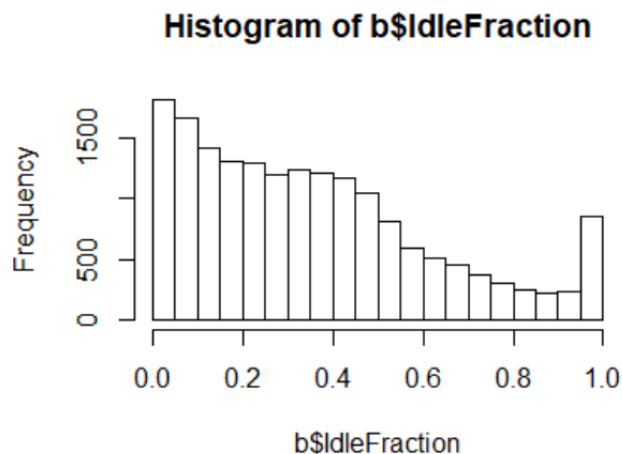
2) Una vez limpios los errores y nulls, vamos a proceder con el análisis estadístico de la database

Primeramente, nos interesa conocer cuán representativo es el monto de Ralentí en una flota de vehículos reales.

Si a primera vista hiciéramos un análisis del porcentaje de tiempo de Ralentí en función del tiempo de operación total por vehículo, vemos que existe una dispersión enorme entre los resultados:

**Figura 8**

*Histograma sobre Tiempo de Ralentí*



```
> summary(b$IdleFraction)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.0000 0.1353  0.3127  0.3566 0.5074  1.0000
```

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Para ello, realizamos un análisis de Pareto para incluir los vehículos cuyas mediciones han sido mayores, de manera de poder acotar el rango de Ralentí. Observamos que no necesariamente tomando los vehículos con mayor cantidad de mediciones, el Promedio de % de Tiempo de Ralentí en función del Tiempo Operativo del vehículo se reduce significativamente. Observamos que realizando un Pareto donde tomamos los vehículos que representan un 70% de las mediciones tomadas, el valor promedio no disminuye significativamente con respecto al considerar todas las mediciones decreciendo de un valor de 37,9% a 34,4%. En contraste, si observamos que el Máximo de dicha medición si disminuye significativamente, entendiendo que un vehículo que estuvo 92% en Ralentí, prácticamente significa que estuvo detenido con el motor encendido en su gran mayoría del tiempo. No obstante, se observan prácticas en la

logística donde esto ocurre y por desconocimiento del costo del Ralentí (tanto ambiental como económico), puede ocurrir que un driver mantenga su vehículo encendido aguardando por cargas en el puerto por ejemplo, donde las esperas suelen ser largas.

**Tabla 6**

*Promedios de Porcentajes de Ralentí en función del volumen de DataSet*

	% Idle	Min % Idle	Max % Idle
Promedio Dataset	37,9 %	3%	92%
Promedio 70% Pareto	34,6 %	3%	71%
Promedio 90% Pareto	35,50%	3%	76%

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Para el caso, tomaremos los valores de Ralentí correspondientes al 15% de vehículos con mayor cantidad de mediciones, como valores representativos del grupo, de manera de no sobredimensionar el peso del Ralentí en una flota y realizar un análisis conservador. Los casos como el anterior donde podemos identificar los valores mayores de Ralentí, podemos identificarlos por separado, estudiar estos casos por separado (analizando rutas, vehículos utilizados, entendiendo el comportamiento de los drivers, condiciones particulares del recorrido, etc), para poder tomar estrategias a fin.

- 3) Para este 15% de vehículos, por medio de la variable "MilesDriven", podemos obtener el total de millas recorridas, el total de tiempo operativo y el total de tiempo de Ralentí. Observamos un extracto de la tabla:

**Tabla 7**

*Análisis de variables principales para el Top 15% de vehículos del DataSet*

Top 15% Vehicles	Operation Time (seg)	% Idle	Idle Time (seg)	Total Miles (Mls)
1	14482734	0,52	7503298	12496
2	14346440	0,49	6992636	16004
3	9518492	0,46	4396038	11273
4	12631475	0,50	6334231	15306
5	15096525	0,54	8156137	18378
6	12489415	0,47	5816939	12300
8	1041420	0,54	565634	13565
12	1491172	0,34	510735	12070
15	1462305	0,07	103972	16609
21	1608681	0,12	187432	18348
22	2050629	0,35	725448	15975
28	2394975	0,06	133184	19861
29	2647450	0,30	803187	16464

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Podemos agregar las columnas de las emisiones de CO2 generadas en función de las millas recorridas (tomando el valor de las Tablas 12 y 14 de la WRI GHG Emissions Factors

Compilation), y obtenemos las emisiones de cada vehículo en dichas rutas, basadas en las Millas recorridas:

**Tabla 8**

*Cálculo de emisiones de CO2 en función de millas recorridas para el top 15%*

Top 15% Vehicles	Operation Time (seg)	% Idle	Idle Time (seg)	Total Miles (Mls)	Emissions KgCo2 (Miles)
1	14482734	0,52	7503298	12496	21369
2	14346440	0,49	6992636	16004	27367
3	9518492	0,46	4396038	11273	19277
4	12631475	0,50	6334231	15306	26173
5	15096525	0,54	8156137	18378	31427
6	12489415	0,47	5816939	12300	21033
8	1041420	0,54	565634	13565	23196
12	1491172	0,34	510735	12070	20639
15	1462305	0,07	103972	16609	28401
21	1608681	0,12	187432	18348	31375
22	2050629	0,35	725448	15975	27317
28	2394975	0,06	133184	19861	33963
29	2647450	0,30	803187	16464	28153
32	1431379	0,11	158202	17837	30502
35	3211994	0,29	939391	20630	35278
37	2556768	0,30	777406	16025	27403
39	1636222	0,13	216624	13051	22318
40	1169722	0,17	200460	15342	26235
45	1959193	0,36	703965	19125	32704

Fuente: Análisis basado en la base de datos de <sup>2014</sup>Heavy-Duty Vehicle Activity\* (EPA-MOVES, 2021)

Como sabemos, en términos de Data Gathering el seguimiento de las millas recorridas por vehículo representa un reto significativo, manual, con baja capacidad de escalar y baja fiabilidad da calidad de datos, y además no es el proceso más recomendable según IPCC para el cálculo de emisiones de carbono. Aquí ya de primera mano, podemos observar las ventajas de los sistemas de Telematics en cuanto a calidad de información sobre Millas junto con su fácil adquisición.

Por tanto, de la Tabla de los vehículos “Datos de Vehículos” proporcionada por NREL, donde se encuentra el detalle de las características de los vehículos, tomaremos dos valores promedios para dicho tipo de camiones Heavy Duty en términos de Fuel Efficiency de 5,5 Mpg y 8.5 Mpg de manera de estimar los consumos de combustible (existen equipo de Telemática que tienen la posibilidad de tener seguimiento del nivel de tanque y por tanto de obtener estos datos via API de manera muy sencilla, pero esta base de datos fue construida con Data Loggers que no poseían dichas capacidades), y posteriormente calcular las emisiones de CO2 basadas en dicho consumo con las tablas de WRI GHG Emissions Factors Compilation:

$$CO2\ Emissions\ per\ Gallon_{On-Road\ Diesel\ Fuel} = 10,13 \frac{Kg\ CO2}{Gallon}$$

**Tabla 9**

*Diferencia de cálculo de emisiones de CO2 en función de millas recorridas y en función del consumo de combustible.*

Top 15% Vehicles	Operation Time (seg)	% Idle	Idle Time (seg)	Total Miles (Mls)	Emissions KgCo2 (Miles)	Max Gallons (5.5 MPG)	Min Gallons (8.5 MPG)	Max Emissions by Gallons (KgCo2)	Min Emission by Gallons (KgCo2)
1	14482734	0,52	7503298	12496	21369	2272	1470	23016	14893
2	14346440	0,49	6992636	16004	27367	2910	1883	29476	19073
3	9518492	0,46	4396038	11273	19277	2050	1326	20763	13435
4	12631475	0,50	6334231	15306	26173	2783	1801	28191	18241
5	15096525	0,54	8156137	18378	31427	3342	2162	33850	21903
6	12489415	0,47	5816939	12300	21033	2236	1447	22655	14659
8	1041420	0,54	565634	13565	23196	2466	1596	24984	16166
12	1491172	0,34	510735	12070	20639	2195	1420	22230	14384
15	1462305	0,07	103972	16609	28401	3020	1954	30591	19794
21	1608681	0,12	187432	18348	31375	3336	2159	33794	21867
22	2050629	0,35	725448	15975	27317	2905	1879	29423	19038
28	2394975	0,06	133184	19861	33963	3611	2337	36581	23670
29	2647450	0,30	803187	16464	28153	2993	1937	30323	19621
32	1431379	0,11	158202	17837	30502	3243	2099	32853	21258
35	3211994	0,29	939391	20630	35278	3751	2427	37997	24586
37	2556768	0,30	777406	16025	27403	2914	1885	29516	19098
39	1636222	0,13	216624	13051	22318	2373	1535	24038	15554
40	1169722	0,17	200460	15342	26235	2789	1805	28257	18284
45	1959193	0,36	703965	19125	32704	3477	2250	35225	22793
61	2013234	0,42	853803	18738	32041	3407	2204	34511	22331

Fuente: Análisis basado en la base de datos de <sup>21</sup>"Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

Podemos observar que al calcular las emisiones mediante los Gallons consumidos en vez de las Millas recorridas, el valor puede fluctuar en función del Fuel Efficiency, por tanto es lógica la recomendación del IPCC de buscar medir las emisiones basadas en consumo de combustible real.

Por otro lado, vamos a hacer foco en el peso de las emisiones por Ralentí, valor que solo puede obtenerse de utilizar un Data Logger conectado al motor del vehículo.

Para ello, conociendo el valor del IdleTime, en otras palabras, la suma de segundos en los que el motor del vehículo estuvo en Ralentí, podemos obtener las horas de Ralentí de dicho vehículo en total. Siguiendo el valor de EPA, por el cual en Ralentí un Heavy Duty Vehicle consume 1 Gallon por hora, fácilmente estimamos los galones de combustible consumidos durante el Ralentí de cada vehículo:

**Tabla 10**
*Consumo estimado de combustible durante el tiempo de Ralentí*

Top 15% Vehicles	Operation Time (seg)	% Idle	Idle Time (seg)	Total Miles (Mils)	Emissions KgCo2 (Miles)	Max Gallons (5.5 MPG)	Min Gallons (8.5 MPG)	Max Emissions by Gallons (KgCo2)	Min Emission by Gallons (KgCo2)	% Max Emissions Gallon vs Min	% Min Emissions Gallon vs Mile	Gallons Idle (hs)
1	14482734	0.52	7503298	12496	21369	2272	1470	23016	14893	1,08	0,6969	2084
2	14346440	0.49	6992636	16004	27367	2910	1883	29476	19073	1,08	0,6969	1942
3	9518492	0.46	4396038	11273	19277	2050	1326	20763	13435	1,08	0,70	1221
4	12631475	0.50	6334231	15306	26173	2783	1801	28191	18241	1,08	0,70	1780
5	15096525	0.54	8156137	18378	31427	3342	2162	33850	21903	1,08	0,70	2266
6	12489415	0.47	5816939	12300	21033	2236	1447	22655	14659	1,08	0,70	1616
8	1041420	0.54	565634	13565	23196	2466	1596	24984	16166	1,08	0,70	157
12	1491172	0.34	510735	12070	20639	2195	1420	22230	14384	1,08	0,70	142
15	1462305	0.07	103972	16609	28401	3020	1954	30591	19794	1,08	0,70	29
21	1608881	0.12	187432	18348	31375	3336	2159	33794	21867	1,08	0,70	52
22	2050629	0.35	725448	15975	27317	2905	1879	29423	19038	1,08	0,70	202
28	2394975	0.06	133184	19861	33963	3611	2337	36581	23670	1,08	0,70	37
29	2647450	0.30	803187	16464	28153	2993	1937	30323	19621	1,08	0,70	223
32	1431379	0.11	158202	17837	30502	3243	2099	32853	21258	1,08	0,70	44
35	3211994	0.29	939391	20630	35278	3751	2427	37997	24586	1,08	0,70	261
37	2556768	0.30	777406	16025	27403	2914	1885	29516	19098	1,08	0,70	216
39	1636222	0.13	216624	13051	22318	2373	1535	24038	15554	1,08	0,70	60
40	1169722	0.17	200460	15342	26235	2789	1805	28257	18284	1,08	0,70	56
45	1959193	0.36	703965	19125	32704	3477	2250	35225	22793	1,08	0,70	196
61	2013234	0.42	853803	18738	32041	3407	2204	34511	22331	1,08	0,70	237

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

El primer hallazgo interesante tiene una arista fuertemente económica, que es qué peso tiene el Ralentí en el consumo total de combustible. Podemos observar que en promedio el consumo de combustible por Ralentí representó entre un 12% y un 19% del total consumido por el vehículo para recorrer las millas que debía hacer en su ruta. Esto quiere decir que en promedio los vehículos gastan entre un 12% y un 19% más de combustible de lo que debían haber utilizado. Para este cálculo se contemplaron rangos de Fuel Efficiency de 5.5 MPG y 8.5 MPG que serían rangos razonables en operaciones del tipo de Long-Haul.

Solamente para estos 54 vehículos (recordemos que estamos tomando el 15% del total del Dataset cuyas mediciones fueron mayores), se consumió en Ralentí una cantidad de 18300 Gallons en una semana. Tomando un precio aproximado de 4 USD/Gallon del diesel para Agosto de 2022, podemos decir que el Ralentí le costó a la empresa un valor promedio de USD 15 por día de operación por cada vehículo:



**Tabla 11**
*Costo anual por combustible consumido durante tiempo de Ralentí*

Top 15% Vehicles	Gallons Idle (hs)	Operation Days	Gallon per Day	USD Total Cost per Day	Yearly Cost
1	2084	299	6,97	\$27,88	\$8.337,00
2	1942	299	6,50	\$25,99	\$7.769,60
3	1221	241	5,07	\$20,27	\$4.884,49
4	1760	256	6,87	\$27,49	\$7.038,03
5	2266	277	8,18	\$32,72	\$9.062,37
6	1616	278	5,81	\$23,25	\$6.463,27
8	157	21	7,48	\$29,93	\$628,48
12	142	50	2,84	\$11,35	\$567,48
15	29	29	1,00	\$3,98	\$115,52
21	52	53	0,98	\$3,93	\$208,26
22	202	55	3,66	\$14,66	\$806,05
28	37	51	0,73	\$2,90	\$147,98

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity (EPA-MOVES, 2021)

Este cálculo lo pudimos realizar, dado que en la tabla Instrumented Times, conocemos por vehículo la cantidad de días de medición del Data Logger, y de allí se encuentra la cantidad de días de operación por vehículo. Esto representa para una flota solamente de 54 vehículos un costo anual promedio de USD 1355 anuales por vehículo, o un total de USD 73202 anuales.

Lo interesante, es poder entender qué peso tiene el costo del Ralentí frente al gasto total de combustible de la flota. Tomando el rango de Fuel Efficiency antes mencionado para las rutas realizadas por estos camiones, para el mismo costo de Combustible obtenemos como resultado que el costo promedio anual en combustible para que cada vehículo realice las rutas se encuentra entre USD 7856 y USD 12141, llegando en el mejor de los casos de rendimiento a un costo anual en combustible de cerca de USD 430.000. Esto quiere decir que el consumo de Ralentí representó entre un 11 y 17% el costo de combustible anual. Pensando que un sistema de Telematics, que puede tener un costo de entre USD 20 y USD 30 por mes por vehículo (que se acerca al costo diario por Ralentí de cada vehículo), no solo nos facilita la gestión de Data Gathering, sino que también nos podría conseguir detectar sobrecostos en Ralentí y sobrecostos en bajos niveles de Fuel Efficiency por vehículo, nos podría colaborar con la determinación de estrategias que se repaguen al cabo de menos de 1 año.

**Tabla 12**

*Peso total del consumo durante Ralentí vs consumo por millaje recorrido*

Top 15% Vehicles	Max Gallons (5.5 MPG)	Min Gallons (8.5 MPG)	Max USD Cost	Min USD Cost	% Idle Fuel vs Mile Fuel
1	2272	1470	\$9.088	\$5.881	0,92
2	2910	1883	\$11.639	\$7.531	0,67
3	2050	1326	\$8.199	\$5.305	0,60
4	2783	1801	\$11.132	\$7.203	0,63
5	3342	2162	\$13.366	\$8.649	0,68
6	2236	1447	\$8.946	\$5.788	0,72
8	2466	1596	\$9.865	\$6.383	0,06
12	2195	1420	\$8.778	\$5.680	0,06
15	3020	1954	\$12.079	\$7.816	0,01
21	3336	2159	\$13.344	\$8.634	0,02
22	2905	1879	\$11.618	\$7.518	0,07
28	3611	2337	\$14.445	\$9.346	0,01

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity\* (EPA-MOVES, 2021)

Es notorio ver la ventaja de poder invertir en sistemas de Telematics, que contemplando estos sobrecostos de Ralentí, se repagan en períodos completamente accesibles para poder medir el Ralentí y Fuel Efficiency trabajar con los drivers de los vehículos con mayor consumo de combustible en reducir esos costos y generar eficiencias económicas en la flota.

Volviendo a nuestro foco de la temática Ambiental, también es fácil notar las ventajas ambientales en cuanto a la medición y reporting de la Huella de carbono. Ahora que conocemos el valor de Gallons de combustible por Ralentí consumidos, fácilmente poder multiplicar por el factor de emisión utilizado anteriormente para obtener las emisiones en CO2:

**Tabla 13**

*Emisiones de CO2 generadas durante Ralentí*

Gallons by Idle (hs)	% Idle Gallons vs Gallons per Drive (8.5 MPG)	% Idle Gallons vs Gallons per Drive (5.5 MPG)	Emissions by Idle (KGC02)	% Emissions Idle vs Total Max	% Emissions Idle vs Total Min	% Emissions Idle vs Mile
2084	1,42	0,92	21113	0,92	1,42	0,99
1942	1,03	0,67	19677	0,67	1,03	0,72
1221	0,92	0,60	12370	0,60	0,92	0,64
1760	0,98	0,63	17824	0,63	0,98	0,68
2266	1,05	0,68	22950	0,68	1,05	0,73
1616	1,12	0,72	16368	0,72	1,12	0,78
157	0,10	0,06	1592	0,06	0,10	0,07
142	0,10	0,06	1437	0,06	0,10	0,07
29	0,01	0,01	293	0,01	0,01	0,01
52	0,02	0,02	527	0,02	0,02	0,02
202	0,11	0,07	2041	0,07	0,11	0,07
37	0,02	0,01	375	0,01	0,02	0,01
223	0,12	0,07	2260	0,07	0,12	0,08
44	0,02	0,01	445	0,01	0,02	0,01

Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity\* (EPA-MOVES, 2021)

En las últimas columnas podemos observar que el porcentaje de las emisiones de CO2 por Ralentí en comparación con las emisiones calculadas por las millas recorridas en promedio representa casi un 13%. Por tanto, de no medir el Ralentí en términos ambientales, estaríamos perdiendo una gran parte de la medición de emisiones, de la cuál no podríamos definir una estrategia para mitigar.

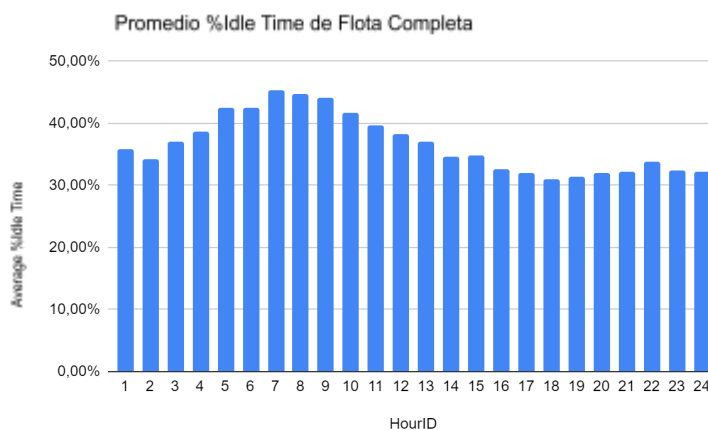
En la tabla se observa que si bien el promedio es 13%, casi el 10% de los vehículos tienen valores de emisiones por Ralentí superiores al 70% que con respecto a las emisiones por Millaje y esto es debido a la gran cantidad de tiempo operativo ocupado solamente en Ralentí (pensemos que los camiones de Long-Haul pueden sufrir enormes esperar para aguardar cargas). La posibilidad de medir el consumo de combustible por nivel de tanque, también nos puede facilitar la detección de bajos rendimientos de Fuel Efficiency para vehículos con mayor capacidad de rendimiento, y también contribuir en una reducción de emisiones (así como de costo en combustible).

Pensemos que de mantener mediciones manuales solamente sobre distancia recorrida, no solo estamos ante un enorme desafío de Data Gathering (baja escalabilidad, baja calidad de datos, posibilidad de errores, dependencia absoluta del proceso manual humano, baja posibilidad de confluir los datos a bases de datos de manera de procesarlos, etc), sino también ante la invisibilidad de costos económicos y ambientales ocultos por las operaciones de Ralentí y de niveles bajos de Fuel Efficiency.

Además, la posibilidad de tener datos real-time, nos permite detectar por hora los momentos en el día en el cuál el tiempo de Ralentí es mayor, lo cual permite que podamos trabajar por vehículo y por driver específicamente en las zonas y tiempos donde estén consumiendo más combustible por Ralentí y definir estrategias para mitigarlas (cambios de ruta, modificación de rutinas del driver, etc)

**Figura 9**

*Promedio de porcentaje de tiempo de Ralentí para el total de la flota de vehículos*



Fuente: Análisis basado en la base de datos de 2<sup>na</sup> Heavy-Duty Vehicle Activity" (EPA-MOVES, 2021)

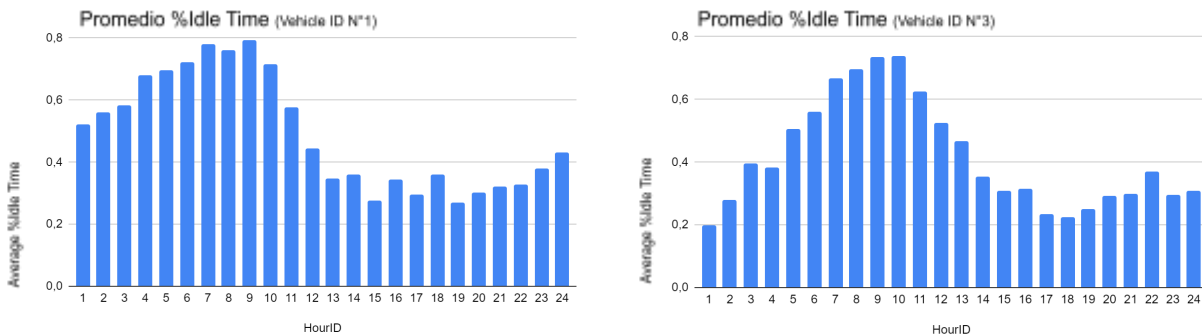
En el gráfico anterior se detecta un caso agregado entre todos los vehículos, el cuál nos indica que si bien hay un gran porcentaje base de Ralentí del 30% que a lo largo de todos los trayectos se debería buscar mitigar, hay momentos en la mañana donde el mismo es mayor.

Podemos tomar el ejemplo de los siguientes dos vehículos (ID 1 y 3) con los cuales se puede observar patrones cercanos de comportamiento en cuanto al Ralentí, llegando al mediodía. (Un ejemplo puede

ser dejar prendido el camión en las paradas para descansar, comer ó ir por una ruta muy congestionada durante la mañana. Otro punto interesante es que vemos que ambos vehículos mantuvieron constantemente el motor encendido 24hs, lo cuál implica que no se haya apagado en ninguna parada ni lo que se conoce como "housing" que es cuando el vehículo se detiene para continuar al día siguiente:

**Figura 10**

*Promedios de tiempos de Ralentí para vehículos con similares características*

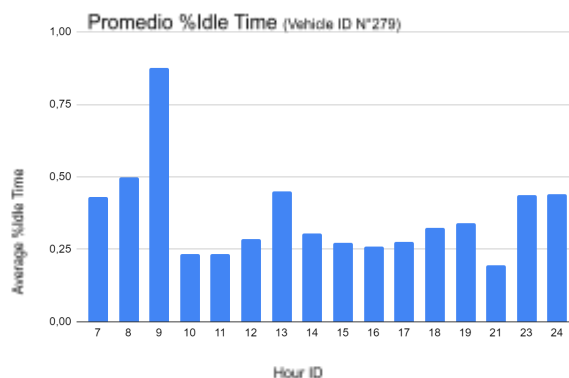


Fuente: Análisis basado en la base de datos de 2<sup>da</sup> Heavy-Duty Vehicle Activity\* (EPA-MOVES, 2021)

Si contemplamos el caso del Vehículo con ID 279, podemos observar un patrón distinto. En primer lugar observamos que el motor del vehículo no se mantiene encendido 24hs al día, lo cual implica que en los periodos de Housing se haya apagado. Si contemplamos las millas recorridas en promedio, observamos que los tres vehículos tienen un promedio similar entre 245 y 330 millas, por lo que si bien las rutas podrían ser disímiles, deberíamos poder esperar patrones similares (como el apagado del motor durante "Housing"). En este caso notamos dos subas del % de Tiempo de Ralentí en horarios que nos dicen mucho del comportamiento del driver, a media mañana (primera parada del día), a medio día (horario de almuerzo) y a fin de día, probable cena.

**Figura 11**

*Promedios de tiempos de Ralentí para vehículo con patrón diferente a la media*



Fuente: Análisis basado en la base de datos de 2<sup>da</sup> Heavy-Duty Vehicle Activity\* (EPA-MOVES, 2021)

Aquí podemos notar que con respecto al vehículo 1, 3 y aquellos que tengan comportamientos similares, podemos compartir los buenos rendimientos de otros vehículos comparativamente que durante el

periodo de Housing no mantienen el motor encendido. Por otro lado, podemos trabajar en los tiempos de descanso y paradas, para que el motor se apague, entendiendo que esto genera gastos de combustible innecesarios.

Poder medir estos casos representativos es de suma importancia, tanto para reportar y medir correctamente emisiones, como para tomar estrategias con los drivers de dichos vehículos en pos de reducir los consumos de Ralentí que derivarán netamente en beneficios económicos como ambientales.

Queda claro que las ventajas de los sistemas de Telematics son muchas:

- 1) En términos de **Data Gathering**:
  - a) Fácil Escalabilidad a nuevas y diversas flotas de vehículos
  - b) Fácil gestión de datos
  - c) Reducción de errores humanos y datos faltantes
  - d) Facilidad en la adquisición de datos via API o sistemas de descarga automática
  - e) Facilidad en el post-procesamiento de datos
  - f) Posibilidad de confluir los datos a Bases de Datos para procesamiento y warehousing
- 2) En términos **Ambientales**:
  - a) Facilitar la adquisición de datos
  - b) Asegurar datos precisos y certeros
  - c) Obtener datos que de manera manual son imprácticos o imposibles (Tiempos de Ralentí, Niveles de Tanque, Consumo de Combustible, Millaje exacto)
  - d) Poder afinar la medición de la Huella de Carbono
  - e) Definir estrategias correctas para mitigar y reducir correctamente los impactos ambientales de la flota
- 3) En términos **Económicos**:
  - a) Posibilidad de reducir costos por Ralentí
  - b) Posibilidad de reducir costos por Fuel Efficiency
  - c) Realizar una inversión baja con rápido periodo de repago y de ahorro a corto plazo

En párrafos anteriores describimos los pilares de ESG que las empresas están abordando, debido a la necesidad por reportar emisiones y comunicar estrategias efectivas de reducción, ya sea por exigencia de clientes, normativas vigentes o inversionistas. El poder medir el Ralentí y Consumo de Combustible, abre las puertas a un abanico de posibilidades mayor que el de solo medir distancia, dado las estrategias que podríamos tomar para mitigación junto con sus resultados. Si la empresa se encuentra en la necesidad de reducir emisiones de logística, ya sea por exigencia de inversores o por nuevas normativas que colocan precio al CO2 emitido (Hoy en día 27 países ya han incluido impuestos al carbono llamado "Carbon Tax" en ciertas ciudades), el poder llegar a datos de Ralentí, Fuel Efficiency, estilos de manejo de los drivers, permiten tomar estrategias de mitigación, como:

- Capacitación a los Drivers sobre apagado del motor en paradas y tiempos de descanso. (En Nueva York está prohibido mantener el motor en Ralentí por más de 3min con el vehículo estacionado o parado)
- Evaluación de Fuel Efficiency por tipo de vehículo, buscando capitalizar las eficiencias de de drivers que han logrado mayor MPG para vehículos similares y estandarizar way-of-workings para dichos tipos de flotas.

- Rediseño de rutas en función de entender si hay drivers que siempre mantienen demoras por encarar ciertos tipos de rutas. (Con mayor tránsito, en reparación, con baja facilidad de manejo, etc)
- Entender hábitos de manejo y estandarizar buenas prácticas para eficientizar la conducción.

### **5.1.3) Smart Metering - Medición de consumo de energía en línea**

Como mencionamos en los apartados anteriores, uno de los principales inputs para medición y cálculo del Scope 2 de la Huella de Carbono es el consumo de energía.

Dado que la gran mayoría de las empresas compran esta energía a la distribuidora local, ó en algunos casos a los generadores directos de la energía vía contratos del tipo PPA (Power-Purchase Agreement), y en contrapartida dichas empresas no generan su propia energía, se considera que dentro de los alcances de la Huella de Carbono, dicho consumo se debe representar en el Scope 2 y no en el Scope 1.

En el mundo actual, toda empresa que desarrolle y/o comercialice algún bien o servicio consume energía, sea para alimentar una línea de producción como para cargar una computadora de sus empleados. Si bien dependiendo del rubro y la intensidad energética, este valor de la Huella generado debido a los consumos energéticos puede tomar distintas magnitudes, podemos garantizar que es un foco importante para prácticamente cualquier empresa que quiera medir y reportar su huella de Carbono.

Si pensamos en una pequeña empresa con una sola oficina/planta productiva, el desafío no representaría mucha complejidad, dado que se podrían obtener los datos de consumo energético mensual (medidos en Kwh o Mwh) en las facturas de electricidad enviadas por la distribuidora.

A priori, en términos de Data Gathering, no parecería problema alguno. Sin embargo, pensemos que la mayoría de las empresas que buscan reportar sus emisiones e impactos ambientales, son grandes empresas con más de un sitio (ya sean oficinas, plantas de producción, centros de distribución, etc). Rastrear esta cantidad de facturas puede representar un gran trabajo manual en términos de adquisición de todos esos datos a tiempo. Pensemos también que muchas empresas cuentan con fábricas o centros de distribución en parques industriales que no cuentan con fronteras comerciales (equipos de medición), propios para cada edificio, por tanto esto requiere un doble trabajo de que un externo (dueño del parque) comparta dicha información, cuya precisión será baja dado que al no contar con equipos de medición individuales en la mayoría de los parques industriales, estos consumos serán estimados y por tanto la Huella será estimada.

Otro punto en términos de calidad de datos, es que al igual que en el caso anterior de obtener los kilometrajes de manera manual de la logística empresarial, en el caso de obtener los consumos energéticos de las facturas de electricidad, corremos el riesgo de que al procesar estos datos se cometan errores dado que el input de dichas facturas es manual. Por tanto, al tratarse de un proceso manual, corremos el riesgo de tergiversar datos por error humano. Si bien existen softwares o sistemas sencillos que pueden “escanear” datos de archivos tipo PDF, necesitaremos mandatoriamente una persona que supervise que este escaneo fue realizado correctamente debido a las propia eficacia de dichos softwares.

Por tanto, vemos como el desafío de data gathering y calidad de la información comienza a asomarse.

Adelantemonos unos pasos en el proceso de desarrollo de una estrategia ambiental; supongamos que recibimos está gran cantidad de datos en tiempo y forma, que tenemos la capacidad ociosa para procesarla (descontando el aspecto económico) y que además contamos con personal idóneo cuya tasa de errores es nula... en este caso, seríamos capaces de poder medir la Huella de Carbono de nuestro consumo energético (multiplicando por los factores de emisión correspondientes, como recapitulamos en los apartados anteriores), sin embargo la gran pregunta es: ¿qué estrategia ambiental podemos

tomar para reducir la huella de nuestros consumos energéticos solo conociendo el valor total de los mismos?

Existirían dos caminos. El primero sería reducir el factor de emisión, para lo cuál sería necesario aumentar el consumo de energías renovables, que puede ser onsite (instalando paneles solares en las terrazas de nuestros sitios por ejemplo), ó puede ser vía off-site en un contrato tipo PPA donde acordamos con un generador de energías renovables la compra de dicha energía para abastecernos con energía limpia.

Recordemos que las emisiones de CO2 a partir del consumo de energía, se calculan de la siguiente manera, donde los factores de emisión por país se pueden obtener de fuentes reconocidas como la IEA (International Energy Agency):

$$CO2 \text{ Emissions per Kwh}_{\text{Energy Consumption}} = EC [Kwh] * EF \frac{g \text{ CO2}}{Kwh}$$

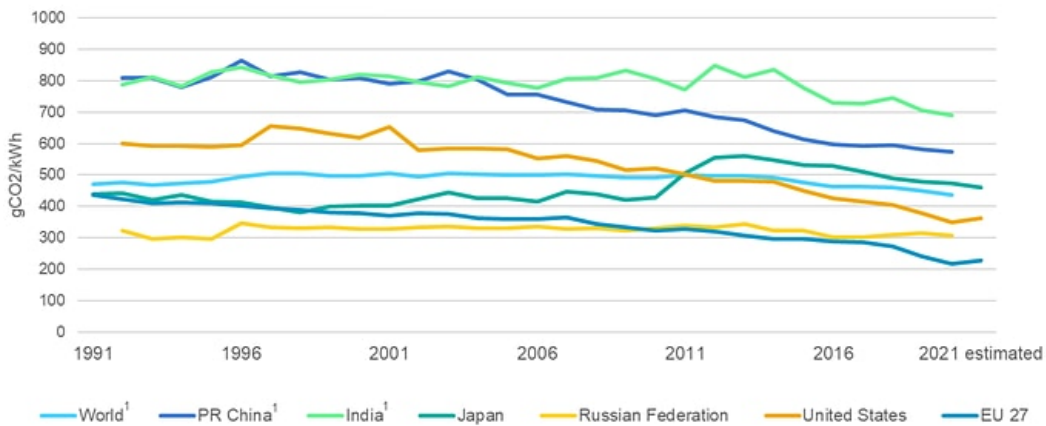
*EC [Kwh]: Energy Consumption*

*EF [ $\frac{g \text{ CO2}}{Kwh}$ ]: Emission factor for Country's energy matrix*

El valor del factor emisión será menor, en países cuyo porcentaje de energías renovables sea mayor en su matriz energética y será mayor en países con mayor porcentaje de generación fósil, como se puede observar en el siguiente gráfico.

**Figura 12**

*Factores de emisión por País según IEA*



<sup>1</sup>Data not available for 2021

Source: <https://energy.northern.com/2022/09/26/iea-releases-emissions-factors-and-greenhouse-gas-emissions-from-energy-databases/>

Como podemos imaginar, lejos está de la capacidad de acción de una empresa que la matriz energética de los países donde opera aumente su capacidad instalada en energías renovables. Por tanto, la opción es abastecerse de manera privada con dichas energías limpias via instalación o contratos privados, como ya mencionamos.

Sin embargo, esto tiene sus limitaciones. En primer lugar, muchos países tienen reglas estrictas (y en muchos casos, complejas de cumplir) para migrar sus consumos a contratos privados, dado que se busca priorizar la estabilidad de la red. En Chile, por ejemplo, para poder salir del mercado regulado y abastecerse de manera privada con un generador, es necesario que la empresa tenga una capacidad

instalada mínima de 500 kw; en Colombia se necesita un consumo mensual de al menos 55.000 Kwh para salir al mercado no regulado.

Como observamos, estos montos altos de consumo energético implica que no todas las empresas puedan migrar fácilmente al mercado no regulado y esto se pueda volver una limitante para hacerlo. La otra dificultad, es que si bien existe una creciente demanda de energías renovables por el sector público y privado, al día de hoy la oferta disponible del mercado no es suficiente para abastecer las necesidades de todos los privados que quisieran migrar sus operaciones.

Por otro lado, en caso de cumplir con las condiciones y que dicha oferta existiera, dichos contratos del tipo PPA se constituyen por al menos un periodo de 5 años. Esto significa, que la empresa tiene que poder hacer proyecciones de consumo energético hacia adelante con suficiente predicción para no cerrar un contrato por un monto de energía menor al necesario o excesivo y pagar un gran sobrecosto. Para ello la cantidad de datos que necesitamos en pos de realizar esta proyección futura con suficiente precisión es mucho mayor que solamente un consumo mensual (que en algunos casos puede ser estimado, como mencionamos).

El segundo camino para buscar reducir las emisiones, sería reducir el consumo eléctrico (o sea el primer término de la ecuación). Pero si necesitamos desarrollar una estrategia de eficiencia energética en nuestra empresa, en pos de reducir los impactos ambientales y generar beneficios económicos, no podemos hacerlo conociendo solamente los consumos globales (y en algunos casos poco precisos) mensuales de la empresa. Es necesario poder tener segmentaciones, por tipos de edificios, por máquinas, por sectores, según los consumos para entender dónde podemos generar prácticas sustentables.

Todas estas características anteriormente mencionadas, nos llevan a la necesidad de contar con mediciones de energía precisas que nos permitan:

1. Facilitar la adquisición de datos de consumo energético
2. Asegurar la calidad y precisión de los datos de consumo
3. Facilitar el almacenamiento y procesamiento posterior de los datos.
4. Tener una granularidad de información suficiente par poder realizar análisis de datos
5. Permitir la segmentación de consumos por sectores, máquinas, turnos de producción, días, horarios, etc.
6. Sentar las bases para generar predicciones de consumo de manera de poder migrar a consumos de energías renovables.

La tecnología de **Smart Metering** podría ser la solución a todos estos desafíos.

Analizaremos entonces el Dataset *"Energy Consumption"*, donde tenemos un registro de datos de Consumo Energético (en Kwh) de distintas áreas de una empresa (Logística, Oficinas, Fábrica, etc), donde además tenemos registro de datos interesantes cómo los horarios de los consumos energéticos y los datos de áreas de cada sitio, entre otros. A continuación observamos un detalle de algunos datos de dicha base:

**Tabla 14**

*Resumen de DataSet "Energy Consumption"*



```

> head(Energy_consumption)
  building_id  timestamp meter_reading primary_use square_feet year_built air_temperature
1           1 2016-01-01 00:00:00           0 Logistics      2720         2004             25
2           2 2016-01-01 00:00:00           0 Logistics      5376         1991             25
3           3 2016-01-01 00:00:00           0 Logistics     23685         2002             25
4           4 2016-01-01 00:00:00           0 Logistics    116607         1975             25
5           5 2016-01-01 00:00:00           0 Logistics      8000         2000             25
6           9 2016-01-01 00:00:00           0 Office       27000         2010             25

  cloud_coverage dew_temperature precip_depth_1_hr sea_level_pressure wind_direction wind_speed
1              6              20              NA              1019.7              0              0
2              6              20              NA              1019.7              0              0
3              6              20              NA              1019.7              0              0
4              6              20              NA              1019.7              0              0
5              6              20              NA              1019.7              0              0
6              6              20              NA              1019.7              0              0
  
```

Fuente: Análisis basado en la base de datos de <sup>31</sup>“Great energy predictor” (ASHRAE, 2020)

Aquí podemos empezar a observar, que la facilidad de obtención de datos es completamente sencilla, dado que no necesitamos una interpretación de facturas, ni desafíos de gata gathering manuales en términos de obtención de facturas de electricidad, ni recaer en cálculos poco precisos en casos de sitios donde diferentes empresas compartan los mismos medidores y no tengan frontera comercial independiente.

En términos de Data Gathering, observamos que con mucha facilidad podemos observar si un sitio cumple o no con los requisitos de consumo para migración a Energías Renovables de distintos países.

La otra gran diferencia con respecto a solo medir una variable global medida de manera periódica como el consumo total de un mes, es que no podemos segregar ni entender los consumos por hora, día o siquiera mes. Al tener mediciones continuas en tiempo real, podemos analizar los datos de consumo en sectores y tiempos más específicos, que nos permitan obtener métricas que nos pueden contribuir con desarrollar estrategias de Eficiencia Energética.

Para ello, comenzaremos con el procesamiento de datos de la variable “*timestamp*” la cuál nos indica el año, mes, día y horario exacto de la medición de consumo energético tomada por los medidores. Este sistema nos permite obtener hasta el minuto y segundo exacto en el cual estas mediciones fueron tomadas, pero a fines de los análisis que buscamos realizar extraeremos Año, Mes, Día y Hora, dejando los minutos y segundos de lado:

### Tabla 15

*Desglose de TimeStamp para los consumos de energía eléctrica*

```

> Energy_consumption$hour_reading <- format(as.POSIXct(Energy_consumption$timestamp),format = '%H')
> tail(Energy_consumption)
  building_id      timestamp meter_reading primary_use square_feet year_built air_temperature
698670      94 2016-12-31 23:00:00    51.1920  Logistics      29469      1968      20.6
698671      96 2016-12-31 23:00:00    414.1430   Factory      200933      2006      20.6
698672      97 2016-12-31 23:00:00    420.6280   Factory      187647      2007      20.6
698673      98 2016-12-31 23:00:00    409.3650   Factory      200933      2007      20.6
698674      99 2016-12-31 23:00:00     8.7504   Office        2260      2016      20.6
698675     100 2016-12-31 23:00:00    201.4920   Factory      24456      1968      20.6
  cloud_coverage dew_temperature precip_depth_1_hr sea_level_pressure wind_direction wind_speed
698670          NA              10              0              1021.1              0              0
698671          NA              10              0              1021.1              0              0
698672          NA              10              0              1021.1              0              0
698673          NA              10              0              1021.1              0              0
698674          NA              10              0              1021.1              0              0
698675          NA              10              0              1021.1              0              0
  year_reading month_reading day_reading hour_reading
698670      2016           12           31           23
698671      2016           12           31           23
698672      2016           12           31           23
698673      2016           12           31           23
698674      2016           12           31           23
698675      2016           12           31           23
  
```

Fuente: Análisis basado en la base de datos de <sup>31</sup>Great energy predictor" (ASHRAE, 2020)

Esto nos permitirá segregar información por horarios. Pensemos en una empresa que tiene fábricas esparcidas por distintas regiones del globo y quisiera entender cuáles están trabajando de manera más eficiente en cuanto al consumo de energía. Podemos fácilmente obtener esta información analizando datos de diversos edificios con las mismas características, para los mismos horarios/meses con similares características.

Pero antes de ello, vamos a crear otra variable adicional de mucha utilidad para el análisis energético de una compañía, la "Intensidad energética". La Intensidad energética (medida en Kwh/m2), es una unidad de medida muy útil para entender cuánto estamos consumiendo por m2 de cierta operación (Ej: fábrica, centro de distribución, etc), y para sitios similares nos permite entender cuáles son más eficientes, entender por qué y buscar desarrollar estrategias de eficiencia energética. Además, veremos más adelante cómo este indicador nos permite proyectar a futuro los consumos de energía, pensando en migraciones a Energías Renovables, donde los plazos de contrato son de al menos 5 años:

## Tabla 16

*Cálculo de intensidad energética*

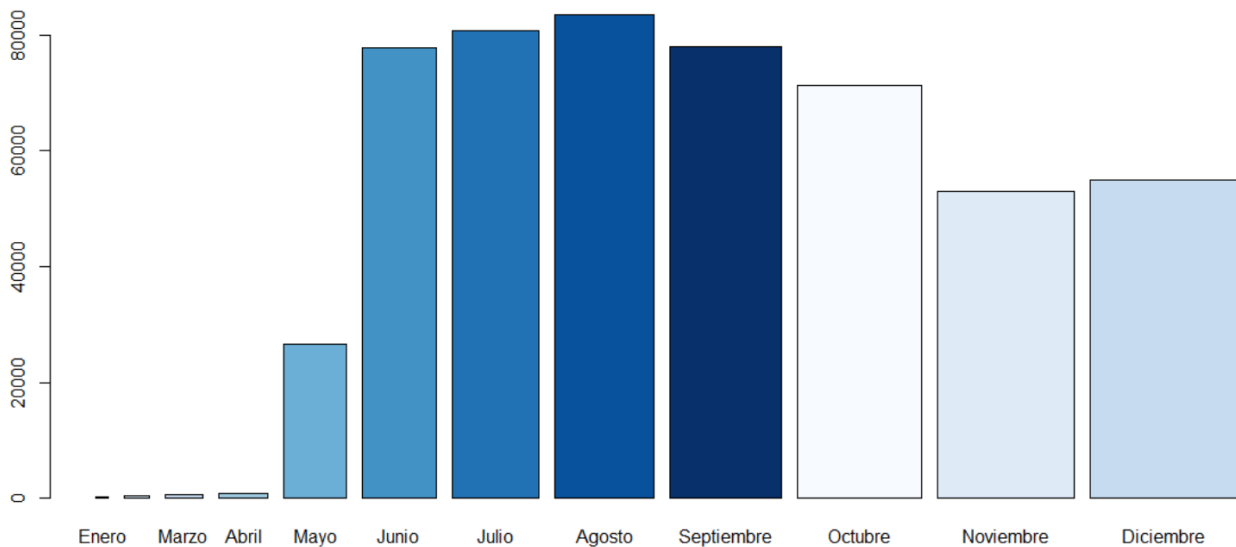
```
> # Creamos la variable de "Intensidad Energetica"
> Energy_consumption$energy_intensity <- Energy_consumption$meter_reading/Energy_consumption$square_feet
> tail(Energy_consumption)
  building_id      timestamp meter_reading primary_use square_feet year_built air_temperature
698670      94 2016-12-31 23:00:00    51.1920 Logistics      29469      1968          20.6
698671      96 2016-12-31 23:00:00   414.1430  Factory      200933      2006          20.6
698672      97 2016-12-31 23:00:00   420.6280  Factory      187647      2007          20.6
698673      98 2016-12-31 23:00:00   409.3650  Factory      200933      2007          20.6
698674      99 2016-12-31 23:00:00     8.7504  Office        2260      2016          20.6
698675     100 2016-12-31 23:00:00   201.4920  Factory      24456      1968          20.6
  cloud_coverage dew_temperature precip_depth_1_hr sea_level_pressure wind_direction wind_speed
698670          NA              10              0              1021.1              0              0
698671          NA              10              0              1021.1              0              0
698672          NA              10              0              1021.1              0              0
698673          NA              10              0              1021.1              0              0
698674          NA              10              0              1021.1              0              0
698675          NA              10              0              1021.1              0              0
  year_reading month_reading day_reading hour_reading energy_intensity
698670      2016           12           31           23      0.001737148
698671      2016           12           31           23      0.002061100
698672      2016           12           31           23      0.002241592
698673      2016           12           31           23      0.002037321
698674      2016           12           31           23      0.003871858
698675      2016           12           31           23      0.008238960
```

Fuente: Análisis basado en la base de datos de <sup>31</sup>Great energy predictor" (ASHRAE, 2020)

En este caso, dado que el Área de los sectores, está medido en pies cuadrados, la unidad de la Intensidad Energética será medida en Kwh/ft2. Tomemos como ejemplo uno de los edificios, donde fácilmente podemos visibilizar el consumo energético de un año de operación entendiendo el consumo por mes:

**Figura 13**

*Consumo energético de un año de operación de Oficinas*

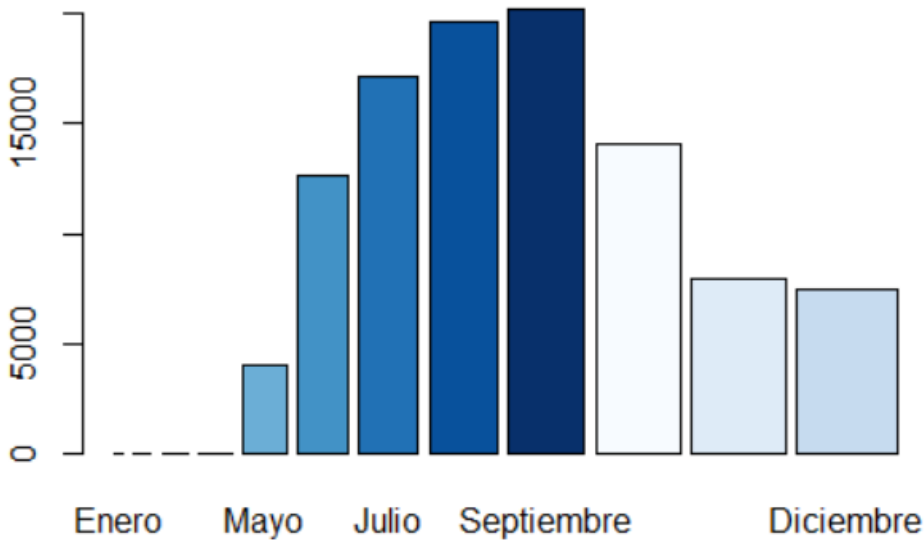


Fuente: Análisis basado en la base de datos de <sup>31</sup>Great energy predictor" (ASHRAE, 2020)

Podemos comparar con otro building\_id de las mismas características (Logistics) en este caso, para ver cómo se comporta el consumo de energía de manera anual. Podemos observar como las curvas de consumo se asemejan notablemente en cuanto a forma, siendo los meses de mayor consumo los meses entre Julio y Septiembre. Sin embargo, notamos que los valores de consumo energético son ampliamente diferentes, por lo cuál es interesante analizar a fondo cuáles pueden ser las diferencias entre ambos.

**Figura 14**

Consumo energético de un año de operación para el edificio de Logística



Fuente: Análisis basado en la base de datos de <sup>51</sup>"Great energy predictor" (ASHRAE, 2020)

A priori, uno de los datos más llamativos, es que el segundo edificio tiene una antigüedad de casi 15 años más que el primero y además tiene una superficie de casi la mitad. Sin embargo los consumos del primero son más altos.

Uno de los sistemas de mayor consumo de energía en este tipo de sitios, suelen ser los sistemas HVAC (Calefacción y Aire Acondicionado). Los datos obtenidos del sistema de Smart Metering, nos dan la posibilidad de realizar un análisis con respecto a las temperaturas de ambos edificios para entender si puede ser un factor influyente.

**Tabla 17**

Comparación de temperaturas para dos Centros de Distribución

```
> #Comparacion de temperaturas en los últimos meses del año para ambos Centros de distribucion
> mean(Energy_consumption$air_temperature[Energy_consumption$building_id == 1 & Energy_consumption$month_
reading ==12])
[1] 19.92393
> mean(Energy_consumption$air_temperature[Energy_consumption$building_id == 2 & Energy_consumption$month_
reading ==12])
[1] 19.95825
> max(Energy_consumption$air_temperature[Energy_consumption$building_id == 1 & Energy_consumption$month_r
eading ==12])
[1] 30.6
> max(Energy_consumption$air_temperature[Energy_consumption$building_id == 2 & Energy_consumption$month_r
eading ==12])
[1] 30.6
```

Fuente: Análisis basado en la base de datos de <sup>51</sup>"Great energy predictor" (ASHRAE, 2020)

Como podemos observar, las temperaturas medias y máximas de ambos edificios son prácticamente idénticas, todo indicaría que estarían en lugares muy cercanos, por tanto el uso de HVAC debería ser similar. Ahora si analizamos un mes con mayor consumo energético para ambos edificios:

**Tabla 18**

Comparación de temperaturas para dos Centros de Distribución en un mes de mayor consumo energético

```

> mean(Energy_consumption$air_temperature[Energy_consumption$building_id == 1 & Energy_
consumption$month_reading == '08'])
[1] 27.58648
> mean(Energy_consumption$air_temperature[Energy_consumption$building_id == 2 & Energy_
consumption$month_reading == '08'])
[1] 27.6592
> max(Energy_consumption$air_temperature[Energy_consumption$building_id == 1 & Energy_c
onsumption$month_reading == '08'])
[1] 35.6
> max(Energy_consumption$air_temperature[Energy_consumption$building_id == 2 & Energy_c
onsumption$month_reading == '08'])
[1] 35.6
  
```

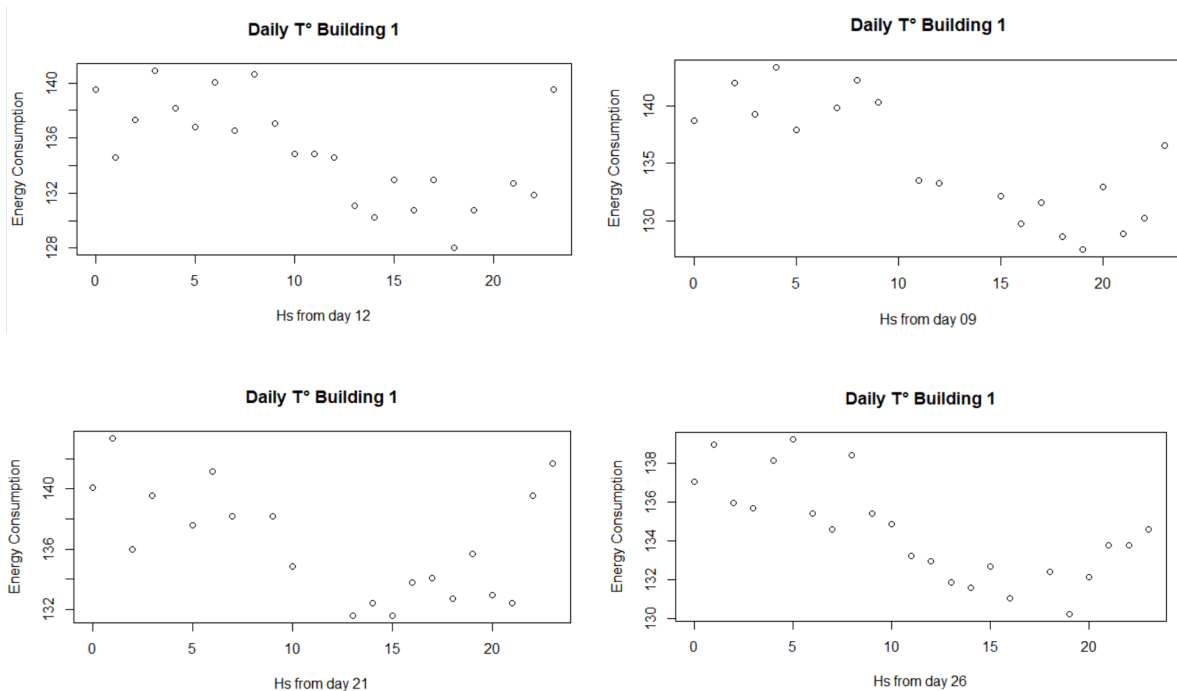
Fuente: Análisis basado en la base de datos de <sup>31</sup>"Great energy predictor" (ASHRAE, 2020)

En este caso podemos observar como para el mes de Agosto, en donde ambos tuvieron fuertes aumentos en consumos de energía, también aumentaron de manera significativa los valores promedios y máximos de temperaturas durante dicho mes. Todo esto indicaría que los sistemas de HVAC están siendo utilizados de manera más intensiva.

Si analizamos el consumo de energía diario, podemos observar patrones de consumo en cuanto la curva de energía. Principalmente observamos como durante la madrugada y mañana el consumo de energía aumenta exponencialmente y se mantiene constantemente alto hasta la tarde donde coincide que la gente suele dejar de trabajar:

**Figura 15**

*Patrones de consumo energético en función de hora de operación*



Fuente: Análisis basado en la base de datos de <sup>31</sup>"Great energy predictor" (ASHRAE, 2020)

Muchos equipamientos están preparados para arrancar al comienzo del día, (esto es debido a que están conectados en línea con los tableros principales) es por eso que notamos que los consumos se disparan a las 0hs de cada día. Un punto interesante para trabajar, es en el sistemas de SCADA que contribuyan con encender los equipos según su necesidad. Sería interesante que los equipos internos de esta

compañía, puedan indagar sobre qué consumos se están disparando en esos horarios y si realmente están contribuyendo con la logística de la compañía.

Por ejemplo, para el caso del aire acondicionado y calefacción de los edificios, pensando que la jornada laboral comienza entre las 6am y 9am dependiendo del Building type (Centro de distribución, Producción por lo general con horarios cercanos a 6am y las Oficinas entre 8am y 9am), podemos dejar programados los sistemas HVAC para que arranquen solo unas horas antes de la llegada del personal y no desde las 0hs, dado que esto genera un gran consumo de energía innecesario. Con encender apenas unas 2hs, los equipos de HVAC más modernos (e incluso algunos ya antiguos), tienen capacidad de poder enfriar o calentar rápidamente los sitios aguardando a la llegada del personal. Es importante que estos datos se contrasten con los departamentos de Mantenimiento y Recursos Humanos, para encontrar eficiencias energéticas que traerán ahorros tanto ambientales como económicos.

Aquí podemos observar como Smart Metering es una herramienta que nos facilita el poder llegar a conclusiones más profundas que un simple detalle de consumo mensual de una factura de electricidad, donde no podríamos realizar estos análisis. El poder encontrar eficiencias en la gestión energética de una compañía implica los desafíos en términos de datos que ya mencionamos, por tanto utilizar herramientas que nos faciliten los procesos de data gathering y procesamiento son muy útiles para la gestión ambiental de una compañía.

El precio de estos equipos en el mercado ronda en los 5 USD, donde las empresas se encargan de su instalación, mantenimiento y la empresa contratante paga un fee mensual por la adquisición de dicha herramienta. Sin considerar que tenemos un equipo HVAC tipo Rooftop de un fabricante Tier 1 con variadores de frecuencia (lo cual hace que dichos equipos sean más eficientes), podemos considerar una potencia de dicho equipo de 20 kW.

Suponiendo que logramos que debido a esta gestión, los equipos de aire acondicionado se enciendan a las 2am en vez de las 0hs, lo cual nos generaría un ahorro de 40 Kwh/día. Suponiendo además que logramos cerrar un acuerdo PPA por Energía Renovable, y nuestro suministro ronda los 70 USD/MWH (precio medio del mercado Latinoamericano en 2021-2022)

0,07 USD/Kwh

$$HVAC \text{ Energy reduction}_{[Kwh/día]} = 40 \frac{Kwh}{Día}$$

$$Monthly \text{ HVAC Energy reduction}_{[Kwh/mes]} = 40 \frac{Kwh}{Día} * 30 \frac{Día}{Mes} = 1200 \frac{Kwh}{Mes}$$

$$PPA \text{ Price}_{[USD/Kwh]} = 0,07 \frac{USD}{Kwh}$$

$$Monthly \text{ HVAC Energy Cost reduction}_{[USD/mes]} = 1200 \frac{Kwh}{Mes} * 0,07 \frac{USD}{Kwh} = 84 \frac{USD}{Mes}$$

Como podemos observar, solo con reducir 2hs de utilización de HVAC podemos generar un ahorro estimado de 84 USD/Mes, lo cuál no solo repaga instantáneamente el sistema Smart Metering sino que además deja ahorros económicos en la empresa.

Esto sumado a que además de acopiar un ahorro económico, estamos reduciendo nuestro consumo eléctrico innecesario, que se traduce en reducción de nuestra Huella de Carbono. Estos es solo un ejemplo de las infinitas posibilidades que existen con los sistemas Smart Metering, si sensorizamos las maquinarias con consumos más importantes de nuestros sitios, podemos empezar a encontrar eficiencias en las mismas que se traducen en ahorros económicos y ambientales. En muchos casos, los análisis requerirán un estudio más profundo de la situación, con parámetros más complejos como uso de aire comprimido, consumos de agua en bombas hidráulicas, utilización de refrigeradores o aparatos de cocinas en las oficinas, entre otros; pero sin ninguna duda que comenzar a cuantificar estos datos nos

va a permitir generar ahorros, además de facilitar nuestra misión de medir los impactos ambientales de nuestras operaciones.

## **5.2) Herramientas implementadas en el proceso de Data Warehousing**

### **5.2.1) API/Web Scraping para la adquisición de factores de emisión**

Dentro de los análisis de Data Gathering anteriores, pudimos observar como la utilización de sistemas de Telematics que nos pueden reportar en tiempo real datos relevantes para el cálculo de las emisiones, en este caso por logística, de la empresa, pueden ser de mucha utilidad.

En el caso anterior, se han utilizado valores fijos de factores de emisión proporcionados por las tablas del WRI. Sin embargo, pensemos que a lo largo del tiempo los factores de emisión varían. Su frecuencia también varía entre ellos, por ejemplo para el caso de los factores de emisión para electricidad (TonCO<sub>2</sub>/Kwh) varían anualmente en función de cómo varíe su matriz energética, siendo reportados de manera anual por cada país a la Agencia Internacional de Energía (IEA), los cuales los publica en sus reportes. Si la empresa tiene una fábrica en España, para un año “x” el factor de emisión tendrá un cierto valor, y si España en el año “x+1” decide incorporar una gran cantidad de Parques Solares y Eólicos a su matriz, se verá reflejado que dicho factor de emisión habrá bajado en dicho año, afectando también el cálculo de las emisiones generadas en Scope 2 por “Purchased Energy”.

En otros casos, como el factor de emisión por transporte, encontramos actualizaciones más frecuentes como la realizada por el United Nations Framework Convention on Climate Change (UNFCCC), cuya bases de datos online se actualiza de manera cuatrimestral.

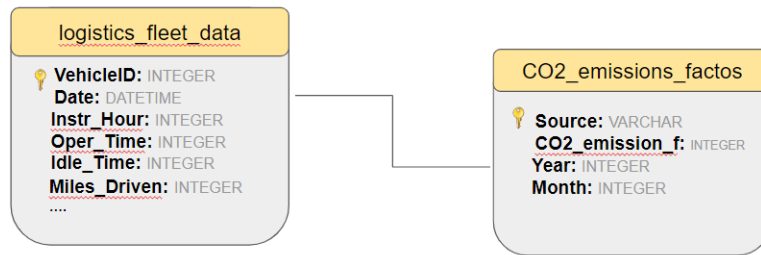
Por tanto es importante, que en la gestión ambiental de la compañía se consideren estos períodos de actualización de factores de emisión para asegurar que los datos sean los más reales y actualizados posibles.

En términos de Data Gathering, observamos que la automatización del proceso de adquisición de datos de consumo de combustible, millaje y otros datos relevantes de la logística puede facilitar de gran manera el trabajo de recolección de datos para facilitar el cálculo de la huella de carbono por la logística, entre otros indicadores.

Pensando en términos de Data Warehousing, sería interesante que pudiéramos generar dos tablas para el cálculo de huella de logística: una tabla con los datos obtenidos via telemática sobre la flota logística y otra tabla con los factores de emisión que se actualizara con las frecuencias antes mencionadas, de manera de quedar:

**Figura 16**

Creación de Tablas para el cálculo de huella de CO2 logística



El problema radica en que, la disponibilidad online de factores de emisión es escasa y en un formato no sencillo de poder conectar via API o Web Scrapping. Existen bases de datos online libres de acceso y gratuitas con datos de factores de emisión como el Greenhouse Gas Inventory Data de las Naciones Unidas (UNFCCC) y el Emission Factor DataBase del Panel Intergubernamental de Cambio Climático (IPCC), los cuales de manera online podemos exportar fácilmente a formato csv los factores de emisión actualizados sobre combustión, electricidad, entre otros.

**Figura 17**

Inventario de Gases de Efecto Invernadero de las Naciones Unidas

Fuente: <sup>11</sup>Greenhouse Gas Inventory Data, (UNFCCC,2022)

Sin embargo, notamos que no tienen una fácil conexión via API o Web Scrapping, pensando en que la frecuencia para la actualización de datos es baja (cuatrimestralmente o anualmente), vemos que el desarrollo de un script de código para intentar descargar estos datos puede volverse muy complejo y poco práctico. Para el caso de la obtención de millones de datos de consumos de combustible, millaje, consumo energético, entre otros, sí torna con sentido porque el manejo manual de estos datos es prácticamente imposible, de baja calidad y escalabilidad (además de los otros atributos que concluimos anteriormente).



Figura 18

Base de datos de IPCC sobre factores de emisión

EF ID	IPCC 1996	IPCC 2006	Gas	Fuel	Type of parameter	Description	Technologies / Practices	Parameters / Conditions	Region / Regional Conditions	Abatement / Control Technologies	Other properties	Value	Unit	Source of data	Action
02930	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			United States of America			2.0	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail
02940	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			Canada			1.81	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail
02941	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			Australia			1.20	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail
02942	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			New Zealand			1.33	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail
02943	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			United Kingdom of Great Britain and Northern Ireland			1.0	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail
02944	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			Ireland			0.85	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail
02945	SA - Solid Waste Disposal on Land	4.A - Solid Waste Disposal	METHANE		1000 IPCC default	Municipal Solid Waste (MSW) Generation Rate			Austria			0.92	kg/cap/day	Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories (Table 5-1 on Pages 6.5-6.7 of the Reference Manual)	Detail

Fuente: "Emission factor database (EFDB), IPCC

Si buscamos generar una base de datos propia con los factores de emisión de manera de facilitar el post-procesamiento de datos de Scope 1, 2 y 3 para el cálculo de huella, basta con realizar una descarga unitaria del histórico de datos de manera manual y actualizar de manera anual o cuatrimestral. En este sentido, no vemos que implique una carga de trabajo excesiva, lo que por el contrario sí podría implicar en caso de implementar un sistema API o Web Scrapping para tal estilo. Pensemos que el tiempo en desarrollar y mantener actualizada una API o Web Scrapping para descarga con frecuencia anual o cuatrimestral, prácticamente se torna innecesario y especialmente debido a que estos sistemas sufren de modificación, nuevos tipos de datos, gran cantidad de campos variables que generan que las queries fallen, por tanto para el caso de la adquisición de datos de factores de emisión no se recomienda la implementación de este tipo de automatizaciones.

### 5.2.2) API/Web Scrapping para la adquisición de costo diario de combustible

Como vimos en el caso anterior, utilizar sistemas de adquisición de datos automáticos como APIs o Web Scrapping, no se torna rentable para datos de baja frecuencia de actualización como los factores de emisión, pero si es de aplicación interesante para los casos de consumo de combustible, distancia recorrida, entre otras de la flota de vehículos de una compañía, dado que se trata de una enorme cantidad de datos que de manera manual sería casi imposible de obtener, y además tiene otras desventajas ya mencionadas.

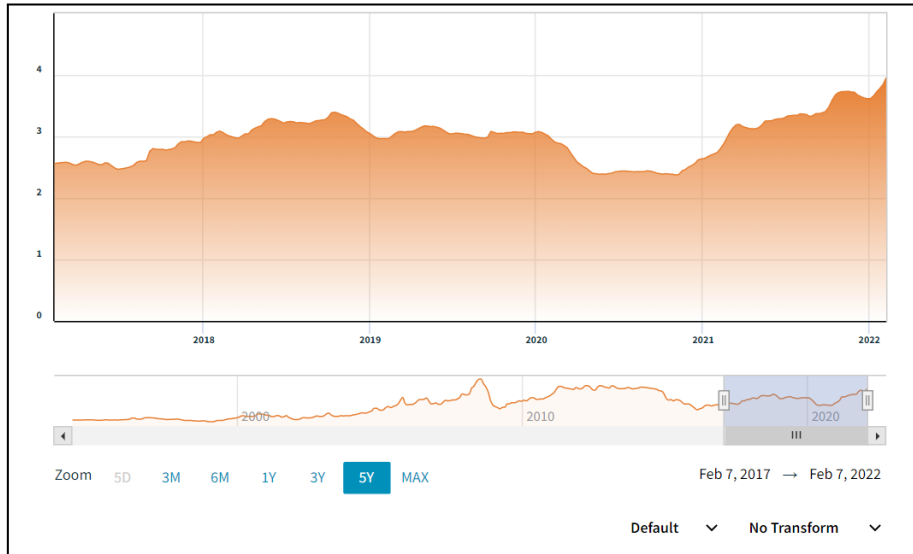
Dentro del análisis económico anteriormente realizado para el caso del Ralentí, se contempló un valor único de precio de Diesel para estimar un posible costo de combustible debido al consumo de combustible con motor encendido pero sin recorrer distancias.

Sin embargo, debemos tener en cuenta que el costo del combustible varía diariamente, incluso en un periodo de año como el contemplado anteriormente, puede tener saltos imprevistos y desproporcionados según las condiciones económicas locales y mundiales así lo presenten.

En el siguiente gráfico de NASDAQ, podemos observar la variación del precio del Diesel en los Estados Unidos en los últimos 5 años:

Figura 18

Precios semanales de combustible Diesel según NASDAQ



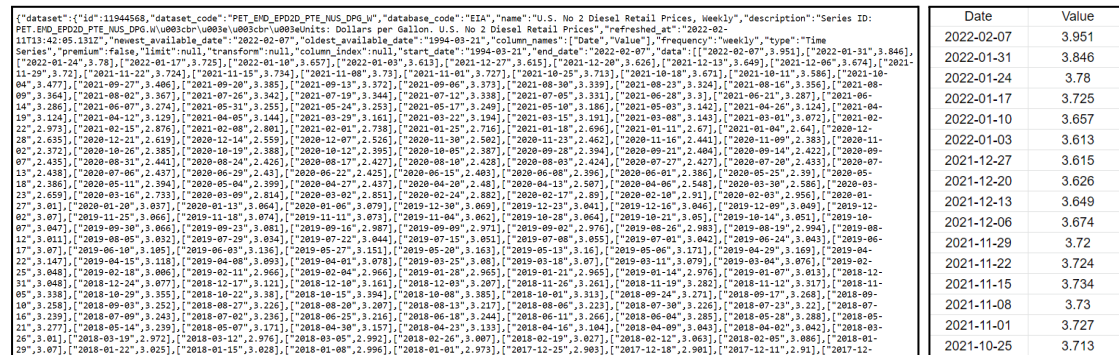
Source: Diesel Retail Prices Weekly, Nasdaq Data Link

El NASDAQ también tiene un sitio web llamado NASDAQ Data Link donde es posible descargar los valores promedios semanales del precio del Diesel y a su vez cuenta con la posibilidad de conectarse via API con dichos datos.

En la siguiente imagen podemos observar la descarga en formato JSON de la API de NASDAQ con los últimos 5 años del precio del combustible Diesel en USD/Gallon, y también en formato tabla para facilitar su visualización:

Figura 19

Descarga via API de precios semanales de combustible Diesel según NASDAQ



Source: Diesel Retail Prices Weekly, Nasdaq Data Link

Este proceso requiere una etapa previa de Data Cleansing, dado que se puede observar que si bien todos los valores contienen el punto para dividir los decimales, en la mayoría de los casos el mismo ha sido omitido y en otros si se contempla, lo cual en caso de no realizar una revisión de los datos y limpieza de los mismos, podríamos obtener desviaciones importantes en los cálculos realizados.

De la base de datos de "Instrument Times" sobre el monitoreo via Telematics de la flota de vehículos logísticos, podemos observar los las fechas en que se inició la medición de los dispositivos de IoT para

medición de datos junto con la cantidad de días que los mismos estuvieron prendidos y la cantidad de aquellos días en los que el vehículo estuvo encendido.

**Figura 20**

Base de datos de "Tiempos Instrumentados"

VehicleID	InstrumentStartDate	InstrumentStartTime	InstrumentEndDate	InstrumentEndTime	InstrumentedDays	Instrumented Weekday	Instrumented Weekend	OperationDays	Operation Weekday	Operation Weekend
1	10/14/2006	11:28:07	10/27/2007	11:11:28	379	270	109	299	232	67
2	10/14/2006	14:25:26	11/3/2007	14:24:06	386	275	111	299	228	71
3	10/14/2006	15:18:46	11/3/2007	11:51:25	386	275	111	241	189	52
4	10/14/2006	14:28:57	10/18/2007	23:01:17	370	264	106	256	186	70
5	10/14/2006	13:44:21	10/27/2007	10:47:01	379	270	109	277	210	67
6	10/16/2006	13:10:14	11/3/2007	13:17:59	384	275	109	278	218	60
7	4/1/2015	9:07:16	4/20/2015	13:20:10	20	14	6	16	14	2
8	4/3/2015	16:49:58	4/23/2015	13:51:15	21	15	6	21	15	6
9	4/1/2015	9:55:23	4/30/2015	23:14:28	30	22	8	22	20	2
10	4/1/2015	11:48:07	4/22/2015	12:24:59	22	16	6	20	16	4
11	4/1/2015	3:34:10	4/22/2015	23:42:42	22	16	6	20	16	4
12	2/1/2014	2:10:11	4/1/2014	0:08:39	60	42	18	50	34	16
13	4/13/2015	18:19:41	4/20/2015	1:49:28	8	6	2	7	5	2
14	4/1/2015	10:23:29	4/30/2015	16:21:35	30	22	8	20	17	3
15	4/1/2015	9:35:56	5/1/2015	23:48:39	31	23	8	29	21	8
16	4/1/2015	10:35:51	4/15/2015	23:52:51	15	11	4	15	11	4
17	4/1/2015	13:51:12	4/27/2015	14:25:49	27	19	8	23	19	4
18	4/1/2015	3:02:23	4/20/2015	21:14:01	20	14	6	17	13	4
19	4/1/2015	10:16:09	4/20/2015	13:31:12	20	14	6	15	12	3
20	2/20/2014	16:59:58	4/1/2014	1:47:14	41	29	12	40	28	12
21	2/3/2014	17:43:14	3/31/2014	23:16:24	57	41	16	53	40	13
22	2/1/2014	2:13:57	4/1/2014	0:12:18	60	42	18	55	39	16
23	2/1/2014	1:22:37	3/22/2014	16:13:26	50	35	15	45	32	13

A partir de estos datos, podemos colocar en nuestra Tabla sobre el costo Anual de vehículos, dichas fechas y calcular la cantidad de semanas promedio. Dado que desde la página del NASDAQ, obtenemos el promedio semanal del precio del Diesel, podemos utilizar estos datos para calcular el promedio del costo de combustible en la cantidad de semanas en las que el vehículo estuvo operando:

**Tabla 19**

Promedio de costo de Combustible con input API de NASDAQ

USD Total Cost per Day	Yearly Cost	Start Date	End Date	Operation Days	Operation weeks	Avg Nasdaq USD/Gal	Yearly Cost	Diferencia
\$4,00	\$8.337,00	10/14/2006	10/27/2007	379	54	2,740	\$5.710,84	0,32
\$4,00	\$7.769,60	10/14/2006	11/3/2007	386	55	2,740	\$5.322,17	0,32
\$4,00	\$4.884,49	10/14/2006	11/3/2007	386	55	2,740	\$3.345,87	0,32
\$4,00	\$7.038,03	10/14/2006	10/18/2007	370	53	2,730	\$4.803,46	0,32
\$4,00	\$9.062,37	10/14/2006	10/27/2007	379	54	2,740	\$6.207,73	0,32
\$4,00	\$6.463,27	10/16/2006	11/3/2007	384	55	2,740	\$4.427,34	0,32
\$4,00	\$628,48	4/3/2015	4/23/2015	21	3	2,917	\$458,32	0,27
\$4,00	\$567,48	2/1/2014	4/1/2014	60	9	3,98	\$564,65	0,005

Se puede observar como la posibilidad de descargar de manera automática los datos de precios de combustible es una herramienta muy poderosa para hacer análisis prácticamente en tiempo real. Existen herramientas pagas con mayor capacidad de análisis, llegando a permitir descargar de datos diarios, los cuales habilitarían a poder hacer análisis cuantitativos en cortos intervalos de tiempo, se pueden analizar tendencias para entender cómo está afectando el consumo de cierto combustible (pensemos en flotas mixtas con vehículos a Diesel, GNC, GNL, entre otros), y a su vez descargar los precios de dichas variedades de combustible para no solo rastrear los gastos de ralentí, sino también para entender posibilidades y beneficios económicos (y ambientales) en posibles cambios de combustible.

Es común por ejemplo, la utilización de otros combustibles alternativos al Diesel para las flota de media milla, como lo son el GNC (Gas Natural Comprimido) y GNL (Gas Natural Licuado). Ambos se utilizan en distancias largas, ya que permiten grandes autonomías, en muchos casos ahorros importantes en costo de combustible y definitivamente una gran reducción de emisiones.

Si volvemos a la tabla de Factores de emisión de la WRI, observamos que el GNL es un 44% menos contaminante que el diesel con un factor de emisión de 6,1 kgCO<sub>2</sub>/Gallon. Recordando el valor para el Diesel:

$$CO2 \text{ Emissions per Gallon}_{On-Road Diesel Fuel} = 10,13 \frac{Kg CO2}{Gallon}$$

$$CO2 \text{ Emissions per Gallon}_{LNG} = 4,46 \frac{Kg CO2}{Gallon}$$

$$CO2 \text{ Emissions reduction per Gallon}_{Replacing Diesel for LNG} = \frac{4,46 \frac{Kg CO2}{Gallon}}{10,13 \frac{Kg CO2}{Gallon}} \approx 44\%$$

Tomando los valores de precio y rendimiento de combustible para LNG de la siguiente tabla):

**Tabla 19**

*Comparación de consumos y precios según tipo de combustible*

Tipo de Combustible	Consumo de combustible por Km	Precio del Combustible
LNG	0,46 m <sup>3</sup> /km	0,7 USD/m <sup>3</sup>
Diesel	0,10 Gallon/km	4 USD/Gallon

Fuente: <sup>19</sup>Fuel Switch to LNG in Heavy Truck Traffic (Smajla et Al. 2019)

Tomando estos valores de Precio y consumo de combustible de ambos combustibles, podemos calcular el costo promedio por km recorrido de la siguiente manera:

$$Cost \text{ per Km}_{Diesel} = 0,10 \frac{Gallon}{Km} * 4 \frac{USD}{Gallon} = 0,4 \frac{USD}{Km}$$

$$Cost \text{ per Km}_{LNG} = 0,46 \frac{m3}{Km} * 0,7 \frac{USD}{m3} = 0,32 \frac{USD}{Km}$$

Como podemos observar el costo por kilómetro recorrido de un vehículo que utilice LNG en vez de Diesel es aproximadamente un 20% menor.

Por tanto la posibilidad de hacer tracking de los precios de los combustibles de manera periódica vía APIs es muy interesante para entender posibilidades económicas de cambio de combustible. Esto aparejado a que combustibles como en LNG, no solo pueden traer aparejado una reducción de costos, sino también un impacto ambiental positivo mediante una reducción de las emisiones de carbono.

En materia de estrategia ambiental, para tomar este tipo de decisiones, es vital contar con datos fidedignos de precios y consumos de combustible en intervalos cortos de tiempo que permitan que la empresa pueda tomar decisiones correctas frente a su gestión. Además, complementando con Telematics, via por la cuál obtenemos datos de rutas y consumos podemos conocer los consumos de Ralentí para entender cuales son las posibilidades de reducción de emisiones y costo asociados también por el cambio de combustible.

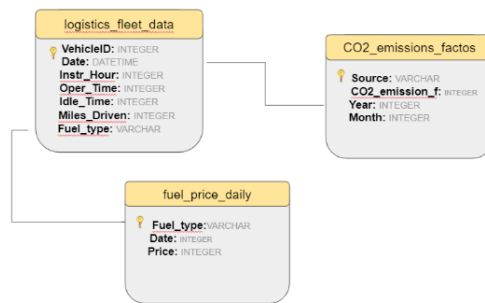
Estos sistemas además contienen la posibilidad de seguimiento de GPS y se pueden obtener detalles precisos de las rutas realizadas. Pensemos que la decisión de cambiar de combustible Diesel a LNG va a traer como consecuencia, no solo el cambio del vehículo a utilizar o su adaptación, sino también asegurar que en dichas rutas se encuentren disponibles estaciones de carga para dicho combustible. Por lo tanto, para tomar efectivamente la decisión de cambiar un combustible como estrategia ambiental para reducir emisiones, también se deben contemplar otros aspectos de costos e infraestructura de carga que podemos obtener de dichas fuentes de datos.

**5.2.3) Bases de datos integradas con APIs para KPIs de flota logística**

Volviendo a los desafíos de Data Warehousing, vemos que estas diversas fuentes de datos deben confluir en bases de datos que nos permitan hacer estos análisis. Es importante entonces sumar una base de datos donde podamos almacenar periódicamente los datos que se descarguen vía API sobre precios de combustibles de manera de poder llevar a cabo los análisis anteriormente mencionados:

**Figura 21**

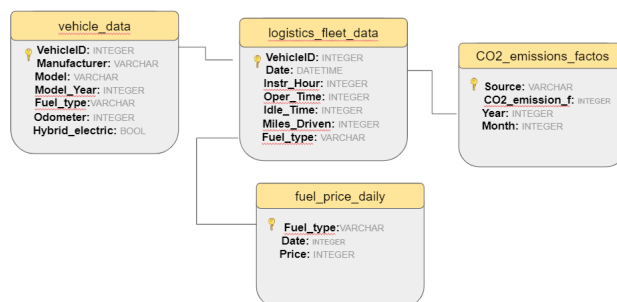
*Incorporación de bases de datos de precios de combustibles*



Para ello incorporamos el dato del tipo de combustible en los datos de flota, dato que se puede parametrizar por ruta según el tipo de sistema de Telematics que se utilice. Sin embargo, dado que la flota de vehículos puede ser variable, también consideramos que es interesante poder segregar los datos de las características de los vehículos, de manera de poder hacer otros tipos de análisis, como porcentaje de flota con emisiones reducidas o vehículos eléctricos. También esto permitiría analizar el promedio de “edad” de la flota, entendiendo que los motores más antiguos no tienen el performance que vehículos más modernos.

**Figura 22**

*Base de datos completa*



Esta gestión de base de datos, permite una mayor facilidad de la gestión ambiental de la flota logística de una empresa. Mediante sistemas API y Telematics, podemos hacer descargas automáticas de datos

que se vuelquen en las tablas de rutas recorridas por cada vehículo, así como también el precio de combustible. A su vez, contamos con otras tablas que requieren ser mantenidas de manera manual como la actualización de los datos de la flota (si bien existen sistemas de Telematics que debido a su conexionado con el motor, pueden obtener estos datos directamente, asumimos también para el caso que no sea así) y la tabla de factores de emisión.

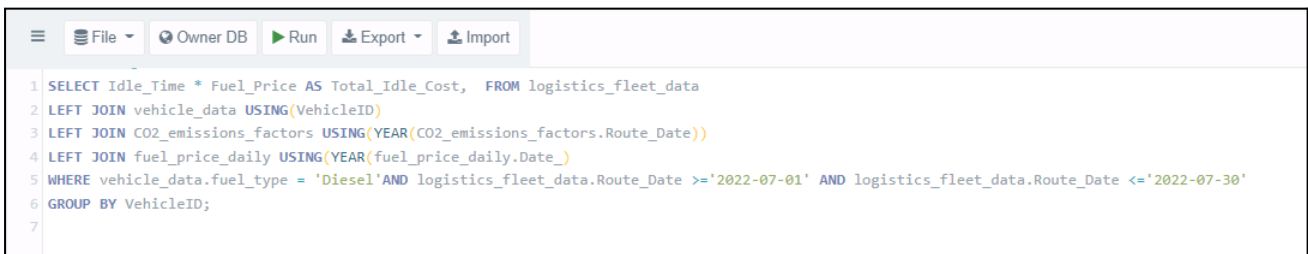
### 5.2.4) Scripts de SQL para cálculo de KPIs Ambientales

En los scripts anexos se encuentran detalles de cómo se pueden calcular fácilmente KPIs importantes para la gestión ambiental de la compañía, realizando sencillas Queries de SQL.

En el siguiente ejemplo, vemos como fácilmente podemos escribir un script en SQL que nos ayude a construir de manera sencilla las tablas que realizamos anteriormente, como por ejemplo calcular el costo total en Ralentí segregado por ID de Vehículo, para el mes de Julio de 2022:

**Figura 23**

*Script de SQL para el cálculo de costo de Ralentí, según bases de datos creadas*



```

1 SELECT Idle_Time * Fuel_Price AS Total_Idle_Cost, FROM logistics_fleet_data
2 LEFT JOIN vehicle_data USING(VehicleID)
3 LEFT JOIN CO2_emissions_factors USING(YEAR(CO2_emissions_factors.Route_Date))
4 LEFT JOIN fuel_price_daily USING(YEAR(fuel_price_daily.Date_))
5 WHERE vehicle_data.fuel_type = 'Diesel' AND logistics_fleet_data.Route_Date >='2022-07-01' AND logistics_fleet_data.Route_Date <='2022-07-30'
6 GROUP BY VehicleID;
7
  
```

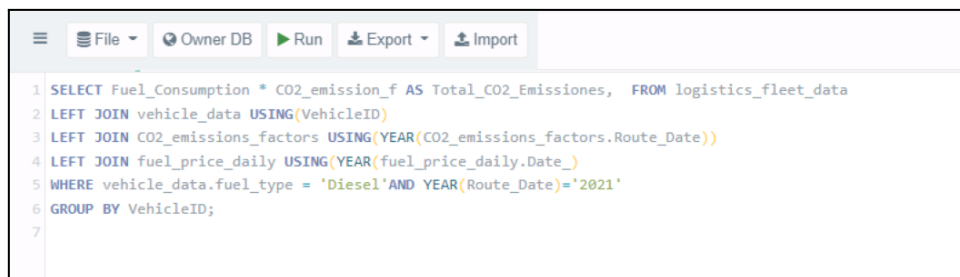
Recordar, que como tomamos el concepto de EPA donde se establece que se consume 1 Gallon de combustible por cada hora de Ralentí, podemos multiplicar directamente tiempo de Ralentí (medido en Hs) y equivalente a Gallons por lo dicho anteriormente y multiplicando por el precio en USD/Gallon descargado directamente de las APIs mencionadas anteriormente.

Vemos como mediante el uso de bases de datos y scripts en SQL, con facilidad podemos realizar cálculos que nos permitan medir y calcular nuestros impactos ambientales y definir estrategias que contribuyan con reducir costos y emisiones.

Lo mismo podemos realizar para dejar establecido un script que nos permita entender las emisiones de la flota para un año establecido:

**Figura 24**

*Script de SQL para el cálculo de emisiones de CO2 generadas en un año agrupado por vehículo*



```

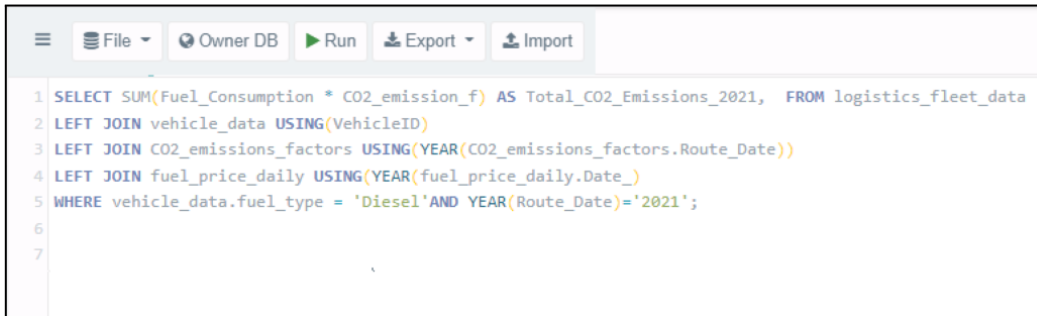
1 SELECT Fuel_Consumption * CO2_emission_f AS Total_CO2_Emissiones, FROM logistics_fleet_data
2 LEFT JOIN vehicle_data USING(VehicleID)
3 LEFT JOIN CO2_emissions_factors USING(YEAR(CO2_emissions_factors.Route_Date))
4 LEFT JOIN fuel_price_daily USING(YEAR(fuel_price_daily.Date_))
5 WHERE vehicle_data.fuel_type = 'Diesel' AND YEAR(Route_Date)='2021'
6 GROUP BY VehicleID;
7
  
```

Como mencionamos en los primeros capítulos, las empresas tienen la responsabilidad conforme a las políticas de ESG de reportar anualmente las emisiones en Reportes públicos de Sustentabilidad. Aquí podemos observar cómo de manera sencilla, podemos conocer el valor de las emisiones generadas por

la flota logística conociendo los datos de consumo de combustible (en caso de que el sistema de Telematics lo pueda medir, en caso contrario se puede hacer con la medida de millaje de dicho sistema), multiplicado por los factores de emisión mantenidos en nuestras tablas de factores. Podemos agrupar el monto por vehículo, y también podemos obtener el valor total de las emisiones por año quitando la sentencia de agrupación y sumando ambas columnas de consumo de combustible y emisiones multiplicadas:

**Figura 25**

*Script de SQL para el cálculo de emisiones de CO2 generadas en un año para el total de flota*



```

1 SELECT SUM(Fuel_Consumption * CO2_emission_f) AS Total_CO2_Emissions_2021, FROM logistics_fleet_data
2 LEFT JOIN vehicle_data USING(VehicleID)
3 LEFT JOIN CO2_emissions_factors USING(YEAR(CO2_emissions_factors.Route_Date))
4 LEFT JOIN fuel_price_daily USING(YEAR(fuel_price_daily.Date_))
5 WHERE vehicle_data.fuel_type = 'Diesel' AND YEAR(Route_Date)='2021';
6
7

```

Es notorio como podemos contribuir con simplificar y automatizar los cálculos y reportes ambientales, con bases de datos sólidas alimentadas con APIs y sistemas de Telemetría, y que además nos contribuyan con poder realizar análisis de datos en pos de definir estrategias ambientales que faciliten la reducción de impactos ambientales así como de costos.

### **5.3) Herramientas implementadas en el proceso de Data Visualization**

Dentro de los procesos anteriormente mencionados, fuimos observando como ciertas tecnologías digitales y de la ciencia de datos, nos permitieron resolver las dificultades que conllevan el Data Gathering, Data Warehousing y Análisis de datos dentro de la gestión ambiental de una compañía, con diversos casos aplicativos a industrias que miden y reportan sus impactos ambientales.

Observamos como tecnologías como Telematics o Smart Metering, nos permiten fácilmente adquirir datos de manera extensiva y sencilla, así como tecnologías de Bases de Datos y SQL nos permiten resolver los desafíos del almacenamiento de datos para su posterior análisis y utilización.

Estas herramientas nos permiten explorar una arista extra dentro de Data Analytics para seguir en un camino de mejora continua y control de nuestra gestión ambiental, la Visualización de Datos.

Durante el proceso de análisis de estas herramientas, fuimos utilizando diversas gráficas y visualizaciones que fueron posibles gracias a que los primeros procesos de la gestión de datos fueron ordenadamente corregidos: recibimos los datos en el formato y frecuencia que la empresa requiere; y fueron almacenados en bases de datos cuyo formato lo elegimos para facilitar nuestro entendimiento, análisis y cálculo de otros indicadores importantes a los cuales buscamos arribar.

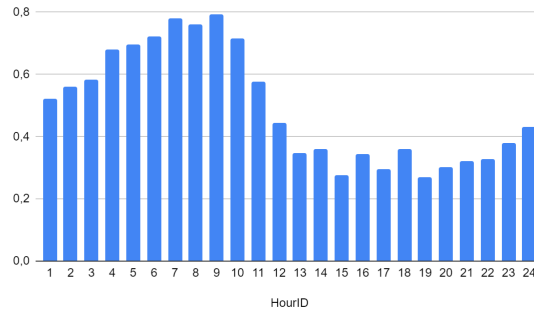
Es notorio como la toma de decisiones y posibilidad de encontrar eficiencias y estrategias efectivas, se multiplica al momento de realizar una correcta visualización de datos.

#### **5.3.1) Visualización de principales indicadores**

En el caso de Ralentí en la flota logística, observamos como un simple gráfico a partir de los datos del sistema de Telematics nos permite tener una herramienta para comenzar a indagar sobre las posibilidades de eficiencia en consumo de combustible con Drivers:

**Figura 26**

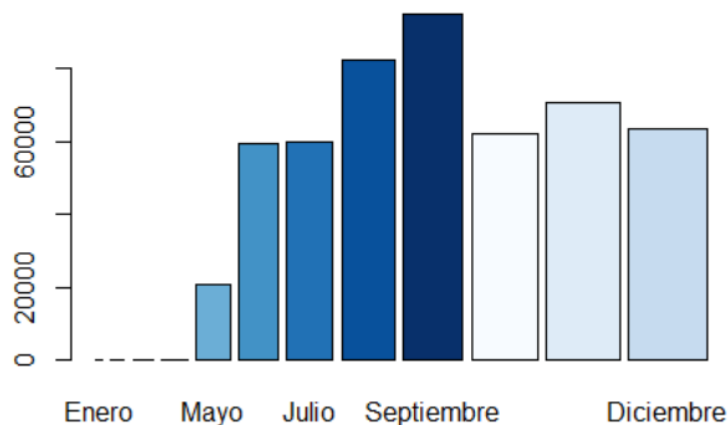
*Porcentaje de tiempo de Ralentí por hora de operación de flota*



A su vez, en el caso de medición de consumo energético via Smart Metering, podemos armar fácilmente gráficos de consumo mensual que nos permitan observar los meses de mayor consumo y buscar eficiencias allí:

**Figura 27**

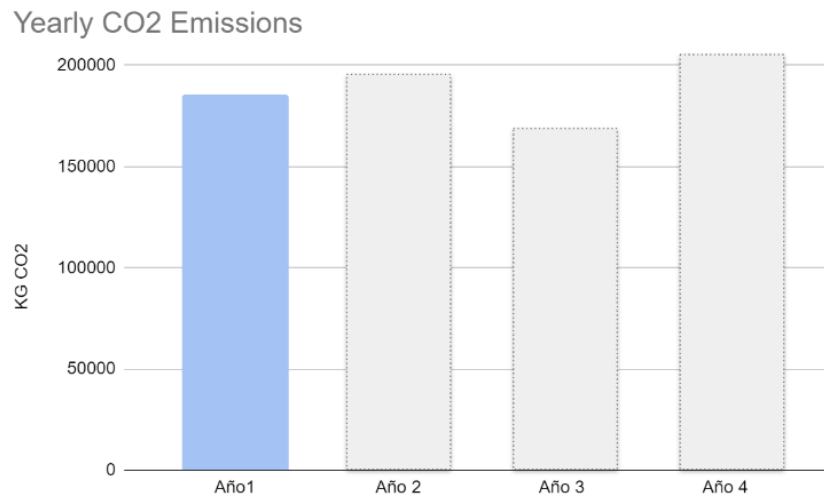
*Consumo energético mensual por mes*



Fácilmente podemos “hacer doble-click” sobre estos datos y entender la variación de consumo mensual por día. Lo cuál vuelve el análisis mucho más interesante. Podemos pensar en colocación de paneles solares para los edificios cuyo consumo energético es ampliamente mayor durante el día.

De cara a los aspectos más estratégicos de la gestión Ambiental, a partir del cálculo del total de emisiones por Ralentí, podemos hacer una proyección en función de datos como crecimiento de flota, kilometraje promedio, rendimiento promedio, fuel efficiency promedio, entre otros, lo cuál nos permite ver a futuro cuál será la base de emisiones y definir estrategias afines. Existe la posibilidad de desarrollar proyectos de regeneración de biomas o de adquisición de bonos de carbono, y para ello contar con un control anual de las emisiones de manera de generar estrategias, es una metodología necesaria para poder llevar a cabo una correcta gestión ambiental dentro del área de una compañía con dichos fines.



**Figura 28**
*Proyección de emisiones de CO2*


En el gráfico anterior podemos observar como fácilmente podemos cuantificar y visualizar las emisiones del Año 1, y hacerlo consecuentemente los próximos años para entender si las estrategias fueron efectivas, y para entender las razones por los cuales las mismas han crecido o reducido, tomando eficiencias a tal fin.

## **5.4) Herramientas implementadas en el proceso de Predictive Analytics**

### **5.4.1) Sentar las bases para predicción de consumo de energía**

Como mencionamos anteriormente en el caso de la medición de energía, sabemos que una de las metodologías más utilizadas para reducción de emisiones de Scope 2 dentro del procurement de energía eléctrica, es la migración hacia fuentes renovables de energía.

Pero como ya comentamos en párrafos anteriores, para abastecer una operación de energías renovables necesitamos tener proyecciones de consumo eléctrico lo más fehacientes posibles, de manera que podamos cerrar un acuerdo PPA (tanto offsite, con generación externa, como onsite, con colocación de paneles solares en edificios propios) a largo plazo de manera que sea económica y técnicamente viable y eficiente.

Para ello, el tener una buena adquisición de datos, como la que nos permiten los sistemas de Smart Metering, desemboca en que podamos utilizar dichos datos para realizar análisis de datos con mayor profundidad para tomar mejores decisiones y desarrollar estrategias con mayor alcance.

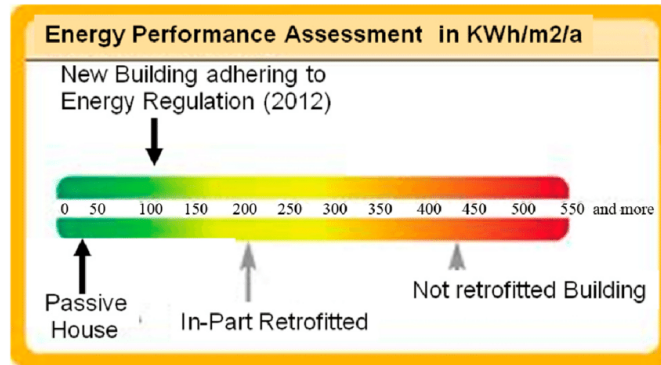
A partir de la implementación de herramientas dentro de los primeros procesos de Data Gathering y Data Warehousing, nos encontramos habilitados para realizar análisis de ciencias de datos y obtener insights interesantes. Estos pasos fueron críticos para poder sentar las bases que van a permitir que se desarrolle la analítica predictiva con mayor profundidad.

Por ejemplo, el clustering vía el algoritmo de K-means puede ser un desarrollo muy interesante para una compañía con diversas fábricas, centros de distribución u oficinas de similares características, pero en distintos países. Dado que como pudimos obtener en el caso de Smart Metering, podemos fácilmente calcular los índices de Intensidad Energética (Kwh/m2), y compararlo entre fábricas de similares características. Podemos encontrar eficiencias entre las mismas operaciones, que por distintas características operativas y constructivas permiten que los consumos sean menores.

Existen escalas como la EPC Energy building (2012), la cual permite cuantificar y observar como los edificios antiguos a los cuales no se le han realizado mejoras constructivas para mejorar sus parámetros de consumo energético tienen un índice de intensidad energética mayor. Esto se traduce simplemente en mayor cantidad de emisiones de CO2 y mayor costo operativo.

**Figura 28**

*Script de SQL para el cálculo de emisiones de CO2 generadas en un año para el total de flota*



Fuente: "Generation of a Tropically Adapted Energy Performance Certificate for Residential Buildings (Wagener, 2014)

Ciertamente a partir de un correcto desarrollo de Data Gathering y Data Warehousing, tenemos las bases para explorar cálculos más complejos que nos permitan indagar y desarrollar mejores eficiencias económicas y ambientales dentro de nuestra estrategia de sustentabilidad.

## 6) Recomendaciones

Hemos recorrido por la implementación de diversas herramientas digitales y su posterior análisis de efectividad, en pos de facilitar los desafíos que existen hoy en día en las empresas y organizaciones en cuanto al desarrollo, control, mejora continua y reporting de una estrategia de sustentabilidad y ambiental.

Sin ninguna duda, las herramientas digitales de la Industria 4.0 como las que hemos estado implementando están generando mejoras sustanciales en los desarrollos tecnológicos de hoy en día y en resolver algunos de los desafíos de la humanidad del siglo XXI. No obstante este abanico importante de herramientas y soluciones, no todas las tecnologías a disposición de la ciencia y la ingeniería resuelven los desafíos que actualmente tenemos.

En un mundo donde las dinámicas organizacionales son cada vez más complejas, con niveles de competencia altos, con velocidades de comunicación tendiendo a instantáneas y exigentes demandas técnicas y profesionales, las organizaciones se ven cada día desafiadas a entender en profundidad cuáles de estas herramientas son las más apropiadas para afrontar sus desafíos.

No todas las organizaciones ni desafíos tienen el nivel de "readiness" necesario para implementar cualquier tecnología, ni cualquier solución tecnológica tendrá el susodicho para resolver cualquier desafío de la actualidad. En organizaciones que son cada día más complejas, no podemos agregar "layers" de complejidad a problemas y estructuras, que ya de por sí en su constitución, son complejas.

Después de haber abordado distintas herramientas digitales en distintas aristas de una estrategia de sustentabilidad, vemos que lo antes dicho no es la excepción.

Es por lo anterior, que detallaremos, en los cuatro estadíos de Data Management analizados en este estudio, cuáles son las recomendaciones sobre las herramientas digitales de la Industria 4.0 más apropiadas para resolver los desafíos ya mencionados de las organizaciones en términos de su Estrategia Ambiental y de Sustentabilidad.

### 6.1) Recomendaciones para el proceso de Data Gathering

Dentro del proceso de **Data Gathering**, a partir de los estudios anteriormente analizados, arribamos a las siguientes recomendaciones:

- Definición clara de indicadores ambientales relevantes dentro de la Estrategia de Sustentabilidad:
  - Si bien esto puede parecer a priori racional y algo alejado de la temática de datos, dentro del desarrollo de esta tesis de maestría pudimos notar cómo dentro de indicadores como la medición de la huella de carbono de una compañía, se desprenden una gran cantidad de métricas e indicadores necesarios para medir. Tener indicadores que sean relevantes en términos ambientales y económicos es fundamental para construir desde allí una fuerte estrategia ambiental y adaptar una gestión de datos tenaz que acompañe las necesidades reales de la organización.

Pudimos observar esto en el caso de medición de Ralentí, que podría ser un indicador muy interesante para empresas cuya flota logística, y por tanto cuyo impacto ambiental de la misma, sea relevante. Si una empresa denota que su huella logística es importante, pudimos observar como si se tomara la simple decisión de realizar mediciones sin sistemas de Telemetría y basados simplemente en datos manuales completados por los drivers, no solo tendríamos un enorme desafío en términos de adquisición de esos datos (ni imaginar su consecuente almacenamiento y tratamiento), sino que además en términos ambientales no estaríamos reportando una huella con gran “accuracy”, por lo que tomar una estrategia ambiental para su reducción con datos de tal manera básicos, es prácticamente imposible.

Además, vemos como se pierden indicadores intrínsecos como la medición de Ralentí, que pueden tener un impacto ambiental y económico muy importante (recordemos que los valores del proyecto MOVE son reales). Adicionalmente, observamos como la implementación de un sistema de Telemetría que contribuya con el proceso de Data Gathering y que nos permita desarrollar una estrategia ambiental sólida para reducir estos costos e impactos, tiene un periodo de pago ridículamente bajo, por tanto su implementación es sumamente recomendable.

Pero para todo esto, es necesario conocer y definir estratégicamente indicadores sólidos y relevantes.

- Conocer las características de los datos que necesitamos (frecuencia, calidad, precisión, etc) y comparar la dificultad de implementación de herramientas digitales en función de los beneficios asociados de mejorar dichas características (Ej: mayor frecuencia, mejor precisión, etc):

Aquí tenemos un caso donde una misma tecnología, dependiendo el contexto, puede ser de mucha utilidad o puede agregar un layer de complejidad. Esto lo observamos mediante el análisis de la implementación de un sistema API para la

adquisición automáticas de factores de emisión. Los sistemas API son una tecnología muy utilizada y que generan muchos beneficios en términos de Data Gathering. Sin embargo, en términos de buscar agilizar y mantener fuentes de factores de emisión, vemos que su implementación es demás compleja. Tenemos que entender el contexto de nuestra organización y como hablamos anteriormente si el dinamismo y exigencias que tienen pueden ser resueltas con estas tecnologías. En este caso observamos, que dado la frecuencia de actualización de estos factores y la complejidad de los sistemas donde se descargan, no existe una gran oportunidad de implementación.

Caso contrario es el de la utilización de APIs para descargar de manera diaria o mensual, los valores de precio de distintos combustibles que podemos tener en nuestra flota logística. En este caso, sí tiene más sentido, en términos de realizar cálculos precisos sobre costos de flota y armado de Business Case para cambios de combustible con emisiones reducidas (como Gas Natural Comprimido, Licuado y/o Biometano). En este caso, debido a la frecuencia en que varían estos valores, y la cantidad que se podrían manejar (pensemos en países donde los valores de combustible tienen fluctuaciones semanales e incluso diarias), herramientas como las APIs pueden ser de gran utilidad.

- Buscar entender el impacto ambiental y económico de nuestras operaciones:
  - Obtener datos con mayor precisión y frecuencia, nos permite arribar a conclusiones que de otra manera no podríamos realizar, y tener una visión completa de la gestión ambiental de nuestra empresa. Sin ninguna duda que aquí se hace reflejo el dicho “lo que no se mide, no se controla”, en el sentido de que necesitamos poder medir variables críticas (que en el caso de flota logística, podrá ser consumo de combustible ó tiempo de ralenti, y en el de una fábrica podrá ser Intensidad Energética), con una precisión que nos permite arribar fácilmente a estrategias de reducción y mitigación que generen eficiencias ambientales y económicas.
  - Es importante entender cuáles son las métricas que realmente tienen un impacto en nuestro negocio, tanto en materia ambiental como económica (aunque vimos que están altamente relacionadas) y definir estrategias de Data Gathering que nos permitan tener los datos que necesitamos en la frecuencia y precisión deseada.
- Entender los costos de la “no implementación”:
  - En las organizaciones actuales, donde la búsqueda de eficiencias económicas y reducción de gastos suelen primar con respecto a otras aristas, es importante poder entender los costos ocultos de la no información, y entender las eficiencias que las empresas podrían estar adquiriendo. Vimos tanto en el caso de Smart Metering como de Telematics, como encontramos repagos en tiempos muy bajos, con inversiones bajas para la organización media. Entender cómo la mejora de procesos de adquisición de datos me abre un mundo de posibilidades de optimizaciones de procesos y costos es vital para poder traccionar estas iniciativas y desarrollar proyectos.
- Comenzar con el “End-Game” en la mira:
  - Nuestros desarrollos de estrategia, ahorros de costos, reducciones de impactos y proyecciones futuras serán tan eficientes como lo sea la calidad de nuestro Data Gathering. Es importante entender, que el proceso de adquisición de datos es la piedra fundamental sobre la cuál vamos a montar toda nuestra estrategia de datos para la gestión ambiental. Si queremos encontrar eficiencias ambientales y económicas, es vital entender qué KPIs buscamos reducir, y para ello que datos

necesitamos. Fácilmente pudimos ver en el caso de gestión de energía, que si queremos a largo plazo desarrollar estrategias de migración hacia Energías Renovables, necesitamos conocer consumos en plazos de 5 años (que son lo que se manejan los contratos PPA) y si queremos encontrar eficiencias de reducción será importante tener otros parámetros como por ejemplo la temperatura, con la que fácilmente podemos entender consumos de equipos de aires acondicionados.

- Asimismo, los procesos de Visualización de Datos serán fuertemente dependientes del ingreso de los mismos. Es importante que desarrollemos las visualizaciones que queremos tener y las que nos ayuden a encontrar eficiencias y navegar aguas arriba hacia la adquisición de datos para definir aquellos parámetros en los que vale la pena invertir debido a las eficiencias que nos otorgarán.

## **6.2) Recomendaciones para el proceso de Data Warehousing**

Dentro del proceso de **Data Warehousing**, a partir de los estudios anteriormente analizados, arribamos a las siguientes recomendaciones:

- Comenzar con el “End-Game” en la mira:
  - Comenzamos por la última recomendación para Data Gathering, dado que aplica para este caso también. Es importante entender cuáles serán los cálculos, análisis y cálculos que buscamos realizar para después volver a navegar aguas arriba hacia definir cómo será el proceso de almacenamiento de datos. Observamos que al planificar nuestro sistema de bases de datos, pudimos encontrar eficiencias y puntos de mejora en nuestro proceso a medida que fuimos analizando los KPIs principales que queríamos calcular o visualizar. Esto también nos va a impactar en el desarrollo de Data Gathering, dado que necesitaremos que los datos arriben de una cierta manera y frecuencia, para definir su almacenamiento óptimo.
- No buscar digitalizar la totalidad del proceso sólo por digitalizar.
  - Un punto importante que observamos, es que no toda solución tecnológica genera eficiencias para el estado de madurez del proceso que abordamos. Observamos cómo el almacenamiento de ciertos datos se tornaba complejo, y que no aportaban ninguna eficiencia al mismo.
  - Trabajando en pos de la agilidad, detectamos también como el desarrollo de Bases de Datos en SQL, es una herramienta realmente ágil para interactuar con nuestras bases de datos y poder calcular y reportar indicadores de manera sencilla.

## **6.3) Recomendaciones para el proceso de Data Visualization**

Dentro del proceso de **Data Visualization**, a partir de los estudios anteriormente analizados, arribamos a las siguientes recomendaciones:

- Generar Visualizaciones con impacto:
  - Pudimos observar durante los casos analizados, cómo ciertas visualizaciones son claves para encontrar eficiencias y desarrollar estrategias ambientales. Es importante que podamos entender cuáles son los puntos de apalancamiento dentro de nuestros impactos ambientales, y cómo podemos visualizarlos de manera de encontrar eficiencias.

- Asegurar la calidad y frecuencia de los datos, es importante para que nuestras visualizaciones sean realmente efectivas y nos permitan tomar decisiones. El estudio de los datos es crítico para entender si las visualizaciones son correctas, si hay desvíos, faltantes de información o se está sesgado por algún input desconocido. Las herramientas analizadas como APIs, Smart Metering y telematics, contribuyen en este sentido.
- El lenguaje técnico de la Sustentabilidad, suele ser un terreno desconocido para muchos practicantes de las industrias y organizaciones. Las visualizaciones de datos, pueden ser una herramienta excepcional para fácilmente explicar por ejemplo a un equipo técnico de Mantenimiento o Finanzas, por qué es conveniente invertir en un sistema de medición de energía en real-time, o por qué es conveniente realizar una migración del suministro a energías limpias.
- Buscar patrones que nos lleven a conclusiones fundamentales.
  - Como observamos en los casos de Medición de Ralentí, pudimos encontrar patrones de comportamiento dentro de los drivers y su consumo de combustible. Ahondar en los datos, explorar distintas visualizaciones y utilizar las posibilidades que nos brinda la gestión de datos es importante para sencillamente encontrar eficiencias que no conlleven ningún tipo de sobre costo o inversión.

#### **6.4) Recomendaciones para el proceso de Predictive Analytics**

Dentro del proceso de **Predictive Analytics**, a partir de los estudios anteriormente analizados, arribamos a las siguientes recomendaciones:

- Sentar las bases hoy para encontrar eficiencias mañana:
  - Pudimos observar durante los casos analizados, como el tener una sólida base de ingreso y almacenamiento de datos, nos permite que podamos explorar estrategias a largo plazo. En el caso de la compra de energía renovable via PPA, observamos que si queremos predecir consumos en base a ciertos algoritmos de Machine Learning o similares, tenemos que poder conocer nuestros datos primarios, obtener una gran cantidad de ellos para llegar a conclusiones válidas, en pos de habilitar estrategias que necesitan de gran fundamento de datos para poder ejecutarse en el largo plazo.
  - Observamos que en muchos casos, los costos de inversión en dejar estas bases sólidas, son bajos y con repagos rápidos. Es importante tener en mente las predicciones y necesidades de nuestra estrategia ambiental a largo plazo, para invertir no solo en eficiencias puntuales, sino en desarrollos a mediano y largo plazo que nos conllevarán a implementar mayor cantidad de proyectos generando reducciones en nuestros impactos ambientales y eficiencias económicas.
  - Es importante que los equipos de Sustentabilidad, comiencen a formarse en temáticas de datos, dado que observamos que existe un sinfín de herramientas prácticas que pueden aprovecharse y capitalizar rápidamente para su exploración.

#### **7) Conclusiones y trabajo a futuro**

Hemos explorado diferentes herramientas digitales en pos de entender cuáles son las mejores alternativas y cómo puedes explotar la capacidad de la tecnología actual para fortalecer y optimizar nuestra gestión ambiental.

Es visible, a partir de los análisis realizados y alineado a la bibliografía estudiada, cómo la digitalización y la sustentabilidad son dos temáticas que caminan de la mano y pueden beneficiarse mutuamente. Dicho esto, también entendimos que no todas las tecnologías son aptas y prácticas para encontrar eficiencias, pudimos encontrar desarrollos que fácilmente generaron ahorros ambientales y económicos; a la vez que ciertas tecnologías incorporan “layers” de complejidad.

Pudimos explorar desde los procesos de Data Gathering, Data Warehousing y Data Visualization, aquellas prácticas que generan eficiencias en las estrategias de sustentabilidad corporativas, en diversas industrias e impactos ambientales, como lo son las operaciones de flota logística y los consumos energéticos. Pudimos validar, que las empresas deben incorporar especialistas en datos y tecnologías digitales a sus estrategias ambientales, para encontrar eficiencias que desemboquen en ahorro de costos y reducción de emisiones.

Entendimos cómo también existe un mundo por explorar, el de la analítica predictiva, en donde para encontrar eficiencias en un universo donde se necesita una cantidad y precisión de datos de niveles exponenciales, es primordial dejar correctas bases sentadas que así lo permitan.

Vivimos en un mundo hoy en día en el cuál las crisis de distintos tipos son moneda corriente día a día. La Crisis ambiental es una problemática que hay que abordar seriamente y de la cuál profesionales de distintas temáticas pueden colaborar en encontrar soluciones practicables, económicas y con gran impacto; sin ninguna duda que las herramientas digitales tendrán un lugar clave para la contribución con los objetivos del IPCC.

Como trabajo futuro, quedará seguir explorando las posibilidades de la analítica predictiva, entendiendo que las bases sólidas están echadas. Seguir explorando eficiencias en todas las industrias, y continuar con el compromiso de ciudadanos, empresas, profesionales y organismos gubernamentales de todo el mundo para seguir explorando, desarrollando y fomentando soluciones que contribuyan con la sustentabilidad global.

## 8) Bibliografía

1. Digitalisation as an Enabler of Circular Economy. 10th CIRP Conference on Industrial Product-Service Systems, 31 May 2018, Linköping, Sweden (Maria Antikainen et al., 2018)
2. A data-driven approach to sustainability, TCS Data analytics Report, 2021 (Tata Consultancy Services, 2021)
3. Big data and open data as sustainability tools, NU. CEPAL, October 2014 (ECLAC, 2014)
4. Artificial intelligence and the circular economy - AI as a tool to accelerate the transition. <http://www.ellenmacarthurfoundation.org/publications> (Ellen MacArthur Foundation, 2019)
5. Sustainable Competitive Advantage Driven by Big Data Analytics and Innovation. Appl. Sci., EISSN 2076-3417, Published by MDPI (Ramadan et al., 2020)
6. Big Data, Big Data Analytics Capability, and Sustainable Innovation Performance. Sustainability, EISSN 2071-1050, Published by MDPI (Hao et al, 2019)
7. An Empirical Study on Visualizing the Intellectual Structure and Hotspots of Big Data Research from a Sustainable Perspective. Sustainability, EISSN 2071-1050, Published by MDPI (Hu et al, 2018)
8. Management theory and big data literature: from a review to a research agenda. December 2018, International Journal of Information Management 43:112-129, DOI: 10.1016/j.ijinfomgt.2018.07.005 (P. de Camargo Fiorini et al., 2018)
9. Harvesting big data to enhance supply chain innovation capabilities: an analytic infrastructure based on deduction graph. International Journal of Production Economics Volume 165, July 2015, Pages 223-233 (K.H. Tan et al., 2015)
10. Role of big data analytics in developing sustainable capabilities. Journal of Cleaner Production 213, 2019, 1264e1273 (Singh et al., 2019)
11. Sustainability as a dynamic organizational capability: a systematic review and a future agenda toward a sustainable transition. Journal of Cleaner Production, 20 January 2017, DOI:10.1016/J.JCLEPRO.2016.07.103 Corpus ID: 53959128 (L.B.L. Amui et al., 2017)
12. From green to sustainability: information Technology and an integrated sustainability framework. March 2011, The Journal of Strategic Information Systems 20(1):63-79 (V. Dao et al., 2011)
13. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol. Forecast. Soc. Change. Technological Forecasting & Social Change (Y. Wang et al., 2018)
14. Measuring technology incorporation/infusion. J. Prod. Innovat. Manag. (R.W. Zmud et al., 1992)
15. Developing A sustainability framework for Industry 4.0. January 2021 Procedia CIRP 98(5):430-435 (Jamwal et al, 2021)
16. Dual Digital: Data Analytics as an Enabler for Sustainability. <https://cib.bnpparibas/dual-digital-data-analytics-as-an-enabler-for-sustainability/> (BNP Paribas, 2018)
17. Investigating the Enablers of Big Data Analytics on Sustainable Supply Chain. Lineth Rodríguez & Mihalis Giannakis & Catherine da Cunha, 2018. "Investigating the Enablers of Big Data Analytics on Sustainable Supply Chain," Post-Print hal-01982533, HAL. (Rodriguez et al, 2018)



18. A framework of sustainable supply chain management: moving toward new theory. June 2008, International Journal of Physical Distribution & Logistics Management, 38(5):360-387 (Carter, C.R. & Rogers, 2008. )
19. A fuzzy multi criteria approach for measuring sustainability performance of a supplier based on triple bottom line approach. Journal of Cleaner Production, Volume 47, May 2013, Pages 345-354 (Govindan, K., 2013)
20. Green supply chain management: pressures, practices and performance within the Chinese automobile industry. Journal of Cleaner Production. Journal of Cleaner Production, Volume 15, Issues 11-12, 2007, Pages 1041-1052 (Zhu, Q., Sarkis, J. & Lai, K., 2007)
21. Modelling the enablers and alternatives for sustainable supply chain management. Concordia University, Department of Concordia Institute for Information Systems Engineering, CIISE, March 2011, Montreal, Quebec, Canada (Hussain, M., 2011)
22. Association Between Supply Chain Glitches and Operating Performance. Management Science Vol. 51, No. 5 (May, 2005), pp. 695-711 (17 pages) (Hendricks, K.B. & Singhal, V.R., 2005)
23. Analytics for Sustainable Business. Deloitte Central Europe Sustainability Report 2012-2013 (Deloitte, 2013)
24. Climate Change 2022, Impacts, Adaptation and Vulnerability. Cambridge University Press, Cambridge, UK and New York, NY, USA, pp.37-188, doi: 10.1017/9781009325844.002 (IPCC, 2022)
25. GHG Protocol Corporate Standard Revised. A Corporate Accounting and Reporting Standard Revised Edition, WRI, Washington DC, USA (World Resource Institute, 2015)
26. Corporate Value Chain (Scope 3) Accounting and Reporting Standard. World Resources Institute and World, Business Council for Sustainable Development, September 2011, ISBN 978-1-56973-772-9 (WRI/WBCSD, 2011)
27. FACTORES DE EMISIÓN DE CO<sub>2</sub> y COEFICIENTES DE PASO A ENERGÍA PRIMARIA DE DIFERENTES FUENTES DE ENERGÍA FINAL CONSUMIDAS EN EL SECTOR DE EDIFICIOS EN ESPAÑA. Reglamento de Instalaciones Térmicas en los Edificios, Ministerios de Industria, Energía, Turismo y Ministerio de Fomento, 14 de Enero 2016, Gobierno de España (RITE, 2016).
28. Guidelines for National Greenhouse Gas Inventories. IPCC 2006, 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. and Tanabe K. (eds). Published: IGES, Japan. (IPCC, 2006)
29. Zhang, Chen; Kotz, Andrew; Kelly, Kenneth (2021): Heavy-Duty Vehicle Activity for EPA MOVES. National Renewable Energy Laboratory. <https://data.nrel.gov/submissions/168>
30. Fuel Switch to LNG in Heavy Truck Traffic (Smajla et Al, 2019)
31. Great energy predictor, ASHRAE DataSet. <https://github.com/energeeks/ashrae-energy-prediction> (ASHRAE, 2020)
32. Greenhouse Gas Data Inventory. [https://di.unfccc.int/time\\_series](https://di.unfccc.int/time_series) (UNFCC, 2022)
33. Emission factor database, EDBF. (IPCC, 2022)
34. Generation of a Tropically Adapted Energy Performance Certificate for Residential Buildings. MDPI Journal Sustainability, ISSN 2071-1050.(Wagener, 2014)