****

Instituto Tecnológico de Estudios Superiores Monterrey

CAMPUS QUERÉTARO

Herramientas computacionales: el arte de la analítica

Grupo 120

**Profesor:**

Pedro Oscar Pérez Murueta

Actividad

Repositorio

**PRESENTA:**

Equipo 4

Jordana Betancourt Menchaca - A01707434

Leonardo Santiago Ramos Pérez A01707812

Stefano Herrejón - A00571720

Fecha:

26 de octubre de 2021

**Repositorio Link :**[**Equipo4**](https://github.com/Leonardo812/Equipo4.git)

**Base de datos seleccionada :**

Link : [Base de Datos](https://www.kaggle.com/selfvivek/environment-impact-of-food-production)

**Motivos :** La razón principal de elegir esta base de datos recae en la idea de poder analizar alimentos que consumimos diariamente, viendo aquellos que causan mayor contaminación ambiental y sustituyendolos por otros que tengan menor impacto.

Resultará de gran trascendencia ser conscientes de que cualquier cosa que comamos aunque no sea ultra procesada causa daño ambiental.

**Información esperada :** Poder establecer relaciones entre las comidas, calorías,kilogramos y la contaminación qué estos producen.

Saber qué carnes, frutas y verduras son las que más dióxido de carbono producen y las que menos.

**¿Celdas vacías, caracteres extraños, signos de puntuación?**

En la primera columna “Greenhouse emmisions(Kg CO2 - equivalents per kg product)”, algunos alimentos cuentan con valores negativos, al analizarlos, se cae en cuenta que estos productos vienen de árboles, por lo cual es correcto que sean negativos, ya que, de hecho, absorben C02.

Alimentos en diferentes columnas tienen valores 0, de igual manera, al examinar si eran correctos estos datos, se concluyó que efectivamente eran apropiados al contexto. Por ejemplo, a las manzanas no se les alimenta con otra cosa que no sea agua o las moras no requieren ningún procedimiento extra, como cera (en caso de las manzanas).

La base de datos no contiene datos NULL, además al estar en inglés no contiene acentos o letras que no puedan ser procesadas.