

Laboratorio 9

Elisa Samayoa y Julio Avila

Para la realización de este laboratorio, se utilizaron dos bases de datos: Food Share Emissions.csv y continents-according-to-our-world-in-data.csv para obtener información sobre las emisiones de carbono en los sistemas de alimentos a nivel mundial. Para la realización de las gráficas, se utilizó PowerBi.

Se obtuvo un total de 5772 registros de 220 países.

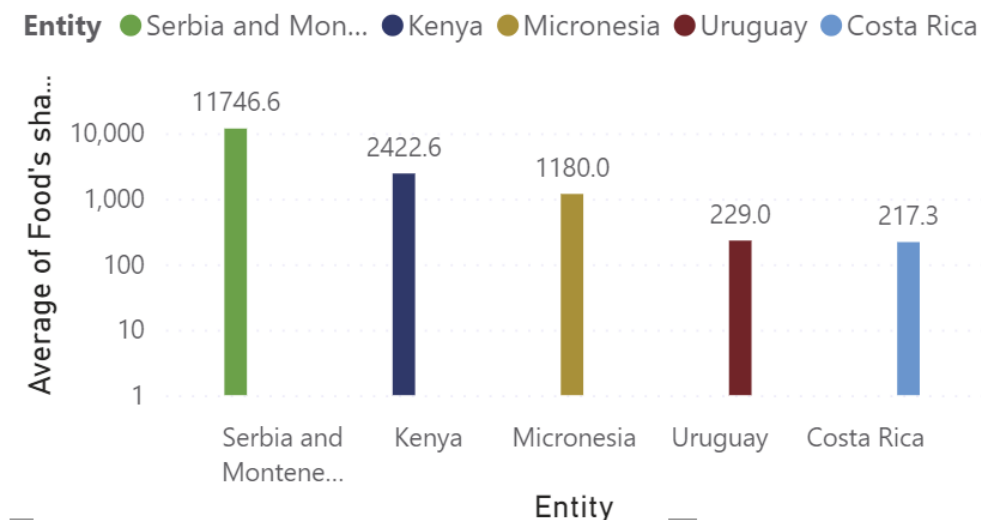
5772

Count of Code

220

Count of Code

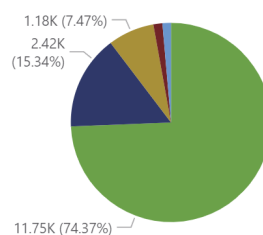
Los cinco países con más emisiones promediadas en todos los años fueron: Serbia y Montenegro, Kenya, Micronesia, Uruguay y Costa Rica. Serbia y Montenegro superó los 11mil emisiones, mientras que para el resto de países hay una mayor diferencia, posicionándose entre los 2500 y 1000 para Kenya y Micronesia y por debajo de 250 para Uruguay y Costa Rica



De esos cinco países, el 74.37% del total de las emisiones promedidas es cubierto por Serbia y Montenegro, Kenya ocupó el 15.34%, por lo que los otros 3 países significan menos del 10%

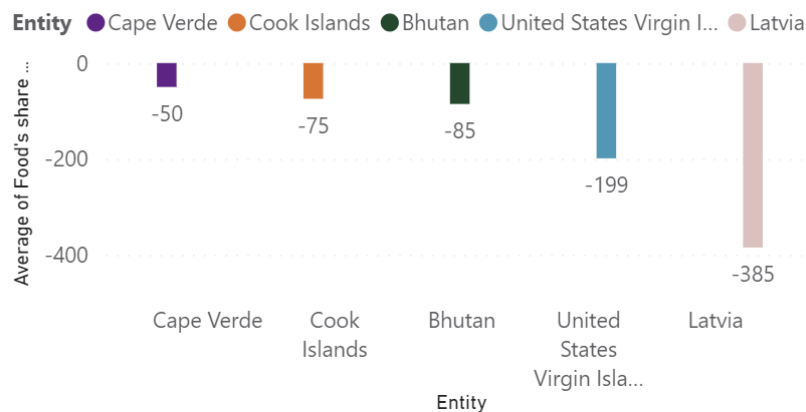
Entity ● Serbia and Mon... ● Kenya ● Micronesia ● Uruguay ● Costa Rica

Porcentaje de emisiones de países del Top 5



Por otro lado, los cinco países con menores emisiones promediadas fueron: Cape Verde, Islas Cook, Bhutan, Las Islas Vírgenes y Latvia. Estos países tienen emisiones negativas que van desde -50 hasta -385, debido a que las emisiones que generan son menores a las que absorben (Goering, 2021).

Top 5 Países con menos emisiones promediadas



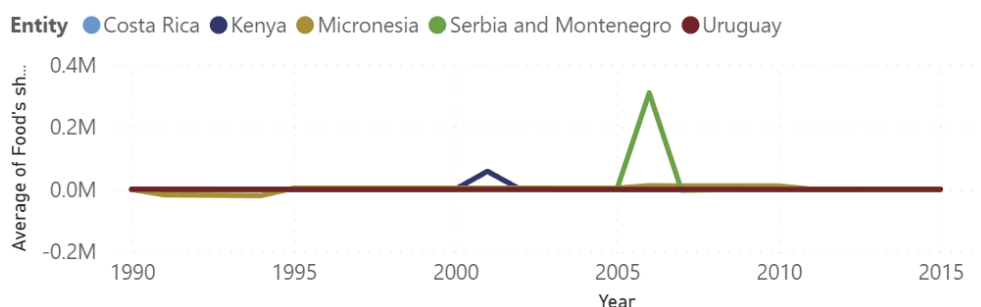
Como se puede observar a continuación, desde 1990 hasta 1995, las emisiones de carbono eran en su mayoría negativas, a partir de 1995 hasta principios de los 2000, a pesar de haber incrementado un poco, se mantuvo bastante estable, por debajo de las 100 emisiones. Sin embargo, alcanzó un pequeño pico entre 2000 a 2002, mientras que a partir de 2005 a 2007 se encontró un segundo pico, con un mayor número de emisiones, superando a las 1000 emisiones al año. A partir de entonces, no se ha encontrado un mayor cambio en las emisiones.

Evolución de las emisiones a lo largo de los años



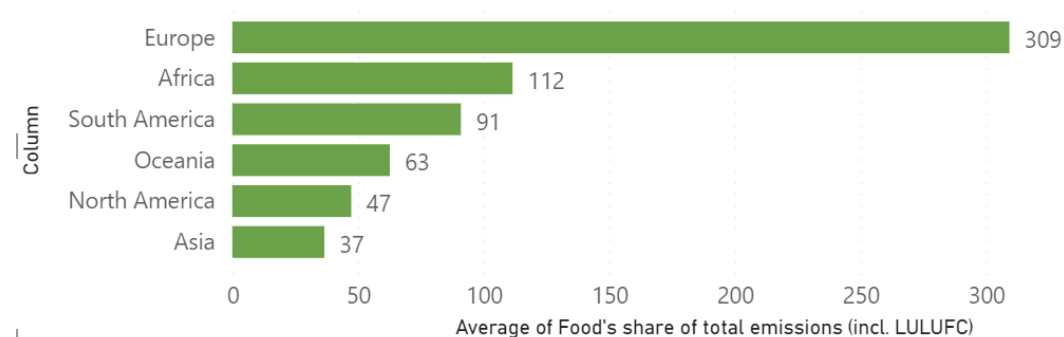
Con el Top 5 que se analizó con anterioridad, se puede observar que Kenya fue el único país que también fue afectado por el primer pico de 2001, mientras que Serbia y Montenegro fue afectado por el segundo. El resto de países, en el resto de los años tienen un comportamiento bastante estable y no se puede observar alguna tendencia.

Evolución de las emisiones a lo largo de los años para el Top 5



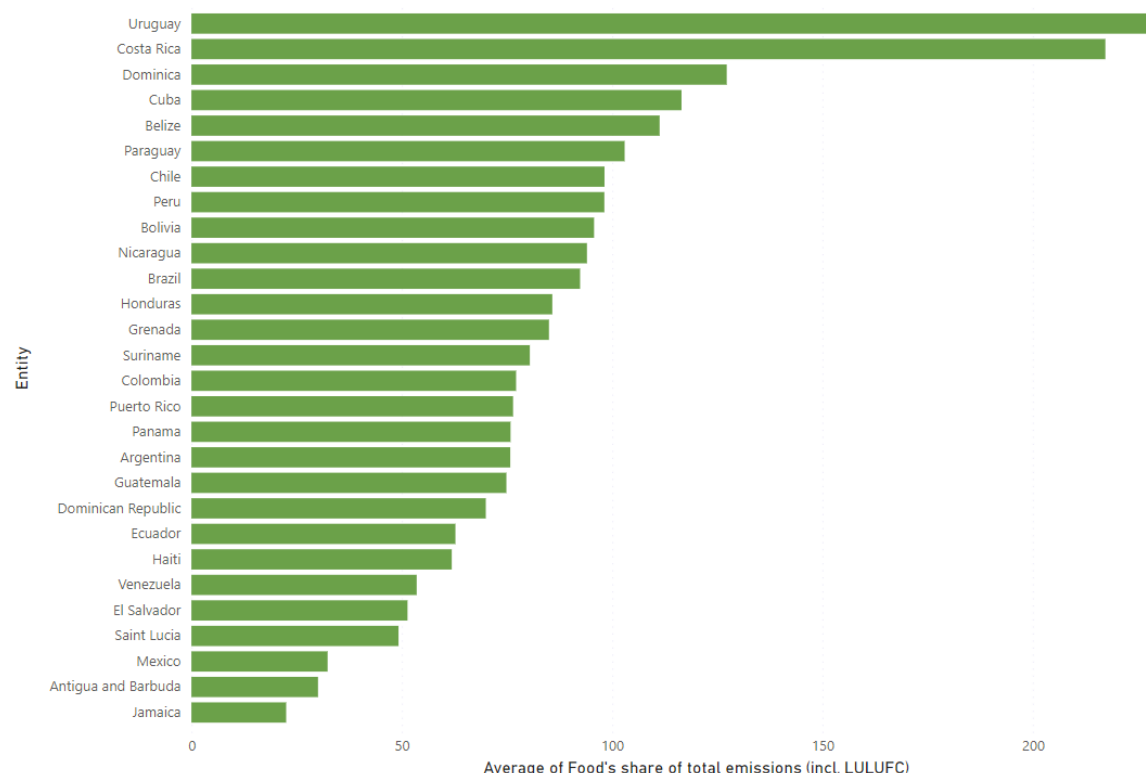
Al analizar el promedio de emisiones por continente, puede observarse que Europa es el que más emisiones ha generado, superando las 309. Mientras que Africa se ubica en segundo lugar, junto con América del Sur. Asia es el último continente con emisiones por debajo de los 40.

Promedio de emisiones por continente



En América Latina, el mayor productor de emisiones de carbono es Uruguay, sobrepasando las 200 emisiones, seguido por Costa Rica y Dominica. Guatemala se encuentra en el puesto 19°, justo debajo de Argentina y sobre República Dominicana, con emisiones entre los 70-80, siendo el segundo más bajo de Centro América. Jamaica es el país con menos emisiones, entre los 20-30.

Promedio de emisiones por países de Latinoamérica



Reflexión

Ambos ya habíamos tenido la oportunidad de trabajar con herramientas de visualización como PowerBi o Tableau, pero siempre es útil aprender y practicar al generar visualizaciones, puesto que es una de las partes más importantes del trabajo de un científico de datos. De muy poco sirve todo el trabajo realizado, si no se puede comunicar bien la información.

- Julio Avila: En lo personal, solo he trabajado con PowerBi, por lo que realizar este trabajo fue más que todo de práctica. Sin embargo, para poder conectar la base de datos externa, tuvimos que realizar algunas conexiones extra en PowerBi, por lo que, al menos desde ese aspecto, eso fue algo nuevo para mí. Considero que una de las más grandes razones por las cuales es importante tener ética en las visualizaciones es para evitar la desinformación. Muchas personas solo ven las gráficas en internet y no leen por completo los informes o estudios que las sustentan, por lo que al generar visualizaciones engañosas, es muy probable que existan malos entendidos que, dependiendo del tema, puedan generar graves consecuencias. Algo relacionado a esto, puede ser el tema del COVID19. En los primeros meses se estaban haciendo muchos análisis de datos, algunos correctos y otros no, por lo que había bastante desinformación al respecto de la situación. Para ser más específicos, las visualizaciones que reportaban las muertes, los casos activos y los casos totales eran las que generaron mucha especulación porque había personas que no podían interpretarlas de manera idónea o porque los datos que presentaban se actualizaban muy seguido, por lo que era complicado entender cuál era la correcta.
- Elisa Samayoa: A pesar de que no es la primera vez que trabajo con una herramienta de visualización (he manejado tableau, powerbi, etc), uno nunca termina de aprender en este tema. Siempre hay maneras más efectivas y útiles para presentar los datos y hacerlos atractivos para la audiencia. Si bien hay ciertas prácticas que uno puede seguir (como agregar siempre títulos en ejes, en gráficas, usar colores atractivos, etcétera), mucho de esto también depende de la personalización que se le quiera dar. Como futura científica de datos, es indispensable saber realizar visualizaciones y contar historias, ya que, las personas realmente necesitan entender lo que se va a contar de una manera bastante simple y memorable. Considero que la ética en este aspecto es importante para preservar la confianza también, ya que, cuando alguna persona o entidad empieza a generar o presentar datos que son incorrectos o que tienden a malinterpretarse, se pierde el interés y la credibilidad por ellos, por lo que es probable que se busquen otras fuentes. Un caso que dio de qué hablar fue respecto a los resultados que presentaron en la televisión sobre una encuesta del gobierno de India. Ya que, aunque el 35% había contestado que sí y el resto que no, la proporción en el gráfico de pie del “sí” era considerablemente mayor al otro, por lo que, era de esperarse que hubiera manipulación.

