

Visualización de emisiones de CO_2

Víctor López Iglesias

Diseño y Evaluación de Sistemas interactivos
4º curso de InDat

Estructura del contenido

1	Introducción	2
2	Análisis de los datos	2
2.1	Origen y tipología de los datos	3
2.2	Transformaciones	3
3	Planificación de la visualización	4
3.1	Descripción de los objetivos	4
3.2	Función	4
3.3	Tono	5
3.4	Efectos deseados	5
3.5	Enfoque	5
4	Diseño	6
4.1	Parte 1: Emisiones totales de CO_2	6
4.2	Parte 2: Emisiones per cápita de CO_2	6
4.3	Decisiones comunes a ambas partes	7
4.3.1	Colores	7
4.3.2	Interactividad	8
4.3.3	Layout y componentes de la visualización	8
4.4	Justificación de la representación bipartita	9
4.5	Bocetos de la propuesta	9
4.5.1	Parte 1 de la visualización	10
4.5.2	Parte 2 de la visualización	11
5	Conclusiones	12
6	Apéndices	12
6.1	Apéndice 1: Análisis de los datos	12

1 Introducción

El ritmo de consumo de la sociedad actual ha provocado serias repercusiones en el medioambiente debidas, en su mayor parte, al efecto invernadero. El efecto invernadero es resultado de las elevadas emisiones de gases dañinos para la capa de ozono de la atmósfera. El principal de estos gases es el CO_2 , el cual se produce principalmente como resultado de procesos de combustión (coches, fábricas, producción de energía, calefacción ...).



A colación de estos resultados y del paulatino perjuicio que esto supone para el planeta, se llegan a proponer y firmar varios tratados como solución a este problema. Uno de ellos y el más conocido es el protocolo de Kyoto, en 1997, que establece ciertos criterios para lograr el descenso de la emisión de CO_2 a la atmósfera.

El objetivo de este trabajo es, por tanto, la realización de una visualización que nos permita mostrar al público general la situación actual de las emisiones de CO_2 , gas que todo ciudadano conoce por su perjuicio medioambiental, y concienciarlo acerca de la situación crítica, así como explorar la tendencia de la evolución de las emisiones de CO_2 tanto globalmente como a posteriori del protocolo de Kyoto.

El objetivo del trabajo así como las decisiones tomadas para llegar a cumplirlo se encuentran detalladas en mayor profundidad en la sección 3.

2 Análisis de los datos

Para la realización de la visualización se utilizarán dos ficheros de datos. Los datos de ambos ficheros se encuentran en formato CSV, lo cual no será un impedimento para integrarlos en la herramienta de desarrollo de visualizaciones D3, dado que dicha herramienta lo soporta.

2.1 Origen y tipología de los datos

El primer fichero de datos utilizados (1) contiene información acerca de las emisiones de CO_2 per cápita para los distintos países y continentes del mundo de manera anual. Las variables que contiene el fichero son: entidad (entity), nombre del país o continente, código (code), identificador unitario de 3 caracteres del país en cuestión (ausente para los continentes y regiones especiales), el año (year) de los datos tomados y las emisiones de CO_2 per cápita en toneladas de dicho año y país.

El segundo fichero (2) contiene información acerca de las emisiones de CO_2 totales por año para los distintos países y continentes del mundo. Igual que en el fichero anterior tenemos las variables entidad, código y año, y como es lógico aparece una nueva variable en sustitución de la anterior que contiene las emisiones totales de CO_2 , en millones de toneladas, para el año y país en cuestión.

Dado que las variables que vamos a utilizar para el desarrollo de la visualización son el nombre del país, el año del dato, las emisiones anuales y las emisiones per capita anuales es importante conocer la naturaleza de las mismas en mayor profundidad. El nombre del país/región es una variable categórica nominal y en total hay 241 diferentes. La variable año es cauntitativa de tipo intervalo, en el caso de los datos per capita el rango es 1785-2016 y en el de emisiones totales 1751-2016. Tanto las emisiones per cápita como las totales son datos numéricos continuos y positivos (el 0 incluido) pero están en diferentes escalas (lo cual es importante), los primeros en toneladas y los segundos en millones de toneladas.

2.2 Transformaciones

Como ya se verá en la sección 4 se utilizarán los datos para un número determinado de países (el número de países utilizado dependerá de razones estéticas, por lo cual puede variar, pero se encontrará entre 5-10) que tengan la mayor emisión global del gas en cuestión para el último año registrado (2016). Esta selección se mantendrá constante en ambas partes de la visualización por razones explicadas posteriormente. Por tanto, aquí aparece la primera transformación que tendremos que realizar sobre los datos, debido a que en ellos se encuentra la lista completa de países y regiones, de la cual deberemos extraer dichos países, además del valor mundial para el caso de las emisiones totales.

El rango de años utilizado será desde 1950 hasta 2016, por lo cual extraeremos únicamente, para los países seleccionados, datos en este intervalo de tiempo.

La selección preliminar de países (incluido el mundo) se expone en la siguiente tabla, en un fragmento perteneciente a los datos de emisiones totales:

Entity	Code	Year	Annual emissions
Saudi Arabia	SAU	2016	634.37
Iran	IRN	2016	655.89
Germany	DEU	2016	801.51
Japan	JPN	2016	1209.16
Russia	RUS	2016	1634.71
India	IND	2016	2430.80
United States	USA	2016	5311.69
China	CHN	2016	10150.82
World	OWID_WRL	2016	36182.59

Table 1: Países con mayor emisión anual de CO_2 en 2016.

El análisis de los datos se ha realizado con la herramienta R y el código utilizado se puede observar en la sección 6.1.

3 Planificación de la visualización

3.1 Descripción de los objetivos

El objetivo principal de la visualización es mostrar como los niveles de emisión de CO_2 en ciertos países han incrementado de manera drástica y preocupante con el paso del tiempo, buscando una mentalidad de cambio en el receptor.

Otro objetivo es explorar la situación actual y cómo ha evolucionado ésta a partir del famoso protocolo de Kyoto, apreciando las diferencias entre países.

También se busca observar si cierta parte del mundo (pequeño porcentaje de países) contribuye en un elevado porcentaje a la emisión mundial de este gas, localizarlos en dicho caso, y ver si se debe en exclusiva a un alto volumen de población o se trata de una población cuya cultura es consumista por excelencia.

3.2 Función

La función es en parte explicativa y en parte exploratoria. Explicativa porque, en general, es de conocimiento público que los niveles de emisiones de CO_2 han crecido desmesuradamente y el propósito es mostrarlo de una manera más visual. La parte exploratoria de la visualización comprende la observación de la tendencia general y de manera posterior al protocolo de Kyoto, además de la exploración subyacente a la participación de cada una de

las regiones/países en la contaminación y su causalidad (elevado volumen de población o consumo excesivo).

3.3 Tono

El tono empleado es pragmático ya que los datos de por sí ya causan el impacto necesario como para hacernos reflexionar y no queremos perder su utilidad exploratoria centrándonos demasiado en darle un tono emotivo.

3.4 Efectos deseados

Los efectos deseados de la visualización se han podido intuir del análisis realizado hasta ahora pero se explicarán a continuación a modo de resumen y ampliación de los mismos:

1. Buscar una actitud de cambio en el individuo.
2. Conocer la tendencia global y patrones de la misma
3. Conocer la tendencia a posteriori del protocolo de Kyoto
4. Conocer las diferencias entre los países con mayores emisiones de CO_2
5. Conocer la contribución de los países con mayores emisiones de CO_2 al total
6. Ayudarnos a distinguir los países con una cultura de consumo y contaminación excesiva

3.5 Enfoque

A continuación se expone el enfoque que se ha tomado para llevar a cabo los objetivos de la visualización propuesta:

1. Mostrar la evolución de las emisiones de CO_2 per cápita de los países más contaminantes.
2. Mostrar la evolución de las emisiones de CO_2 totales de los países más contaminantes, buscando una perspectiva de proporción mundial.
3. Permitir permutar entre las dos partes de la visualización.
4. Permitir limitar el rango de años mostrado en las emisiones per cápita (evolución hasta un límite).
5. Permitir eliminar países de la visualización.
6. Mostrar la fecha del protocolo de Kyoto.
7. Mostrar la evolución de las emisiones globales del mundo como referencia.
8. Permitir resaltar visualmente un país en concreto.
9. Permitir obtener información numérica.

4 Diseño

En primer lugar, para realizar un análisis de la estructura global de la visualización, se debe destacar que la visualización va a constar de dos partes debido a que el correcto cumplimiento de los objetivos así lo requiere, como justificaremos más adelante.

En ambas partes se ha utilizado una selección de los países con mayores emisiones totales de CO_2 en la actualidad, esto se ha decidido a causa de los límites visuales, se busca una representación limpia y que permita explorar los datos, y en consecuencia, se han seleccionado los más representativos ya que son los que tienen un mayor margen de mejora. Que en la representación se hayan utilizado países y no regiones (ó continentes) se debe a la búsqueda de una mayor localización y efectividad del impacto.

4.1 Parte 1: Emisiones totales de CO_2

La primera parte utiliza los datos de emisiones anuales de CO_2 y se ha seleccionado como método de representación un *Stacked Area Chart*. Dicha elección se debe a que los objetivos que tratamos de satisfacer con esta parte son poder discernir la aportación de los países escogidos sobre el total a lo largo del tiempo y transmitir al receptor una mentalidad de contribución (todos estamos contribuyendo a esta situación desfavorable). Este método de representación nos permite representar las tres variables propuestas: una cuantitativa de tipo intervalo (año), una categórica (país) y una cuantitativa (emisiones totales de CO_2).

Por tanto, en este gráfico se incluirá en el eje de abscisas el año y en el de ordenadas las emisiones de CO_2 de cada país. La altura de cada area representa en consecuencia el aporte en emisiones de CO_2 de cada país que se suma a la altura de los anteriores, mostrando las emisiones acumuladas de CO_2 . Para representar cada país se utilizará el color como variable visual, asociado mediante una leyenda. Además se incluirá el valor de las emisiones mundiales (línea) para tenerlo como referencia sin añadir más ruido visual.

4.2 Parte 2: Emisiones per cápita de CO_2

En la segunda parte se representan las emisiones per cápita de CO_2 . Para representar dichos datos se utilizará un *Line Chart*. Se ha elegido este gráfico dado que es idóneo para explorar la tendencia y valorar la diferencia entre los distintos componentes del gráfico, lo que le hace una opción muy interesante para visualizaciones exploratorias.

Se representará en el eje de abscisas el año (variable cuantitativa de tipo intervalo) y en el eje de ordenadas las emisiones per cápita (cuantitativa). Es común que una de las variables en un *Line Chart* sea cuantitativa de tipo ratio pero el resultado con una variable cuantitativa es perfectamente válido. El color como variable visual se utilizará para representar el país (variable cualitativa).

Se incluirá una interacción exclusiva (manipulación de datos) de esta parte que busca aislar

la tendencia hasta cierto límite temporal con el fin de no influir nuestra visión con lo que sucede a partir de cierto punto. Este elemento interactivo es un *slider* que nos permite decidir cuál es el último año representado dentro del rango permitido (figura 6). Esto, por ejemplo, nos facilita observar cuál es la tendencia antes y después del protocolo de Kyoto de una manera más eficaz.

4.3 Decisiones comunes a ambas partes

Para poder explorar la tendencia posterior al protocolo de Kyoto se incluirá una marca temporal en los gráficos (eje abscisas). Esto servirá para tener de referencia en todo momento dicho hito.

4.3.1 Colores

Se ha decidido utilizar como color de fondo de la visualización el blanco, que da sensación de limpieza y claridad y facilita la exploración de los datos.

Los colores utilizados para representar cada uno de los países serán en la medida de lo posible asociados a su respectiva bandera. Por ejemplo, para Japón no se podría utilizar el blanco (su color principal) si se utiliza un fondo blanco para la visualización, ya que no habría contraste y causaría confusión. Si hay repetición o imposibilidad de utilización para alguno de los colores se buscarán otros que contrasten con el resto.

En consecuencia se ha elegido esta paleta colores:

	#EE2E2E	China
	#9B54C3	Estados Unidos
	#FF9933	India
	#3574EE	Rusia
	#FFDD00	Japón
	#000000	Alemania
	#C38865	Irán
	#006C35	Arabia Saudí

Figure 1: Paleta de colores de los países

Los colores empleados, se mantendrán constantes en ambas partes de la visualización, para no romper la asociación cognitiva (color-país) ya creada y no causar confusión.

Se puede ver el uso de los colores en las figuras de la sección 4.5.

4.3.2 Interactividad

La visualización debe permitir el cambio entre ambas partes. Esto se satisfecerá con un selector de 'modo' que constará de dos *Radio Button*, dado que se trata de una estructura interactiva que nos permite elegir una de varias opciones y en todo momento sabemos cuál está seleccionada, lo cual es importante. Esta es una interacción del tipo manipulación de datos, que permite variar la variable representada en el eje de abscisas y con ello la representación.

El cambio de emisiones globales a emisiones per cápita se llevará a cabo mediante una animación que estrechará los áreas hasta convertirlos en líneas y dichas líneas se reposicionarán. El cambio contrario será otra animación en la que las líneas se dirigen hacia el lado derecho hasta desaparecer y se extienden desde abajo los áreas hasta llegar a su posición. Es importante que resulten fluidas y suaves para guardar cohesión entre ambas partes. Esto no es una interacción como tal, pero es parte de la interacción de cambio de 'modo' por esto se ha decidido incluirlo en esta sección.

Otra funcionalidad importante es la exclusión de países con el fin de poder eliminar ruido de la visualización y observar únicamente lo que queremos. Al hacer click en el nombre del país en la leyenda podemos eliminarlo de la visualización (se tachará su nombre en la leyenda). Si vuelves a hacer click se vuelve a representar. El tipo de interacción al que da lugar es una manipulación de datos (exclusión).

Con el fin de hacer énfasis y destacar un país sin perder de vista el resto (comparaciones de uno con los demás), se añadirá una interacción de modificación que se basa en que al pasar el ratón por el nombre del país en la leyenda el resto disminuya su alfa, haciéndolos menos visibles, logrando resaltar el primero. Se puede observar en la Figura 5 para el caso de las emisiones per cápita.

Para poder obtener datos numéricos concretos, pasar el ratón por encima del 'punto' de una línea del *Line Chart* o en una determinada posición horizontal dentro del gráfico (en el *Stacked Area Chart*) despliega un tooltip con información detallada, como se puede ver en las figuras 3 y 4. El uso de una estructura visual que organice la información es muy importante dada la facilidad de percepción de las mismas. En el caso del gráfico de emisiones per cápita se mostrará para ese país el dato de las emisiones totales y el de las per cápita (a modo de comparación). Para el gráfico de las emisiones totales la información mostrada será para dicho año los valores de cada uno de los países. Esta interacción se corresponde con una del tipo anotación.

4.3.3 Layout y componentes de la visualización

El layout de la visualización consta del título, la zona de despliegue del gráfico, la leyenda y el selector. Para el caso de la visualización de los datos per cápita aparece además el limitador temporal. Cada uno de los componentes se diferenciará del resto debido a agrupaciones creadas por proximidad (principio Gelstat). La disposición de estos se puede observar en la

sección 4.5.

En cuanto a los ejes, el temporal no incluye título dado que los valores dejan bastante claro que se trata de años, además de que el título principal no deja lugar a dudas. No aparecerán todos los años representados en el eje debido a que lo que se busca con estos valores es tener una referencia (de 10 en 10 años). En el eje de ordenadas (emisiones de CO_2) se mostrará un título para clarificar la naturaleza del dato y las unidades (aunque no se mencionará si es per cápita o total dado que lo sabemos por el selector). En un principio no tomará valores dicho eje dado que creemos que no es necesario para las tareas exploratorias deseadas y siempre se puede particularizar la información (tooltip).

Se ha decidido no mostrar *tick marks* en ninguno de los dos ejes dado que la visión exploratoria deseada se basa más en tendencias y proporciones que en datos precisos (estas sirven de guías visuales para precisar los datos) y así logramos una visualización más limpia. La rejilla no es necesaria, su representación recargaría en exceso la visualización y se podría confundir con la marca temporal utilizada (Protocolo de Kyoto) por lo que se ha decidido no incluirla.

4.4 Justificación de la representación bipartita

Como ya se ha podido observar en secciones anteriores, la utilización de este modelo de representación bipartito nos permite cumplir los objetivos marcados para la visualización.

Sin embargo, la decisión de utilizar estos dos tipos de datos conjuntamente va más allá. Todo nace bajo la consideración de que la presentación unilateral de cualquiera de las partes sin la presencia de la otra es incompleta.

Podrían mostrarse únicamente los datos de emisiones globales, pero esta información podría llevar a equívoco, ya que es posible la existencia de entidades cuyo volumen de población sea muy superior al resto y que sin embargo la contribución de cada individuo sea extremadamente baja.

Por la otra parte, en la representación de las emisiones per cápita, puede ser que existan países extremadamente industrializados formados por una única ciudad, con un elevadísimo valor per cápita pero que aportan una insignificante parte del total ya que tienen un volumen de población muy bajo.

Por todo esto, se ha considerado este modelo, que encaja con lo buscado a nivel de objetivos y facilita el contraste de la información mostrada, evitando una visualización fácilmente malinterpretable.

4.5 Bocetos de la propuesta

Esta subsección está dedicada a mostrar de manera visual el diseño preliminar de la visualización.

4.5.1 Parte 1 de la visualización

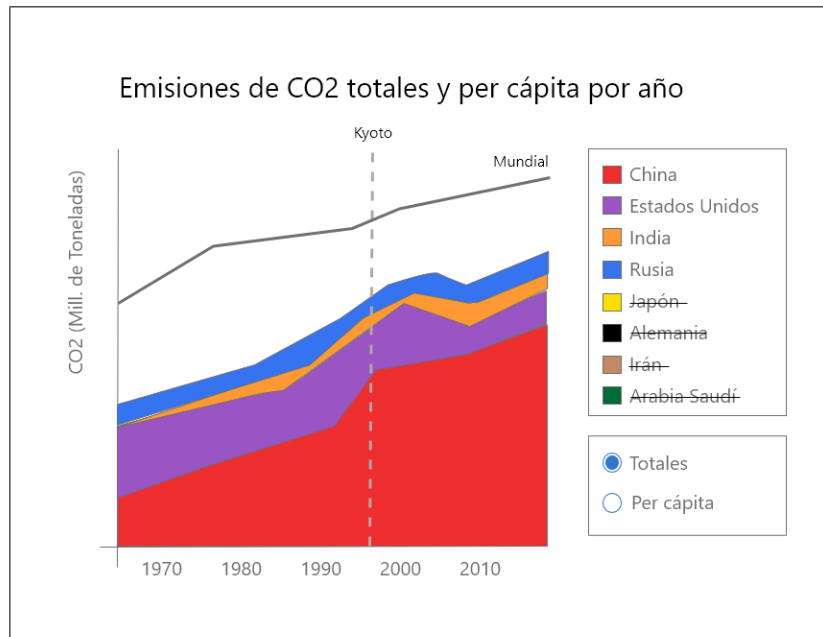


Figure 2: Visualización de emisiones totales

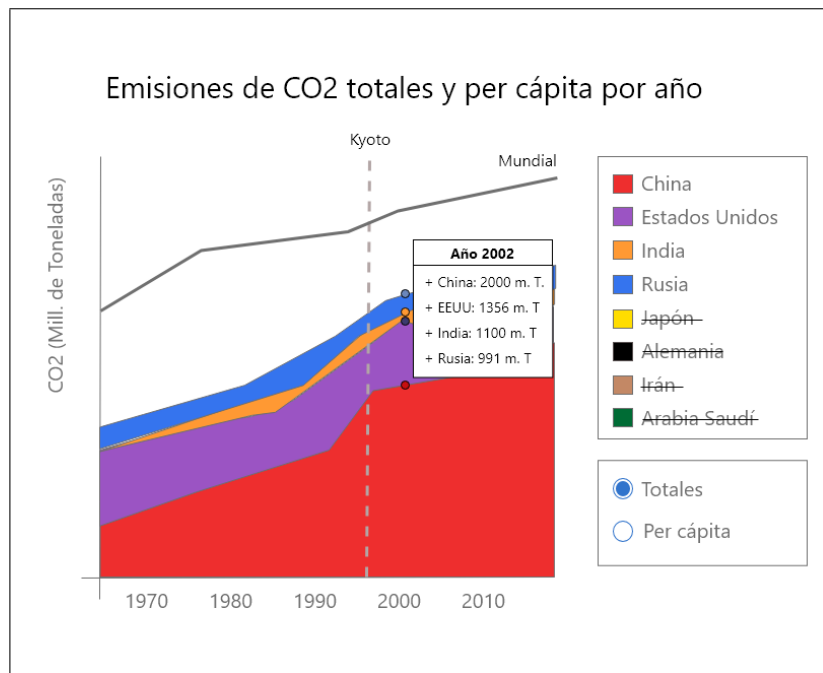


Figure 3: Visualización de emisiones totales con tooltip

4.5.2 Parte 2 de la visualización

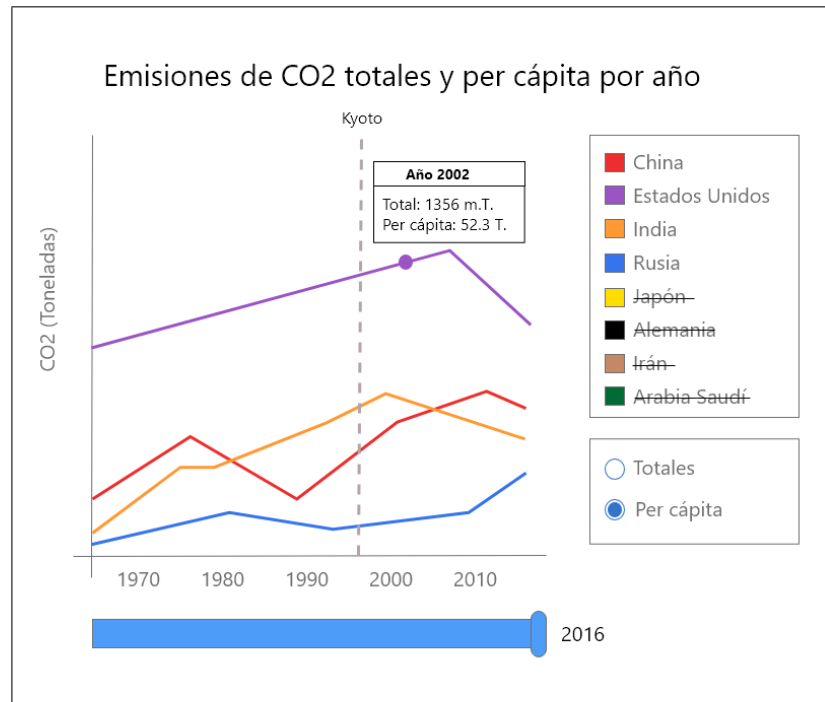


Figure 4: Visualización de emisiones per cápita con tooltip

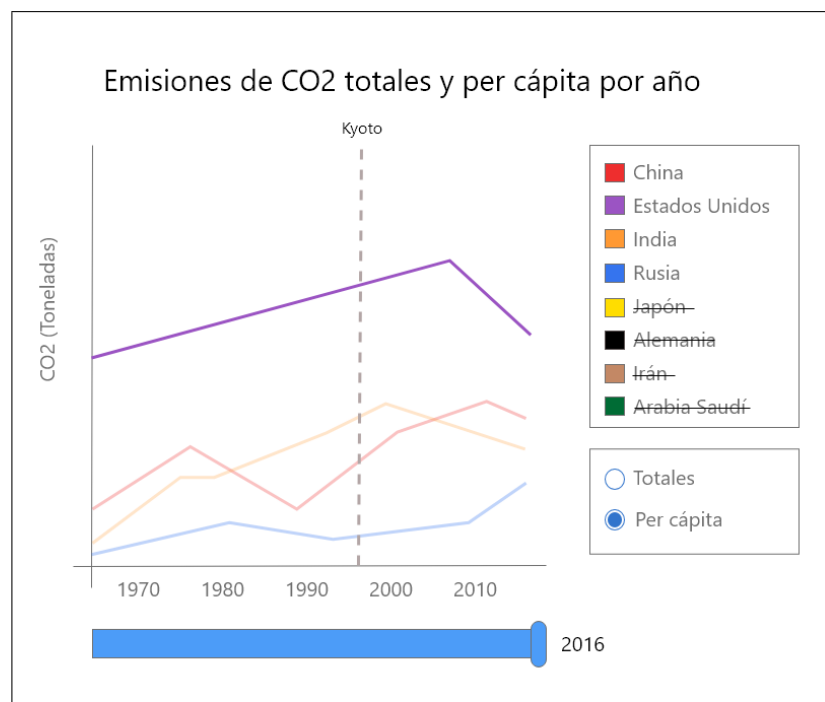


Figure 5: Visualización de emisiones per cápita resaltando un país

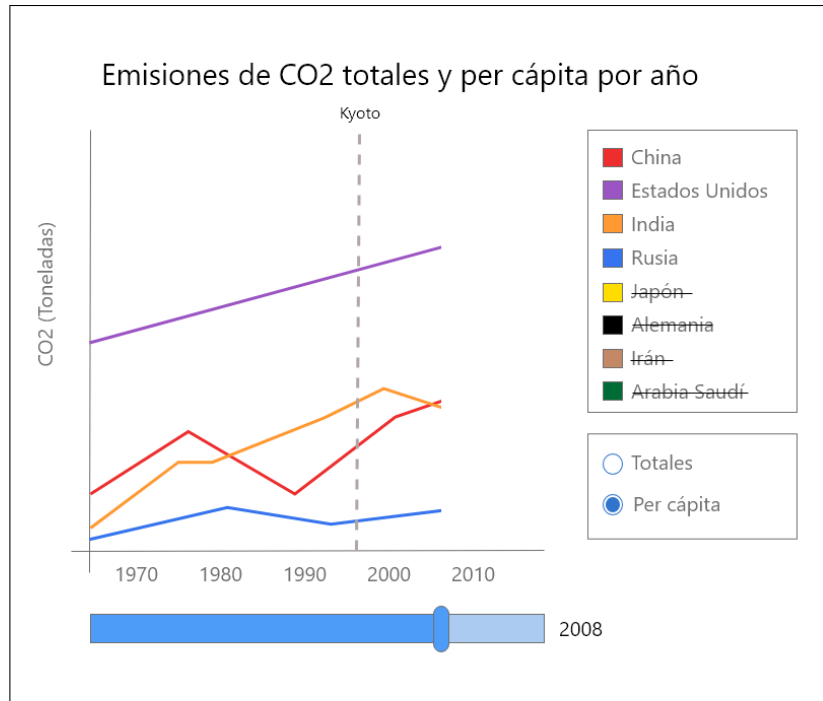


Figure 6: Visualización de emisiones per cápita limitando tiempo

5 Conclusiones

Como comentario adicional, cabe destacar que los bocetos son aproximaciones de lo que se busca y por tanto podrían existir pequeñas variaciones de lo mostrado en estos en la visualización final, que en parte pueden deberse al feedback recibido por los profesores de la asignatura. Estos están desarrollados con la herramienta de prototipado Adobe Illustrator XD, la cual no soporta subíndices y por tanto CO_2 aparece como CO2, lo cual se corregirá en la visualización final.

6 Apéndices

6.1 Apéndice 1: Análisis de los datos

El fragmento de código utilizado (lenguaje R) para el análisis de los datos es el siguiente:

```
global<-read.csv('annual-co-emissions-per-country.csv')
colnames(global)[4] <- 'annual.emissions'
perCapita<-read.csv('co-emissions-per-capita.csv')
colnames(perCapita)[4] <- 'per.capita.emissions'
```

```
# TOP 10 Países con mayor emisión de CO2
```

```

lastYear<-global[global$Year == 2016,]
orden<-order(lastYear$annual.emissions)
final<-lastYear[orden,]
n<-dim(final)[1]
final<-final[(n-20):n,]
colnames(final)[4] <- 'annual.emissions'
final<-final[final$Code != '',]
final
perCapita[perCapita$Year == 2016 & perCapita$Entity %in% final$Entity,]

#Rango variable year

c(min(global$Year),max(global$Year))
c(min(perCapita$Year),max(perCapita$Year))

#Numero de paises

length(global[! duplicated(global$Entity),]$Entity)

```

Referencias

- [1] Datos de emisiones anuales de CO_2 per cápita:
<https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions#per-capita-co2-emissions>
- [2] Datos de emisiones anuales de CO_2 :
<https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions#annual-co2-emissions>