

Proyecto de Visualización de Información con D3.js

Índice:

1. Objetivo
2. Análisis de datos
 - Adquisición
 - Tipo de los datos
 - ¿Faltan datos?
 - Mejora de la calidad
 - Transformaciones
3. Planificación de la visualización. Puntos a considerar
 - Propósito de la visualización
 - Objetivo: función y tono
 - Efectos deseados
 - Factores que afectan al proyecto
 - Enfoque
4. Diseño
 - Datos a usar
 - Formas de representación a usar
 - Decisiones de presentación
5. Implementación / Bocetos
6. Conclusiones

1. Objetivo:

Realizar una visualización de los datos de la producción energética en Castilla y León, especialmente la evolución de las energías renovables.

2. Análisis de datos:

- **Adquisición:**

Datos obtenidos de la página oficial de la Junta de Castilla y León acerca de los indicadores de sostenibilidad y aspectos ambientales, gracias a su proyecto de OpenData. Nos centraremos únicamente en los datos acerca de la producción energética.

- **Tipo de los datos:**

Los datos vienen almacenados como un fichero CSV en forma de tabla, en nuestro caso cada columna representa:

- **Indicador:** en nuestro caso es el tipo de energía sobre la que vamos a tratar. Es un valor categórico y sin orden.
- **Provincia:** es la provincia sobre la que se han hecho las mediciones. Es un valor categórico y sin orden.
- **Fecha de validez:** año en que se tomaron las mediciones. Es un valor cuantitativo (limitado a un intervalo).
- **Valor:** es la cantidad de energía del tipo seleccionado que se produjo un año en una provincia. Es valor cuantitativo, discreto y natural.
- **Unidad:** unidad en la que se proporciona el valor anterior, en nuestro caso siempre es toneladas de petróleo. Es un valor categórico y sin orden.
- **Frecuencia:** frecuencia con que se toman las mediciones, en nuestro caso siempre es anualmente. Es un valor cuantitativo, discreto y natural.

- **¿Faltan datos?**

Sí, hay ciertos tipos de energía sobre la que no hay datos en algunas provincias, como el carbón o la energía nuclear.

- **Mejora de la calidad:**

El problema anterior se puede solucionar introduciendo ceros en las provincias donde no hay datos, puesto que se supone que solo se produce ese tipo de energía en las provincias indicadas (Ej.: Solo hay centrales nucleares en Burgos).

- **Transformaciones:**

Eliminaremos los datos que sobran, hay ciertos indicadores que no son útiles para el análisis de la producción energética, como datos de la calidad del aire o la agricultura ecológica. Una vez eliminados los datos anteriores, transformaremos la columna indicador para que solo indique el nombre del tipo de energía.

3. Planificación de la visualización. Puntos a considerar:

- **Propósito de la visualización:**

Nuestro propósito es hacer más accesibles los datos de la producción de energía en Castilla y León, y centrarnos sobre todo en mostrar la tendencia que siguen estos datos hacia las energías renovables y como se reparten actualmente estas energías en la comunidad.

- **Objetivo: función y tono:**

La **función** de la visualización que propongo es la explicación de los datos, ya que es bastante probable que la gente ya conozca la situación y las tendencias de la producción energética actualmente, nuestro objetivo es hacer más accesible estos datos en nuestra comunidad.

El **tono** que buscaremos principalmente es el pragmático, ya que no buscamos en ningún momento crear emociones en el receptor, sino hacer una visualización lo más legible posible.

- **Efectos deseados:**

Los principales efectos deseados son *aumentar el conocimiento* de los datos de producción energética en la comunidad, con una visualización que sea lo más *agradable* posible para los lectores, así como *enfaticar problemas*, como las provincias donde más electricidad se produce mediante energías No Renovables.

- **Factores que afectan al proyecto:**

Usaremos D3 y javascript, con lo cual el formato que usaremos es web, en cuanto al coste, no es un factor a tener en cuenta actualmente, los principales factores a tener en cuenta son la escasez de tiempo del que se dispone para llevar a cabo las implementaciones junto con las capacidades técnicas actuales con javascript, debido a estas dos limitaciones no podemos esperar hacer gráficos excesivamente complejos.

- **Enfoque:**

Como he comentado con anterioridad, nos vamos a centrar en la evolución de la producción de energía en la comunidad en los últimos años, con un especial enfoque en cómo estas fluyen hacia las energías renovables.

4. Diseño:

- **Datos a usar:**

Debido a la cantidad de fechas, nosotros usaremos únicamente datos en intervalos de dos años, lo cual simplificará los gráficos que dibujemos.

- **Formas de representación a utilizar:**

Como se pueden extraer bastante información de los datos, y para no sobrecargar una sola gráfica, he decidido diseñar un “*Dashboard*”, en el cual dispondremos de 3 gráficas, aprovechándonos de una distribución similar a la proporción áurea, puesto que está demostrado que es más agradable a la vista. Los tipos de gráficas escogidas son:

1. La gráfica más importante será una *Radial Chart*, en el cual se muestra la información de forma concéntrica, en nuestro caso, cuanto más lejos del radio estemos, más cercano será ese dato en el tiempo (como hemos comentado con anterioridad, este se muestra en intervalos de dos años) los sectores representarán a las provincias y los colores a la proporción de energías Renovables vs No Renovables.
2. La siguiente gráfica nos permitirá ver la evolución global de la comunidad autónoma por tipo de energía. Para ello me he decantado de un *Line Chart* debido a que nos permite visualizar la evolución sin sobrecargar mucho la visualización, ya que el Radial Chart va a ser un gráfico bastante complejo. En nuestro caso, el Eje X representará el tiempo y el Eje Y la proporción de uso de esa energía en la comunidad.

3. El tercer gráfico será un *gráfico de sectores (Pie Chart)* para visualizar cuanto aporta actualmente cada provincia a la producción de energía global de la comunidad, lo he escogido debido a que permite visualizar mejor la diferencia entre los datos que otros gráficos y además al ser similar al Radial Chart mantendrá cierta uniformidad.

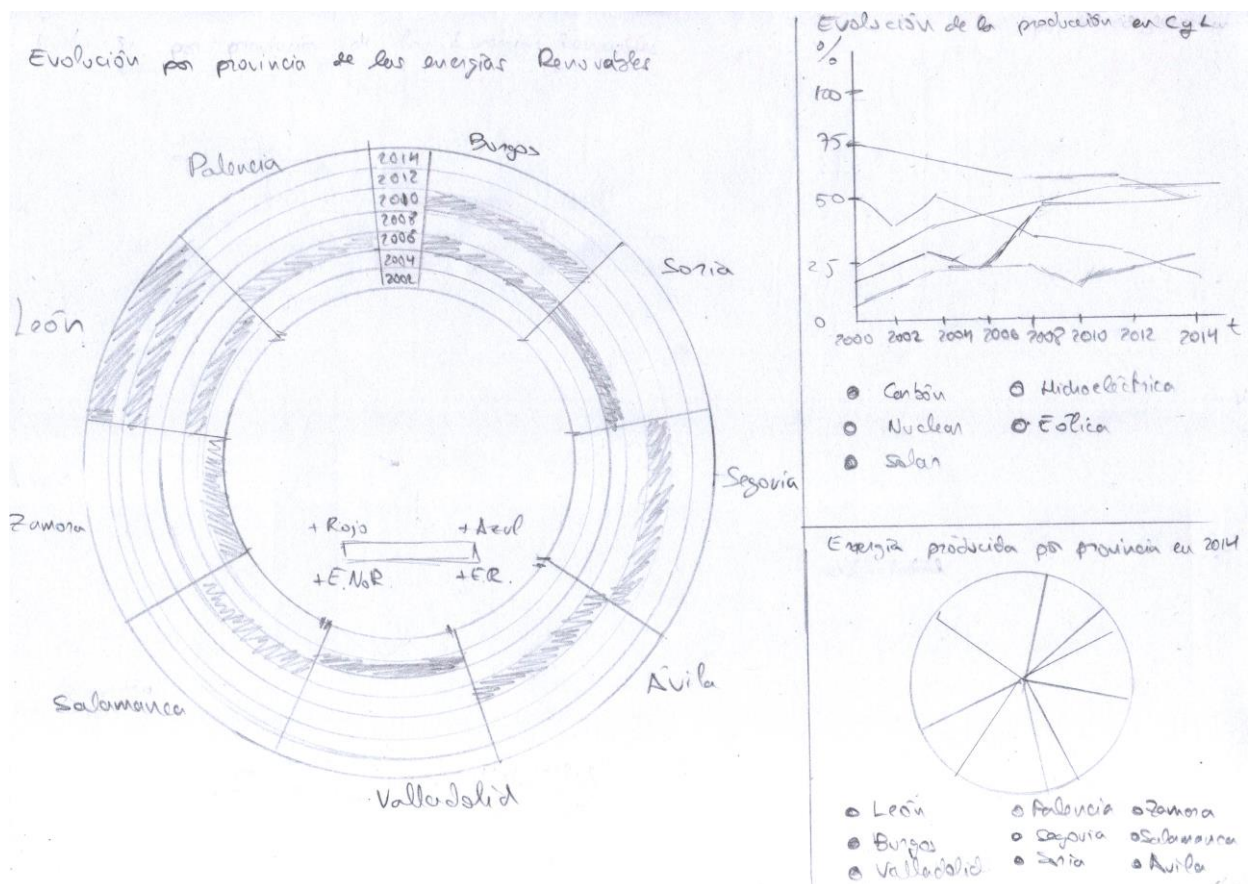
- **Decisiones de presentación:**

En cuanto a los **colores** escogidos, reservaremos los tonos rojos para las energías no renovables, y los colores azules para las energías “limpias”, ya que normalmente se asocia el color rojo a lo negativo (metáfora). Además, como azul y rojo son colores complementarios, podremos distinguir mejor las escalas y la separación entre Renovables y No renovables.

Para añadir cierta **interacción**, intentaremos que en el gráfico principal (Radial Chart), si situamos el cursor por encima de un sector (provincia), las demás provincias pasen al “background”, mandándolas al fondo gris, para poder así, ajustar el foco de atención.

Sobre **anotaciones**, intentaremos mantenerlas al mínimo para no sobrecargar la visualización, aunque manteniendo ciertos títulos y leyendas para ayudar a la comprensión de los gráficos. En el caso de los títulos como resumen, y en el de las leyendas para poder identificar los datos.

5. Implementación: boceto/s de la visualización



6. Discusión o conclusiones – Principales conclusiones, aspectos a resaltar, cuestiones no tenidas en cuenta, etc.

Como conclusión, he buscado bastante la simplicidad de los gráficos y no sobrecargar las visualizaciones, a pesar de esto sí creo que se consigue visualizar uno de nuestros objetivos, que era remarcar la evolución hacia las energías renovables, aunque puede que en los gráficos no se visualice finalmente como se ha evolucionado por provincia y tipo de energía, o no visualice cuanto aporta cada provincia en cada instante.