

Por qué el periodismo de datos no es magia

En este primer módulo van a aprender a encontrar los datos y cómo descargarlos. Y luego, en las próximas semanas, aprenderán a limpiar estos datos, cómo explorarlos, cómo transformarlos, cómo visualizarlos, cómo contar historias basadas en estos.

Es mucho trabajo, o al menos suena como un montón de trabajo.

Y una de mis experiencias como profesor de periodismo de datos y visualización de datos es que muchos estudiantes la primera vez que se enfrentan y empiezan a aprender sobre periodismo de datos se sienten un poco abrumados por la cantidad de habilidades que se requieren para producir periodismo de datos y la cantidad de pericia que se requiere.

El periodismo de datos implica aprender sobre estadística, aprender a veces sobre codificación, aprender sobre herramientas, aprender los conceptos y principios, aprender cómo hacer frente a las fuentes, cómo consultar a expertos, cómo hablar con científicos estadísticos y de datos que saben mucho más que nosotros acerca de los datos que manejamos cada día.

Pero mi objetivo aquí es principalmente decirles que comenzar no es tan difícil.

Para lograr entusiasmarlos con el periodismo de datos y la visualización de datos, les mostraré, con un ejemplo, lo fácil que es simplemente ponerse en marcha.

En 2006, Hans Rosling, un profesor de Salud Internacional en el Instituto Karolinska en Suecia, dio una charla famosa en las charlas TED, en la conferencia TED presentando datos sobre la salud mundial. Él comparó por ejemplo las expectativas de vida en diferentes países de todo el mundo a las tasas de mortalidad y a las tasas de mortalidad infantil. Se refirió a varios patrones en la demografía de todo el mundo, y en esta charla le mostró al público toneladas de hermosas visualizaciones interactivas y animadas.

Mi objetivo, en esta parte de la clase, es mostrarles qué tan fácil es hoy en día producir visualizaciones tales como las que Rosling usó en esa presentación.

La página web del profesor Hans Rosling es gapminder.org. En la sección de videos de esta página pueden ver todas las conferencias que él ha dado a lo largo de los años y también se pueden descargar algunos de los datos que ha utilizado en sus presentaciones. Para esta pequeña demostración, yo de hecho descargué un conjunto de datos de este sitio web que reproduce, o es muy similar, al conjunto de datos que él utilizó en su famosa presentación de 2006.

Así que en este conjunto de datos, que se puede descargar de la plataforma del curso, tenemos una columna con los nombres de países, otra columna con los nombres de

continentes, otra columna con los años. Y luego tenemos las tasas de fertilidad, expectativa de vida, y la población de cada uno de estos países.

La herramienta que voy a utilizar para crear una visualización es muy similar a la que Rosling utilizó en su famosa presentación. Voy a utilizar Flourish, una de las herramientas que van a aprender a lo largo de este curso. Si desean utilizar Flourish, no necesitan aprender y seguir este video, esto es apenas una demostración, pero si quieren empezar solo tienen que registrarse en la herramienta y están listos para comenzar, la herramienta es gratuita.

Así que yo ya hice eso, ya creé un perfil en Flourish. Por lo tanto, voy a ir a “nueva visualización” porque quiero crear mi propia visualización Hans Rosling sobre salud mundial. Voy a crear un diagrama de dispersión, que es un tipo básico de visualización de datos, voy a hacer clic en el gráfico de dispersión y después de haber hecho esto voy a ir a “Datos”, voy a importar mis datos y me voy a encontrar mi conjunto de datos que se llama GapminderDATA_MOOC.

Solamente subo eso, una vez lo hago, lo importo públicamente, comienzo a seleccionar los datos que voy a visualizar. Así que en primer lugar, le tengo que decir a Flourish qué quiero poner, qué quiero poner en el eje x, en el eje horizontal.

Por lo tanto en el eje horizontal voy a poner las tasas de fertilidad país por país. Entonces estará el número de columna, columna D, así que voy a escribir aquí “D”. En el eje y, que es un eje vertical, voy a poner la expectativa de vida, voy a poner la columna E. Nombre, que será el nombre de cada uno de los países, esa será la columna A. Color, quiero que estos puntos que van a aparecer aquí abajo estén coloreados por continente. Así que voy a poner la columna B, que será el color. El tamaño de las burbujas. Quiero que el tamaño de las burbujas sea proporcional a la población de cada uno de estos países, de manera que será la columna F. Y luego la forma, no voy a cambiar eso. Voy a desplazarme hacia aquí abajo y voy a ir a un campo que dice “Time” (tiempo). La columna de tiempo en mi conjunto de datos es la columna de año la cual es la columna C. Voy a hacer clic en C. Después de hacer clic en “return” (regresar), cuando vuelva a la vista previa necesito hacer un ajuste por aquí para que las burbujas sean muy, muy grandes, ¿cierto?

No se preocupen si no pueden acompañarme. Aprenderán esta herramienta más adelante a lo largo de este curso. Pero el resultado es básicamente una visualización que es casi idéntica a la que Hans Rosling utilizó en su famosa presentación.

Aquí tienen el mundo en 1952, cada una de estas burbujas corresponde a un país, el eje horizontal es el número de hijos por mujer, es decir la tasa de fertilidad. Y luego, el eje vertical es la expectativa de vida en años, la expectativa de vida al nacer. Y luego aquí tienen un deslizador de tiempo. Pueden desplazarse hacia la derecha y pueden ojear el mundo en 2012. Noten la diferencia. En 1952 la mayoría de países está por aquí, un montón de hijos por mujer. Altas tasas de fertilidad, bajas expectativas de vida. Pero luego cuando nos movemos a 2012,

muy, muy cerca al presente, notarán que la mayoría de países está más cerca de la esquina superior izquierda que significa muy altas expectativas de vida y muy bajas tasas de fertilidad.

Este es el poder del periodismo de datos, este es el poder de la visualización de datos, y también es lo fácil que es empezar en este mundo.