



Lluvia de preguntas

Usa WTFcsv para identificar qué preguntas se le pueden hacer a tus hojas de cálculo

Resumen

WTFcsv te da una visión general rápida de lo que está en tu archivo .csv. Te ayuda a pasar rápidamente a preguntarle cosas a tus hojas de cálculo que te pueden ayudar a encontrar historias que contar. Estas actividades prácticas ayudan a los participantes a familiarizarse con el manejo de datos mientras miran conjuntos de datos divertidos y sugieren preguntas que hacer sobre ellos.

Metas de aprendizaje

- Entendimiento de que hay varios tipos de datos, incluyendo números, textos y fechas.
- Habilidad mejorada de formar hipótesis sobre los metadatos y encontrar historias a partir de ellos.
- Entendimiento de que se pueden encontrar muchas historias en un conjunto de datos.
- Entendimiento de que los conjuntos de datos pueden combinarse para contestar preguntas más complicadas.

Realizar la actividad

Resolver un problema

Abre una hoja de cálculo en Excel o un documento similar y encuesta al grupo preguntándole cuántas personas trabajan con datos como éstos. Presenta el archivo CSV como una forma habitual de guardar datos de hojas de cálculo. Encontrar historias en un CSV puede ser difícil, especialmente cuando hay muchas columnas y filas. Es útil que empieces a pensar sobre las preguntas que puedes hacerle a tus datos. WTFcsv te ayuda a analizar rápidamente un CSV para entender qué hay en él y así puedas comenzar a hacerle preguntas.

Comparte ejemplos motivadores

Muestra un portal de datos de libre acceso de un lugar cerca de la ubicación del taller (por ejemplo, <http://data.gov> en los EEUU). Habla sobre la amplia variedad de conjuntos de datos que están disponibles en portales como el que estás mostrando. Descarga nuestros datos de muestra de OVNIS (<https://databasic.io/wtfcsv/ufo.csv>) y ábrelo en Excel. Muestra que cada columna tiene diferentes tipos de información sobre cada fila.

Tiempo total

30 minutos

Tamaño del grupo

3 - 100 personas. Edades 12+.

Diseñamos esto para cerca de 30 personas, pero se podría realizar con más o con menos. Está diseñado para grupos de educación media y media superior, organizaciones de noticias, organizaciones sin fines de lucro y talleres comunitarios. No se necesita experiencia previa en el manejo de datos.

Espacio

- Un proyector conectado a una computadora
- Capacidad para separarse en grupos pequeños de 3 reunidos alrededor de una computadora
- Mesas largas o piso adecuado para poner papel sobre él, o cinta para pegarlo a las paredes, para que los participantes puedan dibujar

Materiales

- Computadoras
1 por cada 3 participantes
- Copias del Folleto de Preguntas de WTFcsv
- Bolígrafos

Términos

Tipos de datos

A menudo, los datos se clasifican en tipos. Algunos tipos de datos frecuentes incluyen números, fechas y texto.

Realizar la actividad (continuado)

Presenta la herramienta

Abre WTFcsv (<https://databasic.io/wtfcsv>) con los datos de muestra "Avistamientos de OVNIS". Explica que cada tarjeta en la página de resultados es un resumen de la información en todas las filas para una columna. Haz la demostración de voltear una tarjeta para cambiar entre la información visual y los resúmenes en texto. Da ejemplos de preguntas, de otros datos que necesitarías para contestar esas preguntas y de cómo los obtendrías. Por ejemplo, una pregunta que le podrías hacer a los datos de los OVNIS es "¿Las ciudades con más gente tienen más avistamientos?" Para contestar esto, adicionalmente a los datos de los OVNIS también necesitarías obtener los datos poblacionales de todas las ciudades. Puedes obtener esta información de los registros censales de EEUU.

Da comienzo a la actividad

1. Los participantes tienen 15 minutos
2. Los participantes trabajan en equipos de tres
3. Cada equipo tiene un Folleto de Preguntas de WTFcsv
4. Cada equipo usa WTFcsv para analizar uno de los CSV de muestra y sugerir qué preguntas hacerle
5. Cada equipo escribe las preguntas más interesantes que se les ocurran

Intercambio de experiencias

Conforme rondas por los equipos, mantente atento a los diferentes tipos de preguntas que se puedan destacar. Pregunta qué grupo quiere compartir su mejor pregunta. Límtalos a 1 minuto, para que no se pierdan en el ejemplo de un grupo. Mientras los participantes intercambian experiencias, asegúrate de llamar la atención de cualquier grupo que hayas notado previamente que tenía preguntas diferentes u originales. Enfócate en estas preguntas y tópicos:

- ¿Todas las respuestas a tus preguntas están contenidas en el conjunto de datos?
Probablemente no, algunas veces necesitas encontrar otro conjunto de datos o contactar a alguien para conseguir la información que necesitas.
- ¿De dónde vienen los datos? Si la fuente de los datos no se revela, hay que ser escéptico.
Nosotros descargamos estos datos del Centro Nacional de Reportes OVNIS
- ¿Notas lugares donde faltan valores? ¿Los valores faltantes pueden hacer que los datos se "ensucien"?
Pídele a los participantes que digan formas en que los datos pueden estar "sucios", y qué podría significar que los datos estén "limpios".
- WTFcsv te muestra los resúmenes de datos agregados, pero a veces puedes preguntar cosas interesantes sobre una sola fila del conjunto de datos, o buscar un "dato atípico". ¿Los participantes encontraron historias sobre los datos atípicos?

Términos (continuado)

Metadatos

Los metadatos resumen la información básica sobre tus datos.

CSV

Un archivo de "valores separados por comas"; una forma estándar de guardar tablas de información en un archivo de texto plano. Cada fila está en una nueva línea, y cada columna está separada por una coma.

Dato atípico

Un elemento de tu conjunto de datos que es muy diferente al resto.

A recordar

- No estamos tratando de contestar las preguntas, sólo hacer una lista de ellas
- No hay preguntas tontas
- Algunos datos tienen que ser "normalizados" para poder compararlos. Por ejemplo, para comparar el número de avistamientos por ciudad debes normalizarlos por población para que puedas comparar los avistamientos por persona.



Folleto de Preguntas

Nuestra pregunta

¿Cuál es la pregunta más interesante que podemos hacer de estos datos?

Otros datos que necesitamos

¿Necesitamos algunos otros datos para responder a esta pregunta?

Fuentes de los otros datos

¿Como vamos a conseguir estos datos?