

Practica del curso Visualización de Datos

Flavia Felletti

Fuente

El dataset escogido proviene de la fuente:

Haerpfer, C., Inglehart, R., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano J., M. Lagos, P. Norris, E. Ponarin & B. Puranen (eds.). 2022. World Values Survey: Round Seven - Country-Pooled Datafile Version 5.0. Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat. [doi:10.14281/18241.20](https://doi.org/10.14281/18241.20)

Y es disponible gratuitamente por objetivos no comerciales o académicos (enseñanza e investigación). La documentación del dataset se encuentra al link:

<https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp>

En la documentación se encuentra, entre otras cosas, la correspondencia de las variables. Ej., la variable “a0001” corresponde a “importancia de la familia en la vida” (“Life Importance: Family”).

Breve descripción del conjunto de datos

El conjunto de datos proviene de un survey que se ha conducido a nivel mundial para investigar los valores culturales, la percepción y actitud hacia la familia, la religión, el género, la educación, el estado social, la tolerancia y otros factores.

El survey se ha recopilado a través de diferentes métodos, como por ejemplo entrevistas personales, encuestas postales y online. La versión utilizada en esta practica es la séptima versión del survey, que se ha publicado en el 2022.

1. [10%] Justificad brevemente vuestra selección, ya sea por motivos personales o profesionales.

El conjunto de datos se ha seleccionado por diferentes razones. Primero, el survey trata una temática de interés común: los valores de las personas en diferentes culturas. Entender los valores que las personas tienen en diferentes países del mundo nos puede ayudarnos a entender como relacionarnos con diferentes culturas y también puede facilitar los procesos de integración de inmigrantes en otros países, lo cual es muy importante en el mundo actual donde lamentablemente muchas personas se encuentran obligadas a huir de su propio país por conflictos armados, dictaduras, o situaciones económicas.

Además, el survey contiene diferentes variables de vario tipo, continuas y categóricas, lo cual nos permite decidir el enfoque de la visualización de datos y el uso de diferentes técnicas. De todas las variables disponibles, se seleccionarán algunas que sean adecuadas por las técnicas de visualización que queremos utilizar.

2. [10%] La relevancia del conjunto de datos en su contexto. ¿Son datos actuales? ¿Tratan un tema importante por algún colectivo concreto? ¿Se ha tenido en cuenta la perspectiva de género?

Los datos son actuales, y se han recogido entre el 2018 y el 2022 (la séptima edición del survey). Los datos pueden ser relevantes por diferentes colectivos que se ocupan de temas de política y de

integración cultural, pero también de antropología, y hasta nos puede ayudar a entender los cambios sociales y culturas a través de los años, si comparamos los resultados de este survey con las ediciones precedentes.

El genero ha sido uno de los factores que se han tomado en cuenta en le survey, así que es posible también comparar los valores de los hombres y de las mujeres en las diferentes culturas

3. [25%] **La complejidad (medida, variables disponibles, tipos de datos, etc.).** ¿Tiene del orden de centenares o miles de registros? ¿Tiene del orden de decenas de variables? ¿Combina datos categóricos y cuantitativos? ¿Incluye otros tipos de datos? Evitad los conjuntos excesivamente simples.

El conjunto de datos consiste de 153716 observaciones y 231 variables y combina variables de tipo categórico (ej. Género y nacionalidad) y cuantitativo. Hay variables numéricas continuas así como variables discretas. No todas las variables del conjunto de datos se utilizarán en nuestras visualizaciones, sin embargo, es un dato positivo que el dataset contenga una gran cantidad de variables para que podamos escoger las variables que consideramos más apropiadas por cada técnica de visualización.

4. [25%] **La originalidad. No repetid los conjuntos de datos clásicos.** Podéis, por ejemplo, combinar o mejorar visualizaciones existentes. ¿hay otras visualizaciones basadas en este conjunto de datos? ¿Es una evolución o actualización de un conjunto anterior? ¿Habéis enriquecido un conjunto de datos ya existente?

Un website conocido que permite al usuario de visualizar los datos basados en este dataset es: <https://www.atlasofeuropeanvalues.eu>. El usuario puede, por ejemplo, mirar mapas interactivos que muestran los valores de diferentes países. Es posible también que haya otras visualizaciones publicadas en revistas académicas de las cuales no tengo conocimiento. Sin embargo, hay muchas posibilidades de crear visualizaciones originales que no se encuentran en este sitio web, por ejemplo utilizando técnicas diferentes y más complejas.

5. [30%] **Las cuestiones que responderéis con la visualización de datos,** ¿Tienen en cuenta los puntos anteriores? ¿Están bien planteadas? ¿Son adecuadas por el conjunto de datos elegido?

En nuestro proyecto exploraremos la relación entre el sentimiento individual de felicidad, la importancia que se da a la política, la satisfacción con el sistema político, y el sentimiento de libertad y control sobre la propia vida.

Nos enfocaremos en las preguntas:

1. ¿Influye el nivel de importancia que se da a la política en el sentimiento de felicidad de un individuo?
2. ¿Influye el nivel de importancia concedido a la política en la percepción de la libertad individual?
3. ¿Cómo se correlaciona la satisfacción con el sistema político con la felicidad general?
4. ¿Qué relación existe entre la satisfacción con el sistema político y la libertad individual percibida en general?
5. ¿Están correlacionados el sentimiento de felicidad y la libertad individual percibida?