

PRAC2- VISUALIZACIÓN DE DATOS

NOMBRE: DIEGO REFOYO MATELLÁN

1 PRESENTACIÓN DE LA VISUALIZACIÓN

1.1 Título de la visualización

El título de la visualización es el siguiente:

ANÁLISIS DE LOS TRATADOS DE PAZ POR REGIÓN

1.2 URL DE LA VISUALIZACIÓN Y DEL CÓDIGO

Tanto los documentos utilizados para generar la visualización como el código empleado para ello pueden consultarse en el siguiente repositorio libre de GitHub.

<https://github.com/DiegoRM8/PeacemementAgreement>.

En esta URL se puede encontrar:

- README: un documento explicativo en el que se comenta cómo se ha elaborado la visualización y dónde se puede consultar.
- Código: carpeta que contiene el código del software R con el que se ha tratado la base de datos original hasta obtener la deseada.
- Visualizaciones: url donde se puede consultar la visualización.
- Documentos: carpeta que contiene los datasets utilizados y un pdf en el que se encuentra una definición más extensa de las variables.

1.3 DESCRIPCIÓN DEL DOCUMENTO

El documento original, del cual se parte, se llama PEACE AGREEMENTS DATABASE AND DATASET V4, y ha sido descargado de la siguiente URL (<https://www.peaceagreements.org/search>) . Dicho dataset, contiene información acerca de los metadatos de diversos tratados de paz. En él se puede encontrar información acerca de los tratados de paz tales como la región sobre la que trata o el estado del mismo. A continuación, se ofrece una breve descripción de las variables utilizadas para realizar las visualizaciones.

- Region: Región en la que el conflicto está teniendo lugar.
- Date: fecha en la que el tratado fue firmado.
- Agtpt: variable que describe la naturaleza primaria del acuerdo y el conflicto.

- Contp: Tipo de conflicto (gubernamental, territorial, gubernamental-territorial, entre grupos, otros).
- Stage: estado del tratado.
- StSym: variable binaria que toma valor 1 cuando en el acuerdo se hace referencia a algún símbolo de estado (0 en caso contrario).
- GeLgbtPos: variable binaria que toma valor 1 cuando en el tratado se hace una mención positiva al colectivo Lgtbi (0 en caso contrario).
- GOthAntid: variable binaria que toma valor 1 cuando en el tratado se compromete a tomar medidas contra la discriminación (0 en caso contrario).
- GRefOth: variable binaria que toma valor 1 cuando en el tratado se compromete a tomar medidas favorables a los refugiados (0 en caso contrario).

Para más información acerca de estas variables o del resto que aparecen en el dataset consultar el documento “PEACE AGREEMENTS_V4.pdf” adjunto en la carpeta .zip.

El objetivo de las visualizaciones que se realizan a partir de este documento es aportar información al usuario acerca de los tratados de paz. Tiene como eje central la región a la que pertenece cada uno de ellos, agrupándolos por esta variable principalmente. Pero, además de esto, permite al usuario elegir entre un abanico de opciones, las otras variables con las que complementar el estudio. En esta práctica final, se incluyen dos visualizaciones:

- Visualización 1: esta visualización se centra en ofrecer al usuario una visión del número de caracteres de los tratados agrupados por la región a la que pertenecen. La parte interactiva de la visualización, además de diversos filtros, consiste en elegir la otra variable por la que se dividirá cada región. Se ofrecen tres posibilidades: Agtp, Contp y Stage. La visualización consiste en un gráfico de barras agrupado.
- Visualización 2: en esta visualización se continúa agrupando los tratados dependiendo de la región de la que se trate. En este caso, el estudio consiste en analizar cuántos de esos tratados hacen referencia a diferentes temas de interés en la actualidad. Se ofrece la posibilidad al usuario de elegir entre GeLgbtPos, GOthAntid, GRefOth y StSym ya que se considera que son temas de mucha importancia a nivel social.

2 Explicación de la visualización

2.1 ¿Qué preguntas responde la visualización?

Las preguntas que responde la primera visualización podrían ser:

- ¿Qué tipo de conflicto presenta mayor número de caracteres en la región de America?
- ¿En qué región los tratados que se encuentran en estado “Cea” presentan un mayor número de caracteres?

- ¿De los conflictos de origen interno, cuál presenta mayor número de caracteres para la región de Europa y Euro-Asia?

Para esta primera visualización, se ha elegido un gráfico de barras agrupadas ya que se ha considerado que era el más ideal, debido a que se quiere representar una variable numérica (Nº de caracteres del tratado) en función de una variable cualitativa (Región) y, dentro de esta, agrupada por otra variable cualitativa (elegida por el usuario).

La segunda visualización podría responder a las siguientes preguntas:

- ¿En qué región se hacen más referencias a símbolos del estado?
- ¿Cuántas menciones positivas se hacen al colectivo LGTBI en la región de Asia?
- ¿Se hacen más referencias a medidas favorables a los refugiados en los tratados de Europa y Euro-Asia o en los de América?
- ¿En cuántos tratados se comprometen a tomar medidas contra la discriminación en total de todas las regiones?

Para esta segunda visualización, se ha elegido un gráfico de burbujas agrupadas ya que el número de regiones no es muy elevado y esto permite entender fácilmente el gráfico. Para la variable elegida por el usuario, la región con mayor número de referencias a esa variable en los tratados será la que tenga una burbuja más grande.

2.2 Usos miedo un usuario tipo imaginario

Las dos visualizaciones podrían servir a muchos tipos de usuarios, como por ejemplo, a un estudiante que esté realizando una investigación para la universidad acerca de la concienciación social sobre el colectivo LGTBI.

3 Descripción técnica del proyecto

El proyecto se ha dividido en dos fases:

- Primera fase: preparación del dataset para poder ser utilizado posteriormente. En esta fase, se ha realizado una comprobación de que todas las variables se encuentran en el formato deseado, de que no existen valores perdidos ni valores extremos que no tengan coherencia. Para poder llevar a cabo esta tarea, se ha utilizado el software R, el cual permite realizar una limpieza y preparación de los datos de manera rápida, sencilla y eficaz. El código empleado en R puede consultarse en la correspondiente carpeta “Código”, del enlace de Github.
- Segunda fase: elaboración de las diferentes visualizaciones. Una vez preparada la base de datos con la que se va a trabajar, el siguiente paso ha sido realizar las visualizaciones mediante el software Tableau. Se ha elegido este software ya que es uno de los más potentes a nivel de visualización y permite depositar las visualizaciones en un servidor público y de acceso

gratuito. Para elaborar las visualizaciones, ha sido necesario generar alguna variable auxiliar que no estaba contenida en el dataset original ya que se buscaba conseguir cierta interactividad para con el usuario final. Des esta forma, se permite al usuario elegir entre diferentes variables para adaptar las visualizaciones a sus necesidades. Además, las dos visualizaciones generadas con Tableau, pueden ser consultadas desde un dispositivo móvil, sin perder legibilidad.

4 Visualización de datos

Las visualizaciones creadas pueden ser consultadas en la carpeta visualizaciones, mediante los enlaces que aparecen en el archivo visualizaciones.txt.