

# Proyecto de visualización de datos

## PEC4: Creación de la visualización (A9)

Luis Antón López Sobrado

luilop@uoc.edu

### 1. Título de la visualización

El título escogido para la visualización realizada es: **Reconstrucción económica de los acuerdos de paz.**

### 2. URL de la visualización y del proyecto

La visualización la podemos ver en Tableau Public a través de la siguiente url: [https://public.tableau.com/profile/luis.l.perez#!/vizhome/PEC4\\_15896411188850/PEC3](https://public.tableau.com/profile/luis.l.perez#!/vizhome/PEC4_15896411188850/PEC3).

En cuanto a la información técnica del proyecto (documentos, datos, README y licencia) se encuentra disponible en el repositorio: <https://github.com/luisllopez/PeaceAgreements>

### 3. Descripción

En este proyecto nos hemos centrado en estudiar y analizar los datos de la reconstrucción socioeconómica de las diferentes regiones a los acuerdos de paz después de un conflicto. En la visualización vemos cómo varían los acuerdos de paz entre las regiones según la fase, el tipo o naturaleza de conflicto, las cuales tratamos de reflejar mediante unos conjuntos de gráficas en la visualización realizada. Además de comparar los datos, localizarlos, o detectar posibles patrones en relación a los acuerdos de reconstrucción socioeconómica de las diferentes regiones a los acuerdos de paz, según la fase, el tipo o la naturaleza de conflicto, llegando a ser útil a los integrantes del proyecto o stakeholders para la toma de decisiones.

### 4. Proyecto.

Este fue el primer proyecto de visualización de datos que hemos realizado, en el cual hemos encontrado algunas dificultades y limitaciones al tratar con un alto volumen de datos y de gran complejidad. Por ese motivo nos hemos centrado en estudiar y analizar los datos de la

reconstrucción socioeconómica de las diferentes regiones a los acuerdos de paz después de un conflicto.

Además el hecho de trabajar con un software como *Tableau Desktop* añadió una dificultad extra al nunca haber trabajado con dicha herramienta. Durante la fase de aprendizaje en la PEC2, desconocíamos realmente todo el potencial de la herramienta. Una vez aprendido a utilizarla fue de gran utilidad tanto para la exploración de datos, como para la construcción de la visualización.

Con respecto a la PEC2 realizada, las hojas y tipos de hojas son las mismas que en la PEC3 y PEC4, pero diferente en cuanto a el numero de datos. Durante en desarrollo de la PEC3 observamos que para la representación de los datos en el mapa utilizando como base la variable *Con*, se producían valores nulos a raíz de que se muestran varios países en los que tuvo lugar el conflicto en el mismo campo separados por barras laterales. Una vez aplicadas las transformaciones correspondientes, las cuales explicaremos en mayor detalle en el siguiente apartado, generamos una nueva variable llamada *Country*. Esta se utilizó para representar los datos en el mapa según la naturaleza y el tipo de conflicto, así como en el resto de gráficas mostrando mayor información y datos de otros países que hasta la fecha estaban ocultos en la visualización, evitando así obtener o tomar decisiones erróneas.

No obstante pese a todas estas dificultades, hemos adquirido una gran cantidad de conocimientos tanto a nivel de diseño y análisis, así como para la elaboración de una visualización. La exploración de los datos nos permitió descubrir muchas relaciones y tendencias interesantes, además de aprender lo importe que resulta durante la fase de exploración centrar nuestros esfuerzos en un único aspecto que sea importante, o bien que se quiera realizar un estudio.

En general la visualización realizada nos permite mostrar de manera clara patrones y tendencias de los datos que queríamos estudiar. Al tratarse de nuestro primer proyecto en esta disciplina, hemos tenido que invertir muchos esfuerzos y dedicar muchas horas que no siempre han resultado productivas. Sin embargo nos han permitido entender que los datos son esa información poco amigable, donde entenderlos mediante una correcta visualización mejora experiencia como usuarios, desarrollador o cliente, y nos ayuda a tomar decisiones de manera correcta.

## 5. Descripción técnica del proyecto: lenguajes, librerías, licencias, descripción técnica del proyecto.

El conjunto de datos utilizado (pax\_data\_933\_agreements\_14-04-20.xlsx) contiene 933 filas o registros de datos con los acuerdos de paz firmados en todo el mundo desde 1990 y, 265 variables o columnas (país, región, fecha, naturaleza y tipo de conflicto, tipo de acuerdo, fase del acuerdo de paz, Loc1ISO, Loc2ISO, etc.). El significado de cada una de las variables lo podemos encontrar en PA-X codebook Version3.

Para la construcción de la visualización nos centramos como medida en el número de registros (variables numéricas), y como dimensiones para nuestro estudio utilizamos las siguientes variables categóricas: Con, Reg, Dat, Contp, Agtp, Status, Stage, Loc1ISO y Loc2ISO.

Inicialmente para la representación de los datos en el mapa inicialmente pensamos en utilizar la variable Con, el problema de esta variable es que produce valores nulos a raíz de que se muestran varios países en los que tuvo lugar el conflicto en el mismo campo separados por barras laterales. Viendo el problema con la variable Con, nos fijamos en el código ISO de los países que firman el acuerdo. Estos se encuentra distribuido en dos columnas: Loc1ISO y Loc2ISO.

El atributo Loc1ISO contiene 18 valores vacíos, mientras que el atributo Loc2ISO tiene 112 valores correspondientes a conflictos donde interviene más de un país; además podemos ver que hay conflictos entre más de dos países que no contienen todos los códigos correspondientes. El primer paso consintió en insertar todos los códigos que faltaban en Loc1ISO y Loc2ISO teniendo en cuenta los datos del atributo Con; en segundo lugar para los conflictos entre países fue necesario añadir hasta 7 columnas nuevas para almacenar cada uno de los códigos (Loc13SO,..., Loc19SO). Por último necesitamos modificar la estructura de los datos para tener el código ISO en una única columna, es decir, duplicar la fila y añadir una nueva columna Country con los datos de cada país que forma parte en el conflicto para tenerlo así en cuenta en el total de los registros. De esta manera es como la variable Country nos permitió representar los datos en el mapa según la naturaleza y el tipo de conflicto.

Para todas estas transformaciones utilizamos LibreOffice Calc realizando la primera exploración de los datos, los atributos disponibles y sus posibles valores, además de transformar el documento de

formato .csv a .xlsx. Posteriormente para el análisis visual de los datos utilizamos Tableau Desktop. Con esta herramienta asignamos la función geográfica (País/Región) a la variable Country para poder visualizar los datos en el mapa y, en segundo lugar, transformamos el campo fecha de tipo cadena a tipo fecha, para poder visualizar la evolución de los acuerdos con el paso del tiempo en función de la naturaleza del conflicto.

Una vez efectuado todo el proceso de limpieza y transformación de datos obtuvimos un nuevo archivo xlsx ( `pax_data_933_agreements_08-05-20_modificado.xlsx`) con datos de las variables: Contp,Agtp,Status,Stage,Loc1ISO,Loc2ISO,Loc3ISO,Loc4ISO,Loc5ISO,Loc6ISO,Loc7ISO,Loc8ISO,Loc9ISO,Country,Total\_Loc\_ISO,Número registros.

Estos datos del dataset son los que utilizamos para construir las diferentes gráficas que explicamos en la PEC3, mediante la herramienta Tableau Desktop donde cambiamos los tipos de las variables, construimos las diferentes hojas con las gráficas y sus filtros, para terminar generando el dashboard con la visualización.

Para que el usuario pudiera explorar a fondo los datos, y así poder realizar comparaciones y descubrir tendencias o diferencias importantes, se añadieron a través de Tableau Desktop una serie de filtros e interacciones. Inicialmente en cada uno de las gráficas muestran por defecto todos los registros de datos indiferentemente de la naturaleza o el tipo de conflicto. No obstante, estos resultados se pueden acotar a través de leyendas de colores, así como de los paneles de filtrado, donde el usuario puede seleccionar una opción del conjunto predefinido de opciones. Teniendo en cuenta que ambas opciones aplican el filtro a todas las gráficas. Además para una mayor información, al situar el cursor sobre un país, columna, etc. Se mostrará una ventana con un mayor número de detalle.

Por último, este trabajo lo hemos protegido con una licencia MIT. Esta licencia permite la distribución, modificación y uso comercial o privado de esta visualización.