

Proyecto de visualización de datos
Derechos humanos e igualdad en los acuerdos de paz

ALUMNO: PABLO MAS CAYUELAS
FECHA: 10/06/2020

1. INTRODUCCIÓN	2
El título de esta visualización es “ACUERDOS DE PAZ: DERECHOS HUMANOS E IGUALDAD”.	2
2. CONTENIDO ESPECÍFICO PEC 2	3
2.1. DESCRIPCIÓN DEL DOCUMENTO	3
2.2. PRESENTACIÓN DE LOS DATOS	3
2.3. EXPLORACIÓN DE LOS DATOS	6
2.4. LOS DATOS, PROCEDIMIENTOS Y HERRAMIENTAS	7
2.5. VISUALIZACIÓN SOBRE LOS DATOS	8
3. CONTENIDO ESPECÍFICO PEC 3	15
3.1. ABSTRACT/RESUMEN	15
3.2. INTRODUCCIÓN	15
3.3. DATOS	17
3.3.1. Fuentes de datos	17
3.3.2. Limpieza y transformación de los datos	18
3.3.3. Información utilizada de la PEC2	19
3.3.4. Tipos de relaciones entre los datos	20
3.3.5. Gráficos elegidos	20
3.3.6. Gobernanza de datos	21
3.4. DISEÑO	21
3.4.1. Colores	21
3.4.2. Tipografía	22
3.4.3. Estructura y bloques	22
3.4.4. Mockup	23
4. PROYECTO FINAL	25
4.1. EXPLICACIÓN DEL PROYECTO FINAL	25
4.2. DESCRIPCIÓN TÉCNICA DEL PROYECTO	26
4.3. DESCRIPCIÓN DE LOS GRÁFICOS	27
5. BIBLIOGRAFÍA	28

1. INTRODUCCIÓN

El título de esta visualización es “ACUERDOS DE PAZ: DERECHOS HUMANOS E IGUALDAD”.

Este trabajo muestra el proceso de elaboración de una visualización de datos a partir de la información de los acuerdos de paz de todo el mundo proporcionado por el [proyecto](#) de acuerdos de paz de la Universidad de Edimburgo. La visualización se encuentra publicada en [Tableau Public](#), y su objetivo es representar el número de acuerdos de paz que incluyen mención a los derechos humanos e igualdad, mostrando aspectos auxiliares como el número de acuerdos alcanzado por año, el estado en el que se encuentran dichos acuerdos o la correlación entre diversas variables.

En el presente documento se muestra un recorrido de todas las fases del proyecto: en primer lugar (contenido específico de la PEC2) se muestran las primeras consideraciones tenidas en cuenta para la elaboración del mismo: las variables escogidas y primeras representaciones. En segundo lugar (contenido de la PEC3) se realiza una segunda criba de los atributos y se seleccionan aquellos que son útiles para nuestro propósito. Por último (información final del proyecto) se realizan unas últimas modificaciones y se elabora la representación en *Tableau* junto a las transformaciones de datos realizadas.

2. CONTENIDO ESPECÍFICO PEC 2

2.1. DESCRIPCIÓN DEL DOCUMENTO

En este trabajo se analizan los acuerdos de paz relacionados con los derechos humanos e igualdad, en cada uno de los territorios y regiones del planeta. De esta manera, se tratará de dar una idea de cuáles son los temas que más tienen en cuenta estos continentes (por ejemplo, si el acuerdo está relacionado con las clases raciales o no). Para ello, se han extraído más de mil documentos en la web www.peaceagreements.org, asociando a cada uno de ellos más de doscientos atributos.

Así, la estructura del documento es de la siguiente manera: en primer lugar se presentan los datos distinguiéndolos por tipo y estructura (al menos para los atributos que queramos destacar de la gran cantidad de datos). Tras ello, se intentará evidenciar las principales tendencias localizadas en los datos escogidos, exponiendo las principales relaciones entre los atributos anteriormente descritos. Por último, se explicarán los métodos utilizados para la extracción de las conclusiones anteriores (en este caso, la utilización del programa Python), y se mostrarán algunas visualizaciones sencillas que evidencien esto.

2.2. PRESENTACIÓN DE LOS DATOS

En el conjunto de datos escogido se tienen en cuenta un total de 1249 documentos, de los que se han extraído 265 atributos. Por tanto, se cuenta de 1249 filas y 265 columnas. Estos atributos incluyen características generales y categóricas del acuerdo, como el nombre, el país o los países que intervienen en el mismo o el tipo de acuerdo (en este caso el tipo es exclusivamente “derechos humanos e igualdad”). Las variables categóricas se han utilizado casi en su totalidad para este análisis.

En cambio, la mayoría de los atributos son numéricos, y hacen referencia al grado de relación del acuerdo con temáticas o grupos concretos, como pueden ser los grupos raciales o si tienen en cuenta o no la asistencia humanitaria. Por tanto, estos atributos numéricos servirán para relacionarse con la variable “región”, lo que se podrá observar en las visualizaciones.

En este análisis se ha escogido las variables numéricas que más variedad tienen en base al tema del que tratan, y otras que han resultado interesantes para el análisis, siendo prácticamente imposible agrupar el total de más de 240 variables de este tipo, además de contraproducente. Las variables categóricas y numéricas utilizadas se describen en la siguiente tabla:

VARIABLES CATEGÓRICAS DEL ACUERDO

Con	Nombre del país o países que intervienen en el acuerdo
Reg	<p>Región asociada a los países que intervienen (normalmente continentes):</p> <ul style="list-style-type: none"> • <i>África (excl MENA)</i> • <i>Américas</i> • <i>Asia y Pacífico</i> • <i>Cross-regional</i> • <i>Europa y Eurasia</i> • <i>África del norte y del este</i>
Dat	Fecha del acuerdo en formato AAAA-MM-DD.
Agtp	<p>Naturaleza del conflicto por el que se establece el acuerdo:</p> <ul style="list-style-type: none"> • <i>Inter</i> (entre estados) • <i>InterIntra</i> (entre estados pero referido a asuntos propios) • <i>Intra</i> (en un solo estado) • <i>IntraLocal</i> (problemas locales de un solo estado)
Contp	<p>Tipo de conflicto por el que se establece el acuerdo:</p> <ul style="list-style-type: none"> • <i>Government</i> (político o ideológico) • <i>Territorial</i> (territorial) • <i>Government/territory</i> (afecta a ambos) • <i>Inter-group</i> (conflictos entre grupos no estatales)
Status	<p>Tipo de acuerdo establecido:</p> <ul style="list-style-type: none"> • <i>Multiparty signed</i> (firmado por varios estados) • <i>Unilateral agreement</i> (formado por un estado como respuesta a un acuerdo de otro estado) • <i>Status unclear</i> (no está claro quién lo firma) • <i>Agreement with Subsequent Status</i> (acuerdo no formal)
Stage	<p>Fase en la que se encuentra el acuerdo de paz:</p> <ul style="list-style-type: none"> • <i>Pre</i> (prenegociación) • <i>SubPar</i> (acuerdo de algunas cuestiones) • <i>SubComp</i> (se acuerdan temas esenciales) • <i>Imp</i> (implementación de un acuerdo anterior) • <i>Ren</i> (renovación de un acuerdo anterior) • <i>Cea</i> (el acuerdo está ligado a un alto al fuego)

	<ul style="list-style-type: none"> • <i>Other</i> (otra fase)
StageSub	Dentro de las anteriores fases hay subfases, que no nos detenemos a explicar en la presente tabla por su extensión.

VARIABLES NUMÉRICAS ESCOGIDAS

EqGen	<p>Indica si el acuerdo incluye referencias a la igualdad o no discriminación. Toma valores enteros entre 0 y 3, asociando por medio de estos números el grado de referencias realizadas.</p> <ul style="list-style-type: none"> • 0 (ninguna mención) • 1 (mención breve) • 2 (mención con detalles) • 3 (detalles concretos, como plazos establecidos y planes elaborados)
HrDem	Indica si el acuerdo incluye referencias a la democracia, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
Prot	Indica si el acuerdo incluye referencias a medidas de protección, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
HrFra	Indica si el acuerdo incluye referencias a un marco de referencia de derechos humanos, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
GRa	Indica si el acuerdo incluye referencias a los diferentes grupos raciales de la población, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
GRe	Indica si el acuerdo incluye referencias a los diferentes grupos religiosos de la población, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
GCh	Indica si el acuerdo incluye referencias a los niños, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
DevHum	<p>Indica si el acuerdo hace referencia a la asistencia humanitaria:</p> <ul style="list-style-type: none"> • 0 (no se menciona) • 1 (se menciona)

Ele	Indica si el acuerdo hace referencia a la celebración de elecciones, tomando valores enteros entre 0 y 3, dependiendo del grado de la mención.
-----	--

2.3. EXPLORACIÓN DE LOS DATOS

En un primer lugar, para visualizar de una forma más general los datos, podemos realizar una representación del número de acuerdos relacionados con derechos humanos e igualdad en cada año desde que se tienen datos (Figura 1). Podemos comprobar que el año en el que más acuerdos de paz se consiguieron en este sentido fue el año 1994. El segundo pico lo encontramos en los años 1999 y 2014. Por otro lado, el año 2019 fue el más débil en este número, lográndose alrededor de diez acuerdos, frente a los setenta que se consiguieron en 1994.

Por otro lado, en la Figura 2, podemos comprobar que la gran mayoría de los acuerdos se corresponden con el *Status* “*Multiparty signed*”, es decir, firmado por varios estados. De hecho, corresponden al 92.23 % de todos los acuerdos de paz relacionados con derechos humanos e igualdad. Además, si representamos el tipo de conflicto que se quiere resolver con los acuerdos, se aprecia que el 85 % de los acuerdos se corresponden con conflictos políticos o políticos y territoriales (Figura 3). También podemos representar el estado en el que se encuentra el acuerdo (Figura 4), viendo así que la mayoría se encuentran, o en la prenegociación, o con algunas ideas en común. Además, dentro de cada fase, se puede representar en qué subfase se encuentra el estado; por ejemplo, podemos representar las subfases asociadas a la fase de prenegociación (Figura 5).

Podemos distinguir otro tipo de visualizaciones, correspondientes a la representación de atributos en función de la región. Por ejemplo, en la Figura 6 se aprecia el grado de mención a la igualdad por cada región. Así, podemos ver que América es la que más menciona este tema, mientras que los acuerdos que se dan entre varias regiones (*cross-regional*) es la que menos. También podemos representar el grado de democracia (Figura 7), las medidas de protección (Figura 8) y las alusiones al marco de referencia de los derechos humanos (Figura 9), aunque no son demasiado significativas. Diferencias más significativas encontramos en la naturaleza del conflicto por el que se establece el acuerdo (Figura 10). En ella apreciamos que, mientras las regiones africanas, americanas, asiáticas y pacíficas destacan por conflictos de naturaleza intraestatal (es decir, en el propio estado), la mayoría de los acuerdos firmados entre varias regiones son de carácter interestatal (como es de esperar). Además, la región del este y del norte de África es la que más variedad tiene en esta naturaleza, con un buen número de acuerdos intraestatales e intralocales (entre estados pero referentes a asuntos propios).

Así mismo, en la Figura 11, se contempla la mención o no en los acuerdos por región de la asistencia humanitaria. Así, vemos que los acuerdos realizados entre varias regiones no hacen referencia alguna a la asistencia humanitaria, mientras que Europa y Eurasia es la región que más referencia hace a ello (el 23.21 % la contempla en sus acuerdos).

Por último, podemos realizar representaciones cruzadas de los atributos de nuestra base de datos, sin tener en cuenta la región a la que pertenece el acuerdo. Así, por ejemplo, en la Figura 12, vemos que la mayoría de los documentos que mencionan a los diferentes grupos raciales también mencionan a los grupos religiosos (el 63.93 %), mientras que la ausencia de mención de un grupo también implica la ausencia de mención del otro en el 84.98 % de los casos. De esta misma forma, también podemos representar (Figura 13) de manera cruzada la incidencia en grupos de niños en función de los grupos raciales, evidenciando que la mayoría de los acuerdos que contemplan a los niños también son acuerdos que contemplan a los diferentes grupos raciales.

Adicionalmente, podemos ver (Figura 14) la relación entre la mención de la igualdad en los acuerdos con la mención de la democracia. Así, la mayoría de los acuerdos cuyo grado de mención a la igualdad es máximo, se corresponden también con un alto grado de mención a la democracia (el 36.67 %). Todavía más acusada es la correlación entre la mención a un marco de referencia de los derechos humanos con la mención a posibles medidas de protección; en la Figura 15 se aprecia que casi el 70 % de los acuerdos en los que se menciona abundantemente las medidas de protección, también se menciona con énfasis un marco de referencia de los derechos humanos.

2.4. LOS DATOS, PROCEDIMIENTOS Y HERRAMIENTAS

Para realizar estas primeras visualizaciones se ha utilizado el programa Python, y en concreto el paquete *Seaborn*. En primer lugar, tras importar las librerías necesarias, se ha cargado todos los datos en un archivo *pandas*, el cual se ha ido modificando para obtener los resultados pertinentes. Además, se ha hecho una criba de los atributos a tener en cuenta; se han escogido la mayoría de los atributos correspondientes a variables categóricas, mientras que la gran mayoría de los atributos correspondientes a variables continuas se han eliminado (para simplificar la base de datos, puesto que cuenta con más de 250 atributos).

Algunos de las variables numéricas se han escogido debido a la relación con la base de datos concreta que tenemos, que son los acuerdos de paz asociados a los derechos humanos e igualdad. Por tanto, hemos escogido los atributos *EqGen*, *HrDem*, *Prot* y *HrFram* arriba explicados por su relación con el tema. Los demás atributos numéricos se han escogido en base a la suma de sus valores en toda la columna. Escogiéndolos de esta manera nos aseguramos de que nos van a dar variedad en nuestras representaciones, puesto que un atributo cuya columna sume un valor grande no va a tener excesivos valores nulos y nos va a ser útil.

Para la primera de las representaciones (acuerdos realizados por año, *Figura 1*), se ha tenido que separar del atributo *Dat* las cuatro últimas cifras, para crear un nuevo atributo, *Year*, y poder realizar esta representación.

Todas las Figuras mostradas en el siguiente apartado están realizadas con la librería *Seaborn*, y el comando *catplot*, que es bastante versátil en estas representaciones y cuyos resultados son bastante estéticos. El procedimiento seguido, además, se ha basado en la elección de aquellas visualizaciones más descriptivas, donde se aprecie relaciones directas entre los atributos. Por tanto, se han descartado varias visualizaciones en base a este criterio.

2.5. VISUALIZACIÓN SOBRE LOS DATOS

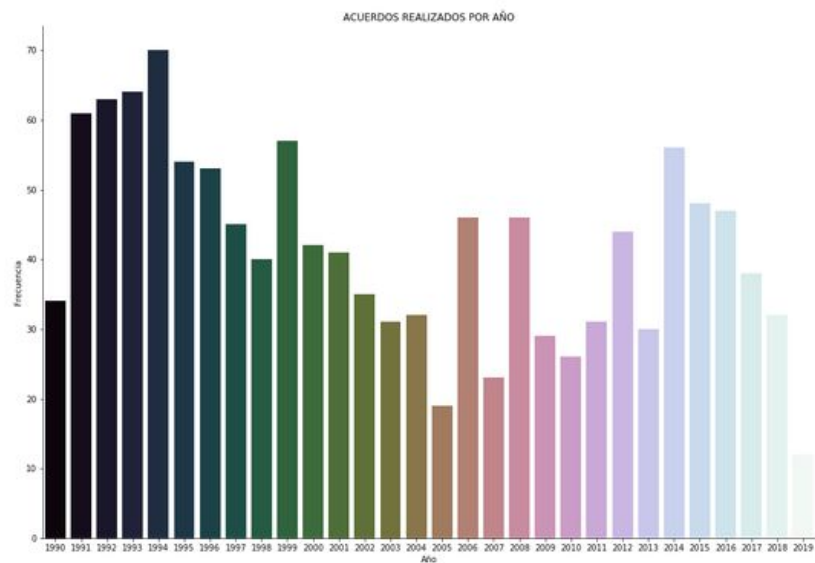


Figura 1

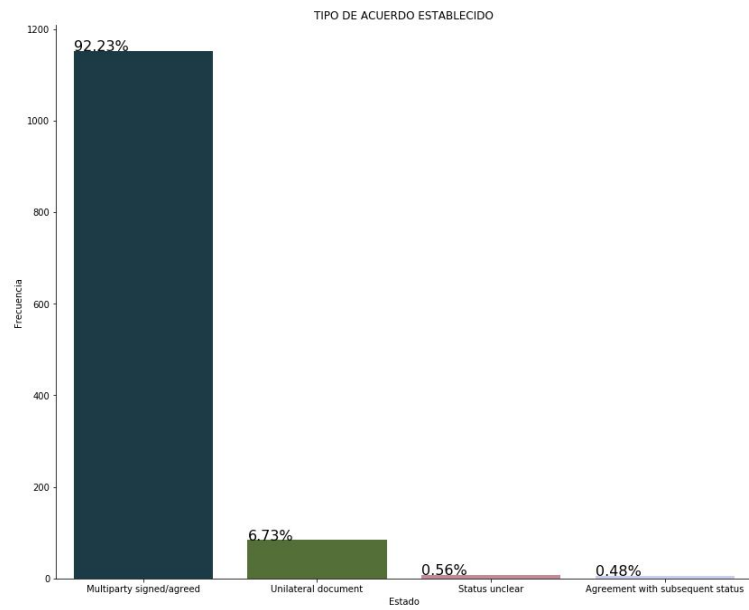


Figura 2

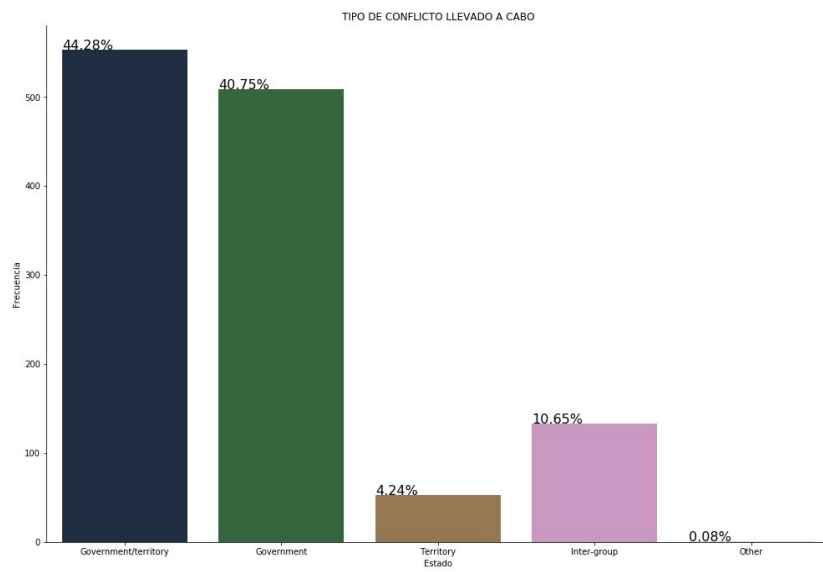


Figura 3

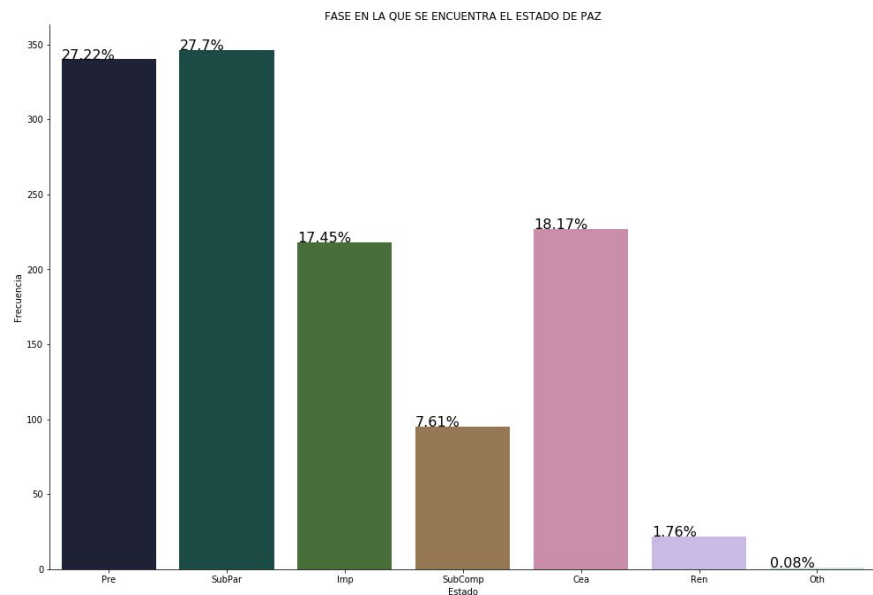


Figura 4

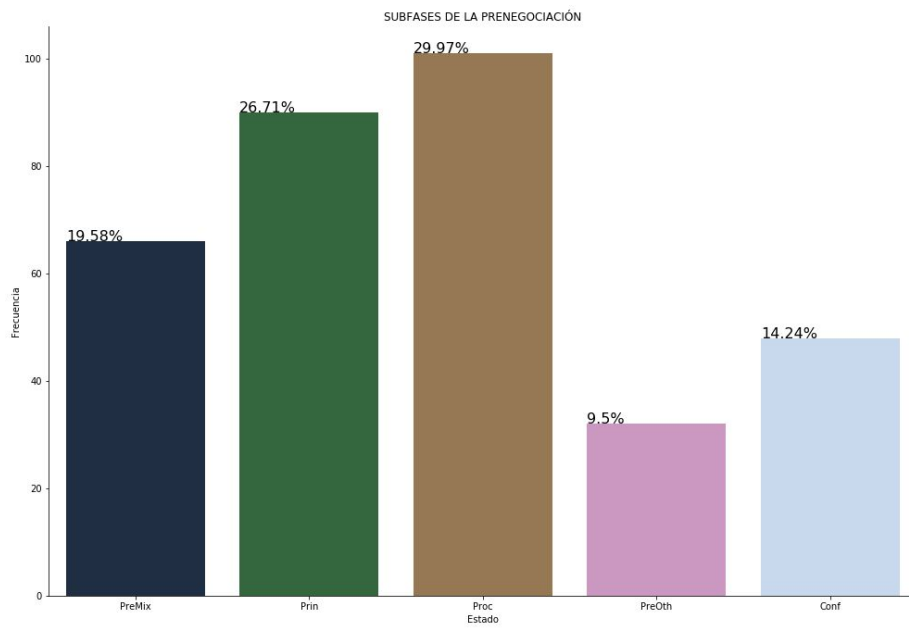


Figura 5

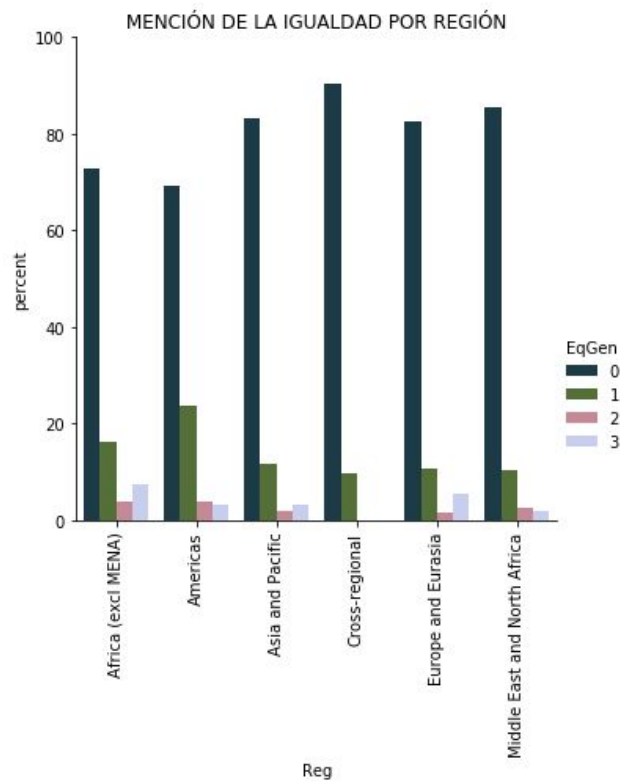


Figura 6

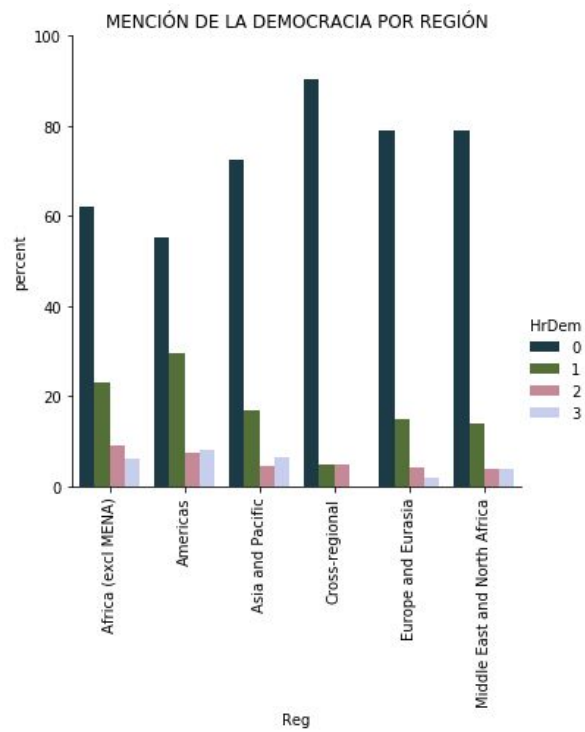


Figura 7

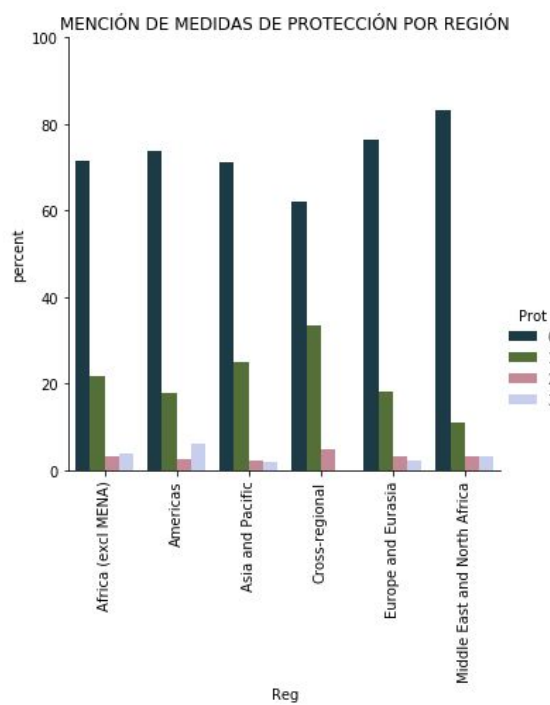


Figura 8

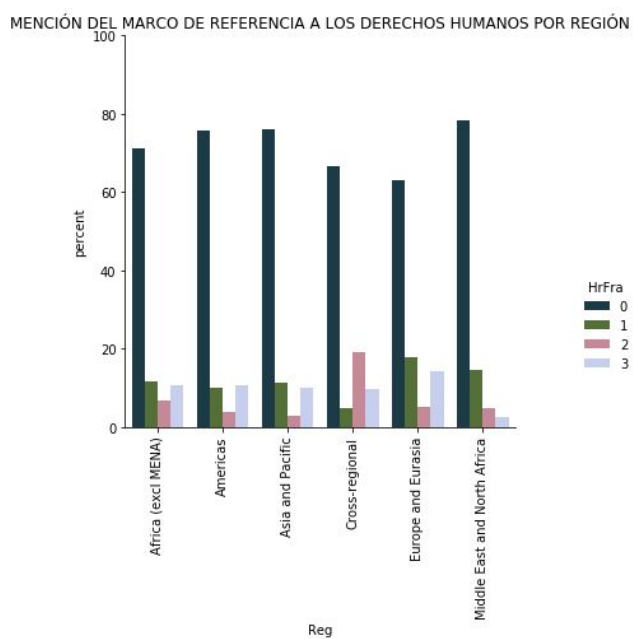


Figura 9

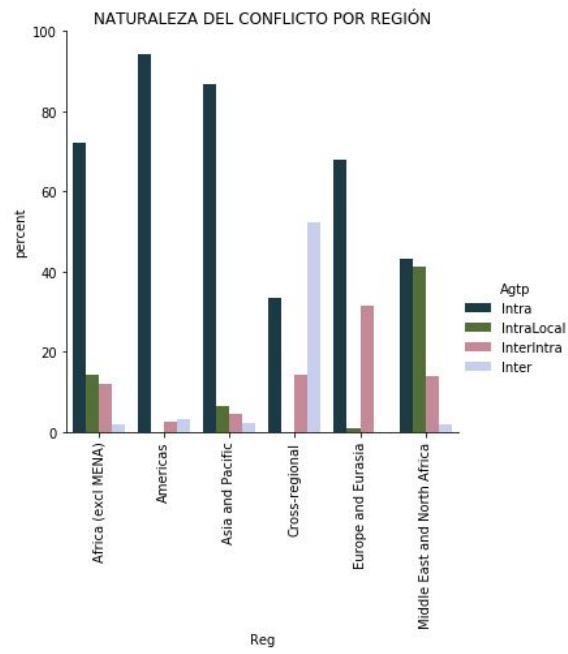


Figura 10

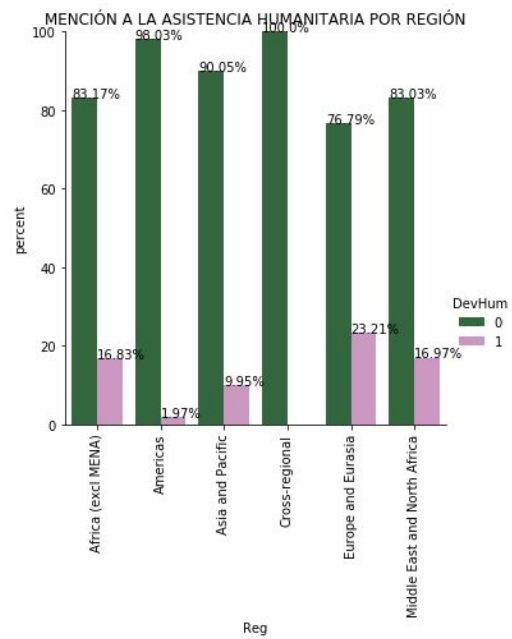


Figura 11

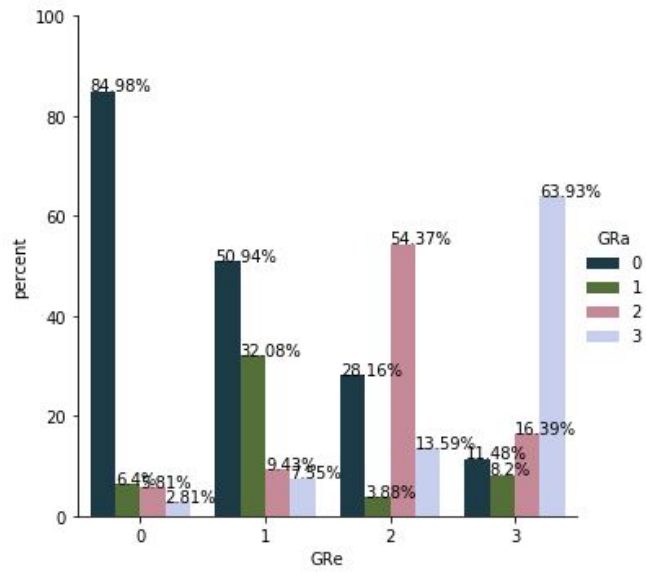


Figura 12

INDICENCIA EN GRUPOS DE NIÑOS EN FUNCIÓN DE LOS GRUPOS RACIALES

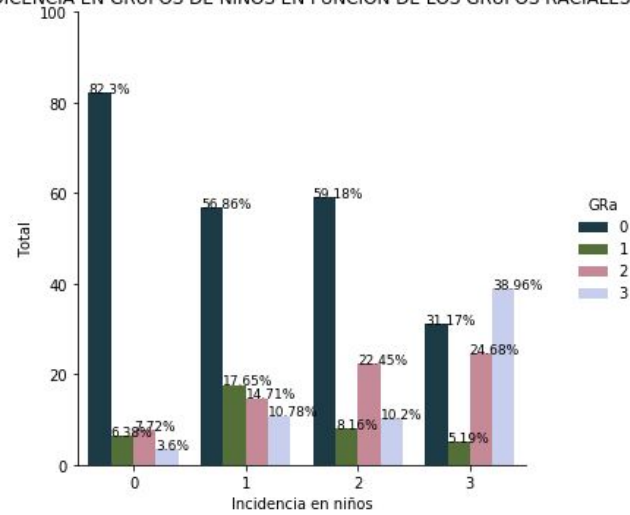


Figura 13

GRADO DE MENCIÓN DE IGUALDAD Y GRADO DE MENCIÓN DE LA DEMOCRACIA

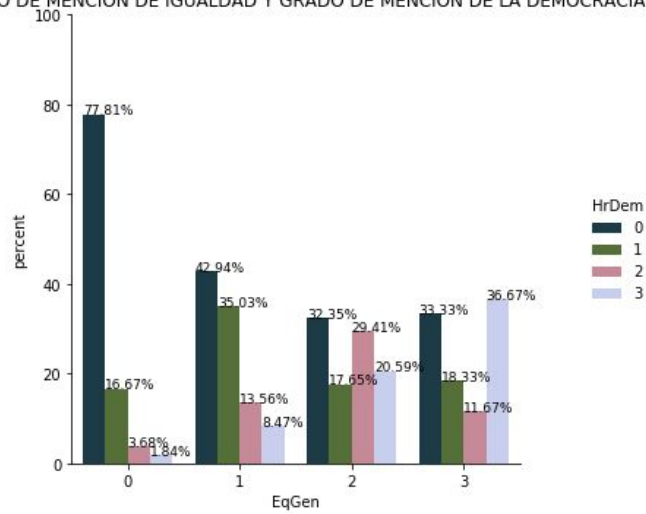


Figura 14

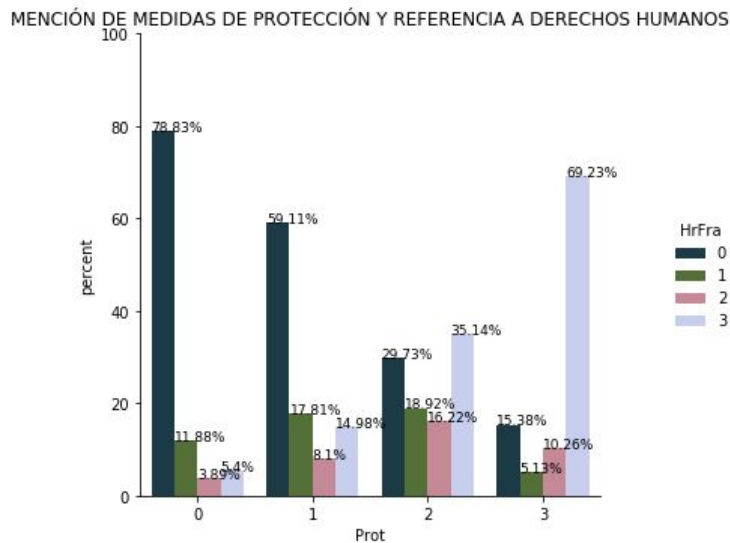


Figura 15

3. CONTENIDO ESPECÍFICO PEC 3

3.1. ABSTRACT/RESUMEN

Este trabajo es la segunda parte de la elaboración de una visualización correspondiente a los acuerdos de paz entre diferentes países del mundo relacionados con derechos humanos e igualdad. La base de datos contiene acuerdos desde el año 1990 hasta la actualidad y nos lo proporciona la Universidad de Edimburgo. La idea principal de la visualización es relacionar los atributos más importantes, como puede ser la asistencia humanitaria o la mención en los acuerdos a la igualdad, con cada uno de los países y regiones. En esta parte del trabajo, mostraremos las fuentes de datos, las principales ideas para los gráficos de la visualización y un diseño final que es parecido a la visualización que expondremos en la siguiente parte del proyecto.

3.2. INTRODUCCIÓN

En este trabajo continuamos con el estudio de la base de datos correspondiente al proyecto de los acuerdos de paz (*Peace Agreements*) al que podemos acceder en la página www.peaceagreements.org. En esta página recordemos que hay una base de datos con más de doscientos atributos, los cuales pueden ser aprovechados para el estudio de diferentes tipos de acuerdos de paz, involucrando así a los países y comparando la calidad de unos acuerdos con la de otros. El objetivo principal de este trabajo es utilizar este conjunto de datos para generar una visualización útil sobre los acuerdos de paz.

El epicentro del proyecto son los acuerdos de paz asociados a tema *derechos humanos e igualdad*. Así, se intentará comparar el alcance de estos acuerdos en cada uno de los territorios y regiones del planeta, viendo cuáles son los temas que más se tienen en cuenta dependiendo de las coordenadas geográficas en las que nos encontremos. El **motivo** de selección de este tema concreto es la voluntad de evidenciar las diferencias entre las diferentes localizaciones del planeta en estos aspectos, algo que se puede intuir pero que no suele estar reflejado en datos.

El paso anterior a este trabajo fue la visualización de las principales tendencias encontradas en la base de datos. Apoyándonos en ellas, intentaremos dar forma en este documento a cómo poder representarlas en base a los datos que tenemos, tras la creación de una hipótesis firme para trabajar, y a la que dar respuesta.

Así pues, tras una breve introducción a los datos y la elaboración de preguntas claves que harán de hipótesis, pasaremos a presentar los datos cuyas tendencias queramos evidenciar, para finalmente realizar el diseño de un boceto para mostrar la forma que queremos que tenga la visualización. En base a las tendencias observadas en la PEC2, las preguntas óptimas que tenemos como **propósito** resolver son:

- ¿En qué años ha habido más acuerdos de paz asociados a los derechos humanos e igualdad?
- ¿En qué fase se encuentran los acuerdos de paz asociados a los derechos humanos e igualdad?
- ¿Cuáles son las regiones y países que más mencionan la igualdad en sus acuerdos de paz?
- ¿Cuáles son las regiones y países que más mencionan los derechos humanos en sus acuerdos de paz?
- ¿Cuáles son las regiones y países que más mencionan la asistencia humanitaria?
- ¿Cómo están relacionadas las menciones a los grupos raciales, grupos religiosos y niños?

De estas preguntas esperamos que las respuestas vayan ligadas al grado de desarrollo de cada país o región, así como al tipo de sistema político y cultural que tienen las diferentes localizaciones geográficas. Además, también se espera que el número de acuerdos de paz ligados a este tema abunde a comienzos del siglo XX, y que la relación entre varios de los atributos, como la mención a los grupos raciales y grupos de niños, sea directa.

La temática inicial del proyecto estaba ligada a la representación de todos los atributos analizados en la PEC2. Así, nos interesaba hacernos en primer lugar una idea de qué variables iban a ser decisivas a la hora de elaborar el trabajo (por ejemplo, queríamos ver si las referencias a medidas de protección en los acuerdos iba a ser clave para diferenciar unas regiones de otras; o si se incluyen menciones de la democracia en los mismos). Tras haber realizado un sesgo inicial para las representaciones de la PEC2, ahora es el momento de reducir todavía más el grupo de atributos, por lo que la temática queda más concretada. Así pues, nos quedaremos en este caso con los atributos que den

respuesta a las preguntas arriba realizadas, los cuales, tras los análisis de la PEC2, son claves para el trabajo. Éstos se expondrán a continuación:

3.3. DATOS

3.3.1. Fuentes de datos

El conjunto inicial de datos, correspondiente a los acuerdos de paz relacionados con derechos humanos e igualdad, cuenta con un total de 1249 documentos, de los que se han extraído 265 atributos. Estos atributos incluyen características generales y categóricas del acuerdo, con el nombre, el país o los países que intervienen en el mismo, o el tipo de acuerdo llevado a cabo (el tipo de esta base de datos es único: “derechos humanos e igualdad”). La mayoría de los atributos son numéricos. Tras el análisis realizado en la PEC2 de los mismos, y viendo los resultados obtenidos, se exponen a continuación los atributos categóricos y numéricos más representativos de la base de datos.

Además, para la correcta construcción de la visualización, utilizaremos la información de la localización de los países que firman el acuerdo, así como atributos directamente relacionados con el tema “derechos humanos e igualdad”:

Atributo	Tipo	Descripción
Con	Texto	Nombre del país o países que intervienen en el acuerdo
Reg	Texto	Región asociada a los países que intervienen (normalmente continentes): <ul style="list-style-type: none">● África (excl MENA)● Américas● Asia y Pacífico● Cross-regional● Europa y Eurasia● África del norte y del este
Loc1ISO	Categórico	Código ISO de uno de los países en el que se ha producido el conflicto (3 letras)
Loc2ISO	Categórico	Código ISO del segundo país, en el caso de que sea un conflicto entre países (3 letras)
Dat	Categórico	Fecha del acuerdo en formato AAAA-MM-DD
Stage	Texto	Fase en la que se encuentra el acuerdo de paz:

		<ul style="list-style-type: none"> • Pre (prenegociación) • SubPar (acuerdo de algunas cuestiones) • SubComp (se acuerdan temas esenciales) • Imp (implementación de un acuerdo anterior) • Ren (renovación de un acuerdo anterior) • Cea (el acuerdo está ligado a un alto al fuego) • Other (otra fase)
EqGen	Categórico	<p>Indica si el acuerdo incluye referencias a la igualdad o no discriminación. Toma valores enteros entre 0 y 3, asociando por medio de estos números el grado de referencias realizadas.</p> <ul style="list-style-type: none"> • 0 (ninguna mención) • 1 (mención breve) • 2 (mención con detalles) • 3 (detalles concretos, como plazos establecidos y planes elaborados)
HrDem	Categórico	Indica si el acuerdo incluye referencias a la democracia, tomando valores enteros entre 0 y 3, dependiendo del grado de mención.
DevHum	Binario	<p>Indica si el acuerdo hace referencia a la asistencia humanitaria:</p> <ul style="list-style-type: none"> • 0 (no se menciona) • 1 (se menciona)
GRa	Categórico	Indica si el acuerdo incluye referencias a los diferentes grupos raciales de la población, tomando valores enteros entre 0 y 3, dependiendo del grado de mención.
GRe	Categórico	Indica si el acuerdo incluye referencias a los diferentes grupos religiosos de la población, tomando valores enteros entre 0 y 3, dependiendo del grado de mención.
GCh	Categórico	Indica si el acuerdo incluye referencias a los niños, tomando valores enteros entre 0 y 3, dependiendo del grado de mención.

3.3.2. Limpieza y transformación de los datos

Tras la presentación de los atributos con los que nos quedamos, tendremos que transformarlos para una correcta representación de los mismos. Veamos uno a uno los atributos y qué transformaciones son convenientes realizar.

Tanto el nombre del país o países que intervienen (*Con*), como las regiones que lo hacen (*Reg*), están expresados de forma correcta; no contienen ningún valor nulo y no será necesario modificarlos. Además, el código ISO de los países que firman el acuerdo se encuentran distribuidos en dos columnas, como hemos descrito anteriormente. Este código facilitará la representación gráfica de los datos y permitirá la consulta y agregación de los valores por país.

El principal dato transformado es la fecha (*Dat*). Como para nuestro estudio nos interesa únicamente el año, y no el día y el mes que se corresponde al acuerdo de paz, crearemos un nuevo atributo llamado *Year* que contenga únicamente el año del acuerdo de paz, lo cual nos permitirá una representación gráfica más adecuada.

El resto de atributos (*HrDem*, *DevHum*, *GRa*, *GRe*, *GCh*) no es necesario modificarlos. Tal y como están son útiles para nuestro propósito.

3.3.3. Información utilizada de la PEC2

Como es de esperar, las conclusiones obtenidas de la PEC2 deben ser útiles para esta parte del trabajo. Así, las relaciones más importantes observadas entre los atributos asociados a nuestra base de datos serán claves para el enfoque final de la visualización. Así, hay características generales que sería interesante reflejar en la visualización final, como la distribución de los acuerdos de paz en materia de derechos humanos e igualdad a lo largo de los años. Además, hay diferencias significativas entre la situación actual de los acuerdos (prenegociación, preacuerdo...), lo cual también es útil llevar a la visualización final.

En lo que respecta a la distribución de las distintas regiones y países, hay tres características que, en base a los gráficos de la anterior PEC, sería útil visualizarlas. La primera de las características es la mención a la igualdad (*EqGen*) en los acuerdos de paz, distinguiendo el grado de mención con números del cero al tres. Tras ver su distribución en el diagrama de barras del documento de la PEC2, sería más útil representarla de una forma más visual (quizá con un gráfico de mapa). La segunda de las características que sería interesante representar es la mención a la democracia (*HrDem*) en los acuerdos, tomando también valores desde el cero hasta el tres. Lo mismo ocurre con la mención a la asistencia humanitaria (*DevHum*), que en este caso es una variable binaria estando disponible solo los valores cero y uno. Podríamos, tanto la variable *HrDem*, como *DevHum*, representarlas más visualmente.

Por último, podríamos representar relaciones importantes entre los datos. En la anterior PEC hemos apreciado que el grado de mención a los grupos raciales (*GRa*), como a los grupos religiosos (*GRe*), como a los grupos de niños (*GCh*), es bastante homogéneo en las diferentes regiones. A pesar de ello, sí nos hemos dado cuenta que existen relaciones directas entre estas variables (vistas en los histogramas cruzados de estos atributos en la PEC2), por lo que sería también interesante su representación.

3.3.4. Tipos de relaciones entre los datos

De esta forma estamos realizando tres tipos de relaciones entre los datos:

- **Distribuir:** esto lo estamos haciendo cuando realizamos la distribución de los acuerdos por año, o cuando queremos representar de la fase en la que se encuentran los acuerdos de paz.
- **Comparar:** esto lo realizamos cuando queremos comparar las diversas regiones y países en base a un atributo concreto. De esta forma, comparamos las distribuciones de los atributos *HrDem*, *DevHum* y *EqGen* por localización geográfica.
- **Componer:** esto se corresponde con la composición de las variables *GCh*, *GRe* y *GRa*. De esta forma, cuando queremos representar de forma cruzada si los acuerdos están asociados con grupos raciales o no, y a la vez si están asociados a grupos religiosos o no, estamos realizando una composición de estos dos atributos.

3.3.5. Gráficos elegidos

Para la visualización utilizaremos principalmente dos tipos de gráficos: gráfico de mapa e histogramas. El primero lo utilizaremos para la representación de las diferentes regiones, tal y como hemos explicado con anterioridad, y el segundo para los análisis cruzados y distribuciones simples.

- **Diagramas de barras:** es el tipo de gráfico adecuado para la representación de distribuciones de una única variable. por tanto, los utilizaremos para representar la cantidad de acuerdos sobre derechos humanos e igualdad por año, y también para graficar las distintas fases en las que se encuentran los distintos acuerdos.
- **Gráfico de mapa:** como es de esperar, utilizaremos este tipo de visualización para representar las principales variables de nuestra representación: los diferentes países en función de las menciones a la igualdad (*EqGen*), a la democracia (*HrDem*) y a la asistencia humanitaria (*DevHum*), las cuales están relacionadas con la temática general de la base de datos: *derechos humanos e igualdad*. Por tanto, lo ideal es mostrar un *mapamundi* donde, por medio de colores, se muestre el grado de mención a estas variables en los acuerdos de paz asociados a cada región o país.
- **Histogramas por grupos:** utilizaremos este tipo de representación gráfica para la composición de las variables *GCh*, *GRe* y *GRa*, ya que queremos ver cómo depende una de otra. Por medio de este tipo de histogramas seremos capaces de visualizar

una variable por rangos, los cuales están asociados a otra de las variables. Por ejemplo, representaremos por colores el grado de mención a los grupos religiosos, y mezclaremos estos colores en el histograma asociado al grado de mención a los grupos raciales.

3.3.6. Gobernanza de datos

Utilizaremos el programa *Tableau* para realizar las visualizaciones. Por tanto, las acciones escogidas para realizar a los datos están relacionadas con este programa. Así, una buena opción para interactuar con los datos es la utilización de filtros. Por medio de ellos, utilizaremos los datos de una vista para filtrar datos en otra para el análisis.

Otra opción parecida pero diferente a la de filtrar es la de resaltar una determinada información en la visualización. En este caso, seleccionaremos un país de nuestra base de datos, o una determinada región, y solo se resaltará en el mapa o en la visualización aquellas regiones o países seleccionados.

Otra opción es la de añadir URL de diferentes sitios webs en nuestra visualización. De esta forma, podríamos ofrecer un enlace de la página de los acuerdos de paz, o algunos enlaces de la wikipedia o con información más precisa de las regiones o países.

Estas serán las formas más directas para interactuar con los datos. En lo que respecta a los diferentes gráficos, podemos distinguir entre primarios y secundarios. De esta forma, consideramos gráficos primarios los diagramas de mapa asociados a las principales variables: *DevHum*, *HrDem* y *EqGen*. En cambio, los diagramas de barras y los histogramas por grupos los podemos considerar gráficos secundarios.

3.4. DISEÑO

3.4.1. Colores

Tableau dispone de múltiples paletas de colores. Desde mi punto de vista, creo que lo mejor es utilizar una paleta de un único color de forma gradual, sobre todo para la visualización que quiero realizar en este trabajo. De esta forma, he escogido una paleta de colores verdes como la siguiente:

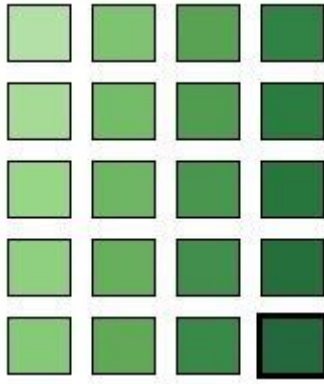


Figura 1

3.4.2. Tipografía

Para estas visualizaciones creo que debemos escoger una tipografía sencilla, como *Arial* o la que nos proporciona *Tableau: Tableau Book*. Además, el tamaño de la misma debe ser adecuada al tipo de representación que queremos realizar. Por ejemplo, para la representación principal se puede utilizar un tamaño de 15, mientras que para los títulos de representaciones secundarias un tamaño de 9.

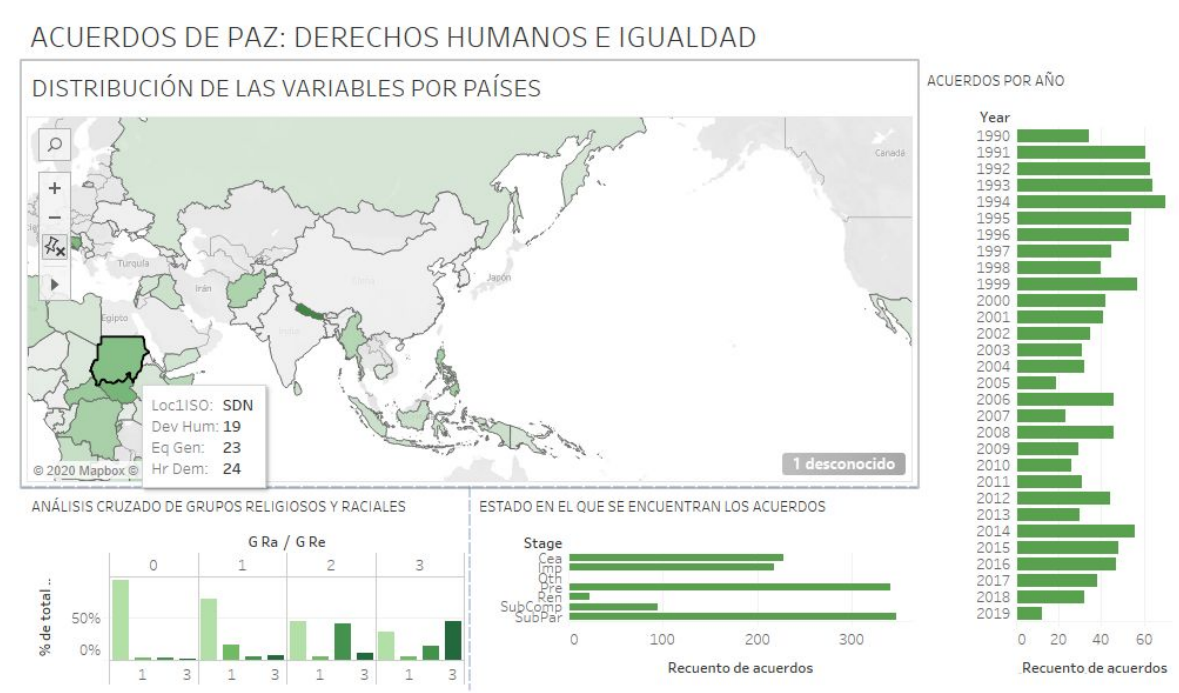
3.4.3. Estructura y bloques

La estructura de la visualización será la siguiente:

- En la parte **superior izquierda** tendremos la visualización principal de nuestro proyecto: un diagrama de mapa con los diferentes países en función de las características básicas de nuestro proyecto: *EqGen*, *HrDem* y *DevHum*. Será un mapa interactivo donde podemos seleccionar un país y que nos muestre características del mismo.
- En la parte **derecha** de la visualización tendremos un diagrama de barras con el número de acuerdos en materia de derechos humanos e igualdad alcanzados por año.
- En la parte **inferior izquierda** tendremos dos gráficos secundarios: por un lado, una comparativa de los atributos *GRe* y *GRa*, o *GCh* entre sí, para visualizar directamente la correlación entre ellos. A su derecha, además, tendremos otro diagrama de barras que muestre el estado en el que se encuentran los acuerdos que estamos estudiando.

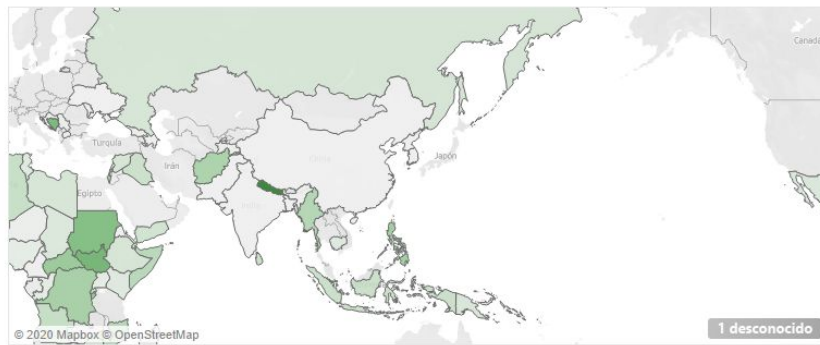
3.4.4. Mockup

En base a la estructura y características anteriores, tendremos un diseño de nuestra visualización parecida a lo siguiente:

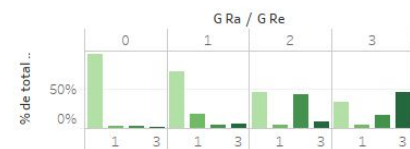


Al seleccionar un país tenemos la suma de los grados de mención de cada una de las características estudiadas (*DevHum*, *HrDem* y *EqGen*), estando el color del país de acorde a estas características. Del mismo modo, podemos seleccionar un año y obtener información del mismo:

DISTRIBUCIÓN DE LAS VARIABLES POR PAÍSES



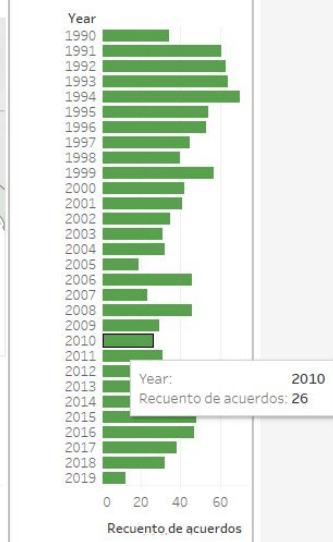
ANÁLISIS CRUZADO DE GRUPOS RELIGIOSOS Y RACIALES



ESTADO EN EL QUE SE ENCUENTRAN LOS ACUERDOS



ACUERDOS POR AÑO



4. PROYECTO FINAL

Este proyecto final contiene la visualización correspondiente a los acuerdos de paz entre diferentes países del mundo relacionados con derechos humanos e igualdad, con diversos aspectos como la localización geográfica, el tipo de conflicto o si el acuerdo tiene en cuenta determinadas características, como la mención en los mismos de la igualdad de género o grupos raciales y religiosos.

La visualización se encuentra en el enlace: [acuerdos de paz](#)

4.1. EXPLICACIÓN DEL PROYECTO FINAL

Este proyecto se ha centrado en ver la distribución de los distintos acuerdos de paz relacionados con los derechos humanos e igualdad en todos los países y regiones del planeta. En la visualización tenemos varias partes. La parte principal es un mapa con los diferentes países en el que aparece la cantidad de acuerdos alcanzados por los mismos. Además, podemos acceder a la wikipedia de cada país poniendo el cursor sobre ellos. Por otro lado, tenemos otras representaciones, como una distribución de la cantidad de acuerdos por año, el estado en el que se encuentran dichos acuerdos, o análisis cruzados de varias variables, como grupos raciales y religiosos.

Es interesante describir qué cumple el proyecto final de las hipótesis iniciales planteadas, y qué cambios hemos aplicado a la idea inicial. Algunas ideas han ido modificándose a lo largo del proyecto. En primer lugar, en la PEC2 solo realizamos diagramas de barras y análisis cruzadas de determinados atributos; mientras que al proyecto final le hemos añadido gráficos de mapa. Además, la temática queda más concretada, quedándonos para la visualización final con los atributos que den respuesta a las preguntas realizadas en la PEC3, los cuales, tras los análisis de la PEC2, son claves para el trabajo.

Ha habido diversos **retos** a los que me he enfrentado a lo largo del proyecto. En primer lugar, una dificultad ha sido la extensión de la base de datos iniciales. Más allá del número de filas, el problema principal han sido los 265 atributos presentes en la misma, los cuales hemos tenido, prácticamente, que analizar uno a uno para ver cuáles pueden resultar útiles para nuestra representación. Ha sido de gran ayuda la descripción de los atributos presentes en la [página](#) de la que hemos sacado los datos, pues nos ha ayudado a saber qué variables son adecuadas para nuestra representación.

Otro de los retos ha sido el proceso de limpieza de datos. Con la ayuda de *Python*, hemos sido capaces de quedarnos con una nueva base de datos para la representación de la misma en el programa *Tableau*. Si introducimos la base de datos sin modificar en

Tableau para la representación del mapa, teníamos un problema: los acuerdos que se producían entre varios países no se tenían en cuenta para la representación, pues el atributo “country” aparecía con los nombres de los diversos países que intervenían en el acuerdo. Para solucionar este problema, con ayuda de *Python*, hemos tenido que duplicar o triplicar estas filas, dependiendo de la cantidad de países del acuerdo, creando así una nueva variable que haga referencia a cada uno de estos países en las filas creadas.

De esta forma, tenemos dos entradas de datos en *Tableau*: por un lado, la base de datos inicial descargada de la página de los acuerdos de paz, donde el atributo correspondiente a los países contiene la información de cada uno de los países del acuerdo. Este fichero nos sirve para representar la distribución de los acuerdos por año, y también el estado en el que se encuentran los mismos, así como los análisis cruzados de variables, como el grado de mención a grupos raciales y religiosos. En cambio, el nuevo fichero tras realizar las modificaciones de *Python* nos resulta bastante útil para la representación del mapa, pues ya tiene en cuenta, con el nuevo atributo correspondiente a los países, los acuerdos realizados entre países.

Una dificultad añadida es el uso de *Tableau* por primera vez, aunque gracias a toda la información presente en su página <https://www.tableau.com/es-es> he sido capaz de crear todas las visualizaciones de una forma satisfactoria. El programa es bastante descriptivo y útil para la visualización de datos, resultando sencillo el seguimiento de los pasos para la realización de cada representación.

Estoy satisfecho con el resultado que he obtenido en la visualización, habiendo solventado todas las dificultades y retos iniciales adecuadamente, y respondido a las preguntas planteadas al inicio con este tipo de representaciones, que son bastante descriptivos para nuestro propósito.

4.2. DESCRIPCIÓN TÉCNICA DEL PROYECTO

En el conjunto de datos escogido se tienen en cuenta un total de 1249 documentos, de los que se han extraído 265 atributos. Por tanto, se cuenta de 1249 filas y 265 columnas. Estos atributos incluyen características generales y categóricas del acuerdo, como el nombre, el país o los países que intervienen en el mismo o el tipo de acuerdo (en este caso el tipo es exclusivamente “derechos humanos e igualdad”). Para la elaboración de la visualización he seleccionado el nombre y el código identificador de los países y regiones (*Con*, *Reg*, *Loc1ISO* y *Loc2ISO*). Además, he creado el atributo *Year* cogiendo las últimas cifras del atributo *Dat* de la base de datos. Para nuestro propósito, además, hemos seleccionado los metadatos *Stage*, *EqGen*, *HrDem*, *DevHum*, *GRa*, *GRe* y *GCh*.

El proceso de limpieza y transformación de datos se encuentra en el archivo adjunto *transform.py*. En él, se ha transformado el archivo inicial descargado de la página de acuerdos de paz: *pax_data_1249_agreements_18-04-20.csv*. En primer lugar, hemos

guardado las variables categóricas y numéricas recurriendo a la normalización de las variables numéricas. Esto es necesario realizarlo para que todas las variables tengan el mismo peso: hay algunas variables numéricas que toman valores enteros de 0 a 3, y otras que solo pueden tomar 0 y 1.

En segundo lugar, dentro del proceso de transformación, hemos sustituido espacios en blanco y comas presentes en el atributo de los países, los cuales dificultaban la representación. La parte principal de la transformación la forma la duplicación de algunas filas. Esto es necesario realizarlo para guardar la consistencia en el gráfico de mapa, pues había determinados acuerdos que no se contabilizaban para los países debido a que en los mismos intervenían más de un país: estos datos no se utilizaban porque *Tableau* no asociaba este atributo al país en cuestión. Para solucionarlo, hemos cargado un fichero adicional (*Country.xlsx*) en el que se asocia cada país con su código ISO. Con el código presente en *Transform.py* hemos sido capaces de duplicar estas filas con la ayuda del fichero de los códigos, y de crear un nuevo atributo “*Con*” que tenga en cuenta a cada uno de los países de “*country*”. Tras eliminar algunas filas que no eran de interés, hemos exportado el fichero como “*peace.xlsx*” y lo hemos cargado en *Tableau* para la representación del mapa.

4.3. DESCRIPCIÓN DE LOS GRÁFICOS

En la parte **superior izquierda** de la visualización tenemos la parte principal del proyecto: un diagrama de mapa con los diferentes países en función de la cantidad de acuerdos llevados a cabo por los mismos. Es un mapa interactivo, en el que, seleccionando un país, podemos ir a la wikipedia el mismo y observar datos adicionales.

En la parte **derecha** de la visualización tendremos un diagrama de barras con el número de acuerdos en materia de derechos humanos e igualdad alcanzados por año. Seleccionando las barras, obtenemos información del número de acuerdos en el año en cuestión.

En la parte **inferior izquierda** tendremos dos gráficos secundarios: por un lado, una comparativa de los atributos *GRe* y *GRa*, o *GCh* entre sí, para visualizar directamente la correlación entre ellos, pues en la parte de exploración de datos nos hemos dado cuenta de que existe correlación entre estas variables. A su derecha, además, tendremos otro diagrama de barras que muestre el estado en el que se encuentran los acuerdos que estamos estudiando, encontrándose la mayoría en estado *SubPar* (acuerdo en algunas cuestiones) o en *Pre* (en la prenegociación).

5. BIBLIOGRAFÍA

- SKAU, D (2012). Visually: *How to use Maps in Data Visualization*.
<https://visual.ly/blog/you-are-here-using-maps-in-data-visualization/>
- <https://help.tableau.com/current/pro/desktop/es-es/actions.htm>, consultado el 05/05/2020.