
LOS REFUGIADOS EN LOS ACUERDOS DE PAZ

INFORME FINAL DE VISUALIZACIÓN DE DATOS



AUTOR: MARIO MARTÍNEZ GARCÍA

10 DE JUNIO DE 2020

ÍNDICE

1.	INTRODUCCIÓN.....	4
1.1.	TÍTULO DEL PROYECTO.....	4
1.2.	URL DE LA VISUALIZACIÓN	4
1.3.	DESCRIPCIÓN DEL DOCUMENTO	4
2.	PEC2: HIPÓTESIS Y EXPLORACIÓN INICIAL DE LOS DATOS.	5
2.1.	DESCRIPCIÓN.....	5
2.2.	LOS DATOS: PRESENTACIÓN.....	5
2.3.	LOS DATOS: EXPLORACIÓN.....	9
2.4.	LOS DATOS: PROCEDIMIENTO Y HERRAMIENTAS.....	13
3.	PEC3: INFORME DEL PROYECTO.....	15
3.1.	INTRODUCCIÓN. HIPÓTESIS DEL PROYECTO.	15
3.2.	BLOQUE DE DATOS.....	16
3.2.1.	DESCRIPCIÓN DEL DATASET A UTILIZAR EN LA VISUALIZACIÓN.....	16
3.2.2.	TRANSFORMACIONES SOBRE EL DATASET	17
3.2.3.	EVIDENCIAS OBTENIDAS SOBRE LA PEC2.	18
3.2.4.	¿CÓMO VAMOS A ENFOCAR LA VISUALIZACIÓN?.....	19
3.2.5.	GOBERNANZA DE DATOS.....	21
3.3.	DISEÑO MOCKUP.....	21
3.3.1.	COLORES.....	21
3.3.2.	TIPOGRAFÍA.....	22
3.3.3.	ESTRUCTURA Y FORMATO.....	22
3.3.4.	PROPUESTA DE DISEÑO.....	23
4.	PEC4: CREACIÓN DE LA VISUALIZACIÓN.....	25
4.1.	DESARROLLO DEL PROYECTO	25
4.2.	EXPLICACIÓN DEL PROYECTO. ¿QUÉ CUMPLE Y QUÉ NO CUMPLE DEL PROYECTO INICIAL?.	
	DESCRIPCIÓN DE LOS GRÁFICOS EMPLEADOS.....	26
4.2.1.	MAPA GEOGRÁFICO.....	26
4.2.2.	MAPA DE CALOR.....	28
4.2.3.	ACUERDOS POR AÑO.....	29
4.2.4.	NIVEL DE IMPLICACIÓN POR REGIÓN	30
4.3.	DESCRIPCIÓN TÉCNICA DEL PROYECTO	30
4.3.1.	PYTHON.....	31

4.3.2. TABLEAU.....	32
4.3. ESFUERZO, RETOS Y RESULTADOS OBTENIDOS.....	33
5. BIBLIOGRAFÍA.....	34
A. APÉNDICE VISUALIZACIONES.....	35

1. INTRODUCCIÓN

A lo largo de este informe se recopilarán 3 de las 4 pruebas de evaluación continua propuestas en el máster de ciencia de datos de la UOC, más concretamente sobre la asignatura de “Visualización de datos”.

La PEC2 en la que se establece una primera familiarización de los datos será descrita en profundidad sobre el apartado dos y el apéndice A.

La PEC3 constituye el grueso del informe del proyecto llegando a describir una propuesta de diseño. Esta PEC la encontraremos sobre el apartado tres.

La PEC4 abarca la presentación de la visualización final y la encontraremos sobre el apartado cuatro.

1.1. TÍTULO DEL PROYECTO

El título establecido para la visualización intenta mostrar las dos piezas claves que encontramos en ella: los refugiados y los acuerdos de paz. De esta forma recibirá el título de: **“LOS REFUGIADOS EN LOS ACUERDOS DE PAZ”**.

1.2. URL DE LA VISUALIZACIÓN

La visualización puede encontrarse sobre la plataforma “Tableau Public”, un software gratuito de visualización de datos que a partir de simples ficheros te permite generar visualizaciones de alto impacto gráfico e interactivo. El enlace a la visualización será el siguiente:

Visualización final: <https://public.tableau.com/profile/mario.mart.nez.garc.a#!/vizhome/Losrefugiadosenlosacuerdosdepaz/Dashboard1?publish=yes>

Por otra parte, con el fin de tener todos los documentos en un mismo espacio, se creará un repositorio en GitHub donde podemos encontrar los ficheros de datos utilizados, el código y el informe final.

GitHub: <https://github.com/mariomartgarcia/RefugeesPeaceAgreements>

1.3. DESCRIPCIÓN DEL DOCUMENTO

A lo largo de este informe se recopilan las directrices seguidas para la creación de una visualización sobre los acuerdos de paz. Los datos serán extraídos de <https://www.peaceagreements.org/search> donde seleccionaremos como grupo de interés a los refugiados y tendremos como objetivo fundamental identificar la identidad de los acuerdos firmados como el país, el número de acuerdos o sus características básicas con el fin de modelar una visualización final sobre la plataforma Tableau. Este informe describe la presentación de los datos donde se narra el proceso de elección del fichero de datos, el análisis exploratorio, el enfoque sobre la visualización tras el análisis, la primera propuesta de diseño y la explicación del proyecto final. La visualización obtenida será capaz de mostrar el lugar, la cantidad y las características básicas de los acuerdos de paz con referencias a refugiados firmados entre 1990 y 2019.

2. PEC2: HIPÓTESIS Y EXPLORACIÓN INICIAL DE LOS DATOS.

2.1. DESCRIPCIÓN

A la hora de buscar la base de datos adecuada que sustenta esta práctica se ha optado por elegir una de las múltiples bases de datos sobre los documentos de acuerdos de paz en todo el mundo desde 1990 hasta hoy. Recurriremos a la página de [“Peace Agreements”](#) [1] donde podremos elegir una base de datos en función de diferentes características como la región, el tipo de acuerdo, la categoría o el tema entre otros. Estos datos han sido recopilados y publicados por la Universidad de Edimburgo, donde cada uno de los acuerdos proviene de repositorios existentes, páginas específicas de países y documentación de organizaciones oficiales entre otras. Cabe destacar que la página trata cerca de 2000 documentos legales y cada documento tiene más de 260 metadatos.

En nuestro caso, hemos optado por centrarnos sobre los acuerdos de paz de un tema concreto, los refugiados. Con esta base de datos se plantearía la hipótesis inicial: **obtener conclusiones sobre la identidad de los acuerdos de paz firmados.**

- Lo que involucraban.
- ¿A quién involucraban?
- ¿En qué regiones?

Las personas refugiadas son aquellas que se han visto obligadas a huir de su país por sufrir persecuciones por motivos de raza, religión, nacionalidad y opiniones políticas. En la actualidad, Siria es el primer país de origen de las personas refugiadas del mundo debido al conflicto que dura más de 8 años. Además, Afganistán, Irak, Somalia, R.D. Congo, Sudán, Sudán del Sur, Myanmar o Palestina con conflictos interminables están entre los principales países de origen de las personas refugiadas. Por otro lado, los principales 5 países donde estos refugiados son acogidos son: Turquía, Pakistán, Uganda, Sudán y Alemania. [2]

La **principal motivación** que nos ha llevado a elegir este fichero es sin duda la gran repercusión política que actualmente provocan los refugiados (principalmente de la guerra de Siria) sobre los países europeos más desarrollados.

Mediante esta primera toma de contacto con la base de datos se pretende entender hacia dónde van enfocadas las medidas establecidas en los acuerdos de paz que referencian a los refugiados. Este segundo apartado del documento corresponde a la PEC2 de la asignatura y quedará dividido en: presentación de los datos, exploración de los datos y, procedimientos y herramientas. La visualización de los datos para este apartado aparecerá recopilada sobre el apéndice A. Cabe destacar que el código utilizado en esta parte de la práctica aparece en el documento adjunto “PEC2_CODE_MMG.ipynb”.

2.2. LOS DATOS: PRESENTACIÓN

Sobre la base de datos tomada nos encontraremos con un total de 265 metadatos, categóricos y numéricos, rellenos por 476 acuerdos de paz diferentes. En este apartado trataremos de dar una visión general sobre los 265 metadatos y exponer los que serán considerados como relevantes para su posterior estudio. La explicación completa de cada uno de los atributos presentes en la base de datos podemos encontrarla en el [codebook](#), en este apartado realizaremos una breve descripción sobre los conjuntos de atributos:

Elementos sobre el contexto (Información básica)

Encontramos elementos principalmente categóricos que serán usados en un posterior estudio como el país, tipo de conflicto, fecha, estado del acuerdo, región, proceso de paz... Cómo vemos esta información nos sirve para situar el acuerdo dentro de un determinado contexto y será vital para nuestro estudio.

Elementos sobre las características del acuerdo

Estos atributos nos darán información sobre las medidas que se toman en el acuerdo de paz. Encontraremos atributos que recibirán valores del 0 al 3 donde se tomará el 0 si no tiene ninguna referencia en el acuerdo y el 3 si se referencia completamente en el acuerdo. Además, tendremos otro tipo de atributos que toman el valor 0 si no tiene ninguna referencia en el acuerdo y el 1 si se referencia sobre el acuerdo. Veremos cada uno de estos conjuntos de atributos:

- Grupos.

Encontramos atributos en función de la relación con un determinado grupo: niños, personas discapacitadas, personas con avanzada edad, trabajadores inmigrantes, razas/etnias/grupos nacionales, personas religiosas, personas indígenas y otros grupos.

Como hemos dicho antes, cada uno de estos atributos tendrá una serie de sub-atributos para especificar con más claridad sobre qué trata el acuerdo.

- Género.

Aparecen atributos que se clasifican en función del sexo (Hombre o mujer), Lgtbi y Familia. Además poseerán una especie de sub-atributos.

- Definición de estado.

Diferentes atributos que nos muestran información si el acuerdo hace referencia a aspectos del estado como la determinación, símbolos, fronteras...

- Gobierno.

Nos muestra la relación del acuerdo de paz con procesos gubernamentales como elecciones, reformas constitucionales, instituciones políticas...

- Reparto de poder.

Aparecen atributos como el reparto de poder político, reparto de poder territorial, reparto de poder económico o reparto de poder militar.

- Derechos humanos e igualdad.

Estos atributos nos aportarán una información básica en los acuerdos de paz sobre los derechos humanos, la igualdad, la democracia, medidas de protección, derechos políticos y civiles y otra serie de derechos que involucren a la sociedad afectada por el acuerdo.

- Sector de reforma jurídica.

Nos mostrará el nivel de las reformas jurídicas alcanzadas en el acuerdo con diversos atributos como la justicia criminal, provisiones del estado de emergencia, prisiones y detenciones...

- Reconstrucción socio-económica.

Encontraremos atributos que tratan de mostrar el desarrollo de la reconstrucción, la “vuelta a la normalidad”, la referencias a negocios, bancos, tasaciones, planes económicos...

- Tierras, propiedad y medio ambiente.

Aparecerán atributos como reformas agrarias, derechos nómadas, patrimonio cultural, medio ambiente y acceso a agua.

- Sector de seguridad.

Destinado a las citaciones sobre los diferentes cuerpos de seguridad que pueden intervenir en el acuerdo, terrorismo, drogas, crimen, corrupción, desmilitarización o reintegración entre otros.

- Transición jurídica.

Nos mostrará atributos como la amnistía, las víctimas, la reconciliación... Con esto se pretende conocer el nivel transicional del acuerdo.

- Ejecución.

Diferentes atributos nos muestran la forma en la que se ha ejecutado el acuerdo.

Como vemos tenemos una inmensa cantidad de datos, en este apartado de la práctica realizaremos un primer filtro considerando aquellos atributos que han sido considerados más interesantes. Disponiendo finalmente de un total de 17 atributos. Cabe destacar que sobre la base de datos se ha realizado una modificación, añadiendo una nueva columna que nos proporciona únicamente el año del acuerdo.

ATRIBUTOS CONTEXTUALES

ATRIBUTO	CARACTERÍSTICAS
Contp	Tipo de conflicto: <ul style="list-style-type: none"> • Government • Government/Territory • Territory • Inter-group
Reg	Región donde el conflicto tiene lugar: <ul style="list-style-type: none"> • Europe and Eurasia • Africa (excl. MENA) • Asia and Pacific • Americas • Middle East and North Africa

	<ul style="list-style-type: none"> • Cross-regional
Date	Fecha en el que el acuerdo fue firmado en formato YYYY-MM-DD
Year	Nuevo atributo obtenido a partir de "Date" que nos muestra exclusivamente el año en el que fue firmado el acuerdo.
Status	Tipo del acuerdo firmado: <ul style="list-style-type: none"> • Multiparty signed/agreed: Firmado por varios estados. • Unilateral agreement: Acuerdo unilateral. • Status nuclear: No está clara su composición. • Agreement with subsequent status: Acuerdo no formal.
Agtp	Tipo de conflicto entre estados, regiones o países: <ul style="list-style-type: none"> • Inter: Entre estados • InterIntra: Entre estados pero evaluado sobre asuntos de un único estado • Intra: Dentro del mismo estado • IntraLocal: Dentro del mismo estado a nivel local.
Stage	Estado del acuerdo: <ul style="list-style-type: none"> • Pre: Acuerdo de prenegociación. • SubPar: Se acuerdan algunas cuestiones para resolver el conflicto. • SubComp: Se acuerdan cuestiones substanciales. • Imp: Ejecución de un acuerdo anterior. • Ren: Renovación de compromisos anteriores. • Cea: Acuerdos de alto al fuego. • Other: Acuerdos que no se pueden clasificar.
StageSub	Sub estado del acuerdo, diferentes estados que se cubren en función del estado general del acuerdo.

ATRIBUTOS SOBRE LAS CARACTERÍSTICAS DEL ACUERDO

- Grupos.

ATRIBUTO	CARACTERÍSTICAS
GCh	Referencia a niños. Valores: 0, 1, 2, 3.
GRa	Referencia a razas/etnias/grupos nacionales. Valores: 0, 1, 2, 3.
GRe	Referencia a grupos religiosos. Valores: 0, 1, 2, 3.
GRef	Referencia a grupos de refugiados. Valores: 0, 1, 2, 3.

- Derechos humanos e igualdad.

ATRIBUTO	CARACTERÍSTICAS
HrGen	Derechos humanos. Valores: 0, 1.
EqGen	Igualdad. Valores: 0, 1, 2, 3.
Med	Menciones a los medios de comunicación. Valores: 0, 1, 2, 3.

- Reconstrucción socio-económica.

ATRIBUTO	CARACTERÍSTICAS
Dev	Desarrollo de la reconstrucción socio económica. Valores: 0, 1, 2, 3.

- Sector de seguridad.

ATRIBUTO	CARACTERÍSTICAS
SsrDdr	Mención sobre la “vuelta a normalidad”, es decir, se tienen en cuenta procesos de desmilitarización y reintegración. Valores: 0, 1, 2, 3.

Estos serán los atributos seleccionados en el primer estudio de los datos. Para su elección se ha tenido en cuenta 2 parámetros:

1. Que la información aportada tuviese o pudiera tener relación con el grupo de refugiados.
2. De ser un atributo numérico, la suma sobre sus filas fuese elevada. Esto indicaría una gran presencia de dicho atributo sobre los acuerdos de paz.

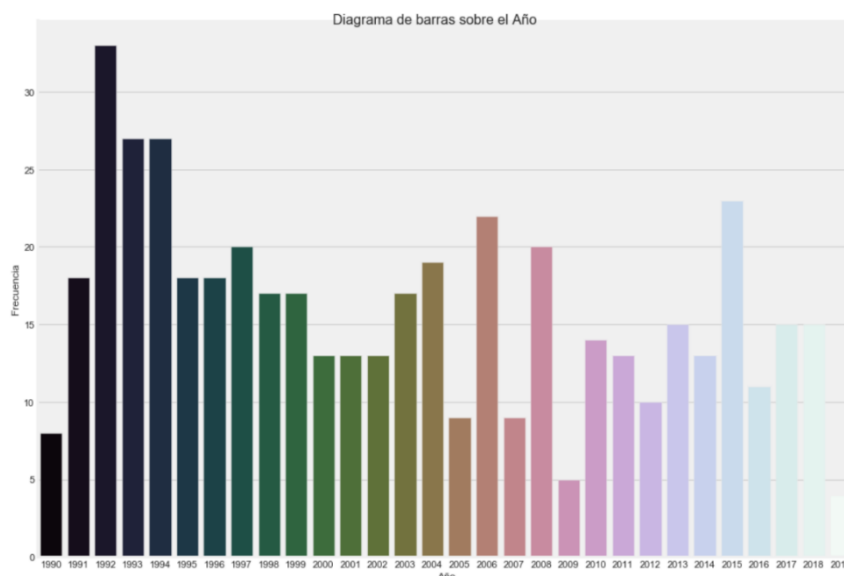
Una vez seleccionados los datos que intervendrán en el estudio, podemos centrar nuestra atención sobre ellos y ver qué características pueden aportarnos. Cabe destacar que este primer subconjunto de datos (no definitivo) nos ayudará a establecer tendencias y relaciones entre los datos, de esta forma si durante este apartado se descubre que es necesario añadir o eliminar algún atributo, lo describiremos sobre el tercer apartado donde estableceremos el conjunto final de datos usado sobre la visualización. Pasaremos a la exploración de los datos.

2.3. LOS DATOS: EXPLORACIÓN

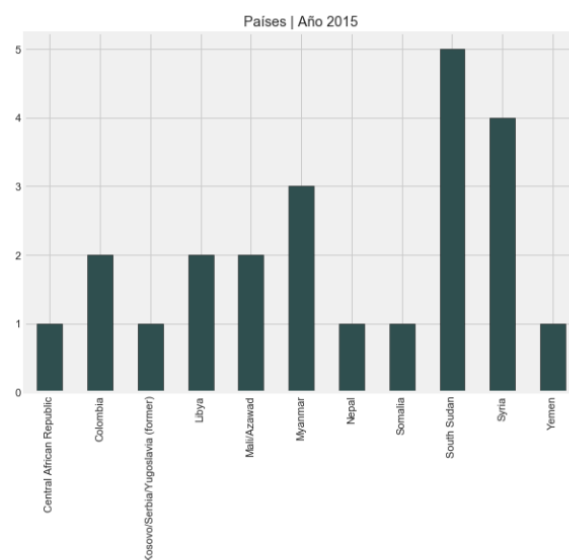
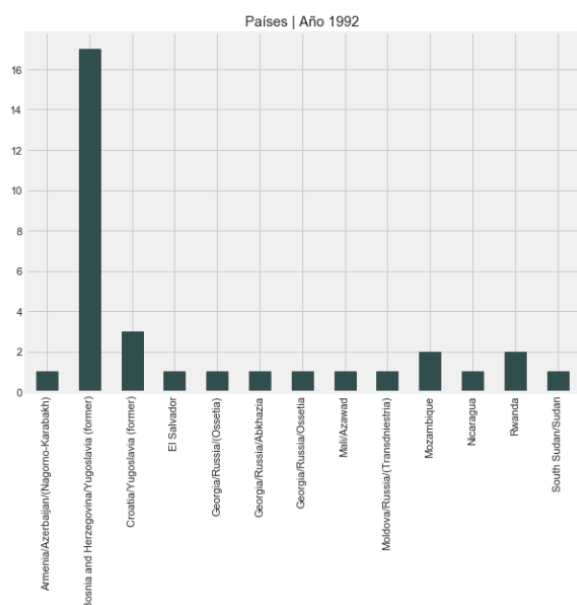
En este apartado trataremos de explicar los hechos más evidentes e interesantes encontrados sobre la base de datos. Aunque en el último apartado aparecen todas las visualizaciones, iremos mostrando los gráficos que se han considerado más importantes para que la explicación sea directa y coherente.

Evolución de acuerdos sobre refugiados desde 1990 hasta 2019

La gráfica principal que nos muestra esta evolución es la siguiente (Figura 3, ult. apartado):



Tras ver esta gráfica, lo que más llama la atención sin lugar a dudas son sus picos, alcanzados en 1992, 2006 y 2015. Es decir, en esos años el número de acuerdos de paz donde se referenciaban a los refugiados fue excesivamente alto. Estudiando que países realizan más acuerdos de paz pueden sacarse conclusiones interesantes sobre el contexto histórico en el que nos movemos. Poniendo el ejemplo con las gráficas de 1992 y 2015 sobre los acuerdos alcanzados con los países tendremos lo siguiente.



Como vemos en imagen a partir de estos gráficos podemos sacar conclusiones directas y evidencias de los datos en función de los picos.

- Pico alcanzado en 1992.

Marcando el final de la guerra Fría tenemos la disolución de la Unión Soviética a finales de 1991 que dejaría múltiples acuerdos sobre refugiados en los que aparecerían Bosnia Herzegovina y Yugoslavia. Durante este periodo más de un millón de personas abandonaron su lugar de residencia huyendo de los conflictos interétnicos que se desarrollaron en Yugoslavia desde el verano de 1991. [3] (Figura 16, ult. apartado)

- Pico alcanzado en 2006.

Aunque no se muestra la gráfica en esta página, esta puede visualizarse en la figura 17 del último apartado. El país con mayores acuerdos sobre los que aparecían referencias a los refugiados es Nepal. El 2006 marca el final de la guerra civil nepalesa y es por esto por lo que el aumento de acuerdos de paz aumenta en este país. [4]

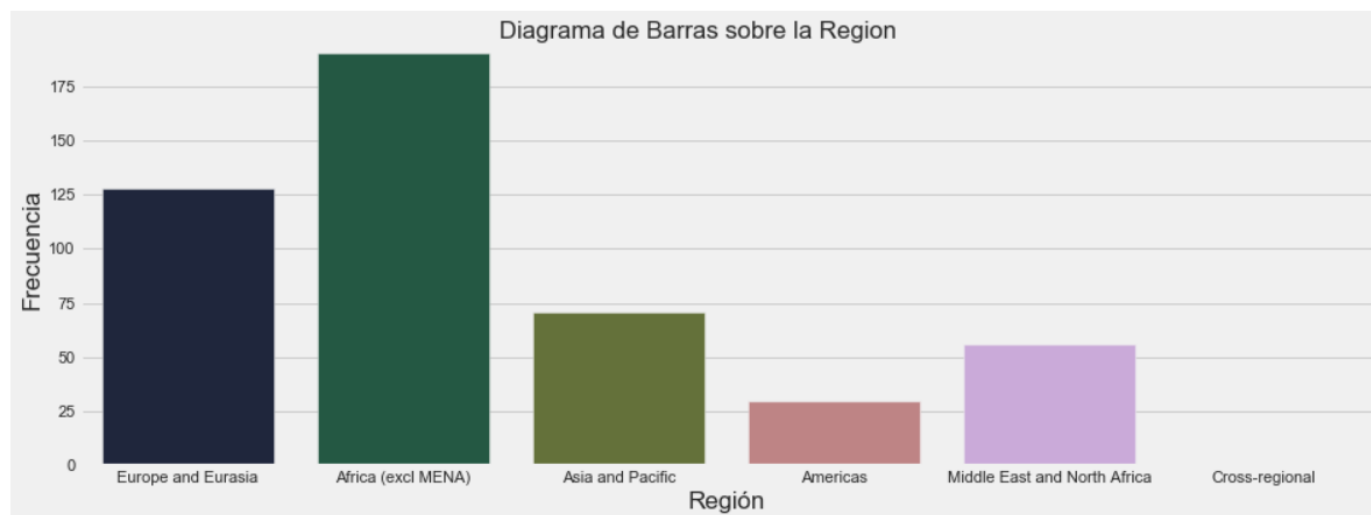
- Pico alcanzado en 2015.

En este punto vemos que los países donde se han alcanzado un mayor número de acuerdos de paz son Siria y Sudán del Sur. La tristemente famosa guerra de Siria se desarrolla durante este periodo y también lo hace la guerra civil de Sudán del Sur. (Figura 18, ult. apartado)

Tal y como hemos apreciado en cada uno de los picos, los acuerdos de paz relacionados con refugiados vienen de la mano de una guerra, ya sea civil o entre diferentes estados. Estos valores pueden considerarse outliers al ser superiores al resto, además, nos aportan tres evidencias especialmente interesantes sobre el contexto de los datos.

Acuerdos de paz en referencia a refugiados en función de las regiones

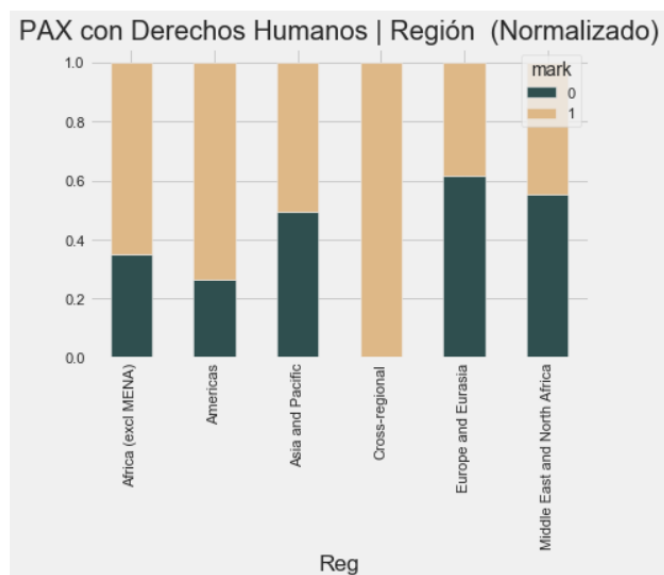
En esta gráfica podremos apreciar cómo se distribuyen (Figura 2, ult. Apartado):



Como vemos desde 1990 la región con mayor número de refugiados es África(excl. Mena) debido principalmente a guerras civiles como la de Sudán del Sur o Somalia y conflictos como el de Chad-Sudán o el de Nigeria. África es seguida por Europa y Eurasia debido al final de la Guerra Fría y la Guerra de Siria entre otros.

Mención sobre los derechos humanos en los acuerdos

Podremos ver la mención hacia los derechos humanos en función de las regiones.



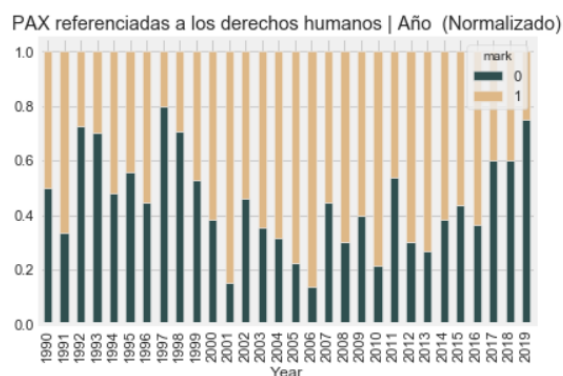
En primer lugar, la región “Cross-Regional” no la tendremos en cuenta al disponer de pocos casos de estudio. En proporción vemos que la región donde más se mencionan los derechos humanos serán las américas, seguida por África y Asia. En último lugar encontraremos sorprendentemente a Europa y Eurasia.

(Figura 9. Ult. apartado)

Por otra parte, si nos fijamos en la figura de la derecha (Figura 14, ult apartado) podemos ver como en 2006 los acuerdos sobre refugiados en los que se incluían referencias sobre los derechos humanos ascienden hasta más del 80%, en pleno final de la guerra civil nepalesa se tendrán mucho en cuenta los derechos humanos.

Además, para los otros dos picos encontrados podemos observar ciertas diferencias. El pico de 1992, marcado por el final de la Guerra Fría, nos muestra que los acuerdos de paz firmados en los que se veían involucrados refugiados tan solo referencian los derechos humanos alrededor del 30% de ellos. Sin embargo, sobre el pico de 2015 los acuerdos con referencias a derechos humanos alcanzan casi un 60%.

Como vemos no se puede sacar una tendencia clara sobre la presencia de los derechos humanos en estos acuerdos, dependerá del tipo de situación, guerra y región que estén involucradas.



Repaso sobre las gráficas del último apartado

En la figura 1 observamos como principalmente el tipo de conflicto predominante sobre este tipo de acuerdos de paz son los referidos al gobierno o gobierno/territorio. Teniendo en mente los datos vistos en páginas anteriores vemos que esto tiene sentido al tratarse en la mayoría de los casos de guerras civiles, conflictos, disputas de territorios... La figura 2 y la figura 3 ya han sido explicadas en las páginas anteriores.

El tipo de acuerdo alcanzado sobre cada uno de estos documentos puede verse reflejado sobre la figura 4, vemos como sobre todo estos acuerdos son firmados por varios estados. Además, la figura 5 nos muestra que principalmente la naturaleza del conflicto es en un solo estado como puede ser una guerra civil.

Sobre la figura 6 encontraremos el estado en el que se encuentra acuerdo, donde la mayoría de ellos se encuentran en estado “SubPar”, es decir, se acuerdan algunas cuestiones para resolver el conflicto.

A partir de esta figura comenzaremos a ver parejas de gráficas donde apreciaremos una representación de los datos sin normalizar y otra normalizada que nos permitan ver cómo se comportan las variables elegidas. Cabe destacar que en los siguientes comentarios excluiré la columna dedicada a “cross-regional” ya que los datos en esta región son muy pocos no pudiéndose extrapolar conclusiones. Sobre la figura 7 aparece el grupo de refugiados clasificado por las diversas regiones de estudio, los casos con nivel 3 donde el acuerdo trata con bastante peso este grupo aparecen repartidos de la misma forma sobre las regiones. Además, África (excl. MENA) y Europa-Eurasia tienen una disposición muy similar en su normalización.

Explorando los niños y las regiones de la figura 8, observamos como en África (excl. MENA) y en las Américas encontramos más referencias a niños en los acuerdos. Europa-Eurasia poseen el menor porcentaje de referencias a niños sobre sus acuerdos.

Estudiando otro de los grupos que aparecen sobre la figura 9, en este caso las razas, etnias o grupos nacionales, vemos como todos poseen un porcentaje de referencias a este grupo de entre un casi 50% hasta un 30% aproximadamente. Otro tipo de grupo considerado en el estudio son los grupos religiosos (figura 10) donde en general sobre todas las regiones se tienen menos referencias en los acuerdos que los grupos dedicados a las

razas, etnias o grupos nacionales. Esto ocurre para todas las regiones salvo en Asia y el Pacífico donde los grupos religiosos poseen más referencias en los acuerdos de paz.

La figura 11, junto con la figura 16 fueron explicadas en el apartado anterior. Para ver el peso de la igualdad en este tipo de acuerdos debemos fijarnos en la figura 12, como vemos, la región África (excl. MENA) y las Américas son las que más porcentaje abarcan respecto a su total, esto puede ser debido a la presencia de países subdesarrollados donde la igualdad es generalmente inferior que en los países desarrollados por lo que debe ser un punto a tener en cuenta sobre los acuerdos. Mediante la variable “SsrDdr” podremos ver la distribución sobre las regiones de las referencias a la “vuelta a la normalidad” en los acuerdos de paz (figura 13). La región con más referencias a este hecho, África (excl. MENA), es también la más afectada por las sucesivas guerras civiles y conflictos que asolan.

Sobre la figura 14 podremos visualizar el desarrollo de la reconstrucción-socioeconómica. Esta variable, muy importante de estudiar en los acuerdos de paz alcanzados después de guerras, aparece referenciada con bastante frecuencia sobre los acuerdos de todas las regiones tratadas. Las menciones a los medios de comunicación son tratados en la figura 15 donde África (excl. MENA) y las Américas encabezan las referencias.

Si estudiamos el grupo de refugiados en función del año obtendremos la figura 17. Lo primero que llama la atención es la desaparición de medidas de “nivel 3” a partir de 2016 justo en el periodo de la guerra de siria y la guerra civil de Sudán del Sur.

Las figuras 18, 19, 20 nos muestran la distribución por países del número de acuerdos sobre los años donde el número de acuerdos fue relativamente alto. (Estas aparecen estudiadas con detenimiento en páginas anteriores). Para finalizar, la figura 21 nos establece una relación entre la variable correspondiente a la “vuelta a la normalidad” y el grupo de refugiados. Si la “vuelta a la normalidad” se encuentra en el nivel 0, 1 o 2 predominará el nivel 1 para los refugiados, pero si la “vuelta a la normalidad” se encuentra en el nivel 3 tendremos aproximadamente el 50% de los acuerdos con refugiados al nivel 3.

2.4. LOS DATOS: PROCEDIMIENTO Y HERRAMIENTAS

El procedimiento seguido en la pregunta anterior se ha basado básicamente en ir probando diversos tipos de gráficas para ver cómo se distribuían los datos, que outlayers obteníamos, curiosidades... Tras un profundo estudio se optó por utilizar los diagramas de barras como elemento idóneo para plasmar la información al ser rápidamente entendibles y comparables.

A la hora de realizar esta PEC se ha usado el lenguaje Python a través de la plataforma de Jupyter Notebook y con el apoyo de los paquetes “matplotlib” y “seaborn”. Con más detalle el código de cada una de las gráficas mostradas por imagen aparecen en el archivo .html adjunto. En primer lugar, tras importar las librerías involucradas en el estudio, se cargan los datos obtenidos en un archivo “pandas”. Al tener un total de 265 atributos se realizará una selección de los que serán utilizados donde tomaremos la mayoría de atributos categóricos que nos muestran información sobre el contexto del acuerdo y realizaremos una previa selección de atributos numéricos que posteriormente se decidirá si son o no incluidos en la visualización en función de la información que logren transmitir. Tras esto, los atributos categóricos seleccionados serán los primeros en ser representados aportándonos una visión sobre el contexto de los datos. Cabe destacar que el nuevo atributo “Year” ha sido obtenido a partir de los cuatro primeros valores del atributo “Dat”. Una vez tenemos representados los datos categóricos en pantalla pasaríamos con el estudio de los atributos numéricos previamente clasificados a partir de los cuales poder sacar resultados concluyentes de los acuerdos. Para

finalizar, se trataría con mayor detenimiento los valores outlayers para estudiar qué ocurría sobre una determinada región, por ejemplo, los picos de acuerdos alcanzados en 1992, 2006 y 2015.

A lo largo del máster hemos ido aprendiendo a utilizar con cierta soltura las herramientas de Rstudio y Jupyter Notebook (Python) con las que poder realizar visualizaciones de los datos que nos permitan sacar conclusiones relevantes. Los ya mencionados paquetes “matplotlib” y “seaborn” de Python nos permiten obtener una gran variedad de gráficos con los que representar los datos en función de sus características. Por otro lado, la librería “ggplot2” de Rstudio tampoco se queda atrás aportándonos gráficas de alta calidad en esta plataforma.

Una de las herramientas datavis que espero saber manejar con solvencia tras finalizar esta asignatura es Tableau. Esta herramienta a diferencia de Rstudio y Jupyter está enfocada especialmente en facilitar el análisis visual e interactivo de los datos, favoreciendo que otras personas exploren y analicen conjuntos de datos de forma interactiva.

3. PEC3: INFORME DEL PROYECTO

3.1. INTRODUCCIÓN. HIPÓTESIS DEL PROYECTO.

Como ya vimos sobre el apartado 2.1 la hipótesis inicial marcaba: **obtener conclusiones sobre la identidad de los acuerdos de paz firmados.**

- Lo que involucraban.
- ¿A quién involucraban?
- ¿En qué regiones?

La hipótesis inicial previa al análisis exploratorio realizado sobre la PEC 2, nos muestra preguntas obvias pero necesarias que debíamos realizar sobre el fichero de datos.

Con respecto a aspectos técnicos del análisis exploratorio realizado en la práctica anterior, pudimos simplificar nuestro fichero de datos y de los 265 metadatos que aparecían en el fichero nos quedaríamos con los 17 más importantes sobre este tipo de acuerdos. A partir de este punto se estudiarían las diversas relaciones que podríamos establecer en los metadatos para contestar a la hipótesis inicial.

Gracias al análisis exploratorio, se conoció sobre qué países ha estado presente la figura del refugiado y cuántos acuerdos de paz se han tratado sobre este grupo. Por otra parte, podremos contrastar la calidad de los diferentes acuerdos establecidos en función de aspectos básicos para los refugiados como los derechos humanos, la igualdad, el desarrollo de la construcción socio-económica o la vuelta a la normalidad entre otros. Además, no nos dejaremos de lado el contexto histórico vital para poder entender completamente el fichero. De esta forma puede justificarse la elección de este tema como un conjunto de datos capaces de generarnos información sobre diversos acuerdos con referencias a refugiados. En nuestro caso, como ya hemos dicho antes, trataremos de identificar la identidad del acuerdo a partir de diferentes características.

La hipótesis inicial sobre el proyecto apenas cambió a la largo del análisis exploratorio, se continuaría con la idea general de identificar la identidad de los acuerdos de paz firmados. Por otra parte, para poder respaldar de forma consistente esta hipótesis inicial surgieron **nuevas preguntas** como:

- ¿Hacia dónde van enfocadas las medidas establecidas?
- ¿Cuál es el número de acuerdos obtenido en cada año?
- ¿Qué países están involucrados en el acuerdo?
- ¿Por qué hay repuntes de acuerdos sobre un cierto año?
- ¿Qué variaciones sufren las regiones en función de los atributos clasificadores como: la igualdad, el grupo de refugiados, el desarrollo socio económico...?

Estas son algunas de las diversas preguntas que surgieron en el análisis exploratorio y esperamos responder tras la visualización. Finalmente y tras un completo análisis exploratorio sobre 17 metadatos se optaría por analizar una parte de estos 17 metadatos e introducir algunos nuevos necesarios para la elaboración de la visualización final con la que poder responder a las preguntas anteriores.

Con la visualización final intentaremos que los usuarios sean capaces de obtener información de manera eficiente sobre la identidad de los acuerdos alcanzados para que cada uno pueda sacar sus propias conclusiones. Por otra parte, también se introducirán elementos que aumenten la interactividad del usuario con la visualización con el objetivo de mejorar la experiencia.

3.2. BLOQUE DE DATOS.

3.2.1. DESCRIPCIÓN DEL DATASET A UTILIZAR EN LA VISUALIZACIÓN.

Sobre el fichero de datos nos encontraremos con un total de 265 metadatos, categóricos y numéricos, rellenos por 476 acuerdos de paz diferentes. Sobre el análisis realizado en la PEC2 se haría una primera selección de cara al proyecto final donde se tuvo en cuenta dos parámetros:

1. Que la información aportada tuviese o pudiera tener relación con el grupo de refugiados.
2. De ser un atributo numérico, la suma sobre sus filas fuese elevada. Esto indicaría una gran presencia de dicho atributo sobre los acuerdos de paz.

Teniendo en cuenta estos criterios se seleccionarían 17 metadatos, que dividiríamos en atributos contextuales o atributos sobre las características del acuerdo. Tras el análisis exploratorio de los datos realizado en la práctica anterior, me percaté de que ciertos atributos contextuales como “Date”, “Status” o “Stage” entre otros, no serían necesarios sobre la visualización final. Por otra parte, otros elementos que no había utilizado como “Con”, “LOC1ISO”, “LOC2ISO” o “AgtId” serían útiles de cara a la visualización final. De esta forma obtendríamos el siguiente dataset:

ATRIBUTOS CONTEXTUALES

Atributo	Características
Con	Territorios, principalmente países donde el conflicto es originado. También se incluyen regiones
Reg	Región donde el conflicto tiene lugar: <ul style="list-style-type: none">• Europe and Eurasia• Africa (excl. MENA)• Asia and Pacific• Americas• Middle East and North Africa• Cross-regional
Year	Nuevo atributo obtenido a partir de “Date” que nos muestra exclusivamente el año en el que fue firmado el acuerdo.
LOC1ISO	Localización del conflicto entre estados. (Asignada aleatoriamente)
LOC2ISO	Localización del conflicto entre estados. (Asignada aleatoriamente)
AgtId	Código identificativo del acuerdo.

ATRIBUTOS SOBRE LAS CARACTERÍSTICAS DEL ACUERDO

- Grupos.

Atributo	Características
GCh	Referencia a niños. Valores: 0, 1, 2, 3.
GRa	Referencia a razas/etnias/grupos nacionales. Valores: 0, 1, 2, 3.
GRe	Referencia a grupos religiosos. Valores: 0, 1, 2, 3.
GRef	Referencia a grupos de refugiados. Valores: 0, 1, 2, 3.

- Derechos humanos e igualdad.

Atributo	Características
HrGen	Derechos humanos. Valores: 0, 1.
EqGen	Igualdad. Valores: 0, 1, 2, 3.
Med	Menciones a los medios de comunicación. Valores: 0, 1, 2, 3.

- Reconstrucción socio-económica.

Atributo	Características
Dev	Desarrollo de la reconstrucción socio económica. Valores: 0, 1, 2, 3.

- Sector de seguridad.

Atributo	Características
SsrDdr	Mención sobre la “vuelta a normalidad”, es decir, se tienen en cuenta procesos de desmilitarización y reintegración. Valores: 0, 1, 2, 3.

Este sería el dataset final sobre el que trabajaremos en las siguientes prácticas, tendremos un total de 15 atributos que nos permitirán obtener la información para responder a las preguntas planteadas sobre los datos. En el siguiente apartado veremos si es necesario realizar alguna transformación sobre los datos.

3.2.2. TRANSFORMACIONES SOBRE EL DATASET

Los datos tratados sobre el análisis exploratorio realizado en la práctica anterior se encuentran listos para su representación, es decir, ya se realizó su análisis y limpieza correspondientes. Cabe destacar la introducción del atributo “Year” sobre la práctica anterior que nos permite mostrar la evolución anual del número de acuerdos de paz con referencias a refugiados.

Sobre el dataset presente en el apartado 2.1 se añadirán 9 columnas nuevas donde dispondremos de todos los atributos sobre las características del acuerdo ('GCh', 'GRa', 'GRef', 'GRe', 'HrGen', 'EqGen', 'Med', 'Dev', 'SsrDdr') pero normalizados. Esta normalización nos ayudará a extraer resultados coherentes de ciertas visualizaciones donde será necesario trabajar con los datos dentro de un mismo rango, sin embargo, los datos sin normalizar también serán de utilidad sobre ciertas gráficas y por ello se ha optado por su conservación. De esta forma tendremos los atributos normalizados y sin normalizar listos para ser representados.

Continuando con los atributos sobre las características del acuerdo, hay que resaltar que aunque se tratan de valores numéricos, pueden ser tratados como categóricos ya que se encargan de clasificar la presencia del atributo sobre un determinado acuerdo. Es importante tener en cuenta que dependiendo del programa y el gráfico a representar tendremos que optar por la categorización de estos datos o dejarlos como valores numéricos.

Los nuevos atributos introducidos sobre esta visualización serán limpiados y analizados antes de su uso para evitar problemas. Por ejemplo, se eliminarán los paréntesis innecesarios que aparecen sobre la columna “Con”, logrando una mayor homogeneización. Una de las posibles transformaciones futuras a realizar sería la extracción

de cada uno de los países involucrados en cada acuerdo, este apartado será estudiado sobre la siguiente práctica.

Por último, con el objetivo de evitar aquellos valores perdidos, eliminaremos el único valor que se dispone sobre la región “Cross-Regional” ya que con tan solo un acuerdo es imposible obtener conclusiones relevantes.

3.2.3. EVIDENCIAS OBTENIDAS SOBRE LA PEC2.

Sobre la PEC2 logramos obtener evidencias de gran calado que podrán ser de gran ayuda a la hora de responder nuestra hipótesis inicial.

Atributos Contextuales

En primer lugar se representaría la evolución del número de acuerdos sobre refugiados desde 1990 hasta 2019, donde sus picos alcanzados en 1992, 2006 y 2015 llamaban mucho la atención. Se estudiaría la distribución de del número de acuerdos para esos determinados años en función de los países implicados en el estudio obteniéndose resultados bastante curiosos.

- En 1992, Bosnia Herzegovina y Yugoslavia generan múltiples acuerdos donde se ven involucrados los refugiados como consecuencia del final de la guerra Fría. [3]
- En 2006, tras la finalización de la guerra civil de Nepal, este país generaría una serie de acuerdos de paz donde los refugiados serían referenciados. [4]
- En 2015, la tristemente conocida guerra de Siria y la guerra civil de Sudán del Sur serán las causantes, en su mayoría, de este tipo de acuerdos.

Como ya se referenciaba sobre la introducción, la contextualización histórica juega un papel muy importante a la hora de entender estos acuerdos de paz.

Por otra parte, los acuerdos de paz en referencia a refugiados en función de las regiones, nos dan una visión general sobre la presencia de este tipo de acuerdos. África(excl. Mena), es la región con mayor número de acuerdos de este tipo debido principalmente a guerras civiles como la de Sudán del Sur o Somalia y conflictos como el de Chad-Sudán o el de Nigeria. África es seguida por Europa y Eurasia debido al final de la Guerra Fría y la Guerra de Siria entre otros.

Además, se representarían otros atributos contextuales como el tipo de conflicto, la región, el tipo de acuerdo, el tipo de conflicto entre regiones y el estado del acuerdo. En principio, aunque estos atributos pueden aportarnos cierta información de utilidad no se tendrán en cuenta a la hora de construir la visualización final ya que se pretende focalizar la atención en el lugar del acuerdo, el año y sus características principales.

Atributos sobre las características del acuerdo

De la práctica anterior podemos concluir que estas características son las que diferenciarán a un acuerdo de otro, su estudio en función de las diferentes regiones nos ha permitido establecer tendencias sobre las características de los acuerdos. Sin lugar a dudas, estos atributos estarán presentes sobre la visualización final tratando de responder a la identidad del acuerdo.

3.2.4. ¿CÓMO VAMOS A ENFOCAR LA VISUALIZACIÓN?

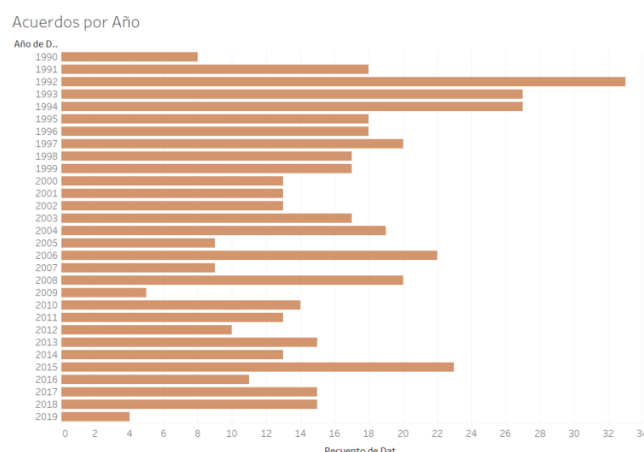
Sobre la visualización final debemos tener muy presente que la relación entre los datos debe basarse en dos aspectos básicos:

- **Relacionar.** Las gráficas presentes sobre los atributos deben estar enlazadas permitiéndonos extraer información sobre todos (o la mayoría) los elementos presentes sobre la visualización.
- **Comparar.** La comparación será otro de los elementos clave, nos interesa ver en qué se diferencian los acuerdos de determinadas regiones, sobre cuáles hay más acuerdos de paz...
- **Distribuir.** Dispondremos diversas gráficas que nos permitirán ver la distribución del atributo en cuestión. Un ejemplo es la distribución del número de acuerdos alcanzados cada año.

Estas relaciones son las que se presentan sobre el primer diseño, en la visualización final podremos tener ciertas modificaciones. Una vez conocidas las relaciones y evidenciado el tipo de atributos que formarán parte de nuestro estudio final, podremos pasar a determinar los gráficos que aparecerán sobre la visualización final.

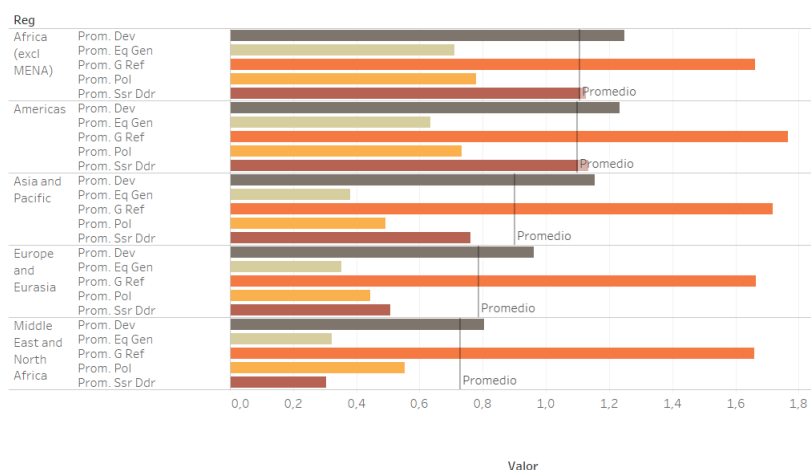
Diagramas de barras

Tal y como se vio sobre la práctica anterior, a pesar de la sencillez, este tipo de gráficos logran cumplir su función con creces y mostrarnos las variables categóricas de forma muy visual. Al disponer de una gran cantidad de variables categóricas y de otras tantas numéricas que pueden convertirse en categóricas, el diagrama de barras tendrá gran presencia sobre nuestra visualización final. La representación de la evolución del número de acuerdos a lo largo de los años quedará representada tal y como se muestra sobre la imagen de la derecha.



Por otra parte se ha optado por recopilar la media de diferentes atributos que recopilan información sobre las características del acuerdo en función de las regiones. Obtendremos lo siguiente:

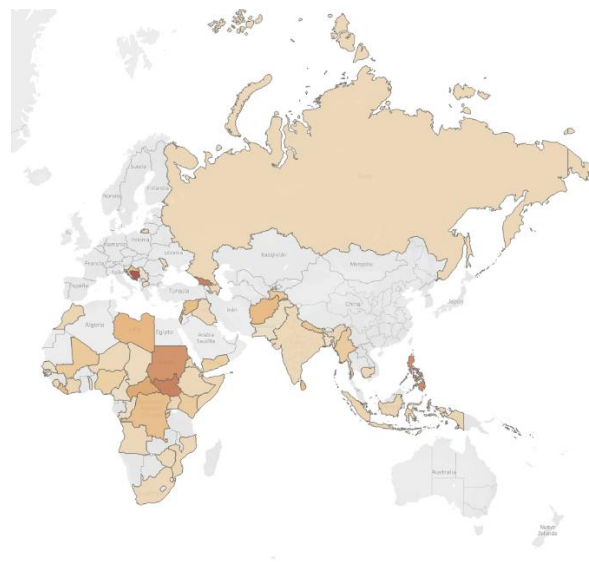
ATRIBUTOS POR REGIÓN



De esta forma podrá compararse como varían las características principales de los acuerdos en función de las diferentes regiones sometidas al estudio. (Sobre la imagen anterior no aparecen todas las características recogidas sobre el dataset, es un simple modelo)

Mapa

Al disponer de las diferentes regiones donde se han llevado a cabo los acuerdos, podría representarse un mapa mundial donde los países estuviesen coloreados en función del número de acuerdos sobre ellos. Un ejemplo podría ser algo así:

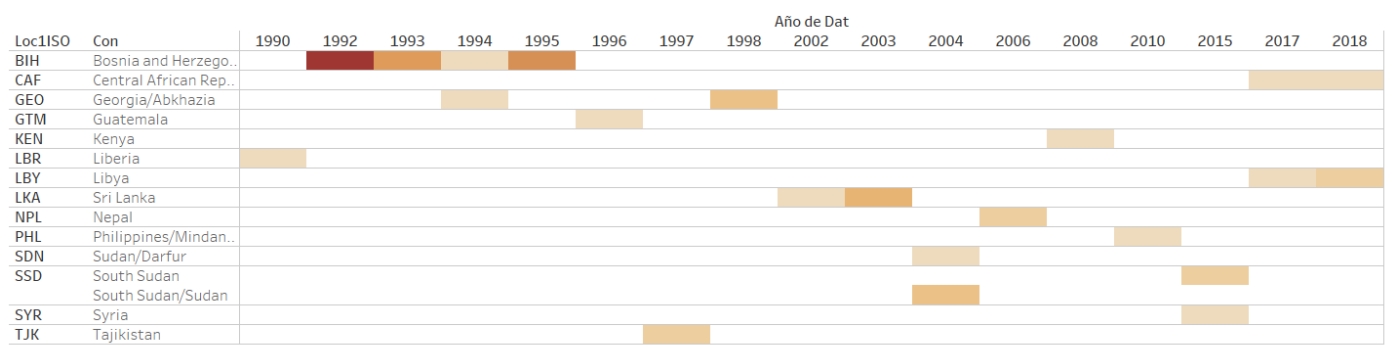


De forma directa puede observarse la distribución del número de acuerdos sobre las diferentes regiones, a mayor oscuridad del color, mayor número de acuerdos.

Mapa de calor

Un mapa de calor puede ayudar a visualizar la relación entre dos variables categóricas. Por ejemplo, si representamos el país o países involucrados en el acuerdo frente al año obtendremos un mapa con diferentes colores que nos indicará la cantidad de acuerdos sobre dicho país y año.

Países con más de 4 acuerdos de paz



Cabe destacar que no solo disponemos de estas herramientas para configurar la visualización y aunque estas son las que se han pensado como principales para la creación del dashboard, también podrían aparecer algunos gráficos de estos tipos:

- Diagrama de barras apiladas. A la hora de representar dos variables categóricas es de utilidad utilizar este tipo de diagramas que nos permitirá representar la distribución de una variable categórica sobre una de las categorías de la otra variable.
- Diagrama de barras paralelas. Del mismo modo que ocurría en el caso anterior podremos representar dos variables categóricas apoyándonos en este tipo de diagrama.
- Mapas de árbol. Este tipo de visualizaciones pueden ser útiles si queremos mostrar en forma de área el tamaño de una determinada categoría.

Con estos seis ejemplos (entre muchas más visualizaciones) quería poner sobre la mesa las múltiples opciones de enfoque que tenemos sobre la visualización.

3.2.5. GOBERNANZA DE DATOS.

Como bien se ha comentado en el apartado anterior, una de las relaciones que queremos fomentar en la visualización es la relación y comparación de los atributos, esto se conseguirá a través de una buena interacción del usuario con los datos. Para llevar a cabo esta interacción nos ayudaremos de la herramienta *Tableau* sobre la que crearemos un dashboard que nos permitirá añadir los siguientes elementos:

- **Filtros:** Los filtros nos permitirán, como bien indica la palabra, filtrar sobre las gráficas presentes en el dashboard la información correspondiente a un elemento concreto.
- **Resaltar:** Este elemento es similar al anterior, pero en lugar de filtrar la información, nos resalta la información correspondiente al elemento sobre el que interactuamos.
- **URL:** Podemos añadir URL a una determinada gráfica que ayuden a completar el dashboard.
- **Cuadros de información:** La introducción de cuadros de información explicativos permite aumentar la facilidad de comprensión hacia el usuario.
- **Ventanas emergentes:** La aparición de ciertas ventanas explicativas tras clicar sobre un elemento también aumentarán el grado de interacción con el usuario

A partir de estas acciones que nos proporciona *Tableau* se podrán relacionar los gráficos descritos en el apartado anterior para mejorar la experiencia del usuario. Además, cada una de las hojas que constituyan el dashboard podrá contar con filtros propios que nos permitan seleccionar la información deseada sobre una determinada gráfica.

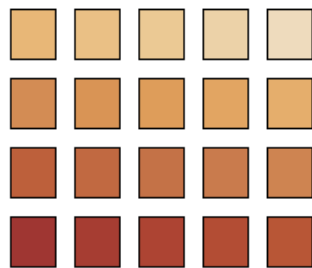
La idea principal sobre el diseño es que el mapa adquiera la reseña de gráfico principal sobre el que construir todos los demás aunque por otro lado, todos los gráficos presentes en la visualización estarán conectados para mejorar la interactividad y la experiencia del usuario. En definitiva, sobre el modelo final se intentará obtener un equilibrio entre las interacciones y la eficacia de la visualización.

3.3. DISEÑO MOCKUP.

3.3.1. COLORES.

Con el objetivo de que la visualización adquiera un toque profesional y homogéneo se trabajará sobre una única gama de colores. La **paleta color café**, que nos permitirá representar los colores más oscuros como aquellos con mayor número de datos y los colores más claros como aquellos con menor representación, será utilizada

tanto para el mapa geográfico como para el heat map. Siguiendo con esta gama de colores los acuerdos por año también adquirirán este color.



Paleta/Degradado color café (Tableau)

Por otra parte, se propondría la paleta de colores “Piedra Miller” sobre la PEC3 pero finalmente no será considerada en el análisis final.



Paleta “Piedra Miller”

De esta forma sobre la visualización final únicamente consideraremos la paleta color café.

3.3.2. TIPOGRAFÍA.

La tipografía usada en los títulos será la **SAVEUR SANS SEMI-BOLD**, esta fuente es limpia por lo que nos facilitará una buena legibilidad, da un aspecto moderno a la visualización y su aspecto simple proporciona fuerza sobre la misma.

Por otra parte sobre los títulos de las pestañas emergentes al situar el cursor sobre algún elemento, elegiremos la tipografía **Trebuchet MS** al no poder mostrarse la tipografía **SAVEUR SANS SEMI-BOLD** en estos títulos. Cabe destacar que de añadir algún comentario secundario sobre la visualización se optaría por la utilización de la fuente “Tableau Book” o similares.

3.3.3. ESTRUCTURA Y FORMATO.

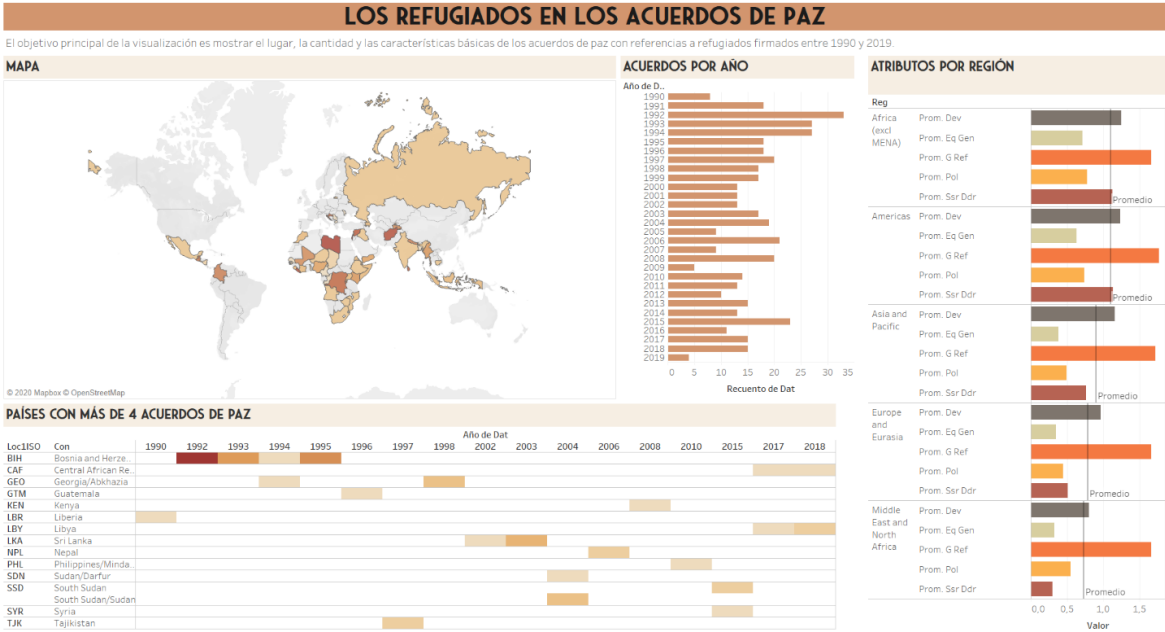
La propuesta de visualización presente en esta PEC, estará formada de 4 bloques.

- El **mapa** que se encontrará en la parte superior izquierda y será una de las partes principales de la visualización.
- El **número de acuerdos por año**, representado como un diagrama de barras que nos ayuda a ver la evolución de este tipo de acuerdos. Estará situado en la parte superior central.
- El **mapa de calor**, se encontrará ocupando la parte inferior izquierda y central de la visualización. Se dispone de esta forma para que puedan apreciarse todos los años y países involucrados.
- Las **características del acuerdo en función de las regiones** vendrán representadas de forma alargada en forma de diagrama de barras sobre la parte lateral derecha.

El proyecto se presentará de forma digital a través de un enlace a *Tableau Public* permitiendo a los usuarios interactuar con el mismo como ya se ha expresado en apartados anteriores. Por otro lado, quiero destacar que esta estructura es provisional y representa la propuesta de diseño, el diseño final podría sufrir modificaciones.

3.3.4. PROPUESTA DE DISEÑO.

Una propuesta de diseño podría ser la siguiente, donde como vemos se apreciarán los títulos con la tipografía elegida, los colores adecuados y las gráficas descritas sobre el apartado anterior.



Esta propuesta no es más que una simple maqueta de prueba construida única y exclusivamente para tener referencias en cuanto a la relación, estructura y colores de la visualización final. No obstante, está sujeta a sufrir modificaciones con el fin de mejorar la experiencia hacia el usuario.

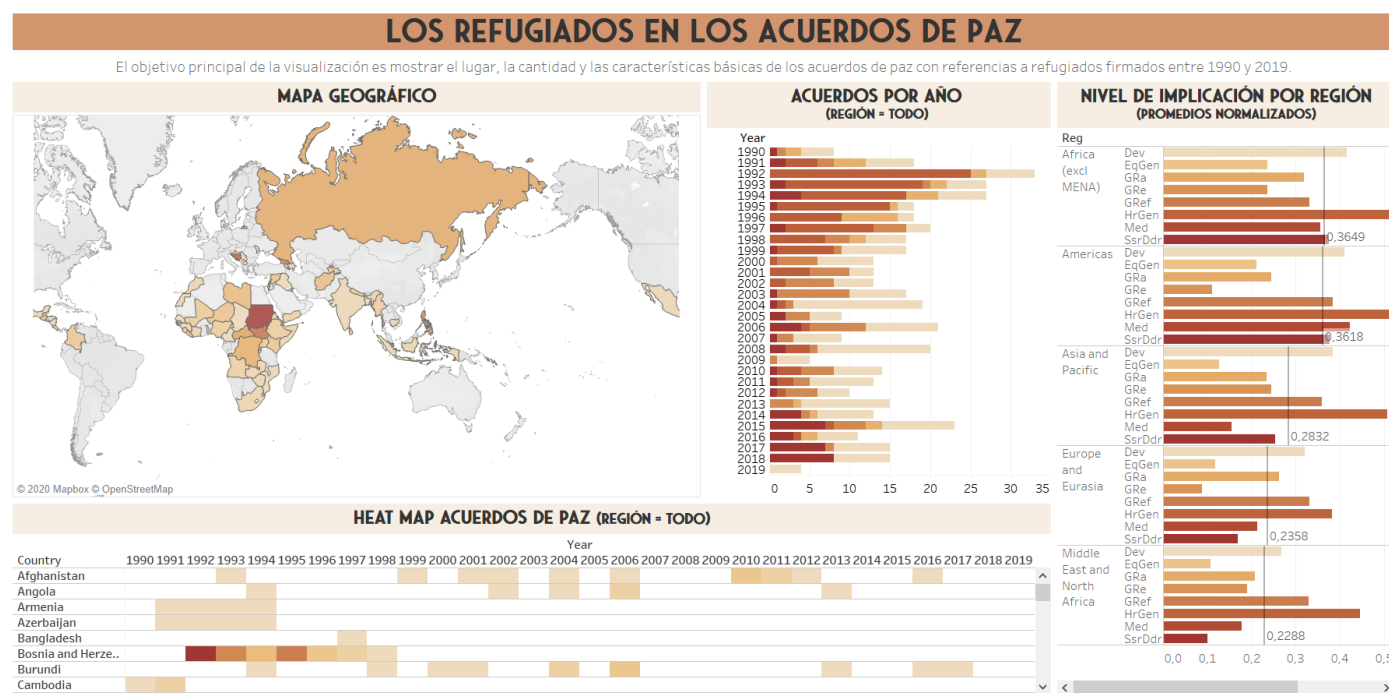
4. PEC4: CREACIÓN DE LA VISUALIZACIÓN

4.1. DESARROLLO DEL PROYECTO

Tras comenzar con más dudas que certezas, empezábamos el proyecto sin saber muy bien donde poner la siguiente piedra pero poco a poco hemos ido moldeando los datos tomados hasta lograr una visualización bastante interesante. Al comienzo nos centramos en realizar diagramas de barras de diferentes formas utilizando exclusivamente Python para las representaciones, con ello queríamos ver principalmente como se distribuían los datos que ponían en contexto al acuerdo. Además, se intentaría descifrar alguna tendencia llegando a algunas conclusiones bastante interesantes como los refugiados de la Guerra Fría, la Guerra Civil de Nepal o la Guerra de Siria. En esta primera PEC, se tratarían a los atributos que proporcionaban las características del acuerdo como categóricos e intentaríamos encontrar tendencias relacionando diferentes campos entre sí. Entre otras cosas, una de las grandes conclusiones obtenidas de la PEC2 sería la importancia del contexto sobre los datos para enriquecer la información.

Con ya una parte del camino recorrida, afrontaríamos la siguiente PEC con la idea de focalizarnos más en aquellos datos que habíamos encontrado más útiles. Por ello, haríamos una nueva selección de los atributos que se someterían a estudio siendo estos ya los definitivos. Además, en esta PEC3 decidiríamos utilizar los atributos que proporcionaban las características del acuerdo como numéricos y no como categorías, por lo tanto se normalizarían todos los datos entre 0 y 1 para poder manejarlos a todos por igual. Sería en esta PEC3, cuando comenzaríamos a representar los datos con Tableau y descubriríamos diversas formas de representación como mapas geográficos o mapas de calor entre otros. Sobre esta PEC estableceríamos una primera propuesta de diseño muy similar al diseño final.

A lo largo de esta parte de la práctica nos centraremos en describir la visualización final y los cambios que se han introducido con respecto a la última PEC. El diseño final es el siguiente:



DISEÑO FINAL DE LA VISUALIZACIÓN

El proyecto final es elaborado íntegramente mediante Tableau y publicado en Tableau Public. A través de este enlace puedes acceder a la visualización mediante este enlace:

- URL: <https://public.tableau.com/profile/mario.mart.nez.garc.a#!/vizhome/Losrefugiadosenlosacuerdosdepaz/Dashboard1?publish=yes>

En esta última parte de la asignatura nos centraremos en explicar profundamente la visualización final, remarcar con exactitud qué puntos han sido modificados sobre la misma y mostrar una comparación entre los gráficos de la propuesta inicial y los finales en el caso de haber sufrido varias modificaciones. Tras ello, abordaremos una descripción técnica del proyecto donde nos detendremos en el procedimiento seguido tanto en Python como en Tableau. Por último, trataremos de describir los retos que han ido apareciendo en la construcción de la PEC, el esfuerzo que ha supuesto y los resultados obtenidos.

4.2. EXPLICACIÓN DEL PROYECTO. ¿QUÉ CUMPLE Y QUÉ NO CUMPLE DEL PROYECTO INICIAL?. DESCRIPCIÓN DE LOS GRÁFICOS EMPLEADOS.

Cabe destacar que la visualización final conserva la misma estructura que se estableció sobre la propuesta de diseño y por tanto, se compondrá exclusivamente de un único dashboard donde primará la sencillez, la claridad y las interacciones para que el usuario pueda sacar sus propias conclusiones de forma rápida y eficaz. Por otro lado, contaremos con 4 hojas que darán forma a la visualización: el mapa, la distribución de los acuerdos según el año, el mapa de calor, y el nivel de implicación de las regiones en los acuerdos. A pesar de la más que evidente continuidad seguida sobre la propuesta de diseño se han realizado una serie de cambios, la mayoría de ellos internos, que nos ayudarán a aumentar drásticamente la calidad de nuestra visualización mostrando aquellos datos con los que realmente queremos trabajar.

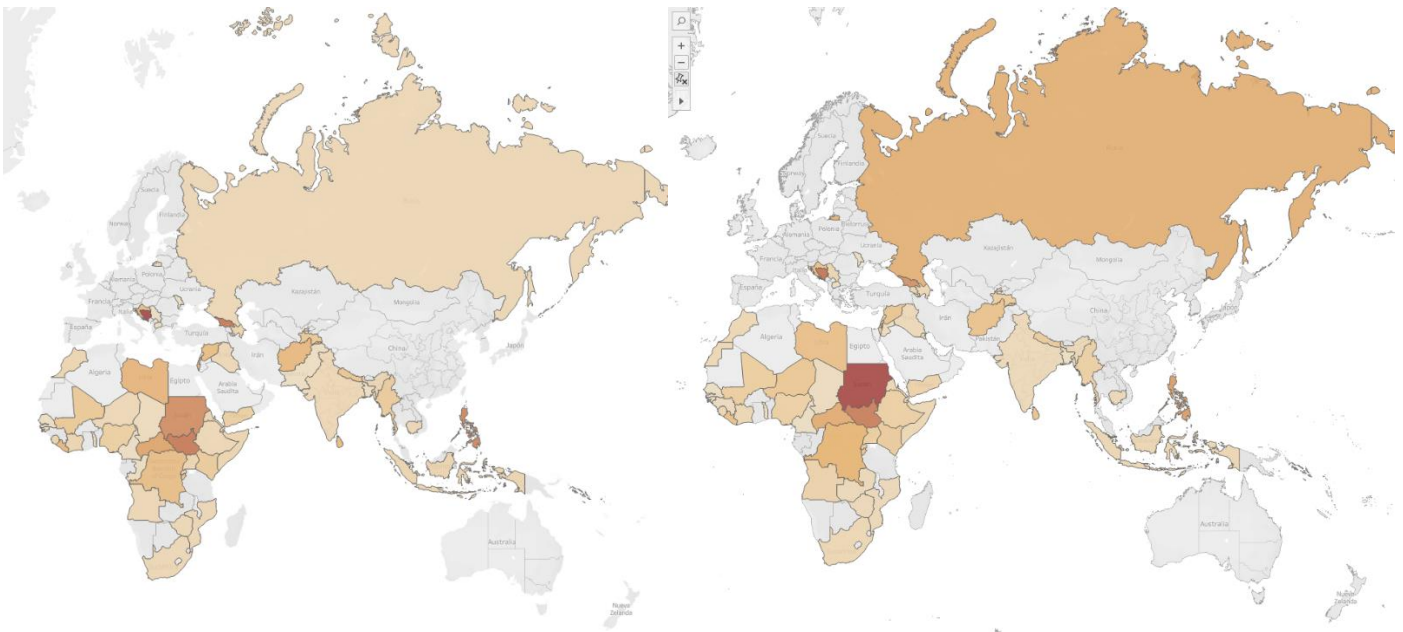
Por otro lado, la tipografía seguirá siendo la misma dando protagonismo a **SAVEUR SANS SEMI-BOLD** para intentar transmitir la idea de visualización moderna e innovadora. Además, la gama de colores de la paleta café continuará aportándonos la seriedad y fortaleza que lograba transmitirnos sobre la propuesta de diseño y se conseguirá aumentar el dinamismo mediante la animación simultánea de las hojas presentes en la visualización evitando los cambios repentinos ante cualquier interacción.

A lo largo de este apartado nos centraremos en mostrar las nuevas aportaciones realizadas sobre cada una de las hojas de la PEC3 que darán lugar a las hojas definitivas de la visualización. Además, en este apartado se desarrollará una descripción completa de las interacciones construidas en forma de *acciones* sobre cada una de las hojas o del dashboard.

4.2.1. MAPA GEOGRÁFICO

- **Localización.** Este mapa lo encontraremos en la esquina superior izquierda de la visualización, siendo una de las partes principales de la misma
- **Modificaciones.** Sobre el mapa geográfico representado en la propuesta de diseño, no aparecía realmente el número de acuerdos que había firmado cada país que es lo que en un principio quería

mostrarse. Nuestro fichero de datos transformado “datatransform.xlsx” (Contiene los atributos seleccionados para el estudio) nos muestra una lista con el conjunto de acuerdos, pero no tiene en cuenta si sobre un acuerdo aparecen involucrados varios países. Las localizaciones “LOC1ISO” y “LOC2ISO”, nos aportaban información parcial que nos permitían construir un irreal mapa del número de acuerdos del país. Por ello, se optaría por extraer cada uno de los países involucrados en cada acuerdo y con esto construir un nuevo mapa donde poder apreciar el número de acuerdos con referencias a refugiados que cada país había firmado realmente.



MAPA IRREAL: CONSTRUIDO A PARTIR DE LOC1ISO

MAPA REAL: CONSTRUIDO A PARTIR DEL N° REAL DE ACUERDOS DE CADA PAÍS
(DEFINITIVO)

De esta forma pasaríamos de un mapa poco representativo a uno donde verdaderamente podríamos apreciar la distribución real de los acuerdos. Los detalles técnicos sobre la extracción de datos serán explicados sobre el apartado “4.2. Descripción técnica del proyecto”

- **Interactividad.** Son diversos los elementos interactivos que podemos encontrar sobre la visualización, uno de ellos es la presencia de descripciones emergentes al situar el cursor sobre alguno de los países marcados en color café. Esta descripción contiene el nombre del país, la región a la que pertenece, el número de acuerdos del país y los promedios de los atributos GRef y HrGen normalizados.

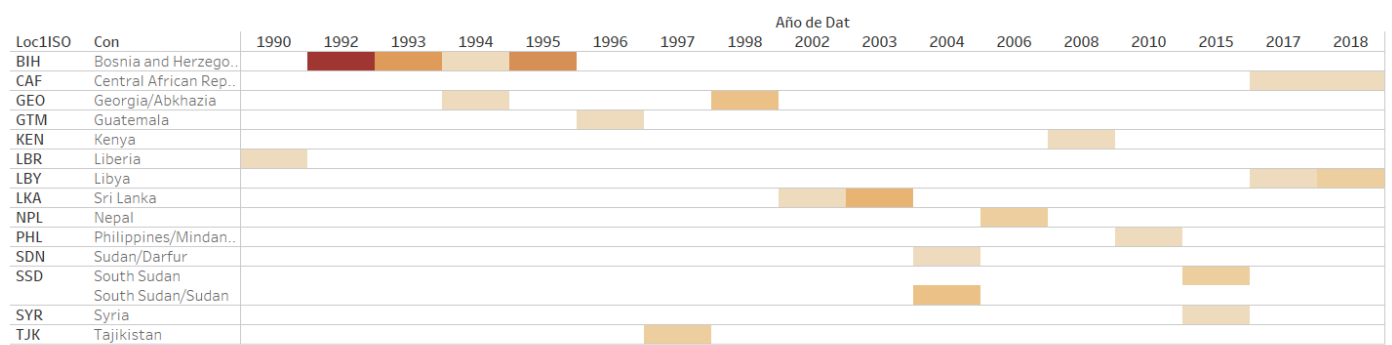
Por otro lado, añadiríamos algo que demandábamos sobre la PEC3, las URLs que serían incluidas mediante una acción sobre la hoja de trabajo. Tras seleccionar un determinado país en el mapa, nos mostrarían sobre la descripción emergente un nuevo elemento que actuaba de enlace hacia la Wikipedia del país. Un ejemplo de este enlace es “Wikipedia Russia”, donde el nombre del país cambia en función del país que seleccionemos. De esta forma logramos enriquecer nuestra visualización mediante enlaces externos que ayudan a ponernos en contexto sobre la situación vivida en un determinado país.

Por último, se crearían acciones en el dashboard donde estableceríamos filtros que se ejecutasen al seleccionar un determinado país en el mapa y filtrasen dicha información sobre el resto de hojas presentes en la visualización.

4.2.2. MAPA DE CALOR

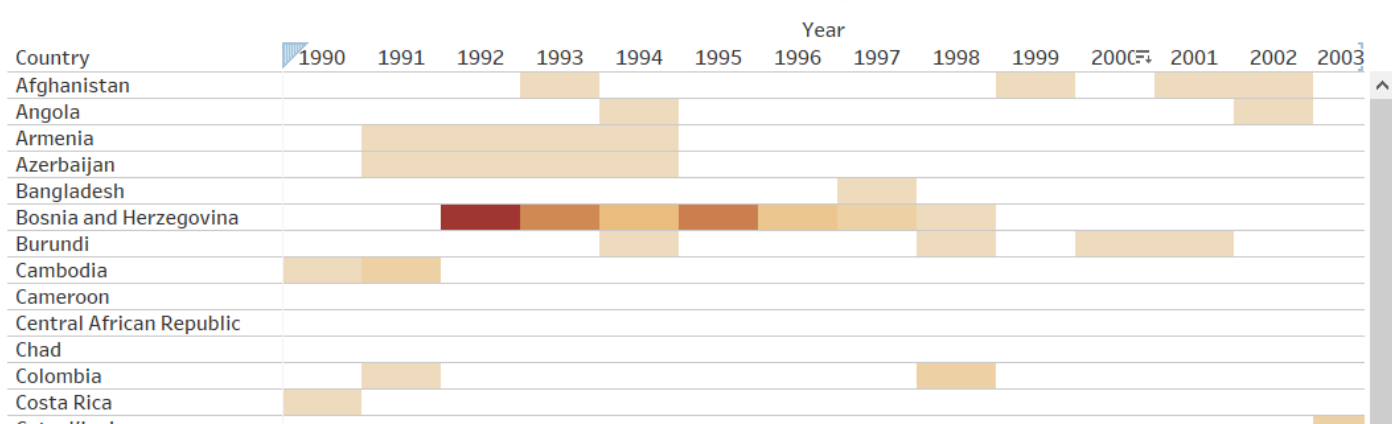
- **Localización.** El mapa de calor continuará situado en la región inferior izquierda, actuando de soporte para conocer con mayor profundidad la distribución de los acuerdos de paz sobre un determinado país.
- **Modificaciones.** En primer lugar, pasaremos de un mapa de calor donde se representaban exclusivamente los países (o conjuntos de ellos) con más de 4 acuerdos de paz a un mapa de calor donde tendrán hueco todos los países del fichero. La problemática de los países también la encontramos sobre el mapa de calor, al no disponer de los países debidamente separados aparecían todos los países que estaban involucrados en un mismo acuerdo. Por ejemplo: 'Cameroon/Nigeria'. Con la corrección realizada sobre los países se consigue un mapa de calor real sobre la distribución del número de acuerdos firmados en cada país.

Países con más de 4 acuerdos de paz



HEAT MAP CON VARIOS PAÍSES INVOLUCRADOS SOBRE UN DETERMINADO ACUERDO.

HEAT MAP ACUERDOS DE PAZ (Región = Todo)



HEAT MAP EXCLUSIVO PARA CADA PAÍS. (DEFINITIVO)

La corrección sobre los países nos ha permitido estudiar el número de acuerdos firmados sobre cada uno de ellos en función del año. Pasamos de un mapa de calor donde aparecen varios países o regiones por fila a uno que nos clasifica exclusivamente cada uno de los países.

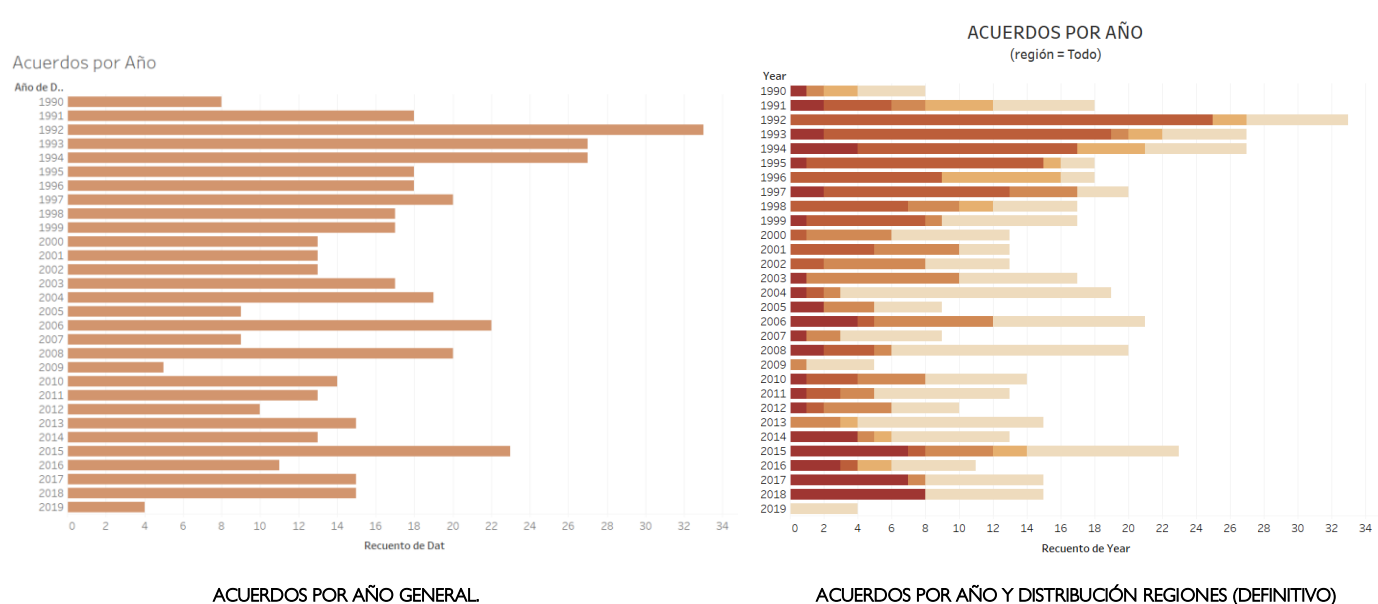
- **Interactividad.** El mapa de calor también dispondrá de descripciones emergentes, al situar el cursor sobre cualquier elemento de la hoja nos mostrará el nombre del país, la región y el número de acuerdos de paz firmados sobre el año seleccionado.

Al continuar trabajando con países, optaremos por incluir las URLs que nos enlazan con la Wikipedia de un determinado país. Seleccionando cualquier elemento de la fila correspondiente en un país aparecerá sobre la descripción emergente el enlace a la Wikipedia del país de la misma forma que en el mapa geográfico. Con esta simple acción conseguimos dar importancia al contexto, considerado de vital importancia para entender la gran cantidad de acuerdos en determinadas regiones.

Además de estos elementos, no nos olvidaremos de las acciones en el dashboard que nos permitirán crear filtros de apoyo entre el mapa de calor y el resto de hojas obteniendo representaciones diferentes en función de lo seleccionado.

4.2.3. ACUERDOS POR AÑO

- **Localización.** El diagrama de barras continuará conservando su posición en la parte superior central, y se dispondrá de forma alargada. Aunque es el gráfico que menos presencia tiene sobre la visualización final es indispensable para la comprensión de los datos.
- **Modificaciones.** Este diagrama de barras apenas sufrirá modificaciones, su estructura seguirá siendo la misma pero incluiremos divisiones sobre sus barras que nos mostrarán la distribución de acuerdos entre las regiones sobre un determinado año.



Así, sobre cada una de las barras se representarán hasta 5 colores de la paleta color café en función de cada una de las regiones que tuvo acuerdos de paz con referencias a refugiados sobre un determinado año.

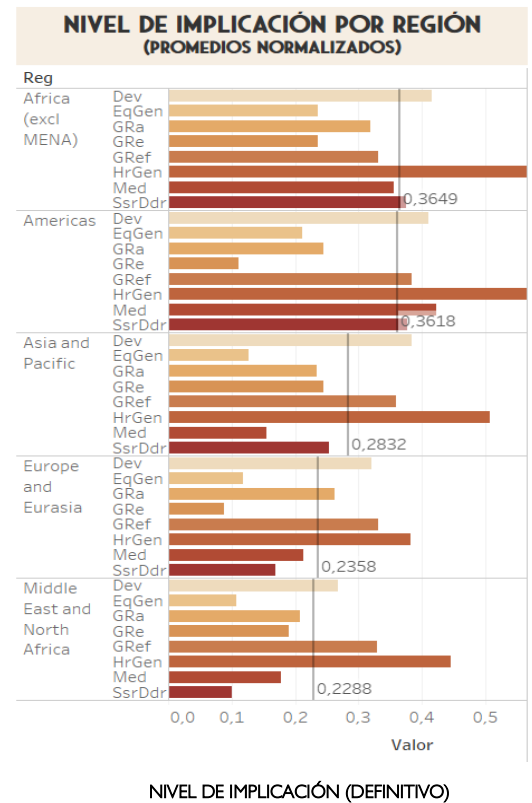
- **Interactividad.** Aunque el diagrama pueda parecer algo caótico en un principio, gracias a la interactividad sobre la gráfica se elimina esa cierta incertidumbre. El diagrama dispondrá de descripciones emergentes que nos mostrarán el año, la región y el número de acuerdos al superponer el cursor sobre alguno de los bloques.

Por otra parte, también se incluirán acciones en forma de filtros sobre el dashboard para favorecer al conectividad con el resto de hojas presentes en la visualización.

4.2.4. NIVEL DE IMPLICACIÓN POR REGIÓN

- **Localización.** Este compendio de diagramas de barras conservará su posición lateral derecha abarcando gran parte de la visualización al ser un elemento vital para contrastar las características del acuerdo entre las regiones.
- **Modificaciones.** Sobre esta gráfica añadiremos todos los atributos normalizados seleccionados en nuestro estudio calculando la media de cada uno de ellos sobre cada región. Además añadiremos una línea negra que nos determine la media global del conjunto de atributos sobre una determinada región.

Por otra parte, cambiaremos la paleta seleccionada para esta gráfica, utilizaremos la paleta color café para homogeneizar la visualización y aumentar la estética



- **Interacciones.** Dispondremos de una descripción emergente que nos mostrará la región y el valor medio del campo al situar el cursor sobre cualquiera de las barras. En esta última hoja tampoco nos olvidaremos de las acciones sobre el dashboard, añadiendo filtros que nos conectarán con el resto de hojas al seleccionar un determinado elemento sobre esta.

4.3. DESCRIPCIÓN TÉCNICA DEL PROYECTO

A lo largo de este amplio proyecto hemos usado dos herramientas que nos han permitido descubrir tendencias, representar información o en algunos casos hasta realizar modificaciones sobre los datos. Estas dos herramientas de las que hablo son Python y Tableau que además dividirán el apartado en dos partes.

4.3.1. PYTHON

Jupyter Notebook mediante el lenguaje Python, sería la primera herramienta utilizada sobre este conjunto de prácticas. A partir de este lenguaje se modelaron dos ficheros de gran relevancia que dan forma a la práctica y explicaremos a continuación: “PEC2_CODE_MMG.ipynb” y “TRANSFORM_MMG”

“PEC2_CODE_MMG.ipynb”

En este archivo encontraremos toda la información extraída de la PEC2. Mediante el uso de librerías de Python como “pandas”, “numpy”, “matplotlib” o “seaborn” se estructuraría una práctica de gran bagaje que nos permitió conocer los datos por primera vez y establecer nuestras primeras conclusiones sobre ellos.

La librería “pandas” nos permitiría leer el csv donde aparecerían los datos recopilados con 476 filas y 265 campos a tratar. Tras disponer de los datos sobre la plataforma estudiaríamos las variables categóricas mediante diagramas de barras a través de la librería “seaborn” y añadiríamos la variable año al conjunto de datos final. Una vez conseguidos los diagramas sobre las variables categóricas, se estudiarían las variables numéricas que fueron consideradas más importantes en los acuerdos con referencias a refugiados. Las representaciones se harían en función de las regiones y alguna variable numérica que realmente se trataba como categórica al disponer como mucho de 4 números enteros (0,1,2,3). Una vez más la librería “seaborn” junto con “matplotlib” nos servirían de apoyo para representar diferentes diagramas de barras. En último lugar, se intentaría representar la distribución de países sobre los picos obtenidos en la gráfica de acuerdos por años. Debido a que el fichero no muestra realmente el número de acuerdos de cada país, esta representación sería aproximada mostrando el conjunto de acuerdos firmados por el mismo grupo de países.

Sobre el siguiente archivo ipynb se intentaría corregir esta irregularidad sobre los países con el fin de poder utilizar este elemento sobre la visualización final y poder responder a las preguntas iniciales.

“TRANSFORM_MMG.ipynb”

En este archivo encontramos las transformaciones realizadas sobre los datos del fichero original que permitieron conseguir un fichero más manejable con las características de cada uno de los acuerdos y otro fichero donde se mostraran todos los acuerdos que había firmado un país. En este fichero utilizaremos algunas de las librerías anteriores como “pandas” o “numpy” y otras librerías nuevas como “sklearn.preprocessing” o la librería “re”.

En primer lugar, volveríamos a leer los datos del fichero original y reconstruiríamos la variable “Year” ya creada sobre el archivo anterior. Además, seleccionaríamos todos los atributos categóricos y numéricos considerados relevantes para el estudio descartando todos los demás. Tal y como se describe sobre la PEC3, los atributos numéricos serán normalizados mediante “MinMaxScaler” e incluidos en el nuevo dataset. Por otra parte, al solo existir una fila con referencias a “Cross-regional” será eliminada del estudio, quedando finalmente un total de 475 filas y 24 campos donde se juntaban tanto los atributos numéricos originales como los normalizados. Este sería el fichero final que utilizaríamos en el estudio de cada uno de los acuerdos y recibirá el nombre de “datatransform.xlsx”.

El siguiente paso dado en el fichero sería la extracción de los países firmantes sobre cada uno de los acuerdos de paz. Comenzaríamos esta acción mediante la eliminación de paréntesis innecesarios sobre los datos. Una vez tenemos los datos un poco más limpios, pasaríamos a la extracción real de los países.

Para empezar, elaboraríamos un excel con una lista de los países y su ISO-CODE asociado apoyándonos en la página iban.com [5]. Este excel puede encontrarse en el GitHub adjuntado mediante el nombre de

“Country.xlsx” y será clave en la extracción de los países. Sobre Python comenzaríamos leyendo este fichero y creando una serie de listas vacías donde iremos introduciendo los elementos de cada país que encontremos. Básicamente, el código se basa en un bucle que va recorriendo el fichero “Country.xlsx” y estudia la presencia de un determinado país en un acuerdo, si el país aparece sobre un determinado acuerdo se guardarán sobre las listas vacías el nombre del país, el código y otros elementos de interés. Finalmente se crea un diccionario con las listas que se han ido llenando durante el bucle y se guarda la información en un dataframe. Tras este proceso se obtendrán un total de 616 filas y 8 columnas correspondientes al fichero excel “peace_con.xlsx” que repetirá un acuerdo tantas veces como número de países estén involucrados en él.

Con esto daríamos por concluida la parte de Python y comenzaríamos a trabajar sobre Tableau.

4.3.2. TABLEAU

A través de Tableau, como ya se ha expresado sobre párrafos anteriores, se construirá la visualización que dará forma al proyecto. En este apartado se intentará describir cada uno de los pasos realizados hasta la obtención final de la visualización.

En primer lugar, cargaríamos dos ficheros de datos excel con los que trabajaríamos a lo largo de la práctica: “datatransform.xlsx” y “peace_con.xlsx”. Estos ficheros no tendrán ninguna conexión entre sí y serán utilizados de forma independiente. Los filtros o acciones entre las hojas creadas con ficheros diferentes si estarán enlazados gracias a la presencia de campos idénticos sobre ambos ficheros como las regiones o el tipo del acuerdo.

Tras la carga de datos, crearíamos cuatro hojas de trabajo: mapa geográfico, mapa de calor, acuerdos por año y el nivel de implicación por región.

- **Mapa geográfico.** Utilizaremos exclusivamente el fichero “peace_con.xlsx” que nos permite representar el número de acuerdos reales firmados por cada país.
- **Mapa de calor.** Utilizaremos exclusivamente el fichero “peace_con.xlsx” para mostrar la evolución del número de acuerdos de un determinado país a lo largo de los años
- **Acuerdos por año.** En este caso usaremos el fichero “datatransform.xlsx” ya que nos interesa que no se repitan los acuerdos firmados independientemente del número de países involucrados.
- **Nivel de implicación por región.** Buscamos obtener medias de los atributos numéricos en función de las regiones para cada uno de los acuerdos de paz firmados por lo que usaremos el fichero “datatransform.xlsx”

Una vez disponemos de todas las hojas de trabajo involucradas en el estudio, se pasará a la elaboración de un **dashboard** explicativo capaz de introducir de manera coherente nuestras hojas de trabajo. Sobre este se empleará cierto tiempo con el fin de aumentar la interactividad, la elegancia y la claridad para mejorar la experiencia sobre el usuario.

En nuestro caso no hemos considerado oportuna la creación de varios dashboard que se agrupasen sobre una historia final. A tenor de los datos, buscábamos algo sencillo capaz transmitir hacia dónde van enfocadas las medidas de los acuerdos y la distribución de los mismos. Por ello, con un único dashboard se cumplen todas las expectativas

4.3. ESFUERZO, RETOS Y RESULTADOS OBTENIDOS

En este largo y amplio proyecto con bastantes horas de dedicación son numerosos los retos que hemos encontrado en cada una de las prácticas que de una forma u otra hemos podido solventar con unos resultados adecuados a cada situación.

Los retos impuestos en la construcción de la visualización comenzarían desde la PEC2 donde nos encontrábamos ante un dataset con 265 atributos, una cantidad muy superior a los atributos con los que solíamos trabajar que impactaba en un primer momento. A pesar de esto, el [codebook](#) que proporciona la misma página web ayudaría drásticamente en la comprensión del dataset para poder seleccionar aquellos atributos que consideraríamos como primordiales en nuestro estudio.

Por otro lado nos encontraremos con Python, un lenguaje versátil y muy conocido que nos aportaría nuevos retos como la búsqueda de visualizaciones con sentido donde poder apreciar tendencias, la normalización de los datos o la creación de un fichero donde se recopilasen los acuerdos firmados por cada país. Aunque sin lugar a dudas este último sería el más enrevesado invirtiendo en el problema buena parte del tiempo utilizado sobre la PEC4, los resultados obtenidos serían muy satisfactorios consiguiendo unos datos que permitirían dar sentido real a las visualizaciones.

Con la PEC3 comenzaríamos a usar Tableau, una herramienta nueva. A pesar de no haber trabajado nunca con ella, no tardaría mucho en entender la dinámica del programa gracias a la buena curva de aprendizaje que posee. Además, la gran cantidad de vídeos que nos ofrece Tableau así como las explicaciones descriptivas que encontramos en su [plataforma](#) ayudan a que el aprendizaje sea rápido, ameno y se disfrute con el proceso.

Tras una ligera interacción sobre la visualización enseguida puede apreciarse que cumple con las hipótesis iniciales marcadas sobre el informe de la práctica. La visualización es capaz de mostrarnos qué involucraban y hacia dónde iban enfocados los acuerdos mediante la gráfica que describe el nivel de implicación de los atributos sobre las características del acuerdo. Me gustaría resaltar que este nivel de implicación no se ha calculado en función de los países porque sería poco representativo para alguno de ellos, por esto se ha optado por una comparación entre regiones que permite obtener conclusiones más robustas. Por otro lado, la visualización nos mostrará el número de acuerdos de cada país y su distribución a lo largo del tiempo. Como puede comprobarse, tanto las hipótesis iniciales como las nuevas preguntas que aparecen sobre las PEC3 pueden ser respondidas a través de la visualización.

En conclusión, la visualización final está marcada por la sencillez y apoyada mediante una misma gama de tonalidades color café que nos aporta seriedad y fortaleza junto con una tipografía moderna e innovadora. Esta sencillez acompañada de la interactividad, aumentarán la efectividad para transmitir las ideas principales que nos planteamos al comienzo del proyecto pudiendo llegar a contagiar un sentimiento de empatía ante la situación vivida en ciertos países.

5. BIBLIOGRAFÍA.

1. Bell, Christine, Sanja Badanjak, Juline Beujouan, Robert Forster, Tim Epple, Astrid Jamar, Kevin McNicholl, Sean Molloy, Kathryn Nash, Jan Pospisil, Robert Wilson, Laura Wise (2020). PA-X Codebook, Version 3. Political Settlements Research Programme, University of Edinburgh, Edinburgh.
www.peaceagreements.org
2. <https://www.cear.es/situacion-refugiados/> consultado el 19/04/2020
3. https://elpais.com/diario/1992/05/19/internacional/706226414_850215.html consultado el 19/04/2020
4. https://es.wikipedia.org/wiki/Guerra_civil_nepalesa consultado el 19/04/2020
5. <https://www.iban.com/country-codes> consultado el 10/06/2020

Puntos de apoyo

- <https://www.youtube.com/watch?v=r9gYdD-REI0> consultado el 04/05/2020
- CAIRO, A. El arte funcional. Vídeos sobre el arte funcional. <https://www.youtube.com/watch?v=r9gYdD-REI0> consultado el 03/05/2020
- <https://help.tableau.com/current/pro/desktop/es-es/actions.htm> consultado el 08/05/2020
- <https://www.tableau.com/es-es/learn/get-started/desktop-viz-design> consultado el 08/05/2020
- https://help.tableau.com/current/pro/desktop/es-es/viewparts_marks_markproperties_color.htm consultado el 08/05/2020
- <https://www.tableau.com/es-es> consultado el 10/06/2020
- <https://github.com/> consultado el 10/06/2020

A. APÉNDICE VISUALIZACIONES

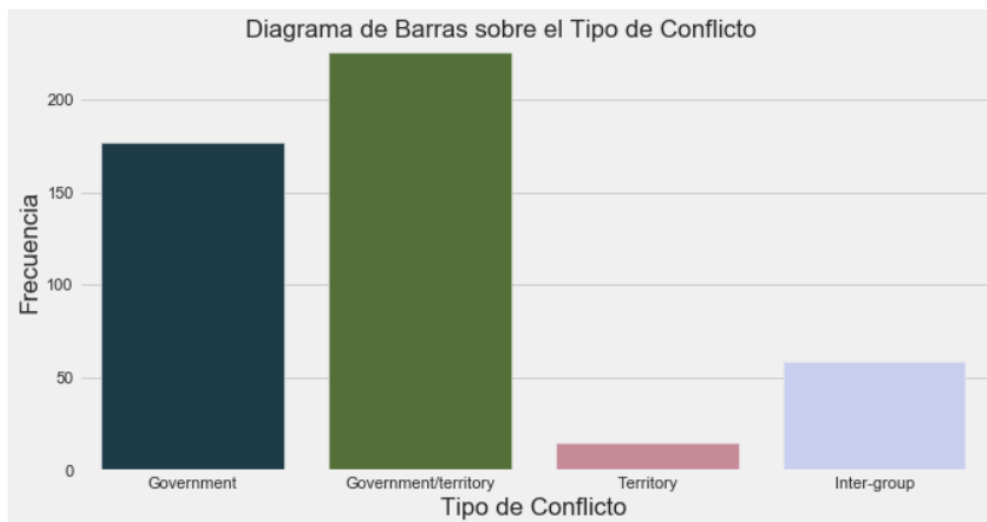


FIGURA 1: Tipo de conflicto (Contp)

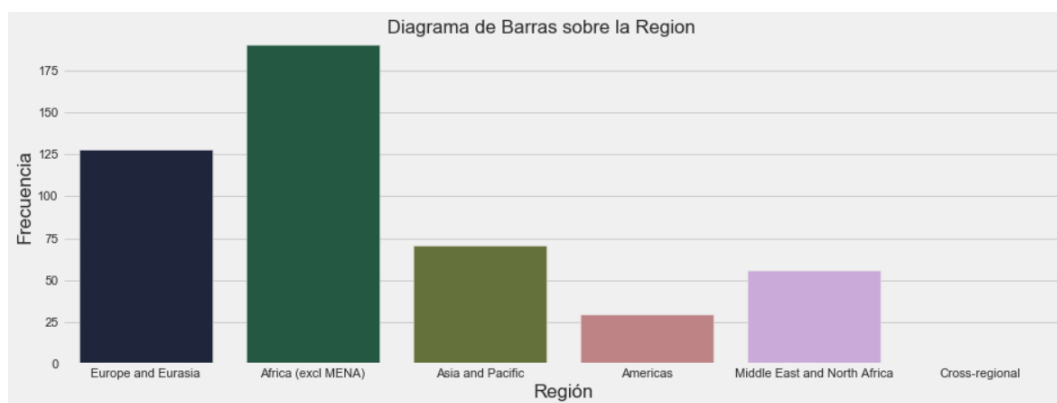


FIGURA 2: Región (Reg)

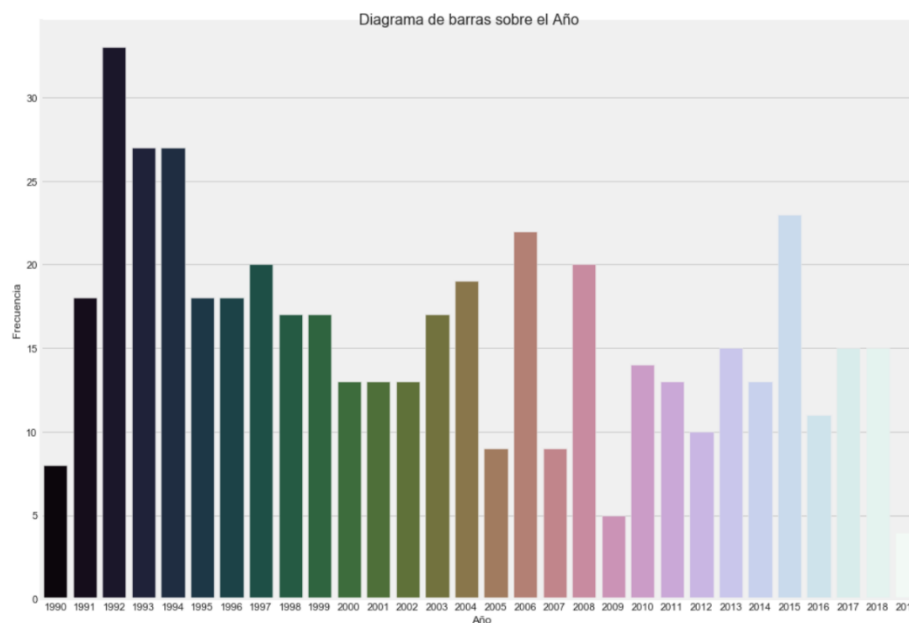


FIGURA 3: Evolución de acuerdos (Year)

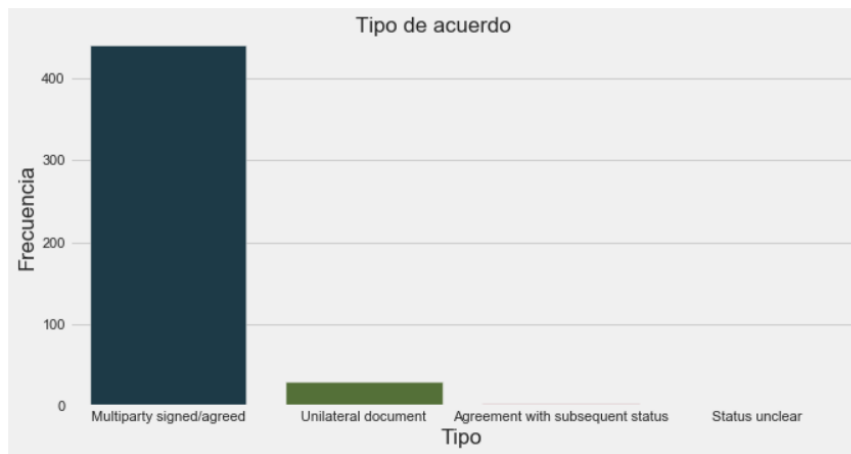


FIGURA 4: Tipo de acuerdo (Status)

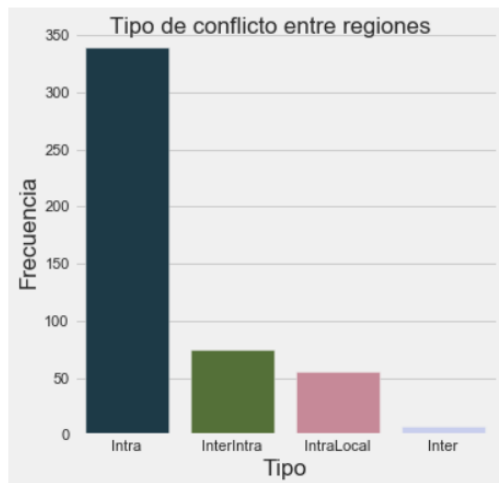


FIGURA 5: Tipo de conflicto entre regiones (Agtp)

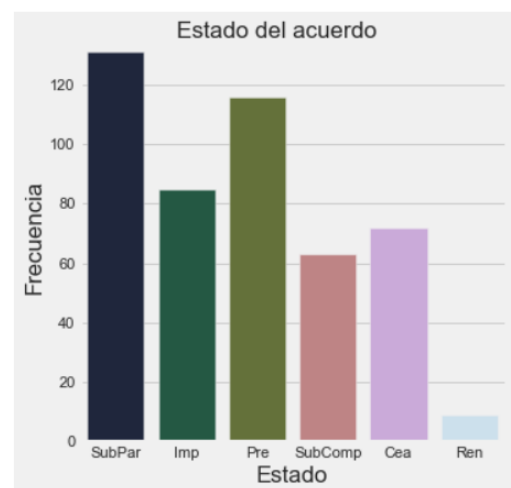


FIGURA 6: Estado del acuerdo (Status)

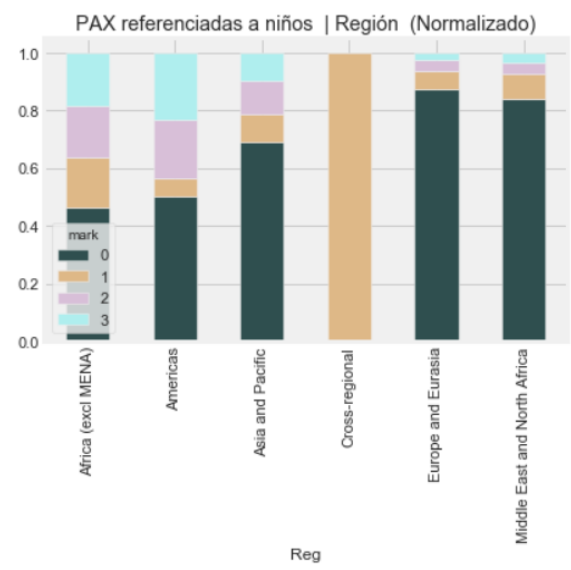
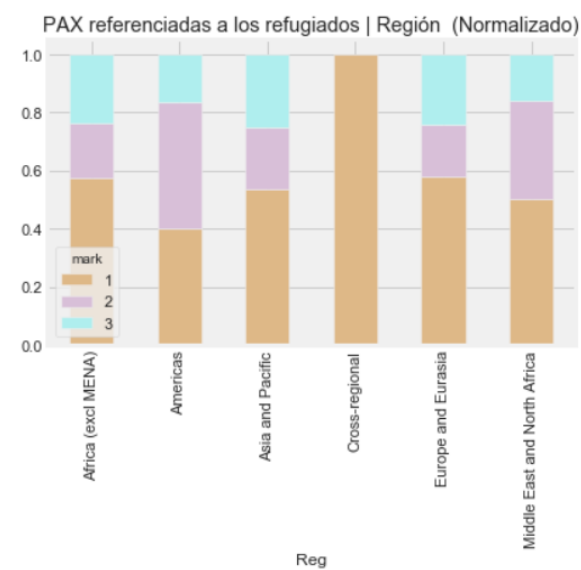
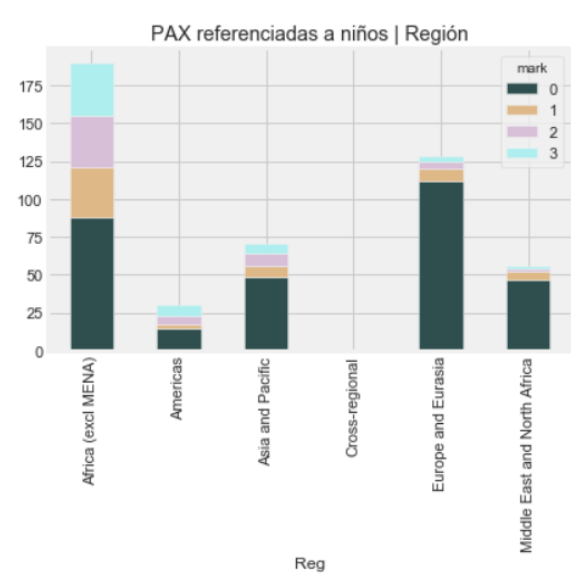
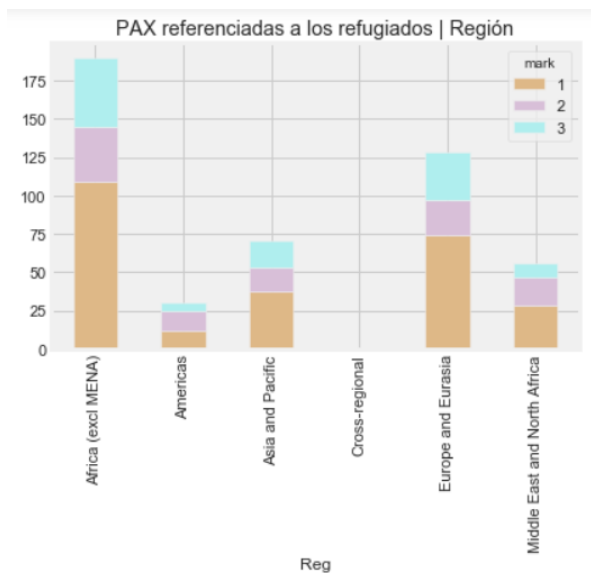


FIGURA 7: Refugiados y regiones (GRef)

FIGURA 8: Niños y regiones (GCh)

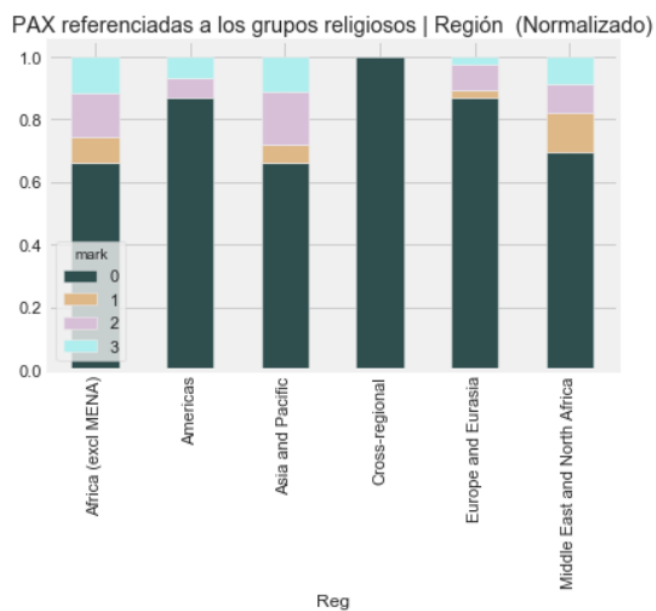
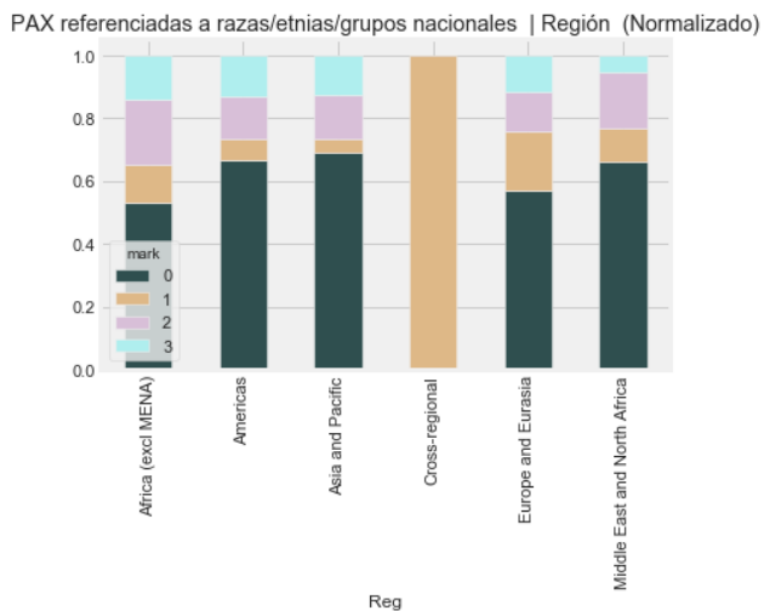
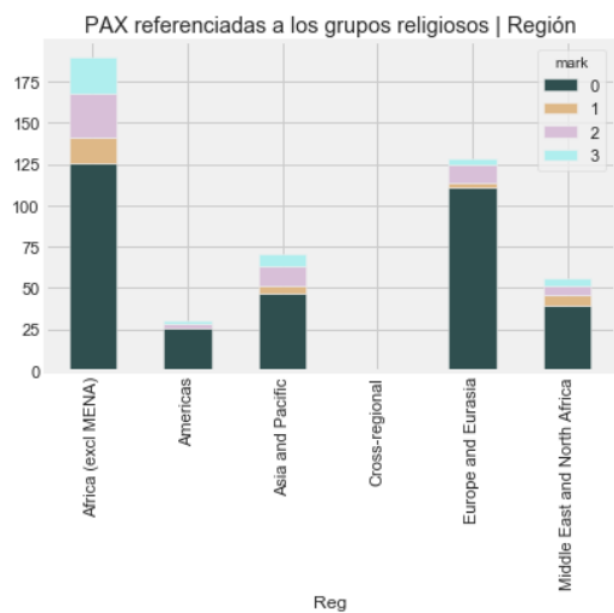
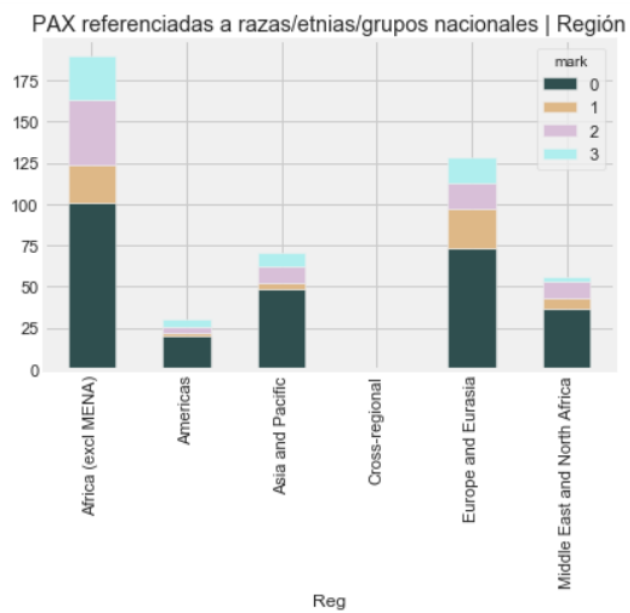


FIGURA 9: Razas/etnias/g.nacionales y regiones (GRa)

FIGURA 10: Grupos religiosos y regiones (GRe)

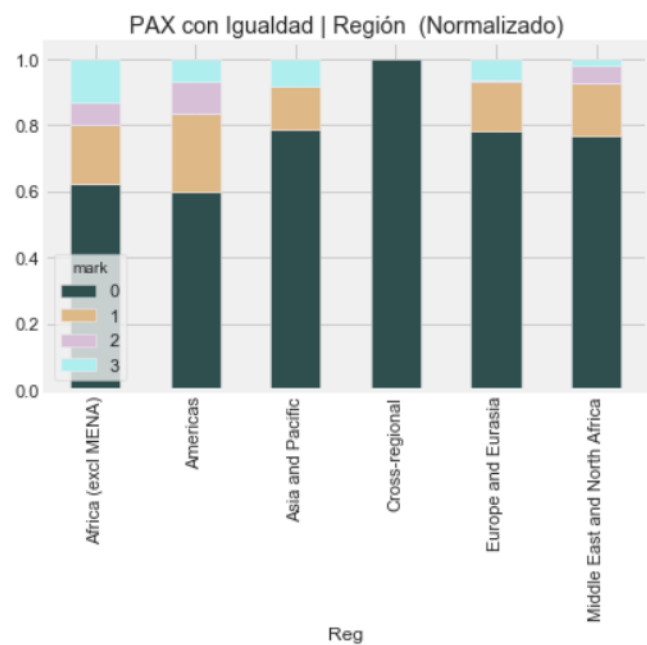
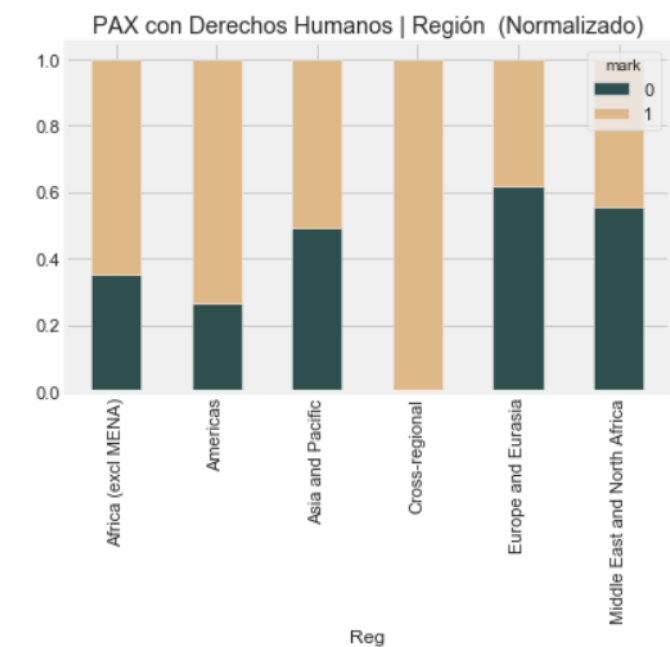
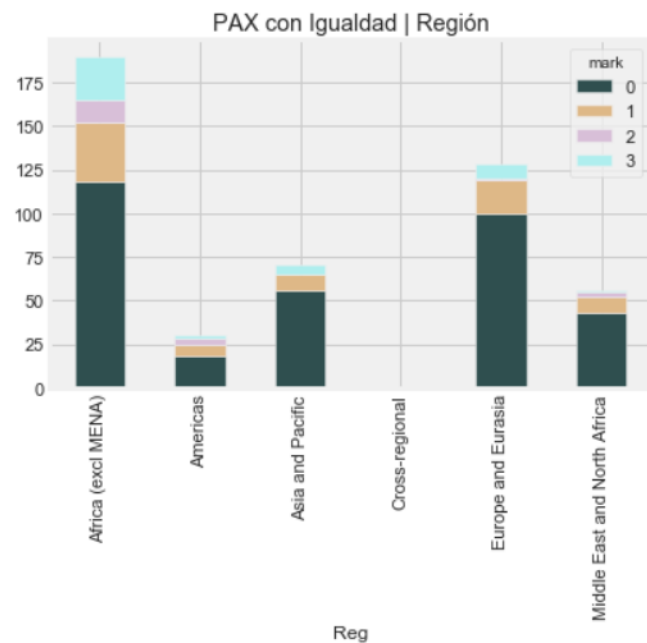
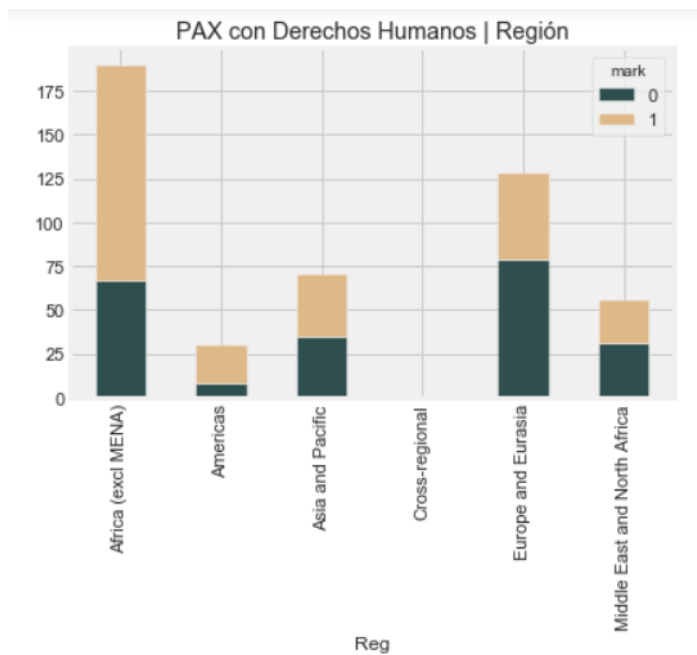


FIGURA 11: Derechos humanos y regiones (HrGen)

FIGURA 12: Igualdad y regiones (EqGen)

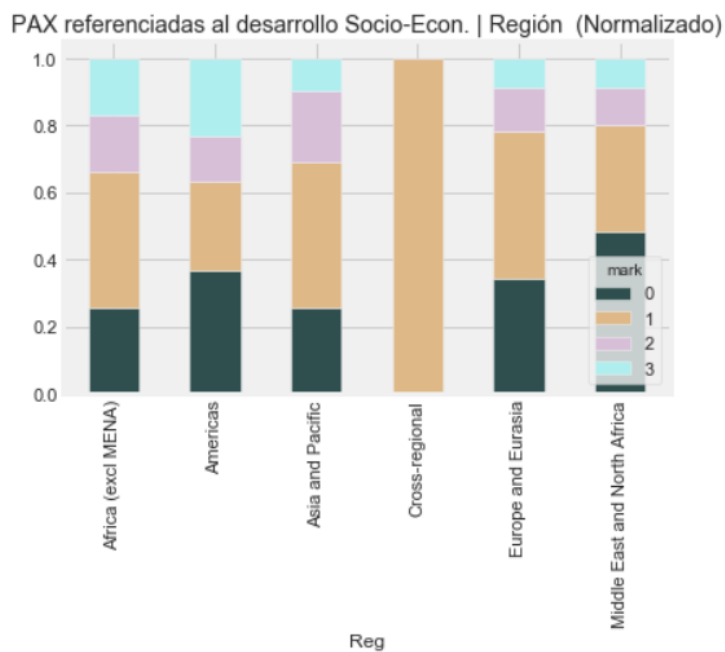
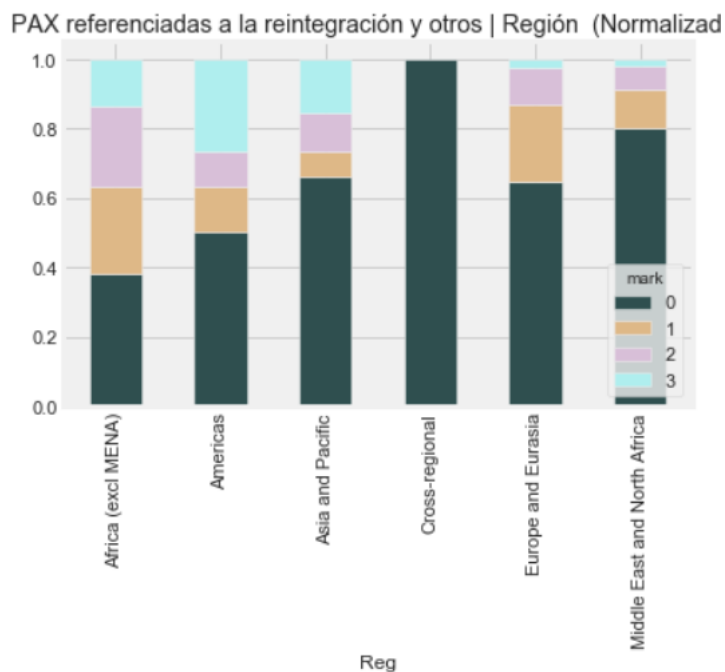
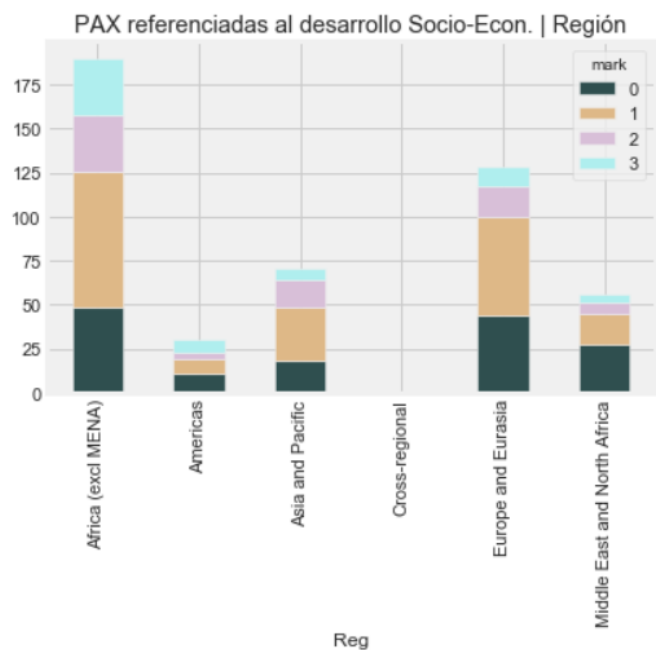
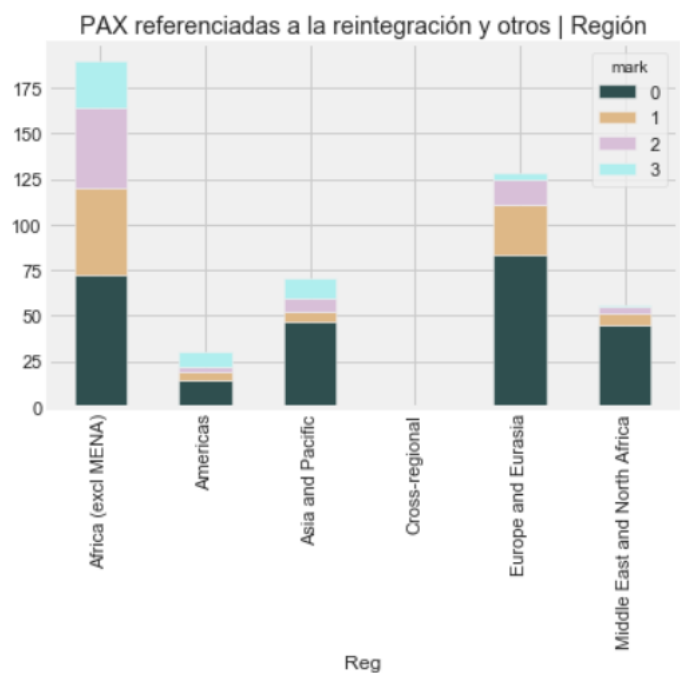


FIGURA 13: “Vuelta a la normalidad “ y regiones (SsrDdr)

FIGURA 14: Desarrollo de la reconstrucción socio-económica y regiones (Dev)

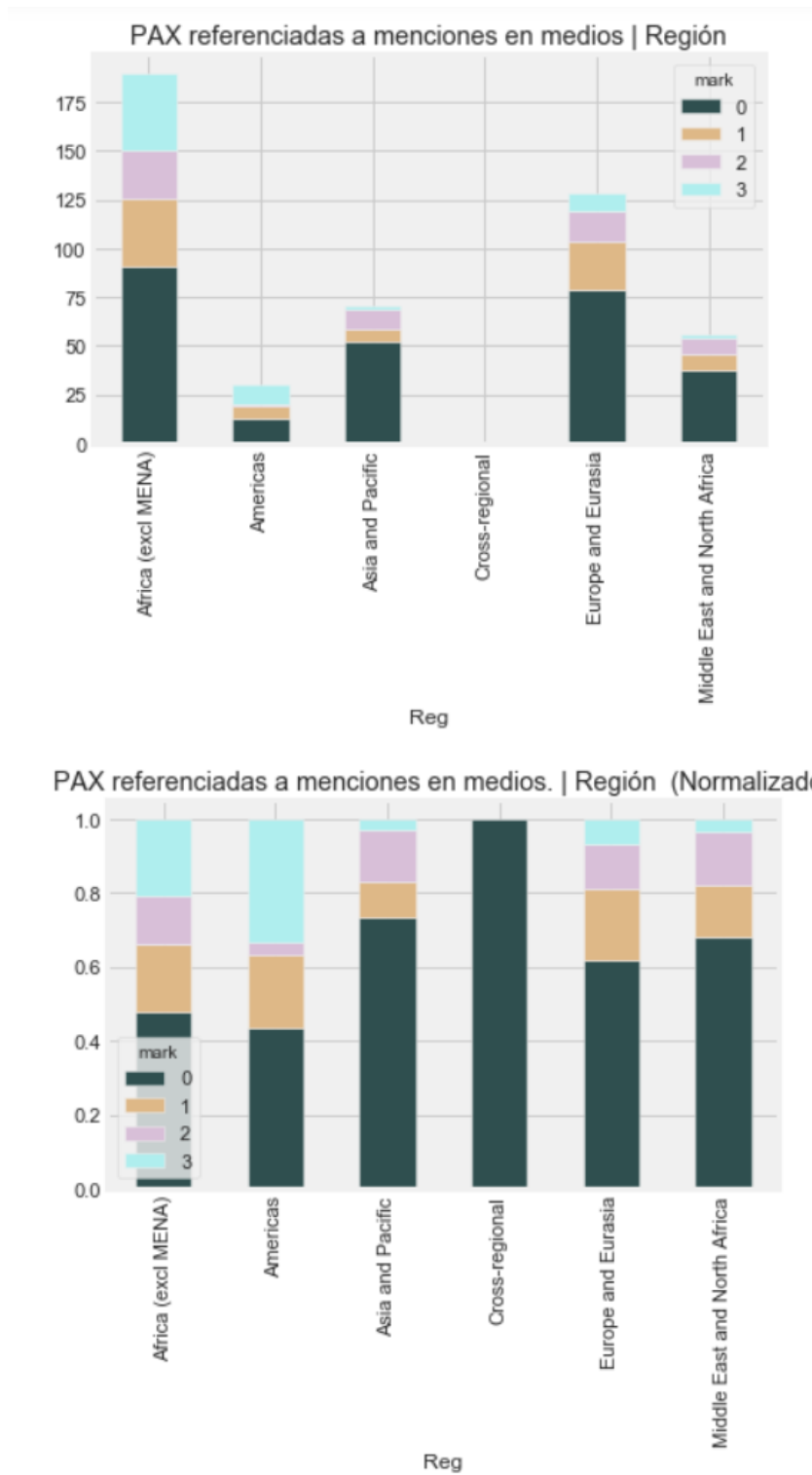


FIGURA 15: Medios de comunicación y regiones (Reg)

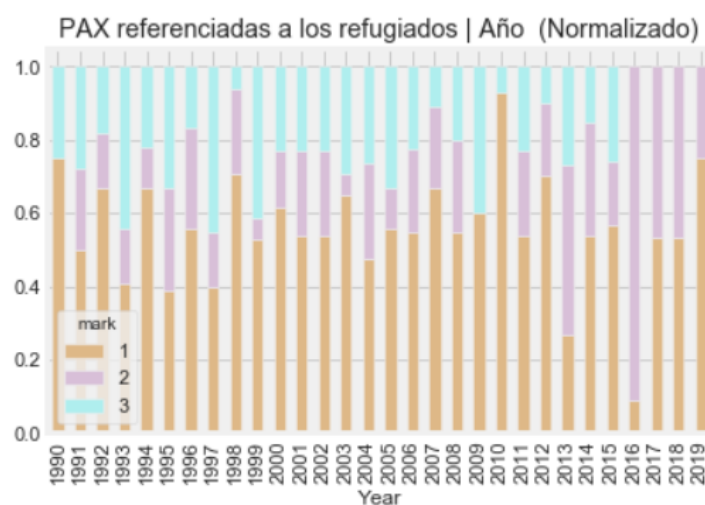
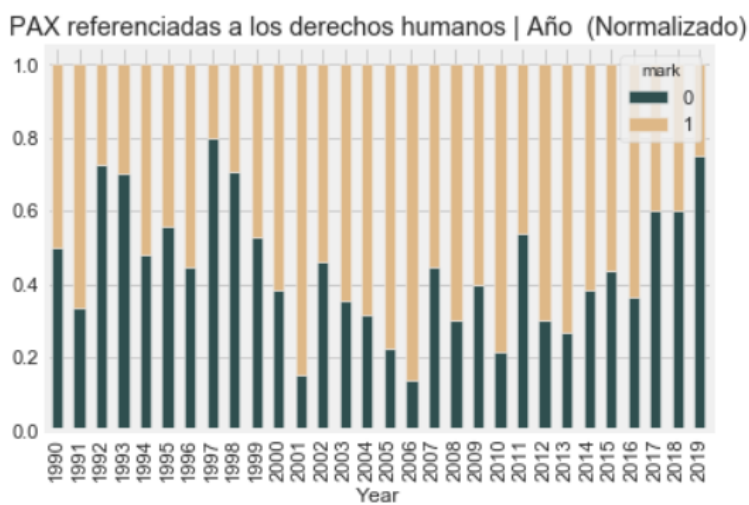
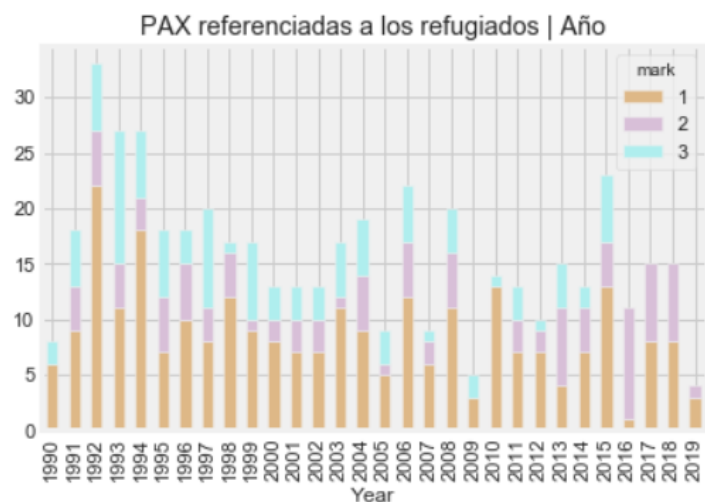


FIGURA 16: Derechos humanos y año (Hr)

FIGURA 17: Refugiados y año (GRef)

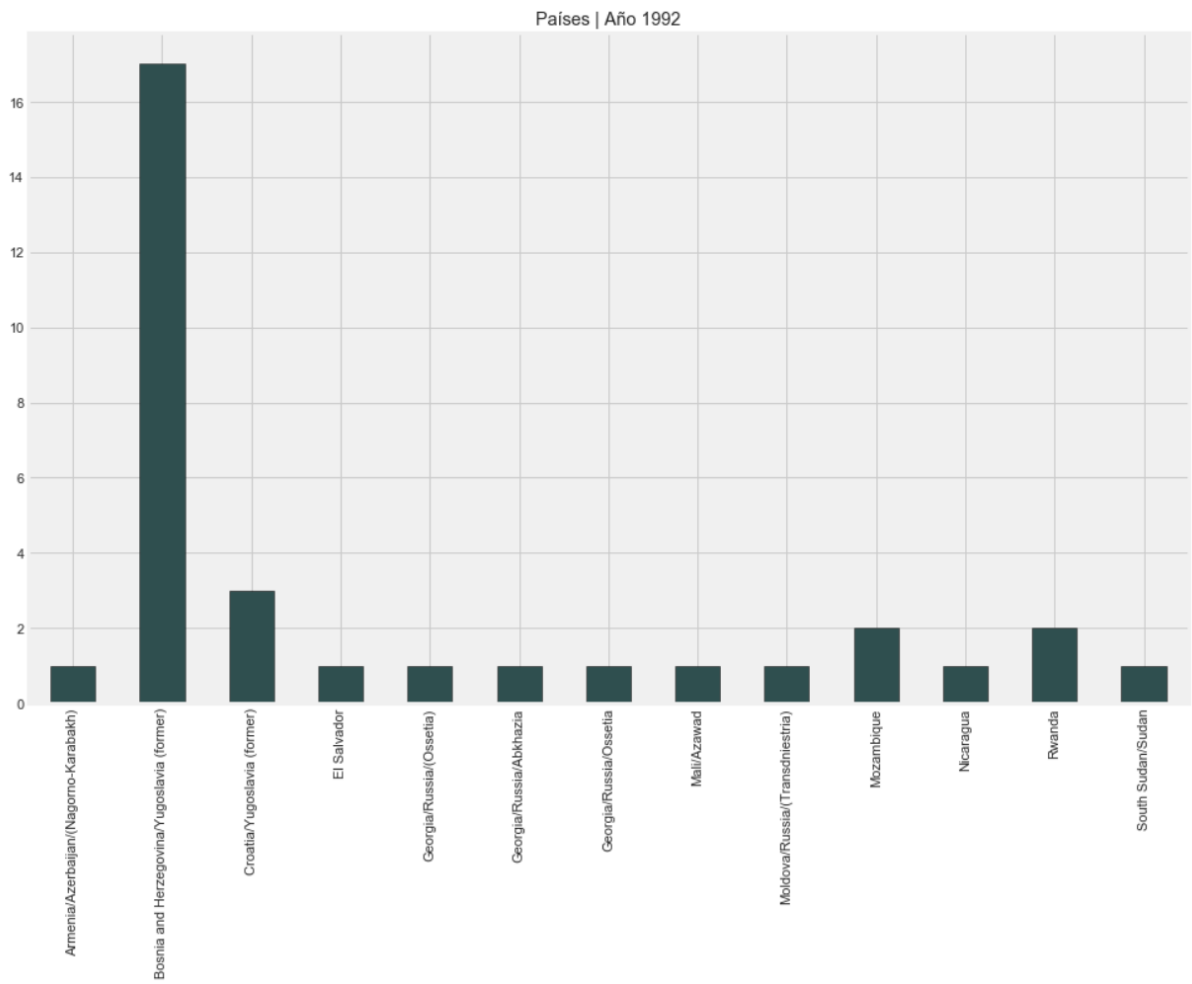


FIGURA 18: Acuerdos año 1992

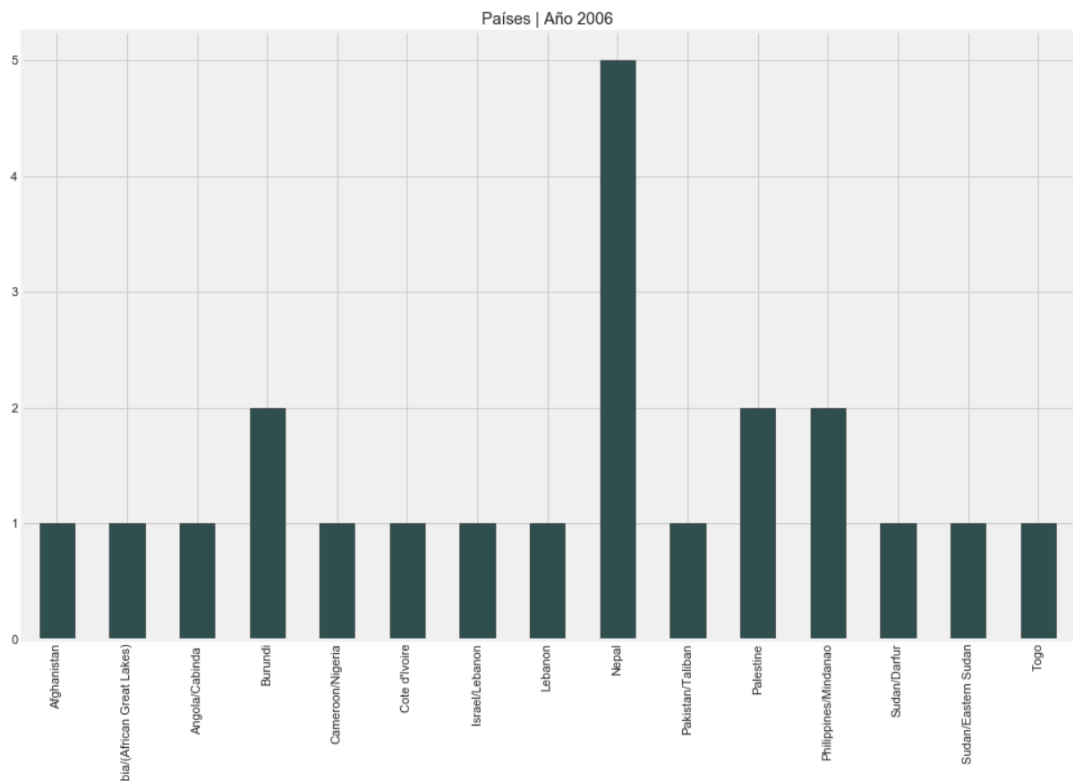


FIGURA 19: Acuerdos año 2006

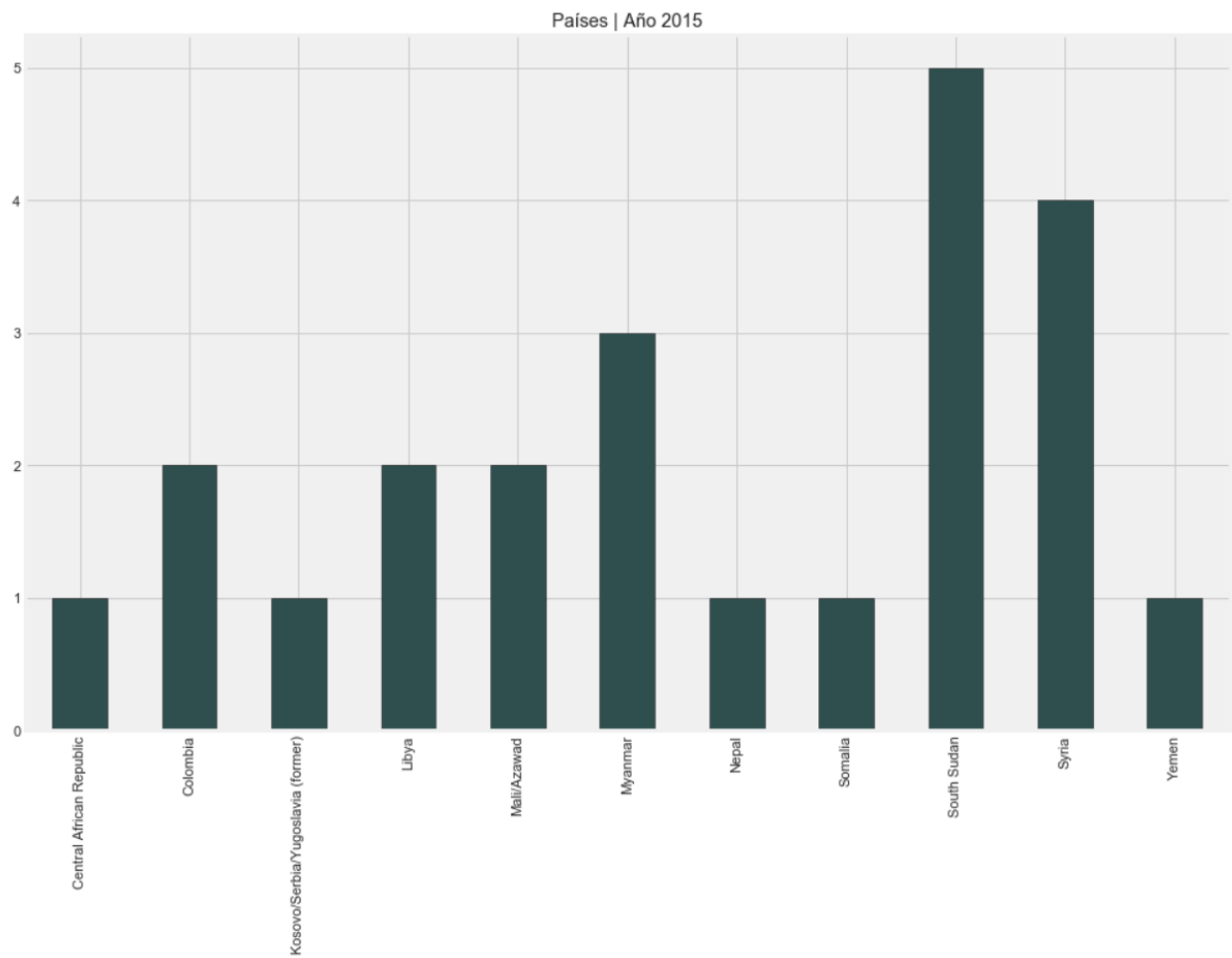


FIGURA 20: Acuerdos año 2015

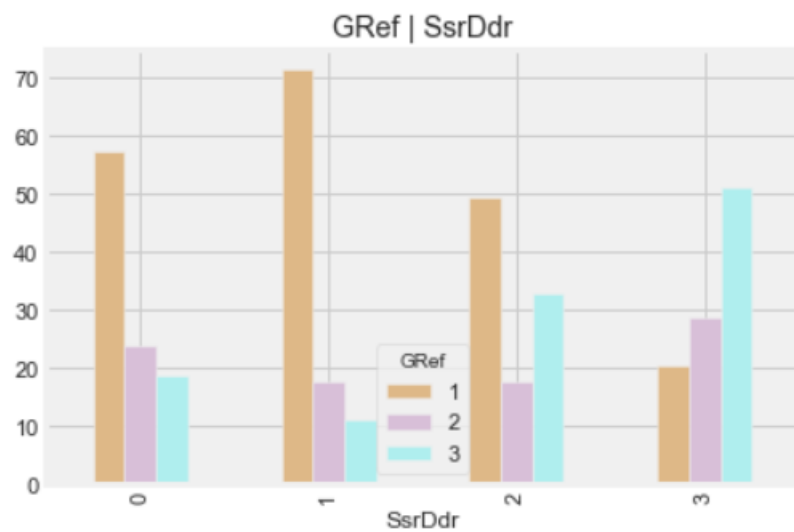


FIGURA 21: Gref frente a “vuelta a la normalidad”