

# PEC2: Proyecto de visualización de datos

14 de abril de 2020

UOC - Máster de Ciencia de Datos Aplicada

Visualización de datos. Aula 1

Alumno: Germán Yepes Calero

**Título de la visualización:**

**Acuerdos de paz para las poblaciones indígenas**

**Descripción:**

El objetivo principal del documento será analizar una selección de tratados de paz extraídos del proyecto *Peace Agreements* de la Universidad de Edimburgo ([peaceagreements.org](http://peaceagreements.org)). El proyecto recoge un gran número de acuerdos de paz de diversa índole, firmados desde 1990 hasta la actualidad. En nuestro caso, abarcaremos únicamente aquellos tratados en los que participen o se tengan en cuenta a comunidades de poblaciones indígenas.

El procedimiento a seguir será el siguiente. Primero, expondremos la base de datos seleccionada y su estructura. Continuaremos escogiendo y describiendo los atributos más relevantes para, posteriormente, proceder a su análisis a través de múltiples herramientas de visualización. Durante esta fase de análisis, trataremos de extraer conclusiones sobre la base de datos: qué tipo de acuerdos son los más comunes, diferencias entre las distintas sociedades, comportamientos interesantes, relaciones entre los tipos de acuerdos... Finalmente, resumiremos lo aprendido en la exploración de los datos exponiendo de manera conjunta todas las visualizaciones.

La totalidad de la práctica, redacción, programación y representación de visualizaciones, será llevada a cabo con lenguaje Látex en Overleaf y lenguaje de programación Python en Jupyter. A lo largo del documento se expondrán con más detalle las librerías y herramientas utilizadas.

## Presentación de los datos:

Comencemos haciendo una evaluación inicial de nuestros datos. Para ello, importaremos el dataframe con la librería pandas (pandas.pydata.org) y mostraremos las cinco primeras líneas correspondientes a cinco tratados de paz donde se tienen en cuenta las poblaciones indígenas.

```
[276]: import pandas as pd
data = pd.read_csv('Indigenous.csv')
data.head()
```

|   | Con                               | Contp                | PP  | PPName  | Reg                | AgtId | Agt   | Dat        | Status                   | Lgt | ... | TJRep | TJRSym | TJRMa | TJNR | ImUN |
|---|-----------------------------------|----------------------|-----|---|--------------------|-------|---|------------|--------------------------|-----|-----|-------|--------|-------|------|------|
| 0 | Bangladesh/Chittagong Hill Tracts | Government/territory | 11  | Bangladesh-Chittagong Hills Tracts: peace process | Asia and Pacific   | 238   | Agreement between the National Committee on Ch... | 1997-12-02 | Multiparty signed/agreed | 14  | ... | 2     | 0      | 1     | 0    | 0    |
| 1 | Central African Republic          | Inter-group          | 155 | CAR: Local Processes                              | Africa (excl MENA) | 2211  | P'Comite de Suivi de la Accord de Paix et de R... | 2018-11-29 | Multiparty signed/agreed | 2   | ... | 0     | 0      | 0     | 0    | 0    |
| 2 | Central African Republic          | Government           | 14  | CAR: coups and rebellions process                 | Africa (excl MENA) | 1442  | Pacte Republicain pour la paix, la réconciat...   | 2015-05-11 | Multiparty signed/agreed | 8   | ... | 1     | 1      | 1     | 2    | 0    |
| 3 | Colombia                          | Government           | 101 | Colombia V - Santos                               | Americas           | 1845  | Final Agreement to End the Armed Conflict and ... | 2016-11-24 | Multiparty signed/agreed | 323 | ... | 3     | 1      | 1     | 3    | 0    |
| 4 | Colombia                          | Government           | 101 | Colombia V - Santos                               | Americas           | 1796  | Agreement on security guarantees and the fight... | 2016-06-23 | Multiparty signed/agreed | 18  | ... | 0     | 0      | 0     | 2    | 0    |

5 rows × 265 columns

En la primera columna "Con", podemos ver el país que llevó a cabo el acuerdo y en algunos tratados, como el de Bangladesh con Chittagong Hill Tracts (línea 0), el nombre de la población indígena que también participó en el pacto. En otros tratados, como los de la República Central Africana (líneas 1 y 2) y los de Colombia (líneas 3 y 4), únicamente aparece el nombre del país. En estos casos, podemos ver quiénes integran la otra parte del acuerdo en la cuarta columna "PPName" y sexta columna "Agt", donde se nos indica el nombre del proceso de paz y el nombre del acuerdo. También se nos incluyen números identificadores para cada tratado, así como su fecha.

En la tabla solo vemos las primeras columnas, utilizadas para identificar los acuerdos. Veamos ahora el número total de columnas de nuestra base de datos o, lo que es lo mismo, cuantos atributos presenta de cada pacto. También queremos conocer el número total de tratados que constituyen el dataframe es decir, el número total de filas del mismo. Utilizaremos la librería numpy (numpy.org)

```
[277]: import numpy as np
print("Número de columnas:", len((data.columns)))
print("Número de filas:", np.shape(data)[0])
```

Número de columnas: 265

Número de filas: 116

Tenemos 265 metadatos o atributos y 116 pactos en nuestra base de datos. El enorme número de atributos nos impide que podamos analizar todos ellos. Por esta razón, seleccionaremos aquellos que puedan proporcionar mejores resultados en la extracción de conclusiones. La mayoría de atributos son numéricos e indican el nivel de repercusión del tratado en un ámbito concreto. El resto de atributos serán de tipo categórico y clasificarán los acuerdos según sus características generales. Los atributos categóricos escogidos en esta práctica son los siguientes:

- Con: Países/regiones/poblaciones que han llevado a cabo el acuerdo.
- Reg: Región del mundo en la que el conflicto resuelto por el acuerdo ha tenido lugar. Tenemos las siguientes regiones:
  - Africa (excl. MENA)
  - Americas
  - Asia and Pacific
  - Europe and Eurasia
  - Middle East and North Africa
  - Cross-regional
  - Other
- Dat: Fecha en la que el acuerdo fue firmado o negociado. Formato: AAAA-MM-DD
- Contp: tipo de conflicto que resuelve el acuerdo. Tenemos los siguientes tipos:
  - Government: Político o ideológico
  - Territory: Territorial
  - Government/territory: Político/territorial
  - Inter-group: Conflictos entre grupos no estatales.
  - Other: Otros tipos de acuerdo
- Agtp: Naturaleza primaria del acuerdo y el conflicto. Tenemos las siguientes naturalezas de acuerdo:
  - Inter: Internacional
  - Intra: Intranacional
  - IntraLocal: Acuerdos relacionados con un conflicto intraestatal, pero con el objetivo de resolver problemas locales.
- Status: Estado del acuerdo. Variable categórica con cuatro tipos de estado de acuerdo distintos:
  - Multiparty signed/agreed: Acuerdo firmado o claramente aceptado por más de un grupo de participantes opuestos (pero no necesariamente todos por los grupos opuestos).
  - Unilateral agreement: Acuerdo unilateral producido por una de las partes del conflicto pero en respuesta a un acuerdo con la otra parte.
  - Status unclear: Si no está claro qué parte del acuerdo se firmó o quién acordó, pero hay alguna indicación de la documentación circundante de que fue firmado o acordado.
  - Agreement with Subsequent Status: Acuerdo propuesto basado en el diálogo entre las partes y una "suposición" de lo que acordarán que, aunque no fue aceptado por las partes como un acuerdo, se convirtió en la base para desarrollos posteriores.
- Stage: Fase en la que se encuentra el acuerdo. Variable categórica con siete fases:
  - Pre: Proceso de prenegociación.
  - SubPar: Acuerdo parcial en algunos aspectos del acuerdo. Algunas partes del acuerdo siguen pendientes de discusión.

- SubComp: Similar a la etapa SubPar pero existe un intento real de resolver el conflicto.
- Imp: Implementación o renegociación con el objetivo de implementar un acuerdo anterior.
- Ren: Renovación con acuerdos cortos que buscan renovar los compromisos anteriores.
- Cea: Fase de alto el fuego o relacionadas. Acuerdos que comprometen total o parcialmente al alto el fuego o similares.
- Other: Otros tipos de acuerdos que se encuentran en estados distintos a los anteriores.

En cuanto a los atributos numéricos, nos centraremos principalmente en los que aparecen un mayor número de veces en nuestro conjunto de acuerdos. La selección aparece explicada con detalle en el apartado de Exploración de los datos. Por lo general, los atributos numéricos aparecen expresados con un número del 0 al 3. Si el acuerdo no contiene referencias al ámbito especificado o similares, el valor se establece en 0. Si existen tales referencias, la variable toma los valores de la siguiente manera:

- 1: Solo se menciona el ámbito sin especificar detalles y este no tiene relevancia real en el acuerdo.
- 2: El acuerdo contiene medidas que hacen referencia al ámbito especificado o similares, y estas brindan más detalles sobre el modo de implementación de las medidas.
- 3: El acuerdo dispone de medidas sustantivas respecto al ámbito especificado, dando detalles e indicando compromiso con su implementación.

Los atributos escogidos de este tipo son los siguientes:

- GRa: Razas/etnias/grupos nacionalistas. Señala las menciones en el acuerdo de raza, grupos étnicos, minorías nacionales, clanes u organizaciones "tribales" similares.
- GRe: Grupos religiosos. Hace referencia a cualquier mención a grupos religiosos, ya sea en términos de esos grupos o en términos de la inclusión de religiones.
- GRef: Refugiados. Hace referencia a los refugiados y a las personas desplazadas (incluida la repatriación) que aparecen en el acuerdo de paz.
- TjVic: Justicia transicional. Incluye cualquier disposición que brinde asistencia específica a las víctimas del conflicto o medidas que guarden alguna relación con aquellos que han sufrido el conflicto, incluidas víctimas desaparecidas

También tenemos un gran número de variables binarias, las cuales toman el valor de 1 si el acuerdo de paz aborda el tema especificado en el atributo. Si en cambio el tema no se aborda en absoluto en el acuerdo de paz, el valor en esta variable es 0. Únicamente escogeremos un atributo de este tipo:

- HrGen: Derechos humanos/Estado de derecho. Variable binaria, toma el valor 1 si el acuerdo de paz incluye referencias a los derechos humanos o al derecho internacional. Si no aparecen tales disposiciones en el acuerdo, el valor de esta variable es 0.

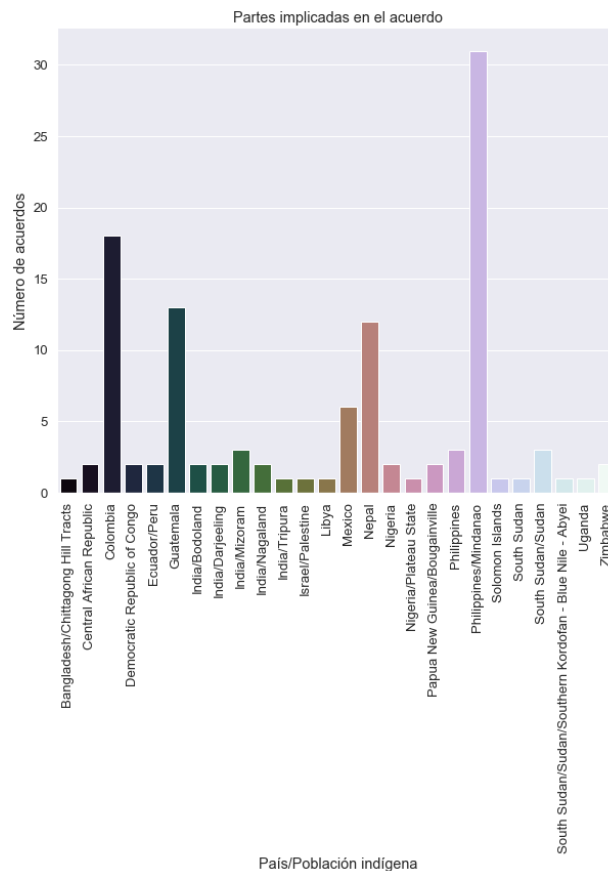
El resto de atributos vienen descritos en el "manual" de la web Peace Agreements.

Manual de atributos: <https://www.peaceagreements.org/files/PA-X>

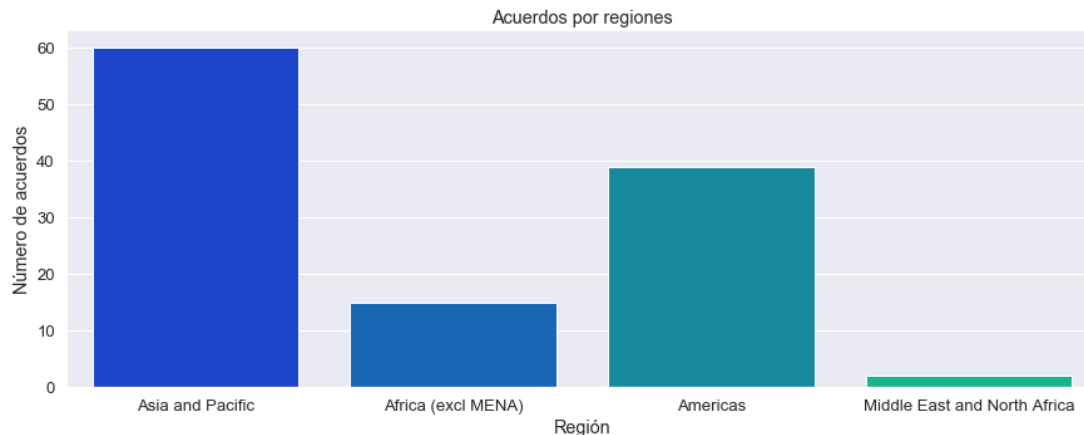
Antes de proceder al análisis queremos conocer las características generales del conjunto de pactos de paz con el que estamos trabajando. Para hacernos una mejor idea de como están distribuidos los tratados y de qué tipo son, visualizaremos en histogramas las siete variables categóricas que hemos considerado más relevantes. Las librerías utilizadas para crear las visualizaciones serán

matplotlib (matplotlib.org) y seaborn (seaborn.pydata.org). Añadimos el código con el que realizamos este tipo de gráficas sns.catplot.

```
[406]: sns.set(font_scale=1.2)
country_plot = sns.catplot(x="Con", kind="count", palette="cubehelix",
    ↪data=data, height=5)
country_plot.set_axis_labels( "País/Población indígena", "Número de acuerdos")
country_plot.fig.set_size_inches(30,8)
country_plot.set_xticklabels(rotation=90)
plt.title("Partes implicadas en el acuerdo")
```

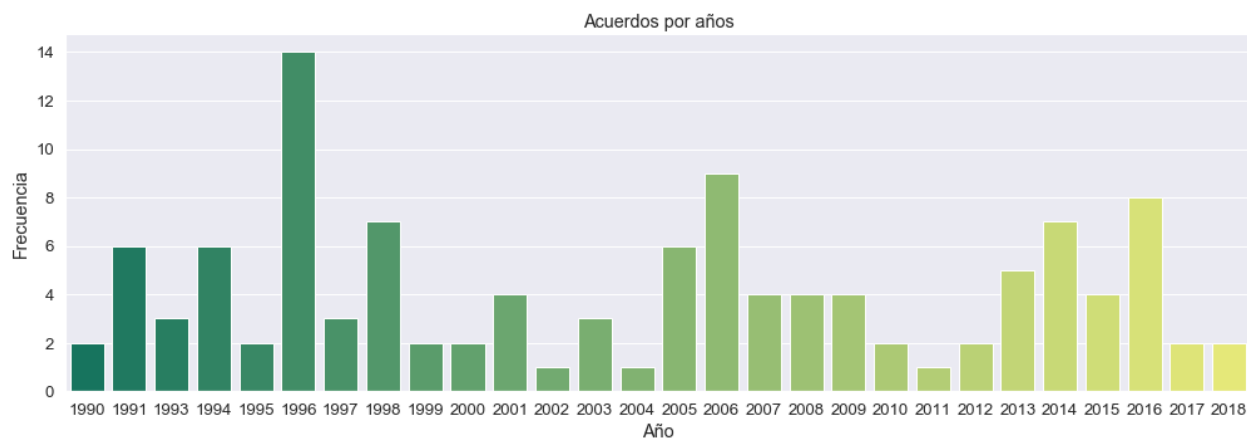


Observando el histograma vemos como la mayor parte de tratados relacionados con poblaciones indígenas incluye países como Filipinas, Colombia, Guatemala, Nepal y México. También cabe destacar a la India, que aparece en múltiples ocasiones pero con tratados que repercuten a zonas de poblaciones indígenas distintas.

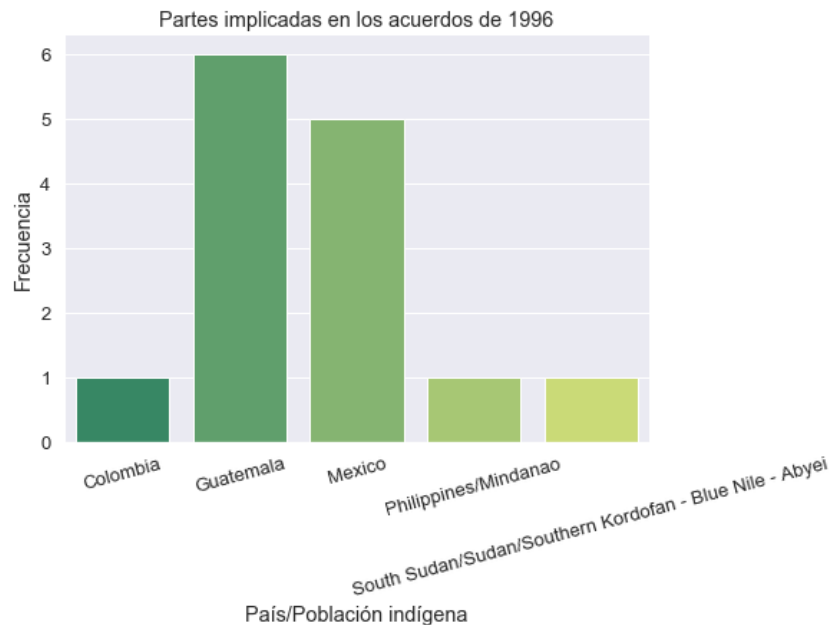


Como era de esperar viendo los resultados del histograma anterior, la mayoría de acuerdos tienen lugar en las regiones de Asia y el Pacífico (India, Nepal, Filipinas...) y en América (Colombia, México...). También tendrán relevancia en África.

Para encontrar distribución de los acuerdos con poblaciones indígenas a lo largo de los años, realizaremos la primera transformación a nuestra base de datos. Nuestro objetivo será obtener una fila que indique únicamente el valor del año en el que se produjo el acuerdo para poder representar el número de acuerdos que se produjeron cada año.



La interpretación de la gráfica nos revela que existen algunas variaciones en el número de acuerdos elaborados cada año. Destaca por encima de todos ellos 1996, con un total de 14 acuerdos, 5 por encima de la segunda mayor marca registrada, en 2006. Se trata de un valor extremo o outlier que nos hace preguntarnos si ese año tuvo lugar algún conflicto o suceso que pudo haber desencadenado tantos pactos. Representemos los acuerdos que tuvieron lugar ese año.



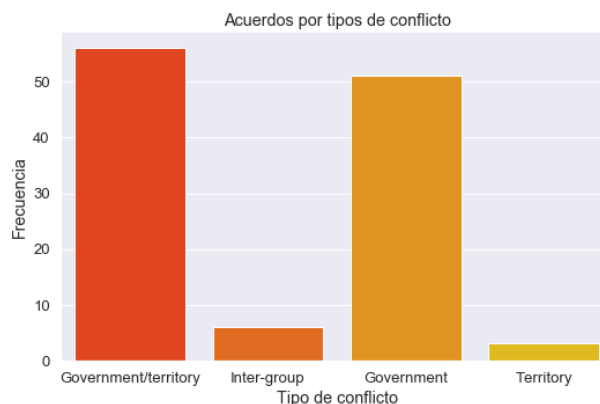
En el histograma podemos ver como los acuerdos con poblaciones indígenas se dispararon ese año en Guatemala y México. Busquemos las razones que expliquen el por qué de tan elevado número de acuerdos.

Por un lado, en Guatemala tenemos el acuerdo "Memoria del Silencio", elaborado a finales de la Guerra Civil de Guatemala (1960-1996). El principal motivo del tratado fue el genocidio guatemalteco, ocurrido entre 1981 y 1983.

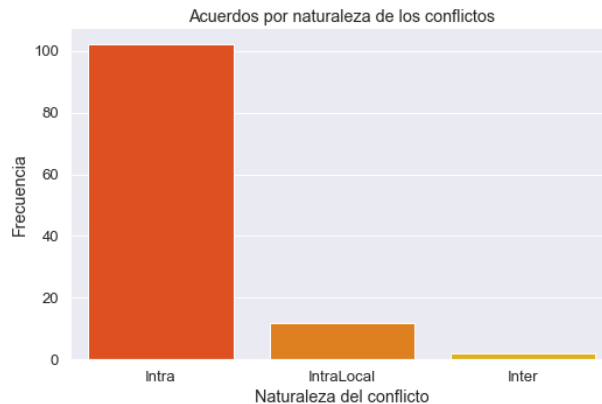
Fuente: [https://es.wikipedia.org/wiki/Genocidio\\_guatemalteco](https://es.wikipedia.org/wiki/Genocidio_guatemalteco)

Por otro lado, en México tenemos "Los Acuerdos de San Andrés sobre Derechos y Cultura Indígena", un documento que el gobierno de México firmó con el Ejército Zapatista de Liberación Nacional el 16 de febrero de 1996 para comprometerse a incluir autonomía a los pueblos indígenas de México.

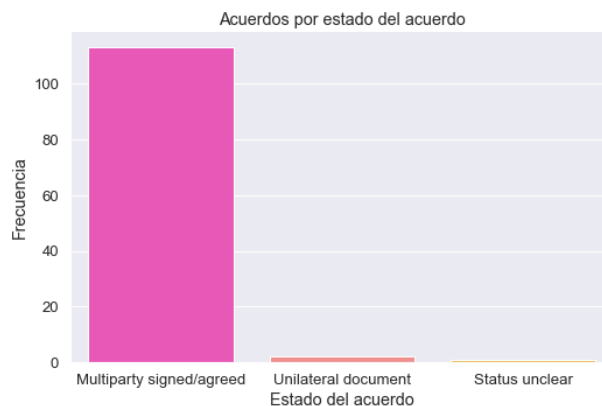
Fuente: [https://es.wikipedia.org/wiki/Acuerdos\\_de\\_San\\_Andr%C3%A9s](https://es.wikipedia.org/wiki/Acuerdos_de_San_Andr%C3%A9s)



En cuanto al tipo de conflicto, trataremos esencialmente con conflictos de tipo político. También tendremos que en parte importante de los acuerdos también habrá un factor territorial importante. Podemos deducir que los acuerdos estarán principalmente relacionados con el reconocimiento político de las poblaciones indígenas y la conservación de sus territorios.

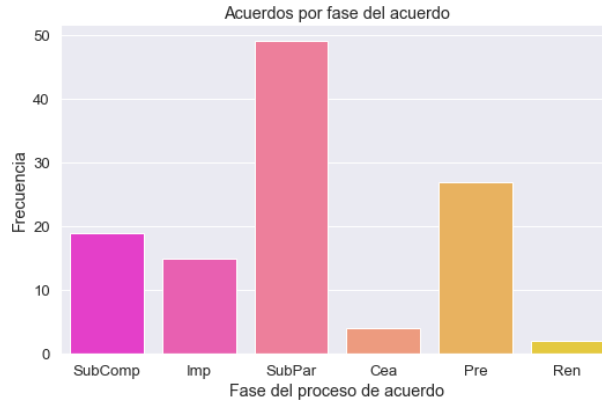


Como cabía esperar, la inmensa mayoría de conflictos tendrán lugar entre la población indígena y el país con el que intentan llegar a un acuerdo. Por lo tanto, serán conflictos de naturaleza intranacional y generalmente no tendrán repercusión fuera del país.



En esta gráfica podemos apreciar que la mayor parte de los acuerdos han sido firmados y acordados por grupos pertenecientes a ambas partes del conflicto. En el conjunto total tendremos un número muy pequeño de acuerdos unilaterales o de estado incierto.





En este caso si habrá una distribución repartida entre las distintas fases en las que se encuentra el acuerdo. Muchos de los tratados se hallarán en fase de acuerdo parcial ("SubPar"), seguidos por tratados en fases de prenegociación, acuerdo parcial comprensivo ("SubComp") y de implementación o negociación. En menor medida tendremos tratados en fases de alto el fuego y de renovación.

### Exploración de los datos:

Una vez comprendemos un poco mejor la selección de tratados con la que estamos trabajando, vamos a proceder a elegir las variables numéricas con las que realizaremos un análisis más exhaustivo.

A la hora de escoger, he decidido elegir aquellas que tuvieran un mayor peso y que, por lo tanto, apareciesen más veces y tomaran valores más altos en los acuerdos. Para ello, he elaborado el siguiente código, donde como resultado aparecen los atributos con mayor relevancia en el cómputo general de todos los acuerdos.

```
[412]: int_attribute=[]
for i in data.columns:
    if data[i].dtypes=="int64" and data[i].sum()>50:
        int_attribute.append(i)
print(int_attribute)
```

```
['PP', 'AgtId', 'Lgt', 'N_characters', 'Loc1GWNO', 'GCh', 'GDis', 'GRa', 'GRe',
'GInd', 'GIndRhet', 'GIndSubs', 'GOth', 'GRef', 'GSoc', 'GeWom', 'StDef', 'Pol',
'Cons', 'Ele', 'Civso', 'Pubad', 'Polps', 'Terps', 'Eps', 'HrGen', 'EqGen',
'HrDem', 'Prot', 'HrFra', 'Med', 'HrCit', 'Dev', 'DevSoc', 'Bus', 'Tax', 'Ce',
'SsrPol', 'SsrArm', 'SsrDdr', 'SsrPsf', 'TjAm', 'TjVic', 'TjRep', 'TjNR', 'ImE',
'ImSrc']
```

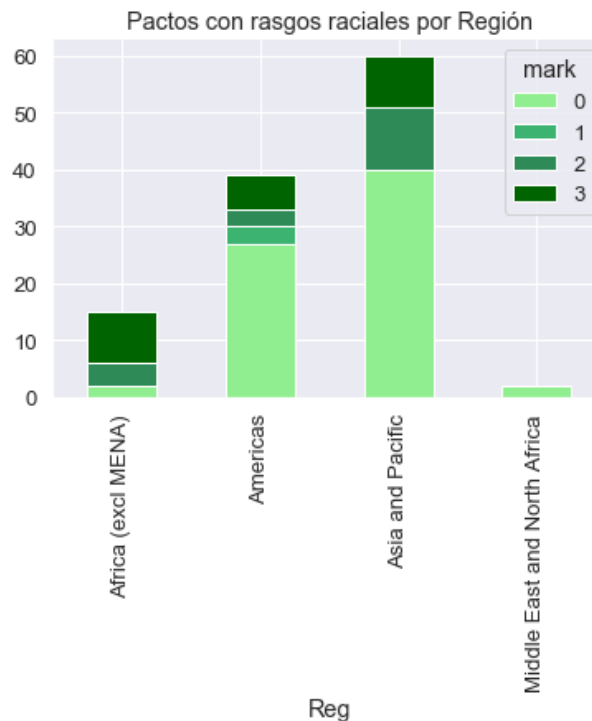
En la lista mostrada podemos observar los atributos que representan los ámbitos que generalmente aparecen en nuestros acuerdos. Los primeros 5 no pertenecen al grupo de atributos numéricos descritos anteriormente, por lo que los descartaremos. Del resto de atributos, he decidido escoger aquellos que me parecían más interesantes a la hora de extraer conclusiones. Los atributos

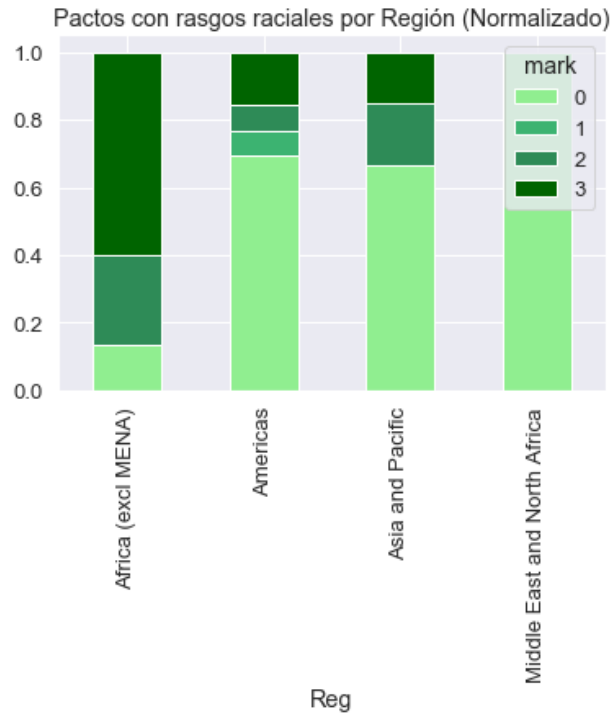
seleccionados son los mostrados en el apartado de presentación de los datos.

A continuación, procedemos a representar los atributos seleccionados con diferentes visualizaciones que nos permitan un correcto análisis y la posterior extracción de conclusiones. Añadimos el código con el que realizamos este tipo de gráficas plot.bar, la segunda de las gráficas será la versión normalizada de la primera.

```
[413]: Ra = pd.crosstab(data["Reg"], data["GRa"])
Ra.plot.bar(stacked=True, color=["lightgreen", "mediumseagreen", "seagreen", "darkgreen"])
plt.title("Pactos con rasgos raciales por Región")
plt.legend(title="mark")

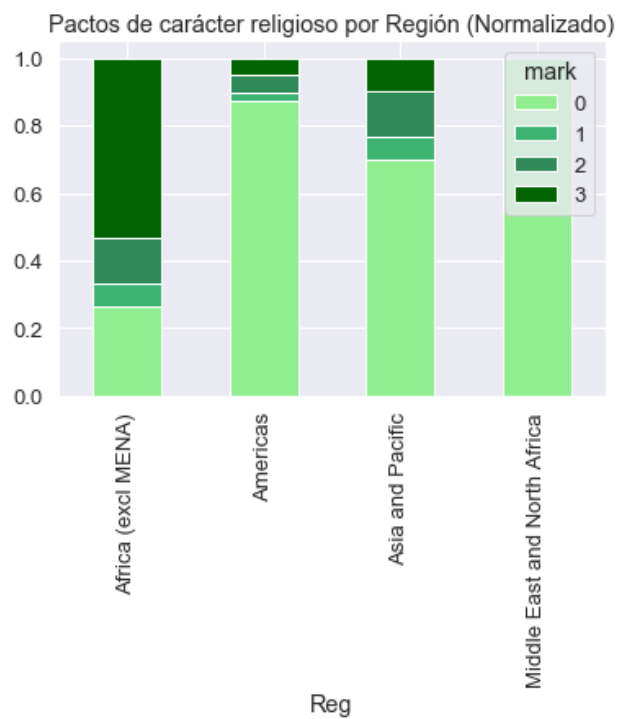
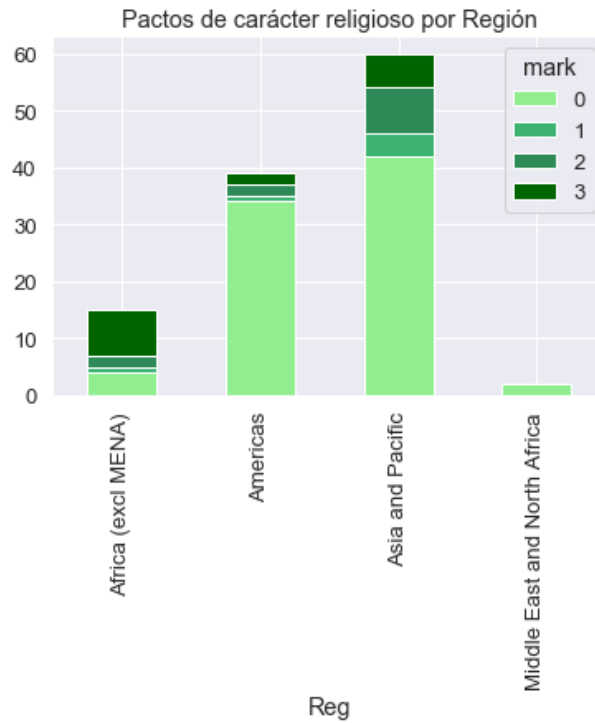
RaN = pd.crosstab(data["Reg"], data["GRa"], normalize= "index")
RaN.plot.bar(stacked=True, color=["lightgreen", "mediumseagreen", "seagreen", "darkgreen"])
plt.title("Pactos con rasgos raciales por Región (Normalizado)")
plt.legend(title="mark")
```





En los histogramas podemos apreciar una clara diferencia entre el grado de implicación racial de los tratados en África con respecto al resto de zonas. Seguramente esto se deba a la mayor densidad de tribus, clanes y grupos étnicos del continente africano en comparación al resto del mundo. Por ejemplo, solo en Ghana, hay entre setenta y cien grupos étnicos en una población estimada de unas 30.115.300 de personas.

Fuente: [https://es.wikipedia.org/wiki/Etnias\\_de\\_Ghana](https://es.wikipedia.org/wiki/Etnias_de_Ghana)



Representando el aspecto religioso del acuerdo llegamos a un comportamiento muy similar al del

diagrama de barras anterior. Comprobemos si existe algún tipo de relación entre las dos variables. Para ver de manera simple si las variables guardan relación haremos uso de la herramienta `crosstab`, cuyo código se muestra a continuación.

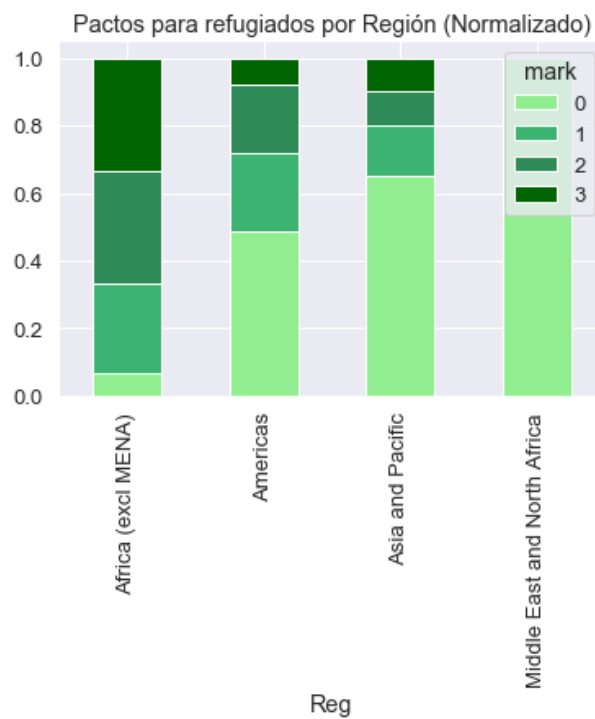
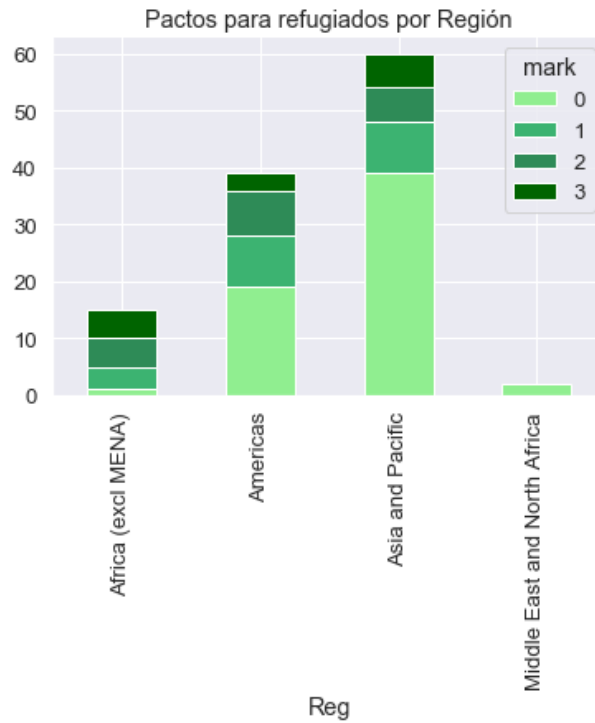
```
[402]: RaRe = pd.crosstab(index=data["GRe"],
                          columns=data["GRa"]).apply(lambda r: r/r.sum() *100,
                                                    axis=1).plot(kind="bar",
→color=["lightgreen", "mediumseagreen", "seagreen", "darkgreen"])

plt.title("Relación entre la raza y la religión en los acuerdos")
```



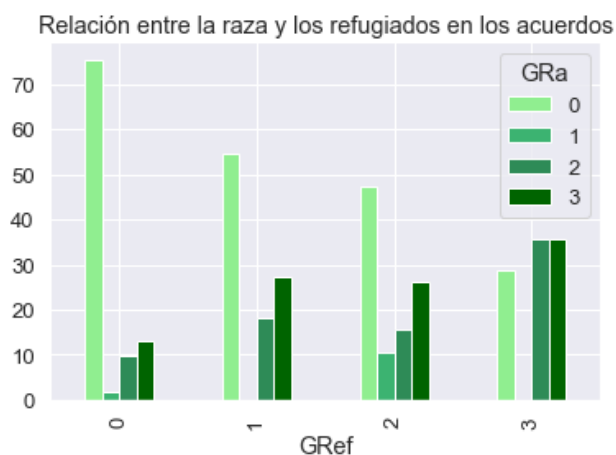
La gráfica muestra de forma evidente un alto porcentaje de casos en los que acuerdos de índole racial tienen que ver también con acuerdos con medidas religiosas. Y del mismo modo, en los acuerdos sin menciones raciales tampoco suelen aparecer implementaciones religiosas.

Al parecer, tenemos una gran correlación entre la raza y la religión de las poblaciones indígenas, siendo la primera la que explica en muchas ocasiones la existencia de la segunda y viceversa.

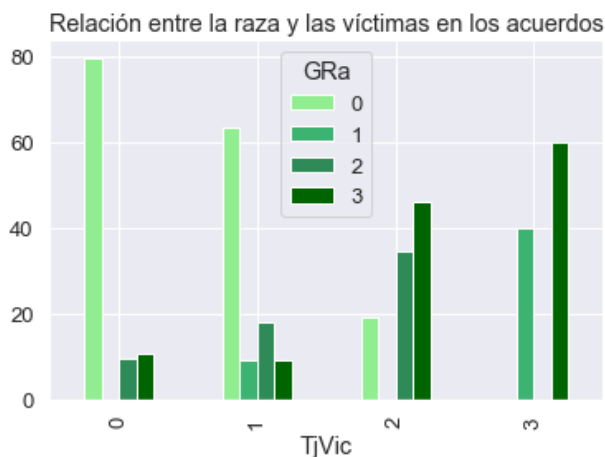


Con las menciones a refugiados volvemos a encontrar un comportamiento similar. Veamos que

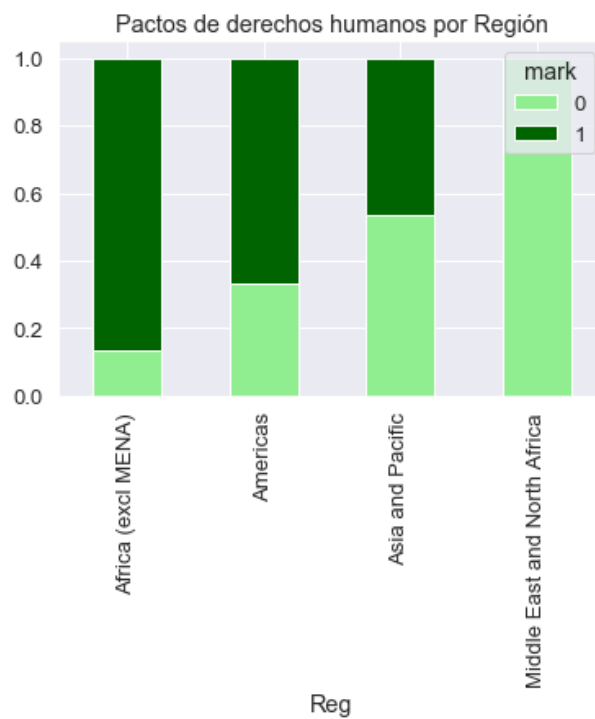
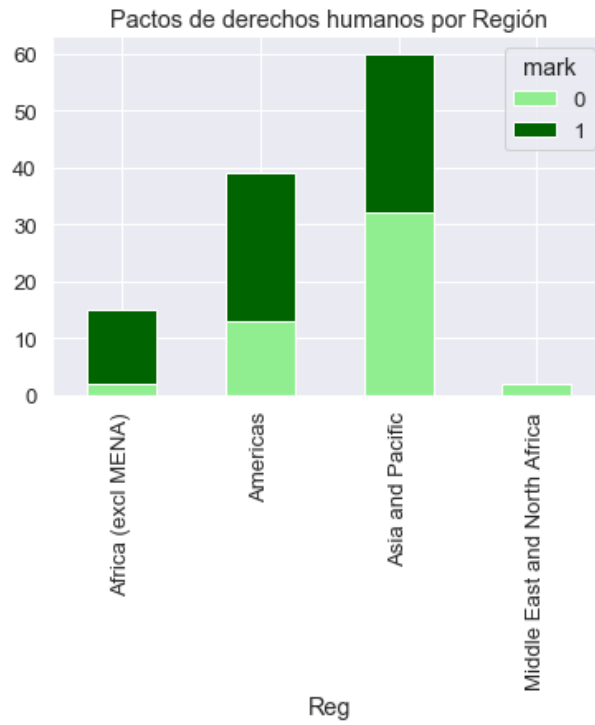
ocurre al representar tanto el grado de raza como el grado de refugiados en una misma gráfica.



En la representación gráfica anterior se intuye una tendencia en los pactos. Cuanto más se tiene en cuenta a los refugiados en el pacto, mayor es la presencia de referencias raciales en el mismo. La explicación obvia de este suceso es que los refugiados son generalmente de razas diferentes a la raza del país de origen o, que su raza no es bien recibida en un territorio determinado.



Siguiendo la línea del análisis, volvemos a centrarnos en el aspecto racial de los pactos. En este caso hemos representado el grado de referencia a víctimas del conflicto y si existen medidas de carácter racial en el mismo. En la gráfica se aprecia un claro aumento en la aparición de acuerdos raciales conforme se incrementa el nivel de importancia que las víctimas del conflicto tienen en el pacto. Esta gráfica, junto con las anteriores, parece mostrar *a priori* la existencia de tendencias racistas que dan lugar a conflictos de odio entre los países y sus poblaciones indígenas.



Por último, tenemos los derechos humanos, los cuales aparecen en un elevado número de acuer-



dos con poblaciones indígenas originales de África. En América, Asia y el Pacífico también tenemos un alto porcentaje de pactos que contienen medidas relacionadas con los derechos humanos.

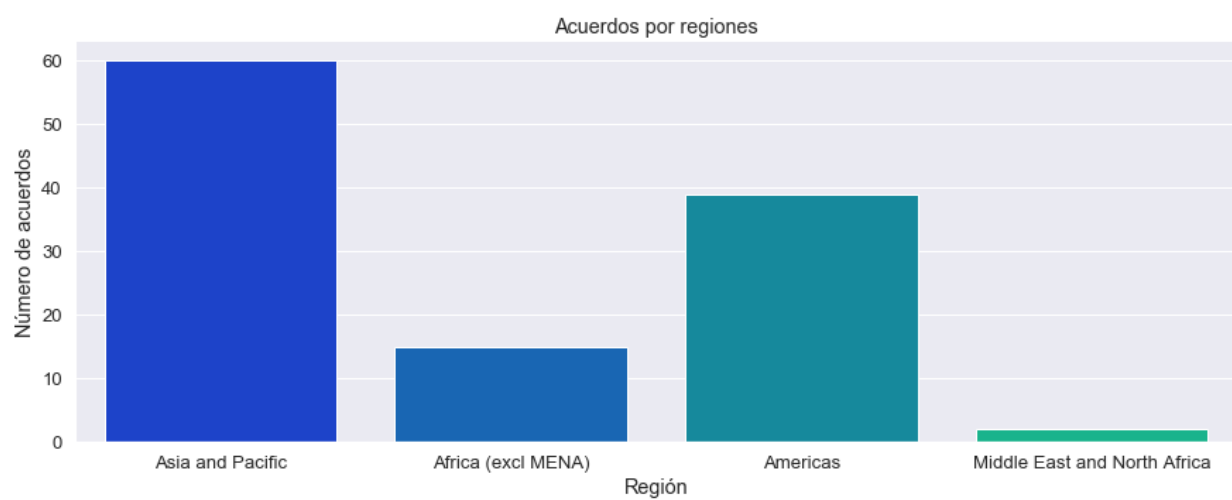
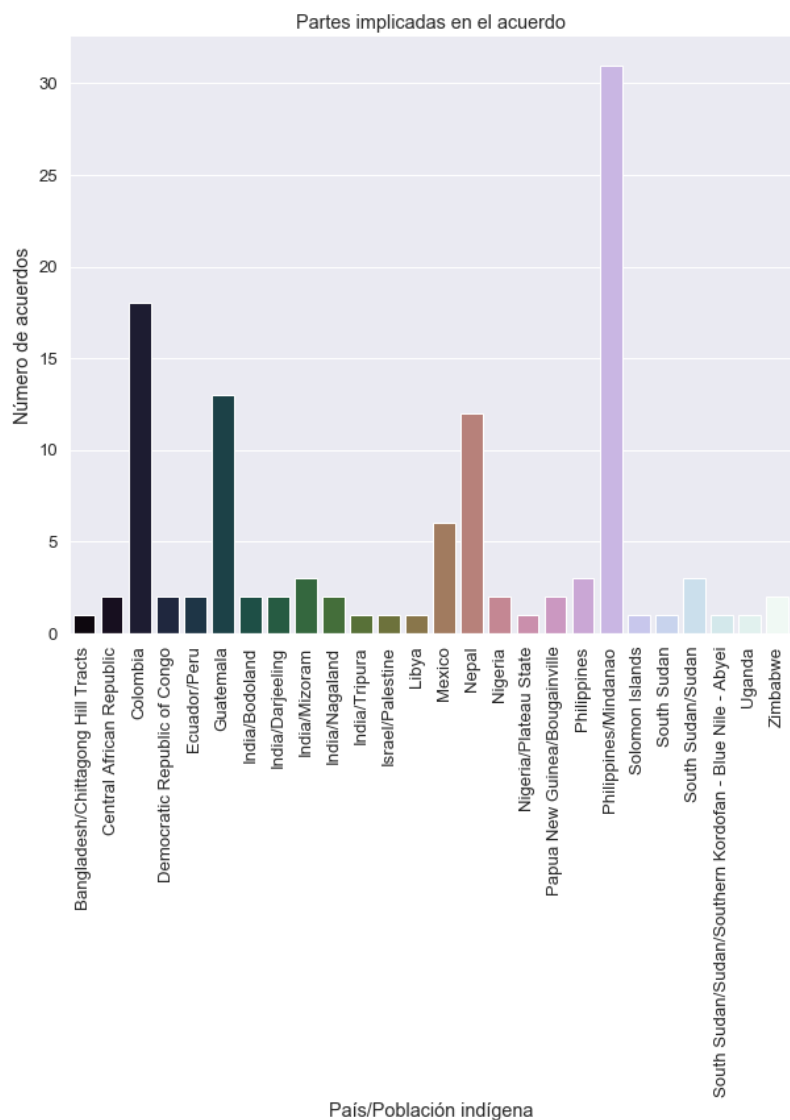
Un mayor número de representaciones han sido realizadas a la hora de elegir cuáles integrarían el documento final. Estas nos han ayudado a analizar mejor las tendencias y valores de nuestra base de datos. No han sido incluidas para evitar una extensión demasiado larga del informe.

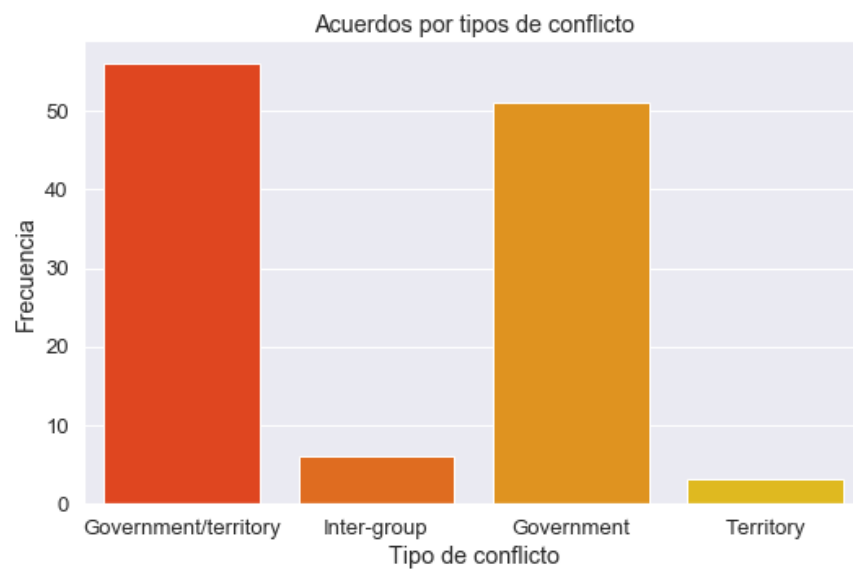
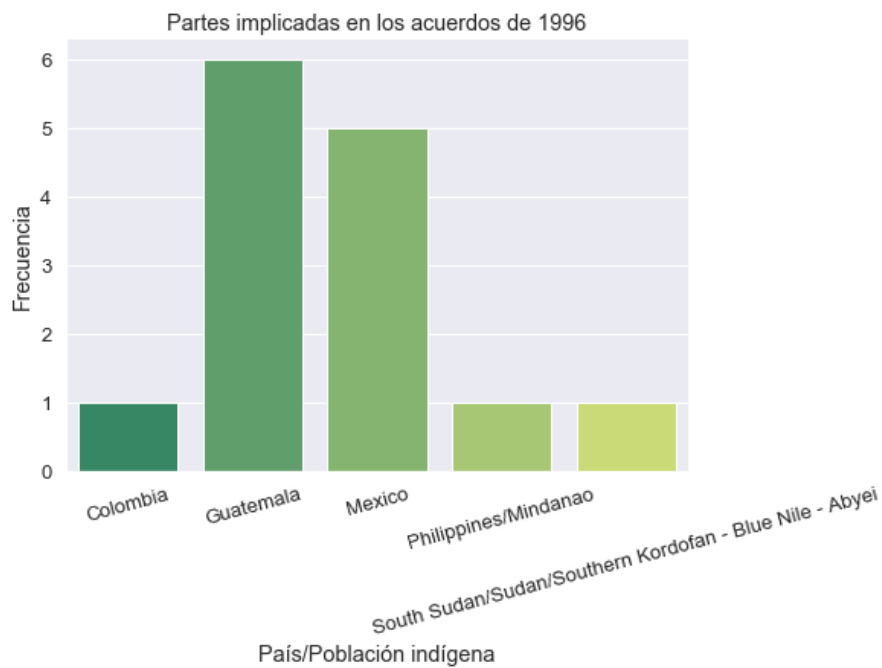
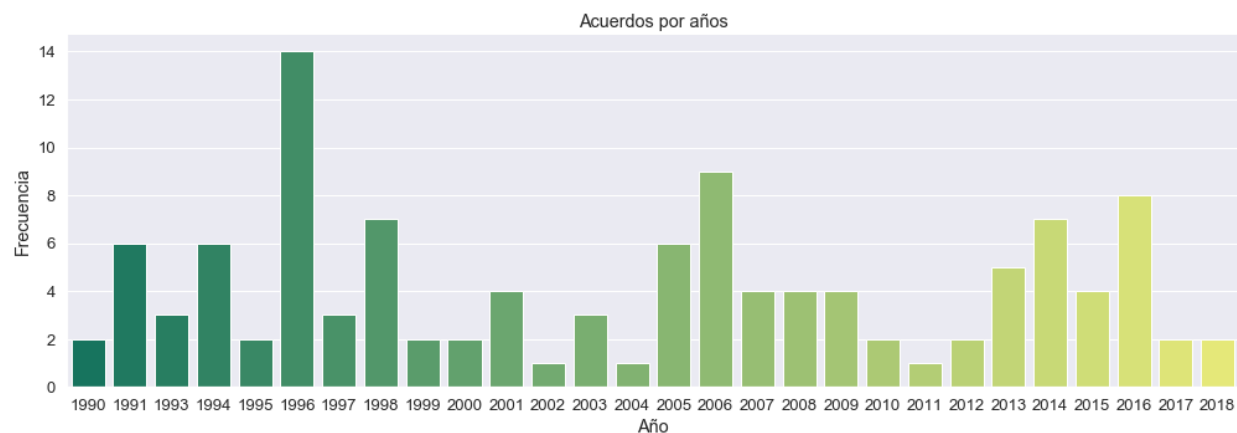
### **Procedimiento y herramientas**

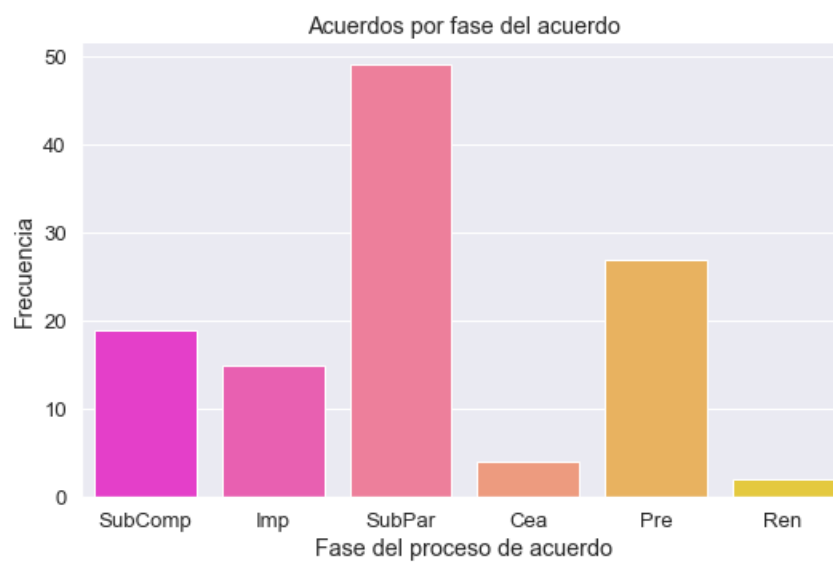
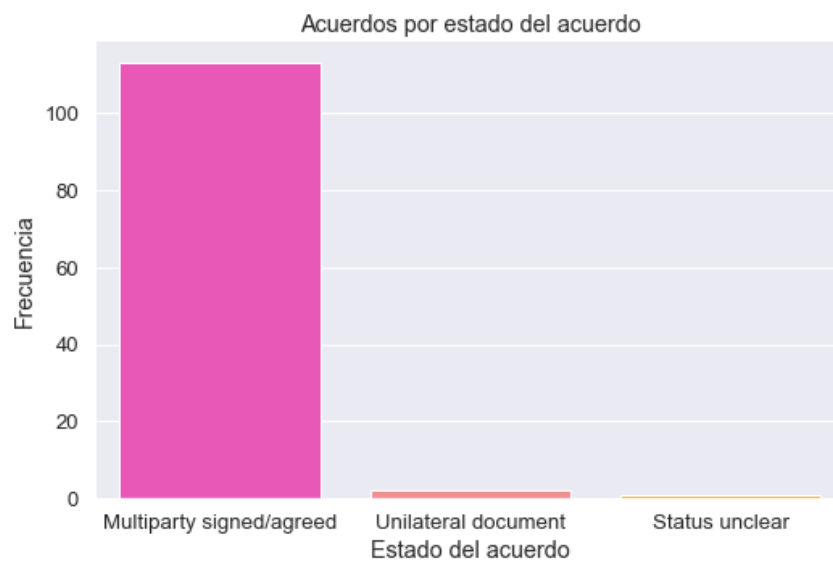
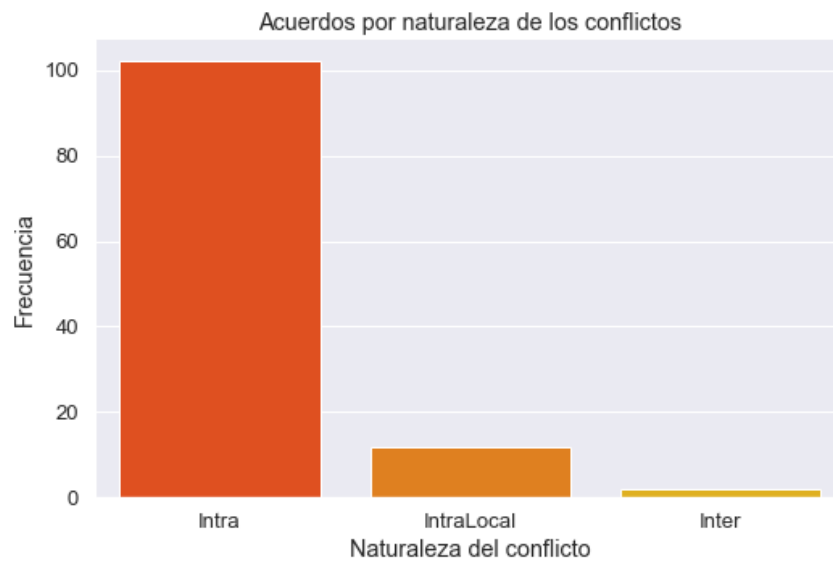
El procedimiento seguido y las herramientas utilizadas han sido expuestos y explicados a lo largo del desarrollo del informe. El documento escrito ha sido elaborado en Overleaf, un editor de documentos  $\text{\LaTeX}$  online. La parte de programación y representación de visualizaciones se ha realizado en Jupyter, con lenguaje Python. Las librerías empleadas han sido matplotlib y seaborn. Otra posibilidad habría sido realizar las visualizaciones en R, programa que hemos usado en otras asignaturas del máster.

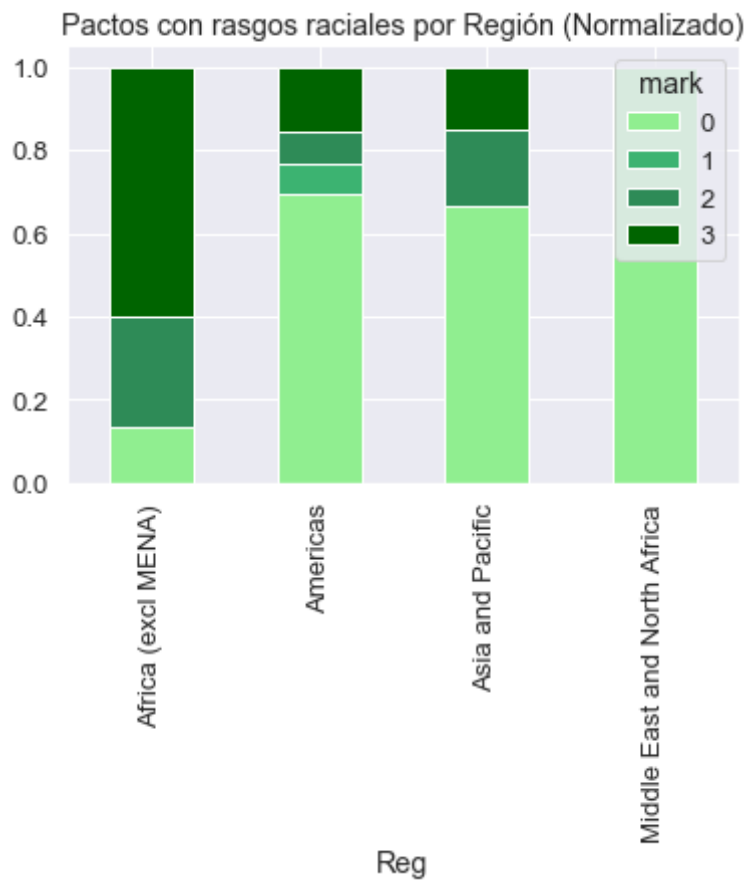
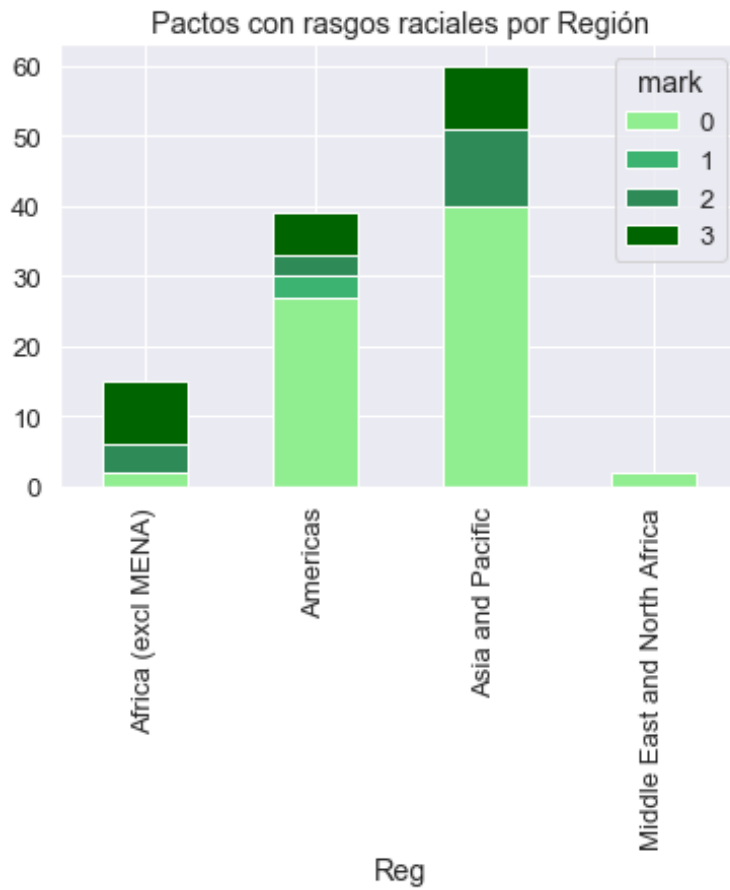
### **Dashboard:**

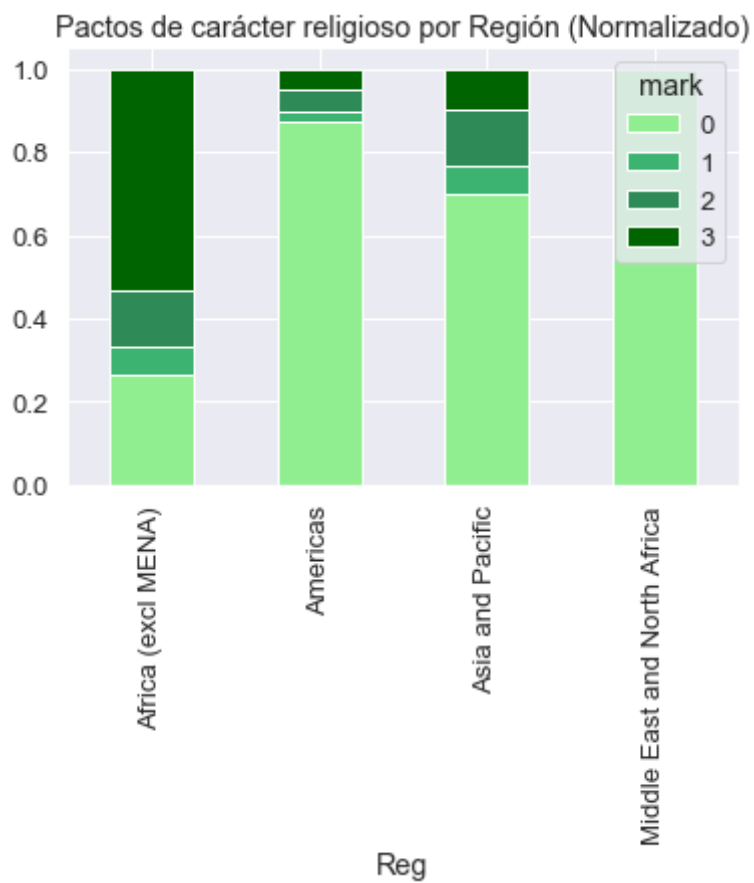
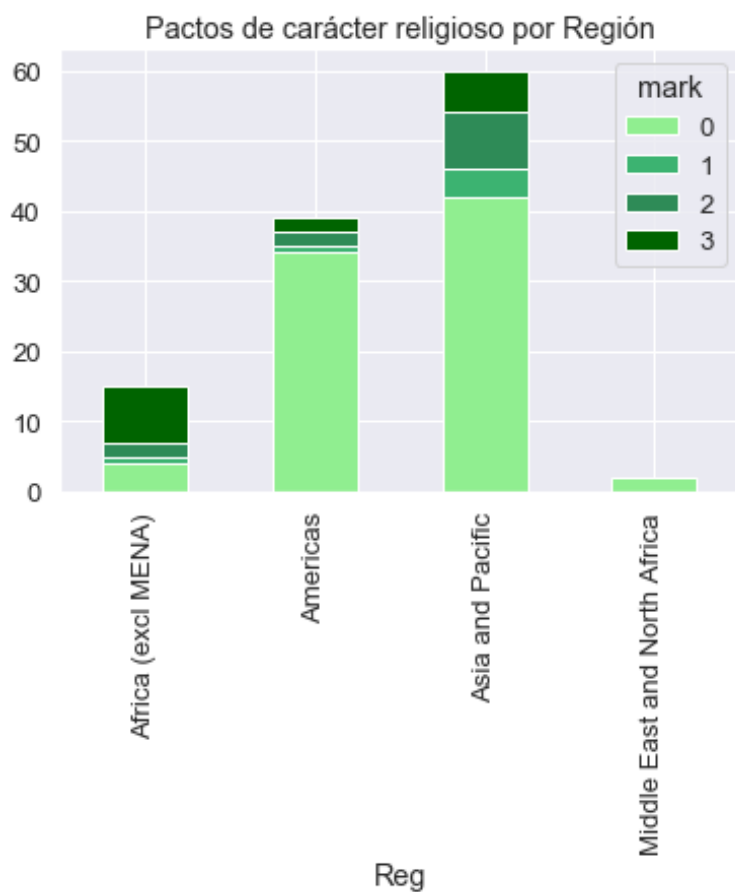
En las páginas siguientes mostramos una exposición conjunta de las visualizaciones mostradas a lo largo del documento.



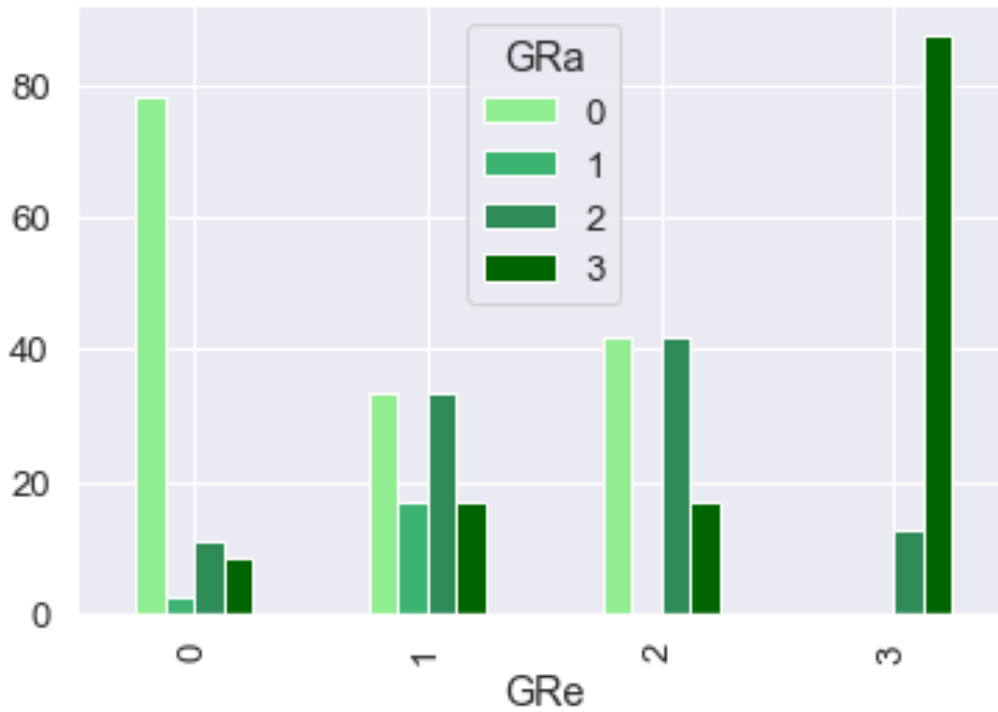




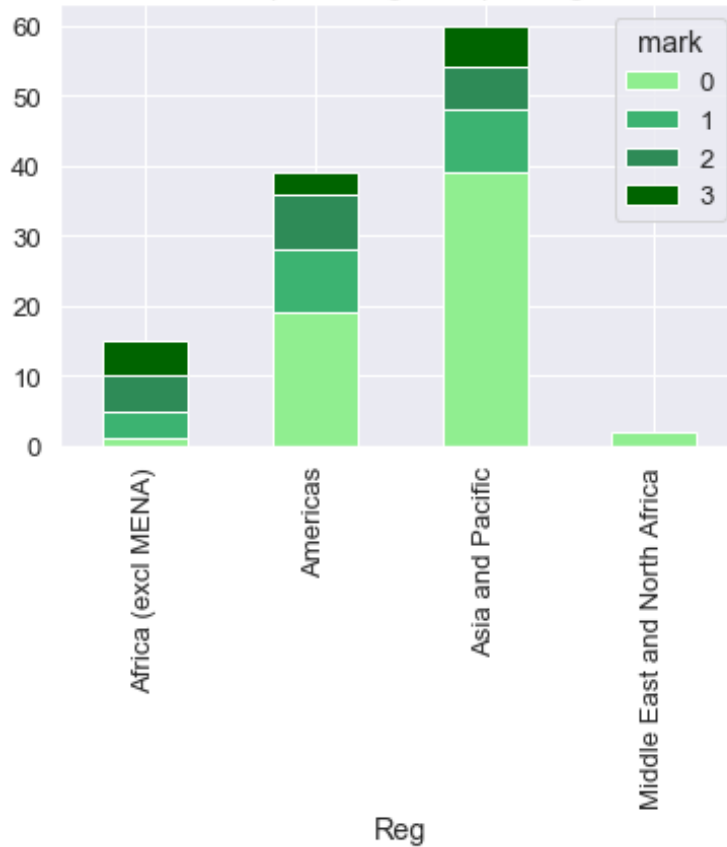




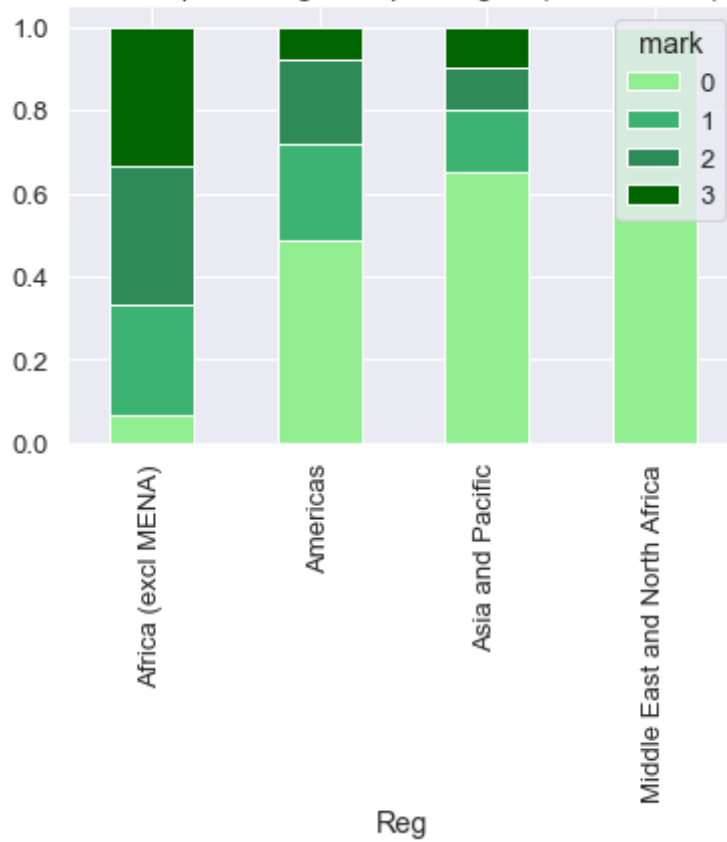
Relación entre la raza y la religión en los acuerdos



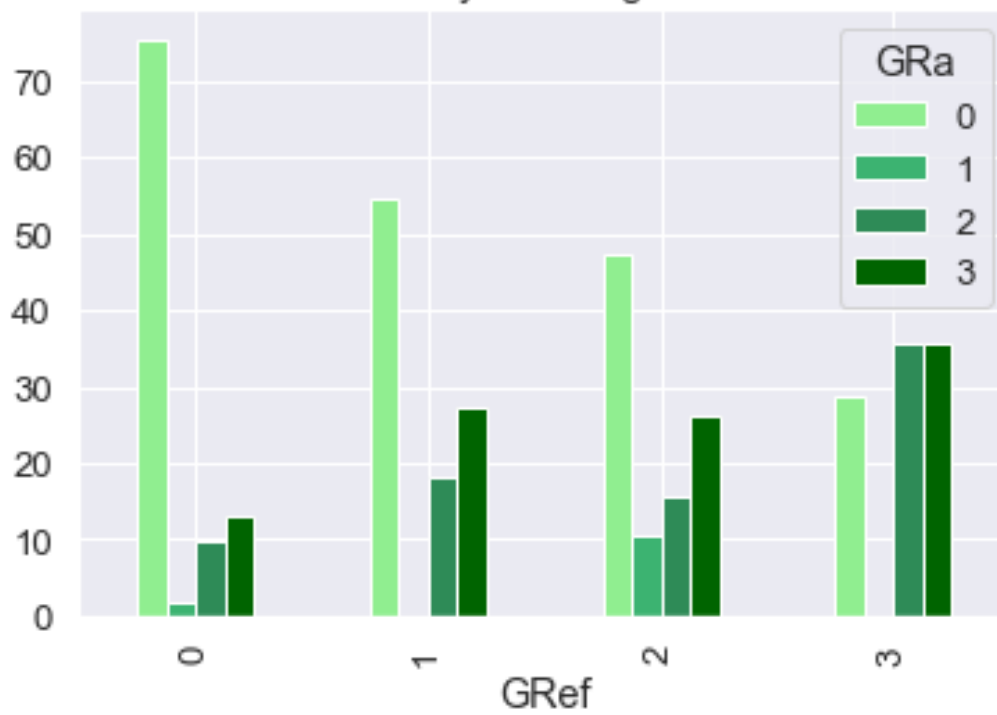
Pactos para refugiados por Región



Pactos para refugiados por Región (Normalizado)

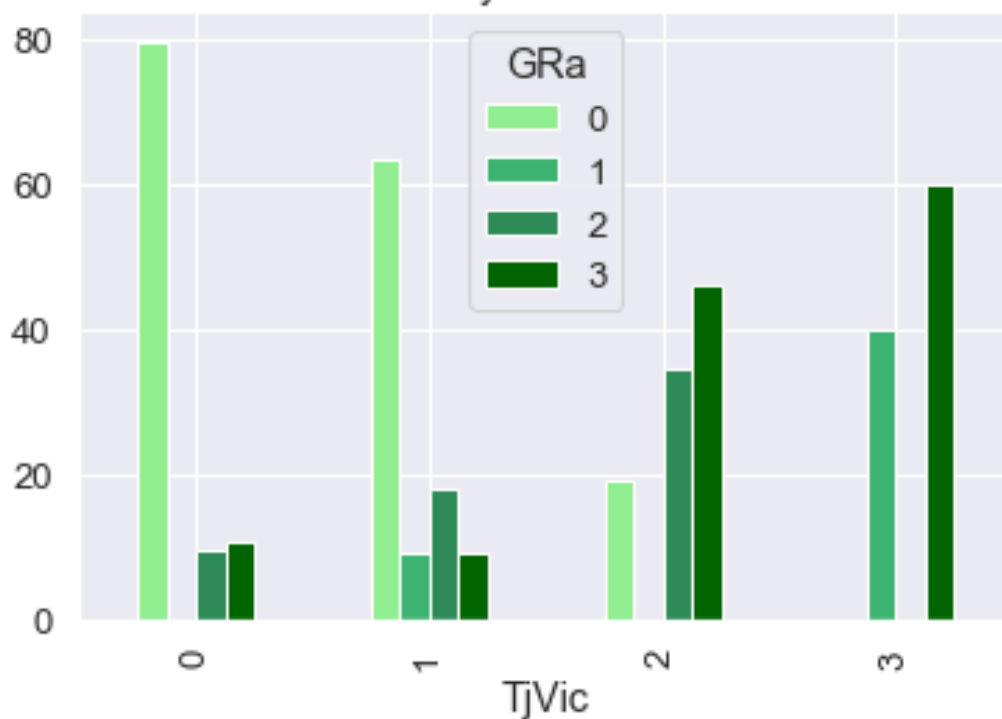


Relación entre la raza y los refugiados en los acuerdos

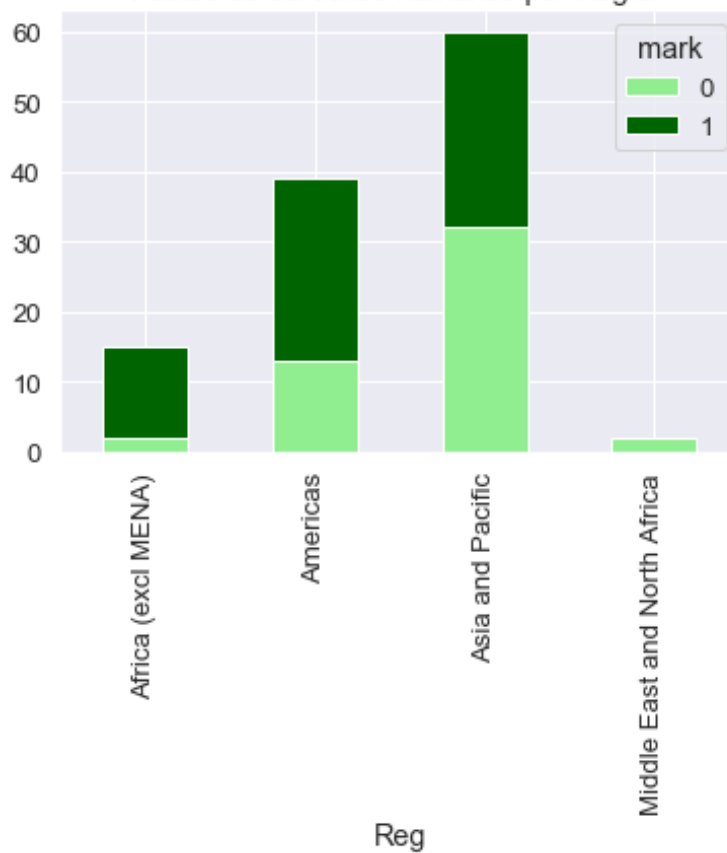




Relación entre la raza y las víctimas en los acuerdos



Pactos de derechos humanos por Región



Pactos de derechos humanos por Región

