

Visualización de datos

PRA1: Selección del conjunto de datos.

Autor: Francisco Sánchez Jiménez

MASTER UNIVERSITARIO EN CIENCIA DE DATOS

Acuerdos de paz para las poblaciones indígenas

1. Selección y justificación de la propuesta

El objetivo principal del documento será analizar una selección de tratados de paz extraídos del proyecto *Peace Agreements* [1] de la Universidad de Edimburgo (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.

La elección de las poblaciones indígenas como tema central guarda relación con mi interés por la diversidad de la especie humana. Me parece apasionante como cada una de estas poblaciones representa una cultura diferente y una forma completamente distinta de ver el mundo. Considero vital la preservación de las pocas comunidades de poblaciones indígenas que quedan hoy día y, por ese motivo, he decidido escoger estas como punto central del proyecto. En la UNESCO [2] podemos aprender más acerca de estas poblaciones indígenas y su situación política y social actual

La elección de esta base de datos se fundamenta también en su riqueza en lo que respecta a la variedad de tipología de datos. Se trata de un fichero de tipo CSV, por lo que se trata de un conjunto de datos estructurado que además presenta una gran cantidad de atributos sobre los cuales entraremos en detalle más adelante. Esta riqueza de formato del dato puede ayudarnos a desarrollar y mejorar la calidad de las visualizaciones que podamos realizar, explotando las posibilidades del dato.

2. La relevancia del conjunto de datos en su contexto

El propósito del proyecto será, por lo tanto, elaborar una visualización que ayude a llamar la atención sobre la cuestión de las poblaciones indígenas, muchas veces olvidada. Con este objetivo en mente, trataremos de extraer conclusiones y curiosidades de la base de datos que puedan ser llamativas para el usuario. De este modo, buscaremos concienciar al usuario de la importancia de preservar la cultura, religión, forma de vida... de estas minorías.

Además, intentaremos hacer comprender al usuario las difíciles condiciones en las que se encuentran y el peligro que corren si no las protegemos.

3. Descripción del *dataset*

El *dataset* recopila un total de 116 acuerdos de paz con diferentes medidas para la convivencia y preservación de comunidades de poblaciones indígenas. Para cada uno de los acuerdos disponemos de 265 metadatos o atributos, generando un total de 30740 registros.

El enorme número de atributos dificulta la descripción e identificación de cada uno de ellos de manera individual. Por esta razón, hemos procedido a agrupar los atributos en grupos, según el tipo de datos que los constituyen. La descripción de la totalidad de los atributos puede ser encontrada en el manual de uso [3] de la base de datos proporcionado por *Peace Agreements*. La clasificación de los atributos queda reflejada en los siguientes apartados.

3.1. Atributos categóricos.

Los atributos categóricos nos ayudarán a identificar el acuerdo y definir sus características generales. Los atributos categóricos más relevantes con los que trabajaremos en el proyecto 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:
 - 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
- LOC1ISO: Localización del conflicto entre estados. (Asignada aleatoriamente)
- LOC2ISO: Localización del conflicto entre estados. (Asignada aleatoriamente)
- 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; los acuerdos pertinentes a menudo tratan con gobiernos revolucionarios provisionales.
 - 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 acuerdo distintos:
- Multiparty signed/agreed: Acuerdo firmado o claramente aceptado por más de un grupo de participantes opuestos (pero no necesariamente todos los grupos opuestos).
 - Unilateral agreement: Acuerdo unilateral producido, por una parte, 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 / acordado.
 - Agreement with Subsequent Status: Acuerdo propuesto basado en el diálogo entre las partes y una suposición de lo que acordarían que, aunque no fue aceptado por las partes como un acuerdo, se convirtió la base para desarrollos posteriores.
- Stage: Fase en el que se encuentra el acuerdo. Variable categórica con siete fases:
- 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.
 - Other: Otros tipos de acuerdos que se encuentran en estados distintos a los anteriores.

3.2. Atributos numéricos.

Los atributos numéricos generalmente indicarán el grado de relación del acuerdo de paz con un ámbito concreto como, por ejemplo, si el acuerdo incluye medidas que tengan en cuenta distintas razas o si existe un factor religioso en el mismo. También tendremos una serie de atributos binarios que tomarán los valores 1 y 0, en función de si el acuerdo de paz aborda el tema especificado o no, respectivamente.

Para evitar extendernos a atributos que no tengan gran peso en nuestra selección de acuerdos indígenas, nos centraremos en aquellos que aparecen un mayor número de veces.

El primer grupo de atributos numéricos, los que indican el grado de relación, irán 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

El resto de variables numéricas serán binarias, las cuales tomarán 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

4. Cuestiones que nos plantearemos.

- ¿Existe alguna tendencia de implicación en los pactos con respecto a algún grupo en concreto?
- ¿Cómo influye la raza y la religión en estos pactos?
- ¿Cómo influye la presencia de los refugiados en los pactos?
- ¿Existen indicios de que muchos de los conflictos de odio sean ocasionados por motivos raciales?
- ¿Cómo se distribuyen los derechos humanos en las diferentes regiones?

Además de estas cuestiones, se podrán plantear y dar respuesta a otras que vayan apareciendo a lo largo de la realización de la Práctica.

A parte de una representación individual detallada para cada uno de los acuerdos de paz, el objetivo fundamental del proyecto de visualización será representar las relaciones existentes entre los pactos de cada una de las regiones. En un principio buscaremos patrones que se repitan dentro de los pactos de una sola región, intentando ver si existe una correspondencia entre el tipo de acuerdos y la región donde tienen lugar.

Posteriormente, se representarán de manera conjunta las distintas regiones en función de los atributos señalados anteriormente. De este modo, se podrán comparar los resultados en busca de tendencias remarcables entre los tratados correspondientes a las diferentes regiones. Por otro lado, en la visualización trataremos de representar la distribución que siguen los acuerdos cuando comparemos dos atributos al mismo tiempo. Uno de ellos consistiría en representar el grado racial frente al grado religioso de los acuerdos, apareciendo cierta relación y proporcionalidad entre ambos atributos.

Aunque no será vital para alcanzar los objetivos establecidos, también se representaran los atributos categóricos para las distintas regiones, facilitando la clasificación del tipo de tratados que predomina en la selección de pactos.

En resumen, el proyecto *tendrá* dos objetivos principales con respecto a los datos a exponer:

1. Identificar las características generales que aparezcan en los acuerdos de unas regiones u otras.
2. Identificar posibles tendencias y relaciones entre los tipos acuerdos según su implicación en los distintos ámbitos (raza, religión, víctimas...)

5. Selección de visualización

Dado que buscamos mostrar diversos factores en una misma visualización, tendremos que recurrir a múltiples tipos de gráficas y herramientas a la hora de representar cada uno de ellos. Según el tipo de hecho que queramos representar tendremos los siguientes tipos de gráficas:

- La exposición de los 116 pactos de paz que integran nuestro *dataset* se hará de manera conjunta en una línea temporal. De este modo, los acuerdos se colocarán en orden cronológico según la fecha en la que fueron firmados. La corta extensión de la base de datos nos permite agrupar todos los pactos y ordenarlos de modo que el usuario pueda ir seleccionando y aprendiendo sobre las características generales de cada uno. La interacción del usuario con los datos se detallará en mayor medida en el siguiente apartado (sección 8). La línea temporal irá acompañada por un histograma que muestre la distribución del número de pactos a lo largo de los años. De este modo, podremos apreciar donde hubo un mayor cúmulo de los mismos, indicando etapas conflictivas o sucesos históricos relevantes. También podrá venir acompañado de un mapa geográfico que nos indique de manera visual dónde se produjo el acuerdo.
- La visualización de las características generales de una región concreta se realizará mediante gráficos de barras en los cuales se representarán los distintos atributos (racial, religioso, refugiados...). Esto nos permitirá ver el grado de implicación de los acuerdos de una región concreta en un ámbito determinado.
- La comparación de las características de varias regiones se hará combinando en una única representación las gráficas de barras anteriores. De este modo, podremos ver en que regiones aparecen más un tipo de pactos que en otras. Añadir un mapa geográfico donde se señale el grado de implicación de los tratados en un ámbito concreto con colores que varíen en intensidad también podría ayudar al usuario. Observando el mapa podríamos saber de manera bastante intuitiva donde se concentra la mayoría de pactos de un tipo u otro.
- Las relaciones entre los atributos se realizarán de manera conjunta para todas las regiones, pues no buscamos analizar si hay un comportamiento concreto en una región, sino si la tendencia puede aplicarse a la mayoría de acuerdos de paz. En este caso, la utilización de un gráfico de barras que represente uno de los atributos en función del otro, por ejemplo, raza frente a religión, nos permitirá una buena comprensión de la relación entre ambos. En este apartado también podría ayudarnos una tabla de contingencia que nos señale si los grados de un atributo se corresponde con los del otro. Es decir, si, por ejemplo, la mayor parte de pactos con rango racial 2 tienen a su vez rango religioso 2. De esta forma, podremos deducir de manera algo más estadística si los atributos guardan relación.
- Por último, los atributos categóricos vendrán representados por gráficas de columnas que nos permitan ver las características generales de la selección de tratados con la que estamos trabajando (tipo de conflicto, naturaleza, estado del acuerdo...). Estas gráficas se realizarán para el conjunto de todas las regiones.

Referencias

- [1] <https://www.peaceagreements.org/>
- [2] <https://es.unesco.org/indigenous-peoples>
- [3] <https://www.peaceagreements.org/search>