

Visualització de dades: PAC 2

Gisela Claret Tortajada

18 d'abril del 2021

1. Títol de la visualització on es presenten el dataset o datasets escollits

EVOLUCIÓ TEMPORAL I SITUACIÓ ACTUAL DELS ACORDS DE PAU DELS DRETS DE LES DONES AL MÓN ENTRE EL 1990 I EL 2020

2. Descripció curta del document i del que s'hi presenta

Aquest document té com a objectiu realitzar un resum descriptiu del *dashboard* creat a partir de les dades dels tractats de pau en l'àmbit mundial entre els anys 1990 i 2020. En concret, el *dashboard* se centra en els tractats de pau que tenen relació amb els drets de la dona; s'hi presenta l'evolució temporal dels últims 30 anys i també es presenta la situació actual d'aquests acords. Per tant, aquesta visualització suposa una primera aproximació a aquest conjunt de dades, amb el que es vol respondre les preguntes: "Com han evolucionat els tractats de pau dels drets de les dones els últims 30 anys?" i "Quina és la situació actual d'aquests tractats de pau?"

El conjunt de dades *Peace Agreements Database* s'ha extret des del web: <https://www.peaceagreements.org/search> pertanyent a la Universitat d'Edimburg el 05/04/2021. El motiu de l'elecció de la creació d'un *dashboard* sobre els acords de pau dels drets de les dones al món és degut a la importància de difondre entre la societat l'evolució i la situació actual d'aquest tema.

3. Presentació de les dades

El conjunt de dades inicial el qual s'ha obtingut des de l'enllaç proporcionat anteriorment consta de 266 variables que descriuen les característiques de 1.868 acords de pau diferents. No obstant això, en el GitHub presentat en l'apartat 6 també s'inclou un conjunt de dades generat en el que solament hi consten les variables necessàries per a la realització del *dashboard*.

Es presenta un conjunt de dades anomenat "dades_genere.xlsx" en el que s'han inclòs les següents 9 variables:

- **Con:** El nom del país o jurisdicció on es va originar el conflicte. Tipus: cadena de caràcters
- **Contp:** El tipus de conflicte pel qual es va crear l'acord de pau. Tipus: cadena de caràcters
- **Reg:** La regió del món en el qual tenia lloc el conflicte. Tipus: cadena de caràcters

- **Dat:** Data en la qual es va signar l'acord, en el format YYYY-MM-DD. Tipus: cadena de caràcters
- **Stage:** Situació actual de l'acord de pau. Tipus: cadena de caràcters
- **Loc1ISO:** Codi ISO del país o regió principal en el qual es genera el conflicte. Tipus: cadena de caràcters
- **Loc2ISO:** Codi ISO del país o regió on es genera el conflicte. Solament existeix aquest valor si hi ha dues localitzacions principals on es desenvolupa el conflicte. Tipus: cadena de caràcters
- **GeWom:** Valor 0 o 1 segons l'acord sigui (1) o no (0) referent als drets de les dones. Tipus: boolean

Obtenim doncs, 8 atributs qualitius les quals algunes serà convenient transformar-les en dades categòriques per facilitar l'anàlisi de les dades, com és el cas de "Contp" o "Reg". En canvi per la variable "Dat" serà convenient, una vegada importada a Tableau, transformar-la a tipus "Fecha" i les variables "Loc1ISO" i "Loc2ISO" caldrà transformar-les a dades geogràfiques per tal de poder realitzar representacions geogràfiques de les dades.

En canvi, pel que fa a la variable "GeWom" no seran necessàries transformacions.

4. Exploració de les dades

En l'exploració inicial de les dades s'ha pogut observar alguns fets evidents com ara:

- Solament 371 dels 1.868 acords del conjunt de dades representen acords de pau referents als drets de les dones, és a dir, solament un 19.86%.
- País amb més acords de pau referent als drets de les dones són Colòmbia (24) seguit per Sudan del Sud (23) i la República d'Àfrica Central (19).
- La major part dels acords de pau que fan referència als drets de les dones són de la regió d'Àfrica (171), un 46%.

Resulta interessant obtenir quins són els països que tenen aquest tipus d'acord de pau amb el conflicte originat en dos països diferents. Hem observat que les combinacions de països que més acords de pau tenen amb aquestes característiques són: Sudan – Sudan del Sud (10) i Irlanda – Gran Bretanya (9) tots dos conjunts de països limítrofs.

A més a més, podem observar que fins a l'any 2015 el nombre d'acords de pau pels drets de les dones signats en un any al món tendia a augmentar, tot i això a partir del 2015 aquesta tendència ha estat negativa.

5. Procediment de l'exploració de les dades

Per la realització d'aquesta PAC s'ha utilitzat el llenguatge de programació Python a través de Jupyter Notebook i l'eina *Tableau*, el motiu pel qual he escollit el primer és perquè és el llenguatge en què tinc més experiència, i el segon l'he escollit perquè ha estat un dels recomanats en aquesta assignatura i he trobat que és una eina molt senzilla d'utilitzar i que permet crear una gran multitud de visualitzacions diverses.

El procés seguit ha estat, en primer lloc, he fet una anàlisi inicial de les dades utilitzant Python, s'han importat les dades i s'ha creat i desat un nou *dataframe*, solament amb el conjunt d'atributs desitjats per les visualitzacions. A més a més, s'ha visualitzat el nombre total d'acords de pau relacionats amb els drets de les dones segons diferents atributs, per exemple, segons el tipus de conflicte, segons la regió o segons l'estat de l'acord.

Un cop desat aquest subconjunt de dades s'ha importat a l'eina Tableau i s'han realitzat les visualitzacions presentades, escollint per cada una el tipus de visualització més adient i modificant els colors o opcions perquè resultin més entenedores. A més a més, en la major part de les visualitzacions s'ha realitzat un filtratge sobre "GeWom" per incloure solament les dades corresponents als acords relacionats amb els drets de les dones o excloure les dades nul·les.

Anteriorment he utilitzat algunes eines de visualització de dades com ara algunes llibreries dels llenguatges de programació Python, R o Matlab. A banda, d'això també he treballat amb eines com *Power BI*, *SAS* i *ArcGis Online*. M'agradaria aprendre a realitzar millors visualitzacions amb *Tableau* i també m'agradaria aprendre a utilitzar algunes de les eines recomanades a l'assignatura, com ara *Flourish* o *Google Data Studio*, però especialment *RAWGraph*, ja que amb un bon tractament de les dades s'aconsegueixen molt bones visualitzacions, també m'agradaria poder crear un storytelling que resulti interessant.

6. Visualització sobre les dades

La versió interactiva del dashboard proposat per aquesta PAC es pot trobar en el següent enllaç:

<https://public.tableau.com/profile/gisela5841#!/vizhome/PAC2-Visualitzacidedades/Dashboard1>

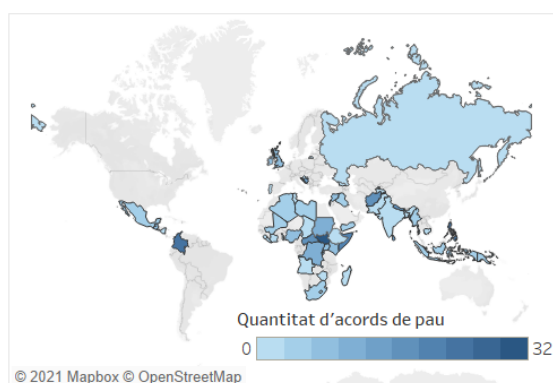
A més a més, també s'ha creat un GitHub que inclou les dades, el codi Python i l'arxiu .twb de la creació del *dashboard*. L'enllaç és el següent:

https://github.com/giselaclaret/PAC2_visualitzacio_de_dades

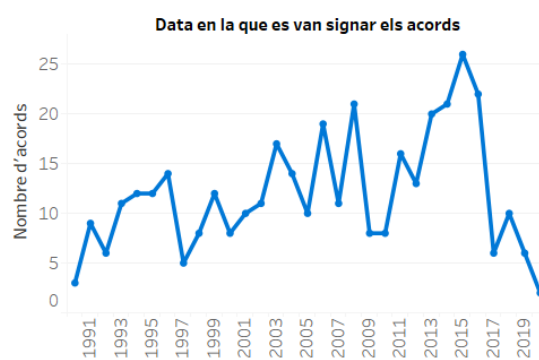
El dashboard presentat en l'enllaç és el següent:

ELS ACORDS DE PAU SOBRE ELS DRETS DE LES DONES

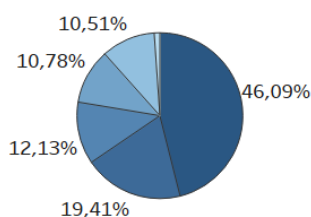
Representació geogràfica de la quantitat d'acords de pau



Nombre d'acords de pau signats per any



Acords segons regions



Conflictes amb l'origen en dos països diferents

Loc. 1	Loc. 2	
SSD	SDN	10
GBR	IRL	9
GEO	RUS	2
IDN	PRT	1
KEN	ETH	1
SOM	ETH	1

Els 5 països amb més acords de pau des del 1990

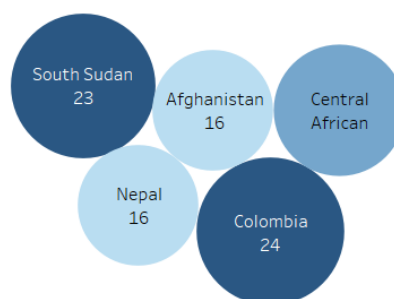


Figura 1: Dashboard dels acords de pau signats sobre els drets de les dones entre el 1990 i el 2020

En ell s'inclouen diferents visualitzacions de les dades que proporcionen una visió global de les dades proposades i també alguns detalls més específics. En concret es planteja una visió dels acords de pau relacionat amb els drets de les dones. Al llarg del *dashboard* es presenten visualitzacions de diferents tipus, però sempre amb la mateixa escala de colors. Aquesta varia entre colors més clars a més foscos segons el nombre total d'acords, essent els elements amb més nombre d'acords els pintats amb una tonalitat de blau més fosca, i els menys amb una tonalitat de blau més clara.

En primer lloc s'inclou un mapa en el qual hi ha una representació geogràfica dels països que han signat com a mínim un acord de pau d'aquest tipus des del 1990 fins al 2020. En ell observem els països que més acords han signat d'aquest tipus com són Colòmbia o Sudan del Sud i també els països que no n'han signat cap, com ara Austràlia o Noruega.

En segon lloc es presenta un gràfic circular que recull quines són les regions amb més i menys acords de pau signats. En aquest gràfic és important utilitzar el recurs d'informació emergent, així, si passem el cursor per sobre dels diferents talls es pot observar de quina regió es tracta, del percentatge total que representa i també el nombre d'acords signats en aquesta regió, tal com es mostra en la Figura 2. En aquest gràfic circular observem que Àfrica és la regió que més acords ha signat amb un total de 46% dels acords.

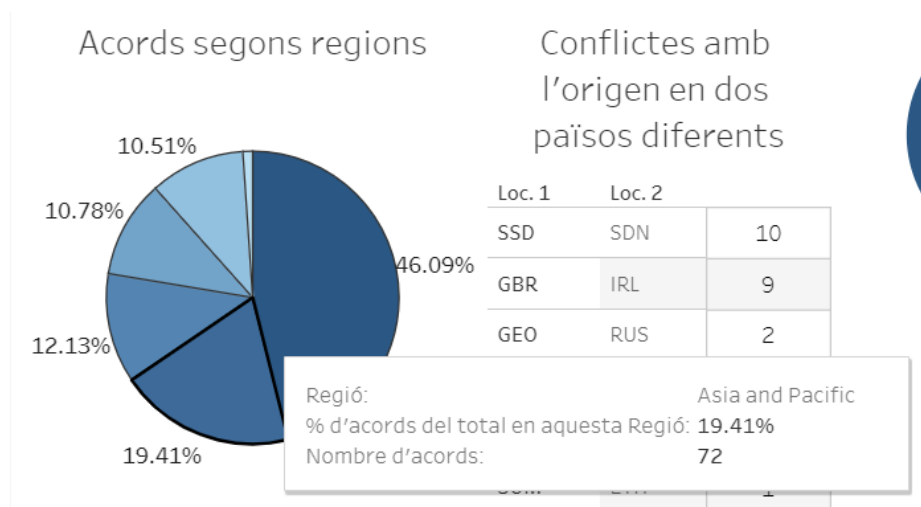


Figura 2: Exemple de la interacció amb el gràfic circular del dashboard dels acords de pau sobre els drets de les dones al món entre el 1990 i el 2020

Seguidament es presenta una taula de text en la que es presenten els països amb conflictes originats en dos països a la vegada. Per exemple Sudan del Nord i Sudan del Sud, o Gran Bretanya i Irlanda.

També s'ha inclòs un gràfic de línies en el que s'inclou una evolució temporal del nombre total d'acords signats que fan referència als drets de les dones. Podem veure com fins el 2015 la tendència era ascendent però a partir d'aquell moment el nombre total d'acords signats en un any disminueix.

Finalment s'inclou la informació dels 5 països que més acords de pau sobre els drets de les dones han signat des del 1990, mitjançant un gràfic de bombolles agrupades. Aquest països són Colòmbia, Sudan del Sud, la República d'Àfrica Central, Afganistan i Nepal.

Cal mencionar que tots ells contenen interacció amb el ratolí que serveix per especificar el nombre exacte o afegir informació rellevant.