

Projecte de visualització de dades

PAC 4

Títol de la visualització

Acords de Pau sobre Violència de Gènere entre Països

URL de la visualització

https://public.tableau.com/views/AcordsdePausobreViolnciadeGnereentrePasos/Historia1?:language=es&:display_count=y&publish=yes&:origin=viz_share_link

Descripció

Per a definir el dashboard que representarà les dades cercades, s'ha utilitzat el programa Tableau. Aquest presenta una sèrie de visuals distribuïts en dues pàgines distingint dues dimensionalitats: Geografia i Temps.

S'ha volgut fer un disseny que representi, per la part de Geografia, mapes amb les diferents regions que participen en els acords de pau, gràfics de barres per mostrar les tipologies dels acords en ordenades per les diferents regions en abscisses i plots per visualitzar la quantitat d'acords segons la dimensió estat. Per la part de temporalitat s'han volgut representar gràfics de línies i de barres que mostren evolucions històriques dels acords.

Si analitzem els visuals obtinguts podem destacar que:

En la visualització es pot observar que, Geogràficament hi ha quantitat considerable d'acords firmats sobretot en les regions d'Àfrica i Àsia i Pacífic, això vol dir que la majoria de conflictes es donen entre països d'aquestes regions.

En quant a l'etapa de l'acord es pot destacar que la majoria són pre-negociacions i que sobretot aquestes es solen donar en les regions de Àsia i Pacífic i d'Europa i Euràsia. En el cas d'Àfrica abunden les etapes d'acords que es troben en "Comprehensive" i "Partial". En el cas de l'estat destaca el "Multiparty signed/agreed, la qual cosa vol dir que la majoria d'acords han estat firmats o acceptats per més d'un grup de protagonistes del conflicte de ambdues parts. A més, podem veure que sobretot en els acords on intervenen països d'Àfrica s'observen que hi ha moltes mencions de crims i també referències a famílies. En la resta de regions la proporció entre aquests dos conceptes es troben bastant igualats, destacant que en la majoria casi no hi ha mencions de sexe masculí.

Pel que fa a la temporalitat es pot veure una fluctuació en el nombre d'acords per any, havent una estacionalitat cada 5 anys aproximadament, mostrant un pic d'acords. En el cas dels mesos s'observa que hi ha una quantitat més alta en els mesos de febrer i desembre.

Finalment, es pot veure que hi ha una evolució més contínua en el cas de les regions d'Àsia, Àfrica i Europa i més discreta pels casos de Amèrica, Middle East i Cross Regional. En tots els casos excepte Amèrica, es veu una tendència a la baixa en quant a la quantitat d'acords firmats.

Què hem après?

Aquest projecte ha constatat de tres pràctiques en les quals s'ha dut a terme tot un procés d'exploració, anàlisi i visualització de dades. En la primera pràctica vam veure com treure dades totalment desestructurades d'una base de dades de documents, les quals contenien moltes metadades i es desconeixia el seu significat; vam aprendre a llegir-les i a buscar tota la informació per saber-les interpretar. En la segona pràctica s'ha realitzat tot el processat d'aquestes dades, s'han convertit en una taula plana, s'han mapejat els seus valors i s'han extret aquelles variables d'interès per fer-ne una visualització. Finalment, en aquesta última pràctica s'ha presentat la visualització d'aquestes dades. S'han pogut interpretar tant a nivell geogràfic com temporal, amb la creació de nous atributs i noves mesures que han permès interpretar el seu significat.

Resultat final

Els objectius inicials d'aquesta pràctica són:

- Obtenir informació de la pàgina peaceagreements.org sobre els acords de pau oficials així com les variables que s'ofereixen i quines són interessants per analitzar o visualitzar
- Processament de les dades que acompanya la correcta visualització d'aquestes
- Plasmar en forma de *dashboard* aquesta informació

Després de la realització de les tres pràctiques es pot afirmar que s'ha arribat als objectius plantejats.

L'esforç ha estat considerable, ja que s'ha partit d'un origen de dades totalment desconegut, tant a nivell informatiu com estructural. S'ha aconseguit interpretar tota la informació que oferia la base de dades de documents, cercant el significat les des gairebé 300 variables que proporcionava el conjunt de dades resultant. De totes aquestes variables s'han seleccionat tan sols aquelles que han semblat interessants per a realitzar una visualització i també se n'han creat de noves per tal d'extraure més informació. Per la part de processat s'ha après a utilitzar el llenguatge python per "netejar" les dades i crear un nou data set amb el qual treballar. Finalment, per representar-les, tot i haver utilitzat en una altra pràctica l'eina Tableau, s'ha après a utilitzar moltes més funcionalitats que ofereix l'eina, tant a nivell de visuals com de càlculs i filtratge de dades.

Descripció tècnica.

Per al processat de dades s'ha utilitzat el llenguatge Python3 i s'ha fet a través del desenvolupador de programari obert Jupyter Notebook, el qual s'ha instal·lat a través de la interfície gràfica Anaconda Navigator. S'han utilitzat sobretot llibreries de numpy i pandas per treballar amb dataframes i obtenir informació ràpida del data set i amb matplotlib per crear visualitzacions de manera fàcil i interpretable.

Per a la part de creació de visualització s'ha aprofitat la llicència de Tableau Desktop que ofereix l'assignatura. S'han creat diferents fitxes – una per cada visual, s'han creat dos *dashboards* – un per la geografia i un per la temporalitat, i finalment s'ha creat un *llibre de treball*, el qual s'ha publicat a Tableau Public.

Tot el codi i els documents es troben disponibles a un repositori GitHub accessible a través de la següent URL:

<https://github.com/Lorena96/PA-X-Dataset-Analysis>