

Beyond Peace Agreements

Josep M^a Espasa Verdés

Màster universitari en Ciència de dades

gener de 2020



[Aquesta obra està subjecta a una llicència de Reconeixement 3.0 Espanya de Creative Commons](https://creativecommons.org/licenses/by/3.0/es/)

Índex

1. Títol.....	4
2. URL de la visualització.....	4
3. Descripció.....	4
4. Coneixements adquirits.....	4
5. Valoració del resultat final.....	5
6. Descripció tècnica del projecte.....	7
6.1 Llenguatges i programari.....	7
6.2 Transformació de dades.....	7
6.3 Principals tasques realitzades a Tableau.....	8
7. Codi.....	9

1. Títol

"Beyond Peace Agreements"

2. URL de la visualització

El projecte de visualització objecte d'aquest document es pot consultar lliurement a la següent url: <https://ilergeta.github.io/PAX-Viz/>

3. Descripció

El present document presenta el projecte de visualització "Beyond Peace Agreements", basat en les dades dels tractats de pau signats entre el 1990 i el 2019, recollides per la base de dades PA-X.

L'objectiu bàsic de la visualització és doble, d'una banda donar una visió general de la distribució territorial i tipològica dels tractats de pau signats i, de l'altra, mostrar quina és la incidència dels conflictes que hi ha darrere dels tractats de pau signats en diferents àmbits del dia a dia d'un país com són l'economia, la salut, la qualitat de vida i l'educació, intentant, així, mitjançant l'anàlisi d'un país en concret (Colòmbia) que s'ha considerat representatiu com a país amb conflictivitat històrica, mostrar i quantificar la naturalesa polièdrica de les conseqüències d'aquests conflictes. Tot això, intentant fer-ho de manera clara i visualment atractiva.

Amb aquest projecte s'espera poder aportar un granet de sorra per tal d'intentar entendre què hi ha darrere d'un conflicte i, així, poder

4. Coneixements adquirits

Aquest projecte de visualització el considero una eina d'aprenentatge molt important, que m'ha ajudat, a nivell tècnic, a conèixer força en profunditat una eina, fins fa uns mesos desconeguda per a mi, com Tableau Desktop i a descobrir tot el seu potencial. No entraré aquí a explicar en detall tots els coneixements adquirits sobre el programa, perquè tampoc considero que sigui l'objectiu d'aquest document, en tot cas, les tasques de recerca inicial a la galeria de Tableau Public, la possibilitat de descarregar-se els fitxers i la gran quantitat d'informació i tutorials disponibles per la xarxa han estat claus per poder assolir els coneixements necessaris per realitzar aquesta pràctica amb la qualitat final obtinguda.

D'altra banda, l'elecció del format de la visualització (DIN-A4 apaïsat) per tal d'afavorir, tant la seva visualització en pantalla com la seva impressió, ha suposat un repte important per tal d'aconseguir sintetitzar tot allò que volia explicar/mostrar en l'espai limitat disponible, tot

aconseguint un layout ordenat i estèticament atractiu. Certament ha estat la part més difícil del projecte i ha suposat un esforç de síntesis important del que, alhora, una vegada vist el resultat final, n'estic més orgullós.

Com a limitació tècnica més important que he trobat a l'hora de realitzar aquest projecte de visualització ha estat el servei de núvol de Tableau Public, efectivament, allò que també ha estat un aspecte positiu, com és compartir fàcilment i de manera gratuïta la feina realitzada sense la necessitat de cap infraestructura addicional, també ha estat una limitació, ja que es disposa del recurs que es disposa, sense possibilitats d'adaptació a les necessitats del projecte. Així, per exemple, s'ha "descobert" com la galeria de Tableau Public no permet visualitzar imatges per usuaris no loggejats a Tableau Public, això ha condicionat, sens dubte el disseny de la visualització i a fer "trampes" per evitar l'ús d'imatges. Una altra limitació que s'ha hagut de superar és la velocitat de processament de la visualització una vegada penjada a Tableau Public, això ha condicionat en gran mesura la visualització del mapa i dificulta la interactivitat de l'usuari amb aquesta visualització, espero que aquest darrer aspecte, que afecta al resultat final de la visualització es tingui en compte en la valoració d'aquest projecte.

A part d'aquestes limitacions tècniques referents al servei de núvol de Tableau Public cal indicar que l'eina de Tableau Desktop ha complert amb escriure les expectatives creades i, almenys pel present projecte, no ha suposat cap limitació.

Analitzant les possibles limitacions conceptuals, cal dir que no s'han considerat com limitacions cap aspecte conceptual del projecte de visualització, al contrari, s'ha intentat veure-ho sempre com una oportunitat i s'ha intentat anar més enllà d'una simple anàlisi visual d'un conjunt de dades de tractats de pau, incorporant-hi fonts externes d'informació amb l'objectiu d'aconseguir un punt de vista el més enriquidor possible.

5. Valoració del resultat final

Respecte a l'esborrany inicial presentat en la PAC3, cal dir que el present projecte de visualització aconsegueix acomplir tots els objectius marcats amb escriure.

Concretament, s'ha aconseguit implementar el mapa de països amb els acords de pau i la seva evolució al llarg dels anys, el diagrama de línies per a cadascun dels aspectes a analitzar i, finalment, però no menys important, el diagrama radial de barres apilades. A més a més, sota el punt de vista de l'autor del projecte, s'ha aconseguit millorar el seu disseny (millora en els colors, utilització de menus amb icones en lloc de textos) i s'han incorporat interaccions entre els diferents gràfics que milloren el resultat final de manera significativa, així, al seleccionar un

sector del diagrama radial, si s'està visualitzant les regions, el mapa filtra la selecció centrant-se en els països i any corresponents, en cas d'estar visualitzant-se la tipologia de conflictes en el diagrama radial, el filtre del mapa únicament és per l'any corresponent; a més, indicar que la selecció dels sectors del diagrama radial pot ser múltiple, fent una finestra de selecció sobre els sectors o prement la tecla "ctrl"; així, el diagrama radial, en certa manera, es pot emprar com assistent per navegar i mostrar informació en el mapa.

L'esforç i la dedicació en aquest projecte de visualització han estat importants, amb moments difícils, sobretot durant la implementació del gràfic de barres apilades radials, considerat un gràfic clau en el projecte de visualització. Així a part del tractament de dades previ que es requereix, la programació del gran nombre de variables necessàries per realitzar el gràfic i l'important nombre de factors i configuracions que cal tenir en compte per la correcta implementació del gràfic, va fer que molts dies es tingués la sensació que no s'avançava res en el projecte, ja que fins que no està totalment ben configurat, no s'aconsegueix cap tipus de visualització "coherent", posant en dubte, fins i tot, la viabilitat general del projecte de visualització plantejat. Finalment, amb la perseverança es va aconseguir obtenir el resultat desitjat.

Indicar, també, que hi ha hagut una important tasca de recerca de visualitzacions que es consideraven interessants o atractives visualment, no només en la fase inicial de disseny d'aquest projecte, sinó posteriorment, al llarg de tot el procés d'implementació, per tal d'intentar aplicar els conceptes i elements que, poc a poc, han anat millorant el present projecte.

En tot moment, durant la realització d'aquest projecte de visualització s'han intentat mostrar les dades de la manera més autoexplicativa possible, en aquest sentit, s'han implementat icones que permeten canviar la tipologia de visualització mostrada, a més a més del text. Igualment, l'elecció de les mateixes icones/formes mostrades ha estat molt personalitzada per tal de facilitar al màxim la comprensió de l'usuari de la visualització, així, la major part d'icones s'han dissenyat expressament per aquest projecte de visualització.

Finalment, sembla lícit, després d'haver-hi dedicat tantes hores a aquest projecte de visualització, aprofitar els coneixements adquirits i indicar cap a on es podria ampliar o millorar el treball presentat. Així, un aspecte interessant de millora d'aquest projecte, seria la incorporació en l'anàlisi detallada de més països, a part de Colòmbia, possiblement, en una primera fase es podria analitzar un país per Regió i, posteriorment, tots els països signants de tractats de pau. Analitzar l'abast d'aquesta millora és important, ja només la primera fase implicaria definir 4 països que es puguin considerar interessants/representatius, un per cada regió addicional, processar les dades dels 4 àmbits, és a dir, 16 fons d'informació addicionals i, a més, afegir un altre selector per tal que l'usuari pugui elegir el país. Com es pot veure, la incorporació d'aquesta informació no és trivial i, per tant, es considera que està fora de l'àmbit

d'aquesta pràctica, podent-se incorporar perfectament en una possible futura tasca de recerca, ja que es considera que, de ben segur, pot aportar informació valuosa.

Amb tot això, la valoració global del resultat obtingut no pot ser altra que molt positiva i enriquidora.

6. Descripció tècnica del projecte

6.1 Llenguatges i programari

Com a eina bàsica de creació d'aquest projecte de visualització s'ha emprat el programari *Tableau Desktop Professional Edition*, concretament la versió 2019.4.0, disponible gratuïtament gràcies a la llicència gratuïta per a estudiants. Igualment, s'ha penjat la visualització final a la plataforma *Tableau Public* per tal de fer-la pública per tothom, sense necessitat de loggear-se, ni tenir instal·lat el programari *Tableau Desktop* i tampoc sense necessitat d'aixecar cap tipus d'infraestructura addicional. Finalment, indicar que s'ha enllaçat la visualització de Tableau Public a una pàgina web per tal de poder mantenir l'enllaç en cas d'actualitzar la visualització, s'ha optat per GitHub pages per la facilitat en la seva configuració i per la possibilitat d'unificar la visualització amb el repositori de codi de l'entrega.

Respecte a tot el procés de neteja i transformació de dades per tal d'aconseguir els datasets, en aquest cas fitxers de tipus csv, que han estat emprats a *Tableau Desktop*, s'ha realitzat mitjançant el llenguatge de programació Python3 i les seves llibreries numpy, pandas i matplotlib. Igualment, s'ha utilitzat Jupyter notebook com entorn de programació.

6.2 Transformació de dades

El preprocessament de les dades que ha requerit aquest projecte de visualització, per tal d'obtenir, a partir dels fitxers descarregats en les diferents fonts emprades, els fitxers que s'han emprat com a font en el programari de *Tableau Public* ha estat:

Càrrega de les dades: en aquesta fase, s'han carregat els diferents fitxers que es poden trobar en el repositori GitHub dins de la carpeta notebook-data/data/in en objectes DataFrame de la llibreria pandas. Indicar que les fonts d'informació han estat: [PA-X Database](#), [The World Bank](#) i [UNDP](#).

Transformació de les dades: a partir dels DataFrames obtinguts en l'anterior fase, s'han realitzat les següents transformacions:

- Afegir el cap d'any, a partir de la data de signatura del tractat de pau.
- Afegir el cap de país únic per tal de passar d'una entrada amb la cadena: Pais1/Pais2/PaisN a N entrades, una per cada país i,

d'aquesta manera poder comptar de manera correcta el nombre de tractats de pau en els que cada país ha intervingut. Indicar que tota la informació del tractat de pau s'ha duplicat per cada nova entrada creada, ja que s'ha considerat que era la manera de poder estudiar correctament la participació de cada país als diferents tractats de pau, ja sigui de manera única o compartida amb altres països o regions.

- Afegir el camp de regió correcte en aquelles entrades amb valor "Cross-regional". Cal indicar que això s'ha realitzat per disposar del valor correcte de Regió per cada entrada de país, però degut a la tipologia de visualització emprada en aquest projecte, aquest camp no s'ha emprat, seleccionant-se, al seu lloc al valor original de regió, per tal de poder-se mostrar de manera correcta els països associats a conflictes "Cross-regional" quan són seleccionats en el diagrama radial de barres apilades segons regió.
- Afegir la informació provinent d'altres fonts diferents a la PAX, mitjançant l'addició de les files i columnes necessàries.

Exportació de les dades: una vegada obtinguts els DataFrame amb la informació i el format desitjats, s'han exportat a fitxers del tipus CSV (comma-separated values), aquests fitxers de sortida es poden trobar en el directori notebook-data/data/out del repositori de GitHub.

Tot aquest procés, tal com s'ha indicat anteriorment, s'ha realitzat amb un fitxer de tipus Jupyter notebook que es pot trobar en el directori notebook-data del GitHub.

6.3 Principals tasques realitzades a Tableau

Una vegada creats els fitxers CSV de l'apartat anterior, s'ha procedit a elaborar els diferents gràfics/elements que componen el present projecte de visualització mitjançant el programari de Tableau Desktop.

Els components principals de la visualització realitzada són: un mapa on s'han representat geogràficament els diferents tractats de pau signat per país i any, un diagrama radial de barres apilades on es mostren, per cada sector angular un any i per cada sector radial una tipologia (podent-se triar regió o tipologia de conflicte) els tractats de pau signats corresponents i un diagrama de línies que mostra pel cas de Colòmbia i de la mitjana de la regió, l'evolució de diferents indicadors d'economia, salut, educació i qualitat de vida, sobreposats amb els tractats de pau signats al llarg de l'any. Igualment, s'han implementat les interaccions entre els diferents gràfics que s'han detallat en l'apartat 5 d'aquest document.

A continuació es descriuen les tasques més significatives realitzades en el programari de Tableau Desktop per tal d'obtenir la present visualització.

Per la creació del mapa s'ha hagut d'assignar manualment la localització geogràfica dels països/regions no reconegudes oficialment com països, això s'ha fet indicant la latitud/longitud de la capital de la regió/país o del país oficial on pertany l'entitat segons s'ha considerat oportú per una correcta representació geogràfica del valor.

D'altra banda, per la realització dels dos diagrames radials de barres apilades, destacar que s'han afegit nombroses variables i camps calculats addicionals, per tal de poder obtenir els valors necessaris per realitzar el graf. Bàsicament, per cada barra que pertany a un sector angular (corresponent a un any) i a un sector radial (corresponent a una regió o tipologia de conflicte, segons el gràfic), s'han hagut de calcular cinc coordenades, corresponents als 4 vèrtexs del paral·lelogram que defineix cada barra més el cinquè punt que és igual al primer per tal de poder definir un polígon tancat. Amb aquests càlculs i emprant la tipologia de gràfic de "Path" s'han pogut realitzar els dos diagrames radials de barres que pertanyen a aquest projecte de visualització.

Finalment, per tal de poder seleccionar, de manera intuïtiva entre les diferents visualitzacions realitzades, s'han creat dos selectors, un per seleccionar entre els dos diagrames radials de barres apilades disponibles (segons regió geogràfica o segons tipologia de conflicte) i un altre per poder triar quin àmbit es vol mostrar en el diagrama de línies (economia, educació, salut o qualitat de vida). Per la realització d'aquests selectors, ha estat necessari crear unes taules addicionals (fitxers xls), que s'han afegit com una consulta tipus join a les fonts de dades originals dels gràfics per, així, poder ser emprades com filtres en els diferents fulls dels gràfics. Igualment, s'ha creat pròpiament el menu selector mitjançant icones personalitzades per, d'aquesta manera, facilitar la usabilitat de la visualització.

7. Codi

Tot el codi que s'ha implementat per tal de realitzar les tasques de transformació de dades descrites a l'apartat 6.2, es pot trobar en el repositori de GitHub següent:

<https://github.com/lldergeta/PAX-Viz/tree/gh-pages/notebook-data>

Igualment, la visualització és accessible en obert a través de l'enllaç URL indicat en l'apartat 2 d'aquest document, on, a més, si es desitja, es pot descarregar el fitxer del WorkBook des de la mateixa visualització que es pot obrir sempre que es disposi del programari de Tableau Desktop.