
Introducció a Tableau

PID_00246774

Víctor Pascual Cid

Temps mínim de dedicació recomanat: 2 hores





Els textos i imatges publicats en aquesta obra estan subjectes –llevat que s'indiqui el contrari– a una llicència de Reconeixement-Compartir igual (BY-SA) v.3.0 Espanya de Creative Commons. Podeu modificar l'obra, reproduir-la, distribuir-la o comunicar-la públicament sempre que en citeu l'autor i la font (FUOC. Fundació per a la Universitat Oberta de Catalunya), i sempre que l'obra derivada quedi subjecta a la mateixa llicència que el material original. La llicència completa es pot consultar a <http://creativecommons.org/licenses/by-sa/3.0/es/legalcode.ca>

Índex

1. Introducció a Tableau	5
1.1. Què és Tableau?	5
1.2. Per què Tableau?	6
1.3. Pros i contres de Tableau	7
1.4. L'ecosistema de Tableau	7
1.4.1. Tableau Desktop	8
1.4.2. Tableau Public	8
1.4.3. Tableau Reader	8
1.4.4. Tableau Server	8
1.5. Conceptes bàsics de Tableau	8
1.5.1. Preparació de les dades	8
1.5.2. La interfície d'usuari	11
2. Visualització de dades sobre la malària	16
2.1. Càrrega de dades a Tableau	16
2.2. Creació d'un gràfic de barres	17
2.3. Creació d'un mapa	18
2.4. Creació d'un <i>heatmap</i>	19
2.5. Creació d'un <i>dashboard</i>	20
Bibliografia	23

1. Introducció a Tableau

Aquest manual comença amb una breu guia d'aprenentatge de Tableau, una de les eines de programari més avançades que existeixen actualment per analitzar dades visualment i generar representacions interactives a fi de narrar històries mitjançant les dades. No pretén ser un manual exhaustiu de Tableau, ja que hi ha nombrosos recursos en obert sobre aquest tema, però sí un document d'introducció que expliqui què és l'eina, per a què serveix i com es pot començar a utilitzar de manera pràctica. Sempre que és possible es fa referència a la documentació original de Tableau, que descriu detalladament cadascuna de les funcionalitats utilitzades en aquest manual.

D'altra banda, hi ha un gran nombre de videotutorials gratuïts que ajuden a dur a terme els primers passos del procés d'utilització de Tableau. No obstant això, amb l'ús de l'eina veurem que no tot és tan senzill com sembla en aquests vídeos, ja que dominar l'eina implica conèixer profundament els elements bàsics de la teoria de visualització de dades. No es tracta de plantejar la creació d'una visualització des del «com es fa?», ni tan sols des del «què haig de fer?», sinó des del «per què?», i la resposta hauria de determinar les altres dues preguntes i les seves consegüents respostes. No obstant això, Tableau pot ajudar-nos a plantejar i respondre totes aquestes qüestions.

La segona part d'aquest manual és un recorregut per cinc vídeos que mostren un exemple pràctic d'utilització de Tableau per a la creació d'una visualització de dades relatives a la infecció pel paràsit de la malària al món. Aquest recorregut revisa els conceptes descrits a l'apartat anterior.

1.1. Què és Tableau?

Tableau és una eina de *business intelligence* enfocada especialment a facilitar l'anàlisi visual i interactiu de dades. Dit d'una altra manera, Tableau es pot veure com una espècie d'Excel o full de càlcul avançat, molt focalitzat i preparat per explorar dades de manera visual, al mateix temps que permet generar visualitzacions o *dashboards* (quadres de comandament) perquè altres persones explorin i analitzin conjunts de dades de manera interactiva.

Així doncs, és molt important tenir clares aquests dues vessants de l'eina: la d'explorar dades i la de comunicar o narrar històries. La facilitat d'ús de l'eina, que bàsicament utilitza menús i interaccions de *drag & drop*, dona una flexibilitat enorme per fer anàlisis exploratòries de dades. Aquest concepte va ser definit per John Tukey el 1962 com:

«Procediments per analitzar dades, tècniques per interpretar els resultats d'aquests procediments o maneres de planejar la recollida de dades per fer l'anàlisi més fàcil, més precís o més exacte.»

Dit d'una altra manera, Tableau permet, per mitjà de simples interaccions, tenir diferents perspectives sobre un conjunt de dades. A més ens ofereix nombroses tècniques d'anàlisi i elements estadístics que podem incorporar a les nostres visualitzacions per entendre millor les dades i extreure'n coneixement a partir de l'ús de representacions gràfiques.

1.2. Per què Tableau?

Gartner realitza anualment un estudi sobre les millors eines de *business intelligence* i analítica del mercat. Des de fa ja uns quants anys, Tableau està posicionada com una de les millors eines en aquest sector, i aquest últim any s'ha situat entre les tres líders, juntament amb Qlik i amb els serveis oferts per Microsoft Cloud (entre els quals destaca especialment Power BI), tal com mostra la figura 1.

Figura 1



Quadre de comparació entre les diferents eines analitzades per Gartner, on l'eix Y representa la capacitat que té un programari de permetre fàcilment la realització d'anàlisis complexos, i l'eix X, la completitud de les eines que cada programari proporciona

A més, Tableau és el fruit de la recerca realitzada per Jock Mackinlay, que va assentar les bases de la teoria que permet a una màquina decidir com representar unes dades de la millor manera possible. Aquest fet és un dels seus aspectes fonamentals, i és que una vegada que se li indica a Tableau el tipus de dades que conté cada columna de la nostra taula, Tableau pot inferir quines són les millors maneres per representar-la, i ajudar l'usuari en el procés de creació de la visualització de dades.

1.3. Pros i contres de Tableau

Com ja s'ha comentat anteriorment, un punt a favor de Tableau és que és una eina fonamentada en una recerca que s'està duent a terme des de ja fa trenta anys, la qual cosa li dona una base teòrica molt potent i justifica la majoria de decisions que s'han pres durant el seu disseny i la seva implementació, i fa que Tableau sigui una eina molt robusta. A continuació, s'enumeren altres avantatges que cal tenir en compte i que també justifiquen que se l'hagi escollit com a eina:

- A causa de la seva constant evolució i qualitat, és una de les eines més sol·licitades al mercat.
- Disposa de molt bona documentació i ofereix tutorials, *webinars* i altres modalitats d'aprenentatge en línia.
- Disposa d'una gran varietat de connectors que permeten treballar amb una immensa diversitat de fonts de dades, que inclou fitxers CSV o XLS, bases de dades, Google Analytics, etc.
- Pel que fa a la seva visualització, tendeix a mantenir al màxim algunes de les bones pràctiques bàsiques de la disciplina, característica que no es pot dir d'eines com Excel (almenys fins a l'última versió).
- Ofereix una versió totalment gratuïta anomenada Tableau Public que permet utilitzar gran part de les funcionalitats de la versió de pagament.

Com tota eina, també té alguns problemes:

- La versió Desktop i les llicències de Tableau Server tenen un preu bastant alt que impedeix l'ús ocasional o *amateur*.
- Alguns connectors a fonts de dades encara no són tan utilitzables com s'hauria d'esperar.
- És una eina cognitivament costosa, és a dir, algunes operacions requereixen una gran concentració i pot ser molt fàcil cometre petits errors que arribin a afectar severament el resultat final, per la qual cosa és necessari tenir molt clars tant l'objectiu de la visualització (el «per què?») com els detalls tècnics de les capacitats de Tableau.

1.4. L'ecosistema de Tableau

Realment Tableau és el nom de l'empresa que desenvolupa el programari. A més, proporciona altres serveis que el complementen, la qual cosa dona lloc a un ampli ventall de productes i opcions.

1.4.1. Tableau Desktop

L'eina principal de Tableau és la seva versió Desktop. És una aplicació nativa a Mac i Windows que ens permet importar dades des d'una gran varietat de fonts de dades. És la versió que tot professional faria servir per al desenvolupament de visualitzacions de dades.

1.4.2. Tableau Public

És la versió gratuïta de Tableau Desktop. Permet utilitzar gran part del potencial del seu germà gran, amb l'excepció que limita el nombre de fonts de dades a les quals un es pot connectar. A més, tot treball realitzat amb Tableau Public (tant la visualització com el *dataset* utilitzat) es guardarà de manera pública als servidors de Tableau i serà compartit entre tots els usuaris.

1.4.3. Tableau Reader

És un lector de visualitzacions fetes a Tableau Desktop. És la manera gratuïta de compartir aquestes visualitzacions en mode de lectura, perquè qualsevol persona explori la nostra visualització sense necessitat de fer-la pública ni que el nostre client hagi de comprar una llicència de Tableau Desktop.

D'alguna manera, Tableau Reader seria una espècie d'Acrobat Reader per als fitxers generats a Tableau Desktop.

1.4.4. Tableau Server

Permet emmagatzemar de manera privada les visualitzacions que vulguem fer confidencials, així com controlar permisos d'usuaris a les visualitzacions i als diferents *datasets*, que podran fer-se «públics» per a tots els membres d'una organització o empresa.

1.5. Conceptes bàsics de Tableau

En aquesta secció comentarem els conceptes bàsics per poder començar a treballar amb Tableau una vegada que hagi estat instal·lat seguint les instruccions sobre aquest tema.

1.5.1. Preparació de les dades

Per començar a treballar de la millor manera possible, és vital seguir els passos següents:

- *Assegurar-nos que el format de les dades és el correcte.* Per poder treballar correctament, Tableau requereix una taula de dades en la qual el nostre ele-

ment d'anàlisi es representi mitjançant files, i les columnes de les quals mostraran els diferents atributs dels quals es disposa per a cada element.

Posem com a exemple el conjunt de dades que es pot veure a la imatge següent. En aquest cas, tenim dades del valor de l'euríbor durant una sèrie d'anys (fins a l'abril de 2016), representats mitjançant les columnes, i durant els mesos d'aquests anys, representats per les files.

Figura 2

# Sheet1 Mes	# Sheet1 2005	# Sheet1 2006	# Sheet1 2007	# Sheet1 2008	# Sheet1 2009	# Sheet1 2010	# Sheet1 2011	# Sheet1 2012	# Sheet1 2013	# Sheet1 2014	# Sheet1 2015	# Sheet1 2016
1	2.312000	2.833000	4.064000	4.498000	2.622000	1.232000	1.550000	1.837000	0.575000	0.562000	0.298000	0.0420000
2	2.310000	2.910000	4.094000	4.349000	2.135000	1.225000	1.714000	1.678000	0.594000	0.549000	0.255000	-0.0080000
3	2.335000	3.100000	4.106000	4.590000	1.909000	1.215000	1.924000	1.499000	0.545000	0.577000	0.212000	-0.0120000
4	2.265000	3.221000	4.253000	4.820000	1.771000	1.225000	2.086000	1.368000	0.528000	0.604000	0.180000	-0.0100000
5	2.286000	3.308000	4.373000	4.994000	1.644000	1.249000	2.147000	1.266000	0.484000	0.592000	0.165000	null
6	2.103000	3.401000	4.505000	5.361000	1.610000	1.281000	2.144000	1.219000	0.507000	0.513000	0.163000	null
7	2.168000	3.538000	4.564000	5.393000	1.412000	1.373000	2.183000	1.061000	0.525000	0.488000	0.167000	null
8	2.223000	3.615000	4.666000	5.323000	1.334000	1.421000	2.097000	0.877000	0.542000	0.469000	0.161000	null
9	2.219000	3.715000	4.725000	5.384000	1.261000	1.420000	2.067000	0.740000	0.543000	0.362000	0.154000	null
10	2.414000	3.799000	4.647000	5.248000	1.243000	1.495000	2.110000	0.650000	0.541000	0.338000	0.128000	null
11	2.684000	3.864000	4.607000	4.350000	1.231000	1.541000	2.044000	0.588000	0.506000	0.335000	0.080000	null
12	2.783000	3.921000	4.793000	3.452000	1.242000	1.526000	2.004000	0.549000	0.543000	0.329000	0.059000	null

Format de dades inadequat per Tableau

Si entenem que el que volem és analitzar l'evolució de l'euríbor durant els anys i els mesos dels quals disposem de dades, aquest *dataset* no estarà en el format adequat, ja que el nostre element principal d'anàlisi, el valor de l'euríbor, no està expressat en files. Així doncs, el que s'hauria de fer és «pivotar» les dades perquè ens quedi un únic valor de l'euríbor per fila, de manera que cada fila tingui tres columnes: una per indicar el valor de l'índex, una altra per indicar el mes i una altra per a l'any, tal com podem veure a la imatge següent:

Figura 3

# Pivot Año	# Pivot Euríbor	# Sheet1 Mes
2006	3.538000	7
2006	3.615000	8
2006	3.715000	9
2006	3.799000	10
2006	3.864000	11
2006	3.921000	12
2007	4.064000	1
2007	4.094000	2
2007	4.106000	3
2007	4.253000	4
2007	4.373000	5
2007	4.505000	6

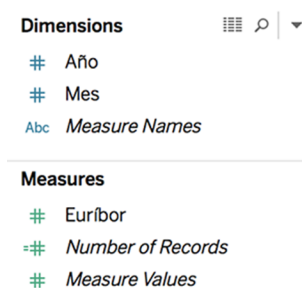
Format adequat de dades per a Tableau; el nostre element d'anàlisi, el valor de l'euríbor, s'expressa per files, i cada columna té informació sobre un atribut de l'element

Per facilitar aquesta transformació del format de les dades, Tableau incorpora una eina que permet «pivotar» les dades, tal com s'explica en aquest article.

És important ser conscient que Tableau no modificarà, en cap cas, la nostra font de dades. D'aquesta manera, totes les transformacions que es realitzin només seran visibles a Tableau, i no a la font d'on provenguin les dades (tant si és un full de càlcul d'Excel com una base de dades, etc.).

- *Identificar el tipus de dades de cada columna.* Tableau, de manera automàtica, intenta identificar el tipus de dades que conté cada columna. No obstant això, per diferents motius, hi ha vegades que no és capaç de fer aquesta identificació correctament, per la qual cosa haurem d'assegurar-nos que cada columna de la nostra taula és del tipus que li correspon. En aquest article podem veure tots els tipus de dades que suporta Tableau. Per exemple, tot i que les dates podrien processar-se com a cadenes, és millor fer servir el tipus específic per poder realitzar operacions entre dates, com calcular el nombre de dies entre dos fets.
- *Renombrar la denominació de les nostres columnes.* Tot i que pot semblar un consell molt banal, aquesta acció és vital per simplificar el treball posterior amb Tableau. Com es veurà a continuació, la nostra taula de dades «desapareixerà de la nostra vista» una vegada que comencem a treballar, i l'única cosa que tindrem a l'abast de la mà seran els noms de les columnes per decidir com representar-les. Per aquesta raó és importantíssim tenir noms de columnes que siguin molt explicatius del seu contingut, tal com veurem a la imatge al final d'aquesta secció, i evitar noms críptics o massa llargs que no en facilitin la manipulació.
- *Identificar si les nostres columnes són mètriques o dimensions.* Les mètriques o mesures són les columnes de les dades que contenen valors numèrics i que, per tant, es poden sumar, és possible calcular la seva mitjana, etc., com per exemple, el nombre d'infectats per malària en un país. D'altra banda, les dimensions són aquelles columnes que contenen informació categòrica que ens permetrà *segmentar* les dades; per exemple, una columna que contingui el nom del país, en la qual veurem el nombre d'infectats per malària.

Figura 4



Dimensions i mètriques del dataset de l'euríbor

Com es pot veure a la imatge anterior, la nostra columna «Euríbor» és l'única mètrica o mesura del conjunt de dades de l'euríbor que hem utilitzat com a exemple, mentre que «Año» i «Mes» són dues dimensions. És important entendre que, tot i que aquestes dues columnes són numèriques, els nombres

que contenen no es poden sumar, multiplicar, etc., ja que no té sentit sumar el mes 1 (gener) i el mes 2 (febrer) per obtenir 3 (març), i el mateix passa amb els anys.

Aquesta divisió entre mesures i dimensions és important, ja que Tableau separa les operacions que es poden realitzar amb cada atribut en funció de la seva naturalesa.

Finalment, la dimensió «Measure Names» és una dimensió creada automàticament que conté el nom de totes les mètriques disponibles, mentre que «Number of Records» i «Measure Values» també són mètriques precalculades pel sistema.

1.5.2. La interfície d'usuari

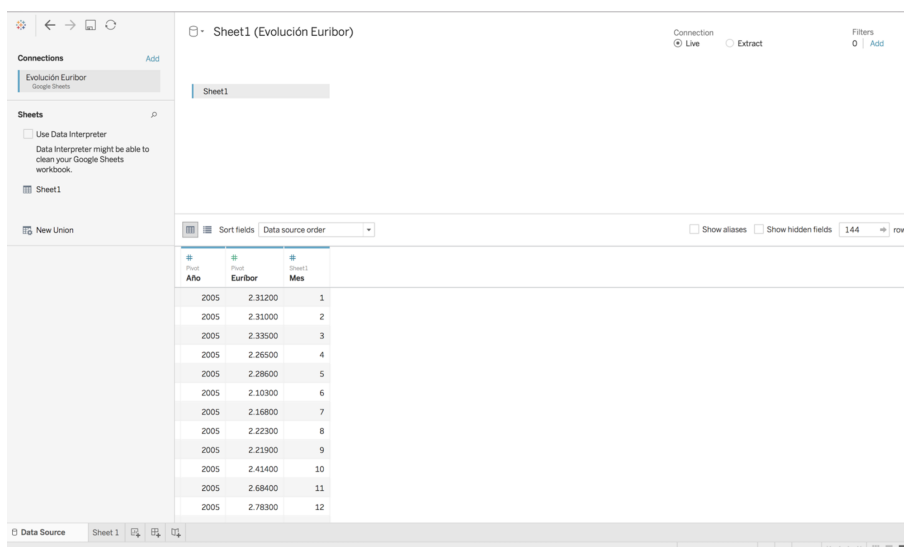
A continuació, es descriuen breument les diferents interfícies d'usuari de Tableau.

La pantalla de font de dades

És la primera pantalla que apareix una vegada que ens hem connectat a les nostres dades. Ens permet triar la pestanya del full de càlcul o taula de la base de dades amb la qual es desitja treballar de la font a la qual s'hagi establert la connexió.

Des d'aquesta pantalla es podran establir unions entre més d'una taula de dades o ocultar columnes que no es desitgin utilitzar en l'anàlisi, simulant les operacions típiques de les bases de dades relacionals. Aquest article aprofundeix més en aquests conceptes.

Figura 5



The screenshot shows the Tableau interface for a data source named 'Evolución Euribor'. The main area displays a table with the following data:

Año	Euribor	Mes
2005	2.31200	1
2005	2.31000	2
2005	2.33500	3
2005	2.26500	4
2005	2.28600	5
2005	2.10300	6
2005	2.16800	7
2005	2.22300	8
2005	2.21900	9
2005	2.41400	10
2005	2.68400	11
2005	2.78300	12

Pantalla de font de dades

La pantalla d'edició de visualitzacions

Com ja s'ha comentat anteriorment, un cop s'ha dut a terme la connexió a les dades amb les quals treballarem, desapareix la vista de taula a la qual estem acostumats amb Excel o altres eines de càlcul de dades. En aquest cas, les nostres dades s'expressen en forma de pastilles que representen cadascuna de les columnes del conjunt de dades.

Figura 6



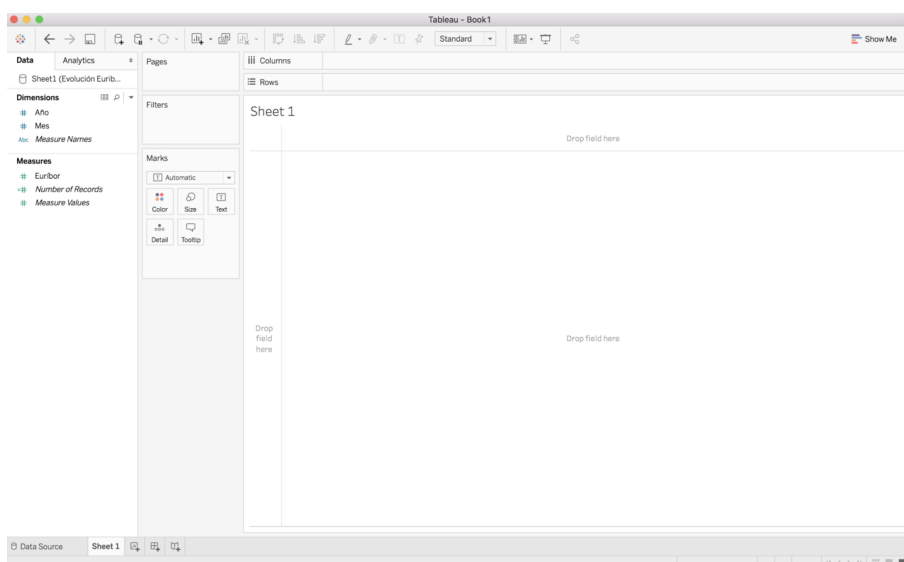
Dues columnes del *dataset* de l'eüríbor representades mitjançant pastilles que es podran arrossegar

Aquestes pastilles tenen dues característiques principals:

- Es poden arrossegar a diferents llocs de la interfície a fi d'utilitzar-les per representar dades o, per exemple, per filtrar-los per algun valor.
- A la dreta, les pastilles sempre contenen una fletxa blanca o selector que permet accedir a opcions que es poden dur a terme amb les dades contingudes en la columna que representa aquesta pastilla.

La pantalla d'edició de visualitzacions permet crear noves representacions visuals. A l'esquerra hi ha dos panells, un per a les mesures i un altre per a les dimensions. Dins d'aquests panells hi haurà les pastilles que representen les columnes de la taula de dades. Per moure una columna de «Mesura» a «Dimensió», n'hi haurà prou a arrossegar la pastilla de la columna en qüestió.

Figura 7



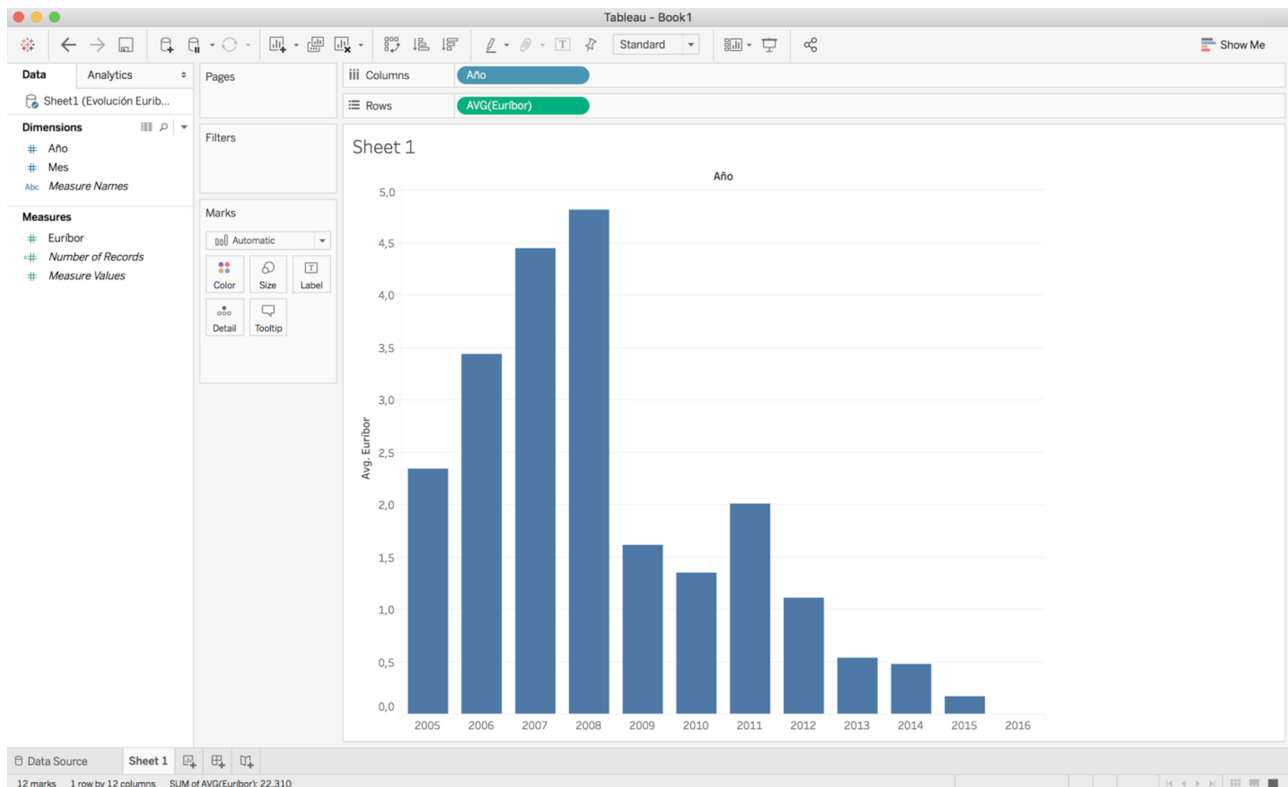
Pantalla d'edició de visualitzacions

Quan es fa doble clic en qualsevol pastilla, tant si és una mesura com una dimensió, Tableau la col·locarà o bé a la secció «Columns» o bé a la secció «Rows», a la part superior de la pantalla. En funció de com s'ordenin les pastilles en aquesta secció, es crearà una visualització o una altra.

A la imatge següent, es pot observar la mitjana del valor de l'euríbor per any, visualitzada perquè s'ha col·locat la pastilla «Euríbor» a la secció «Rows» (utilitzant la funció AVG per al càlcul de mitjanes) i la pastilla «Año» a la secció «Columns».

Com podem veure, Tableau permet fer càlculs amb les columnes numèriques, mentre que les dimensions s'utilitzen per segmentar les dades, i així generar, en aquest exemple, un gràfic de barres.

Figura 8



Si fem doble clic a les pastilles o les arrosseguem a les seccionis «Files» i «Columnes», apareix un gràfic de barres

Entre els panells de dimensions i mesures i la visualització en si, apareixen tres petits panells més:

- *El panell «Pages».* Permet realitzar animacions amb les dades, com s'explica en aquest article. És com afegir una dimensió més lligada a un atribut que pren diferents valors, que es pot fer servir com a eix «temporal» que permeti observar com canvien les altres dades d'acord amb els valors que va prenent.
- *El panell «Filters».* Permet filtrar les dades mostrades seleccionant rangs o bé diferents categories. Aquest article aprofundeix en aquests conceptes.
- *El panell «Marks».* Permet assignar les columnes de les dades als diferents atributs visuals que Tableau suporta. Es pot trobar més informació sobre l'ús d'aquest panell en aquest article.

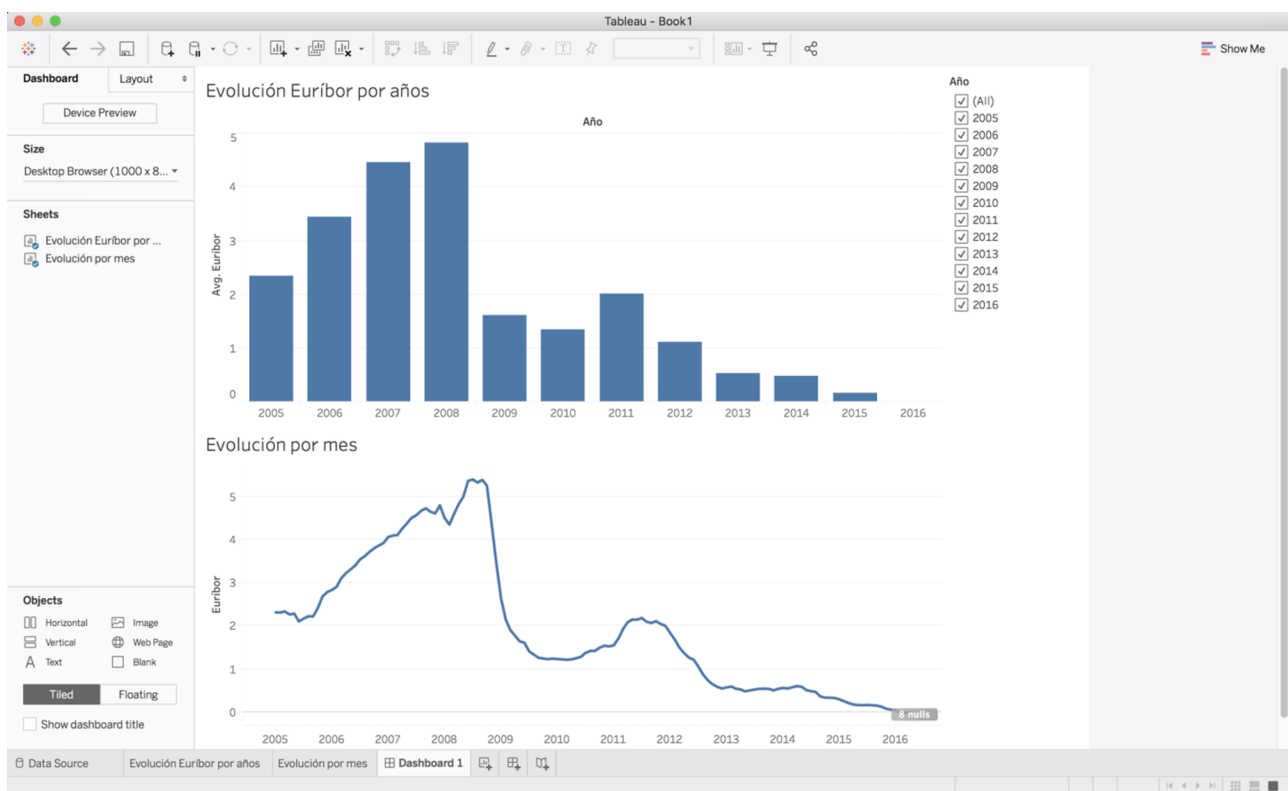
Finalment, a la part superior dreta de la pantalla d'edició de visualitzacions es troba el botó «Mostrar-me», que és el que permet seleccionar el tipus de visualització que es desitja realitzar. En aquest article es pot trobar més informació sobre la funcionalitat d'aquest botó i els diferents tipus de visualitzacions disponibles en funció de la naturalesa de les dades que es volen visualitzar.

La pantalla de creació de *dashboards* o quadres de comandament

A la part inferior de qualsevol pantalla de Tableau apareixen tres icones. Una permet crear una nova visualització, mostrant la pantalla comentada a la secció anterior; una altra permet crear un nou *dashboard*, i finalment, una altra permet crear una nova història.

Fent clic a la creació d'un *dashboard* apareixerà la interfície següent:

Figura 9



Pantalla d'edició de *dashboards*

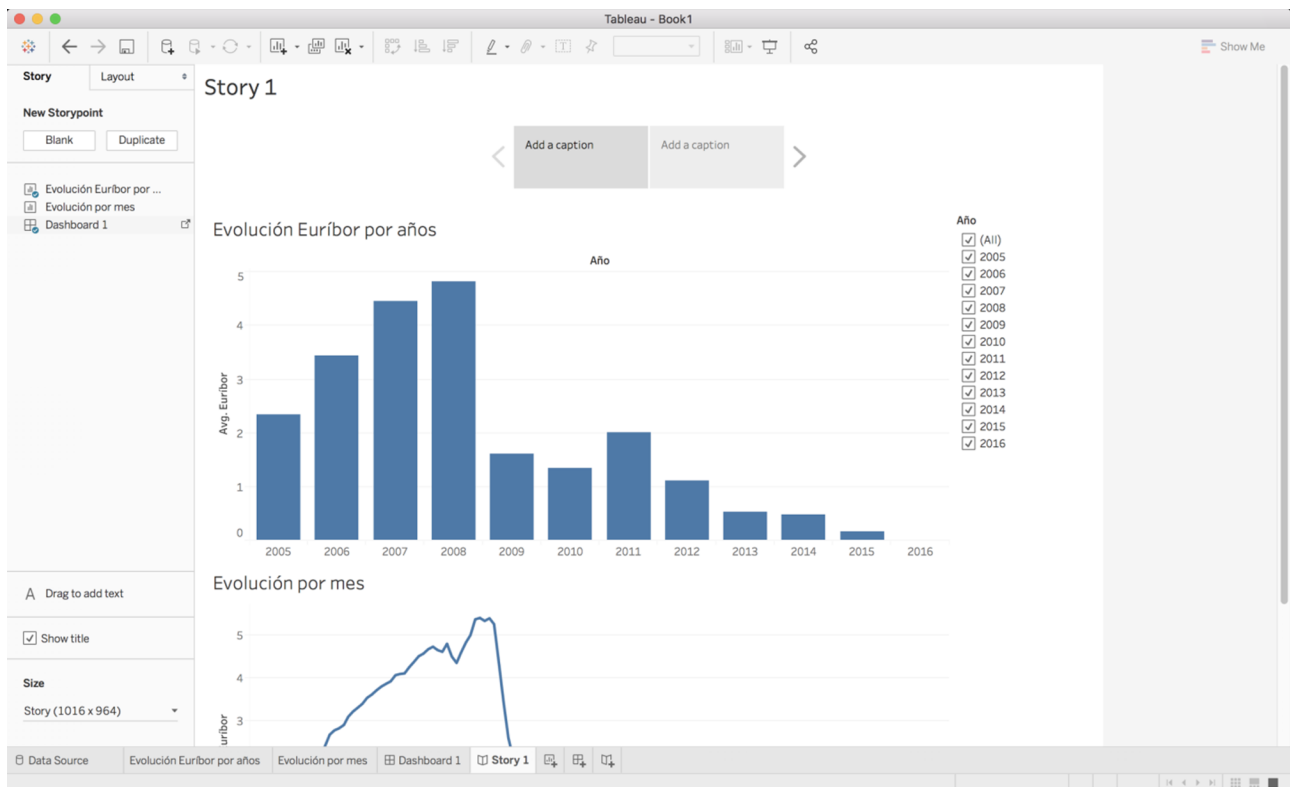
La pantalla d'edició de *dashboards* permet, mitjançant el panell esquerre, arrossegar qualsevol de les visualitzacions ja creades en un espai on es podran sincronitzar i mostrar al mateix temps. En aquest vídeo, així com en els proporcionats en aquest curs, es pot observar com es crea un *dashboard* de manera interactiva.

De fet, un *dashboard* no deixa de ser una agregació de visualitzacions coordinades (o no) entre si, de manera que es presenten una o més visualitzacions associades a diferents conjunts de dades.

La pantalla d'edició d'històries

A més de permetre la creació de visualitzacions i *dashboards*, amb Tableau es pot crear el que es coneix com a *story*. Mentre que un *dashboard* és l'agregació de visualitzacions, una història és l'agregació de visualitzacions i *dashboards*; és a dir, s'afegeix un nivell d'abstracció més. D'aquesta manera es pot crear una història amb les dades, i es mostren tant les visualitzacions com els *dashboards* ja creats de manera seqüencial. Es podrà trobar més informació sobre la creació d'històries en aquest article. Les històries poden narrar-se com una presentació, de manera que es vagin combinant diferents visualitzacions i *dashboards* en una seqüència predeterminada.

Figura 10



Pantalla d'edició d'històries

2. Visualització de dades sobre la malària

Amb l'objectiu de donar una guia més personalitzada a la utilització de Tableau, i com un acompanyament dels videotutorials gratuïts, hem desenvolupat unes instruccions que mostren els passos necessaris per visualitzar dades relacionades amb l'evolució dels infectats pel paràsit de la malària al món.

L'objectiu d'aquest *dashboard* és veure, a partir de la coordinació entre tres visualitzacions diferents, les dades del nombre d'infectats pel paràsit de la malària al món. El resultat final permet contestar dues preguntes: «On hi ha més infectats per malària al món?» i «quin ha estat l'evolució d'aquesta malaltia, a nivell mundial i per països?». Aquestes qüestions responen al «per què?», és a dir, a les raons i objectius pels quals es crea aquesta visualització de dades.

El procés de creació d'aquesta visualització està dividit en cinc vídeos diferents, que cobreixen des de la càrrega de dades a Tableau fins a la creació d'un *dashboard* que conté tots els elements que componen la visualització.

2.1. Càrrega de dades a Tableau

En aquest vídeo es mostra com es carreguen dades de Google Sheets a Tableau. En una primera part, el vídeo explica que cal descarregar Tableau Public, tot i que podeu utilitzar la versió comercial, Tableau Desktop, usant les instruccions que es troben en el document «Ús de Tableau a la UOC», que trobareu a l'aula de l'assignatura.

A més, el vídeo mostra com s'han de preparar les dades disponibles perquè, prèviament a la seva utilització a Tableau, estiguin en un format adequat per mitjà de la pantalla de font de dades, a partir d'un document Excel preparat expressament.

Figura 11

Country	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16
Afghanistan	32	24	36	40	22	32	46	25	null	0	null	null	null	null	null
Algeria	0	3	0	1	5	1	0	null	null	null	null	null	null	1	2
Angola	5,714	7,300	5,736	6,909	8,114	10,530	9,465	9,812	10,220	13,768	12,459	38,598	14,434	9,473	9,510
Argentina	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
Armenia	null	null	null	null	0	0	0	0	0	0	0	0	0	0	0
Azerbaijan	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bahamas	null	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bangladesh	45	15	11	36	37	47	154	228	508	501	505	574	598	470	484
Belize	0	0	0	0	0	0	0	0	1	0	1	null	0	0	0
Benin	1,869	2,288	2,261	1,753	964	1,375	918	1,290	1,226	322	944	560	707	468	null
Bhutan	0	0	1	1	2	4	2	2	7	5	7	14	11	14	15
Bolivia	1	0	0	0	0	0	0	0	0	0	3	1	4	0	11
Botswana	22	7	3	8	8	6	12	6	40	11	19	18	23	29	null
Brazil	36	41	60	70	76	85	68	93	110	123	102	104	95	142	245
Burkina Faso	5,632	6,294	7,963	7,001	9,024	7,982	7,834	6,472	8,083	5,224	4,205	4,860	4,032	4,233	null

Pantalla de càrrega de dades a Tableau en la qual es manipularan les dades per adequar-les a les necessitats de l'eina

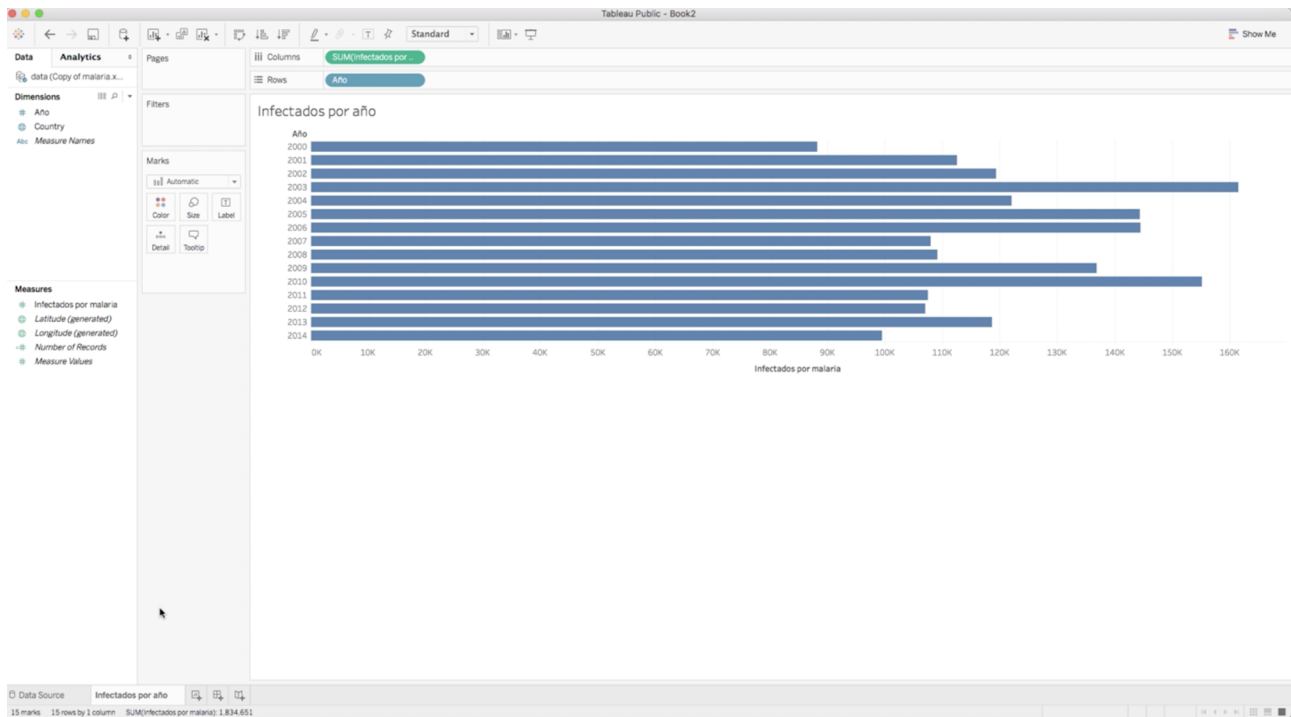
2.2. Creació d'un gràfic de barres

En el segon vídeo s'introdueix la pantalla d'edició de visualitzacions i es detalla com es duu a terme un gràfic de barres.

Especialment, es remarquen els dos conceptes que cal tenir clars per treballar amb Tableau, ja esmentats amb anterioritat: les mètriques, que són aquelles columnes de la taula de dades que contenen dades que poden manipular-se numèricament, i les dimensions, que són aquelles columnes que contenen informació categòrica que permetrà segmentar les dades.

Per fer el gràfic de barres, es podrà observar que n'hi ha prou a seleccionar la mètrica «Infectats», que Tableau agrega automàticament utilitzant la funció SUM. Mitjançant la dimensió «Any», Tableau segmentarà la mètrica per crear una barra per a cada any. El gràfic de barres resultant serveix per poder veure l'evolució temporal dels infectats per malària.

Figura 12



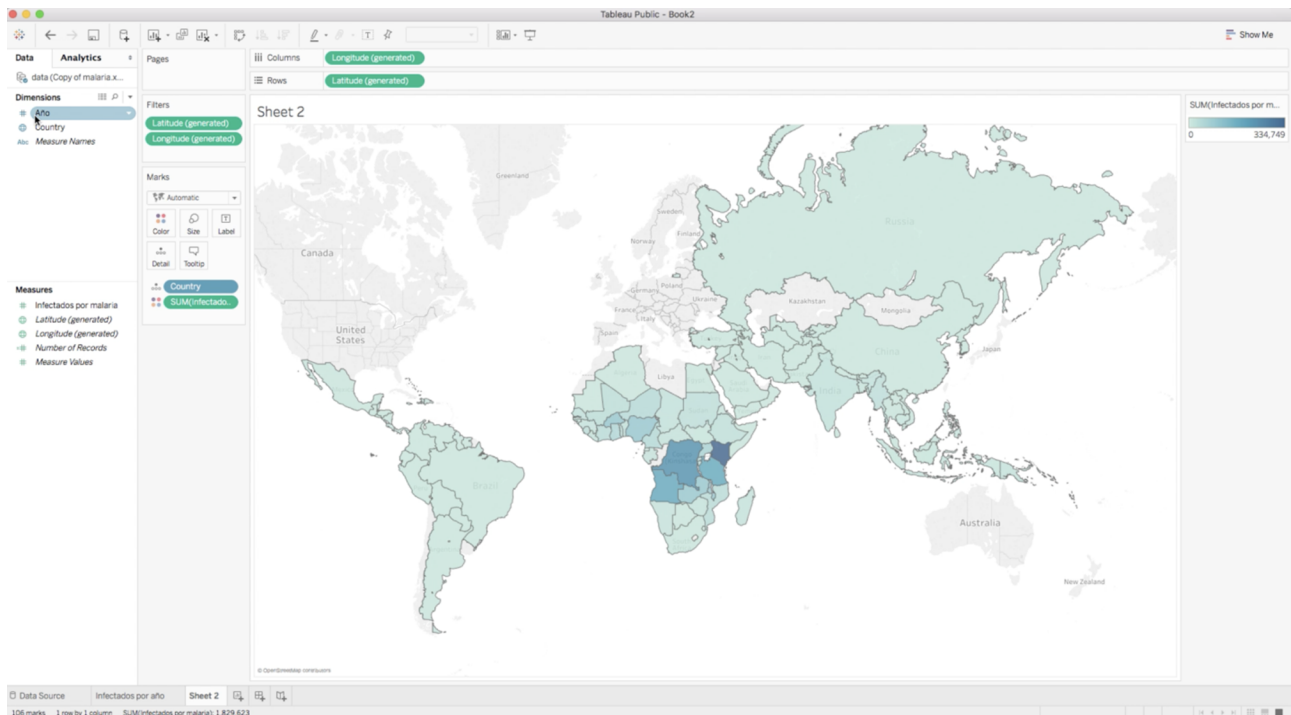
Gràfic de barres horitzontal que mostra l'evolució d'infectats per any per malària

2.3. Creació d'un mapa

En el tercer vídeo es veurà com es crea un mapa en el qual el color representa el nombre d'infectats en una regió. Per a això, es veurà tant l'opció de posar punts damunt del mapa com la utilització de *choropleth maps* o mapes acolorits. El mapa permetrà fer l'anàlisi geogràfica dels infectats per malària, i entendre en quines zones del món hi ha més afectats per aquest paràsit.

El vídeo mostra com n'és de senzill crear mapes amb Tableau, ja que l'eina és capaç de gestionar informació semàntica relativa a noms geogràfics, com poden ser països, ciutats i fins i tot estats o comunitats autònomes.

Figura 13



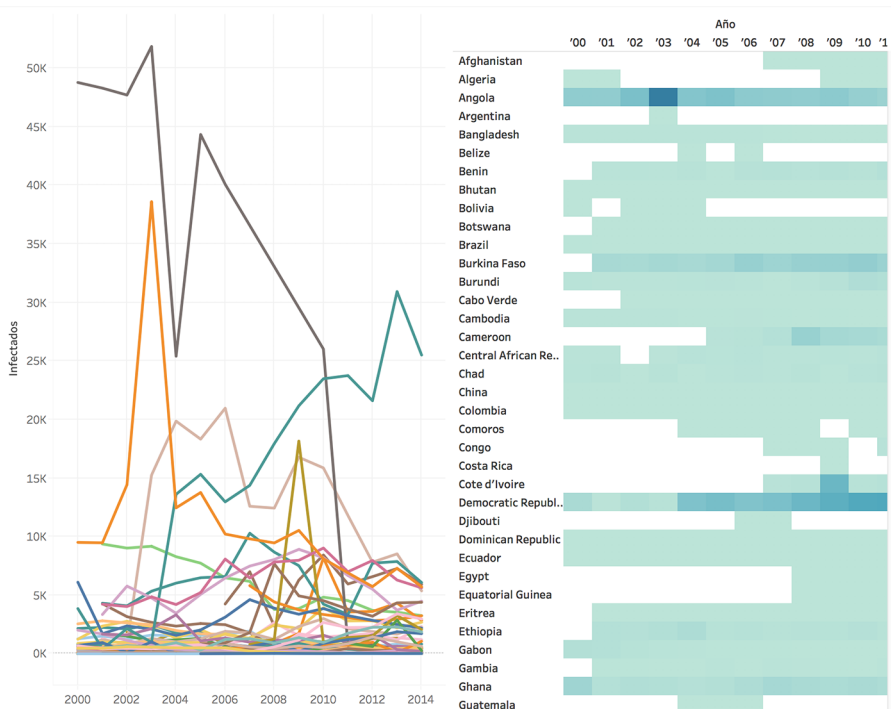
Mapa del món que representa el nombre d'infectats

2.4. Creació d'un *heatmap*

Al quart vídeo es presenta com es desenvolupa un *heatmap*, que és un tipus de visualització molt utilitzada que consisteix a acolorir una taula de dades per destacar els valors en un rang predeterminat. En aquest cas, el color de cada cel·la tindrà una intensitat o una altra en funció del valor d'infectats per malària de cada any.

És molt important entendre que, tot i que la millor manera de veure l'evolució de la malària seria utilitzar un gràfic de línies, crear un gràfic en el qual cada país tingués una línia seria molt complex visualment, a causa del gran nombre de països que es pretén visualitzar. En canvi, el *heatmap* permet utilitzar una fila de la taula de dades resultant per a cada país. Així s'eviten solapaments i s'aconsegueix una visualització més fàcil d'explorar i d'entendre.

Figura 14



Gràfic de línies i heatmap del nombre d'infectats per malària: el gràfic de línies és confús a causa del gran nombre de països; en canvi, el heatmap ens permet veure més fàcilment les dades de cada país

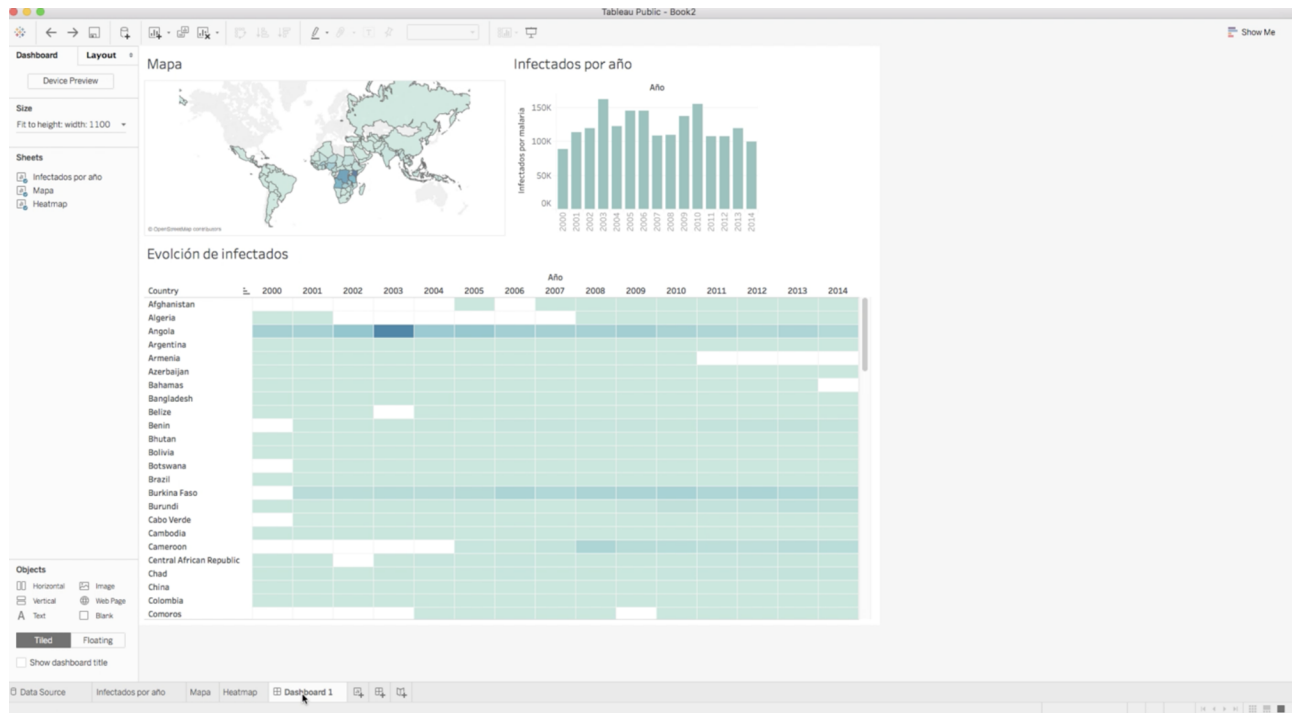
A més, el vídeo presenta breument la utilització de fórmules a Tableau, que són molt semblants a les fórmules que es poden fer en Excel.

2.5. Creació d'un *dashboard*

Finalment, el cinquè vídeo explica com es poden combinar les tres visualitzacions ja generades en un *dashboard* o quadre de comandament, utilitzant accions per crear filtres entre visualitzacions. És important entendre que el *dashboard* és una agrupació de visualitzacions de les mateixes dades, la qual cosa permet tenir a simple vista diferents perspectives del problema que s'està tractant.

També es mostra breument com es pot embellir una visualització modificant el format per defecte. Finalment, s'acaba publicant el *dashboard* final a Tableau Public.

Figura 15



Dashboard final que integra les tres visualitzacions creades

Bibliografia

Mackinlay, J.; Hanrahan, P.; Stolte, C. (2007, núm. 6). «Show me: Automatic presentation for visual analysis» [article en línia]. A: *IEEE transactions on visualization and computer graphics* (vol. XIII). [Data de consulta: 22 de maig de 2017]. <<http://ieeexplore.ieee.org/abstract/document/4376133/>>.

Make Over Monday (2017). *Retos para mejorar visualizaciones utilizando Tableau* [en línia]. [Data de consulta: 22 de maig de 2017]. <<http://www.makeovermonday.co.uk/data/>>.

Make Over Monday (2017). *Dataset utilizado para la visualización de la malaria, accesible desde Google Sheets* [en línia]. [Data de consulta: 22 de maig de 2017]. <http://bit.ly/malaria_data>.

Tableau (2017). *Web de Tableau* [en línia]. [Data de consulta: 22 de maig de 2017]. <<https://www.tableau.com/es-es>>.

Tableau (2017). *Vídeos de capacitación gratuitos* [en línia]. [Data de consulta: 22 de maig de 2017]. <<https://www.tableau.com/es-es/learn/training>>.

Tableau (2017). *Galería pública de ejemplos* [en línia]. [Data de consulta: 22 de maig de 2017]. <<https://public.tableau.com/es-es/s/gallery>>.

Tukey, J. W. (1962, núm. 1). «The future of data analysis» [article en línia]. A: *The Annals of Mathematical Statistics* (vol. XXXIII, pàg. 1-67). [Data de consulta: 22 de maig de 2017]. <http://projecteuclid.org/download/pdf_1/euclid.aoms/1177704711>.

