

Explotación y Visualización

Práctica Tableau

**José Manuel Bustos Muñoz
Santiago Martinez De la Riva**

1. Ejercicio 1 de Uniones y mezclas.

Para este ejercicio se cargan los datos de un fichero excel que contiene los datos de dos aerolíneas a lo largo de dos años y distribuidos por región o continente, y así saber cada aerolínea lo que ha facturado en cada región en los años de 2015 y 2016.

Al cargar los datos los visualizamos para un primer vistazo a los mismos y ver que se han cargado correctamente y conocer los datos cargados.

Se carga el excel dos veces, en cada una arrastramos a la vista de datos una de las tablas para hacer la mezcla de datos.

Conexiones

Añadir

Airline-Comparison

Microsoft Excel

Hojas

Airline1

Airline2

Nueva unión

Airline-Comparison

Airline1

Ordenar campos

Orden de fuente de datos

Abc Airline1 Period	Abc Airline1 Region	# Airline1 Revenue
FY2015	Africa	10.000.000
FY2015	Asia	20.000.000
FY2015	Australia	70.000.000
FY2015	Europe	120.000.000
FY2015	North America	150.000.000
FY2015	South America	60.000.000
FY2016	Africa	8.000.000
FY2016	Asia	16.000.000
FY2016	Australia	56.000.000
FY2016	Europe	96.000.000
FY2016	North America	120.000.000
FY2016	South America	48.000.000

Fuente de datos

Datos aerolinas por año

En la hoja de trabajo una vez cargados todos los datos y mezclados ya nos aparecen las dimensiones y medidas de ambas tablas.

Se puede observar como se ha realizado la unión por los atributos “period” y “region” de cada tabla. Como en un principio “region” no venía en ambas tablas con el mismo nombre, para que hiciera la mezcla por ese atributo hubo que cambiarle en una el nombre para que fuera como en la otra tabla.

Datos	Análisis
<div><div><div>Airline-Comparison</div><div>Airline2 (Airline-Comparis...</div></div></div>	
<div><div>Dimensiones</div><div><div>Period</div><div>Region</div><div>Nombres de medidas</div></div></div>	
<div><div>Medidas</div><div><div>Revenue</div><div>Número de registros</div><div>Valores de medidas</div></div></div>	

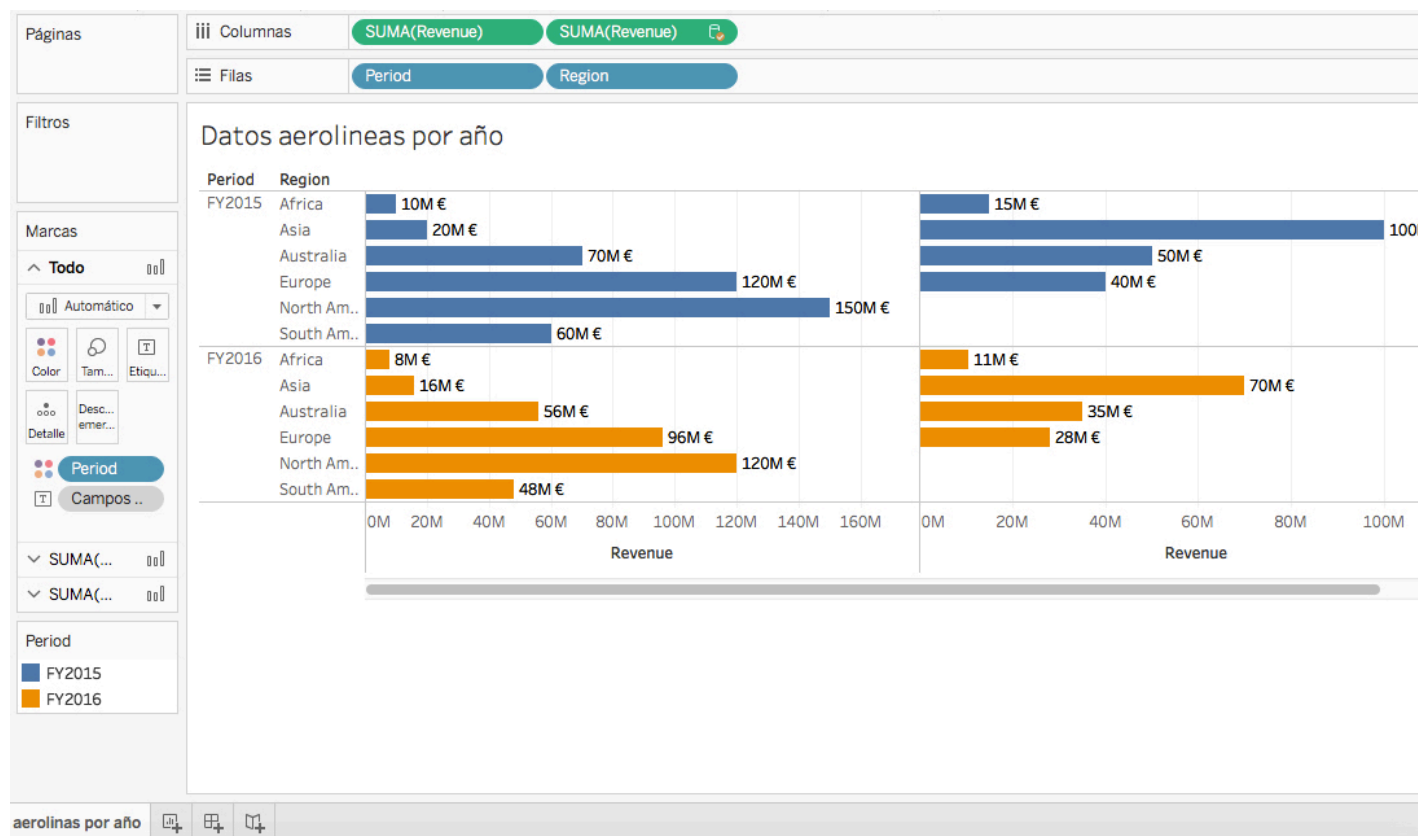
Datos	Análisis
<div><div><div>Airline-Comparison</div><div>Airline2 (Airline-Comparis...</div></div></div>	
<div><div>Dimensiones</div><div><div>Period</div><div>Region</div><div>Nombres de medidas</div></div></div>	
<div><div>Medidas</div><div><div>Revenue</div><div>Número de registros</div><div>Valores de medidas</div></div></div>	

Una vez nos ponemos con el gráfico para realizar un análisis visual de estos datos, lo que hacemos es arrastrar primero los datos de la tabla 1. A las filas arrastramos la dimensión “Region” y a las columnas arrastramos la medida “Revenue”. Así primeramente visualizamos para la tabla 1 los datos de ventas por cada región.

Hacemos lo mismo para la tabla 2, y se añade otra gráfica con los mismos datos pero esta vez para la otra tabla.

Para visualizar los datos distribuidos por año, se añade la dimensión “Period” al eje de las filas, y así la gráfica se divide por año además de por región, y se obtiene la gráfica como se ve en la imagen final partida por año y región, además de por tabla.

Otros formatos que se han cambiado ha sido para visualizar las ventas en millones y en euros. También añadimos color para visualizar en color distinto por año por ejemplo.



Puede observarse como hubo un retroceso general en las ventas para las dos compañías en el año 2016 respecto al 2015, además la aerolínea 1 tiene mayor suma de ventas por norma general que la aerolínea 2, excepto en Asia. Esto nos lleva a pensar que la segunda aerolínea es de procedencia asiática y por eso centra en este continente sus vuelos y por ejemplo parece no operar en todo el continente americano.






La primera compañía opera en todo el mundo, y sus mayores ventas están en Norte América, por lo que es lógico que se trate de una compañía estadounidense o canadiense.

2. Ejercicio 2 de Uniones y mezclas.

En el segundo ejercicio cargamos los datos de otro fichero, que representan distintas ventas de compañías, como ordenes de compras.


Hay tres tablas diferentes: una con la lista de ordenes de cada venta con datos del cliente, otra las ordenes con los datos del producto, y otra tabla que contiene las categorías y los objetivos de venta.

Se realiza una unión entre las ordenes de venta con los datos de cliente y las ordenes de venta por producto, y así tendríamos unida toda la información de las ventas.




Conexiones [Añadir](#)

AmazingMartEU3
Microsoft Excel



Hojas 

- ListOfOrders
- OrderBreakdown
- SalesTargets
- Nueva unión

ListofOrders+OrderBreakdown



ListOfOrdersOrderBreakdown

 Ordenar campos

Orden de fuente de datos

Abc ListOfOrders Order ID	Calendar ListOfOrders Order Date	Abc ListOfOrders Customer Name	World ListOfOrders City	World ListOff Cour
BN-2011-7407039	1/1/2011	Ruby Patel	Stockholm	Swe
AZ-2011-9050313	3/1/2011	Summer Hayward	Southport	Unit
AZ-2011-6674300	4/1/2011	Devin Huddleston	Valence	Fran
BN-2011-2819714	4/1/2011	Mary Parker	Birmingham	Unit
BN-2011-2819714	4/1/2011	Mary Parker	Birmingham	Unit
AZ-2011-617423	5/1/2011	Daniel Burke	Echirolles	Fran
AZ-2011-617423	5/1/2011	Daniel Burke	Echirolles	Fran
AZ-2011-2918397	7/1/2011	Fredrick Beveridge	La Seyne-sur-Mer	Fran
...

Vistazo a los datos de la tabla “ListOfOrders”, con los datos de cliente de cada venta.

Hojas

- ListOfOrders
- OrderBreakdown
- SalesTargets
- Nueva unión

Ver datos: ListOfOrders

4.117 filas

City	Country	Customer Name	Order Date	Order ID	Region	Segment	Ship Date	Ship Mode
Stockholm	Sweden	Ruby Patel	1/1/2011	BN-2011-7407039	North	Home Office	5/1/2011	Economy
Southport	United Kingdom	Summer Hayward	3/1/2011	AZ-2011-9050313	North	Consumer	7/1/2011	Economy
Valence	France	Devin Huddleston	4/1/2011	AZ-2011-6674300	Central	Consumer	8/1/2011	Economy
Birmingham	United Kingdom	Mary Parker	4/1/2011	BN-2011-2819714	North	Corporate	9/1/2011	Economy
Echirolles	France	Daniel Burke	5/1/2011	AZ-2011-617423	Central	Home Office	7/1/2011	Priority
La Seyne-sur-Mer	France	Fredrick Beveridge	7/1/2011	AZ-2011-2918397	Central	Corporate	8/1/2011	Priority
Toulouse	France	Archer Hort	8/1/2011	BN-2011-3248724	Central	Consumer	14/1/2011	Economy
Genoa	Italy	Evie Flockhart	11/1/2011	AZ-2011-6712797	South	Consumer	16/1/2011	Economy
Vienna	Austria	Faith Greenwood	11/1/2011	AZ-2011-4827146	Central	Consumer	15/1/2011	Economy
Murcia	Spain	Summer Hayward	11/1/2011	AZ-2011-6439906	South	Consumer	15/1/2011	Economy
Woking	United Kingdom	Gracie Powell	11/1/2011	AZ-2011-7053593	North	Consumer	11/1/2011	Immediat
Lohne	Germany	Hershel Snyder	12/1/2011	AZ-2011-5702370	Central	Corporate	19/1/2011	Economy
Leicester	United Kingdom	Julia Martell	12/1/2011	AZ-2011-9927716	North	Home Office	17/1/2011	Economy
Sheffield	United Kingdom	Viola Watson	12/1/2011	AZ-2011-2222024	North	Consumer	15/1/2011	Priority
Dordrecht	Netherlands	Julian Dobie	13/1/2011	BN-2011-4913858	Central	Consumer	19/1/2011	Economy
Gothenburg	Sweden	Rose Heap	13/1/2011	BN-2011-2807470	North	Consumer	20/1/2011	Economy
Vienna	Austria	Ella Troy	14/1/2011	AZ-2011-5960662	Central	Home Office	19/1/2011	Economy
Langen	Germany	Everett Dunbar	15/1/2011	AZ-2011-7675351	Central	Corporate	20/1/2011	Economy

Fuente de datos Ventas y diferencia a objetivo

Vistazo a la tabla de “OrderBreakdown” con los datos de las ventas pero con detalle del producto.

Hojas

- ListOfOrders
- OrderBreakdown
- SalesTargets
- Nueva unión

Ver datos: OrderBreakdown

8.047 filas

Category	Order ID	Product Name	Sub-Category	Discount	Profit	Quantity	Sales
Office Supplies	BN-2011-7407039	Enermax Note Cards, Premium	Paper	0,500000	-26	3	45
Furniture	AZ-2011-9050313	Dania Corner Shelving, Traditional	Bookcases	0,000000	290	7	854
Office Supplies	AZ-2011-6674300	Binney & Smith Sketch Pad, Easy-Erase	Art	0,000000	21	3	140
Office Supplies	BN-2011-2819714	Boston Markers, Easy-Erase	Art	0,500000	-22	2	27
Office Supplies	BN-2011-2819714	Eldon Folders, Single Width	Storage	0,500000	-1	2	17
Office Supplies	AZ-2011-617423	Binney & Smith Pencil Sharpener, Water Color	Art	0,000000	21	3	90
Office Supplies	AZ-2011-617423	Sanford Canvas, Fluorescent	Art	0,000000	77	4	207
Furniture	AZ-2011-2918397	Bush Floating Shelf Set, Pine	Bookcases	0,100000	36	1	155
Office Supplies	AZ-2011-2918397	Accos Thumb Tacks, Assorted Sizes	Fasteners	0,000000	2	3	33
Office Supplies	AZ-2011-2918397	Smead Lockers, Industrial	Storage	0,100000	143	4	716
Furniture	BN-2011-3248724	Ikea Classic Bookcase, Metal	Bookcases	0,600000	-1.012	6	987
Office Supplies	BN-2011-3248724	Binney & Smith Sketch Pad, Blue	Art	0,500000	-56	5	116
Furniture	AZ-2011-7053593	SAFCO Executive Leather Armchair, Red	Chairs	0,000000	14	3	1.384
Office Supplies	AZ-2011-7053593	Binney & Smith Canvas, Blue	Art	0,000000	20	2	103
Furniture	AZ-2011-6439906	Bevis Training Table, with Bottom Storage	Tables	0,600000	-342	2	268
Office Supplies	AZ-2011-4827146	Boston Canvas, Fluorescent	Art	0,000000	10	1	55
Office Supplies	AZ-2011-4827146	Smead Trays, Single Width	Storage	0,000000	31	2	97
Office Supplies	AZ-2011-6439906	Novimex File Folder Labels, Alphabetical	Labels	0,000000	6	5	40
Office Supplies	AZ-2011-6712797	Ibico Hole Reinforcements, Recycled	Binders	0,000000	7	3	22
Office Supplies	AZ-2011-2222024	Green Bar Note Cards, Multicolor	Paper	0,500000	-6	2	34
Furniture	AZ-2011-9927716	Hon Chairmat, Adjustable	Chairs	0,000000	70	5	290
Furniture	AZ-2011-5702370	Ikea Stackable Bookrack, Traditional	Bookcases	0,100000	165	5	552
Office Supplies	AZ-2011-5702370	Binney & Smith Canvas, Blue	Art	0,000000	49	5	257

Fuente de datos Ventas y

Vistazo a la tabla de “SalesTargets” con la información de ventas objetivo por cada categoría y periodo.

Category	Month of Order Date	Target
Furniture	1/1/2011	10.000
Furniture	1/2/2011	10.100
Furniture	1/3/2011	10.300
Furniture	1/4/2011	10.400
Furniture	1/5/2011	10.500
Furniture	1/6/2011	10.600
Furniture	1/7/2011	10.800
Furniture	1/8/2011	10.900
Furniture	1/9/2011	11.000
Furniture	1/10/2011	11.100
Furniture	1/11/2011	11.300
Furniture	1/12/2011	11.400
Furniture	1/1/2012	11.500
Furniture	1/2/2012	11.600
Furniture	1/3/2012	11.800
Furniture	1/4/2012	11.900
Furniture	1/5/2012	12.000
Furniture	1/6/2012	12.100
Furniture	1/7/2012	12.300
Furniture	1/8/2012	12.400
Furniture	1/9/2012	12.500
Furniture	1/10/2012	12.600
Furniture	1/11/2012	12.800
Furniture	1/12/2012	12.900

En la hoja de trabajo para visualizar y analizar los datos tendríamos la unión realizada entre las ventas por cliente y las ventas por producto, y además se hace una mezcla con los datos objetivo por cada categoría. Se hace mezcla porque los datos de las tablas no están con la misma granularidad.

Primero vamos a visualizar los datos de la unión de las ventas. Al eje de las columnas arrastramos la dimensión “OrderDate” y se selecciona en la misma que se agrupe o muestre por mes. Y en el eje de las filas se arrastra la medida “Sales”. Seleccionando además la dimensión “Category” que representa a las 3 categorías en las que se engloban los datos, al color para que se puedan ver los datos en 3 colores distintos según la categoría, y creando un filtro conseguimos parte del gráfico final requerido.

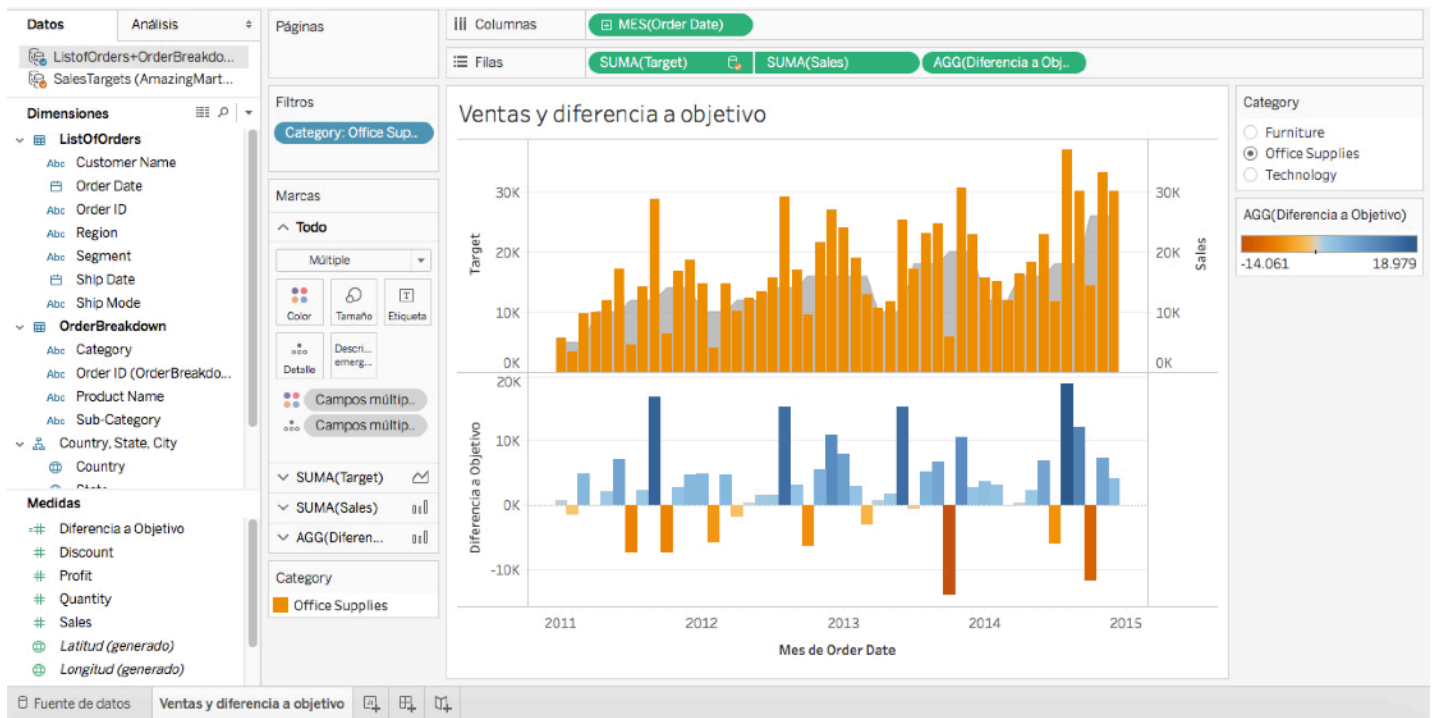
El filtro se configura y tendremos su leyenda donde podrá seleccionarse una categoría y se visualizarán sus datos solamente.

Para la mejor visualización se utiliza una gráfica de barras.

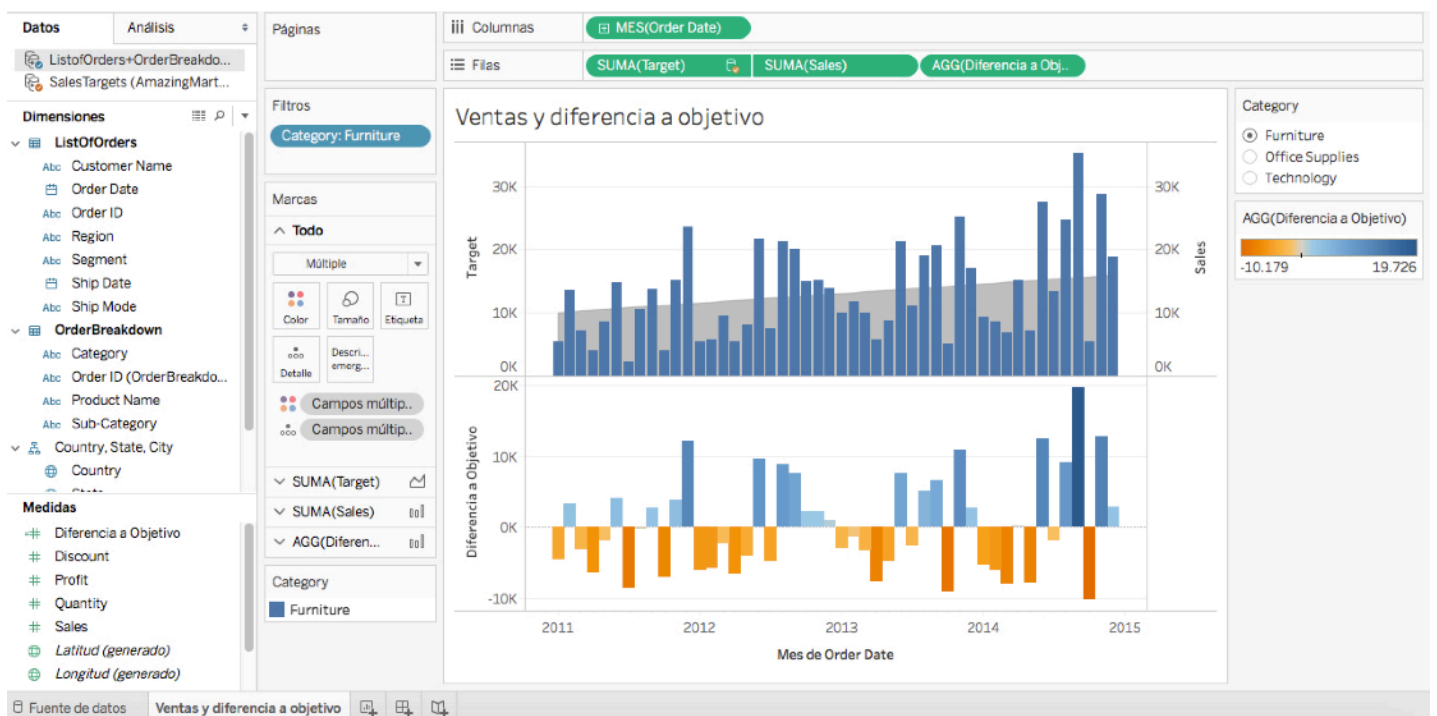
Para añadir granularidad vamos a introducir datos de la tabla de ventas objetivo de cada categoría. Arrastramos la medida “Target” al eje de las filas, y formateándolo de la forma adecuada ya que no está en la misma escala que el resto de datos y creando las relaciones pertinentes, añadiremos el detalle del objetivo a la gráfica que ya teníamos con las ventas mensuales por cada categoría.

Para visualizarlo como se puede ver en las siguientes imágenes, se va a crear una variable calculada que represente la diferencia entre las ventas y las ventas objetivo, y arrastramos esta nueva variable calculada a partir de otras existentes al eje de las filas para que este detalle se introduzca en la gráfica.

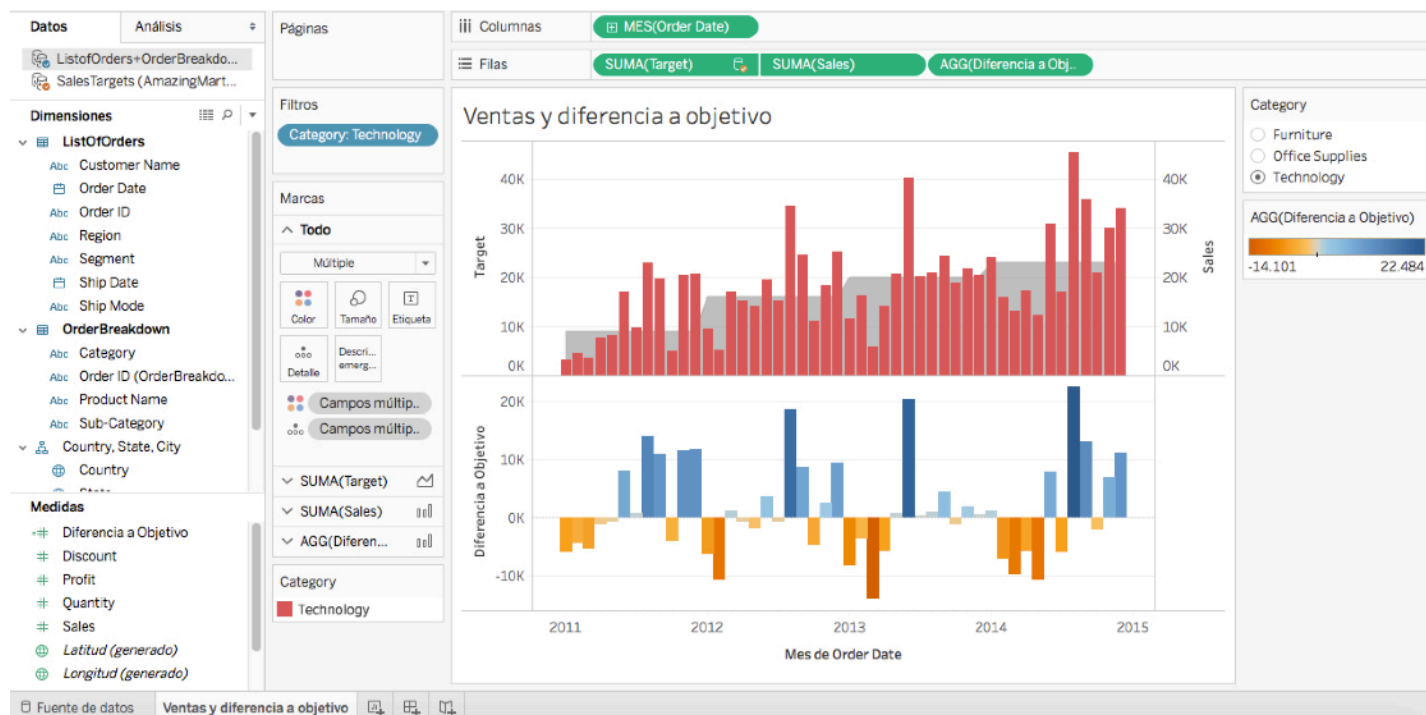
Visualización de los datos de ventas y diferencia a objetivo de la categoría Office Supplies.



Visualización de los datos de ventas y diferencia a objetivo de la categoría Furniture.



Visualización de los datos de ventas y diferencia a objetivo de la categoría Technology.



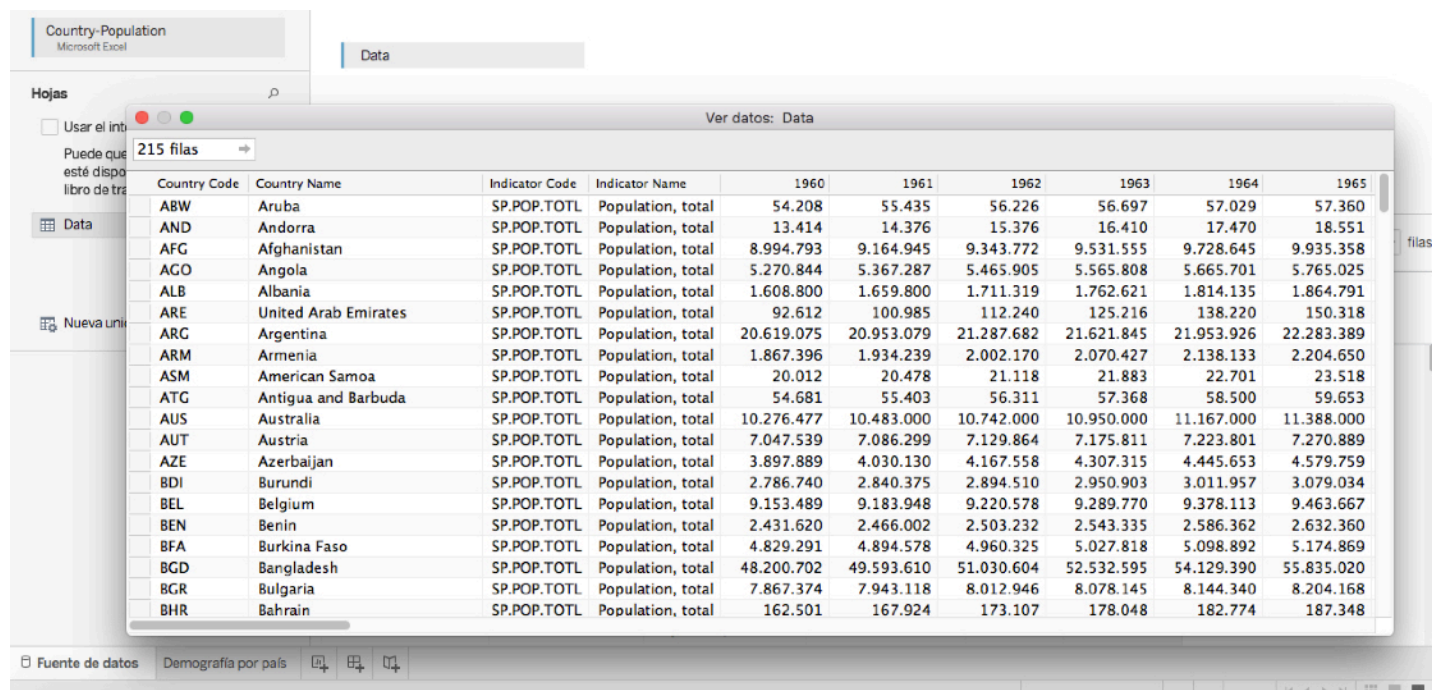
Con esta gráfica podría hacerse un análisis global por ejemplo para ver los meses en los que las tres categorías parecen bajar o subir las ventas y llegar o no al objetivo, para poder tomar medidas globales. Gracias al filtro por categoría, este análisis también puede hacerse por categoría, si se aprecian diferencias que nos deban llevar a tomar medidas distintas según la categoría, aunque sea la misma época del año.

Además vemos la distancia global de cada categoría hasta su objetivo y podemos hacer foco en una si hace más falta que en otras. Aunque como vemos hay alguna con meses más acusados en su diferencia con el objetivo, pero al mismo tiempo vemos otra categoría donde globalmente se llegan menos meses al objetivo pero cuando no se llega se hace de forma menos acusada. Habrá que realizar un estudio a fondo de estos datos y balancear esto.

3. Ejercicio 3 de Evolución demográfica de países.

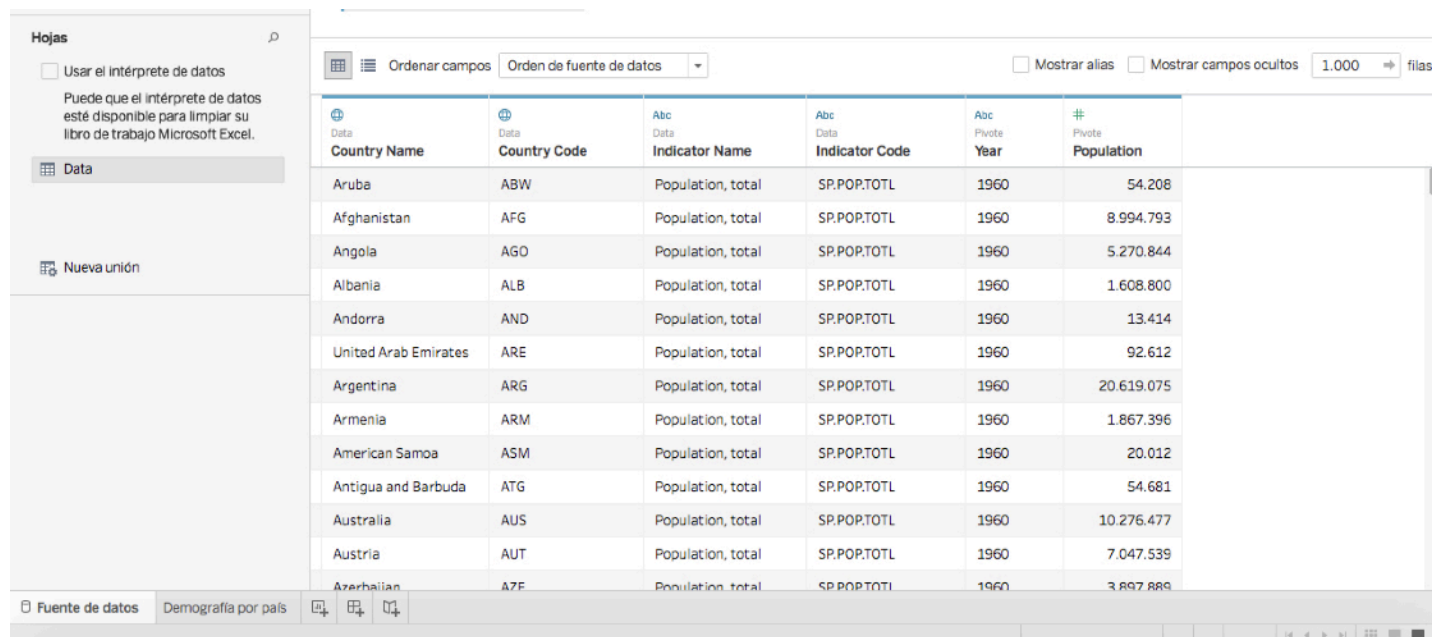
Para el tercer ejercicio se cargan unos datos que contienen información de los distintos países del mundo, en cuestión de población, esperanza de vida y fertilidad o número de hijos por familia. Todos estos datos representados a lo largo de un periodo de tiempo, de año en año, desde 1960 hasta la actualidad. Esto nos va a permitir visualizar y analizar la evolución de cada país a lo largo de los últimos 50-60 años en cuanto al aumento o no de su población, esperanza de vida y fertilidad.

Cargamos los datos de población con su fichero excel y damos un vistazo viendo que efectivamente se tienen los registros para cada país y todos los años del periodo como columnas con la población por año de ese país.



Country Code	Country Name	Indicator Code	Indicator Name	1960	1961	1962	1963	1964	1965
ABW	Aruba	SP.POP.TOTL	Population, total	54.208	55.435	56.226	56.697	57.029	57.360
AND	Andorra	SP.POP.TOTL	Population, total	13.414	14.376	15.376	16.410	17.470	18.551
AFG	Afghanistan	SP.POP.TOTL	Population, total	8.994.793	9.164.945	9.343.772	9.531.555	9.728.645	9.935.358
AGO	Angola	SP.POP.TOTL	Population, total	5.270.844	5.367.287	5.465.905	5.565.808	5.665.701	5.765.025
ALB	Albania	SP.POP.TOTL	Population, total	1.608.800	1.659.800	1.711.319	1.762.621	1.814.135	1.864.791
ARE	United Arab Emirates	SP.POP.TOTL	Population, total	92.612	100.985	112.240	125.216	138.220	150.318
ARG	Argentina	SP.POP.TOTL	Population, total	20.619.075	20.953.079	21.287.682	21.621.845	21.953.926	22.283.389
ARM	Armenia	SP.POP.TOTL	Population, total	1.867.396	1.934.239	2.002.170	2.070.427	2.138.133	2.204.650
ASM	American Samoa	SP.POP.TOTL	Population, total	20.012	20.478	21.118	21.883	22.701	23.518
ATG	Antigua and Barbuda	SP.POP.TOTL	Population, total	54.681	55.403	56.311	57.368	58.500	59.653
AUS	Australia	SP.POP.TOTL	Population, total	10.276.477	10.483.000	10.742.000	10.950.000	11.167.000	11.388.000
AUT	Austria	SP.POP.TOTL	Population, total	7.047.539	7.086.299	7.129.864	7.175.811	7.223.801	7.270.889
AZE	Azerbaijan	SP.POP.TOTL	Population, total	3.897.889	4.030.130	4.167.558	4.307.315	4.445.653	4.579.759
BDI	Burundi	SP.POP.TOTL	Population, total	2.786.740	2.840.375	2.894.510	2.950.903	3.011.957	3.079.034
BEL	Belgium	SP.POP.TOTL	Population, total	9.153.489	9.183.948	9.220.578	9.289.770	9.378.113	9.463.667
BEN	Benin	SP.POP.TOTL	Population, total	2.431.620	2.466.002	2.503.232	2.543.335	2.586.362	2.632.360
BFA	Burkina Faso	SP.POP.TOTL	Population, total	4.829.291	4.894.578	4.960.325	5.027.818	5.098.892	5.174.869
BGD	Bangladesh	SP.POP.TOTL	Population, total	48.200.702	49.593.610	51.030.604	52.532.595	54.129.390	55.835.020
BGR	Bulgaria	SP.POP.TOTL	Population, total	7.867.374	7.943.118	8.012.946	8.078.145	8.144.340	8.204.168
BHR	Bahrain	SP.POP.TOTL	Population, total	162.501	167.924	173.107	178.048	182.774	187.348

Esta no sería una forma de los datos cómoda para trabajar con ella en Tableau, así que vamos a realizar el pivotaje de los atributos que representan los años, eliminando todas esas columnas y pasando a que cada dato correspondiente a un año sea una nueva fila para cada país y año.



Country Name	Country Code	Indicator Name	Indicator Code	Year	Population
Aruba	ABW	Population, total	SP.POP.TOTL	1960	54.208
Afghanistan	AFG	Population, total	SP.POP.TOTL	1960	8.994.793
Angola	AGO	Population, total	SP.POP.TOTL	1960	5.270.844
Albania	ALB	Population, total	SP.POP.TOTL	1960	1.608.800
Andorra	AND	Population, total	SP.POP.TOTL	1960	13.414
United Arab Emirates	ARE	Population, total	SP.POP.TOTL	1960	92.612
Argentina	ARG	Population, total	SP.POP.TOTL	1960	20.619.075
Armenia	ARM	Population, total	SP.POP.TOTL	1960	1.867.396
American Samoa	ASM	Population, total	SP.POP.TOTL	1960	20.012
Antigua and Barbuda	ATG	Population, total	SP.POP.TOTL	1960	54.681
Australia	AUS	Population, total	SP.POP.TOTL	1960	10.276.477
Austria	AUT	Population, total	SP.POP.TOTL	1960	7.047.539
Azerbaijan	AZE	Population, total	SP.POP.TOTL	1960	3.897.889


Se cargan el resto de ficheros con los datos por país de fertilidad o esperanza de vida, y se realiza el mismo pivotaje que se realizó anteriormente y así tendremos todos los datos con la misma estructura para poder visualizarlos de forma más óptima.

Una vez cargados todos los datos, hacemos las relaciones entre cada fuente de datos, por los atributos de año, país y código del país.


En la hoja de trabajo ya podemos ver todas las fuentes de datos con sus dimensiones y medidas, y las relaciones que habría entre ellas, porque dimensiones están hechas.

Datos


Análisis




Data (Country-Population)



Data (Fertility-Rate)





Data (Life-Expectancy-At-Birth)





Metadata - Countries (Country-Metadata)

Dimensiones












Country Code







Country Name




Indicator Code




Indicator Name




Year






Nombres de medidas


Medidas




Population




Latitud (generado)



Longitud (generado)




Número de registros




Valores de medidas

Datos


Análisis




Data (Country-Population)



Data (Fertility-Rate)





Data (Life-Expectancy-At-Birth)





Metadata - Countries (Country-Metadata)

Dimensiones












Country Code




Country Name




Indicator Code



Indicator Name




Year




Nombres de medidas


Medidas




Population




Latitud (generado)



Longitud (generado)



Número de registros



Valores de medidas

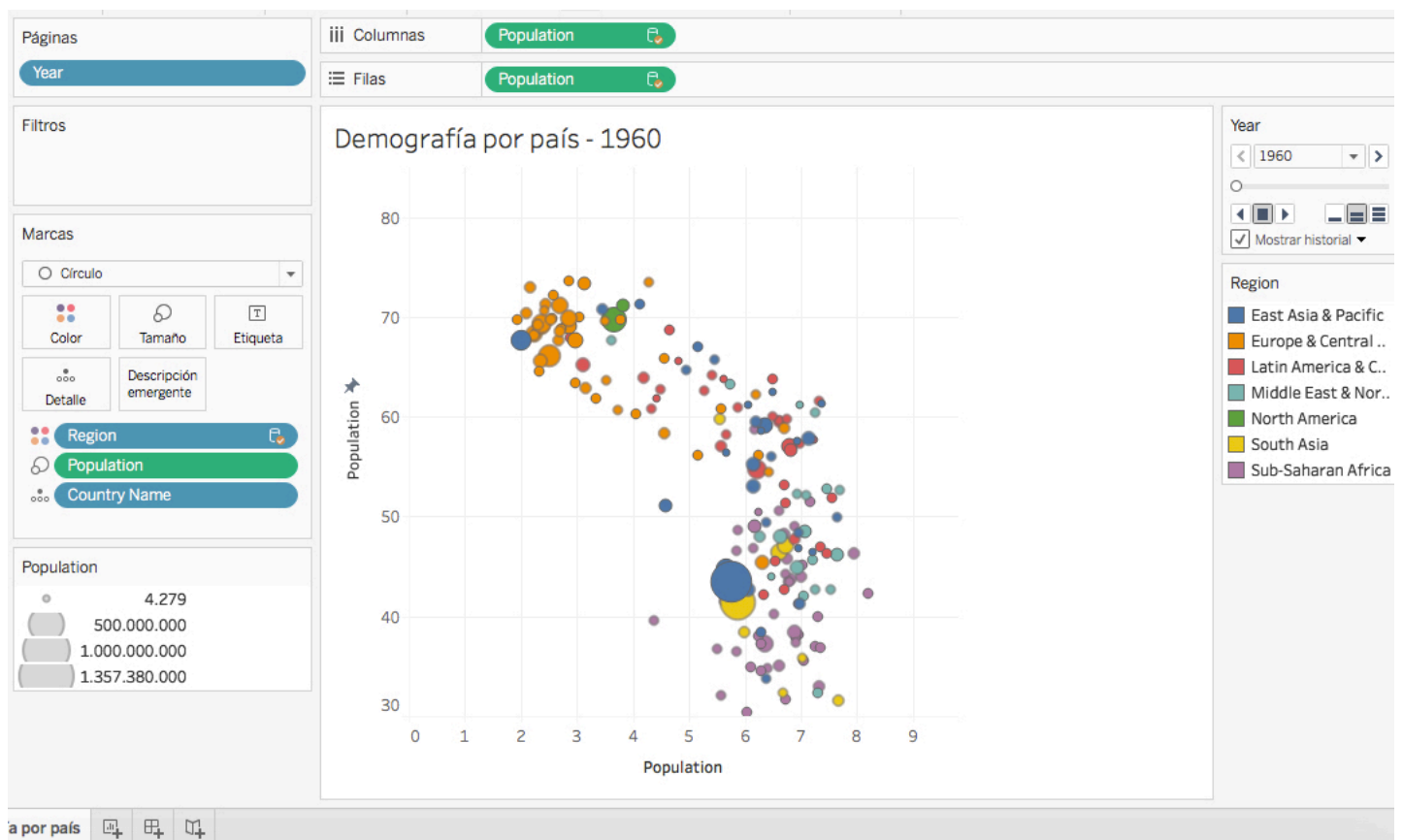
Para visualizar el gráfico que queremos arrastramos “Population” al tamaño, para que visualmente los países con mayor población tengan su elemento (en este caso círculo) de mayor tamaño que el resto.

Al eje de las columnas arrastramos la medida de fertilidad, así en el eje x del gráfico tendremos el número de hijos por familia. Al eje de las filas arrastramos la medida de esperanza de vida, de este modo en el eje y tendremos la esperanza media de vida. Ambas medidas las formateamos para que se representen de la forma que queremos.

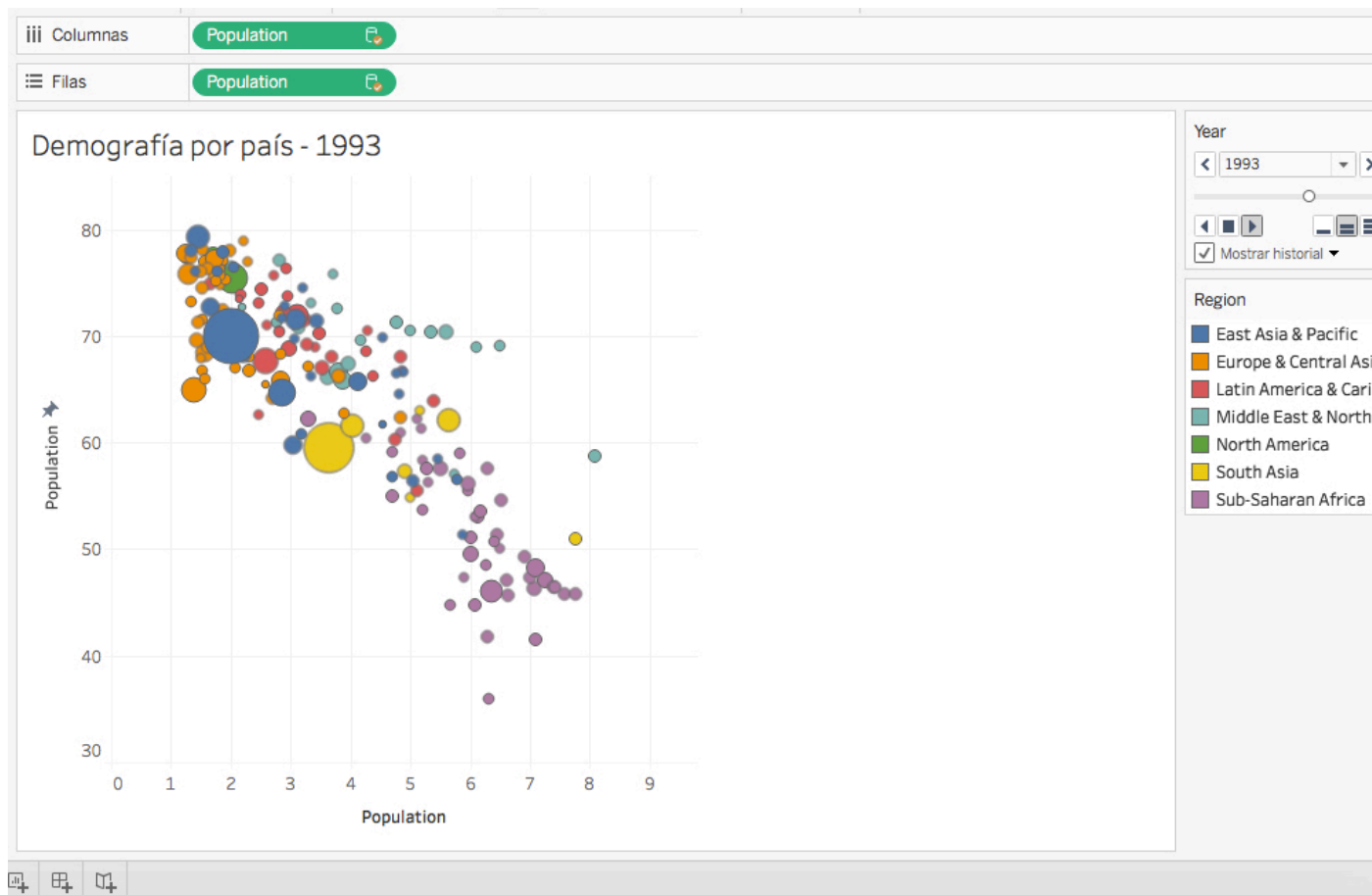
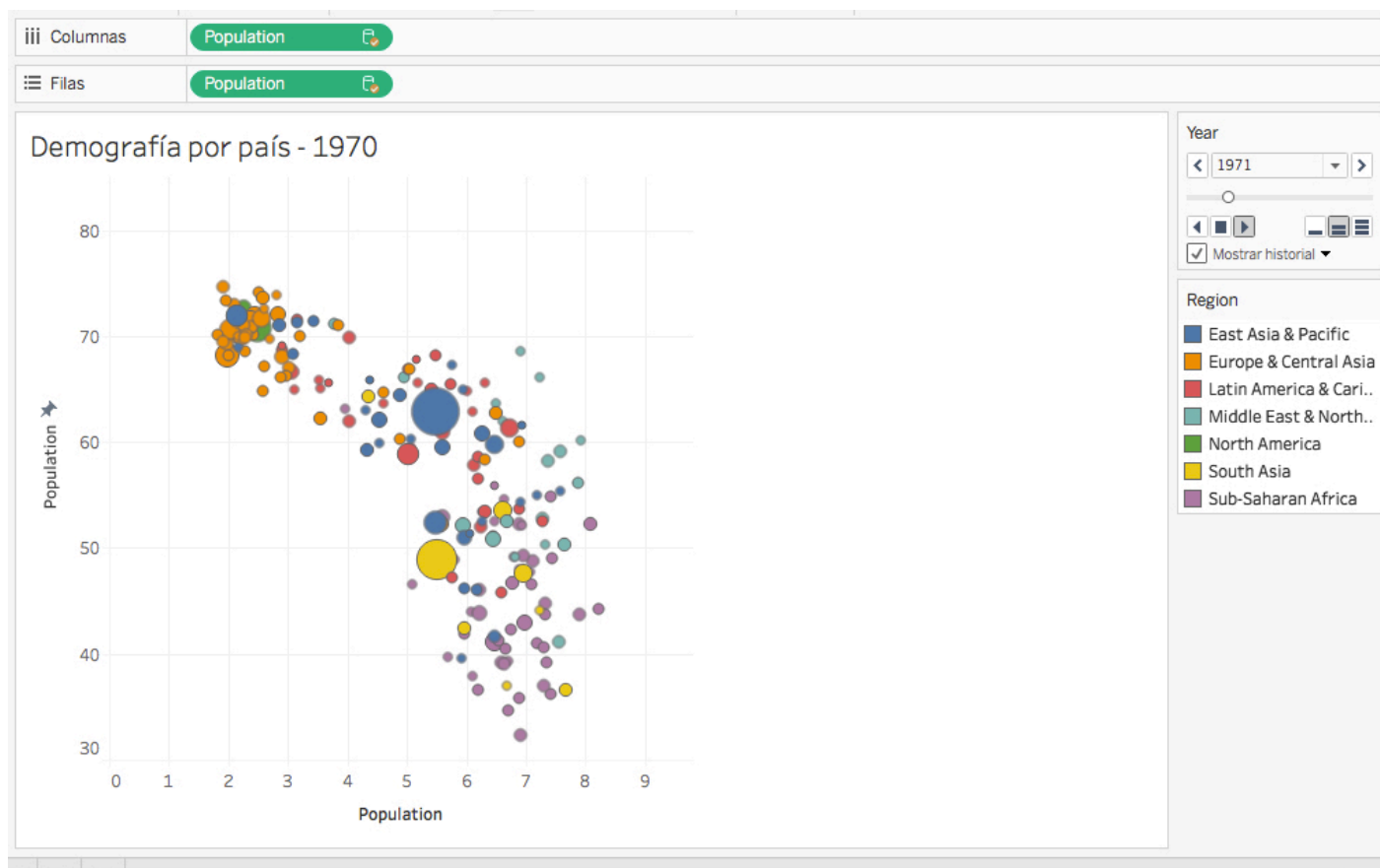
Por último, arrastramos la información de regiones o continentes al color, y así veremos representados los elementos por colores que representan cada región del mundo, y puede hacerse un análisis más profundo ya que también se puede estudiar por continente y no sólo por país.

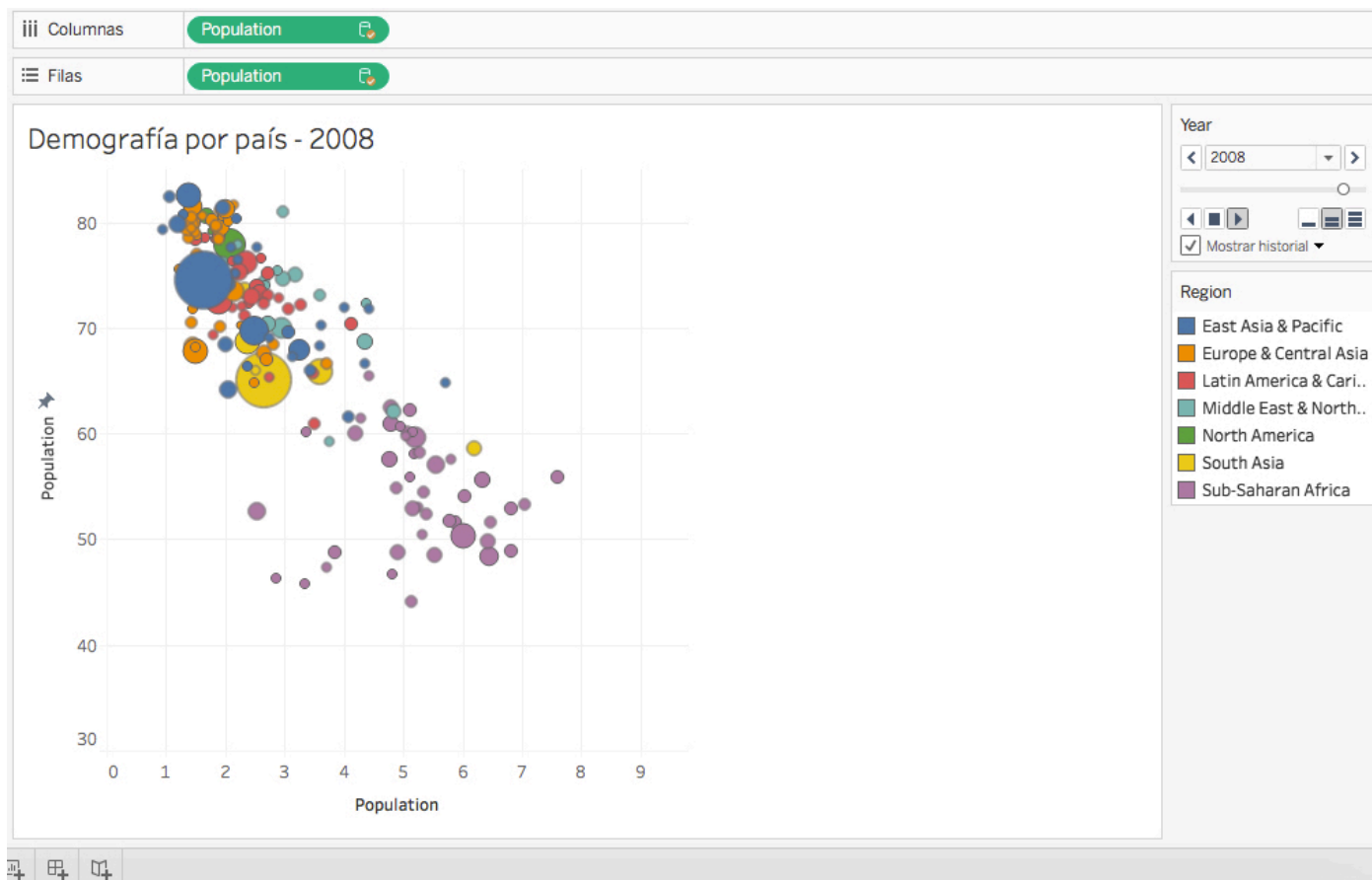
Ya tendríamos toda la granularidad que queremos para la visualización de los datos y poder analizarlos, ya el final es añadir la animación, arrastrando la información del año hasta el área de “Páginas”, y con esto se crea la animación que nos permite darle a un botón de play y que el gráfico se anime reproduciendo la información de cada año desde 1960 hasta el último año que se tengan datos de forma automática.

Gráfico final con la información del primer año.

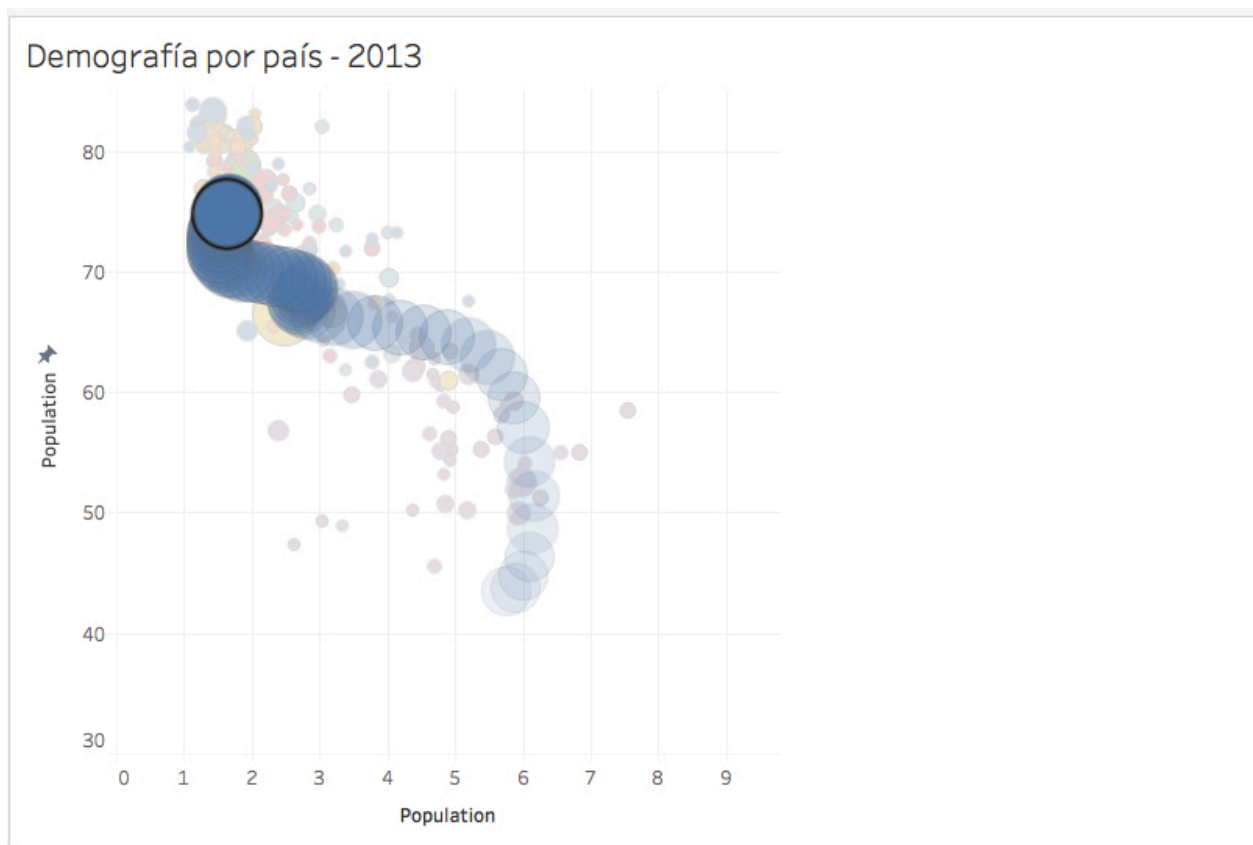


Activando la animación, podemos ir viendo la información año tras año.





Para finalizar activamos en la animación otros parámetros, como sería el rastro, pudiendo hacer que al seleccionar un país y activar la animación se vaya viendo el rastro que va dejando año a año, y así poder ver de forma rápida y concisa su evolución en las distintas características a analizar.



Este tipo de gráfica tiene gran detalle y por tanto se puede hacer multitud de análisis. Podemos analizar por país gracias al rastro que deja al resaltarlo y pulsar la animación, por ejemplo en la imagen vemos el rastro de un país y como a lo largo de estos 60 años ha bajado drásticamente el número de hijos por familia y ha subido más de 30 años la esperanza de vida media del país.

Podemos hacer el análisis por continente o región gracias al color que representa a cada una, y se aprecia como por ejemplo los países del Africa subsahariana se han quedado muy atrás del resto de regiones del mundo en la esperanza de vida, y además también son los países con mayor número de hijos por pareja.

Y ya de forma global sin tener en cuenta regiones, se observa claramente un descenso claro y drástico del número medio de hijos por familia en la mayoría de países y regiones, y un aumento progresivo y grande en la mayoría de regiones en la esperanza de vida a lo largo de los últimos 60 años.