



Microsoft Power BI

Aviso

Se recuerda que esta sesión será grabada, salvo el listado de asistentes y las preguntas del chat, que no serán objeto de grabación. Si no has prestado tu consentimiento, recuerda desactivar tu cámara y micro, ya que si realizas una participación verbal esta constituirá tu consentimiento para la grabación

Índice

Informes con varias páginas

Navegación

Filtros comunes

Influencia del modelo de datos

R y Python en Power BI

Ejercicio repaso

Marcadores

Ejercicio

En

https://raw.githubusercontent.com/RafaelCaballero/tdm/master/powerbi/empresas.csv

tenemos un listado de empresas con sus datos

Queremos que en la primera página se seleccione la empresa y otras dos páginas posteriores muestren información de esa empresa y de las empresas a las que factura

Comenzamos por cargar esta página como fuente de datos

Pag. 1

selección empresa facturas Listado de Empresas ✓ ☐ A la Vista Vinos. ∨ □ Aceites Olé ✓ ☐ AllAccess Tarjetas Gráficas. ✓ ☐ Alta Carnicería. 108 Automóviles A Todo Gas. ✓ ☐ Bisutería Perla. Total de empresas en la BD ∨ □ Bolsos Tándem. ∨ □ Brico DIY. ✓ ☐ Café Smarter. ✓ ☐ Camisetas Jump. ✓ □ Camisetas OnTarget. ✓ ☐ Cats Marketplace. Selección realizada

Pág. 2



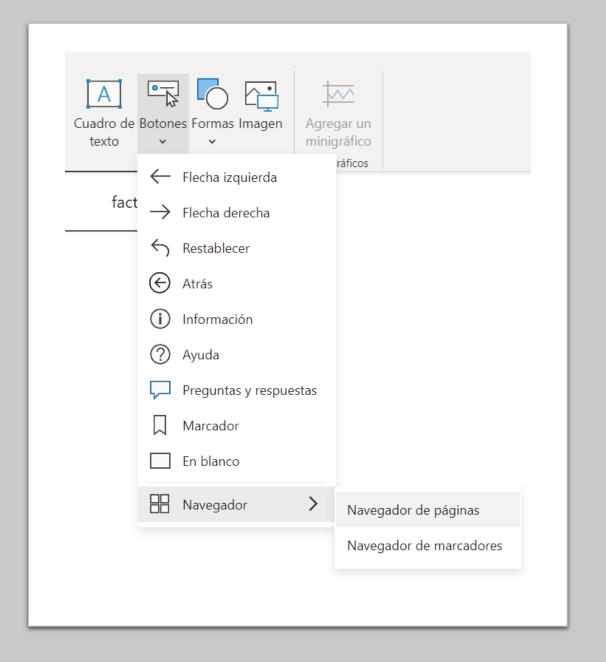
Pág. 3



Paso 1: páginas y su navegación

Creamos tres páginas: selección, empresa y facturas

En cada página elegimos el menú insertar añadimos botones de navegación



Paso 2: selección de empresa

En la página *selección* añadimos un segmentador (filtro) para el nombre de la empresa

Podemos añadir como segunda campo el CIF para el caso de más de una empresa con el mismo nombre

Listado de Empresas

✓ □	A la Vista Vinos.
$\vee \Box$	Aceites Olé
$\vee \square$	AllAccess Tarjetas Gráficas.
$\vee \Box$	Alta Carnicería.
$\vee \Box$	Automóviles A Todo Gas.
$\vee \Box$	Bisutería Perla.
$\vee \square$	Bolsos Tándem.
$\vee \Box$	Brico DIY.
$\vee \square$	Café Smarter.
\vee	Camisetas Jump.
✓ □	Camisetas OnTarget.
$\vee \square$	Cats Marketplace.

Paso 2: mensaje señalando que se debe seleccionar una empresa

Para ayudar al usuario queremos que se muestre bajo el filtro un mensaje animando a seleccionar una empresa, pero solo si no se ha seleccionado aun ninguno

Para eso vamos a añadir una medida que mostraremos en una etiqueta

```
SeleccionarUno =
if(COUNT(empresas[CIF])==1, "Selección
realizada", "Por favor, seleccione una empresa")
```

Listado de Empresas

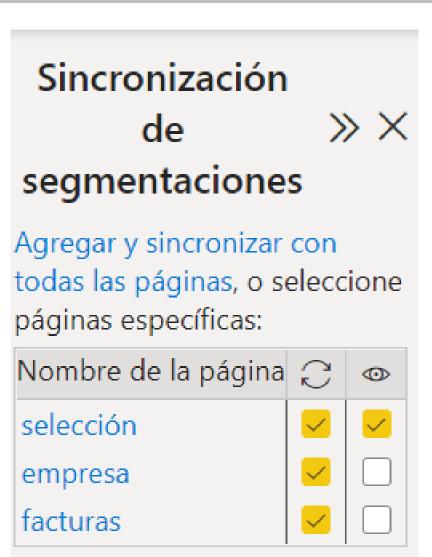
$\vee \Box$	A la Vista Vinos.
$\vee \Box$	Aceites Olé
$\vee \Box$	AllAccess Tarjetas Gráficas.
$\vee \square$	Alta Carnicería.
$\vee \square$	Automóviles A Todo Gas.
$\vee \square$	Bisutería Perla.
\vee	Bolsos Tándem.
$\vee \square$	Brico DIY.
✓ □	Café Smarter.
$\vee \Box$	Camisetas Jump.
✓ □	Camisetas OnTarget.
✓ □	Cats Marketplace.

Selección realizada

Paso 3: compartir el filtro

Para que el filtro afecte a varias páginas y no solo a una vamos a "Ver" + "Sincronización de segmentaciones"

Allí seleccionamos nuestro filtro e indicamos que se utilizará en las 3 páginas pero solo se verá en la primera



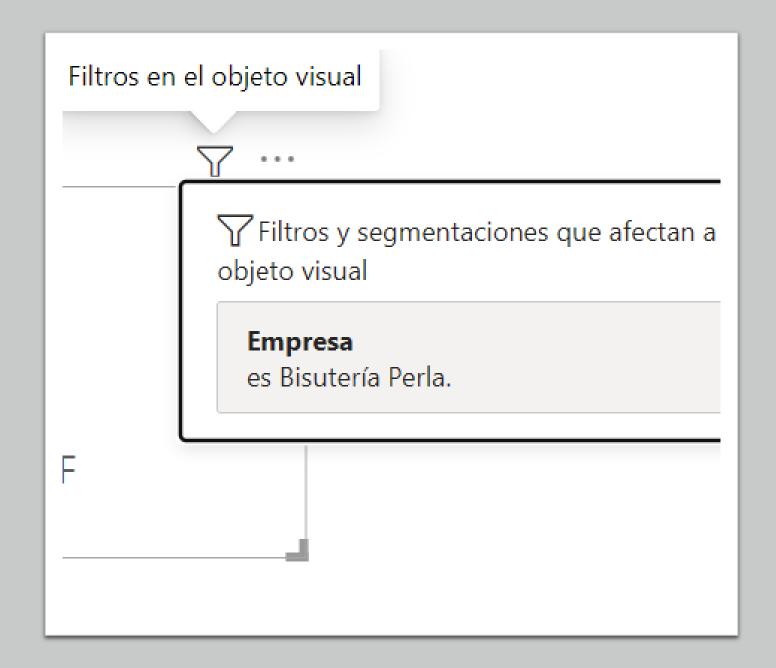
Paso 4: componentes independientes del filtro

Vamos a añadir una etiqueta al lado del filtro que muestre el total de empresas que hay en la base de datos

Para ello seleccionamos como campo el CIF y como resumen "Recuento distintivo"

Se ve afectado por el filtro, aunque no es lo que queremos ¿Por qué? ¿Cómo solucionarlo?

Arriba del componente a la derecha vemos que el filtro le está afectando



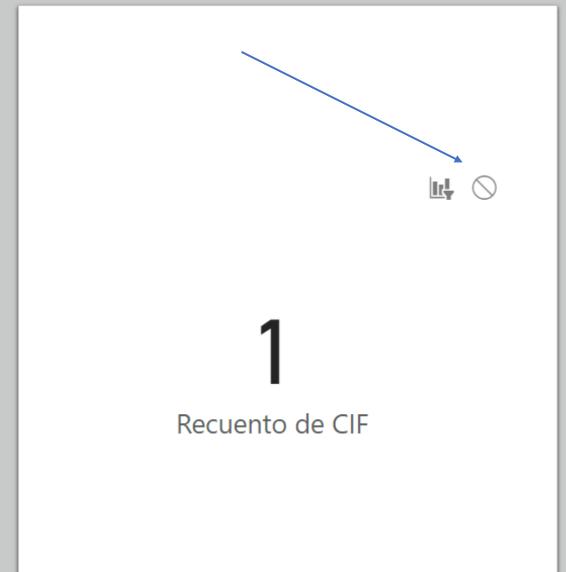
Paso 4: solución

Hacemos click fuera de la etiqueta (no la podemos tener seleccionada)

Entonces hacemos un cambio en el filtro (desmarcar si estaba marcado, o viceversa)

Nos fijamos en la etiqueta debe tener un símbolo similar al de prohibido en la parte superior

Hacemos ahí click para que quede bloqueado el uso de filtros



Paso 5: página empresa

Simplemente mostramos en tarjetas los datos de la empresa

selección facturas empresa CIF Dirección Q55555 Rosal 5 Provincia Nombre de la empresa Girasoles Solete Sevilla

Paso 6:carga de datos de facturas y relación con el modelo

Cargamos el fichero

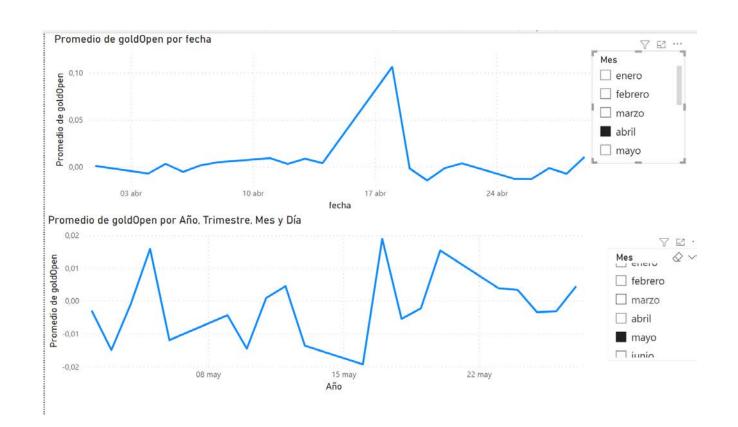
https://raw.githubusercontent.com/RafaelCaballero/tdm/master/powerbi/facturas.csv de facturas

¿Qué cambio(s) en el modelo de datos parece que hay que hacer? Recordemos que:

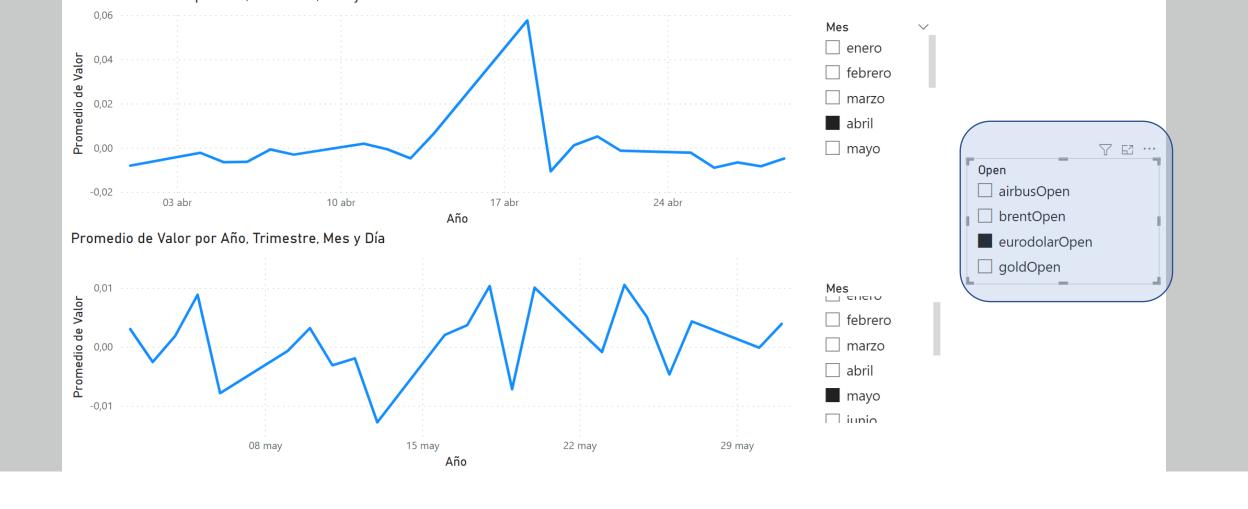
La empresa facturadora es la fijada en la primera página

De la empresa facturada hay que mostrar además el nombre

Ejercicio



En el dataset stockpc del incremento de los valores bursátiles con respecto al día anterior queremos poder comparar una gráfica consigo misma en un mes diferente ¿cómo podemos hacerlo?



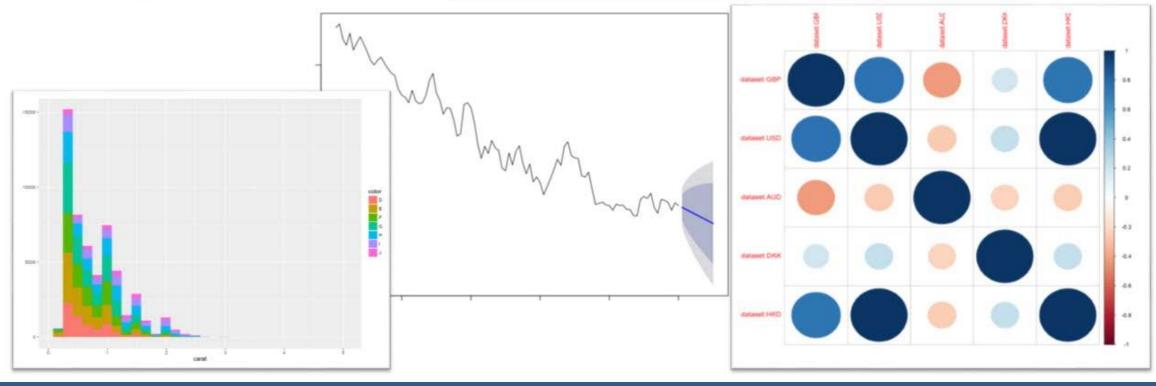
Un paso más

Ahora queremos poder elegir cuál de los 4 valores Open se va a mostrar en ambas gráficas ¿cómo hacerlo?





Power BI



Funcionalidades R

 A pesar de toda su potencia Power BI tiene limitaciones tanto en cuento a transformaciones como la inclusión de gráficos

 La posibilidad de incluir scripts en lenguajes de programación como R o Python eliminan en buena medida estas limitaciones

R y Power Bl

1) Ejecución de scripts desde Power Query para hacer transformaciones complejas

- Al importar datos
- En Power Query

2) Incluir gráficos de R en el report

Se puede programar en RStudio

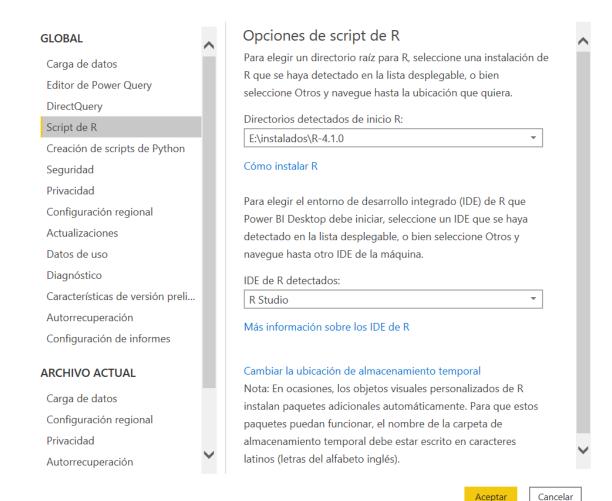
Preparación

Si tenemos ya instalado un entorno de R como Rstudio, Power BI debe ser capaz de detectarlo

Vamos a Archivo + Opciones y Configuración + Opciones

Si no está deberíamos instalarlo y reiniciar Power Bl

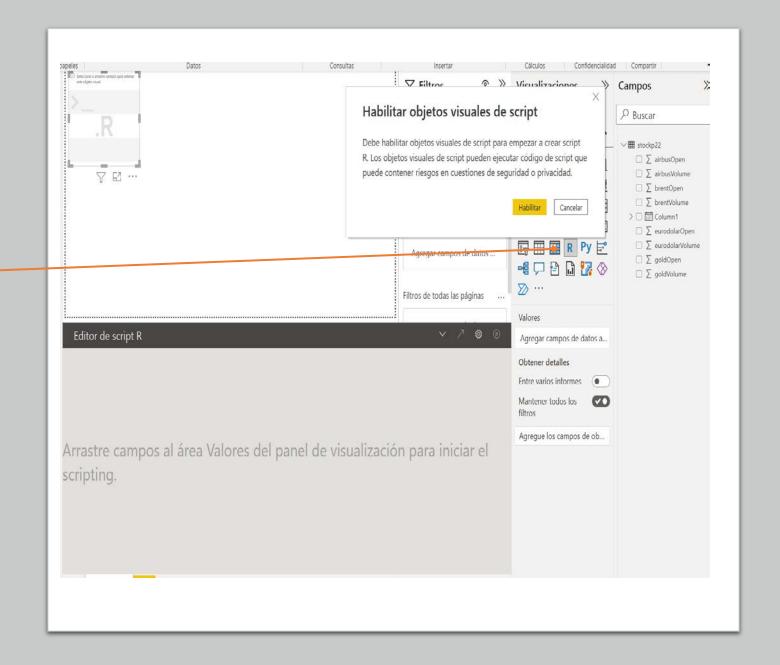
Opciones



 \times

Objetos visuales desde R

- En primer lugar cargamos los datos y preprocesamos para que los tipos estén bien
- En el panel de informe seleccionamos objeto de tipo R
- Es posible que se nos pida la primera vez permiso para su uso
- A continuación seleccionamos los valores a utilizar (en nuestro caso los 4 "open")



Objetos visuales: añadir código

Se nos abre una ventana donde podemos añadir código

En nuestro caso añadimos código para mostrar un diagrama de correlaciones

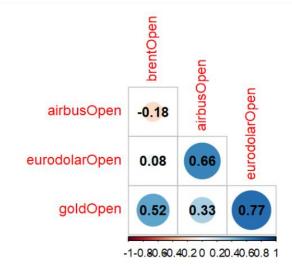
El código debe partir de un dataframe de nombre dataset

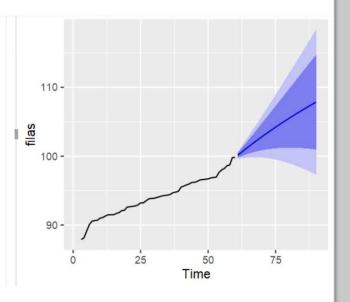
```
library(corrplot)
df2 <- na.omit(dataset)
M<-cor(df2)
corrplot(M, method="circle",addCoef.col = "black", diag=0, type="lower")</pre>
```

Ejercicio

Añadir el gráfico que se genera con el siguiente código

```
library(ggplot2)
library(forecast)
boDiff<- dataset$airbusOpen
filas <- boDiff[1:60]
fit <- Arima(filas, order=c(2,1,1))
autoplot(forecast(fit,h=30))</pre>
```





Objetos visuales en R

- Siempre se eliminan filas repetidas al comienzo; si no se quiere que esto suceda se puede añadir una columna índice en Power Bl

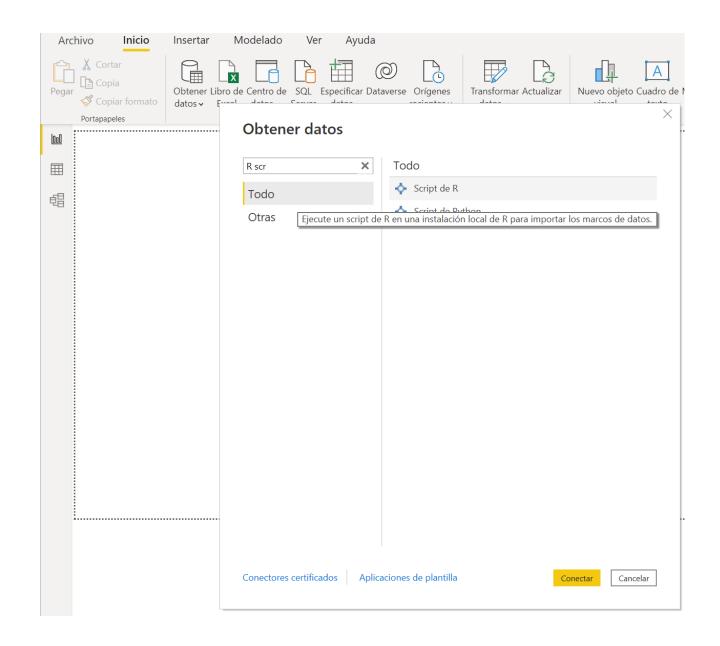
- Limitación a 150 000 filas y 72 ppi

- No son gráficos interactivos

- Los scripts no puede cambiar el nombre de las columnas

Scripts R en Power BI: importación de datos

Podemos elegir R como método para importar los datos en "Obtener Datos"



Importación de datos con R

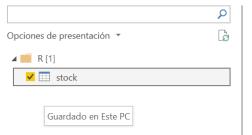
Se abre una ventana para escribir el código. Por ejemplo

```
url =
"https://raw.githubusercontent.
com/RafaelCaballero/tdm/mast
er/datos/stockp22.csv"
```

```
stock <- read.csv(url, sep=";",
dec = ",")</pre>
```

El nombre stock aparecerá como nombre de una tabla a seleccionar

Navegador





Х	airbusOpen	airbus Volume	brentOpen	brentVolume
2022-01-03	113,5999985	1320377	77,94000244	27224
2022-01-04	116,5800018	1588549	78,94999695	31321
2022-01-05	119,3199997	1806200	80,08000183	35152
2022-01-06	118,1999969	1441589	80,15000153	41582
2022-01-07	118	1244642	82,05999756	29501
2022-01-10	118	1099153	81,79000092	30035
2022-01-11	118,5999985	1060555	80,95999908	37934
2022-01-12	118,7200012	1288629	83,69000244	32298
2022-01-13	118	1278350	84,77999878	27107
2022-01-14	116,2799988	1051871	84,12999725	31970
2022-01-17	118,2799988	958258	null	nuli
2022-01-18	116,1999969	1131796	86,5	45875
2022-01-19	115,5599976	1479906	88,44999695	37483
2022-01-20	115,0800018	1161977	87,73000336	30524
2022-01-21	114	2241374	87,34999847	29566
2022-01-24	112,5199966	2524642	87,65000153	34253
2022-01-25	108,8199997	2101749	87,16999817	29496
2022-01-26	111	1868916	88,01999664	29805
2022-01-27	111,1200027	1523610	89,77999878	15201
2022-01-28	113,5	1530319	89,88999939	10824
2022-01-31	112,7600021	1636479	90,80000305	118
2022-02-01	113,5800018	1601492	91,29000092	36248
2022-02-02	115,3799973	1412361	89,29000092	38172
<				>



Transformar datos



 \square \times

2

Añadir pasos en R en power query

Ejemplo: cargamos la tabla de datos de comunidades de esta página

https://es.wikipedia.org/wiki/Anexo:Comunidades y ciudades aut%C3%B3nomas de Esp a%C3%B1a

La tabla presenta numerosos errores de formato

Podemos arreglar mucho, pero por ejemplo el nombre repetido de las comunidades (con excepción de Navarra) es difícil

Para este paso en Power Query elegimos R

Allí tecleamos

library(stringr)

mitades <- str_length(dataset\$Nombre)/2

dataset\$Nombre <- ifelse(str_detect(dataset\$Nombre,"Navarra"), dataset\$Nombre, str_sub(dataset\$Nombre,end=mitades))

output <- dataset</pre>

Privacidad

Nos puede dar un aviso de privacidad

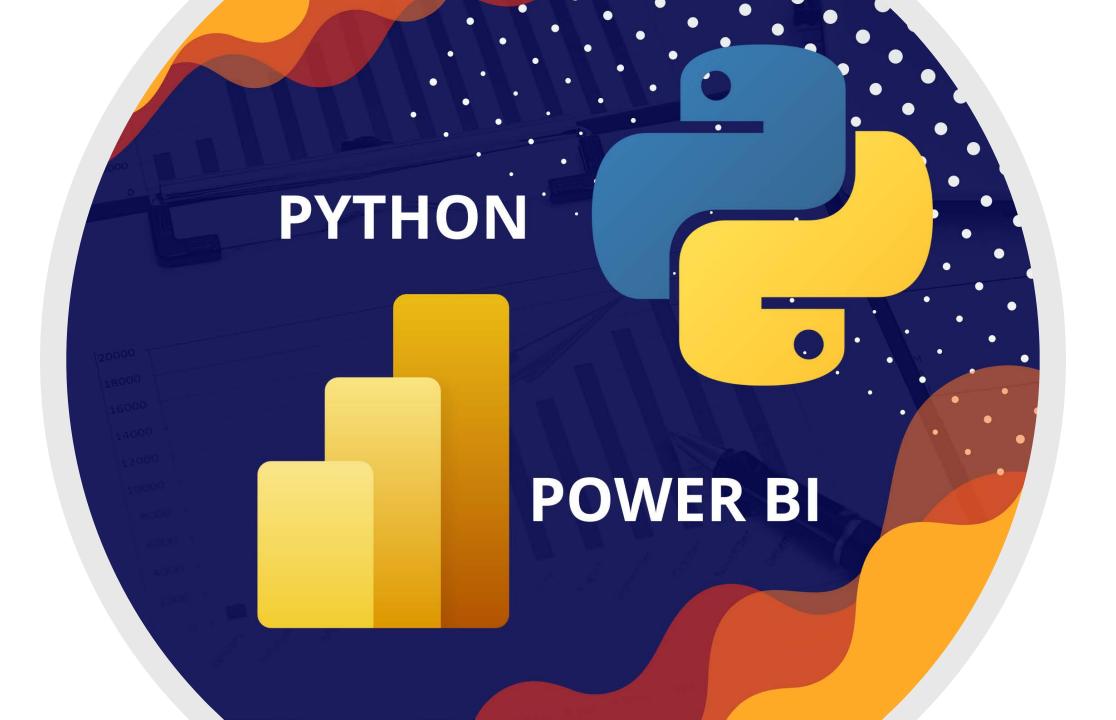


Detalles importantes

1) Los scripts deben ser autocontenidos; en particular deben cargar todas las librerías que se necesiten

2) Las librerías tienen que haber sido instaladas previamente (install.packages("stringr"))

3) La última instrucción output <- dataset es importante porque se espera que el dataset se llame output.

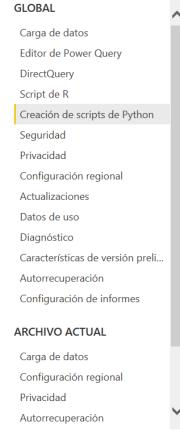


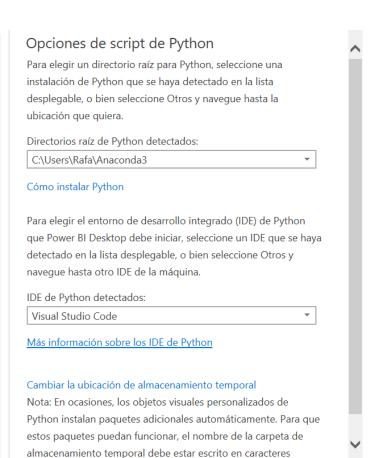
Configuración

- Igual que en el caso de R; ir a Archivo + Opciones y Configuración + Opciones
- Allí en "creación de scripts de Python" Power Bl nos debería mostrar que ya tenemos instalado Anaconda

Opciones

TAREAU TALERINU - CONTROLLICIONO





Aceptar

Cancelar

Carga de tablas con Python

Ejemplo 1: Inicio + Obtener Datos + Mas + Python Script

import pandas as pd

Interesante por ejemplo para generar tablas con valores aleatorios y hacer simulaciones

Carga de tablas con Python

Navegador



tabla

nombre	apellidos	provincia	edad
МРјН	xvzSJD	Guadalajara	63
TiwrcXRPc	EihfRGwLWRnBRfj	Madrid	56
uyfGlvHKYl	fGlvHKYI tbJlWCuCPPzgWWNPAdfTPI Ponte		44
Duintj	xbjhUnCOmwNDRQtlUECtiDPF	Barcelona	19
mgjPbBOgY	PqFvsyxkpCxS	Guadalajara	59
gCVrYO	OfJqldmSTOTnLDQeKsAThL	Cádiz	102
IiAya	CJmxalFLrzvrc	Jaén	4
TZekVh	jSiHWES	Ávila	11
eftfnEDjP	NWbrsMIUGsFjmKI	Cádiz	18
ijQBAUdkE	TmLhRQFCoLlvVxdPEvImMsNVd	Vizcaya	23
rOrrgQ	strupJWGaOjQXQTOr	Navarra	10
VXJKrOqKw	FArdhioxRtBe	Cáceres	19
AseeB	HVcZCzZfyeVHhUzYSy	Albacete	14
AWwgzVPaa	JMXJblNtKQ	Sevilla	81
RGoCETQI	MfmFmvldLAUw	Almería	18
oltqlahwJb	pcmLiqizmFAWDPWKeC	Zaragoza	62
IuSFLfq	cjAbOpZcFAxRPzvWQYyQOI	Alicante	49
RoPNE	ZainHxeli	Jaén	56
rAjMY	fYAwYGPenI	Almería	89
IARIvwY	TNnyesCrEVFwsjw	Soria	65
WsXAbDmVJ	FCjfjRCeYFJNIcYi	Álava	31
gGfg	vlwrRvTvrnIKhuoMfex	Guadalajara	71
hgifKijbY	AqazOWodY	Vizcaya	80
tCXLFtMqxR	cdroazVbeYhyCT	Segovia	97

Cargar

Transformar datos

Cancelar

2

```
import string
import random
import pandas as pd
def random char(y):
  return ". join(random. choice(string. ascii_letters) for x in range(y))
cuantos=random.randint(5,100) # de 5 a 100 filas
provincias=["Álava,..."Zaragoza"];
nombre = []
apellidos = []
provincia = []
edad = []
for i in range(cuantos):
  nombre.append(random char(random.randint(4,10)))
  apellidos.append(random_char(random.randint(6,25)))
  provincia.append(random.choice(provincias))
  edad.append(random.randint(0,110))
tabla = pd.DataFrame({"nombre":nombre, "apellidos":apellidos, "provincia":provincia, "edad":edad})
tabla
```

Scripts en Power Query con Python

Ejemplo: en el fichero https://raw.githubusercontent.com/RafaelCaballero/tdm/master/powerbi/empresas-dirt.csv

Tenemos un fichero con empresas pero algunos nombres tienen caracteres diferentes de letra al principio que queremos eliminar

```
def limpia(n):
    i=0
    while i< len(n) and not n[i].isalpha(): i+=1
    return n[i:].strip() if i<len(n) else None

dataset["Empresa"] = dataset.Empresa.map(limpia)</pre>
```

Scripts en Power Query con Python

• Ejemplo: vamos a utilizar los datos de

https://raw.githubusercontent.com/RafaelCaballero/tdm/master/datos/stockp22.csv

- Vamos a hacer una copia (o referencia?) y quedarnos solo con las columnas fecha, airbusVolume, brentVolume, goldVolume
- El objetivo es estandarizar cada una de las columnas de volumen para que se conviertan en una normal 0,1 y poder comparar mejor
- "Protegemos" la fecha convirtiéndola en texto

Script en Python para estandarizar

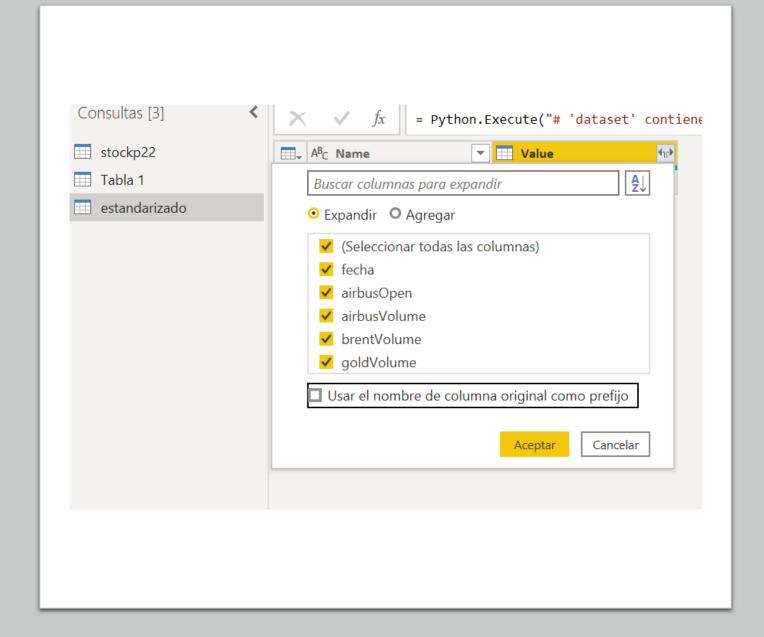
• En Power Query elegimos Python y añadimos el código

```
from sklearn.preprocessing import StandardScaler
cols = ["airbusVolume","brentVolume","goldVolume"]
dataset = dataset.dropna()
data = dataset[cols]
data = StandardScaler().fit(data).transform(data)
for i in range(len(cols)):
    dataset.loc[:,cols[i]] = data.transpose()[i]
```

Script en Python para estandarizar

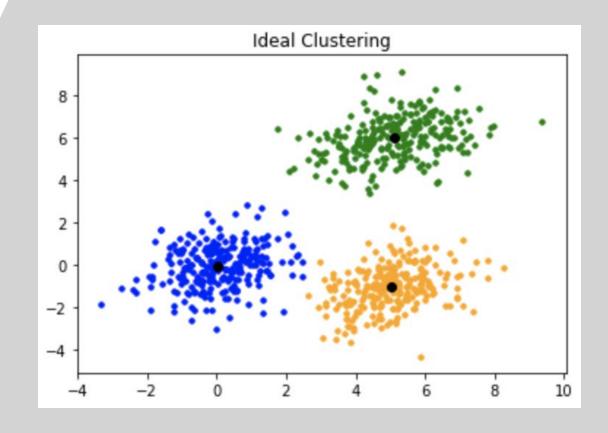
En el resultado expandimos la tabla con el resultado

Todavía habrá que añadir algunos cambios como convertir la fecha



Scripts en Python: machine learning

- Vamos a utilizar la técnica de Kmeans para agrupar los datos del volumen en 4 grupos, cada uno incluyendo elementos con características similares
- El script generará una columna nueva con el número de clúster, de 0 a 3
- El número de clústers, 4, es arbitrario y se puede (y se debe) cambiar



Machine Learning con Python

```
import pandas as pd
from sklearn.cluster import KMeans
from sklearn.preprocessing import MinMaxScaler

columnas = ["airbusVolume","brentVolume","goldVolume"]

df = dataset.dropna() # quitar nulos

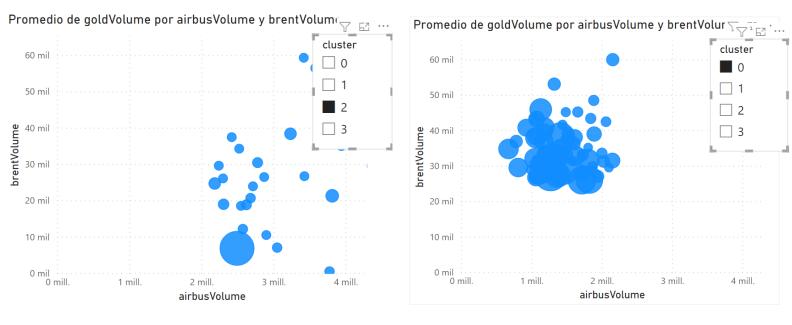
df2 = df[columnas] # seleccionar columnas

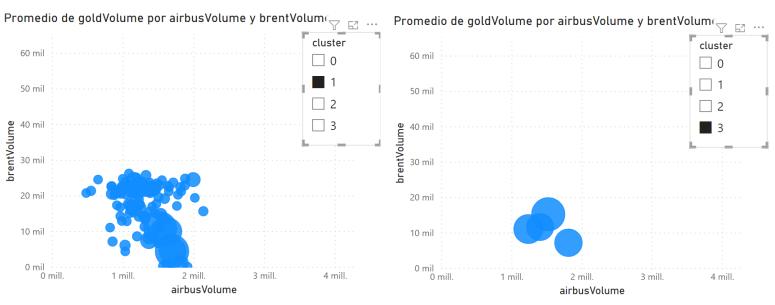
data = MinMaxScaler().fit(df2).transform(df2) # escalar

df["cluster"] = KMeans(n_clusters=4, random_state=0).fit_predict(data) # kmeans
```

Comprobando....

El ultimo cluster es muy pequeño, se puede probar con 3



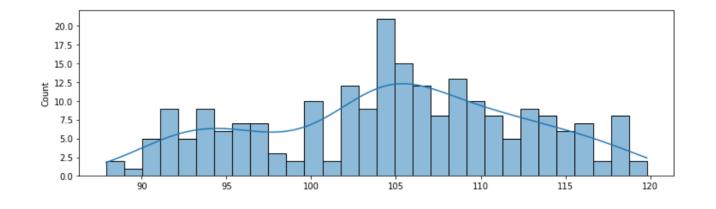


Python para objetos visuales

Ejemplos.

Utilizaremos la librería seaborn para hacer un histograma de airbusOpen a partir de stockp22

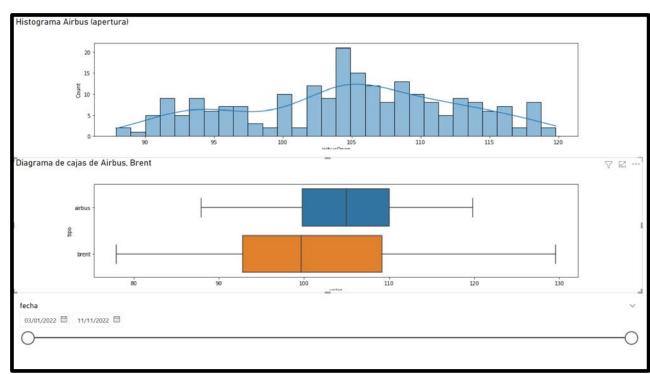
```
import seaborn as sns
import matplotlib.pyplot as
plt
sns.histplot(data=dataset,x
="airbusOpen",bins=30,kde=T
rue)
plt.show()
```



Ejercicio

Añadir el siguiente código para incorporar el box plot de AirbusOpen y de BrentOpen

```
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns
dataset.dropna(inplace=True)
                                            Diagrama de cajas de Airbus, Brent
df1 = dataset[["airbusOpen"]]
df1["tipo"] = "airbus"
df1.columns = ["valor","tipo"]
df2 = dataset[["brentOpen"]]
df2["tipo"] = "brent"
                                            03/01/2022 🖽 11/11/2022 🖽
df2.columns = ["valor","tipo"]
df3 = pd.concat([df1,df2],axis=0)
sns.boxplot(data=df3, x="valor", y="tipo",orient="h")
plt.show()
```



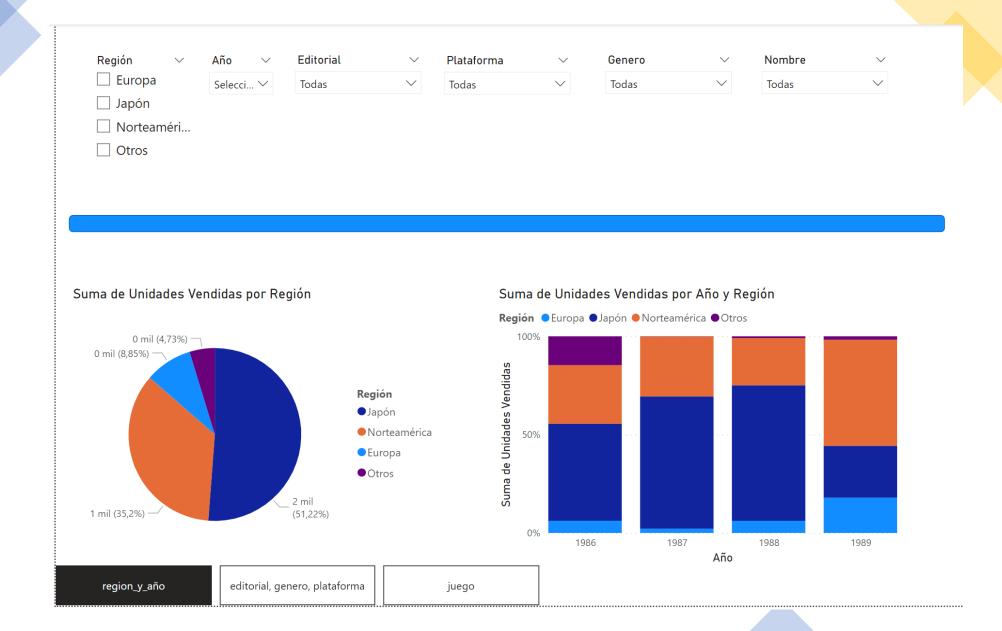
Ejercicio

Consideramos (de nuevo) el <u>dataset de videojuegos</u> con columnas ['Año', 'Editorial', 'Genero', 'Nombre', 'Plataforma', 'Región', 'Unidades Vendidas']

- La primera página permitirá restringir si se desea el año y la región
- La segunda la editorial, el género y la plataforma.
- Finalmente una tercera página nos permitirá elegir el nombre del juego concreto

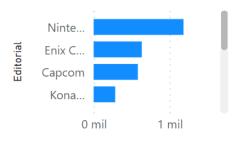
Todos los filtros tienes que afectar a las páginas siguientes:

- los filtros de la pág. 1 afectarán a las pág. 1, 2 y a la 3
- los filtros de la 2 a la 2 y a la 3
- Los filtros de página 3 solo a la página 3

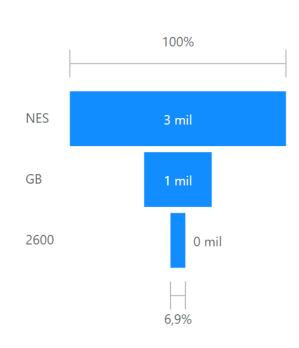




Suma de Unidades Vendidas por Editorial



Suma de Unidades Vendidas por Plataforma



Suma de Unidades Vendidas por Genero



region_y_año editorial, genero, plataforma

juego

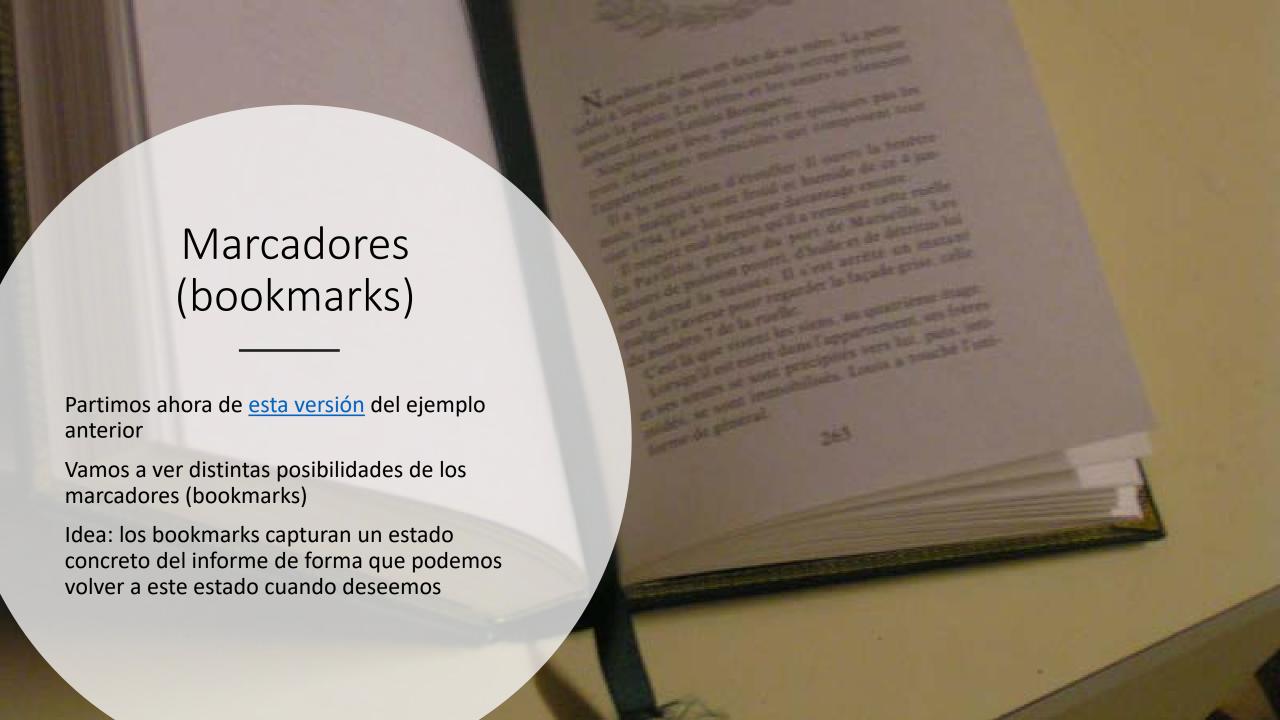


							7 63
Región	Año	Editorial	Plataforma	Genero	Nombre	Suma de Unidades Vendidas	
Japón	1986	Capcom	NES	Action	Commando	25	
Japón	1986	Capcom	NES	Action	Ghosts 'n Goblins	61	
Japón	1986	Capcom	NES	Role-Playing	Dragon Warrior	52	
Japón	1986	Hudson Soft	NES	Action	Doraemon	115	
Japón	1986	Hudson Soft	NES	Platform	Adventure Island	105	
Japón	1986	Konami Digital Entertainment	NES	Platform	Castlevania	62	
Japón	1986	Konami Digital Entertainment	NES	Platform	Ganbare Goemon! Karakuri Douchuu	12	
Japón	1986	Konami Digital Entertainment	NES	Shooter	Gradius	1	
Japón	1986	Nintendo	NES	Platform	Kid Icarus	109	
Japón	1987	Activision	2600	Action	Commando	0	
Japón	1987	Activision	2600	Action	Kung-Fu Master	0	
Japón	1987	Atari	2600	Shooter	Defender II	0	
Lanón	1007	Atori	2600	Chaotar	Docart Falcon	0	
Total					_	3506	

region_y_año

editorial, genero, plataforma

juego



Aplicación 1: estado de filtros



Como en el panel hay muchos filtros puede ser una lata ir borrándolos todos cada vez que queremos volver al estado inicial

Vamos a utilizar un marcador para "recordar" el estado inicial de los marcadores

Pasos

- 1) Borramos a mano todos los filtros
- 2) En Ver + Marcadores se nos abre el panel de marcadores. Allí añadimos uno con el nombre que queramos (reset_filters p.ej.)
- Añadimos un botón que tenga como asociado ese marcador (imagen del botón en Formato+Botón+Icono+Tipo personalizado)

También se pueden crear para recordar combinaciones de filtros determinadas

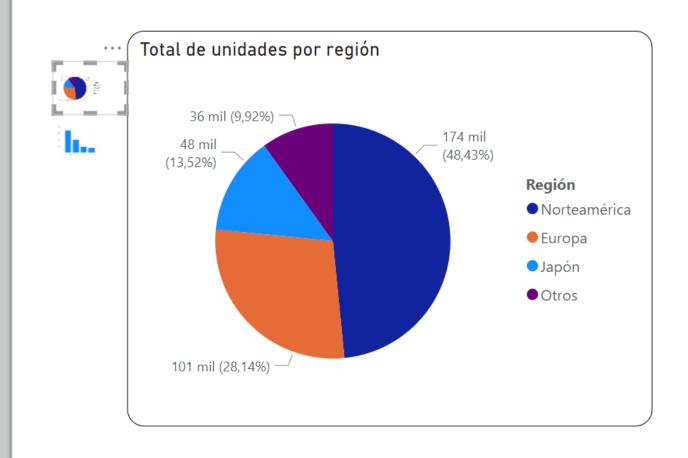
Aplicación 2: alternar gráficos

Vamos a ver + Selección

Hacemos click en el gráfico que queramos y en el panel de selección le ponemos un nombre que podamos recordar

Ahora sobre el gráfico hacemos Ctr+C+Ctr+V para crear un duplicado

Cambiamos el duplicado al tipo que queramos



Ejercicio



Hacer que los filtros se puedan tapar y ver

Incluir encima una forma que contenga un título y señal de menu

Incluir en los filtros un botón para cerrarlos que cambie de uno a otro