

# SPRINT 8

## Nivel 1

### Conexión Power BI con Python:

Lo primero de todo es necesario habilitar la creación de scripts de Python desde opciones y configuración de Power BI. En la pestaña 'Creación de scripts de Python' Debo seleccionar la ubicación de Python y además la de la IDE que yo desee, en mi caso elijo como IDE de Python Visual Studio Code y busco la ubicación del ejecutable en mi ordenador para vincularla.

#### Opciones

The screenshot shows the 'Opciones de script de Python' (Python script options) tab in the Power BI Options dialog. On the left is a sidebar with various categories, with 'Creación de scripts de Python' (Python script creation) selected. The main area contains instructions and configuration fields. It includes a dropdown for 'Directorios raíz de Python detectados' (Detected Python root directories) showing a path to Python. There is a link 'Cómo instalar Python' (How to install Python). Below that, it asks for the 'IDE de Python detectados' (Detected Python IDE) with a dropdown set to 'Otros' (Others). A text field for 'Vaya al IDE de Python que quiera usar:' (Go to the Python IDE you want to use:) contains the path to the Visual Studio Code executable, with an 'Examinar' (Browse) button. At the bottom, there is a link 'Más información sobre los IDE de Python' (More information about Python IDEs).

**Opciones de script de Python**

Para elegir un directorio raíz para Python, seleccione una instalación de Python que se haya detectado en la lista desplegable, o bien seleccione Otros y navegue hasta la ubicación que quiera.

Directorios raíz de Python detectados:

C:\Users\ronch\AppData\Local\Programs\Python\Pyth...

[Cómo instalar Python](#)

Para elegir el entorno de desarrollo integrado (IDE) de Python que Power BI Desktop debe iniciar, seleccione un IDE que se haya detectado en la lista desplegable, o bien seleccione Otros y navegue hasta otro IDE de la máquina.

IDE de Python detectados:

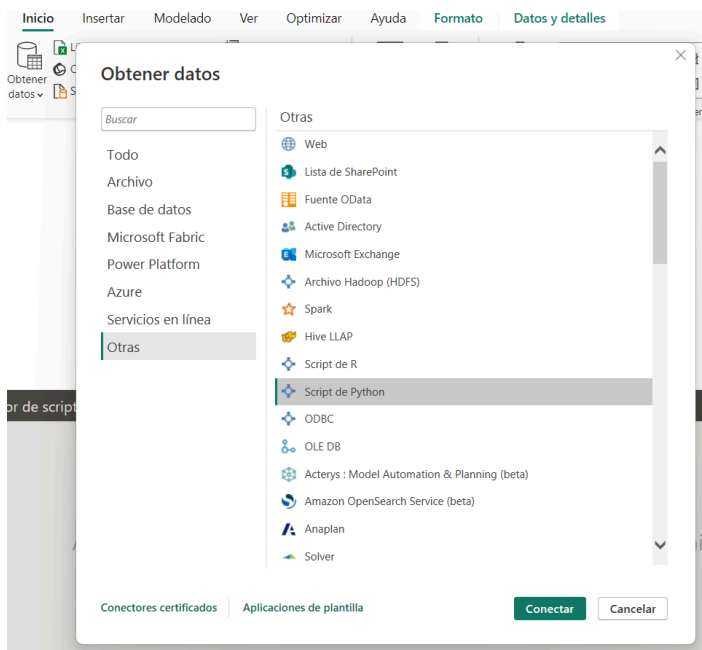
Otros

Vaya al IDE de Python que quiera usar:

C:\Users\ronch\AppData\Local\Programs\Microsoft VS Code\Code.exe **Examinar**

[Más información sobre los IDE de Python](#)

Una vez realizada la configuración, voy a inicio de un nuevo informe de Power BI. Desde la opción, de obtener datos, hago click en 'Más' y aparece, entre otras, la opción 'Script de Python', como se muestra en figura.



Hago una prueba con un script pequeño de prueba para ver si estoy conectada y funciona. Después de aceptar esta pantalla, voy a la visión de tablas y compruebo que se ha cargado el DataFrame importado a través del script de Python.

### Script de Python

Script

```
import pandas as pd
df = pd.DataFrame({
    'Fname': ['Harry', 'Sally', 'Paul', 'Abe', 'June', 'Mike', 'Tom'],
    'Age': [21, 34, 42, 18, 24, 80, 22],
    'Weight': [180, 130, 200, 140, 176, 142, 210],
    'Gender': ['M', 'F', 'M', 'M', 'F', 'M', 'M'],
    'State': ['Washington', 'Oregon', 'California', 'Washington', 'Nevada', 'Texas', 'Nevada'],
    'Children': [4, 1, 2, 3, 0, 2, 0],
    'Pets': [3, 2, 2, 5, 0, 1, 5]
})
print (df)
```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

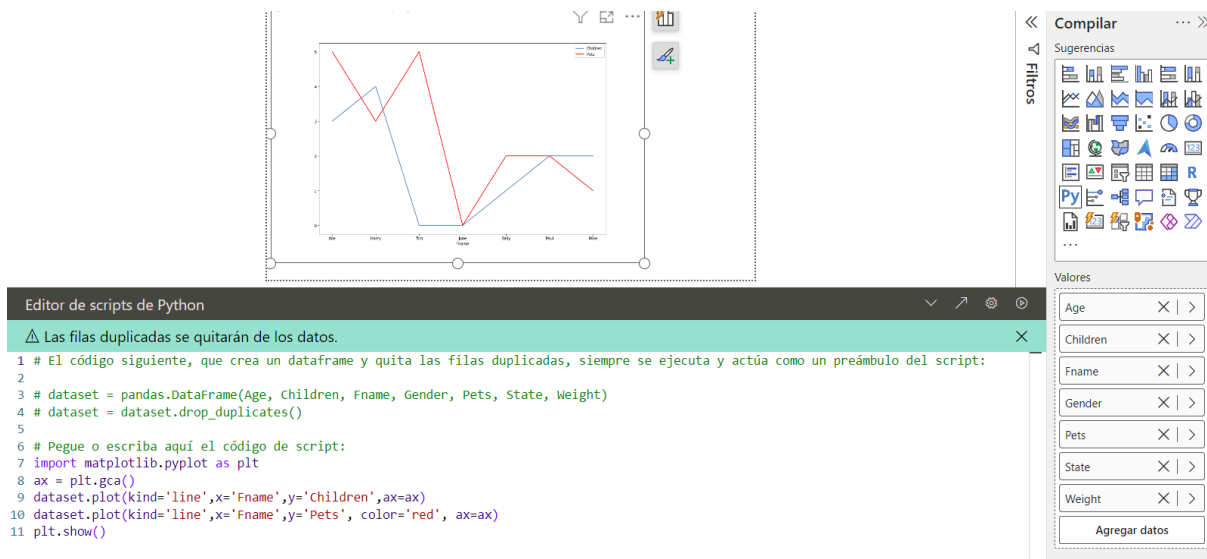
Cancelar

Fname	Age	Weight	Gender	State	Children	Pets
Harry	21	180	M	Washington	4	3
Sally	34	130	F	Oregon	1	2
Paul	42	200	M	California	2	2
Abe	18	140	M	Washington	3	5
June	24	176	F	Nevada	0	0
Mike	80	142	M	Texas	2	1
Tom	22	210	M	Nevada	0	5

Después de esto, desde la vista de informe y desde la ventana de elección de gráficos, habilito el '**objeto visual de Python**', desde el icono 'Py'. Finalmente, se abre una pantalla en la que tenemos un nuevo gráfico y una ventana en la parte inferior titulada '**Editor de scripts de Python**'.



Hago un gráfico de prueba con los datos importados anteriormente.



Ahora paso a conectarme a mi Script de Python. Pruebo de importar todas las tablas a la vez, pero no funciona, puesto que Power BI no admite diccionarios, así que importo una a una las tablas de la base de datos que he usado para mis gráficos que son: 'transaction', 'users', 'products', 'desglose\_transaction' y 'company'.

## Ejercicio 1: UNA VARIABLE NUMÉRICA

### 1A. Conexión a Python y obtención de la tabla 'products'

#### Script de Python

Script

```

conn = connection()

sql_select_Query = "select * from products"
cursor = conn.cursor()
cursor.execute(sql_select_Query)
# .fetchall() = .fetchone() lo uso para leer el resultado de la consulta
records = cursor.fetchall()
print("Total number of rows in table: ", cursor.rowcount)
products = pd.read_sql_query(sql_select_Query, conn)

def close_connection():
    ''' Función que cierra la conexión con la base de datos r2trans'''
  
```

El script se ejecutará con la instalación de Python siguiente:  
 C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.  
 Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar
Cancelar

## 1B. Transformación de datos en Power BI

Elimino del script original de Python, las líneas referentes a la limpieza de los datos, para eliminar el símbolo de \$ porque no funciona correctamente. Y decido hacer estas pequeñas modificaciones en Power BI.

### Navegador

Opciones de presentación ▾

Python [1]
 

☒ products

products

id	product_name	price	colour	weight	wa
1	Direwolf Stannis	\$161.11	#7c7c7c	1	
2	Tarly Stark	\$9.24	#919191	2	
3	duel tourney Lannister	\$171.13	#d8d8d8	15	
4	warden south duel	\$71.89	#111111	3	
5	skywalker ewok	\$171.22	#dbdbdb	32	
6	dooku solo	\$136.60	#c4c4c4	8	
7	north of Casterly	\$63.33	#b7b7b7	6	
8	Winterfell	\$32.37	#383838	14	
9	Winterfell	\$76.40	#b5b5b5	12	
10	Karstark Dorne	\$119.52	#f4f4f4	24	
11	Karstark Dorne	\$49.70	#141414	27	
12	duel Direwolf	\$181.60	#a8a8a8	21	
13	palpatine chewbacca	\$139.59	#2b2b2b	1	
14	Direwolf	\$147.53	#c4c4c4	2	
15	Stannis warden	\$194.29	#dbdbdb	15	
16	the duel warden	\$180.91	#666666	3	
17	skywalker ewok sith	\$91.89	#7c7c7c	32	
18	Karstark warden	\$148.91	#c4c4c4	8	
19	dooku solo	\$60.33	#3f3f3f	6	
20	warden Karstark	\$91.96	#b5b5b5	14	
21	duel Direwolf	\$96.90	#e2e2e2	12	
22	chewbacca mustafar	\$150.02	#fcfcfc	24	
23	riverlands north	\$169.96	#545454	27	

Cargar

Transformar datos

Cancelar

Reemplazo el símbolo \$ por nada. Y hago el cambio de tipo de la columna 'price' a través de la opción 'cambio de tipo' > 'usar configuración regional'

transaction

users

products

desglose

empresa

products\_2

trans\_empresa

country\_age\_count

ventas\_realizadas\_pais

trans\_users

id	product_name	price	colour	weight
1	Direwolf Stannis	\$161.11	#7c7c7c	
2	Tarly Stark	\$9.24	#919191	
3	duel tourney Lannister	\$171.13	#d8d8d8	1
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				
16				
17	skywalker ewok sith	\$91.89	#7c7c7c	3
18	Karstark warden	\$148.91	#c4c4c4	
19	dooku solo	\$60.33	#3f3f3f	
20	warden Karstark	\$91.96	#b5b5b5	1
21	duel Direwolf	\$96.90	#e2e2e2	1
22	chewbacca mustafar	\$150.02	#fcfcfc	2
23	riverlands north	\$169.96	#545454	2

Reemplazar los valores

Reemplace un valor con otro de las columnas seleccionadas.

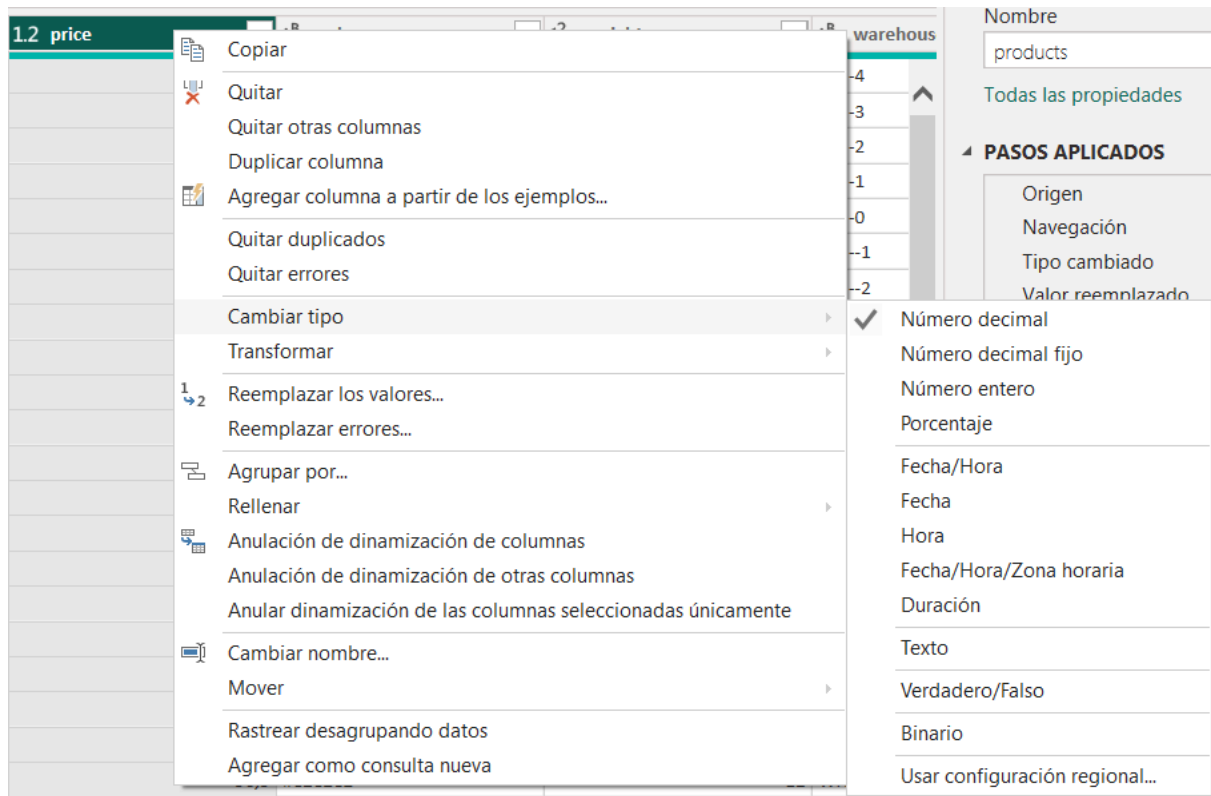
Valor que buscar

\$

Reemplazar con

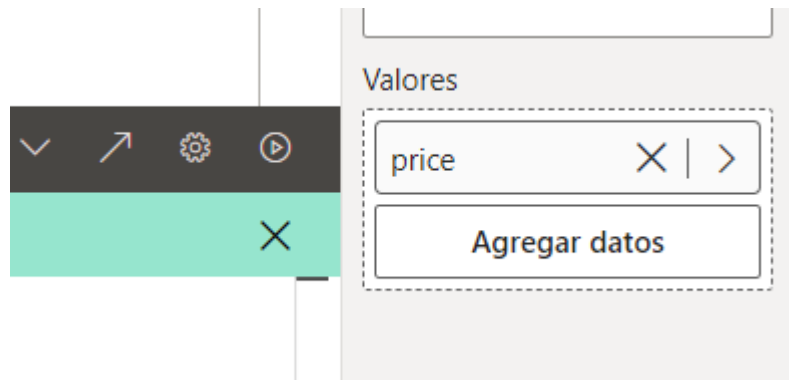
Aceptar

Cancelar



### 1C. Código para crear la figura de Python en Power BI.

Algo muy importante a tener en cuenta para que funcione el código de Python en Power BI, es conseguir anteriormente tener el dataframe a partir del cual realizamos la visualización importado con python. Después se debe introducir los campos que se va a usar en la tabla en la sección 'Valores'

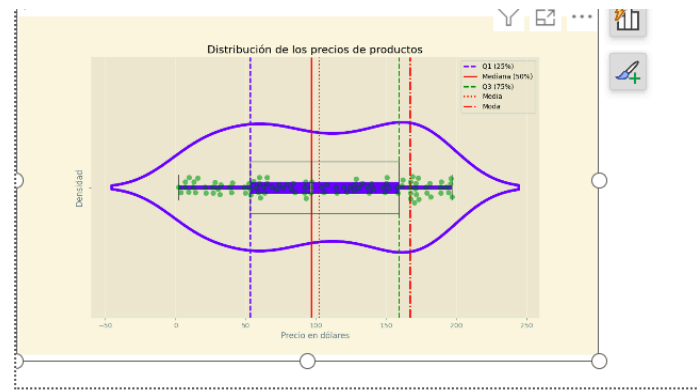


Power BI, va a reconocer estos datos y renombrar con 'dataset', por eso es fundamental, cambiar el nombre de 'dataset' por el nombre que se está utilizando en el código. O al revés, usar dataset en el código que crea la figura. Esto lo hago en todas las figuras.

```

10
11 # cambio el nombre del dataset que genera Power BI automáticamente por el nombre de mi tabla
12 products = dataset
13 # Calculo las medidas de estadística descriptivas generales
14 price_stats = round(products['price'].describe(), 2)
15
16 ## Añado algunas medidas más: Mediana, Moda, Rango, Varianza
17 price_median = round(products['price'].median(), 2)
18 price_range = np.ptp(products['price'])
19 price_modo = multimoda(products['price']).actua('int')

```



Editor de scripts de Python

⚠ Las filas duplicadas se quitarán de los datos.

```
9 from statistics import multimode # para calcular la moda
10
11 # cambio el nombre del dataset que genera Power BI automáticamente por el nombre de mi tabla
12 products = dataset
13 # Calculo las medidas de estadística descriptivas generales
14 price_stats = round(products['price'].describe(), 2)
15
16 ## Añado algunas medidas más: Mediana, Moda, Rango, Varianza
17 price_median = round(products['price'].median(), 2)
18 price_range = np.ptp(products['price'])
19 price_mode = multimode(products['price'].astype('int'))
20 price_varianza = round(np.var(products['price'], ddof=1), 2)
21
```

## Ejercicio 2: DOS VARIABLES NUMÉRICAS

### 2A. Conexión a Python y obtención de la tabla 'products\_2'

#### Script de Python

Script

```
def close_connection():
    ''' Función que cierra la conexión con la base de datos r2trans'''
    # .close() cierra la conexión a la base de datos.
    conn.close()
    print("MySQL connection is closed")

close = close_connection()

cantidad = desglose.groupby(desglose['product_id']).count().reset_index()
cantidad = cantidad.rename(columns= {"id": "quantity", 'product_id': 'id'})
products_2 = pd.merge(products, cantidad, on=None, how='inner')
```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

Cancelar

### 2B. Transformación de los datos en Power BI

En esta ocasión a pesar de que he llamado a varias tablas con el código python, solo necesito la última tabla que es 'products\_2' y es la que cargo en el informe. Hago las mismas transformaciones que en el caso anterior.

## Navegador

Opciones de presentación ▾

- Python [4]
  - ☐ cantidad
  - ☐ desglose
  - ☐ products
  - ☒ products\_2

## products\_2

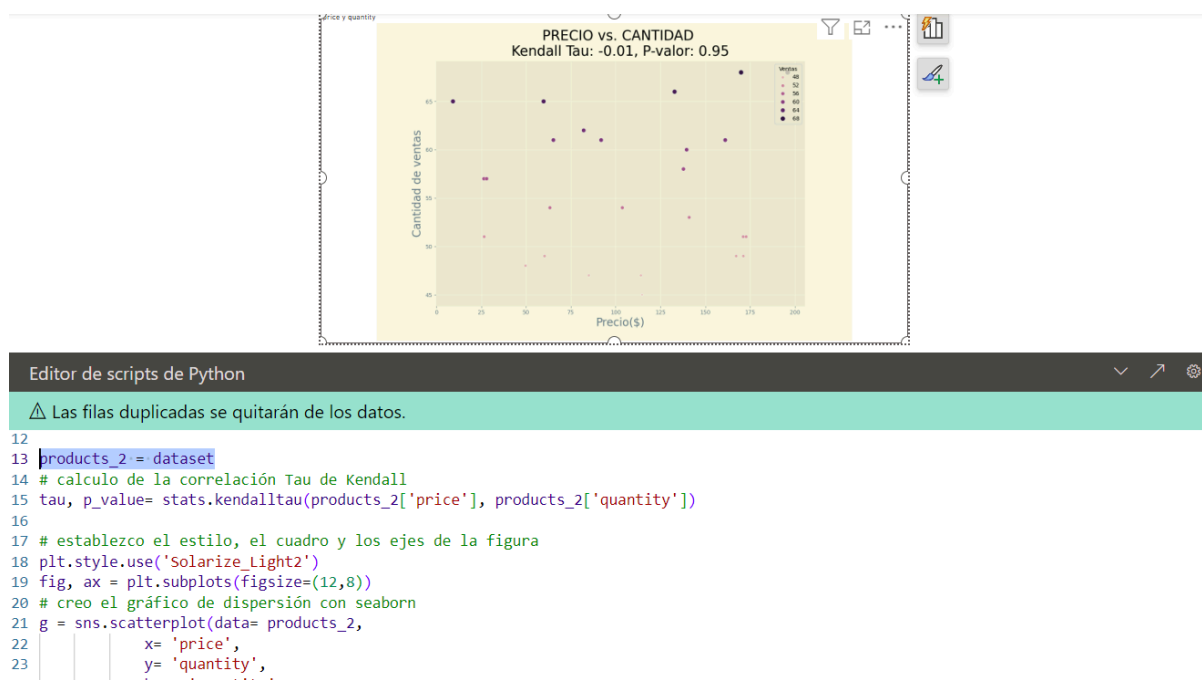
id	product_name	price	colour	weight	wa
1	Direwolf Stannis	\$161.11	#7c7c7c	1	
2	Tarly Stark	\$9.24	#919191	2	
3	duel tourney Lannister	\$171.13	#d8d8d8	15	
5	skywalker ewok	\$171.22	#dbdbdb	32	
7	north of Casterly	\$63.33	#b7b7b7	6	
11	Karstark Dorne	\$49.70	#141414	27	
13	palpatine chewbacca	\$139.59	#2b2b2b	1	
17	skywalker ewok sith	\$91.89	#7c7c7c	32	
19	dooku solo	\$60.33	#3f3f3f	6	
23	riverlands north	\$169.96	#545454	27	
29	Tully maester Tarly	\$167.20	#111111	32	
31	Lannister	\$85.02	#3f3f3f	6	
37	Direwolf Littlefinger	\$26.66	#aaaaaa	1	
41	Lannister Barratheon Direwolf	\$141.01	#fcfcfc	32	
43	duel	\$59.80	#5b5b5b	6	
47	Tully	\$82.15	#919191	27	
53	kingsblood Littlefinger the	\$137.81	#3a3a3a	32	
59	Direwolf Stannis	\$114.77	#bcbcbc	27	
61	Winterfell Lannister	\$28.01	#848484	1	
67	Winterfell	\$195.94	#1c1c1c	6	
71	Tully Dorne	\$103.73	#424242	27	
73	Dorne bastard	\$114.09	#848484	1	
79	Direwolf riverlands the	\$132.86	#b2b2b2	6	

Cargar

Transformar datos

Cancelar

## 2C. Visualización dos variables numéricas y código python



## Ejercicio 3: UNA VARIABLE CATEGÓRICA

### 3A. Conexión a Python y obtención de la tabla 'company'

#### Script de Python

Script

```
return connection

conn = connection()

sql_select_Query = "select * from company"
cursor = conn.cursor()
cursor.execute(sql_select_Query)
# .fetchall() = .fetchone() lo uso para leer el resultado de la consulta
records = cursor.fetchall()
print("Total number of rows in table: ", cursor.rowcount)
empresa = pd.read_sql_query(sql_select_Query, conn)

def close_connection():
```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

Cancelar

### 3B. No transformaciones necesarias

#### Navegador

Opciones de presentación ▾

- Python [1]
  - ☒ empresa

empresa

company_id	company_name	phone	email
b-2222	Ac Fermentum Incorporated	06 85 56 52 33	donec.por
b-2226	Magna A Neque Industries	04 14 44 64 62	risus.done
b-2230	Fusce Corp.	08 14 97 58 85	risus@pro
b-2234	Convallis In Incorporated	06 66 57 29 50	mauris.ut
b-2238	Ante Iaculis Nec Foundation	08 23 04 99 53	sed.dictum
b-2242	Donec Ltd	01 25 51 37 37	at.iaculis@
b-2246	Sed Nunc Ltd	02 62 64 73 48	nibh@yah
b-2250	Amet Nulla Donec Corporation	07 15 25 14 74	mattis.inte
b-2254	Nascetur Ridiculus Mus Inc.	06 26 87 61 84	suspendis
b-2258	Vestibulum Lorem PC	02 02 87 33 40	aenean.mi
b-2262	Gravida Sagittis LLP	03 81 28 33 97	turpis.vita
b-2266	Mus Aenean Eget Foundation	06 25 15 52 43	mi.duis@h
b-2270	Dis Parturient Institute	05 36 29 78 74	purus@pri
b-2274	Sed LLC	01 63 16 26 52	at@outloc
b-2278	Arcu LLP	06 46 04 41 45	dui@aol.c
b-2282	Pretium Neque Corp.	07 77 48 55 28	eleifend.n
b-2286	Fringilla LLC	08 29 15 93 57	gravida@f
b-2290	Quisque Libero LLC	01 45 48 71 11	sapien.mo
b-2294	Auctor Mauris Vel LLP	08 09 28 74 14	nec.tempu
b-2298	Elit Etiam Laoreet Associates	07 69 74 17 45	ultrices@g
b-2302	Nunc Interdum Incorporated	05 18 15 48 13	non@outl
b-2306	Augue Foundation	06 88 43 15 63	mauris@y
b-2310	Non Magna LLC	06 71 73 13 17	nisl.quisqu

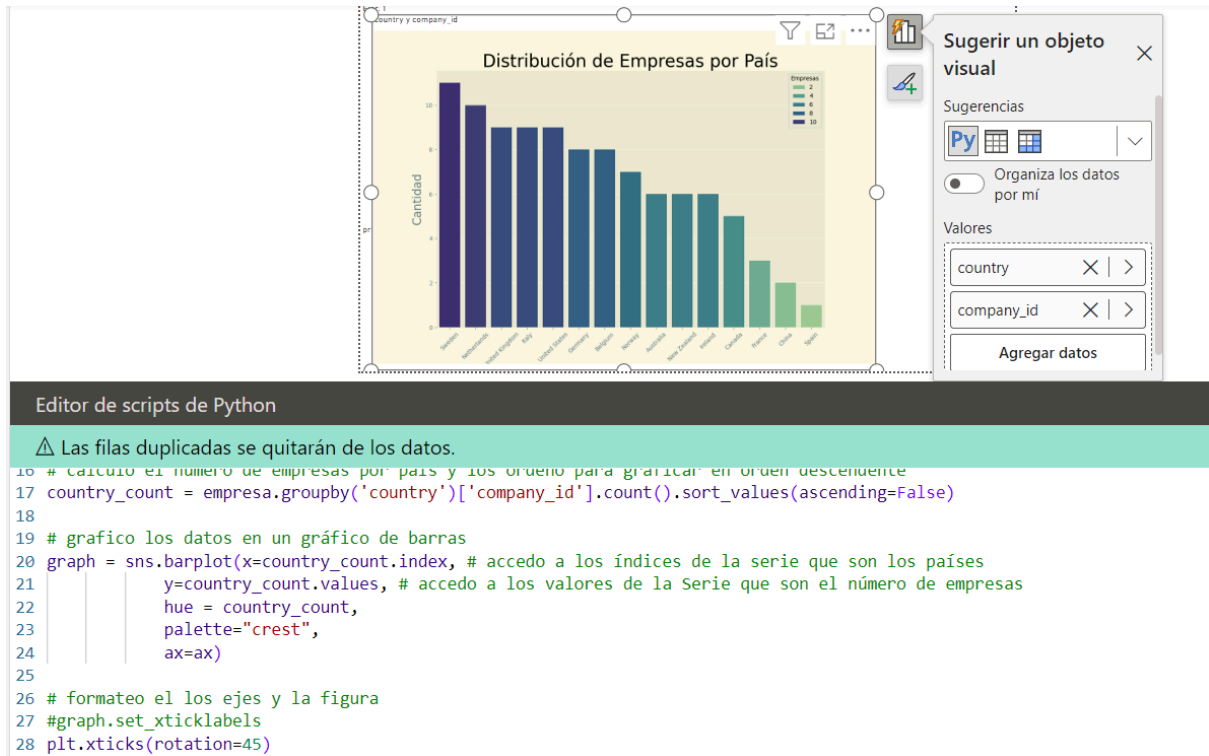
Cargar

Transformar datos

Cancelar



### 3C. Visualización de una variable categórica y código python



## Ejercicio 4 : UNA VARIABLE CATEGÓRICA y UNA NUMÉRICA

### 4A. Conexión a Python y obtención de la tabla 'trans\_empresa'

#### Script de Python

Script

```

def close_connection():
    ''' Función que cierra la conexión con la base de datos r2trans'''
    # .close() cierra la conexión a la base de datos.
    conn.close()
    print("MySQL connection is closed")

close = close_connection()

transaction = transaction.rename(columns= {"business_id": "company_id"} )
trans_empresa = pd.merge(transaction, empresa, on='company_id',how='inner')
  
```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

Cancelar

## 4B. Transformación de datos en PowerBI.

En la tabla importada 'trans\_empresa' el campo 'amount' no lo ha importado correctamente. Aquí necesito ir a Power Query para ver que está ocurriendo. Y si elimino el último paso que realiza de 'Tipo Cambiado', consigo que los datos de 'amount' estén correctamente.

### Navegador

trans\_empresa  
Vista previa descargada el domingo

	card_id	company_id	timestamp	amount	declined
3B02	CcU-2938	b-2362	28/08/2021 23:42:24	46692	
3128	CcU-4219	b-2302	26/07/2021 7:29:18	4953	
i4A5	CcU-2987	b-2250	06/01/2022 21:25:27	9261	
38AE	CcU-3743	b-2618	26/01/2022 2:07:14	39418	
EBA9	CcU-2959	b-2346	26/10/2021 23:00:01	27993	
1109F	CcU-3225	b-2386	28/06/2021 21:11:42	34087	
38E2	CcU-3071	b-2298	11/05/2021 20:40:06	30305	
3539	CcU-4359	b-2302	26/02/2022 20:33:54	43049	
14B5	CcU-3141	b-2338	04/03/2022 14:54:35	28881	
7598	CcU-3309	b-2434	10/04/2021 20:58:41	10344	
4853	CcU-3435	b-2506	02/02/2022 7:29:36	42869	
735A	CcU-2959	b-2234	17/04/2021 5:30:17	25247	
FA0A	CcU-3701	b-2618	01/11/2021 1:02:11	44711	
13D2	CcU-2959	b-2346	07/12/2021 9:30:38	46531	
05DC	CcU-2938	b-2222	07/07/2021 17:43:16	29357	
8037	CcU-3155	b-2346	16/05/2021 21:00:28	2785	
2179	CcU-3981	b-2362	14/07/2021 20:55:48	1572	
BB9A	CcU-4366	b-2302	09/06/2021 6:04:14	17201	
0A5F	CcU-3407	b-2490	02/04/2021 5:17:46	34888	
ED54	CcU-4520	b-2302	29/12/2021 20:38:23	1797	
1E80	CcU-3197	b-2370	02/06/2021 4:10:57	5009	
218A	CcU-3365	b-2466	30/10/2021 13:42:44	8058	

Cargar Transformar datos Cancelar

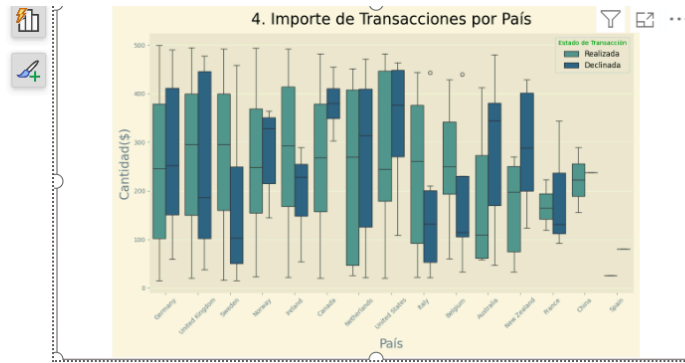
fx = Origen[{"Name":"trans\_empresa"}][Value]

	A <sub>id</sub>	A <sub>card_id</sub>	A <sub>company_id</sub>	A <sub>timestamp</sub>	A <sub>amount</sub>	A <sub>declined</sub>
1	02C6201E-D90A-1859-B4EE-88D2986D3...	CcU-2938	b-2362	2021-08-28 23:42:24	466.92	0
2	0466A42E-47CF-8D24-FD01-C0B689713...	CcU-4219	b-2302	2021-07-26 07:29:18	49.53	0
3	063FBA79-99EC-66FB-29F7-25726D1764...	CcU-2987	b-2250	2022-01-06 21:25:27	92.61	0
4	0668296C-CD89-A883-76BC-2E4C44F8C...	CcU-3743	b-2618	2022-01-26 02:07:14	394.18	0
5	06CD9AA5-9B42-D684-DDDD-A5E394FE...	CcU-2959	b-2346	2021-10-26 23:00:01	279.93	0
6	07A46D48-31A3-7E87-65B9-0DA902AD1...	CcU-3225	b-2386	2021-06-28 21:11:42	340.87	1
7	09DE92CE-6F27-2BB7-13B5-938582B3B...	CcU-3071	b-2298	2021-05-11 20:40:06	303.05	1
8	0A476ED9-0C13-1962-F87B-D3563924B...	CcU-4359	b-2302	2022-02-26 20:33:54	430.49	0
9	0BEB80B7-9D66-1707-CE4B-9DC7E7191...	CcU-3141	b-2338	2022-03-04 14:54:35	288.81	1
10	0C7C3A33-9947-3BC1-846D-7BE3D0D17...	CcU-3309	b-2434	2021-04-10 20:58:41	103.44	1

**PROPIEDADES**  
Nombre  
trans\_empresa  
Todas las propiedades

**PASOS APLICADOS**  
Origen  
X Navegación

## 4C. Visualización una variable categórica y una numérica



Editor de scripts de Python

⚠ Las filas duplicadas se quitarán de los datos.

```

33     nue= declined , # sirve para graficar en funcion de si las transacciones estan declinadas o no
34     palette="crest",
35     ax=ax)
36
37 # Establezco los titulos del gráfico y ejes
38 plt.xticks(rotation= 45)
39 ax.set_title("4. Importe de Transacciones por País", fontsize = 25, color= 'black', pad=15)
40 plt.xlabel('País', fontsize= 20)
41 plt.ylabel("Cantidad($)", fontsize = 20)
42 # Cambiar etiquetas de la leyenda
43 handles, labels = ax.get_legend_handles_labels()
44 legend= ax.legend(handles, ['Realizada', 'Declinada'], title="Estado de Transacción", fontsize=12)
45 plt.setp(legend.get_texts(), color='black') # Cambiar color del texto de la leyenda

```

## Ejercicio 5 : DOS VARIABLES CATEGÓRICAS

### 5A. Conexión a Python y obtención de la tabla 'country\_age\_count'

#### Script de Python

Script

```

# Calculo la diferencia entre fecha actual y nacimiento
# me devuelve un dato tipo 'Timedelta'
# con dt.days obtengo los días como números enteros y lo convierto en años
age = ((actual_date - birthday_date).dt.days / 365.25).astype('int')
age = pd.DataFrame(age).reset_index()
age = age.rename(columns= {'index': 'id', 'birth_date': 'age'})
age['id'] = age['id'].astype('str')
age['age'] = age['age'].astype('int')
users = pd.merge(users,age, on='id', how='left')

# Creo la función necesaria para categorizar las edades de los usuarios
def cat_edad(age):

```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

Cancelar

5B. Transformación de datos en PowerBI.

Navegador

Opciones de presentación ▾

Python [3]

☐

age

☒country\_age\_count

☐users

country\_age\_count

Vista previa descargada el domingo

country	cat_edad	count	percentage
Canada	25-30	19	253
Canada	31-35	20	267
Canada	36-40	20	267
Canada	40-45	16	213
United Kingdom	25-30	15	300
United Kingdom	31-35	13	260
United Kingdom	36-40	11	220
United Kingdom	40-45	11	220
United States	25-30	45	300
United States	31-35	39	260
United States	36-40	36	240
United States	40-45	30	200

Cargar

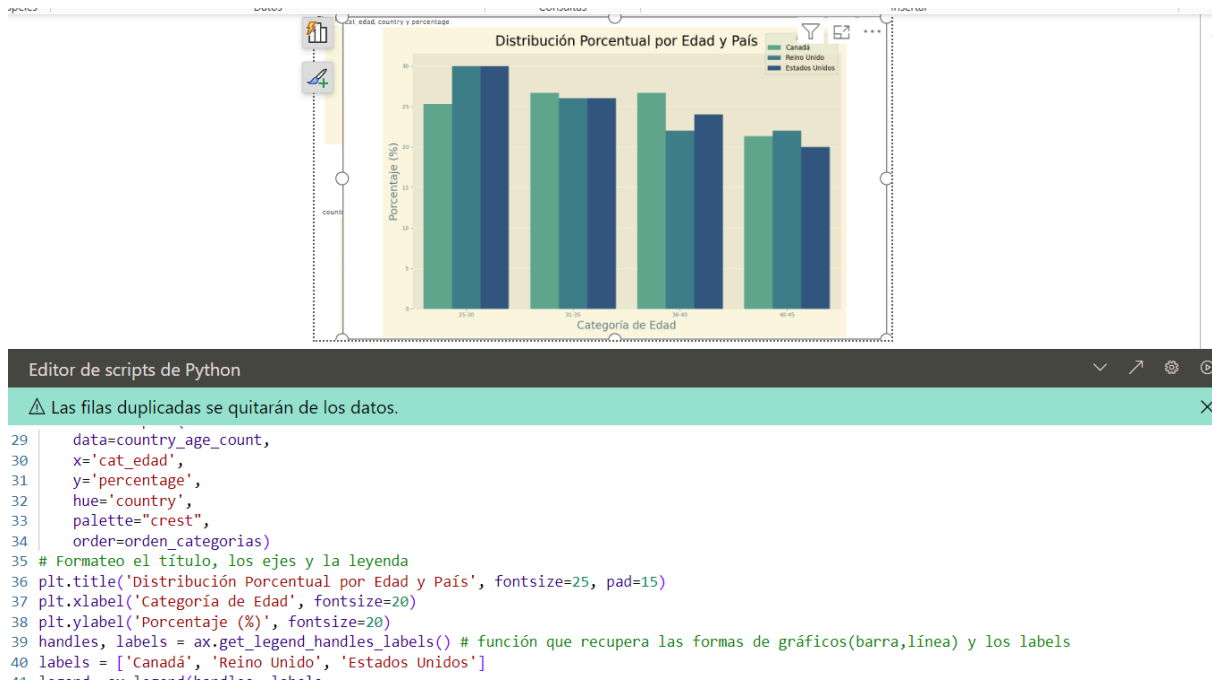
Transformar datos

Cancelar

Consigo importar la tabla pero en esta ocasión no me devuelve los percentage correctamente. Para conseguir el dato correcto, divido la columna por 10.

country	cat_edad	count	percentage
Canada	25-30	19	25,3
Canada	31-35	20	26,7
Canada	36-40	20	26,7
Canada	40-45	16	21,3
United Kingdom	25-30	15	30
United Kingdom	31-35	13	26
United Kingdom	36-40	11	22
United Kingdom	40-45	11	22
United States	25-30	45	30
United States	31-35	39	26
United States	36-40	36	24
United States	40-45	30	20

## 5C. Visualización dos variables categóricas



## Ejercicio 6 : TRES VARIABLES

### 6A. Conexión a Python y obtención de la tabla 'ventas\_realizadas\_pais'

#### Script de Python

Script

```

trans_empresa['declined'] = trans_empresa['declined'].astype('int')
# creo el filtro de transacciones realizadas
filtro_realizadas = trans_empresa['declined'] == 0
# aplico el filtro a la table 'trans_empresa'
trans_realizadas = trans_empresa[filtro_realizadas]
# sumo las ventas totales por país
ventas_totales_pais = trans_realizadas.groupby('country', observed=True)['amount'].sum().reset_index()
# cuento el número de transacciones realizadas por país
trans_pais_realizadas = trans_realizadas.groupby('country', observed=True)['declined'].count().reset_index()
# uno estos dos df para poder graficar
ventas_realizadas_pais= pd.merge(ventas_totales_pais, trans_pais_realizadas, on = 'country', how='left')
ventas_realizadas_pais= ventas_realizadas_pais.rename(columns={'declined':'realizadas'})

```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

Cancelar

## 6B. Transformación de datos en PowerBI.

### Navegador

Opciones de presentación ▾

Python [7]

☐ empresa

☐ trans\_empresa

☐ trans\_pais\_realizadas

☐ trans\_realizadas

☐ transaction

☒ ventas\_realizadas\_pais

☐ ventas\_totales\_pais

ventas\_realizadas\_pais

Vista previa descargada el lunes

country	amount	realizadas
Australia	106399	6
Belgium	204174	8
Canada	1493069	57
China	44448	2
France	50823	3
Germany	2688855	111
Ireland	1600622	56
Italy	219008	9
Netherlands	24094	10
New Zealand	100237	6
Norway	1531801	61
Spain	2622	1
Sweden	1881345	68
United Kingdom	2527438	93
United States	258778	9

Cargar

Transformar datos

Cancelar

Consigo importar la tabla, sin embargo, de nuevo la columna 'amount' no está correctamente. Para conseguir el dato correcto, divido la columna por 100.

fx

= Table.TransformColumns("#Tipo cambiado", {"amount", each \_ / 100, type number})

	country	amount	realizadas
1	Australia	1063,99	6
2	Belgium	2041,74	8
3	Canada	14930,69	57
4	China	444,48	2
5	France	508,23	3
6	Germany	26888,55	111
7	Ireland	16006,22	56
8	Italy	2190,08	9
9	Netherlands	240,94	10
10	New Zealand	1002,37	6
11	Norway	15318,01	61
12	Spain	26,22	1
13	Sweden	18813,45	68
14	United Kingdom	25274,38	93
15	United States	2587,78	9

PROPIEDADES

Nombre

ventas\_realizadas\_pais

Todas las propiedades

PASOS APLICADOS

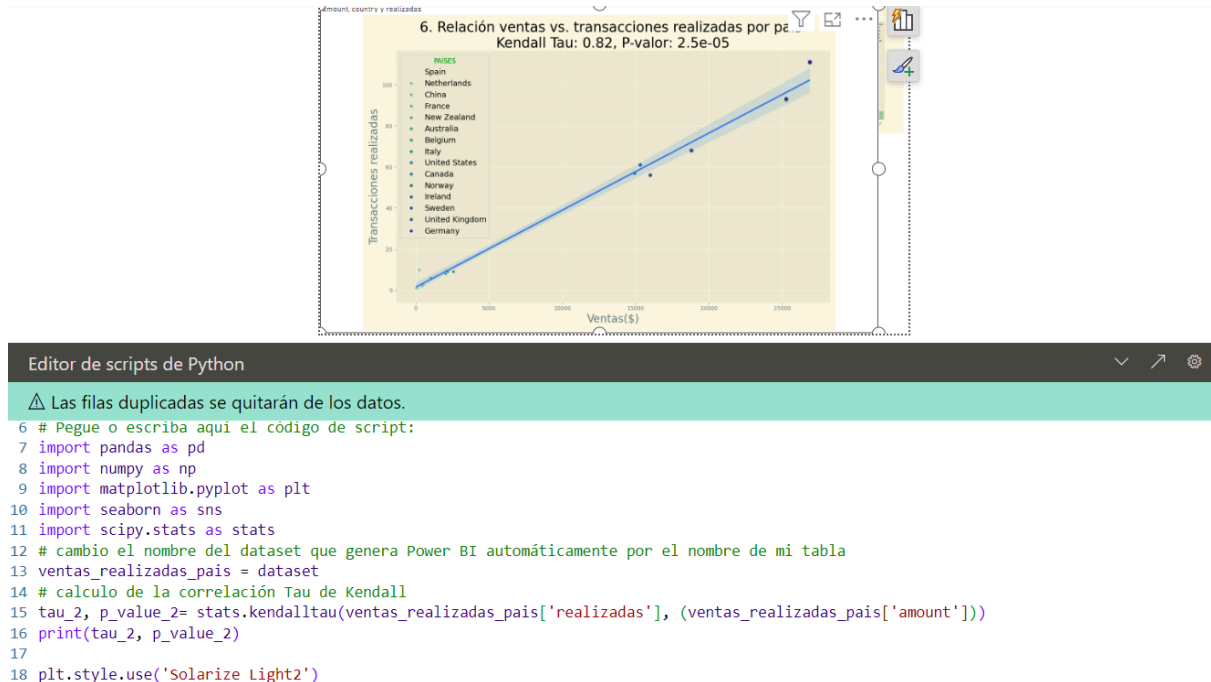
Origen

Navegación

Tipo cambiado

Columna dividida

## 6C. Visualización dos tres variables



## Ejercicio 7 : PAIRPLOT

### 7A. Conexión a Python y obtención de la tabla 'trans\_users'

En este código, decido eliminar la parte del código donde calculo la edad del usuario que no tenía valor en ese campo, y dejo solamente la instrucción para rellenar los valores nulos de la columna 'age' con 42 que es la edad que he calculado anteriormente.

#### Script de Python

Script

```
users = pd.merge(users, age, on='id', how='left')
users = users.rename(columns= {'id': 'user_id'})
trans_users = pd.merge(transaction, users, on='user_id', how='left')
trans_users['amount'] = pd.to_numeric(trans_users['amount'], errors='coerce')
trans_users_mean = trans_users.groupby('user_id')['amount'].mean().reset_index()
trans_users_mean = round(trans_users_mean, 2)
trans_users_mean.rename(columns= {'amount': 'average_amount'}, inplace=True)
trans_users = pd.merge(trans_users, trans_users_mean, on='user_id', how='left')

# hartman tiene 42 años así que voy a sustituir los valores nulos en la edad por '42'
trans_users.fillna({'age': 42.0}, inplace=True)
```

El script se ejecutará con la instalación de Python siguiente:

C:\USERS\RONCH\APPDATA\LOCAL\PROGRAMS\PYTHON\PYTHON312.

Para establecer la configuración y cambiar la instalación de Python que quiere ejecutar, vaya a Opciones y configuración.

Aceptar

Cancelar





## Informe final en Power BI

