

Sprint 8.02 - Power BI con Python

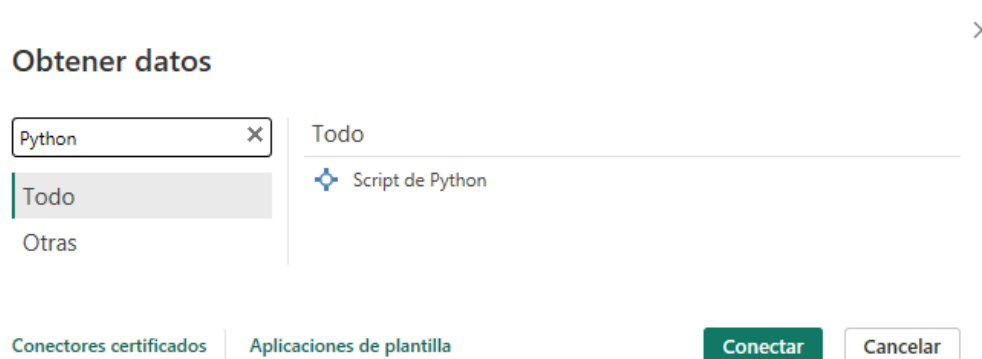
Esta labor consiste en la elaboración de un informe de Power BI, aprovechando las capacidades analíticas de Python. Se utilizarán los scripts de Python creados previamente en la Tarea 1 para generar visualizaciones personalizadas con las bibliotecas Seaborn y Matplotlib. Estas visualizaciones estarán integradas en el informe de Power BI para ofrecer una comprensión más profunda de la capacidad del lenguaje de programación en la herramienta Power BI.

Nivel 1

Los 7 ejercicios del nivel 1 de la tarea 01

Obtención de datos.

Busco y selecciono la opción “Script de Python”.



Uso el mismo script con el que generé la conexión a MySQL y con el que convertí las tablas de mi base de datos a los dataframes que usaría para las visualizaciones. Agrego al script las conversiones de columnas que considere pertinentes para mantener una concordancia de los datos, y también los merge entre dataframes que haya creado en Python para las distintas visualizaciones generadas.



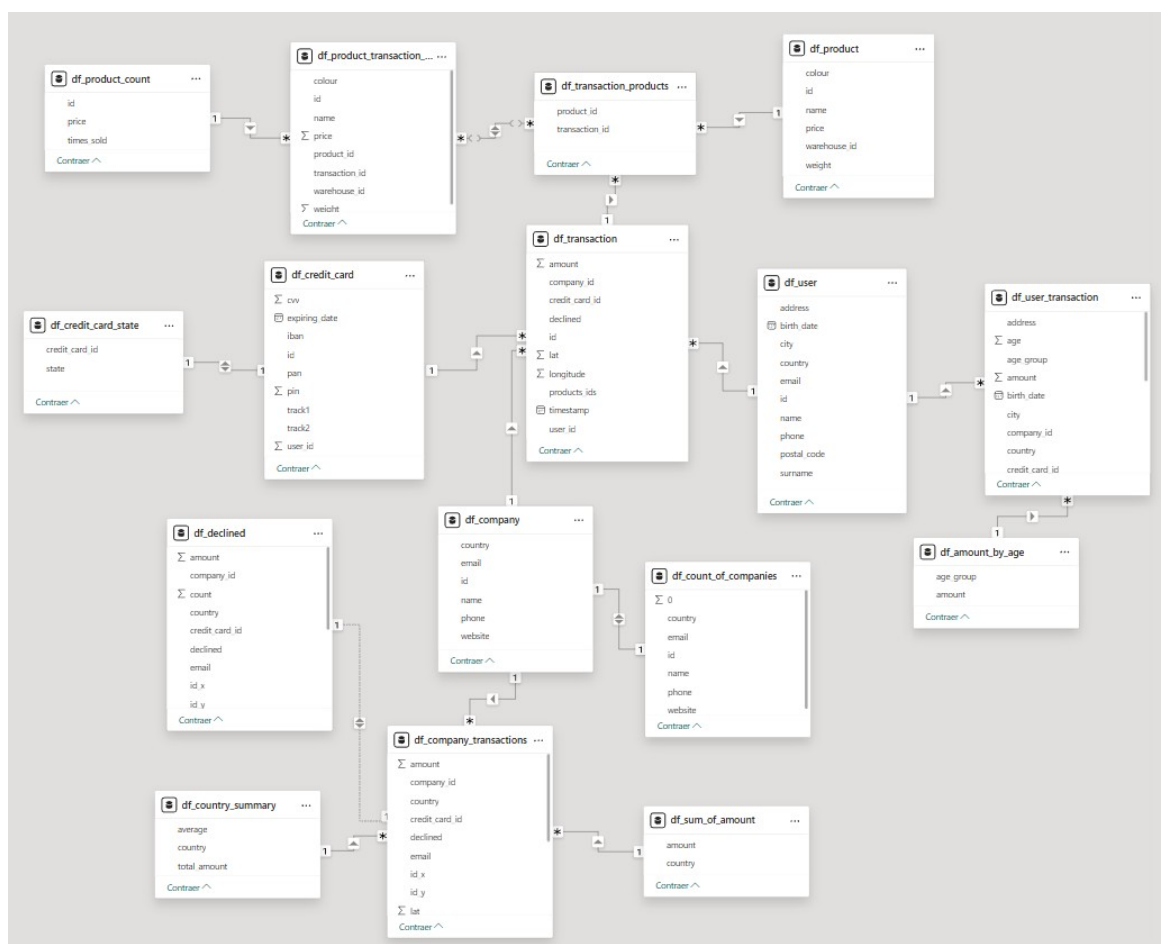
Una vez se inicializo el script en la previsualización veo que los datos de algunas columnas necesitan ser cambiados, como price y weight, y amount de los dataframes df_product y df_transaction respectivamente. Entonces en la transformación de datos primero tendré que convertir las columnas a texto y remplazar los puntos (.) por comas (,), para posteriormente poder tener dichas columnas como números decimales.

df_product

id	name	price	colour	weight	warel
1	Direwolf Stannis	16111	#7c7c7c	10	W
2	Tarly Stark	924	#919191	20	W
3	duel tourney Lannister	17113	#d8d8d8	15	W
4	warden south duel	7189	#111111	30	W
5	skywalker ewok	17122	#dbdbdb	32	W

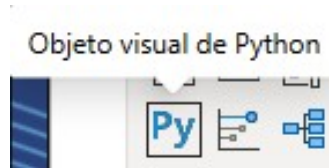
Cargar Transformar datos Cancelar

Verifico las relaciones entre los dataframes, y en caso de que no exista dicha relación creo las que vayan a ser necesarias.



Creación de las visualizaciones.

Para poder generar las mismas visualizaciones que tengo en Python en Power BI selecciono la opción Objeto visual de Python.



Finalmente, selecciono las columnas que utilizaré para la visualización, y en cada visualización cambio en nombre del dataframe que esté utilizando a dataset, ejecuto el código y está todo listo.

```
Editor de scripts de Python

⚠ Las filas duplicadas se quitarán de los datos.

1 # El código siguiente, que crea un dataframe y quita las filas duplicadas, siempre se ejecuta y actúa como un preámbulo del script:
2
3 # dataset = pandas.DataFrame(amount, id)
4 # dataset = dataset.drop_duplicates()
5
6 # Pegue o escriba aquí el código de script:
7 import seaborn as sns
8 import matplotlib.pyplot as plt
9
10 plt.figure(figsize=(16, 8))
11 sns.set_theme(style='darkgrid', palette='inferno')
12 ax = sns.histplot(dataset['amount'], bins=10)
13 plt.title('Histogram of amount', size=24)
14 plt.xlabel('Amount €', size=18)
15 plt.ylabel('Count', size=18)
16 plt.tick_params(axis='both', labelsize=12)
17 plt.show()
```

Captura de pantalla con las visualizaciones del nivel 1 en Power BI.

