



GORKA BONALS SASTRE

Element de qualificació

Ponderació

calculada Qualificació Gamma Percentatge Retroacció

Contribució al total del curs

Data Analytics

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció
-------------------------	------------	-----------	--------------	-------	-------------	------------

Hola Gorka,

Agrego la media entre tu sprint 1 y el 2 anterior.

TASCA

 [Tasca S2.01. Nocións bàsiques SQL](#)

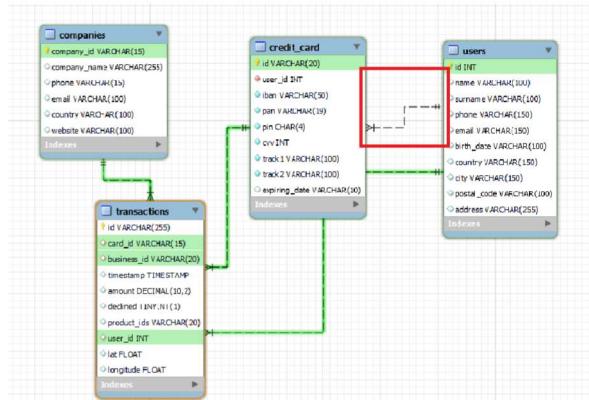
- ✓ 8.75 0-10 87.50 %

Un saludo,

Alana

Element de qualificació	Ponderació	Qualificació	Gamma	Percentatge	Retroacció	Contribució al total del curs
	calculada					
					Hola Gorka,	
					Muy buen trabajo!	
					El ejercicio 1 está muy bien. De la creación de la tabla nueva me hubiese gustado ver el modelo final de lo que tienes en la bbdd antes de iniciar con las consultas, pero bueno con el código está todo bien!	
					En el ejercicio 2 está súper bien! exacto, si usas una condición de where o de limit ya no es necesario modificar el safe_update.	
					En el ejercicio 3 muy bien en buscar una forma para automatizar el proceso! Cuando vas a insertar datos normalmente (cuando no le dices not null) los que no nombras antes del values les asigna NULL por defecto. Este ejercicio está muy muy bien realizado!	
					<pre>93 • USE transactions; 94 • INSERT INTO company 95 (id,company_name,phone,email,country,website) 96 VALUES ('b-9999',NULL,NULL,NULL,NULL,NULL); 97 • INSERT INTO credit_card 98 (id,iban,pan,pin,cvv,expiring_date) 99 VALUES ('CcU-9999','','',',-1,-1,NULL');</pre>	
					El ejercicio 4 es perfecto!	
					El ejercicio 1 del nivel 2 es una buena solución jajajajaja te muestra el código de como se debe hacer.	
					El ejercicio 2 está perfecto!	
TASCA					Solo toma en cuenta que no todas las app tienen la misma opción, hay unas más básicas que no tienen esta opción, mientras entiendas súper bien lo que está sucediendo.	
 Tasca S3.01. Manipulació de taules	✓ 9.50	0-10	95.00 %		El ejercicio 3 está perfecto!	
					El ejercicio 1 del nivel 3 hubiese estado genial que mostraras la imagen del modelo de tu bbdd para poder comparar de forma más sencilla lo que tienes y los cambios a realizar.	
					¿por qué después de que haces el cambio que se te pide vuelves al date?	
					<pre>235 • SET SQL_SAFE_UPDATES=0; 236 • UPDATE credit_card SET expiring_date = DATE_FORMAT(expiring_date_old, '%m/%d/%Y'); 237 • ALTER TABLE credit_card DROP expiring_date_old; 238 • SET SQL_SAFE_UPDATES=1;</pre>	
					Todos los cambios estuvieron bien, excepto que faltó el cambio del nombre de la tabla user:	
					Un saludo,	
					Alana	

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	Contribució al total del curs
TASCA Tasca S4.01. Creació de Base de Dades	-	✓ 9.50	0-10	95.00 %	Hola Gorka, Buen trabajo! Ok ok arrancas con lo que conocías de transactions, es una buena manera de resolver la creación de las tablas. Las tablas users_ca y uk no es que molesten pero sí que ocupan espacio en memoria, entonces si no las usas, que no tiene sentido usarlas porque ya tienes toda esa info en users, elimínalas!	Exactamente, si tuvieses un número que ya no pueda procesar excel tendrías problemas, la mejor opción es cargando los archivos directamente en el sql. Hubiese sido interesante que continuaras con la carga de los archivos. La idea es que todos esos cambios que haces en excel los resuelvas en el sq!l Si se supone debes buscar un modelo estrella ¿crees que está bien esta conexión entre user y credit_card?	-



La consulta si que es correcta! Toma en cuenta que solo el sprint 2 era el que obligatoriamente debían usar subconsultas, de allí en adelante como creas mas conveniente!

El ejercicio 2 también es correcto solo toma en cuenta que cuando hablamos de montos monetarios se manejan con dos decimales únicamente.

En el ejercicio 1 del nivel 2 es una buena opción hacer una vista en vez de una tabla. Muy bien ejecutado!

El ejercicio 1 del nivel 3, no las puedes relacionar directamente porque product_ids en transaction no son valores individuales de los productos sino que puedes tener mas valores en la misma celda.

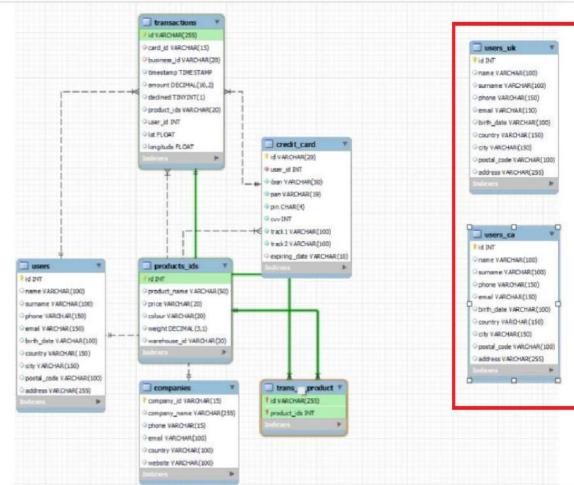
La idea en la programación es siempre optimizar y generalizar procesos, si en este caso, de la cantidad de productos, no tuvieses 4 sino 100 entonces ¿harías 100 columnas a mano? exacto, no nos serviría si hubiese más de 4 ni sería lo mas óptimo.

```

32     trans_x_product_aux_2 AS
33     (
34         SELECT
35             t.id,
36             TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 1), ',', -1)) AS valor_1,
37             TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 2), ',', -1)) AS valor_2,
38             TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 3), ',', -1)) AS valor_3,
39             TRIM(SUBSTRING_INDEX(SUBSTRING_INDEX(t.product_ids, ',', 4), ',', -1)) AS valor_4
40         FROM trans_x_product_aux_1 t),
41     -- Hem de convertir la consulta anterior en una taula de dues columnes

```

Fijate que los users_ca y uk no haces nada con ellos



La consulta está perfecta!

Muy bien!

Un saludo,

Alana

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	
TASCA  Tasca S7.01. Estructuras de datos y de control	-	✓ 9.00	0-10	90.00 %	Hola Gorka,	Muy buen trabajo! <p>El ejercicio 1 está bastante bien, solo ten cuidado con mantener separadas las cosas: aquí haces 2 funciones donde haces todo en ellas, recibes las entradas, haces cálculos, comparaciones, evalúas errores e imprímes. Siempre se busca optimizar los códigos para que sean escalables. A nivel programático, en vez de tener muchos condicionales, se crean funciones que hagan cada proceso por separado y de esta forma también será mucho más fácil el manejo de errores, que mira que los planteas, pero en muchos lugares a la vez.</p> <p>En los cuadros rojos dentro del código te recalco cosas redundantes. En el condicional del if le dices < 18.5 y luego en el siguiente elif le dices >= 18.5, este ya no lo tienes que colocar nunca contrastará esta condición porque si el imc es < 18.5 se queda en la primera nunca continuará. Lo mismo con el siguiente elif con el imc >= 25.</p> <p>El resto de los colores de los cuadros son cosas que puedes separar en funciones.</p> <pre> In [14]: #La función no té paràmetres d'entrada porque es demanat per pantalla que l'usuari els entri. def ingresar_peso_y_altura(): #El usuario entre el peso y la altura en kg y #en metros respectivamente y si no los entra bien se hacen ciertas acciones. #i) Si el usuario se desplaza y entra el peso y la altura con decimales en vez de un punto, #el programa corrige automáticamente el decimal a punto. #ii) Si el usuario entra texto, en el peso y la altura, da un error y devuelve None, None. #iii) Si el peso y la altura no son valores normales, el programa da error y se devuelve None, None. #Error # peso_floated[entre su peso pa kg-1] # peso_floated(peso_str) #except ValueError: # #Averiguar de decidir a punto # peso_floated(peso_str.replace('.','')) if peso_str==new_peso_str: print("Error! Se ha entrado", peso_str, "como peso que es un texto. Repetir el proceso.") return (None,None) peso_floated(peso_str_new) try: #altura_str==[entre su altura en metros] altura_floated(altura_str) except ValueError: #Averiguar de decidir a punto altura_floated(altura_str.replace('.','')) altura_floated(altura_str_new) if altura_str==new_altura_str: print("Error! Se ha entrado", altura_str,"como altura que es en texto. Repetir el proceso.") return (None,None) altura_floated(altura_str_new) if peso_floated or altura_floated or altura_floated==0: print ("Error! El peso", peso_floated, "y/o la altura", altura_floated, "es", "es negativa o está #Repitir el proceso.",sep="") return (None,None) return (peso_floated,altura_floated) #La función inc, se ha preparado para que solo se pueda ejecutar, tal como dice el enunciado, #apenas que un usuario entre los datos del peso en kg y la altura en metros. Así todos los controles #de error están encapsulados en la función ingresar_peso_y_altura def calcular_imc(): (peso,altura)=ingresar_peso_y_altura() if (peso==None and altura==None): #Pongo imc imc=0 if imc<18.5: clas_imc="Bajo peso (menos de 18.5)" elif imc>18.5 and imc<24.9: clas_imc="Peso normal (18.5-24.9)" elif imc>24.9 and imc<29.9: clas_imc="Sobrepeso (25-29.9)" else: clas_imc="Obesidad (más de 30)" print ("El usuario de peso (peso) kg y altura (altura) = tiene un IMC (imc) kg/m2. Una consideración como: \'{}\'".format(clas_imc)) else: return (None,None) #IMC PRINCIPAL# (imc,clas_imc)=calcular_imc() </pre>	-

El ejercicio 2 está super bien, pero sin embargo siempre se busca optimizar los códigos para que sean escalables. Lo que ya te he comentado antes, la refactorización! Muy bien que creas funciones que hagan todo ese proceso de convertir desde cada unidad de temperatura a todas las demás, pero me dejas por fuera de una función todo el código de abajo que lo llamas principal pero por fuer, entonces cada vez que quieras cambiar una temperatura tendrás que copiar y pegar todo eso para poder ejecutar, no tiene sentido! Lo bonito de la programación es que puedas reutilizar. Muy bien con la impresión de errores, aunque lo tienes esparcido por todos lados.

Element de qualificació	calculada	Qualificació	Gamma	Percentatge	Retroacció
					<pre>##### ##Codi Principal##### T_nombres=[“Celcius”, “Fahrenheit”, “Kelvin”, “Rankine”, “Reamur”] T_unidades=[“C”, “F”, “K”, “Ra”, “Re”] (grados_entrada,dig_T_entrada, list_convers)=ingresar_temperatura_y_dos_conversores() list_T_nom_unid=list_T_nom_and_uni(T_nombres,T_unidades) asortido=[[“celcius”, ‘C’], [“fahrenheit”, ‘F’], [“kelvin”, ‘K’], [“rankine”, ‘Ra’], [“reamur”, ‘Re’]]</pre> <p>Si tots els valors són correctes, es mostra el missatge "Los datos han sido ingresados correctamente". Si hi ha un error, es mostra el missatge "Por favor, ingrese los datos nuevamente".</p> <pre> si tots els valors són correctes print("Los datos han sido ingresados correctamente") #print("Los datos han sido ingresados correctamente") else print("Por favor, ingrese los datos nuevamente")</pre> <pre> grados_conv_1=grados_entrada_a_otra(grados_entrada,dig_T_entrada, list_convers[0]) print("grados_entrada "+str(grados_entrada)+" "+str(list_T_nom_unid[dig_T_entrada-1][1])+" "+str(grados_entrada_a_otra(grados_entrada,dig_T_entrada, list_convers[0]))) grados_conv_2=grados_entrada_a_otra(grados_entrada,dig_T_entrada, list_convers[1]) print("grados_entrada "+str(grados_entrada)+" "+str(list_T_nom_unid[dig_T_entrada-1][1])+" "+str(grados_entrada_a_otra(grados_entrada,dig_T_entrada, list_convers[1])))</pre> <p>No cal fer un else, doncs en el moment que hi hagi un error el fer l'ingrés de les dades, ja hi ha un print de error que s'indica que s'havia de repetir tot el procés i no entrarén en aquesta condició: grados_entrada!=None and len(list_convers)==2</p>

Los comentarios, normalmente, se escriben en el markdown a menos que sea un comentario del código como tal, pero me pones demasiados comentarios mezclados: salidas, comentarios que no es de código, etc. Recuerda que puedes ir ejecutando en diferentes líneas.

El ejercicio 3 tiene muchas cosas que ya existen en funciones y métodos:

- La limpieza de los signos de puntuación, espacios, el cambio a minúsculas y divisiones lo puedes integrar todo en una sola línea al aplicar sobre todo el texto que entre y no por separado con el uso del regex por ejemplo! `re.findall(r'\b\w+\b', texto.lower())`

También existe `isalpha()`, `isalnum()` y `isspace()` para filtrar los caracteres no deseados, o con `.translate` y `.maketrans` para hacer un mapeo de traducción, o también con `.replace` (que lo usas para los signos de puntuación) para reemplazar esos caracteres por nada (no por espacio).

```

#es tracta d'una funció que utilitzarem per anar substituint diferents
#tipus de caràcters repetits per un espai, ".
#Per exemple, si troben al text, una separació de més de dos espais entre
#aparells, es substituirà per un espai. Si troben un espai seguit d'un
#; serà el caràcter que vindrà a substituir per un espai. També ho utilitzarem
#pepis salts de línia.
def sub_car_repetits_per_espai(texto, c):
    #Si pot ser " ", o "w", o bé " ", o bé " " o bé " "
    #Si c és un espai, " ", aleshores serà un espai de 15 posicions com a màxim.
    #Avet salt de línia i com a molt 15 salts de línia.
    match c:
        case " ":
            rep=15
        case "w":
            rep=15
        case "b":
            rep=15
        case _:
            rep=1

    for i in range(rep,0,-1):
        #Caracter superior a un espai, amb la condició superior a 1 (texto.count(car_repetits)>0)
        #serà suficient. Es a dir, només ens interessaria substituir els que tinguin més d'un'espai, ja
        #que si tenim un espai entre paraules ja ens estaria bé.
        #Així que si el caràcter està en la posició 15, doncs s'ha saltat de línia i també
        #s'hauria de substituir per un espai. Així que fent 15 ens servir per tots els casos. Pel cas dels espais
        #d'un únic espai, simplement no ens serveix el metode.
        if texto.count(car_repetits)>0:
            texto_n=texto.replace(car_repetits," ")
        else:
            texto_n=texto

    return texto

#Bous un text en format molt general, amb forces casuístiques, et retorna una llista amb totes les
#paraules en cada posició de la llista
def paraules_text_a_llista(texto):
    signes_punt=[".",",",";","(",")","[","]","!","?","`"]
    #print("TEXTO ORIGINAL\n")
    print(texto)

    #print("\n\nTEXTO CONVERTIT EN MINÍMUSCULAS\n")
    #d'alhora de comptar paraules, volem que està i està les compti com la mateixa,
    #que no ho diferencie com paraules diferents sind.
    texto_n=texto.lower()
    #print(texto_n)

    #print("\n\nTEXTO DESPRÉS DE CANVIAR SIGNES PUNTUACIÓ PER ESPAIS\n")
    for element in signes_punt:
        texto_n=texto_n.replace(element, " ")

    #print(texto_n)

    #print("\n\nTEXTO DESPRÉS D'ELIMINAR TOTS ELS SALTS DE PÀGINA PER UN ESPAI\n")
    texto_n=texto_n.replace("\n"," ")
    #print(texto_n)

    #Els millor fer-ho al final, ja que quan substituim signes puntuació per espais, l's'eliminen els
    #secls de pàgina anterior per espais és poden formar espais superiors a un entre paraules. Per això
    #el final del text n'eliminem tots els espais superiors a un.
    #print("\n\nTEXTO DESPRÉS D'ELIMINAR TOTS ELS ESPAIS SUPERIORS A UN ESPAI\n")
    texto_n=subs_car_repetits_per_espai(texto_n, "\n")
    #print(texto_n)

    #Al afegir un texto n=texto.replace(" "," "), em va molt bé per separar frases amb un punt i seguit
    #aper al final del text n'apreix amb un espai que després m'el separarà a part com si fos una paraula.
    #Així que la tècnica del silice n'm'ajuda aquell últim espai.
    texto_n=texto_n[:len(texto_n)-1]

    #Ara es veu el sentit de tot el que hem fet. Passen la cadena de text a una llista, separant els elements
    #que estan separats per espai, que són les paraules
    llista_texto_n=split(texto_n)
    #print(llista_texto_n)

    #print("\n\nLlista de paraules\n")
    #print(llista_texto_n)

    return llista_texto_n

#Cream un diccionari d'una llista de paraules que poden estar repetides amb les paraules com a clau
#i amb el número de vegades que apareixen cada paraula com a valor
def dict_paraules_and_count(llista_texto_n):
    dict_paraules_count={}
    for paraula in llista_texto_n:
        dict_paraules_count[paraula]=1

    dict_paraules_count=Counter(llista_texto_n)

    return dict_paraules_count

#Cream un diccionari d'una llista de paraules que poden estar repetides amb les paraules com a clau
#i amb el número de vegades que apareixen cada paraula com a valor
def dict_paraules_and_count1(llista_texto_n):
    dict_paraules_count={}
    for paraula in llista_texto_n:
        dict_paraules_count[paraula]=1

    dict_paraules_count=Counter(llista_texto_n)

    return dict_paraules_count

```

- o Igual el código que queda fuera no debería estar allí, tienes impresión, entradas y las llamadas al resto de las funciones.

Element de qualificació

Ponderació

calculada Qualificació Gamma Percentatge Retroacció

```
#####
##Codi Principal#####
#Per la última visualització de la solució
from tabulate import tabulate

texto="""Este es un texto de prueba. Algunas palabras tienen más de un espacio.

También existen frases separadas por varios saltos de página. Algunas frases
después del punto, no tienen espacios. Existen palabras separadas por coma sin
espacio, o también por punto y coma; claro.

Intentaremos resolver este caso general."""

list_parcles=parcles_text_a_llista(texto)
list_count_x_ele=[]

#Per cada parcla de la lista_parcles, compta les vegades que apareix a la llista
#amb el mètode count i l'afegelgo a la llista list_count_x_ele.

dic_parcu_count=dict_parcles_and_count(list_parcles)

# Faig pip install tabulate a més
#visualitzaré el diccionari en dues columnes ja que així és més fàcil veure els resultats
tabla = [[clave, valor] for clave, valor in dic_parcu_count.items()]
print(tabulate(tabla, headers=["Clave", "Valor"]))


```

Sin embargo hace lo que se te pide!

El ejercicio 4 está bien, pero no es correcto del todo. Al igual que los anteriores la idea es que hagas una función que puedas ejecutar tantas veces como diccionarios quieras probar. Vuelves a dejar mucho código por fuera, lo bonito de la programación es que puedes optimizar el tiempo y ahorrarte procesos, si es un proceso repetitivo haces una función que puedes utilizarla cada vez que requieras hacer lo mismo.

```
#####
##Codi principal#####
#Fem els dos exemples, un que sí que pot fer l'invers i l'altre que no.
print("Diccionari_enquesta_1",enquesta_1,sep="")
enquesta_inversa_1=diccionari_invers(enquesta_1)
if enquesta_inversa_1!=None:
    print("\nEls resultats del diccionari invers són:")
    print("\ndiccionari_enquesta_inversa_1",enquesta_inversa_1,sep="")
print("\nDiccionari_enquesta_2",enquesta_2,sep="")
enquesta_inversa_2=diccionari_invers(enquesta_2)
if enquesta_inversa_2!=None:
    print("\nEls resultats del diccionari invers són:")
    print("\ndiccionari_enquesta_inversa_2",enquesta_inversa_2,sep="")


```

Cuando tienes este tipo de funciones lo que se hace es comparar el largo de la lista tal cual entra y el largo de la lista consultando los valores únicos `len(set(lista_valors))` y no devuelves nada solo lo evalúas en tu if de `diccionari_invers`. Si es igual es porque no hay repetidos, si no pues si hay repetidos.

```
def llista_valors_poden_ser_claus(llista_valors):
    pot_ser_claus=True
    #Ens va millor fer un for, ja que el mètode count de les llistes
    #no recorre els elements de la llista, i per tant no podríem
    #utilitzar el count amb un index de la llista.
    for element in llista_valors:
        if llista_valors.count(element)>1:
            #Si trobem ja una condició en que hi ha valors repetits,
            #el return fa com de break, i ens estalviem de recorrer tota la llista
            #ja que sortim de la funció.
            return False
    #Si recorrem tots els elements de llista es que No hem entrat a la condició
    #que ens hauria retornat False, i per això retornaríem cert aquí.
    return pot_ser_claus


```

No te compliques tanto!

Esta función estaría perfecta si tomara en cuenta lo que te comenté antes:

```
def diccionari_invers (enquesta):
    (list_claus,list_valors)=llistes_claus_valors_diccionari(enquesta)
    if lliste_valors_poden_ser_claus(list_valors):
        print("\nEl diccionari presenta si pot passar els valors a claus d'un diccionari invers,\nja que el zip ens fa ònics la llista de valors per tal que després es pugui aplicar la funció dict.\n#el zip ens fa ònics la llista de valors per tal que després es pugui aplicar la funció dict.\nreturn(dict(zip(list_valors,list_claus)))
    else:
        print("\nEl diccionari presenta no pot passar els valors a claus d'un diccionari invers,\nja que no cal retorna None, doncs ell per defecte ho farà si no hi ha un return\n#return None


```

Solo los print los haría en una función aparte!

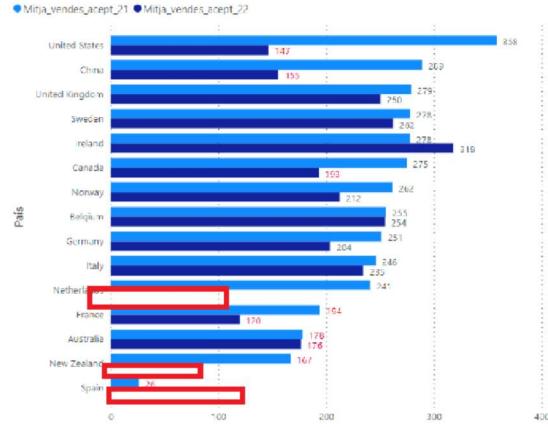
Ahora el ejercicio 1 del nivel 2 es una continuación de éste último ejercicio, la idea es agregar la extensión que necesita para continuar, no es volver a hacer un código nuevo.

El ejercicio 2 también está bien en general. El objetivo de una función siempre es retornar algo, excepto si es una función para imprimir obviamente. No es una buena práctica editar sobre el objeto que envías porque es más fácil tener errores y no darte cuenta de lo que ocurre.

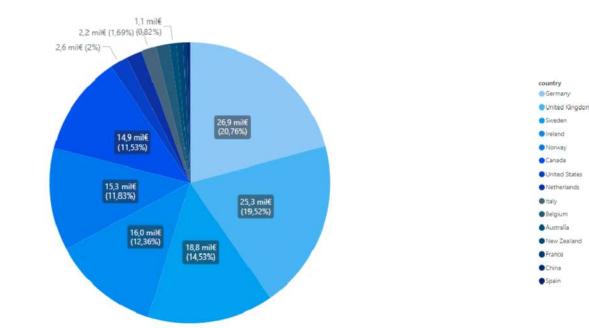
Element de qualificació	Ponderació		
	calculada	Qualificació	Gamma
		Percentatge	Retroacció
			<pre>def asignar_elem_a_list_float_o_list_str(element,list_float,list_str): try: element_float=float(element) list_float.append(element_float) except ValueError: list_str.append(element)</pre>
			<p>Recuerda lo del código fuera de funciones!</p>
			<p>De resto está bien!</p>
			<p>En el ejercicio 1 del nivel 3 pudiste aprovechar el contador de palabras del ejercicio 3 del nivel 1 y solo agregar lo demás para cumplir con lo que te pedían. Siempre recuerda que la idea es optimizar procesos.</p>
			<p>Es un poco mas de lo mismo que ya te comenté en aquel ejercicio, fijate que finalmente es un copy paste y modifcas para generar el diccionario del primer nivel.</p>
			<p>Sigues dejando código por fuera y demasiado texto de comentario, esto no se hace!</p>
			<p>Para ordenar tienes el método sorted que ya utilizas antes, automáticamente el evalúa por orden alfabético las palabras, por ejemplo: si tienes amor y amarillo y las ordenas con sorted te coloca como primera palabra amarillo y luego amor, hace lo que haces tu a mano, revisa el primer valor, es igual? comparo el segundo? es igual? comparo el tercero, quien está mas cerca de la a? la a, entonces amarillo va antes que amor!</p>
			<p>Antes de ponerte a desarrollar código busca a ver si ya existe la función!</p>
			<p>Un while para esto no tiene sentido, para eso existe el for:</p>
			<pre>while i<=(len(list_claus)-2): if list_claus[i][0]==list_claus[i+1][0]: dic_aux[list_claus[i]]=list_valors[i] else: #En el canvi de lletra de dos elements #no entro a la condició anterior, però dic_aux[list_claus[i]]=list_valors[i] #Al diccionari solució, afegeixo com a dic_sol[list_claus[i][0]]=dic_aux #Si he entrat a l'else és que la primer #de claus és diferent. Per tant, he de dic_aux={} i+=1</pre>
			<p>También tienes la función Counter de la librería collections que te cuenta lo que le pases en donde le indiques, tienes las comprensiones de listas y diccionarios que te ayudan a reducir los loops a una línea y así tienes muchas funciones que te pueden ayudar a que todo ese código se conviertan en unas cuantas líneas. Adicionalmente que puedes usar las funciones que ya creaste anteriormente!</p>
			<p>Un saludo,</p>
			<p>Alana</p>

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	Contribució al total del curs
TASCA  Tasca S6.01. Anàlisi avançat de les visualitzacions interactives amb Power BI	-	✓ 9.00	0-10	90.00 %	Hola Gorka,	Porfa recuerda siempre colocar con quién has realizado el P2P. Revisado por: Maria Pijoan Gall.	-

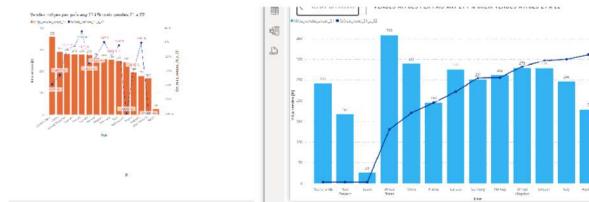
Mitja vendes per país i any



Muy bien el ejercicio 2, espero ansiosamente las interpretaciones. En este caso me muestras unos colores en el pdf y otros en el pbix. Ten cuidado que hay colores que terminan solapándose o mas bien, no se puede encontrar una diferencia clara:



El ejercicio 3 está muy acertado el gráfico, aunque está mucho mejor organizado en el pbix que en el pdf.



El ejercicio 4 está muy bien.

Trata siempre de que en un mismo dashboard no tengas tantas gráficas del mismo estilo porque las personas que lo leen terminan perdiendo el interés por lo que ven.

Muy bien el ejercicio 5, veo que solucionaste lo de Europa! muy bien!

En el ejercicio 6 ya comentas lo del cambio de los colores! Mmmm bueno describes las gráficas muy bien, pero necesito tus sugerencias, la única que tengo es que el gerente debería seleccionar empresas más atractivas del resto del 70% de países. Otra de las cosas que medio asomas es que en el 2022 se deberán implementar acciones para incrementar las ventas, pero ¿qué acciones? ¿qué propones? ¿qué crees conveniente?

Element de qualificació

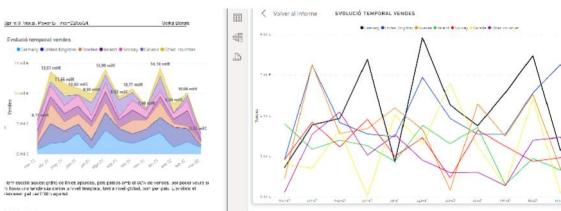
Ponderació

calculada Qualificació Gamma Percentatge Retroacció

Quiero saber cosas más allá de lo que puedes describir de las gráficas, por ejemplo: ¿qué te parecen los objetivos planteados? ¿son acordes a la empresa? ¿son buenos evaluadores? ¿los crees indicados? ¿qué crees que se debe adaptar? ¿qué crees que serían objetivos mas realistas? ¿qué crees que se puede hacer para mejorar los mercados en las grandes potencias económicas? ¿qué podrías sugerir para revertir esas tendencias negativas en las diferencias entre los años para los meses siguientes? Bueno lo que comentamos en la conversación.

El ejercicio 1 del nivel 2 está perfecto.

El ejercicio 2 me muestra dos gráficas algo distintas entre el pdf y el pbix. La del pdf no es objetiva! tienes que tener cuidado con las áreas apiladas! Fíjate que el que está en el archivo es muy diferente al de áreas porque muestra en los ejes reales y sin apilar, aunque con las apiladas podrías ver tendencias generalizadas solamente si tienes claro lo que pasa individualmente y extraes los valores atípicos.

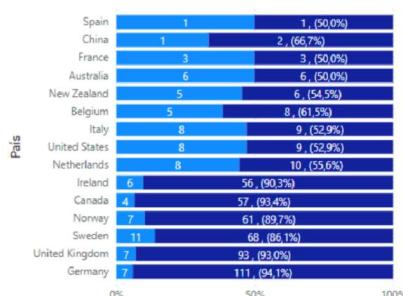


El ejercicio 3 está perfecto.

El ejercicio 4 me encanta lo que hiciste, pero no le das mucha importancia. Esto es realmente lo que hay que tomar en cuenta cuando evaluamos las cantidades de transacciones declinadas! El ratio!

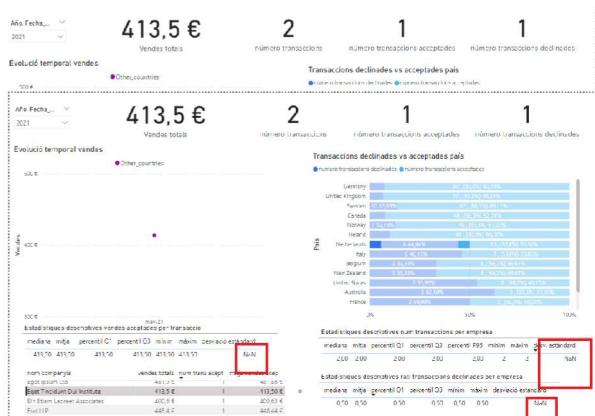
Num transaccions declinades vs acceptades per país

● num_trans_declined ● num_trans_acept



Hem posat etiquetes personalitzades, amb la variable DAX que hem vist anteriorment. Hem posat només el percentatge relatiu en les transaccions declinades, ja que hi cap millor. I les transaccions declinades el percentatge que no hi cabria, es pot deduir.

En el ejercicio 5 de los estadísticos son muy interesantes, pero no me comentas nada con respecto a lo obtenido ¿qué cosas encontraste? Cuidado con el manejo de los nan.

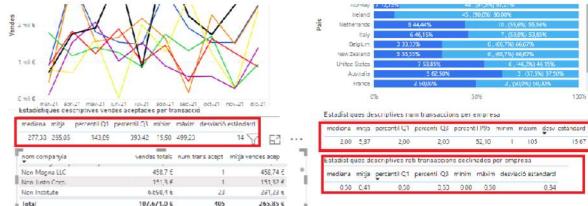


Element de qualificació

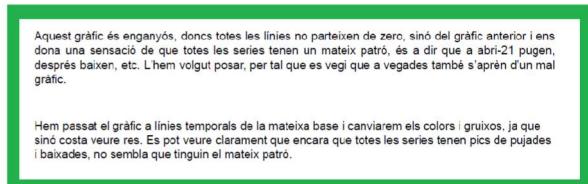
Ponderació

calculada Qualificació Gamma Percentatge Retroacció

No me dices nada que te haya llamado la atención de algún estadístico ¿por qué sería importante mostrar todas esas métricas? ¿pros contras? ¿qué puedes leer de ello? Por ejemplo: en el caso de Non Institute ¿qué significan esos valores? la mediana es mayor que la media ¿tiene sentido esto? ¿por qué? ¿qué puedes decir de la desviación y los percentiles? ¿son valores adecuados? ¿son valores muy grandes o muy pequeños?



En el punto 5.6 de tu pdf, esto es absolutamente certero, todo lo que comentas del gráfico de la evolución temporal de las ventas, que algo te había comentado que no era objetivo.



Supongo que aquí quisiste decir segundo nivel:

Volem senyalar doncs, que el Dashboard final de primer nivell és el següent:



Está muy bien todo lo que me comentas, pero como lo hablamos, necesito las sugerencias y posibles soluciones un poco más pintadas sobre la mesa. Igual que para los siguientes comentarios, expones problemas sin soluciones.

El gerent de vendes, té que prendre una decisió i intentar fer mesures per diversificar les vendes d'Alemanya amb altres empreses, doncs si *Nunc Interdum Incorporated* es donés de baixa de la nostra e-commerce podrem tenir una davallada de vendes d'un 20% aproximadament, un risc massa gran. En el següent nivell s'analitzarà quins usuaris i productes han participat en aquestes transaccions.

Aunque, este tipo de comentarios está muchísimo mejor, esto es realmente lo que vale. Trata siempre de que si muestras problemas comenta posibles soluciones.

5.6.3 Vistos els resultats, és prioritari pel gerent de vendes, en centrar-se en les empreses dels països del 80% de les vendes o en les del 20%?

Hem vist que com que cada país d'aquest top80 pràcticament genera les vendes amb una única empresa. Per tant, el gerent de vendes hauria de fidelitzar el contracte amb aquestes empreses, i retener alguna que ha perdut al 2022. Com que sembla que aquests països semblen més atractius, també seria interessant que alguna empresa més generés vendes, fent més campanyes de publicitat o fent noves captacions.

5.6.4 Quines accions hauria de fer el gerent de vendes amb els països del 20% de les vendes?

Segurament hauria de prioritzar fer més relacions amb Estats Units i Xina, uns mercats amb molt més potencial i que actualment només generen un 1,58% de les vendes i Xina amb unes vendes totals de 444,5 €. Uns resultats molt decepcionants.

5.6.5 Quines altres accions hauria de fer el gerent de vendes per maximitzar els beneficis i reduir el risc?

Una de les altres accions podria ser intentar aplanar les corbes, és a dir, pensar amb accions, on es puguin generar vendes fora de les puntes del final de cada trimestre, sobretot per Alemanya i Regne Unit que tenen un patró similar.

Ponderació

Element de qualificació Ponderació Qualificació Gamma Percentatge Retroacció

Entonces, si esto piensas ¿por qué no planteas un cambio de objetivo o un replanteamiento del objetivo que se tiene estimado?

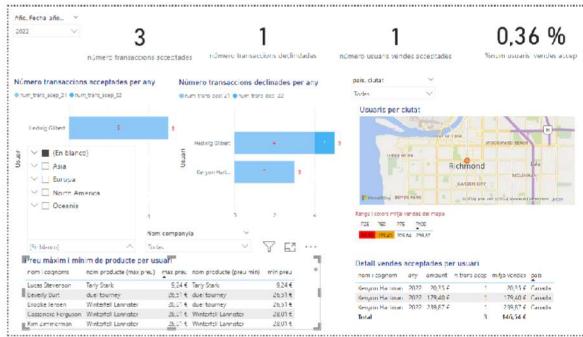
5.6.6 Vistos els resultats de l'histograma que mostra el percentatge de transaccions declinades i acceptades, podem treure alguna una conclusió rellevant?

Clarament els països del 90% de les vendes, tenen un rati de transaccions declinades més baix que la resta de països.

El que realment passa, es que en aquests països top, les transaccions es fan a molt poques empreses, i aquestes empreses generen molt poques transaccions declinades.¹ A priori podríem pensar que són aquestes empreses les que van fer que les que hi hagi poques transaccions declinades, però fins que no veiem a l'últim nivell quins usuaris han fet aquestes transaccions no podrem treure conclusions.

Plantéate los escenarios: podrías crearte una nueva columna con la información que recibiste de otros departamentos de las causas de las transacciones declinadas! ¿qué soluciones planteas conociendo ese detalle?

En el nivel 3 ten cuidado con el manejo de blancos:



Me encantó la distribución geográfica de los usuarios con colores tomando en cuenta la media del amount.

Solo queda raro que los filtros segmentadores no afecten a todo el dashboard. Estos de cobre de continente y compañía puedo entender que es para buscar los usuarios que han comprado a esas empresas ¿podría ser?



Entonces ¿qué solución propones? No eres de marketing como para decir qué campañas o no emplear o implementar, pero estás analizando los datos, puedes decir con mucha certeza si los objetivos se pueden o no lograr, si tienen o no sentido, si son adecuados o no, etc.

Amb aquests resultats, el gerent ha d'estar molt intranquil, doncs quasi bé totes les vendes han estat produïdes per molts pocs usuaris, cosa que augmenta encara més, un futur incert en la previsió d'ingressos si alguns d'aquests pocs usuaris deixen de comprar productes a la nostra ecommerce.

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	Contribució al total del curs
No es que los usuarios compren a una determinada empresa, realmente no importa a quién compren mientras que lo hagan a alguna de las empresas que se tienen registradas.							

Del nivell anterior, em vist que pràcticament el 80% de les vendes s'han generat per una única empresa per país, a excepció del Regne Unit (dos empreses). Ens faltava veure si han comprat molt o pocs usuaris a aquestes empreses tot i que sembla pels resultats del gràfic de transaccions acceptades que són molts pocs usuaris.

Este nivel es solo de usuarios, no era la idea agregar información de las empresas que han vendido los productos a los usuarios.

Mas abajo está mas claro, pero ten cuidado con los comentarios, lee esto sin leer los párrafos siguientes.

També podem obtenir la informació de quins usuaris per país han comprat, gràcies al segmentador del mapa. Si filtrarem per Canadà tenim 25 usuaris, 40 pel Regne Unit i 9 per EUA. Això ens dóna

49

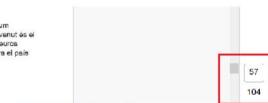
Sprint 6: Visual. Power BI. Inici=22/05/24.

Gorka Bonals

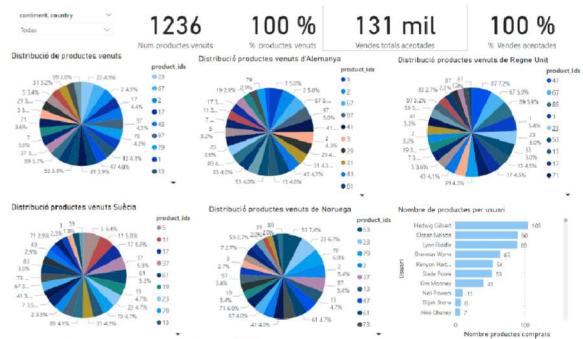
una taxa de participació respecte al clients potencials que tenim (taula d'usuaris) del 33%, 80% i del 6%.

Estoy impresionada de la cantidad de hojas que destinaste a este nivel 3:

Veiem que s'han venut 1236 productors, i del primer diagrama de sectors, que és un resum general, no veiem cap producte que sobrepassi els altres. El que més s'ha venut es el 22 que representa el 2,4% del total i només amb un 4,6% de tots els productes venduts però en euros supera un 8%. El que es curiosa que aquest producte no s'ha fet només a Alemanya, que era el país que havia generat més vendes.



Cuidado con la decisión de qué datos incluir en un gráfico. Si no conoces el detalle de qué id pertenece a qué producto ¿tiene sentido hacer los gráficos con los id?



Muy bien con la información del usuario con id 96. Esta sugerencia está muy bien:

Tal com hem comentat en la Anàlisis anterior, el gerent de vendes, hauria d'investigar perquè aquest usuari ha baixat tant en transaccions al primer trimestre del 2022. Seria interessant que li fessin una trucada comercial, per millorar el nivell de satisfacció. També seria interessant saber perquè sempre productes a empreses de Suècia, quan són productes que també es fabriquen a altres països amb els mateixos preus.

Ok ok de la 63 en adelante son anexos.

Un saludo,

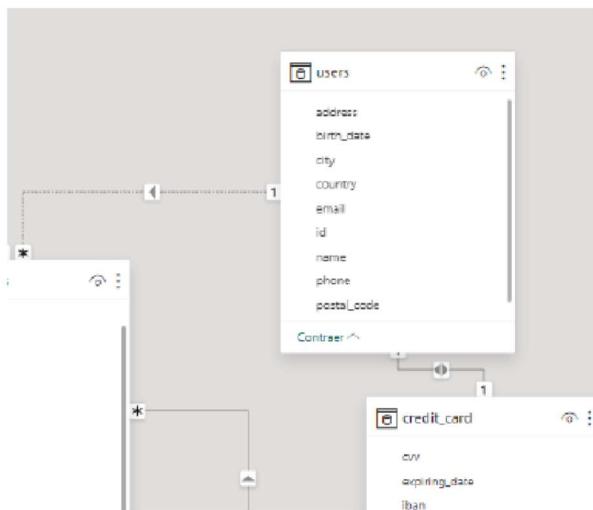
Alana

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	Contribució al total del curs
TASCA Tasca S5.01. Iniciació al anàlisi de dades amb Power BI i indicadors	-	✓ 8.50	0-10	85.00 %	Hola Gorka, Muy buen trabajo! Porfa recuerda que cada nivel esté en un panel y colocar: Revisado por: Albert Batlle Valero		-



Muy bien el ejercicio 1 con la importación, el cambio de timestamp a date y la nueva tabla calendario!

Sigo viendo esa relación entre users y credit_card.



En el ejercicio 2 lo primero que veo es la creación de la tabla de medidas jajajaja borré el comentario en el ejercicio 1, que con tantas medidas y cosas a calcular te podría ser útil una tabla de medidas también.

Entonces, tomando en cuenta que del 2022 solo tienes 2 meses y medio ¿crees que se podría alcanzar el objetivo? ¿cuál es tu opinión del objetivo planteado tomando en cuenta lo que sabes del 2021 y lo que va del 2022?

Per al 2022, encara que sembla que l'empresa no ha complert l'objectiu empresarial de vendes de 25 mil euros quedant-se amb 21,83 mil euros en vendes, no seria correcte afirmar-ho. En realitat pel 2022 només tenim dades pels tres primers mesos, per tant fins que no tinguem les dades dels 12 mesos del 2022 no ens podríem pronunciar.

Fecha_ahor_mes	Suma de amounts
mar-21	6,471,55
abr-21	15,989,75
may-21	15,364,17
jun-21	14,241,19
ju-21	11,373,93
ago-21	15,608,04
sep-21	10,503,32
oct-21	12,281,17
nov-21	8,005,31
dic-21	15,162,07
ene-22	8,980,32
feb-22	10,841,42
mar-22	4,697,52
Total	150,703,75

El ejercicio 3 está muy bien, solo recuerda que debes colocar las unidades de medida de alguna manera sea en el título o en el objeto gráfico.

Seguiment de les vendes mitjanes anuals



Es a dir, la transaccions mitges de les operacions acceptades pel 2021, han estat un 6,34% superiors a l'objectiu que s'havia marcat de 250, amb un valor mig de 265,85 € per transacció.

El ejercicio 4, ten cuidado con los comentarios, no son informaciones falsas, te están diciendo que la media deberá ser de 250€ para el año. Si, no tienes todos los datos del 2022, te entregan solo una parte, pero no por eso dejarás de analizarlo, al igual que con el 2021, no te entregan los datos del año completo y entregas tu análisis de lo que encuentras.

Entonces, te están hablando de media y en estos dos meses y medio,

Element de qualificació

Ponderació

calculada Qualificació Gamma Percentatge Retroacció

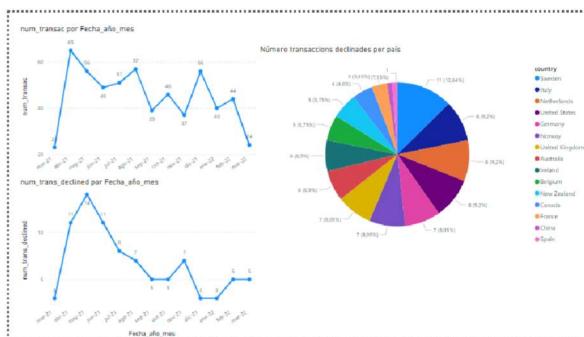
aún no se cumple ¿evaluaste solo 2 meses? ¿qué debería pasar para que se cumpliera este objetivo en 2022? ¿qué le recomiendas a tu cliente que te paga por este informe?

Podriem pensar que l'empresa no ha complert l'objectiu pel 2022, però seria una afirmació falsa. Doncs només tenim tres mesos d'informació pel 2022, tal com es veu en la següent taula, i és per això que no en podem treure cap conclusió.

	Totals	Suma de amount
mai-21	6.217,66	
jun-21	13.890,98	
may-21	15.154,17	
jun-21	14.241,19	
juil-21	11.739,89	
agost-21	15.869,48	
sept-21	10.503,32	
oct-21	12.281,17	
nov-21	8.688,91	
dic-21	16.597,95	
ene-22	8.690,02	
feb-22	10.541,42	
mar-22	4.877,62	
Total	150.793,75	

Muy bien el ejercicio 5, solo sigo insistiendo en mejorar la parte de distribución de los paneles con un panel para cada nivel, por ejemplo.

El ejercicio 6 está perfecto! Muy muy interesante el complementar con este otro gráfico del cual no me dices gran cosa y hay información muy importante. Me estás mostrando la cantidad de transacciones y la cantidad de declinados, que esta parte proporcional sería mucho más interesante que esa proporción que muestras de las declinadas totales. ¿qué puedes comentar de esas dos gráficas a la izquierda del panel auxiliar?



De este ejercicio 7 no me dices mucho de tu interpretación ¿qué piensas entonces de ese objetivo? Si miras ambos años ¿es viable este objetivo? ¿qué crees que se podría sugerir para mejorar el panorama?

La idea del ejercicio 8 no es que me muestres una tabla donde solo estén esos usuarios que cumplen con la condición de que sea $\geq 300€$ y $\geq 320$$, la idea es poderlos ver a todos y que resaltes de alguna manera aquellos que cumplen con esa condición.

Nom i Cognom	Edad	Mitja vendes	Mitja vendes dolar
Sophia Bradford	39	300,77	324,83
Raymond O'Neill	37	301,55	325,67
Chase Ellis	24	304,63	329,00
Uriel Herbert	43	304,95	329,35
Remedios Case	30	305,02	329,42
Donna Rivers	34	306,96	331,52
Germaine Suarez	40	307,79	332,41
Kerry Adkins	41	309,22	333,96
Chloe Keith	25	309,45	334,20
Dane Shepard	25	316,74	342,08
Shellie Valenzuela	26	318,65	344,14
Bruce Gill	34	320,52	346,16
Philip Carey	32	325,08	351,09
Idola Bowen	33	325,34	351,36
Athena Malone	33	326,86	353,01
Damian Mcgee	36	327,56	353,76
Kenneth Morrison	43	328,64	354,93
Beverly Burt	28	329,72	356,10
Brody Talley	33	337,22	364,20
Ainsley Herrera	28	339,62	366,79
Dominique Tillman	25	339,88	367,07
Melissa Cameron	32	346,88	374,63
Emerson Sharp	30	349,66	377,63
Hilary Ferguson	43	350,13	378,14
Wanda Campbell	25	350,28	378,30
Jean Paul	41	353,84	379,99
Total		397,70	429,52

El ejercicio 9 eran solo 50 palabras donde explicas las gráficas HECHAS que quieras, todas, algunas, como tu quieras. No era generar 4 paneles mas.

Exercici 9.1

Exercici 9.2 **X**

Exercici 9.3

Exercici 9.4

Lo que hiciste está muy muy bien porque estás entregando mucho más de lo que te están pidiendo, pero ten cuidado porque este tipo de cosas te pueden llevar mucho más tiempo del que tienes

En el ejercicio 1 del nivel 2 quiero que me comentes mas acerca de lo que piensas ¿qué crees de este objetivo ya conociendo los demás? o incluso, solo tomando en cuenta esta sola gráfica ¿qué opinas del objetivo trazado? ¿debería modificarse? ¿es alcanzable? ¿qué opinas de la cantidad de meses que no se han llegado al objetivo? Muy bien con respecto a la tendencia!

Fecha año mes	Vendes totals accept	Var. vendes mes
mar-21	5.711,35 €	5.711,35 €
abr-21	13.507,68 €	7.796,33 €
may-21	11.659,27 €	-1.848,41 €
jun-21	10.817,81 €	-841,46 €
jul-21	9.931,77 €	-886,04 €
ago-21	13.900,24 €	3.968,47 €
sep-21	9.524,43 €	-4.375,81 €
oct-21	10.765,57 €	1.241,14 €
nov-21	7.675,74 €	-3.089,83 €
dic-21	14.177,16 €	6.501,42 €
ene-22	8.236,31 €	-5.940,85 €
feb-22	10.077,55 €	1.841,24 €
mar-22	3.520,71 €	-6.556,84 €
Total	129.505,59 €	3.520,71 €

El ejercicio 2 está perfecto!

El ejercicio del nivel 3 lo mismo de tener todo en un mismo panel, de tal manera que puedas ver todos los gráficos y la información en el mismo panel. Te comento por ítems:

- Cuidado con los percentiles que no se distingue cual es cual: podría ser Percentil1... Percentil2...

Estadístics per vendes acceptades [€]								
129.505,59	259,01	262,47	15,05	136,19	262,47	390,85	499	
Vendes_total_sce...	Mèdia_vendes_sc...	Mediana_vendes_...	Mitj._vendes_accep...	Percentil_Vendes...	Percentil_Vendes...	Percentil_Vendes...	Maxi...	
Mèdia_num.prod...	Mediana_num.p...	Distr_standariz...	Mitj._num.prod...	Percentil_num.p...	Percentil_num.p...	Percentil_num.p...	4	Maxim...

Estadístics per num productes venuts per transacció								
2,48	2,00	0,99	1	2,00	2,00	3,00	4	
Mèdia_num.prod...	Mediana_num.p...	Distr_standariz...	Mitj._num.prod...	Percentil_num.p...	Percentil_num.p...	Percentil_num.p...	4	Maxim...

Con respecto a los comentarios, si, la media y la mediana están cerca, pero no son las mismas! Antes de hablar si es o no una variable normal tendrías que aplicar algún test estadístico para comprobar la normalidad de esa variable y comentar cuantitativamente si es o no normal.

Una distribución normal tiene una campana simétrica y si te fijas en los cuartiles de productos, que son mas evidentes, ves que el 1 y el 2 son el mismo valor, es decir, hasta el 50% de los datos está en este punto.

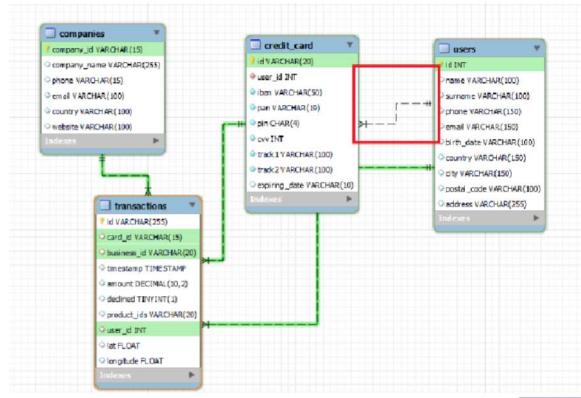
Es relevant que el percentil 75%, que ens dóna un valor de 390,85 € en promig, ens indica que només un 25% de les vendes per transacció estan per sobre de 390,85 € i per sota de 499,23 €. Per tant, sembla que la dispersió de les vendes segueix bastant una forma de campana de Gauss centrada en 259,01 €.

En quan a l'estadístic del nombre de productes venuts per transacció, veiem que en promig s'han venut 2,48 productes per transacció i que com a molt s'han venut 4 productes per transacció. També els percentils ens donen una idea que sota del promig i per sobre del promig la distribució del nombre de productes venuts per transacció segueix una distribució normal.

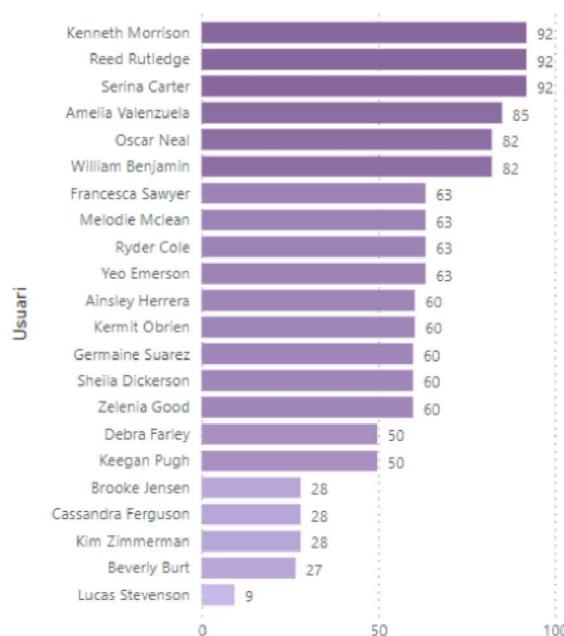
- En el ítem 2 mira que la visualización que seleccionaste no es fácil de entender, es difícil hablar del producto 1, 2, 3, etc etc es mucho mas fácil decir hey marketing! el producto skywalker ewok no se vende, que decirles el producto 25 no se vende, al igual que a los clientes. Ellos no saben los ids de los productos!

- En el ítem 3 si, ese es el problema de haber dejado la conexión entre users y credit_card (que te comenté en el sprint 4) alguna termina siendo irrelevante.

Si se supone debes buscar un modelo estrella ¿crees que está bien esta conexión entre user y credit_card?



- La media por usuario está muy bien!
 - Producto mas caro por usuario muy bien, pero mira que podrían existir varios productos que cuesten lo mismo. Hubiese quedado mucho mejor si le agregaras el nombre.



- Distribución geográfica me gustó muchísimo el gráfico que presentas, es bastante claro, se entiende con facilidad!

En tu caso tienes un error en cómo aplicas el filtro dentro del CALCULATE. Estas contando las filas y guardando este resultado en una variable la cual es un número fijo y no cambia, incluso si cambias el contexto de filtrado.

```

1 num_transac_declined_v2 =
2 VAR num_trans_=COUNTA(transactions[declined])
3 RETURN CALCULATE
4 (num_trans_,
5 transactions[declined]=TRUE
6 )

```

Esto es una forma de resolver: con FILTER creo una tabla temporal que solo incluye las filas de transactions donde la columna declined es true. Esa tabla filtrada la almacena en la variable filtered_transactions. Luego,uento las filas de esa tabla.

```

1 num_transac_declined_v2 =
2 VAR filtered_transactions =
3   FILTER(
4     transactions,
5     transactions
6     [declined] = TRUE
7   )
8 RETURN
9   COUNTROWS
10  (filtered_transactions)

```

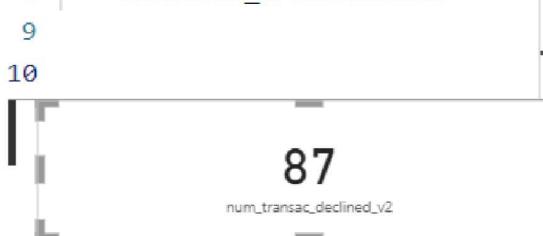


También se puede hacer sin filter: aquí uso CALCULATE para modificar el cálculo, aplico el filtro donde declined es true, yuento las filas. El resultado igual lo almaceno en filtered_transactions.

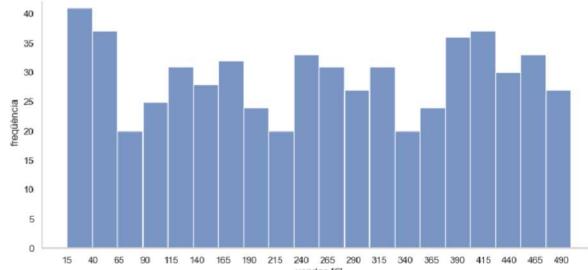
```

1 num_transac_declined_v2 =
2 VAR filtered_transactions =
3   CALCULATE(
4     COUNTROWS
5     (transactions),
6     transactions
7     [declined] = TRUE
8   )
9 RETURN
10 filtered_transactions

```



Element de qualificació	Ponderació calculada	Qualificació Gamma Percentatge Retroacció				Contribució al total del curs
					Sin embargo, la mejor opción es siempre la mas simple, tu primera opción es esta mejor opción. Todo esto es lo mismo pero complicando la cuestión con las variables.	
					Un saludo,	
					Alana	
				Hola Gorka,		
				Muy bien!		
				Solo repasa el tema teórico para que lo tengas claro la diferencia entre copo de nieve y estrella.		
				Cuando tengas errores, antes de continuar y suponer de qué va, léelo para que no pierdas tiempo intentando soluciones.		
TASCA	-	✓ 10.00	0-10	100.00 %	No te acostumbres tanto a la parte gráfica, siempre ten claro cómo se hace en código porque no todas las app tienen la sección gráfica. Así como el workbench tiene algo sencillo, hay otras que son mucho más completas en la parte gráfica pero hay muchas otras que no lo son.	-
				De resto muy bien!		
				Un saludo,		
				Alana		

Element de qualificació	Ponderació					curs
	calculada	Qualificació	Gamma	Percentatge	Retroacció	
TASCA	-	✓ 9.75	0-10	97.50 %	Hola Gorka,	-
 Tasca S8.01.					Muy buen trabajo!	
 Visualitzacions en Python					Muy bien la conexión!	
					Es una buena práctica en este mundo de programación que las librerías, clases, funciones, métodos o cualquier cosa que importes lo hagas siempre al inicio, separado del resto del código.	
					<pre>In [55]: from sqlalchemy import create_engine, inspect #En culpa de la mala contraseña @ import urllib.parse #Dades de connexió usuario='root' contraseña="123" #contraseña_codificada = urllib.parse.quote_plus(contraseña) # Codificar la contraseña para que no se vea en el código fuente host='localhost' puerto=3306 nombre_base_datos='bdd_tasca_4' #Crear la URL de la connexió urlConexion=f"mysql+pymysql://:{usuario}:{contraseña}@{host}:{puerto}/{nombre_base_datos}" try: #Crear el engine de SQLAlchemy engine=create_engine(urlConexion) cnx=engine.connect() print("connexió correcta") #cnx.close() except Exception as e: print(f"Error al conectar con SQLAlchemy: {e}") connexió correcta</pre>	
					<pre>In [56]: #Recordar que inspect del código anterior quan creo la connexió amb sqlalchemy #from sqlalchemy import create_engine, inspect import pandas as pd inspector=inspect(engine) nombre_tablas=inspector.get_table_names() #df=pd.DataFrame()</pre>	
					Muy bien la importación de las tablas a df! Solo ten en cuenta que cuando declares una variable trata de que su nombre te hable del contenido, esto mas bien sería algo tipo dict_df, diccionario de dfs.	
					El ejercicio 1 está bastante bien! porfa sigo insistiendo los comentarios que no son de código déjalos en las celdas markdown.	
					Para ver la distribución de una forma más fácil, entre los parámetros de displot puedes hacer kde = True y te mostrará el histograma y la función de kde. De resto está super bien!	
						
					<p>Comentaris del gràfic: Respecte a la construcció de l'histograma hem de dir, que hem triat una amplada del tamany de l'interval lo suficientment petita per tal que es poguessin treure més conclusions de la forma de la seva distribució. També s'han adaptat les etiquetes numèriques de l'eix x per tal que fossin coincidents amb els tamany's dels intervals, i així facilités l'anàlisis posterior. Si visualitzem la distribució veiem que es forma plana i no té forma de campana de Gauss. La Mitjana és de 256 € per transacció (no viem amb df['transactions']["amount"].describe().També s'observa que les vendes que s'han produït més són transaccions de quantitats baixes entre el rang entre 5 y 40 € amb una freqüència de 40 transaccions, bastant més que altres grups. També veiem que hi ha com tres subpopulacions dins del histograma. S'aprecien com tres muntanyes de grups, hi ha com una petita distribució de transaccions centrada a 150 €, una altra distribució a 275 € i un altre a 425 €. Al ser una una distribució força plana no considerem que hi hagin valors que es puguin considerar outliers.</p>	
					Para que puedas ver esto debes ejecutarlo en celdas separadas o usar print, porque solo muestra el contenido de la última consulta que genere un valor o un objeto de retorno:	
					<pre>In [61]: df["transactions"]["amount"].describe() df["transactions"]["amount"].info()</pre>	
					<pre><class 'pandas.core.series.Series'> RangeIndex: 587 entries, 0 to 586 Series name: amount Non-Null Count Dtype ----- 587 non-null float64 dtypes: float64(1) memory usage: 4.7 KB</pre>	

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	Contribució al total del curs
-------------------------	------------	-----------	--------------	-------	-------------	------------	-------------------------------------

El ejercicio 2 está perfecto!

El ejercicio 3 está perfecto también, me encantó que emplearas lambda! solo separa siempre el código de definición de funciones de los gráficos ya que si quieras modificar algo del gráfico estás ejecutando cada vez la creación de la función.

```
In [66]: #funció per mostrar els valors i porcentatges del pie
def func(pct, allvalues):
    absolute = int(round(pct/100 * sum(allvalues)))
    return f'{.1f}\n{int(absolute)}'
#Especifica una tonalitat de colors blaus per pintar el gràfic de sectors
colors = ['#004d80', '#3366cc', '#66b3ff']

ax=df_seerie_count_x_country["count"].plot.pie(autopct=lambda pct: func(pct,
    pct), colors=colors, startangle=90)
ax.set_xlabel("frequència")
ax.set_title("Nombre d'usuaris per país", fontweight='bold')
```

El ejercicio 4 está muy bien! Perfecto el uso del apply.

Hay muchas líneas que puedes simplificar, no es necesario que declares tantas variables para cada paso.

- Este lo puedes resumir en algo tipo

```
trans_no_decl=df["transactions"][df["transactions"]["declined"]==0]
trans_no_decl.head(2)
```

```
In [76]: #Aplica un filtre per tenir les transaccions no declinades
es_no_declinada=df["transactions"]["declined"]==0
trans_no_decl=df["transactions"][es_no_declinada]
trans_no_decl.head(2)
```

- Este puedes anidar todas las funciones que quieras a la derecha, tomando en cuenta que irá utilizando el resultado hasta el .y adicional añadí el reset_index para que te mantenga en formato de df que es más facil para graficar con cualquier librería y te mantenga los nombres de las columnas.

```
Serie_vendes_x_mes_any=trans_no_decl.groupby("data_mes_any")["amount"].sum().reset_index()
Serie_vendes_x_mes_any
```

```
In [77]: #Calculo la suma de les vendes per data_mes_any i ho agrupem
agrup_mes_any=trans_no_decl.groupby("data_mes_any")
Serie_vendes_x_mes_any=agrup_mes_any["amount"].sum()
Serie_vendes_x_mes_any
```

El ejercicio 5 está muy bien per fíjate que no estás graficando únicamente 2 variables categóricas, las estas acompañando de variables numéricas también, igualmente está genial!

Aquí un poco más de lo mismo que te comenté antes si haces un groupby, esto debe estar siempre acompañado de una función de agregación por lo que debería ser .groupby(..).count() en este caso.

```
In [85]: #aplicació que havíem dit per país d'empresa i per país d'usuari.
country_comp_country_user=tr_no_decl_merge_comp_merge_users.groupby(["country_comp","country_user"])
<ipython-input-85-133a2a2a2a2a>
C:\Users\Usuario\AppData\Local\Temp\ipykernel_34132\3950215164.py:2: FutureWarning: The default of observed=False is deprecated and will be changed to True in a future version of pandas. Pass observed=True to retain current behavior or observed=False to adopt the future default and silence this warning.
  agrup=tr_no_decl_country_comp_country_user=tr_no_decl_merge_comp_merge_users.groupby(["country_comp","country_user"])
In [86]: Serie_country_comp_country_user=agr_tr_no_decl_country_comp_country_user["id_transac"].count()
```

Si le agregas el .reset_index ya tienes en formato df.

Cuando tienes mas columnas que usaras que las que no, lo que se hace es un drop: pivot_df.drop('Total', axis = 1).plot.....

```
In [92]: #No ho fem com pivot_df.plot.barch(stacked=True) per que ens pintaria també el tot
ax=pivot_df[["Canada","United Kingdom","United States"]].plot.barch(stacked=True)
```

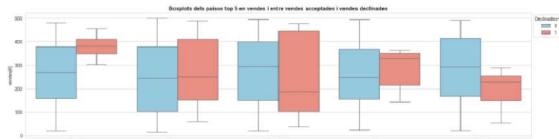
De resto muy bien!

El ejercicio 6 está muy muy bien! ¿qué significa que el rango intercuartil sea mas pequeño o mas grande?

Element de qualificació

Ponderació

calculada Qualificació Gamma Percentatge Retroacció



Comentaris del gràfic: El primer que podem comprovar es que les transaccions NO declinades, (quan declinades=0), totes les transaccions tenen una dispersió força semblant (mida de les caixes), a excepció de Noruega que tant el percentil 25% com el 75% són una mica més reduïts que els dels altres països del top 5. Aquest gràfic també ens dóna molta informació respecte la comparació de les medIANES de les vendes pels països del top 5. Pràcticament, les vendes medianes per país són bastant semblants (ralla intermitja del boxplot), tot i que és curiós que Alemanya que és el país amb més vendes en total, és dels que té una mediana més baixa. Per tant, els usuaris que compren a Alemanya són d'un perfil de gastar menys, però compren molt més sovint.

El ejercicio 7 está muy bien, pero recuerda que puedes hacer merge entre df, que finalmente funcionan tal cual un join de sql.

Por eso es importante dejar en el inicio todas las librerías a importar porque finalmente cada vez que abres el archivo ejecutas esa línea y continuas, o puedes también ejecutar todo.

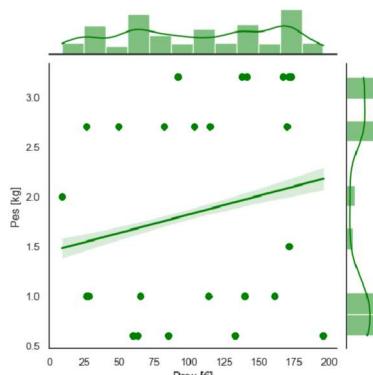
In [96]: `import seaborn as sns`

Muy bien recuerda que también puedes seleccionar las gráficas de la diagonal, de la parte superior a la diagonal y de la parte inferior a ella.

Exactamente, ninguna tiene una distribución normal ni tienen relación entre ellas.

El ejercicio 1 del nivel 2 está muy bien hecho! con esto confirmas cuantitativamente lo que venías comentando antes de la falta de relación entre las variables. Lo cual, dada la naturaleza y significancia de esas columnas tienen sentido.

Muy bien el ejercicio 2, si es como complementario al anterior y a algunas gráficas mas. Exacto el regplot es un modelo de regresión lineal ajustada que nos muestra la tendencia y la relación, la cual pues, vuelve a comprobar que hay una muy pequeña relación entre esas dos variables.



Comentari del gràfic: Aquest gràfic conté molts gràfics comentats anteriorment. El gràfic de dispersió té els mateixos punts en 2D que l'exercici 2, però a més aquí com a valor afegit al haver afegit el paràmetre kind="reg" a la funció de sns.jointplot utilitzem una línia de tendència amb un interval de confiança de la regressió, que generalment és del 95%. Es a dir, si hagués una línia de tendència s'esperaria que caiguis en aquesta zona sombreada amb una probabilitat del 95%. Els gràfics de distribució de l'histograma de les variables pes i del preu s'han explicat quan hem explicat l'exercici 7 del Nivell 1, per tant seria repetir-nos.

El ejercicio 1 del nivel 3 está muy bien, solo que el eje x debería tener su nombre: países de compañías, que veo que lo dejas en blanco. Al violin debes decirle inner=None para que no se vea el boxplot que trae por defecto.

```
In [103]: plt.figure(figsize=(20, 10))
sns.set_style("whitegrid")
custom_palette = ["skyblue", "salmon"]
ax=sns.violinplot(x="country",y="amount",data=df_trans_merge_company_top_5, palette=custom_palette)
ax=sns.swarmplot(x="country", y="amount", data=df_trans_merge_company_top_5,color="black",size=10)
ax.set_xticklabels(["Venda acceptada","Venda declinada"])
#per tal que no apareixi la etiqueta country a l'eix x
ax.set_xlabel("")
ax.set_ylabel("vendes[€]",fontsize=15)
#Aumentar el tamany de los ejes x e y
ax.tick_params(axis='x', labelsize=15) # Tamany de La fuente de las etiquetas del eje x
ax.tick_params(axis='y', labelsize=15) # Tamany de La fuente de las etiquetas del eje y
ax.tick_params(axis='both', which='major', size=8) # Tamany de Los numeros en los ejes x e y
```

El ejercicio 2 con el facetgrid está perfecto!

Element de qualificació	Ponderació calculada	Qualificació	Gamma	Percentatge	Retroacció
-------------------------	-------------------------	--------------	-------	-------------	------------

Un saludo,

Alana

Element de qualificació	Ponderació					-
	calculada	Qualificació	Gamma	Percentatge	Retroacció	
TASCA  Tasca S8.02. Power BI amb Python	-	✓ 8.00	0-10	80.00 %	Hola Gorka, Muy bien el trabajo! La conexión es correcta! Recuerda que las importaciones de las librerías se hacen al inicio como buenas prácticas de programación.	-

```
In [1]: from sqlalchemy import create_engine, inspect
#Dades de connexió
usuari='root'
contraseña='123'
host='localhost'
puerto=3306
nomre_base_datos='bbdd_tasca_4'

#Crear la URL de la connexió
url_conexion=f'mysql+pymysql://:{usuari}:{contraseña}@({host}):({puerto})/({nomre_base_datos}'

try:
    #Crear el engine de SQLAlchemy
    engine=create_engine(url_conexion)
    cnx=engine.connect()

    print("connexió correcta")
    #cnx.close()
except Exception as e:
    print(f"Error al conectar con SQLAlchemy: {e}")

connexió correcta
```



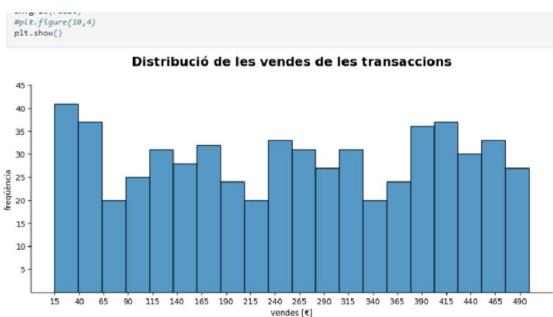
```
In [2]: #Recordar que inspect del codi anterior quan creo la connexió amb sqlalchemy
#necessitarem pandas per utilitzar pd.read_sql
import pandas as pd
from datetime import date, time, datetime
```

Igual las funciones, siempre trata de colocarlas luego de las importaciones:

```
def extr_any(data):
    any_actual=int(datetime.now().year)
    any_naix=int(str(cadena[-4:]))
    any_naix=int(data.year)
    return(any_actual-any_naix)
```

Estas en un jupyter, no es necesario que hagas todo en una misma celda, si quieres cambiar algo tendrás que ejecutar todo lo que tienes allí. No es óptimo!

Aquí la idea era que me mostraras lo que tienes en el powerbi, una captura de tu código y lo que se genera. Esto no:



Sino mas bien algo tipo:

The screenshot shows a Power BI interface with a histogram titled "Distribució de les vendes de les transaccions". The histogram displays the frequency of sales across different value ranges. Below the histogram, there is a script editor window with the following Python code:

```

1 # El código siguiente, que crea un datafram y quita las filas duplicadas, siempre se ejecuta y actua
2 # como un preámbulo del script:
3 # * dataset = pandas.DataFrame(amount)
4 # * dataset = dataset.drop_duplicates()
5 #
6 # Recuer o escribir aquí el código de script:
7 #
8 import matplotlib.pyplot as plt
9 import numpy as np
10
11 #estorna una variable de Seaborn, un objecte de seaborn
12 #apareix(plt)
13
14 #seaborn.set(plt)

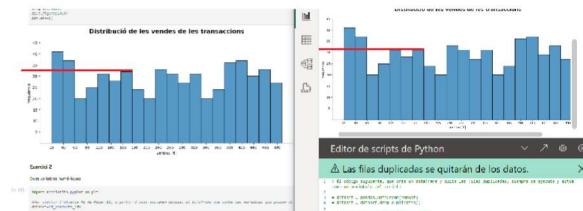
```

Y colocas el código de python en celdas markdown y no de python, porque, si son código, pero que has implementado en powerbi entonces, cuando queremos mostrar código en celdas markdown se colocan dentro de 3 acentos:

~ ~ ~
 $y = ax+b$
 ~ ~ ~

Aquí lo más importante es que entiendas la información que tienes en tus df en python para que puedas adaptarlo a lo que tienes en power bi.

El primer ejercicio no está del todo perfecto. Fíjate que tu primera línea donde se genera el script de python te dice que se van a eliminar todos los duplicados y pues al seleccionar solamente amount, todas las transacciones que tienen el mismo monto las eliminó. Habría que comparar los detalles de los tamaños del dataset que queda pero fíjate que podemos mirar en la gráfica:



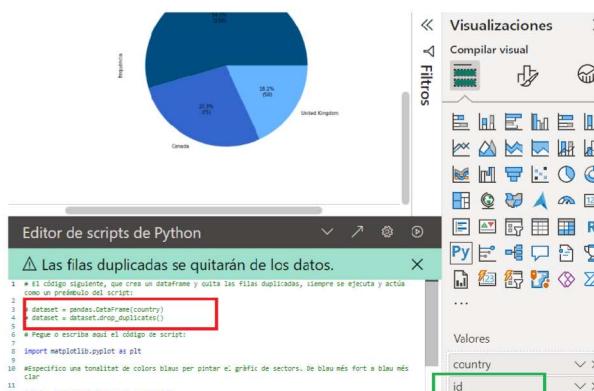
No es la única afectada, pero visualmente es la más notoria.

La idea no es que hagas una adaptación en python de lo que generas en powerby es solo copiar y pegar tal cual lo que tienes en powerbi y ya, como si lo llevaras a un word.

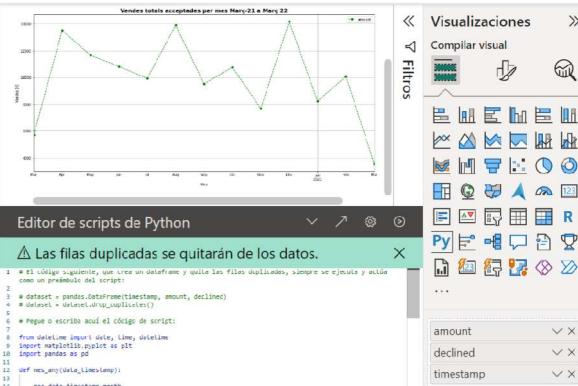
El ejercicio 2 está perfecto a pesar que tampoco tomas en cuenta el id.

El ejercicio 3 está perfecto! en este sí que pusiste lo del id. Sigo insistiendo lo de los comentarios en celdas de código, es una buena práctica de programación colocar solo comentarios necesarios para el código.

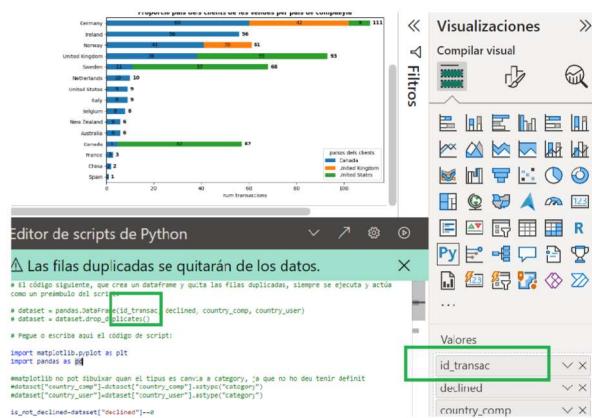
Aunque agregaste el id mucho después de comenzar a escribir el código, mira que en el cuadrito rojo no tienes generado desde el principio el dataset con el id.



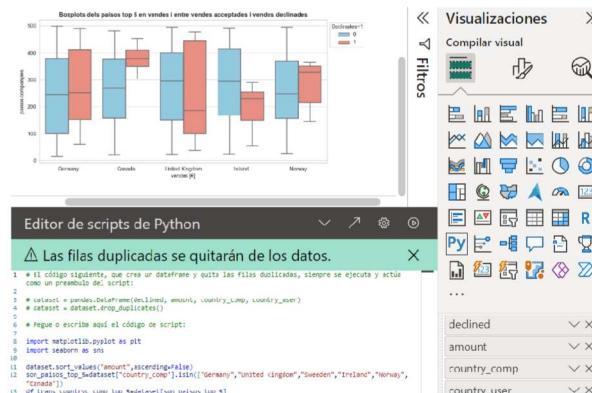
El ejercicio 4 está perfecto, pero no tomas en cuenta el id.



El 5 sí que está perfecto! Este sí que tiene el id_transac desde el inicio:

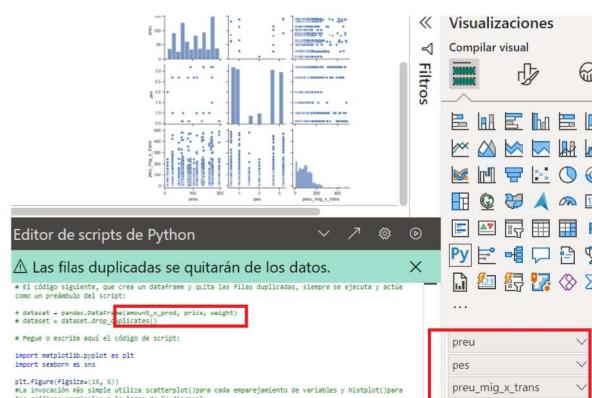


El ejercicio 5 también está como los anteriores, bien pero sin id.



El ejercicio 7 está bien en general pero no incluyes ningún id. Ten cuidado con esto porque puedes estar perdiendo información sin darte cuenta!

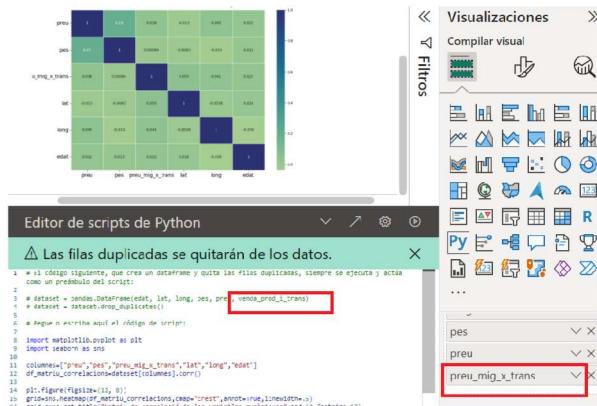
Ahora en este caso declaraste al inicio unas variables y terminaste con otras.



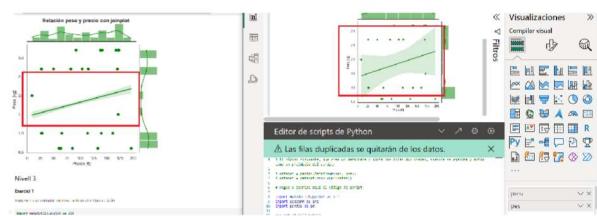
Porfa limpia los códigos, no dejes códigos comentados! Limpia

El ejercicio 1 del nivel 2 un poco mas de lo mismo, está muy bien pero no estás tomando en cuenta los id, si por casualidad hay mas transacciones donde esas columnas casualmente tienen los mismos

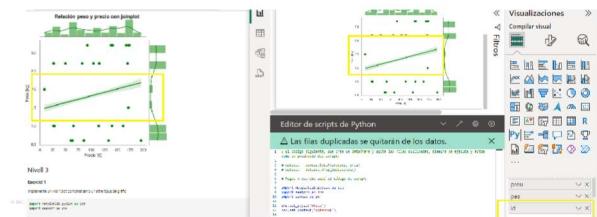
valores no te lo tomará en cuenta.



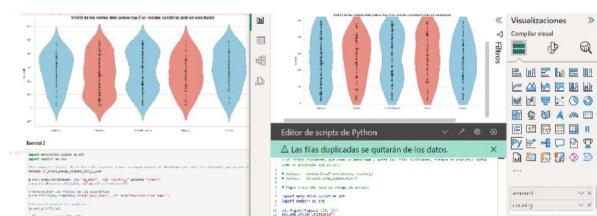
El ejercicio 2 mira que aquí si que le afecta en gran medida visualmente a la gráfica el hecho que no tengas el id seleccionado.



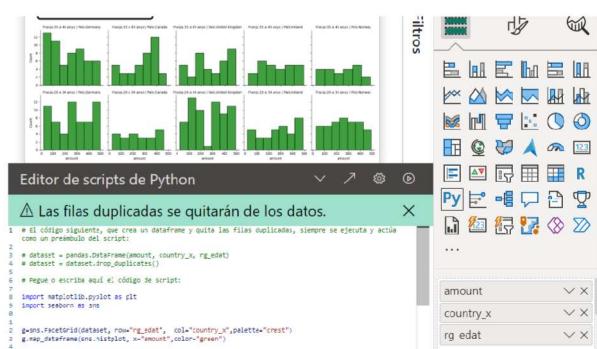
Mira la diferencia si le agrego el id de tu df_prew_pes_amount.



El ejercicio 1 del nivel 3 un poco más de lo mismo, si no seleccionas id te mostrará datos únicos.



Muy bien el ejercicio 1 del nivel 3, faltan los id.

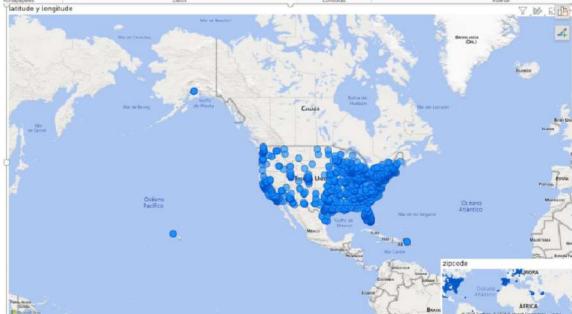


Siempre siempre agrega id!

Un saludo,

Alana

Element de qualificació	Ponderació calculada	Qualificació Gamma Percentatge Retroacció				Contribució al total del curs
TASCA  Avaluació de la prova de nivell II	-	✓ 10.00	0-10	100.00 %	Hola Gorka, Muy buena prueba! Siempre parte de lo mas sencillo y después evalúa si vale la pena o de verdad necesitas complicarlo mas. De resto super bien! Conceptos y fundamentos excelentes!	-

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció	
						Hola Gorka,	Contribució al total del curs
						Muy buen trabajo!	
						El ejercicio 1 te comento por ítems:	
						<ul style="list-style-type: none"> ○ Perfecto! ○ Muy bien! Exacto, siempre hay muchas maneras de solucionar. ○ Perfecto! ○ Muy bien! Exactamente ese sería el primer registro de usuario porque hay que tomar en cuenta el _id como en el 1. ○ Perfecto! Sí, exactamente, solo que no te piden que sea exclusivamente de comedia, sin embargo ambos argumentos son correctos! 	
						El ejercicio 2 está perfecto!	
						El ejercicio 3 está perfecto, mira que de las dos formas te da exactamente lo mismo. El número cambiará si le quitas los corchetes para que no sea exclusivamente de USA.	
						El ejercicio 1 del nivel 2 está bastante bien, jajajaja con lo que comentamos el otro día me esperaba que me entregaras la cantidad de comentarios que tenía cada usuario con ese dominio! No me esperaba la media!	
TASCA						Aquí también puedes hacer el uso del regex y decirle algo tipo {'email': {'\$regex': 'gameofthron.es'}}.	
 Tasca S9.01.		 9.25	0-10	92.50 %		El ejercicio 2 está muy bien pero fíjate que te preguntan cuántos, faltó agregarle a ese group el count!	
Consultes amb MongoDB						El ejercicio 1 del nivel 3 está perfecto!	
						El ejercicio 2 la idea era hacerlo en el mismo mongodb no en Power BI, solo debes ir a la pestaña de esquemas, analizar y buscar dentro de locations la variable geo que tiene latitud y longitud y lo muestra en el mapa.	
							
						Un saludo,	
						Alana	
						Hola Gorka,	
						Agredo tu media entre los demás sprint para que no te afecte el promedio total.	
TASCA						Un saludo,	
 Tasca 10. Projecte final		 9.14	0-10	91.40 %		Alana	

Element de qualificació	Ponderació	calculada	Qualificació	Gamma	Percentatge	Retroacció
-------------------------	------------	-----------	--------------	-------	-------------	------------

AGREGACIÓ

\bar{x} Total del curs
Mitjana ponderada de
les qualificacions.

0-100

[Privacy Policy](#)