



Ampliant la caixa d'eines: *common table expression* i funcions analítiques

i

El dimoni és en els detalls: optimització de la base de dades en funció del seu ús.

NOM I COGNOMS: _____

L'empresa Indústria del Cubito ja disposa de la base de dades per al seu ERP, que hem construït durant el transcurs de l'assignatura **Bases de dades per a *Data Warehousing***. Vol implementar una sèrie de millores i, novament, s'ha posat en contacte amb nosaltres perquè implementeu els requisits que ens han proposat.

Per implementar aquesta PAC, heu de crear una base de dades nova anomenada **dbdw_pac4** i executar l'script adjunt **BBDD_ERP_structure.sql**. El segon script proporcionat (**BBDD_ERP_data.sql**) us donarà un conjunt de dades que us permetrà implementar els components requerits als diferents apartats d'aquesta PAC. **NOTA**: les dades proporcionades no es poden modificar i els exercicis s'han de fer en base a aquestes dades.

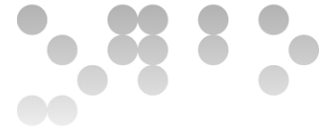
Consideracions per al lliurament i realització de la PAC:

- Tot el que es demana en aquesta PAC està explicat als blocs didàctics 4 i 5.
- Sigueu fidels a l'enunciat. Si alguna cosa no és clara, si us plau comenteu-ho al fòrum o directament al correu del professor col·laborador.
- Es recomana la utilització de **pgAdmin** per a la implementació de tota la PAC. Hi ha una altra alternativa que és **psql** (línia d'ordres), però és preferible que utilitzeu pgAdmin ja que és una interfície gràfica que us permetrà editar i crear sentències SQL (així com mostrar els resultats) de forma més senzilla que **psql**.
- Tal com s'indica a l'enunciat, cada resposta dels exercicis que calgui s'ha de lliurar en un fitxer **.sql** diferent, amb el nom corresponent. S'avaluarà el codi lliurat en aquests fitxers **.sql** i **NO el codi que aparegui al document amb les captures de pantalla**.
- Les captures de pantalla dels exercicis (i explicacions pertinents) s'han de proporcionar en un document a part (es proporciona una plantilla per al cas, **indiqueu el vostre nom al document**, si us plau).
- S'ha de fer el lliurament de tots els fitxers de la PAC (tant els fitxers **.sql** com el document amb explicacions i captures de pantalla) en un fitxer comprimit **.zip**.



Consideracions per a l'avaluació de l'exercici:

- Es tindrà en compte l'aplicació de les bones pràctiques de codificació en SQL, de consultes i de programació de procediments i disparadors. És a dir: codi amb sagnat, ús de clàusules SQL de forma correcta, comentaris, capçaleres en el procediment, etc.
- Els *scripts* proporcionats per l'estudiant amb les solucions dels exercicis s'han d'executar correctament. L'estudiant ha d'assegurar-se que llençant l' *script* complet de cada exercici no produeixi cap error.
- **Important** : Les sentències SQL proporcionades als *scripts* han de ser creades de forma manual i no mitjançant assistents que PostgreSQL/pgAdmin puguin proporcionar. Es pretén aprendre SQL i no utilitzar assistents.
- Les sentències SQL proporcionades en els exercicis han de ser **només** aquelles que demana l'enunciat i cap altra. Qualsevol sentència afegida, si està malament o provoca que l' *script* no s'executi correctament a l'hora de corregir-ho, penalitzarà la puntuació de l'exercici.



EXERCICI 1 (30%)

L'empresa Indústria del Cubito vol realitzar unes millores més a la base de dades. Per això ens han contractat novament per a demanar-nos els següents punts:

- 1) (10%) Actualment la taula *tb_cars* té per clau la matrícula del vehicle. Es requereix tenir una nova clau autonumèrica. Per a resoldre-ho ens demanen:
 - Crear una seqüència anomenada *seq_cars_id* que comenci per 100 i que vagi incrementant en una unitat.
 - Afegir el camp *cars_id* de tipus enter, sense nuls. És molt important que el camp *cars_id* per defecte utilitzi la seqüència *seq_cars_id* quan s'afegeixin nous cotxes.
 - Modificar la clau primària de *cars_registration* a *cars_id*. Per això cal substituir les claus forànies que apunten a *cars_registration* per *cars_id*. Això implica afegir el camp *cars_id* com a clau forània de *tb_cars*, assignar el seu valor al *cars_id* corresponent a *cars_registration*, i finalment eliminar la columna *cars_registration*.
 - Per a continuar amb la resta d'exercicis, tornar a executar els *scripts* per a restablir la base de dades original.
- 2) (10%) Defineix, **mitjançant CTE recursives**, una consulta que retorni els sortidors i el total de litres per sortidor, només d'aquells sortidors que hagin repostat més de 300 litres:

	pm_id integer	parent_list text	liters bigint
1	23	GS01 -> Surtidor N°1 -> Diesel	371
2	25	GS01 -> Surtidor N°2 -> Diesel	362
3	33	GS02 -> Surtidor N°2 -> Diesel	308
4	35	GS02 -> Surtidor N°3 -> Diesel	384

3) (10%) Defineix, **mitjançant funcions analítiques**, la sentència SQL que obtingui el total de litres repostats de cada cotxe. En concret es demana:

- Matrícula i model del cotxe.
- Una columna amb el número de repostatge relatiu al cotxe.
- El codi de la gasolinera.
- Una columna amb el número de repostatge relatiu al cotxe en aquesta gasolinera.
- La data i els litres repostats.
- Una columna amb la suma de litres acumulada del cotxe.
- La informació ha d'estar ordenada per matrícula, gasolinera i data.

Les sentències SQL d'aquests 3 apartats s'han d'entregar en un fitxer anomenat **pac4_ex1.sql**.



cars_registration character (7)	cars_model character varying (50)	position bigint	gas_id character (5)	position bigint	rf_date date	rf_liters integer	sum bigint
0019GVM	MERCEDES-VITO	1	GS01	1	2022-08-01	47	47
0019GVM	MERCEDES-VITO	2	GS01	2	2022-08-26	33	80
0019GVM	MERCEDES-VITO	3	GS01	3	2022-09-05	55	135
0019GVM	MERCEDES-VITO	4	GS01	4	2022-09-10	34	169
0019GVM	MERCEDES-VITO	5	GS02	1	2022-08-06	51	220
0019GVM	MERCEDES-VITO	6	GS02	2	2022-08-31	45	265
0019GVM	MERCEDES-VITO	7	GS02	3	2022-09-15	53	318
0019GVM	MERCEDES-VITO	8	GS03	1	2022-08-11	39	357
0019GVM	MERCEDES-VITO	9	GS03	2	2022-09-20	43	400
0019GVM	MERCEDES-VITO	10	GS04	1	2022-08-16	44	444
0019GVM	MERCEDES-VITO	11	GS05	1	2022-08-21	42	486
0815GYR	MERCEDES-VITO	1	GS01	1	2022-08-08	32	32
0815GYR	MERCEDES-VITO	2	GS01	2	2022-08-14	32	64
0815GYR	MERCEDES-VITO	3	GS01	3	2022-08-18	31	95



EXERCICI 2 (35%)

Donades les següents transaccions T1 y T2:

T1
SELECT inv_amount FROM tb_invoice WHERE inv_id=1;
COMMIT

T2
SELECT inv_amount FROM tb_invoice WHERE inv_id=1;
UPDATE tb_invoide SET inv_amout=20 WHERE inv_id=1;
ROLLBACK

- (10%) Indiqueu els tipus d'interferències que es poden produir entre T1 i T2 i el nivell d'aïllament necessari, segons SQL estàndar, per evitar-les.
- (10%) Que T1 i T2 produeixin interferències implica que sempre es produiran? Justifiqueu la resposta.
- (15%) Detalla tots els possibles horaris d'execució de T1 i T2. I per a cadascun d'ells indica el tipus d'interferència que genera.



EXERCICI 3 (20%)

L'empresa Indústria del Cubito ens demana que realitzem unes tasques d'investigació i pràctica sobre PostgreSQL i el seu complement **postgres_fdw**, com a nous experts que som.

Volen conèixer més **postgres_fdw**, per a poder contestar les següents qüestions que es plantegen:

- (5%) Per a què serveix el complement **postgres_fdw**?
- (15%) Poseu un exemple mínim per a veure el funcionament d'aquest complement.



EXERCICI 4 (15%)

a) (5%) Identifica quin és el pla d'execució que utilitza PostgreSQL per a resoldre la següent consulta. Fes captura de la pantalla del pla d'execució resultant i descriu-lo amb les teves pròpies paraules.

```
SELECT * FROM tb_refueling r INNER JOIN tb_cars c
ON c.cars_registration=r.cars_registration
```

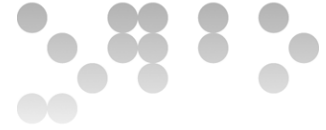
b) (10%) PostgreSQL permet habilitar o deshabilitar l'algorisme *SeqScan* utilitzat en les *join* mitjançant les següents sentències:

```
SET enable_seqscan = off;
SET enable_seqscan = on;
```

Completa la següent taula amb el pla d'execució de la *join* de l'apartat a) en cada cas. Per a completar la columna "Sense índex" haureu d'eliminar la clau principal de *tb_cars*. Per a la columna "Amb Btree" haureu de crear un índex de tipus *btree*, i eliminar-lo abans de completar la següent columna. I per a la última columna haureu de crear un índex de tipus hash.

	Sense índex	Amb Btree	Amb Hash
seq_scan=ON			
seq_scan=OFF			

Finalment justifica els resultats obtinguts.



Criteris de valoració

A l'enunciat s'indica el pes/valoració de cada exercici.

Per aconseguir la puntuació màxima als exercicis, cal explicar amb claredat la solució que es proposa.

Format i data de lliurament

Heu d'enviar la PAC a la bústia de lliurament i registre d'AC disponible a l'aula (apartat Avaluació). El format del fitxer que conté la vostra solució pot ser .pdf, .doc i .docx. Per a altres opcions, si us plau, contacteu prèviament amb el vostre consultor. El nom del fitxer ha de contenir el codi de l'assignatura, el vostre cognom i el vostre nom, així com el número d'activitat (PAC4).

La data límit per lliurar la PAC és el **23/01/2023**.

Nota: Propietat intel·lectual

Al presentar una pràctica o PAC que faci ús de recursos aliens, s'ha de presentar juntament amb ella un document en el qual es detallin tots ells, especificant el nom de cada recurs, el seu autor, el lloc on es va obtenir i el seu estatus legal: si l'obra està protegida pel copyright o s'acull a cap altra llicència d'ús (*Creative Commons*, llicència GNU, GPL etc.).

L'estudiant s'haurà d'assegurar que la llicència que sigui no impedeixi específicament el seu ús en el marc de la pràctica o PAC. En el cas de no trobar la informació corresponent haurà d'assumir que l'obra està protegida pel copyright. A més, serà necessari adjuntar els fitxers originals quan les obres utilitzades siguin digitals, i el seu codi font si així correspon.