

## De la creació a la manipulació d'una base de dades relacional

**NOM I COGNOMS:** \_\_\_\_\_

El Comitè Olímpic UOC vol disposar d'una base de dades per a registrar tota la informació sobre els Jocs Atlètics Olímpics UOC. Per a això, el departament de sistemes d'informació de l'empresa s'ha posat en contacte amb la UOC i ens ha demanat que siguin els alumnes de l'assignatura **Bases de dades per a *Data Warehousing*** els que s'encarreguin d'aquesta feina.

Consideracions per a l'entrega i realització de la PAC:

- Tot allò que es demana en aquesta PAC està explicat al bloc didàctic 2 (excepte que es tracti d'un exercici d'investigació, que en tal cas ho especificarà l'enunciat clarament). No és necessari revisar material d'altres blocs didàctics posteriors per a la realització d'aquesta PAC.
- Es recomana la utilització de **pgAdmin** per a la implementació de tota la PAC. Existeix una altra alternativa que és **psql** (línia de comandes), però és preferible que utilitzeu pgAdmin ja que aquest disposa d'una interfície gràfica que us permetrà editar i crear sentències SQL (així com mostrar els resultats) de forma més senzilla que amb psql.
- Tal com s'indica a l'enunciat, cada resposta als exercicis s'ha d'entregar en un fitxer .sql diferent, amb el nom corresponent. **S'avaluarà el codi entregat dins d'aquests fitxers .sql** i NO el codi que hi hagi al document o amb captures adjuntes.
- Les captures de pantalla dels exercicis (i les explicacions pertinents) s'han de proporcionar en un document separat; es proporciona una plantilla per a poder fer-ho (indiqueu el vostre nom al document).
- S'han de lliurar tots els fitxers de la PAC (tant els fitxers .sql com el document amb les explicacions i captures de pantalla) en un fitxer comprimit .zip.

Consideracions per a l'avaluació de l'exercici:

- Els *scripts* proporcionats per l'estudiant amb les solucions dels exercicis s'han de poder executar correctament. L'estudiant s'ha d'assegurar que l'execució de l'*script* complet de cada exercici no provoca cap error.
- **Important:** les sentències SQL proporcionades als *scripts* han de ser creades de forma manual i no mitjançant assistents de PostgreSQL/pgAdmin. Es pretén aprendre SQL i no la utilització d'assistents.

Les sentències SQL proporcionades als exercicis han de ser únicament aquelles que demana l'enunciat i cap més. Qualsevol sentència "extra" afegida pot penalitzar l'exercici si està malament o provoca que l'*script* no obtingui els resultats esperats.

### EXERCICI 1 (25%)



Es demana proporcionar les sentències SQL per crear els components especificats a continuació. El resultat d'aquest exercici s'ha de lliurar al fitxer **pac2\_ex1.sql**.

## pac2

Nom de la base de dades on s'han de crear l'esquema i els components requerits.

**Nota:** la comanda de creació de la base de dades s'ha de proporcionar comentada a l'script.

## olimpic

Nom de l'esquema de la base de dades on s'han de crear els components especificats a continuació.

## tb\_discipline

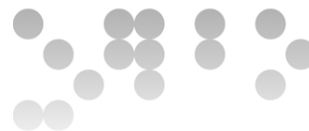
Taula per a guardar informació de les diferents disciplines esportives dels jocs olímpics. Es vol guardar informació segons l'especificació proporcionada a continuació. El tipus de disciplina (*type*) només pot tenir valors RUN, JUMP o THROW.

Taula: tb_discipline Esquema: olimpic					
Nom columna	Tipus de dades	Accepta Nuls	Clau primària	Clau forana	Valor per defecte
discipline_id	Numèric enter	No	Si		
name	Cadena de 50 caràcters variable	No	No		
inventor	Cadena de 50 caràcters variable	No	No		
type	Cadena de 10 caràcters variable	No	No		
object_type	Cadena de 20 caràcters fixe	Si	No		

## tb\_athlete

Taula per a guardar informació dels atletes. guardar información acerca de los atletas. Es vol guardar informació segons l'especificació proporcionada a continuació.

Taula: tb_athlete Esquema: olimpic					
Nom columna	Tipus de dades	Accepta Nuls	Clau primària	Clau forana	Valor per defecte
athlete_id	Cadena de 7 caràcters fixe	No	Si		
name	Cadena de 50 caràcters variable	No	No		
country	Cadena de 3 caràcters fixe	No	No		
substitute_id	Cadena de 7 caràcters fixe	Si	No	tb_athlete	

**tb\_play**

Taula per a guardar informació de les disciplines dels atletes. Es vol guardar informació segons l'especificació proporcionada a continuació.

Taula: tb_play Esquema: olimpic					
Nom columna	Tipus de dades	Accepta Nuls	Clau primària	Clau forana	Valor per defecte
athlete_id	Cadena de 7 caràcters fixe	No	Si	tb_athlete	
discipline_id	Numèric enter	No	Si	tb_discipline	

**tb\_round**

Taula per a guardar informació de les tandes. Es vol guardar informació segons l'especificació proporcionada a continuació.

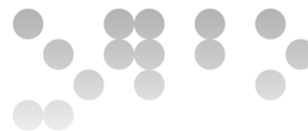
Taula: tb_round Esquema: olimpic					
Nom columna	Tipus de dades	Accepta Nuls	Clau primària	Clau forana	Valor per defecte
round_number	Numèric enter	No	Si		
discipline_id	Numèric enter	No	Si	tb_discipline	

**tb\_register**

Taula per a guardar informació dels resultats dels atletes en una tanda. Es vol guardar informació segons l'especificació proporcionada a continuació.

Taula: tb_register Esquema: olimpic					
Nom columna	Tipus de dades	Accepta Nuls	Clau primària	Clau forana	Valor per defecte
athlete_id	Cadena de 7 caràcters fixe	No	Si	tb_athlete	
round_number	Numèric enter	No	Si	tb_round	
discipline_id	Numèric enter	No	Si	tb_round	
register_date	Data	No	No		Data Actual
register_position	Numèric enter	Si	No		
register_time	Temps	Si	No		
register_measure	Real	Si	No		

Un cop executades les sentències SQL de creació de les estructures que es requereixen, es demana executar l'script **BBDD\_Olimpic\_data.sql** adjunt. Aquest fitxer NO pot ser modificat per a la realització de la PAC i cal que totes les sentències SQL del fitxer s'executin correctament. Si es troben indicis que el fitxer s'ha modificat o que les dades estan incompletes, tota la PAC serà considerada com no vàlida.



## EXERCICI 2 (25%)

Es demana proporcionar, per a cada apartat, **una única consulta en llenguatge SQL**, i el **resultat d'aquesta consulta mitjançant una captura del resultat en pgAdmin**. Per a poder realitzar aquest exercici, el primer exercici ha d'estar finalitzat de forma correcta.

Les consultes SQL d'aquest exercici s'han d'entregar en un fitxer anomenat **pac2\_ex2.sql** (les captures de pantalla, i explicacions si són necessàries, s'han d'entregar en un document adjunt).

Els resultats de les consultes no han de tenir repeticions en cap cas, per la qual cosa l'eliminació de resultats pot ser necessària en alguna de les consultes. D'altra banda, l'ús innecessari de la clàusula DISTINCT serà penalitzat.

**Nota:** La instal·lació per defecte de PostgreSQL és *case sensitive*; això significa que la cadena de caràcters "ALUMNE UOC" és diferent de "Alumne UOC" i de "alumne uoc".

**Nota:** és possible que per a alguna de les consultes existeixi més d'una solució. En aquest cas, qualsevol de les solucions es considerarà vàlida, sempre i quan compleixi amb els aclariments proporcionats a l'enunciat.

- a) Proporcionar els atletes espanyols que el seu nom comenci per PE, ordenats descendentment pel seu identificador. Volem obtenir l'identificador, nom, país i identificador del substitut.

	athlete_id [PK] character (7)	name character varying (50)	country character varying (3)	substitute_id character (7)
1	1358251	PENA Inigo	ESP	1302771
2	1349249	PEREZ Maria Isabel	ESP	1307950
3	1349246	PEREZ Marta	ESP	1324396
4	1349245	PEREZ Maria	ESP	1334678
5	1349243	PELETEIRO Ana	ESP	1304891
6	1325882	PEREZ BLANCO Jimena	ESP	1302772



- b) Proporcionar la llista d'atletes francesos que practiquen alguna disciplina de salt, ordenats ascendentment pel nom de la disciplina i descendentment pel seu nom. Volem obtenir el nom de la disciplina, nom de l'atleta i identificador de l'atleta.

	name character varying (50) 🔒	name character varying (50) 🔒	athlete_id character (7) 🔒
1	BMX Cycling	VALENTINO Manon	1319979
2	BMX Cycling	MAHIEU Romain	1319977
3	BMX Cycling	ETIENNE Axelle	1319978
4	BMX Cycling	DAUDET Joris	1319973
5	BMX Cycling	ANDRE Sylvain	1319971
6	Trampoline	MORANTE Allan	1334953
7	Trampoline	LABROUSSE Lea	1334991

- c) Proporcioneu la disciplina amb més atletes participants. Volem obtenir l'identificador de la disciplina, el nom i el total d'atletes.

	discipline_id [PK] integer	name character varying (50)	count bigint 🔒
1	2	Athletics	2041

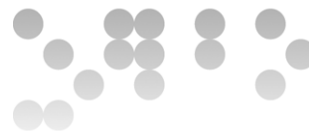
- d) Proporcioneu la llista d'atletes que participen en més d'una disciplina. Volem veure l'id de l'atleta, el nom, el país i el total de disciplines que practica.

	athlete_id [PK] character (7)	name character varying (50)	country character varying (3)	count bigint 🔒
1	1467919	HAVIK Yoeri	NED	2
2	1330503	BRENNAUER Lisa	GER	2
3	1346205	GANNA Filippo	ITA	2
4	1330510	KLEIN Lisa	GER	2
5	1342530	DYGERT Chloe	USA	2
6	1393561	KOPECKY Lotte	BEL	2



- e) Proporcioneu la llista d'atletes que han participat en més tandes. Volem veure l'id de l'atleta, el nom, la disciplina i el total de tandes en les que ha participat.

	<b>athlete_id</b> character (7) 🔒	<b>name</b> character varying (50) 🔒	<b>discipline_id</b> integer 🔒	<b>name</b> character varying (50) 🔒	<b>count</b> bigint 🔒
1	1478152	PROSTOROV Mykola	15	Trampoline	4
2	1429192	KOCHESOK Susana	15	Trampoline	4



### EXERCICI 3 (25%)

Es demana proporcionar, per a cada apartat, **una única sentència en llenguatge SQL sempre que sigui possible**, i els seus resultats mitjançant una captura en pgAdmin. Per a poder realitzar aquest exercici, el 1r exercici ha d'estar finalitzat.

Les sentències SQL d'aquest exercici s'han d'entregar en un fitxer anomenat **pac2\_ex3.sql** (les captures de pantalla s'han d'entregar en un document adjunt).

a) Ens diuen que cal inserir 3 nous atletes segons l'especificació següent:

Nom columna	Atleta 1	Atleta 2	Atleta 3
athlete_id	0000001	0000002	0000003
name	REMBRAND Luc	SMITH Mike	LEWIS Carl
country	FRA	ENG	USA
substitute_id	NULL	NULL	NULL

- b) Ens demanen que afegim una nova restricció per a impedir que hi hagi atletes espanyols sense substitut.
- c) Implementeu, en una **única sentència**, una vista **olimpic.exercise33** sobre la consulta a) de l'exercici 2. Assegureu-vos que els usuaris no puguin inserir valors que inicialment no s'han de mostrar a través de la vista.
- d) Ens demanen que s'afegeix-hi una columna en la taula **tb\_athlete** anomenada **date\_add**, que representa la data d'alta a la federació del seu país. Aquest camp no pot tenir valors nuls i, per defecte, serà la data actual.
- e) S'ha de crear un usuari de sistema (compte, no el confongueu amb **tb\_users**) **registerer** amb accés a l'esquema **olimpic** i contrasenya **1234**. Aquest usuari podrà realitzar lectures, insercions, actualitzacions i esborrats (mai truncaments) de la taula **tb\_register** i permisos de lectura de la taula **tb\_athlete**. No podrà accedir a les altres taules de l'esquema **olimpic**. Assigneu els permisos necessaris a aquest usuari mitjançant SQL. Assegureu també que aquest usuari no pugui assignar permisos a altres usuaris sobre aquestes taules. Referència a la documentació:

<https://www.postgresql.org/docs/12/sql-grant.html>

**Nota:** Podeu provar els permisos creant una nova connexió amb l'usuari creat i mantenir així les dues sessions obertes. Per a això, heu de crear un nou servidor al Browser de pgAdmin amb les credencials de l'usuari nou.



#### EXERCICI 4 (25%)

Des de l'empresa ens demanen que realitzem unes tasques de comprensió i pràctica sobre PostgreSQL, com a nous experts que sou.

Contesteu breument els següents apartats:

- a) Expliqueu les clàusules GROUPING SET, ROLLUP i CUBE de PostgreSQL. L'explicació s'ha de complementar amb exemples d'elaboració pròpia. **(15%)**

Aquestes clàusules permeten definir diferents tipus d'agrupaments (GROUP BY).

- GROUPING SET: equivalent a fer diferents agrupaments simples GROUP BY tot combinant els resultats. El resultat conté tots els camps dels diferents agrupaments amb valors NULL quan no són utilitzats.

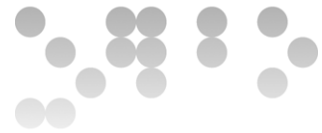
```
SELECT country, substitute_id, COUNT(*)
FROM olimpico.tb_athlete
WHERE country IN ('ESP','FRA','ENG') AND substitute_id IS NOT NULL
GROUP BY GROUPING SETS((country),(substitute_id))
HAVING substitute_id IS NULL OR count(*)>2
```

	country character varying (3) 🔒	substitute_id character (7) 🔒	count bigint 🔒
1	ESP	[null]	132
2	ENG	[null]	1
3	FRA	[null]	188
4	[null]	1889067	4
5	[null]	1341483	3
6	[null]	0000001	4
7	[null]	1307950	3

- ROLLUP: a partir d'una llista de camps genera el GROUPING SET de tots els subconjunts de camps continus començant per l'esquerra. Per exemple:

```
ROLLUP ( e1, e2, e3, ... )
```





```
GROUPING SETS (
    ( e1, e2, e3, ... ),
    ...
    ( e1, e2 ),
    ( e1 ),
    ( )
)
```

```
SELECT country, substitute_id, COUNT(*)
FROM olimpic.tb_athlete
WHERE country IN ('ESP','FRA','ENG') AND substitute_id IS NOT NULL
GROUP BY ROLLUP (country,substitute_id)
HAVING substitute_id IS NULL OR count(*)>1
```

	country character varying (3)	substitute_id character (7)	count bigint
1	[null]	[null]	321
2	ESP	1485432	2
3	FRA	1324460	2
4	ESP	1476208	2
5	ESP	1341483	2
6	ESP	0000001	2
7	FRA	1360801	2
8	FRA	1337490	2
9	FRA	1347044	2
10	FRA	1307950	2
11	ESP	1285817	2
12	FRA	1393269	2
13	ESP	1889067	2
14	FRA	1333346	2
15	FRA	1330891	2
16	FRA	1431591	2
17	FRA	1485427	2
18	FRA	1429019	2
19	ESP	1333472	2
20	FRA	1889067	2
21	ESP	[null]	132
22	ENG	[null]	1
23	FRA	[null]	188



- CUBE: la clàusula CUBE permet generar totes les combinacions dels camps indicants.

```
CUBE ( a, b, c )
```

```
GROUPING SETS (
  ( a, b, c ),
  ( a, b   ),
  ( a,    c ),
  ( a     ),
  (    b, c ),
  (    b   ),
  (      c ),
  (       )
)
```

### Referències:

<https://www.postgresql.org/docs/12/queries-table-expressions.html#QUERIES-GROUPING-SETS>

<https://www.postgresqltutorial.com/postgresql-grouping-sets/>

- b) Explicar les columnes generades (clàusula GENERATED ALWAYS AS) de PostgreSQL. L'explicació ha de ser complementada amb exemples d'elaboració pròpia. **(10%)**

Permet definir columnes calculades en una taula. Una columna calculada és aquella el valor de la qual es calcula a partir d'altres columnes no calculades, i funcions inmutables. No poden formar part de la clau primària però sí d'una clau forana.

La clàusula GENERATED ALWAYS pot aparèixer en les sentències CREATE TABLE o ALTER TABLE.

La clàusula STORED és obligatòria, i indica que els valors de la columna generada es guarden en disc. Altres SGBD com MySQL permeten l'opció VIRTUAL, que calcula el valor de la columna calculada a l'aire.

```
CREATE TABLE olimpique.people (
  id INT NOT NULL,
  first_name CHARACTER VARYING(50) NOT NULL,
  last_name CHARACTER VARYING(50) NOT NULL,
  full_name CHARACTER VARYING(100) GENERATED ALWAYS AS (first_name||' '||last_name) STORED,
  CONSTRAINT pk_people PRIMARY KEY (id)
);
```

```
INSERT INTO olimpique.people VALUES (1, 'John', 'Smith');
```

```
SELECT * FROM olimpique.people;
```



	id [PK] integer	first_name character varying (50)	last_name character varying (50)	full_name character varying (100)
1	1	John	Smith	John Smith

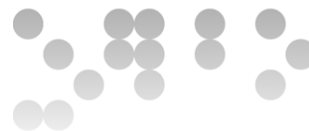
```
ALTER table olimpic.tb_athlete
ADD COLUMN description VARCHAR(100) GENERATED ALWAYS AS (country||'-'||name) STORED;
```

	athlete_id [PK] character (7)	name character varying (50)	country character varying (3)	substitute_id character (7)	date_add date	description character varying (100)
1	1346266	AALERUD Katrine	NOR	[null]	2021-10-04	NOR-AALERUD Katrine
2	1344792	ABASS Abobakr	SUD	[null]	2021-10-04	SUD-ABASS Abobakr
3	1328854	ABDALLA Abubaker Haydar	QAT	[null]	2021-10-04	QAT-ABDALLA Abubaker Haydar
4	1306332	ABDEL LATIF Radwa	EGY	[null]	2021-10-04	EGY-ABDEL LATIF Radwa
5	1306339	ABDEL RAZEK Samy	EGY	1306332	2021-10-04	EGY-ABDEL RAZEK Samy
6	1312038	ABDELRAHMAN Ihab	EGY	1328854	2021-10-04	EGY-ABDELRAHMAN Ihab
7	1305691	ABDELWAHED Ahmed	ITA	1312038	2021-10-04	ITA-ABDELWAHED Ahmed

## Referències:

<https://www.postgresql.org/docs/12/ddl-generated-columns.html>

<https://severalnines.com/database-blog/overview-generated-columns-postgresql>



## Criteris de valoració

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

Per aconseguir la puntuació màxima en els 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 de l'arxiu que conté la vostra solució pot ser **.pdf**, **.doc** i **.docx**. **Per a altres opcions, si us plau, contacteu prèviament amb el vostre consultor.**

La data límit per lliurar la PAC2 és el **8/11/2021**.

### 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.