L'informe descriptiu

La pàgina de botiga de `games_clean` Noms _____

Superset → **SQL → SQL Lab**, base de dades `Dades classe`, esquema `public`, taula `games_clean`. Escriviu l'informe en un document (plantilla del professor): sis parts amb encapçalament, **250 paraules com a màxim**. Per parelles: un executa consultes, l'altre escriu; canvieu a la part 4. Aquest full és per a les vostres notes i per a la revisió.

Referència · les sis parts

Un informe descriptiu és la pàgina de botiga d'un conjunt de dades: el llegiu abans d'"instal·lar" les dades en un gràfic.

Part	Pregunta que respon	A la pàgina de botiga
1 · Font	Qui les ha recollit, on, quan? Qui les ha netejat? D'una API: endpoint, paràmetres, data d'extracció.	Desenvolupador i data de llançament.
2 · Mida	Quantes files i columnes? Què és una fila? Quin període?	Mida de la baixada.
3 · Variables	Quines són les columnes importants, el seu tipus i la seva unitat?	Llista de característiques.
4 · Qualitat	Què falta, què és estimat o dubtós? Quant?	Problemes coneguts.
5 · Xifres clau	Centre, dispersió i distribució del que importa.	Puntuació de les ressenyes i etiquetes.
6 · Límits	Què **no** us poden dir aquestes dades?	Requisits del sistema: on no funcionarà.

Tasca 0 · Dos informes, les mateixes dades (4 min)

Informe A	Informe B
"El conjunt de dades té molts jocs. La majoria són barats. El preu mitjà és 7.91. Falten algunes puntuacions. Els jocs són cada cop més populars."	"26 685 jocs de Steam llançats 2004–2018. La meitat costen 4.99 USD o menys (mediana); la mitjana és 7.91 USD per uns quants productes de fins a 595.99 USD. Només el 10.7% té Metascore."

0.1 De quin us fiaríeu? Subratlleu les paraules exactes que marquen la diferència i després digueu tres coses que té i que a l'altre li falten.

0.2 Una frase de l'altre informe diu una cosa que les dades no poden provar. Quina, i per què?

Tasca 1 · Parts 1 a 3 (7 min)

La font i les variables surten de les vostres notes de la UT2. La mida necessita una consulta:

```sql
SELECT COUNT(*) AS n_games, MIN(release_date) AS first_release,
       MAX(release_date) AS last_release
FROM games_clean;
```

Columnes: _____ Primera data: _____ Última data: _____ Una fila és: _____

Tasca 2 · Parts 4 i 5 (11 min)

Reutilitzeu les consultes de dimecres passat (la fitxa estadística). Patró d'una distribució amb percentatges:

```sql
SELECT price_band, COUNT(*) AS n,
       ROUND(100.0 * COUNT(*) / SUM(COUNT(*)) OVER (), 1) AS pct
FROM games_clean GROUP BY price_band ORDER BY n DESC;
```

El mínim, cada cosa amb la seva **unitat** (USD, jocs, %):

2.1 Una xifra de dades absents amb el seu percentatge: _____

2.2 Centre i dispersió de `price`: _____

Quin centre poseu a l'informe, i per què? _____

2.3 Una distribució (`price_band` o `owners_tier`): _____

Tasca 3 · Part 6, sense cap consulta (5 min)

Què conclouria malament un periodista d'aquesta taula? Penseu en `owners_min`, `metascore` i `price`. Dos límits:

3.1 Aquestes dades no poden dir... _____

3.2 Compte: _____

Revisió entre parelles (11 min)

Doneu aquest full i el vostre informe a una altra parella. **Revisors**: marqueu sí o no. Cada "No" necessita la línia a què es refereix. Sense notes i sense adjectius: només sí, no i on. Revisat per _____

#	Comprovació	Sí / No	Si és no: quina línia?
1	Les sis parts hi són, en ordre, cadascuna amb el seu encapçalament.		
2	La font diu qui ha recollit les dades, d'on, i quins anys cobreix. Per a una API: endpoint, paràmetres i data d'extracció.		
3	La mida dona files **i** columnes, i diu què és una fila.		
4	Cada xifra té una unitat, i la seva consulta és adjunta.		
5	Per a `price`, el centre és la mediana (o totes dues), i el text diu per què.		
6	Qualitat: almenys dos problemes, cadascun amb recompte o percentatge.		
7	Els límits tenen almenys dues coses que les dades no poden dir.		
8	Cap afirmació va més enllà de les dades. 250 paraules o menys.		

R **Autors**, quan torni el full: corregiu **un** "No". Què heu canviat?

_____

Heu acabat quan Sis parts amb encapçalament, 250 paraules o menys · cada xifra té unitat i el seu SQL sota l'informe · l'altra parella ha omplert la llista · el vostre canvi és escrit aquí · informe lliurat.

Encallats?

- **`relation "games_clean" does not exist`**: demaneu `games_clean.sql` al professor i executeu-lo, o ajunteu-vos.
- **`function round(double precision, integer) does not exist`**: convertiu: `ROUND(metascore::numeric, 1)`.
- **`SUM(COUNT(*)) OVER ()` dona un error**: hi falta el `GROUP BY`.
- **Encara executeu consultes quan ja ha passat la meitat del temps?** Deixeu de consultar, comenceu a escriure.

A casa Acabeu l'informe si cal i lliureu-lo. Per a dimecres: rellegiu les vostres notes sobre `COALESCE`, `CASE WHEN`, `CREATE TABLE AS` i les sis parts. Obriu la taula `pokemon` i mireu `type_2` i `generation_id`.