



SQL DML : Data Manipulation Language 1

Fabien Coelho, Claire Medrala

Mines Paris – PSL, MESR

Janvier 2025



Opérations élémentaires sur les données

CRUD

SQL DML 1

FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

insertion de nouvelles données

```
INSERT INTO personnes(nom) VALUES ('calvin');
```

sélection de données selon un critère

algèbre relationnelle, opérations ensemblistes...

```
SELECT nom FROM personnes WHERE nom LIKE 'cal%';
```

modification de données existantes

```
UPDATE personnes SET nom='Calvin' WHERE nom='calvin';
```

effacement de données

```
DELETE FROM personnes WHERE nom='Calvin';
```



Faisons notre cinéma avec 2 tables



Généralités

SELECT

SQL DML 1

FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

but manipulation des données dans des **tables existantes**

résultat pour obtenir **une nouvelle table**

tri des lignes ascendant ou descendant

aggrégation des lignes (somme, comptage)

sous requêtes dans expressions ou tables...

Lien avec l'algèbre relationnelle

π sélection de colonnes (attributs)

```
SELECT FROM
```

σ sélection de lignes (tuples)

```
WHERE
```

\times \bowtie de plusieurs tables

```
JOIN
```

\cup \cap $-$ opérateurs ensemblistes si compatibles

```
UNION INTERSECT EXCEPT
```

SQL DML 1

FC/CM

Intro

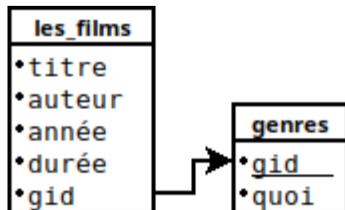
SELECT

Produit

Jointure

Aggrégation

TD



genres

gid	quoi
1	Drame
2	Comédie
3	Comédie dramatique
4	Action

les_films

titre	auteur	année	durée	gid
Citizen Kane	Wells	1936	01:59:00	1
The Dictator	Chaplin	1940	02:07:00	2
Modern Times	Chaplin	1936	01:27:00	2
City Lights	Chaplin	1931	01:27:00	3
Ohayô	Ozu	1959	01:37:00	2
Le Procès	Wells	1963	01:58:00	1



Syntaxe

SELECT



Expressions

SELECT

SQL DML 1
FC/CM

```

SELECT [ALL|DISTINCT] expression [AS colname], ...
FROM tablename [AS alias], ...
WHERE tuple condition
GROUP BY expr/colname/colnum, ...
HAVING aggregate condition
UNION/EXCEPT/INTERSECT ... SELECT ... composition
ORDER BY expr/colname/colnum [ASC|DESC], ...
LIMIT number
OFFSET start;

```

5 / 31

SQL DML 1
FC/CM

Exemples sans table

```

-- constantes
SELECT 1789 AS année,
       'révolution française' AS "évènement";

-- expressions arithmétiques
SELECT 2*PI()*6300 AS périmètre,
       'terre' AS planète;

-- expressions bizarres
SELECT EXTRACT(DOW FROM NOW()) AS jour,
       CURRENT_DATE;

```

année	évènement
1789	révolution française

périmètre	planète
39584.067435231394	terre

jour	current_date
1	2025-01-13

6 / 31



Expressions (suite)

SELECT



Projection de colonnes

SELECT ... FROM ...

SQL DML 1
FC/CM

Types, Constantes, Opérateurs, Fonctions

entiers et flottants **SMALLINT INTEGER REAL FLOAT...**

- constantes 123 1.25E12, opérations + - * / %
- fonctions diverses SQRT(2.0) PI() ABS(-1)

booléens **BOOLEAN**, valeurs **TRUE FALSE NULL**

- opérations logiques **AND OR NOT...**
- résultat des comparaisons = <> > <=...

textes **TEXT CHAR(10) VARCHAR(32)**

- constante 'hello', opérations ||, comparaison **LIKE**

autres temps, binaires, géométrie, argent, extensions...

valeur **NULL** possible pour tous les types !

7 / 31

SQL DML 1
FC/CM

- clause **FROM** précise les tables (schéma, alias)
- sélection des colonnes, avec nommage éventuel
- appliqué à chaque ligne des tables

```

-- simple
SELECT quoi FROM Genres;

```

quoi
Drame
Comédie
Comédie dramatique
Action

```

-- table
SELECT Genres.quoi
FROM Genres;

```

```

-- nommage et alias
SELECT g.quoi AS genre
FROM Genres AS g;

```

genre
Drame
Comédie
Comédie dramatique
Action

8 / 31



Projection de colonnes (suite)

SELECT ... FROM ...



Projection de toutes les colonnes

SELECT ... FROM ...

SQL DML 1
FC/CM

```
-- sélection de colonnes...
SELECT titre, année, auteur AS nom
-- précise le schéma
FROM public.les_films;
```

titre	année	nom
Citizen Kane	1936	Wells
The Dictator	1940	Chaplin
Modern Times	1936	Chaplin
City Lights	1931	Chaplin
Ohayô	1959	Ozu
Le Procès	1963	Wells

9 / 31

SQL DML 1
FC/CM

- * ou nomTable.* ou alias.*...
- si plusieurs tables, attention colonnes homonymes

```
-- toutes les colonnes
SELECT *
FROM Genres;

-- idem avec nom table
SELECT Genres.*
FROM Genres;

-- idem avec alias
SELECT g.*
FROM Genres AS g;

-- avec un raccourci
TABLE Genres;
```

gid	quoi
1	Drame
2	Comédie
3	Comédie dramatique
4	Action

10 / 31



Projection + expressions + nommage

SELECT ... FROM ...



Restriction de lignes

SELECT ... WHERE ...

SQL DML 1
FC/CM

```
-- opérations arithmétiques
SELECT titre, (année+99)/100 AS siècle
FROM les_films;
```

titre	siècle
Citizen Kane	20
The Dictator	20
Modern Times	20
City Lights	20
Ohayô	20
Le Procès	20

11 / 31

SQL DML 1
FC/CM

- clause **WHERE** condition logique
- opérateurs de comparaison < > =... avec les attributs
- opérateurs booléens **AND OR NOT**

```
-- films d'avant guerre
SELECT titre, auteur, année
FROM les_films
WHERE année < 1940;

-- films de Chaplin avant 1940
SELECT auteur, titre, année
FROM les_films
WHERE auteur = 'Chaplin'
AND année < 1940;
```

titre	auteur	année
Citizen Kane	Wells	1936
Modern Times	Chaplin	1936
City Lights	Chaplin	1931

auteur	titre	année
Chaplin	Modern Times	1936
Chaplin	City Lights	1931

12 / 31



Restriction exprimé sur l'entrée

SELECT ... WHERE ...



Restriction du nombre de lignes

LIMIT / OFFSET

SQL DML 1
FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

- condition vérifiée indépendemment de la projection

-- films de l'année 36

```
SELECT titre, auteur
FROM les_films
WHERE année = 1936;
```

titre	auteur
Citizen Kane	Wells
Modern Times	Chaplin

-- auteur d'un film

```
SELECT auteur
FROM les_films
WHERE titre = 'Citizen Kane';
```

auteur
Wells

13 / 31

SQL DML 1
FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

- limitation avec **LIMIT** *nombre* (nombre max affiché)
- décalage initial éventuel avec **OFFSET** *décalage*
- résultat plus pertinent si trié *hit-parade*

-- affiche les trois premiers films

```
SELECT titre, auteur
FROM les_films
LIMIT 3;
```

titre	auteur
Citizen Kane	Wells
The Dictator	Chaplin
Modern Times	Chaplin

-- affiche les deux suivants

```
SELECT titre, auteur
FROM les_films
LIMIT 2 -- montre deux films max
OFFSET 3; -- saute les trois premiers
```

titre	auteur
City Lights	Chaplin
Ohayô	Ozu

14 / 31



Suppression des doublons ?

DISTINCT vs ALL



Tri des lignes

ORDER BY expr/colname/colnum [ASC / DESC]

SQL DML 1
FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

suppression option **DISTINCT** de **SELECT**, ensembliste
conservation par défaut, option **ALL** non ensembliste

-- auteurs avec doublons

```
SELECT ALL auteur
FROM les_films;
```

auteur
Wells
Chaplin
Chaplin
Chaplin
Ozu
Wells

-- auteurs sans doublons

```
SELECT DISTINCT auteur
FROM les_films;
```

auteur
Chaplin
Ozu
Wells

15 / 31

SQL DML 1
FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

- par défaut résultats non déterministes (selon stockage, calculs...)
- tri selon type, lexicographique, ordre croissant ou décroissant

-- ordre de l'unité de l'année

-- puis du nom de l'auteur

```
SELECT auteur, année, titre
FROM les_films
ORDER BY année % 10 ASC, auteur;
```

auteur	année	titre
Chaplin	1940	The Dictator
Chaplin	1931	City Lights
Wells	1963	Le Procès
Chaplin	1936	Modern Times
Wells	1936	Citizen Kane
Ozu	1959	Ohayô

-- ordre alphabétique inverse des titres

```
SELECT auteur, année, titre
FROM les_films
ORDER BY 3 DESC;
```

auteur	année	titre
Chaplin	1940	The Dictator
Ozu	1959	Ohayô
Chaplin	1936	Modern Times
Wells	1963	Le Procès
Chaplin	1931	City Lights
Wells	1936	Citizen Kane

16 / 31



Tri + Restriction + Projection

SQL DML 1

FC/CM

-- films de Charlie Chaplin

```

SELECT titre, année
FROM les_films
WHERE auteur='Chaplin'
ORDER BY année ASC;

```

titre	année
City Lights	1931
Modern Times	1936
The Dictator	1940

-- films ordonnés par année

```

SELECT titre, auteur
FROM les_films
ORDER BY année ASC;

```

titre	auteur
City Lights	Chaplin
Citizen Kane	Wells
Modern Times	Chaplin
The Dictator	Chaplin
Ohayô	Ozu
Le Procès	Wells

17 / 31



Produit cartésien

CROSS JOIN ou ,

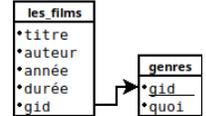
SQL DML 1

FC/CM

```

-- produit cartésien : films X genres
SELECT f.*, g.gid AS gid2, g.quoi
FROM les_films AS f CROSS JOIN Genres AS g
LIMIT 13;

```



titre	auteur	année	durée	gid	gid2	quoi
Citizen Kane	Wells	1936	01:59:00	1	1	Drame
Citizen Kane	Wells	1936	01:59:00	1	2	Comédie
Citizen Kane	Wells	1936	01:59:00	1	3	Comédie dramatique
Citizen Kane	Wells	1936	01:59:00	1	4	Action
The Dictator	Chaplin	1940	02:07:00	2	1	Drame
The Dictator	Chaplin	1940	02:07:00	2	2	Comédie
The Dictator	Chaplin	1940	02:07:00	2	3	Comédie dramatique
The Dictator	Chaplin	1940	02:07:00	2	4	Action
Modern Times	Chaplin	1936	01:27:00	2	1	Drame
Modern Times	Chaplin	1936	01:27:00	2	2	Comédie
Modern Times	Chaplin	1936	01:27:00	2	3	Comédie dramatique
Modern Times	Chaplin	1936	01:27:00	2	4	Action
City Lights	Chaplin	1931	01:27:00	3	1	Drame

18 / 31



Jointure simple (1/2)

JOIN

SQL DML 1

FC/CM

Syntaxes de jointure

générale toujours possible **JOIN ON**
 simplifiée si même nom d'attribut **JOIN USING**
 déconseillée risque d'oubli de conditions **CROSS JOIN WHERE**

-- jointure sur le genre (join on)

```

SELECT titre, auteur, année, quoi
FROM les_films AS f
JOIN Genres AS g ON f.gid = g.gid;

```

titre	auteur	année	quoi
Citizen Kane	Wells	1936	Drame
The Dictator	Chaplin	1940	Comédie
Modern Times	Chaplin	1936	Comédie
City Lights	Chaplin	1931	Comédie dramatique
Ohayô	Ozu	1959	Comédie
Le Procès	Wells	1963	Drame

19 / 31



Jointure simple (2/2)

JOIN USING, CROSS JOIN WHERE

SQL DML 1

FC/CM

-- jointure sur le genre (join using)

```

SELECT titre, auteur, année, quoi
FROM les_films
JOIN Genres USING (gid);

```

-- jointure sur le genre (X et where)

```

SELECT titre, auteur, année, quoi
FROM les_films AS f
CROSS JOIN Genres AS g
WHERE f.gid = g.gid;

```

20 / 31



Jointure + Restriction + Projection + Tri

SQL DML 1
FC/CM

```

-- les genres des films de Chaplin
SELECT titre, année, quoi
FROM les_films AS f
JOIN Genres AS g ON f.gid = g.gid
WHERE auteur = 'Chaplin'
ORDER BY année ASC;

```

titre	année	quoi
City Lights	1931	Comédie dramatique
Modern Times	1936	Comédie
The Dictator	1940	Comédie



Auto-jointure

JOIN

SQL DML 1
FC/CM

```

-- films distincts de même durée
SELECT f1.titre AS film1, f2.titre AS film2
FROM les_films AS f1
JOIN les_films AS f2 ON f1.durée = f2.durée
WHERE f1.titre <> f2.titre;

```

film1	film2
Modern Times	City Lights
City Lights	Modern Times



Aggrégation

COUNT, AVG, SUM...

SQL DML 1
FC/CM

- aggrégation sur un attribut : opérations arithmétiques et statistiques
- COUNT(*) comptage occurrences
- COUNT(DISTINCT ...) comptage occurrences distinctes
- MAX MIN SUM AVG VARIANCE STDDEV fonctions statistiques

nom	mat	eval
Calvin	maths	4
Calvin	phys	2
Calvin	info	3
Hobbes	maths	1
Hobbes	phys	2
Moe	maths	0
Moe	phys	3
Moe	info	2
Susie	maths	6
Susie	phys	8
Susie	info	10

- COUNT(*) 11
- COUNT(DISTINCT nom) 4
- AVG(note) 3.7
- SUM(note) 41
- MIN(note) 0
- ...



Aggrégation d'un attribut

COUNT MIN

SQL DML 1
FC/CM

```

-- nombre de films total
SELECT COUNT(*) AS nfilms
FROM les_films;

-- année du premier film
SELECT MIN(année) AS année
FROM les_films;

```

nfilms
6

année
1931



Regroupement de lignes

GROUP BY



Aggrégation

COUNT MIN MAX

SQL DML 1
FC/CM
Intro
SELECT
Produit
Jointure
Aggrégation
TD

	Nom	Note	
Calvin	Calvin	4	AVG=3.00
	Calvin	2	
	Calvin	3	
Hobbes	Hobbes	1	AVG=1.50
	Hobbes	2	
Moe	Moe	0	AVG=1.66
	Moe	3	
	Moe	2	

■ regroupement d'attributs de valeurs identiques
clause `GROUP BY expr/colname/colnum...`

`GROUP BY nom`

nom	avg
Calvin	3.00
Hobbes	1.50
Moe	1.67
Susie	8.00

`GROUP BY mat`

mat	avg
info	5.00
maths	2.75
phys	3.75

25 / 31

SQL DML 1
FC/CM
Intro
SELECT
Produit
Jointure
Aggrégation
TD

```
-- nombre de films par auteurs
SELECT auteur, COUNT(*) AS nfilms
FROM les_films
GROUP BY auteur
ORDER BY nfilms DESC, auteur;
```

auteur	nfilms
Chaplin	3
Wells	2
Ozu	1

```
-- durée d'activité
SELECT auteur,
       MAX(année)-MIN(année) AS activité
FROM les_films
GROUP BY auteur
ORDER BY activité DESC;
```

auteur	activité
Wells	27
Chaplin	9
Ozu	0

26 / 31



Aggrégation (suite)

COUNT AVG



Aggrégation (encore)

COUNT DISTINCT

SQL DML 1
FC/CM
Intro
SELECT
Produit
Jointure
Aggrégation
TD

```
-- nombre de film et durée moyenne par auteur
SELECT auteur, COUNT(*) AS nfilms, AVG(durée) AS mdurée
FROM les_films
GROUP BY auteur
ORDER BY mdurée DESC;
```

auteur	nfilms	mdurée
Wells	2	01:58:30
Chaplin	3	01:40:20
Ozu	1	01:37:00

27 / 31

SQL DML 1
FC/CM
Intro
SELECT
Produit
Jointure
Aggrégation
TD

```
-- nb films par auteur et genre
SELECT auteur, quoi,
       COUNT(*) AS nfilms
FROM les_films
JOIN Genres USING (gid)
GROUP BY auteur, quoi
ORDER BY 1, 2;
```

auteur	quoi	nfilms
Chaplin	Comédie	2
Chaplin	Comédie dramatique	1
Ozu	Comédie	1
Wells	Drame	2

```
-- nb genres différents par auteur
SELECT auteur, COUNT(DISTINCT quoi) AS ngenres
FROM les_films AS f
JOIN Genres AS g ON f.gid=g.gid
GROUP BY auteur
ORDER BY ngenres DESC, auteur;
```

auteur	ngenres
Chaplin	2
Ozu	1
Wells	1

28 / 31



Aggrégation avec condition

HAVING COUNT



Attention

WHERE vs HAVING

SQL DML 1

FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

- clause **HAVING** *condition*
- condition porte une **aggrégation** projetée ou non

```
-- auteurs avec plusieurs films
SELECT auteur
FROM les_films
GROUP BY auteur
HAVING COUNT(*)>1;
```

auteur
Chaplin
Wells

29 / 31

SQL DML 1

FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

- clause **WHERE** condition sur table en **entrée**
- clause **HAVING** condition sur table agrégée en **sortie**

```
-- auteur et durée moyenne des films avant 1940
-- SEULEMENT pour les auteurs avec plusieurs films
SELECT auteur, AVG(durée) AS mdurée
FROM les_films
WHERE année < 1940
GROUP BY auteur
HAVING COUNT(*) > 1;
```

auteur	mdurée
Chaplin	01:27:00

30 / 31



Exercices

SQL DML 1

FC/CM

Intro

SELECT

Produit

Jointure

Aggrégation

TD

- quelle est la date du jour ?
- quelle heure est-il ?
- les cinéastes dont les noms commencent par C ou W
- les cinéastes ayant fait un film avant 1940
- les titres des films de *Hayao Miyazaki*
- les titres des films dramatiques (*Drame*)
- les titres des films d'après guerre, triés par année
- le nombre de films par an
- le nombre de films par décennie (histogramme)
- le nombre de films par genres
- la moyenne *géométrique* des durées de films
- les cinéastes ayant sorti des films la même année, sans doublons

31 / 31