



HMIN122M

Entrepôts de données et big-data

Rendu TD/TP 2

Auteur :

Gracia-Moulis Kévin
Canta Thomas

Master 1 - AIGLE / DECOL
Faculté des sciences de Montpellier
Année universitaire 2020/2021



Table des matières

1. Coût de plans d'exécution logiques	2
Question 1	2
Question 2	2
Question 3	2
2. Définition de plans d'exécution logiques	3
Requête 1	3
Requête 2	4
Requête 3	4
3. Réécriture de plans d'exécution logiques	5
Question 1	5
Question 2	5
4. Tous les plans d'exécution logiques	6
Question 1 & 2	6

1. Coût de plans d'exécution logiques

Question 1

Cette requête permet d'obtenir le nom des étudiants inscrits au module intitulé "*EDBD*".

Question 2

Chaque plan crée 4200 lignes intermédiaires.

Question 3

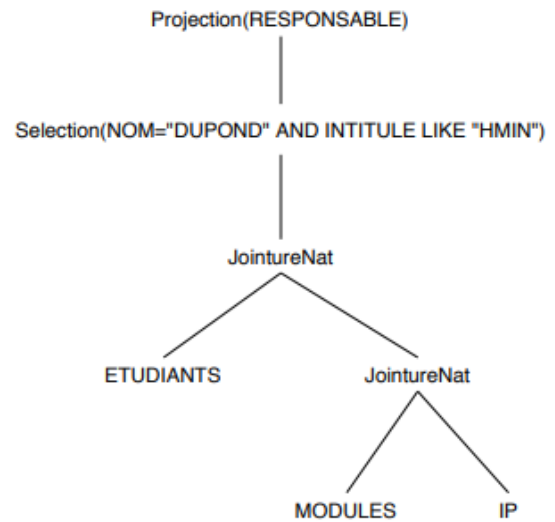
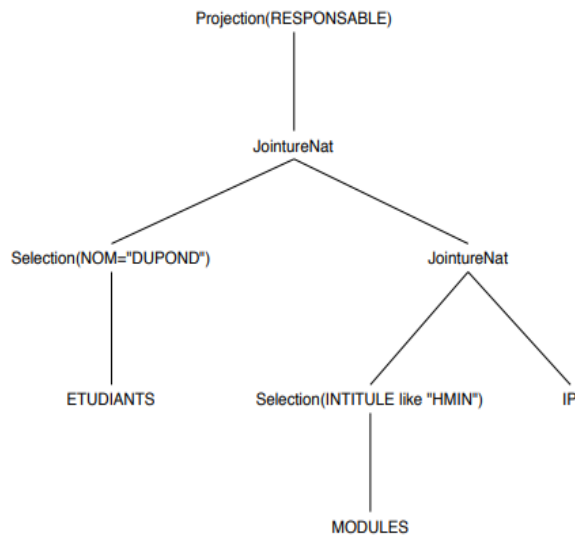
Le plan 3 est le plus optimal car il permet une sélection plus spécifique avant l'appel des jointures, réduisant ainsi le nombre maximum de colonnes.

2. Définition de plans d'exécution logiques

Requête 1

Cette requête permet d'obtenir le responsable des modules commençant par "HMIN" contenant au moins un élève ayant pour nom "Dupond"

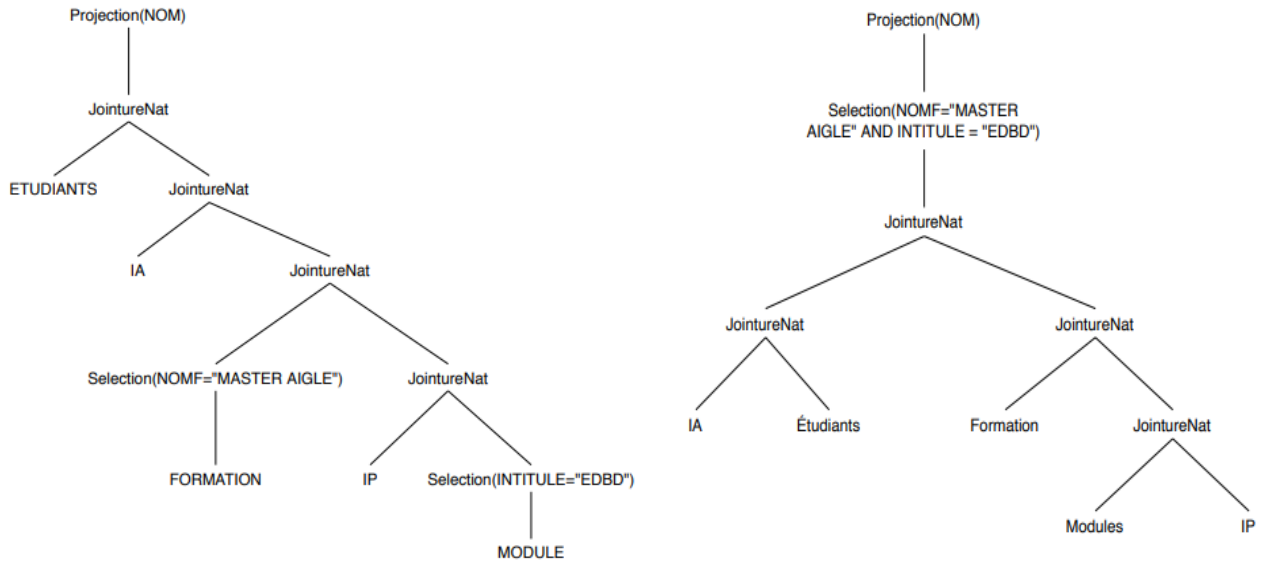
L'arbre de gauche, ci-dessous, est l'arbre le plus optimal.



Requête 2

Cette requête permet d'obtenir le nom des étudiants de la formation "MASTER AIGLE" et qui participent au module ayant pour intitulé "EDBD"

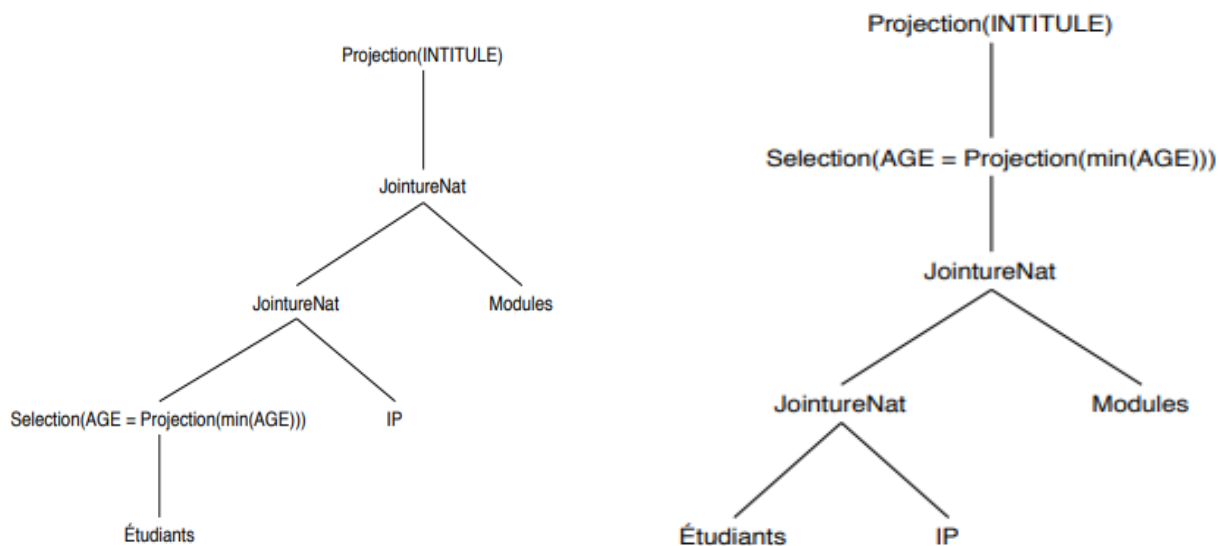
L'arbre de gauche, ci-dessous, est l'arbre le plus optimal.



Requête 3

Cette requête permet d'obtenir l'intitulé des modules possédant un étudiant ayant l'âge minimum

L'arbre de gauche, ci-dessous, est l'arbre le plus optimal.

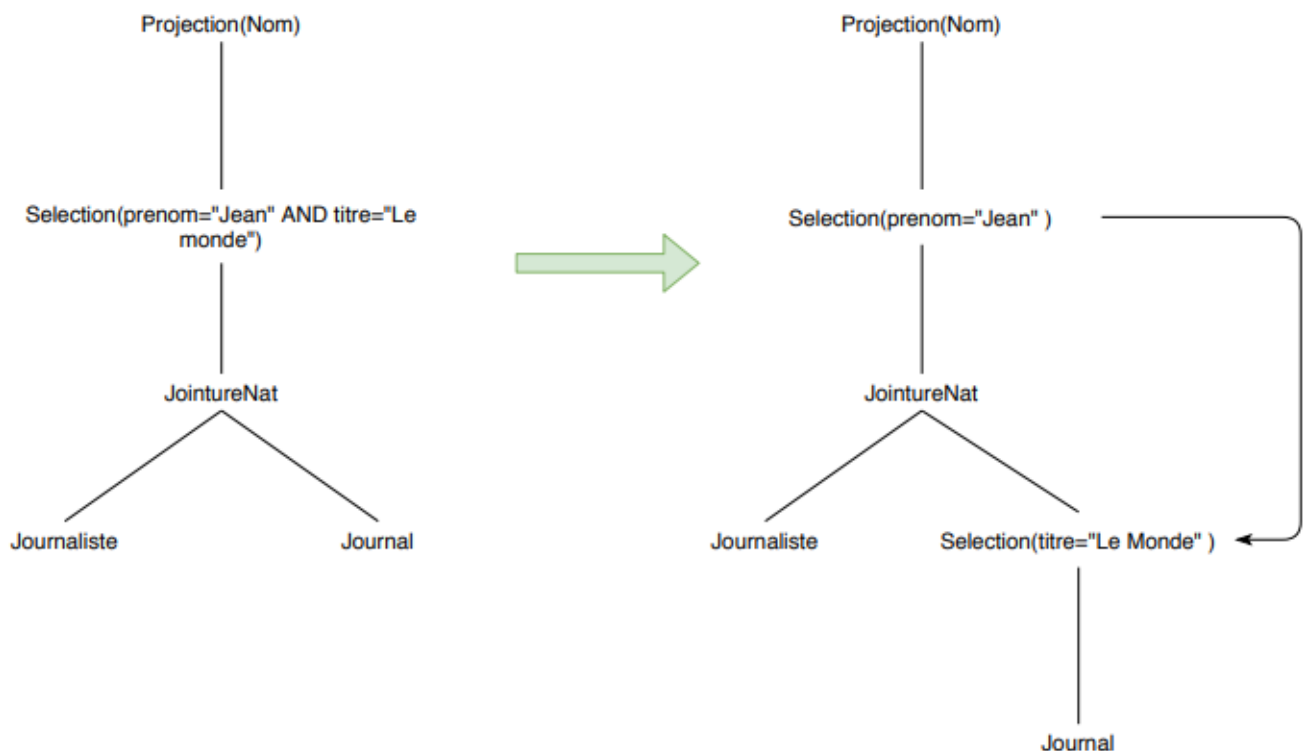


3. Réécriture de plans d'exécution logiques

Question 1

Elles retournent le même résultat car la sélection sur le titre ne s'applique qu'au journal. Si on "descend" la sélection dans l'arbre ceci ne change pas le résultat final.

Une illustration de nos propos ci-dessus



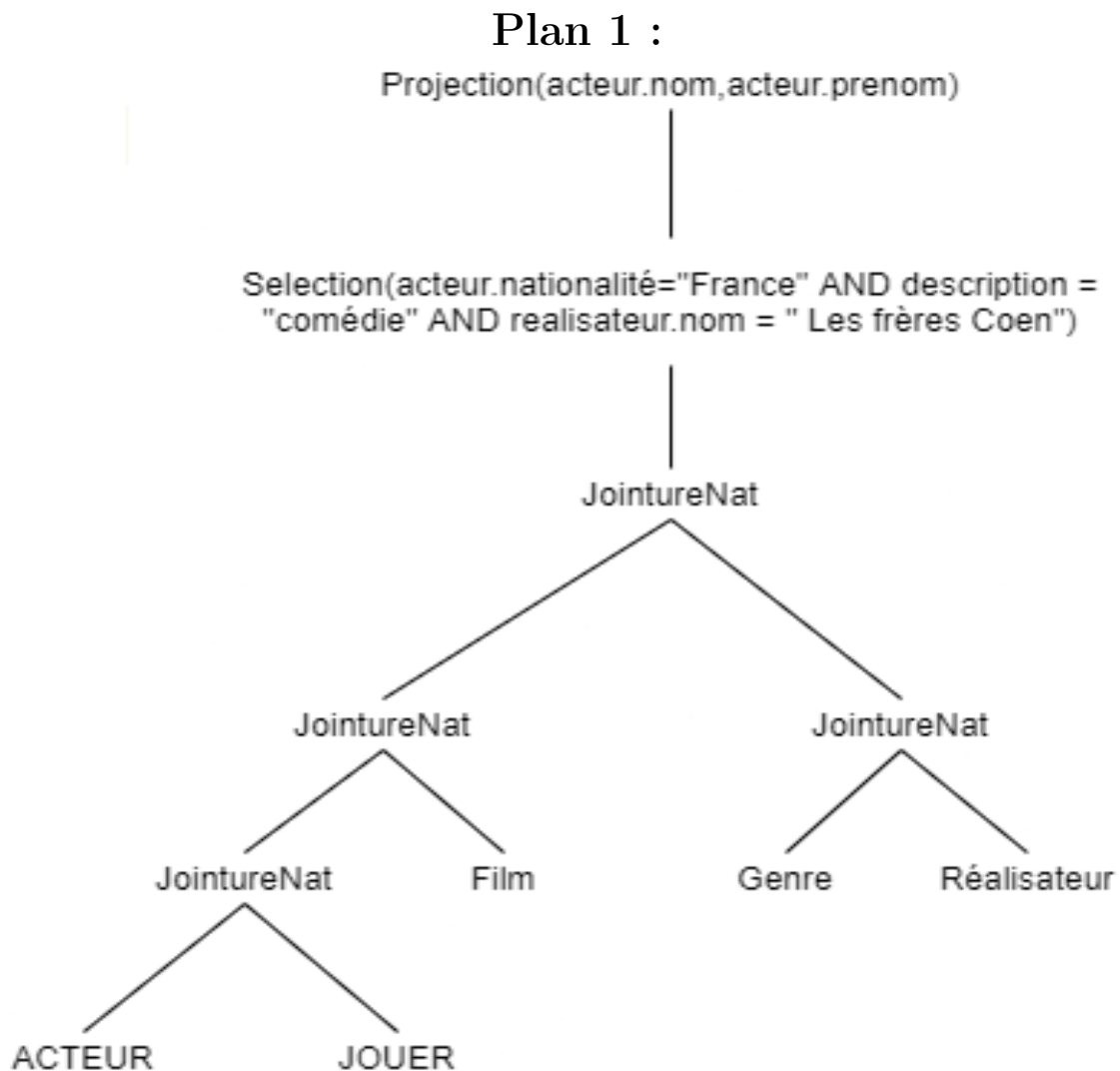
Question 2

Oui, celle de droite semble être plus efficace car la sélection sur les *titre* de *Journals* est faite le plus tôt possible, permettant ainsi de réduire le nombre de lignes. Une autre version plus optimisée serait donc de "descendre" aussi la sélection sur les *prénoms* de *Journaliste*

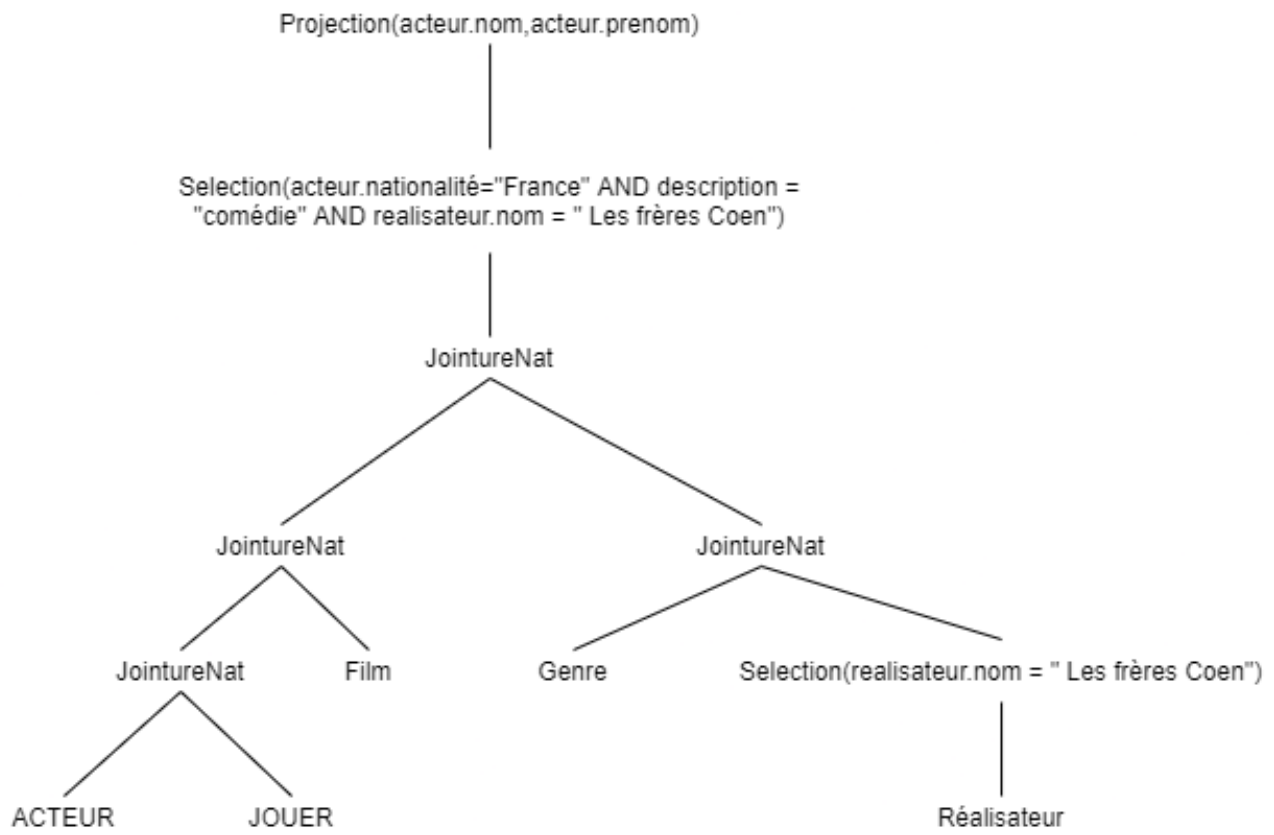
4. Tous les plans d'exécution logiques

Question 1 & 2

Il existe de nombreuses variantes pour représenter ce modèle, on en a donc représentés 3, l'un d'eux représente le plus optimale parmi tous ceux représentables. C'est, ici, le **plan 3** car il permet une sélection plus spécifique avant chaque appel de jointures, réduisant le nombre maximum de lignes.



Plan 2 :



Plan 3 :

