



# MEMOIRE

En vue de l'obtention du diplôme de **Master**  
Filière : **Informatique**  
Spécialité : **Systeme d'Information et Web**

## Thème

---

**Etude et comparaison des techniques de sélection des vues matérialisées dans les  
entrepôts de données**

---

### Présenté par :

- Mr. ABDELLAHI Akram
- Mr. LOKBANI Chouaib

Soutenu le : **05 octobre 2020**

Devant le jury composé de :

- M. MALKI Mimoun      Président
- M. KLOUCHE Badia      Examinatrice
- Dr. KECHAR Mohamed      Encadreur
- M. KHATIR Manel Khansa      Co-Encadreur

Année Universitaire : 2019 - 2020

## Résumé

Dans de nombreux systèmes à forte intensité de données, dans notre cas, les systèmes d'entreposage de données (Data Warehouse (DW)), la sélection des vues est très importante. La sélection des vues consiste à choisir un ensemble approprié de vues à matérialiser, qui optimise le coût total des requêtes, compte tenu d'une quantité limitée de ressources, par exemple, l'espace de stockage et le coût total de maintenance de la vue. La sélection de Vues Matérialisées (VM) est l'un des problèmes de recherche les plus difficiles liés à l'évolution de l'entrepôt de données, connu pour être un problème NP-complet.

Dans ce sujet, au moyen d'une revue de la littérature disponible, beaucoup d'auteurs ont donné différentes méthodologies et stratégies, tiré plusieurs conclusions pour résoudre ce problème (PSVM) de manière efficace ; et des perspectives d'avenir sont prévues pour combler les grandes lacunes constatées dans les méthodes existantes. Dans notre travail, diverses techniques sont identifiées, évaluées et comparées en matière d'espace de stockage en mémoire, de coût et de temps de traitement des requêtes pour déterminer si une approche particulière est supérieure aux autres.

### **Mots-clés :**

Entrepôt de données - Évolution de l'entrepôt de données - Matérialisation des vues - Sélection des vues - Problème de Sélection des Vues Matérialisées

## ملخص

في العديد من الأنظمة كثيفة البيانات ، في حالتنا ، أنظمة مستودعات البيانات يتكون اختيار طرق العرض من اختيار مجموعة مناسبة من طرق العرض المراد تحقيقها ، مما يؤدي إلى تحسين التكلفة الإجمالية للاستعلامات ، نظراً لمحدودية الموارد ؛ على سبيل المثال ، مساحة التخزين وإجمالي تكلفة صيانة العروض. يعد اختيار طرق العرض الفعلية أحد أصعب مشكلات البحث المتعلقة بتطور مستودع البيانات ، والمعروف باسم مشكلة نـمـت. في هذا الصدد ، من خلال مراجعة الأدبيات المتاحة ، قدم العديد من المؤلفين منهجيات واستراتيجيات مختلفة ، واستخلصوا عدة استنتاجات لحل هذه المشكلة بشكل فعال ؛ وآفاق المستقبل مخططة لسد الفجوات المحددة في الأساليب الحالية. في عملنا ، يتم تحديد التقنيات المختلفة من حيث مساحة تخزين الذاكرة والتكلفة ووقت معالجة الطلب وتقييمها ومقارنتها لتحديد ما إذا كان نهج معين متفوقاً على الأساليب الأخرى.

### الكلمات المفتاحية :

مستودع البيانات - تطور مستودع البيانات - تجسيد طرق العرض - اختيار طرق العرض

## **Abstract**

In many systems with a high data intensity, in our case , Data Warehouse (DW), view selection is very important. The selection of views consists of choosing an appropriate set of views to be materialized, which optimizes the total cost of queries, given a limited amount of resources; for example, storage space and total view maintenance cost. The selection of Materialized Views (MV) is one of the most difficult search problems related to the evolution of the data warehouse, known as a complete NP problem.

In this regard, by means of an available literature review, many authors have given different methodologies and strategies, drawn several conclusions to solve this problem (PSVM) effectively ; and prospects for the future are planned to bridge the gaps identified in existing methods. In our work , various techniques in terms of memory storage space, cost and request processing time are identified, evaluated and compared to determine whether a particular approach is superior to the others.

### **Keywords :**

Data warehouse-Data warehouse evolution-View materialization-View selection -Materialized views selection problem