UNIVERSITATEA DIN BUCUREȘTI

FACULTATEA DE MATEMATICĂ ȘI INFORMATICĂ

SPECIALIZAREA INFORMATICĂ

   
 

**DATA WAREHOUSE & BUSINESS INTELLIGENCE**    
**- ANALIZA -**

Autor

Tiberius Coman

George Banica

Daniela Alexandra Constantin

Sebastian Alexandru Velciu

Profesor titular

Lect. Dr. GABRIELA MIHAI

 

   
GRUPA 405

ANUL I MASTER, SEMESTRUL I

7.

Pentru a putea stabili care vor fi obiectele dimensiune pe care le vom creea, vom incepe prin a mention ace sunt acestea si la ce ne pot ajuta.

Dimensiunea este o structură care clasifică datele pentru a le permite utilizatorilor să răspundă la întrebările de afaceri. Dimensiunile utilizate în mod obișnuit sunt clienții, produsele și timpul. De exemplu, fiecare canal de vânzări al unui comerciant poate colecta și stoca date privind vânzările. Managementul lanțului de vânzare poate construi un depozit de date pentru a analiza vânzările produselor sale în toate magazinele de-a lungul timpului și pentru a răspunde la întrebări precum:

* Care este efectul promovării unui produs comparative cu unul care nu este promovat?
* Care sunt vânzările unui produs înainte și după o promoție?
* Cum afectează o promovare diferitele canale de distribuție?

În baza de date Oracle, informațiile dimensionale în sine sunt stocate într-un tabel de dimensiuni. În plus, obiectul dimensiune ajută la organizarea și gruparea informațiilor dimensionale în ierarhii. Aceasta reprezintă relații naturale 1:n între coloane sau grupuri de coloane (nivelurile unei ierarhii) care nu pot fi reprezentate cu condiții de constrângere. Urcarea unui nivel în ierarhie se numește acumularea datelor și coborârea la un nivel în ierarhie se numește detaliere a datelor.

Obiectele dimensiune nu trebuie definite neaparat. Cu toate acestea, creearea lor poate aduce beneficii semnificative, deoarece ajută la rescrierea interogărilor să efectueze tipuri mai complexe de rescrire.

In cazul acestui proiect au fost identificate 2 obiecte dimensiune: locatie, timp. Ambele dimensiuni sunt normalizate(nu este stocata in mai mult de un table, sau altfel zis, nu are referinte la alte tabele).

Aceste obiecte vor avea urmatoarele nivele si ierarhie:

* Locatie: locatie\_id -> strada -> oras -> tara
* Timp: data\_id -> data -> anul

Pentru dimensiunea locatie, clauza ATTRIBUTE ... DETERMINES se referă la a pune in legatura locatie\_id cu celelalte campuri, si anume: strada, oras, tara si site. Aceasta este o determinare unidirecțională. Este doar garantat că pentru oun anumit id, o sa fie găsita exact o valoare potrivită pentru strada, oras, tara si site. Insa nu se poate determina un id pentru un anumit oras spre exemplu.

Similar este si in cazul dimensiunii timp, un singur lucru difera aici, si anume faptul ca putem creea mai multe relatii prin atribute de aceasta data. Motivul pentru care putem face acest lucru este faptul ca atat data\_id cat si data sunt unice in tabelul dim\_calendar. Prin urmare, campurile: id\_data, data, ziua, luna, anul pot fi determinate in mod direct atat prin data\_id cat si prin data.

Nu in ultimul rand, informațiile unui obiect dimensiune sunt doar declarative și nu sunt impuse de baza de date. Dacă relațiile descrise de dimensiuni sunt incorecte, pot apărea rezultate incorecte. Prin urmare, ar trebui să fie verificațe periodic relațiile specificate de CREATE DIMENSION utilizând procedura DBMS\_DIMENSION.VALIDATE\_DIMENSION.