

# ***PROIECT DATAWAREHOUSE***

~ Modul analiza ~

## **1. Descrierea modelului ales și a obiectivelor aplicației.**

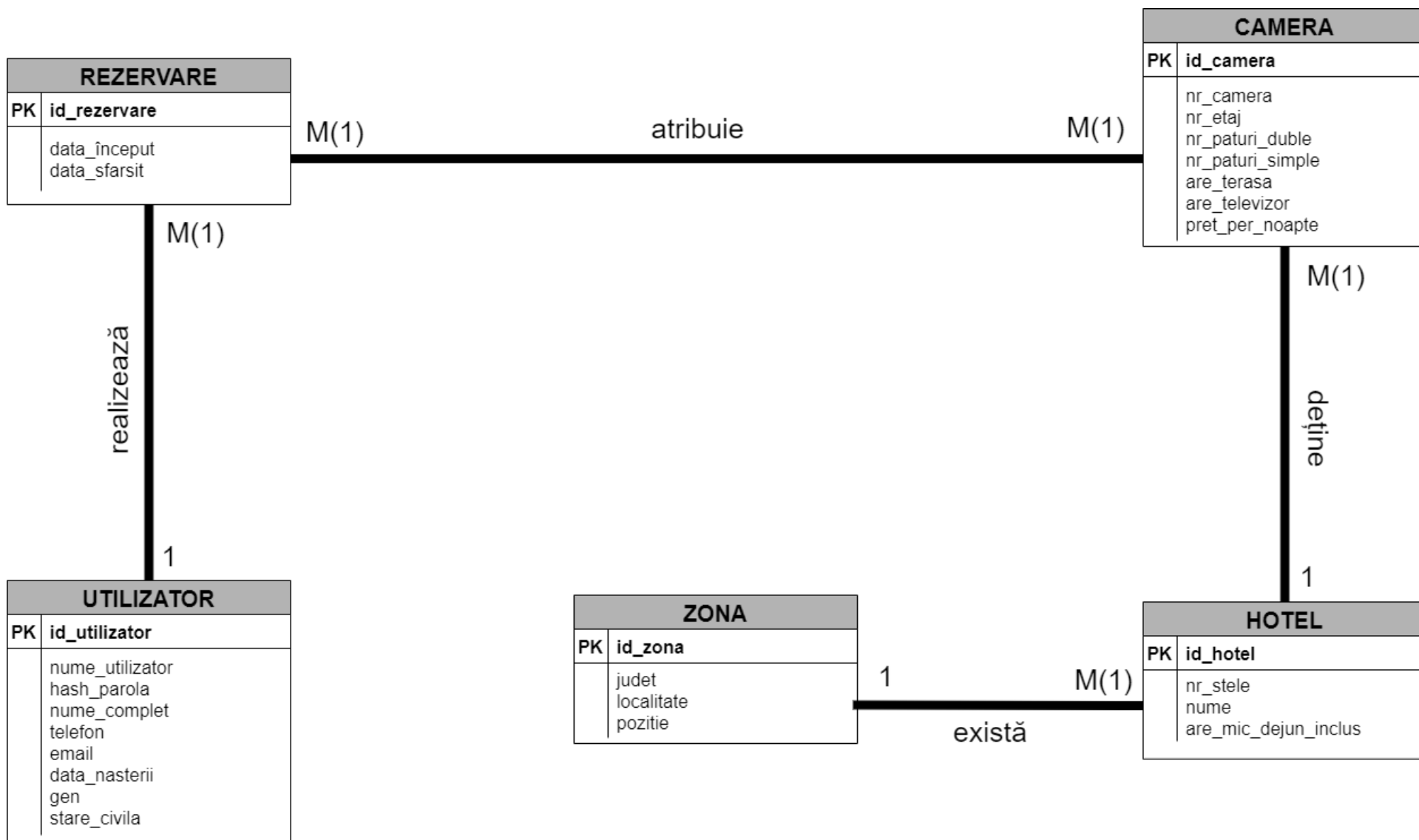
Pentru a înțelege funcționalitatea aplicației noastre, vom începe cu ***descrierea modelului***. Aplicația oferă posibilitatea utilizatorilor de a intra în aplicație prin intermediul completării unor detalii cu caracter personal cum ar fi numărul de telefon, adresa de email, data\_nasterii, genul, starea\_civilă, etc.

Clientul poate realiza una sau mai multe rezervări. Aceste rezervări atribuie una sau mai multe camere în funcție de perioada disponibilă. Un hotel partener al aplicației noastre deține un număr de stele și poate avea mai multe camere. O camera poate avea o serie de caracteristici importante cum ar fi etajul, numărul de paturi duble sau simple, terasă, televizor. La finalul sejurului, clientul trebuie să plătească o sumă de bani în funcție de pretul camerei pe noapte și de numărul de zile de cazare. Fiecare hotel se diferențiază în funcție de zona în care se află.

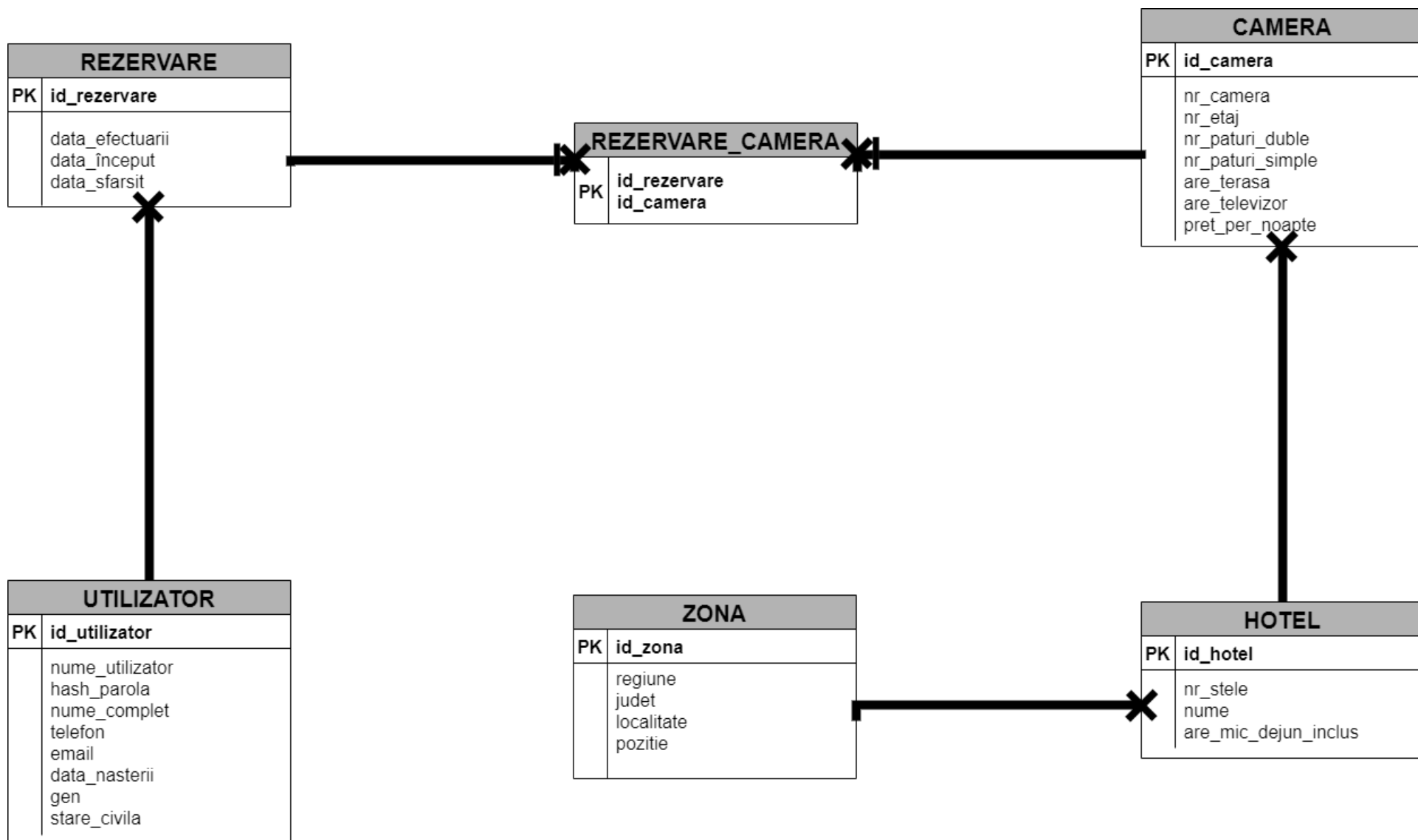
Obiectivele aplicației sunt de a oferi o interfață cât mai intuitivă clientului pentru a putea să își aleagă cu ușurință locul unde dorește să își petreacă vacanța și de a oferi managerilor o serie de rapoarte bogate în informații utile pentru a putea să își înțeleagă mai bine afacerea. Printre aceste rapoarte putem întâlni informații cu privire la perioadele de aglomerare a clienților, vârsta clienților majoritari în funcție de perioadă și informații cu privire la zonele de amplasare a viitoarelor hoteluri, etc.

## 2. Diagramele bazei de date OLTP

### a. Diagrama entitate – relație a bazei de date OLTP.



**b. Diagrama conceptuală a bazei de date OLTP.**



Schemele relaționale corespunzătoare acestei diagrame conceptuale sunt următoarele:

**UTILIZATOR** ( #id\_utilizator, nume\_utilizator, hash\_parola, nume\_complet, telefon, email, data\_nasterii, gen, stare\_civilă)

**REZERVARE** ( #id\_rezervare, id\_utilizator, data\_început, data\_sfârșit, data\_efectuării)

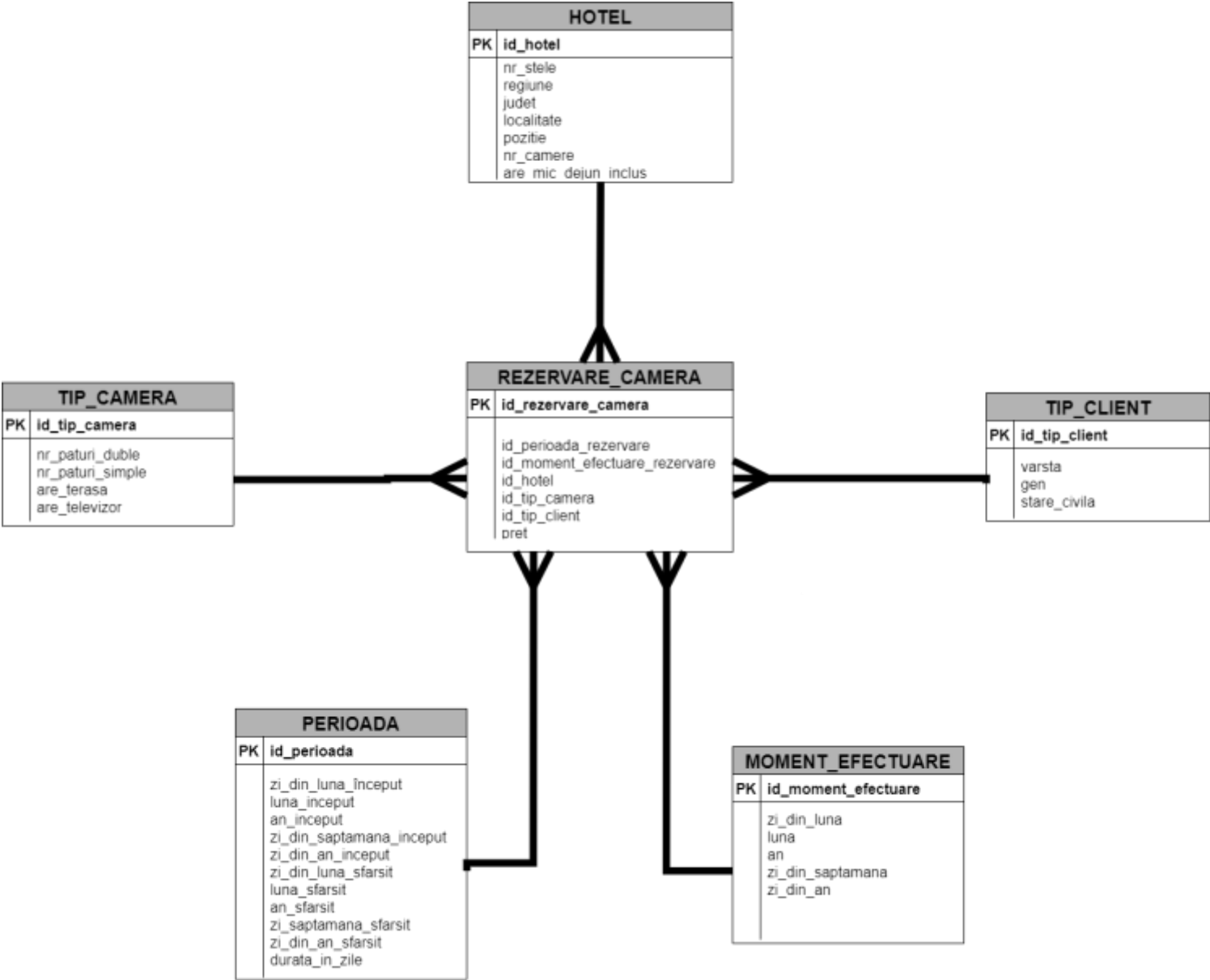
**REZERVARE CAMERA** ( id\_rezervare, id\_camera).

**CAMERA** ( #id\_camera, id\_hotel, nr\_camera, nr\_etaj, nr\_paturi\_duble, nr\_paturi\_simple, are\_terasa, are\_televizor, pret\_per\_noapte)

**HOTEL** ( #id\_hotel, id\_zona, nume, nr\_stele, are\_mic\_dejun)

**ZONA** ( #id\_zona, regiune, judet, localitate, pozitie).

3. Diagrama stea/fulg a bazei de date depozit (un tabel de fapte și cel puțin 5 tabele dimensiune).



**4. (1p)** Descrierea câmpurilor necesare pentru fiecare tabel din baza de date depozit și modul de populare al acestora cu informații din baza de date OLTP

**HOTEL** ( #id\_hotel, nr\_stele, judet, localitate, pozitie, nr\_camere, are\_mic\_dejun\_inclus)

id\_hotel - cheie primara autogenerata

nr\_stele - se va prelua din tabela hotel a bazei de date OLTP

judet - se va utiliza join intre tabela hotel din OLTP si tabela locatie din OLTP

localitate- se va utiliza join intre tabela hotel din OLTP si tabela locatie din OLTP

pozitie - se va utiliza join intre tabela hotel din OLTP si tabela locatie din OLTP

nr\_camere - se va prelua din tabela hotel a bazei de date OLTP

are\_mic\_dejun\_inclus - se va prelua din tabela hotel a bazei de date OLTP

**TIP CAMERA** ( #id\_tip\_camera, nr\_paturi\_duble, nr\_paturi\_simple, are\_terasa,are\_televizor)

id\_tip\_camera - cheie primara autogenerata

nr\_paturi\_duble - se va prelua din tabela camera a bazei de date OLTP

nr\_paturi\_simple - se va prelua din tabela camera a bazei de date OLTP

are\_terasa- se va prelua din tabela camera a bazei de date OLTP

are\_televizor- se va prelua din tabela camera a bazei de date OLTP

**PERIOADA** ( #id\_perioada, zi\_luna\_inceput, luna\_inceput, an\_inceput, zi\_din\_saptamana\_inceput, zi\_din\_an\_inceput, zi\_luna\_sfarsit, luna\_sfarsit, an\_sfarsit, zi\_din\_saptamana\_sfarsit, zi\_din\_an\_sfarsit, durata\_in\_zile)

id\_perioada - cheie primara autogenerata

luna\_inceput - se va prelua din tabela rezervare din OLTP

an\_inceput - se va prelua din tabela rezervare din OLTP

zi\_din\_saptamana\_inceput - se va prelua din tabela rezervare din OLTP

zi\_din\_an\_inceput - se va prelua din tabela rezervare din OLTP

zi\_luna\_sfarsit - se va prelua din tabela rezervare din OLTP

luna\_sfarsit - se va prelua din tabela rezervare din OLTP

an\_sfarsit - se va prelua din tabela rezervare din OLTP

zi\_din\_saptamana\_sfarsit - se va prelua din tabela rezervare din OLTP

zi\_din\_an\_sfarsit - se va prelua din tabela rezervare din OLTP

durata\_in\_zile - se va prelua din tabela rezervare din OLTP

**MOMENT EFECTUARE** ( #id\_moment\_efectuaire, zi\_luna, luna, an, zi\_din\_saptamana, zi\_din\_an)

id\_moment\_efectuaire - cheie primara autogenerata

zi\_luna-

luna-

an -

zi\_din\_saptamana-

**TIP CLIENT** ( #id\_tip\_client, varsta, gen, stare\_civila)

id\_tip\_client -

varsta - diferenta intre sysdate si data nasterii preluata din tabela client OLTP

gen - preluat din tabela client OLTP

stare\_civila - preluat din tabela client OLTP

Pentru popularea bazei de date depozit cu date din baza de date OLTP se va utiliza o procedură. Au fost create funcții ajutătoare pentru obținerea mai multor informații, precum calculul vârstei clientului.



**5. (1p)** Identificarea constrângerilor **specifice depozitelor de date** ce trebuie definite, justificând alegerea făcută

- Pentru toate id-urile prezente în tabele se va aplica constrângerea de cheie primară (valoare unică și nenulă pentru fiecare tuplu).
- Pentru toate datele introduse în tabela “perioada” se vor aplica restricții de tipul NOT NULL, întrucât o rezervare fără date de început și final nu poate fi considerată validă.
- Pentru toate datele introduse în tabela “hotel” se vor aplica constrângeri de tipul NOT NULL. Vor fi necesare toate datele solicitate pentru a se considera un hotel valid. Aceste date sunt utile atât pentru clienți cât și pentru analiștii de date.
- Pentru ca un client să fie considerat valid, este necesar ca toate câmpurile tabelii asociate să aibă restricția NOT NULL. Aceasta este importantă și pentru a evita rezervările false.
- Pentru câmpurile unde este necesară introducerea datelor de tip zi din lună, se poate impune restricția ca valoarea acestora să fie cuprinsă între 1 și 31.
- Pentru câmpurile unde este necesară introducerea datelor de tip zi din an, se poate impune restricția ca valoarea acestora să fie cuprinsă între 1 și 365.

**6. (0,5p)** Identificarea indecșilor **specifci depozitelor de date** ce trebuie definiți asupra modelului (**minim 2** dacă echipa este formată din 4 persoane); formularea unei cereri în limbaj natural care va determina utilizarea indecșilor specificați și va fi implementată în următoarea etapă

- **Indexare după ID-ul rezervării**

Întrucât rezervarea este pionul central al bazei de date, este util să definim un astfel de index pentru a accesa mai ușor detaliile unei rezervări. De asemenea, vrem să vedem câte camere au fost rezervate simultan de către același client. Poate fi de ajutor pentru raportare. Câmpul `id_rezervare` este preluat din tabelul “rezervare camera” din baza de date OLTP și astfel putem identifica ce camere au fost rezervate în cadrul acestui ID.

- **Indexare după luna efectuării rezervării**

Acest index poate fi utilizat pentru un raport în care se dorește observarea lunii din an în care se efectuează rezervările. Spre exemplu, o agenție de turism ar avea nevoie de astfel de date pentru a ști când se pot lansa oferte noi.

- **Indexare după localitatea în care se află hotelul și poziția acestuia față de centrul localității**

Acest index va fi construit pentru a putea sorta și grupa rezervările după locație. Locația are rol semnificativ în rapoartele și statisticile destinațiilor de vacanță preferate de către clienți.

- **Indexare după tip camera**

Acest index este construit pentru a sorta rezervările în funcție de tipul de cameră ales. Poate fi utilizat pentru a vedea mai ușor ce tip de cameră preferă clienții.

**7. (0,5p)** Identificarea obiectelor de tip dimensiune ce trebuie definite asupra modelului (**minim 2** dacă echipa este formată din 4 persoane)

- **Obiect de tip dimensiune pentru tabelul perioada rezervare**

Această dimensiune evidențiază dependențele dintr zi zi (\_inceput/\_sfarsit) - zi\_luna (\_inceput/\_sfarsit) - zi\_an (\_inceput/\_sfarsit). Mai exact, ziua în care începe (/se sfârșește) rezervarea poate determina valorile atributelor zi din lună și zi din an.

- **Obiect de tip dimensiune pentru locație**

Am putea considera ierarhia localitate județ. Totuși, această dimensiune este problematică deoarece același nume de localitate se poate regăsi în mai multe județe. Este un caz bun pentru obiecte de tip dimensiune dar utilizat în raportări poate duce la erori grave.

- **Obiect de tip dimensiune pentru id rezervare**

Atributul id\_rezervare determină attributele id\_perioada\_rezervare, id\_moment\_efectuare\_rezervare, id\_hotel, id\_tip\_client.

**8. (1p)** Identificarea tabelelor care vor fi partiționate și a tipului de partiționare (minim 2 dacă echipa este formată din 4 persoane; formularea unei cereri în limbaj natural care va determina utilizarea lor și va fi implementată în următoarea etapă.

- **Partitionare prin range pret pentru tabelul rezervare\_camera**

Se pot împărți rezervările în funcție de un interval de preț. Aceste intervale pot fi utilizate în raportările pentru bugetele clienților sau raportările pentru stabilirea ofertelor în piață.

- **Partitionare prin listă după numărul de stele pentru tabelul hotel**

Acest tip de partiționare se poate utiliza pentru situațiile în care se dorește obținerea unor informații în funcție de tipul hotelului indicat de numărul de stele. Spre exemplu, “Care este pretul mediu al unei camere cu un anumit număr de locuri la un hotel cu x nr de stele?”.

**9. (0,5p)** Formularea în limbaj natural a unei cereri SQL complexe care va fi optimizată în următoarea etapă, folosind tehnici specifice bazelor de date depozit. Precizarea tehnicilor de optimizare ce ar putea fi utilizate pentru această cerere particulară (avantaje / dezavantaje de utilizare pentru o anumită tehnică)

Construiți o cerere care să afișeze numele hotelurilor ce au valoarea rezervărilor cuprinse între două valori.

**10. (2p)** Formularea în limbaj natural a cel puțin 5 cereri cu grad de complexitate diferit, concretizate în rapoarte (grafice) ce vor fi create în următoarele etape

- Realizați un raport care arată numărul de camere rezervate în fiecare lună.
- Raport grafic care arată câte rezervări au fost făcute în anumite perioade ale anului.
- Raport grafic ce arată cele mai scumpe 5 hoteluri din România.