**Tema III**

O dependență funcțională(FD) este o relație între două atribute, de obicei între PK și alte atribute non-cheie dintr-un tabel. Pentru orice relație R, atributul Y depinde funcțional de atributul X(de obicei PK), dacă pentru fiecare instanță valabilă a lui X, acea valoare a lui X determină în mod unic valoarea lui Y.

* O relație R este în **FN** dacă toate atributele sunt relevante pentru entitate și acestea nu sunt atomice; nu pot primi valori multiple
* O relație R este în **FN2** dacă R se află în FN1 și fiecare atribut care nu este cheie(nu participă la cheia primară) este dependent de întreaga cheie primară. (Atunci când un atribut nu depinde de cheia primară poate fi scos din tabel și pus în alt tabel)
* O relație R este în **FN3** dacă R se află în FN2 și fiecare atribut care nu este cheie(nu participă la o cheie) depinde direct de cheia primară. În forma normală FN3 NU avem dependențe tranzitive.

*Diagrama ER trimisă în prima temă(Jobs Application Interview for Internships Programs) respectă toate cele 3 forme de normalizare.*

Vom considera următoarea schemă (*Orders*) care nu se încadrează în nicio formă de normalizare și care va fi descompusă până vom ajunge în FN3. În cele ce urmează voi prezenta toate cele 3 etape.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| id\_cust | first\_name | last\_name | address | phone | email | ord\_id | ord\_date | prod\_id | q | prod\_type | prod\_price |
| 5 | Noah | Davis | 68th Ave, Pennsylvania, Lancaster, USA, 175884 | +1484473221 | [noah\_d@gmail.com](mailto:noah_d@gmail.com) | 154 | 10/10/2020 16:54:20 | 65 | 1 | clothing | 20,0 |
| 5 | Davis | Davis | 68th Ave, Pennsylvania, Lancaster, USA, 175884 | +1484473221 | [noah\_d@gmail.com](mailto:noah_d@gmail.com) | 154 | 10/10/2020 16:54:20 | 2667 | 1 | clothing | 120,50 |
| 6 | Emma | Smith | 67th Ave, Pennsylvania, Lancaster, USA, 175883 | +1484573361 | [emma\_smith@gmail.com](mailto:emma_smith@gmail.com)  [emmy.smith@gmail.com](mailto:emmy.smith@gmail.com) | 235 | 18/10/2020 12:54:20 | 548 | 1 | clothing | 1000 |
| 2371 | Jayden | Zwald | 7th Oaks, Pennsylvania, Lancaster, USA, 156883 | +1484853361+1544588361 | [jaydenz@gmail.com](mailto:jaydenz@gmail.com) | 785 | 10/12/2020 16:32:20 | 457 | 2 | footwear | 249,99 |
| 2371 | Jayden | Zwald | 7th Oaks, Pennsylvania, Lancaster, USA, 156883 | +1484853361+1544588361 | [jaydenz@gmail.com](mailto:jaydenz@gmail.com) | 785 | 10/12/2020 16:32:20 | 15432 | 1 | clothing | 54,99 |
| 2371 | Jayden | Zwald | 7th Oaks, Pennsylvania, Lancaster, USA, 156883 | +1484853361 +1544588361 | [jaydenz@gmail.com](mailto:jaydenz@gmail.com) | 8745 | 19/12/2020 16:15:20 | 45884 | 2 | clothing | 10,50 |

Această tabelă nu se află în niciuna dintre formele normale. Nu există nicio coloană care să identifice unic liniile. Cele două coloane care identifică în mod unic fiecare rând sunt **order\_id** și **item\_id**: nu există două rânduri care au aceeași combinație de ordin\_id și item\_id. Prin urmare, împreună se califică pentru a fi utilizate ca cheie principală a tabelului.

Avem atributele adresss, phone\_number, email sunt atribute cu valori multiple pentru a aduce această tabelă în FN1 putem construi următoarele tabele:

* + Unii clienți au mai multe adrese de email. Pentru atributul email:

Creem o tabelă nouă în care vom avea 3 coloane: email\_id, cust\_id și email (cust\_id este codul clientului). Cheia tabelului este compusă din cust\_id(cheia tabelului inițial) si email\_id care va reține valori atomice ale atributului cu valori multiple. Cheia email\_id poate fi omisă. Email\_id este o cheie primară artificială ce poate fi adăugată opțional.

|  |  |  |
| --- | --- | --- |
| email\_id | cust\_id | email |
| 1 | 6 | [emma\_smith@gmail.com](mailto:emma_smith@gmail.com) |
| 2 | 6 | [emmy.smith@gmail.com](mailto:emmy.smith@gmail.com) |

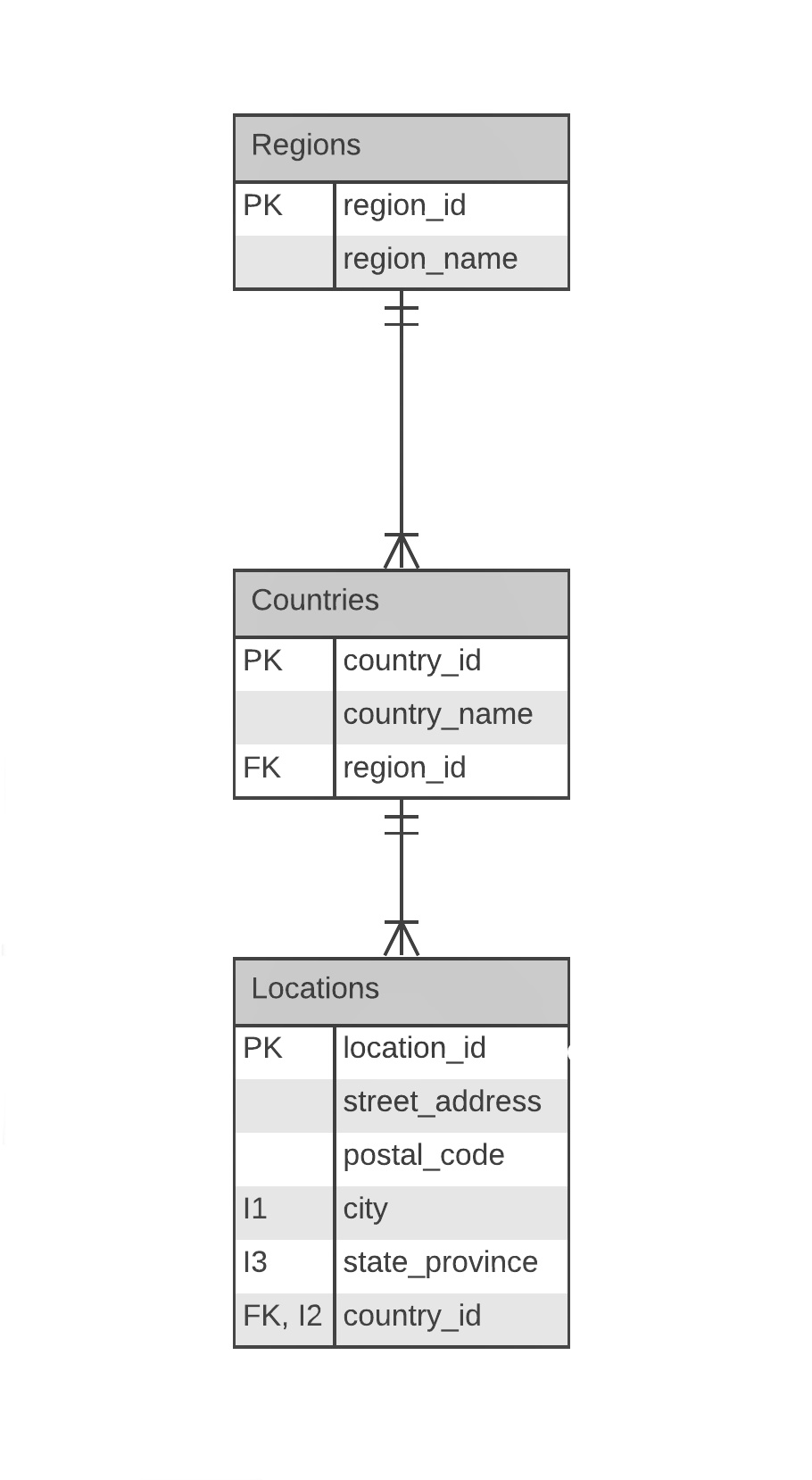
* + Unii clienți au mai multe numere de telefon. Analog atributului email, pentru atributul phone\_number

Creem o tabelă în care vom avea 3 coloane: cust\_id și phone\_number.

Cheia tabelului este compusă din cust\_id(cheia tabelului inițial) și phone\_number\_id care va reține valorile atomice ale atributului cu valori multiple.

|  |  |
| --- | --- |
| cust\_id | phone\_number |
| 2371 | +1484853361 |
| 2371 | +1544588361 |

* + Address nu este un atribut atomic ea poate fi descompusă în Region, Country, County, City, Street etc. Address poate fi descompusă în următoarele tabele:



|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| cust\_id | first\_name | last\_name | location\_id | ord\_id | ord\_date | prod\_id | quantity | prod\_type | prod\_price |
| 5 | Noah | Davis | 1 | 154 | 10/10/2020 16:54:20 | 65 | 1 | clothing | 20,0 |
| 5 | Davis | Davis | 1 | 154 | 10/10/2020 16:54:20 | 2667 | 1 | clothing | 120,50 |
| 6 | Emma | Smith | 2 | 235 | 18/10/2020 12:54:20 | 548 | 1 | clothing | 1000 |
| 2371 | Jayden | Zwald | 3 | 785 | 10/12/2020 16:32:20 | 457 | 2 | footwear | 249,99 |
| 2371 | Jayden | Zwald | 3 | 785 | 10/12/2020 16:32:20 | 15432 | 1 | clothing | 54,99 |
| 2371 | Jayden | Zwald | 3 | 8745 | 19/12/2020 16:15:20 | 45884 | 2 | clothing | 10,50 |

În această formă, tabela noastră se află în **FN1**.

În continuare, testăm fiecare tabelă pentru dependențe parțiale pe o cheie concatenată. Aceasta înseamnă că pentru un tabel care are o cheie primară concatenată, fiecare coloană din tabel care nu face parte din cheia primară trebuie să depindă de întreaga cheie concatenată pentru existența sa. Dacă orice coloană depinde doar de o parte a cheii concatenate, atunci spunem că întreaga tabelă nu este în Forma Normală II. În acest caz, trebuie să eliminăm dependențele parțiale .

În continuare vom lua fiecare atribut ce nu face parte din cheia primară și îl vom analiza:

* + **order\_data** este data la care a fost făcută comanda. Evident, se bazează pe order\_id, dar o anumită comandă într-o anumită dată nu se bazează pe un anumit articol. Deci tabela noastră nu se află în FN2.
  + **cust\_id** este numărul de identificare al clientului care a plasat comanda. Nu se bazează pe **order\_id** deoarece un client poate exista fără să plaseze comenzi și am vrea să reținem informațiile despre clienți chiar dacă nu au momentan comenzi plasate. Evident u se bazează pe **item\_id**. Așadar **cust\_id** nu este identificat de niciun atribut al cheii primare, analog si atributele care descriu clientul.
  + **quantity(q)** se referă la numărul de articole cumpărate de un anumit produs identificat cu **prod\_id** dintr-o anumită comandă. Nu poate exista fără un **prod\_id,** întrucât nu putem vorbi despre „cantitatea de nimic” (cel puțin nu în proiectarea bazei de date). Cantitatea nu poate exista fără un **order\_id,** deoarece o cantitate a unui anumit produs care este achiziționată fără o comandă este lipsită de sens. Deci, această coloană nu încalcă a doua formă normală: **item\_qty** depinde de ambele submulțimi ale cheii noastre primare.
  + **prod\_type** reprezintă o descriere a unui obiect. Depinde în întregime de **prod\_id.** Nu se bazează pe **order\_id** deoarece un produs poate exista și fără plasarea unei comenzi.
  + **prod\_price** reprezintă prețul unui produs. Prețul unui produs depinde atât de produs, cât și de comandă. Deci prod\_price depinde de cheia primară.

Pentru a o aduce în FN2 trebuie să descompunem relația în alte două relații. În primul tabel păstrăm în întregime cheia compusă și atributele care depind în întregime de aceasta. A doua tabelă va conține doar mulțimea de atribute care depind de o submulțime proprie a cheii împreună cu cheia care determină aceste atribute.

**ORDERS**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| cust\_id | first\_name | last\_name | address\_id | ord\_id | ord\_date |
| 5 | Noah | Davis | 1 | 154 | 10/10/2020 16:54:20 |
| 6 | Emma | Smith | 25 | 235 | 18/10/2020 12:54:20 |
| 2371 | Jayden | Zwald | 32 | 785 | 10/12/2020 16:32:20 |

**ORDER\_PRODUCTS**

|  |  |  |  |
| --- | --- | --- | --- |
| ord\_id | prod\_id | quantity | prod\_price |
| 154 | 65 | 1 | 20,0 |
| 154 | 2667 | 1 | 120,50 |
| 235 | 548 | 1 | 1000 |
| 785 | 457 | 2 | 249,99 |
| 785 | 15432 | 1 | 54,99 |
| 8745 | 45884 | 2 | 10,50 |

**PRODUCT**

|  |  |
| --- | --- |
| prod\_id | prod\_type |
| 65 | clothing |
| 2667 | clothing |
| 548 | clothing |
| 457 | footwear |
| 15432 | clothing |
| 45884 | clothing |

În această formă relația este în **FN2**.

Pentru a aduce relația în forma normală III trebuie să eliminăm dependențele tranzitive.

* **id\_cust** nu depinde de **ord\_id**

Vom crea o noua tabelă **CUSTOMER** ce va avea cheia primară **id\_cust** și va reține datele despre clienți(indiferent dacă aceștia au plasat o comandă sau nu).

**CUSTOMER**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| cust\_id | first\_name | last\_name | location\_id | ord\_id |
| 5 | Noah | Davis | 1 | 154 |
| 6 | Emma | Smith | 2 | 235 |
| 2371 | Jayden | Zwald | 3 | 785 |
| 2372 | Alice | Eyre | 24 | null |

Cheia externă(FK) va fi ord\_id(aceasta poate să fie null)**.**

Tabela ORDERS va reține și ea **customer\_id** ca si cheie externă(aceasta NU poate să fie null).

|  |  |  |
| --- | --- | --- |
| ord\_id | ord\_date | cust\_id |
| 154 | 10/10/2020 16:54:20 | 5 |
| 235 | 18/10/2020 12:54:20 | 6 |
| 8745 | 19/12/2020 16:15:20 | 2371 |

Acum relația se află în **FN3**.