Gyakorlati óravázlat 09 – Csoportfüggvények alkalmazása, összekapcsolás gyakorlása

1. Join tapasztalatok
2. Csoportfüggvények
3. Distinct, \*
4. Group by
5. Having

\_\_\_\_\_\_\_\_\_\_\_\_

**Join tapasztalatok**

Átnéztem a múlt órán feltöltött megoldásokat és jajj. Egyrészt eleve kevesen töltik fel. Aztán vannak páran, akik jó poénnak tartják bemásolni a chat-en kapott megoldásokat (ellenkező esetben esélytelen, hogy két karakterre tök azonos kód legyen, pláne 4-5. Egy ügyes pedig ezt is bemásolta: „Chat Conversation End”.

Hívjuk fel újra a figyelmet, hogy hiba külső join használata ott, ahol nincs rá szükség. Az is gyakori tévesztés volt, hogy a feladat megoldásához elég lett volna 2 tábla, de 3-at kapcsoltak össze (azért volt felesleges a 3., mert az abból lekérdezett adat idegen kulcsként benne volt a másik táblában is, más adat meg onnan nem is kellett).

Nem join függő, de gyakori hiba még, hogy a select listába nem kerülnek bele fontos adatok. Pl. ajánljunk szállodát az ügyfeleinknek, és erre csak a szállodák adatai vannak, nem derül ki, hogy kinek ajánljuk. Illetve csak nem egyedi név szerepel, az egyedi azonosító nélkül.

Akiknek egyetem-es adatbázisa volt, azok is a szálláshelyes feladatokat kezdték megoldani. Mondtam, hogy praktikus mindenkinek a saját témájával kezdeni a feladatokat.

**Szinonímák lekérdezése**

A mintapéldák a MINTA.ORSZAGADATOK táblára készült ORSZ szinonímára készültek első körben. Ezek egy táblás lekérdezések. Aztán join gyakorlás miatt használtam ugyanazt a 4 szinonímát is. A Minta sémában található kolcsonzo,nem, dvd és kolcsonzes táblákara a DBA létrehozott ún. publikus szinonímákat mszemely,mnem, mdvd és mkolcsonzes néven.

**Csoportfüggvények**

A csoportfüggvények külön függvénykategóriát alkotnak a dokumentációban (7. Functions/Aggregate functions). Közös tulajdonságuk, hogy nem soronként, hanem sorok csoportjára lehet őket kiszámítani. Legtöbb esetben számfüggvények. A COUNT,MIN,MAX bármilyen adattípusra működik.

**Feladatok:**

1. Számoljuk ki az orsz szinonímából az ott szereplő országok teljes népességét!
2. Számoljuk ki, hogy összesen a népesség mekkora aránya lakik fővárosokban!
3. Mekkora a táblában szereplő országok területének átlaga?

**Darabszám számítás**

A függvény neve COUNT, de többnyire nem mindegy, hogy mit írunk a zárójelbe. A \* a rekordok számát adja eredményül, ha oszlopnevet írunk, akkor az adott oszlop nem null értékeit, ha distinct-et, akkor az adott oszlop különböző értékeinek számát.

Természetesen egy táblás lekérdezésnél az elsődleges kulcs (vagy bármely not null oszlop) darabszámának lekérdezése egyenlő a count(\*)-gal.

select

count(\*) sorok\_száma

,count(orszag) országok\_száma

,count(autojel) autojelek\_száma

,count(distinct autojel) unique\_autojelek\_száma

from orsz;

**Feladatok:**

1. Hány rekord szerepel az orszagadatok táblában? Ez hány ország adata? Hány országnak van megadva az autójele és ez hány különböző autójelet jelent?
2. Hány különböző földrajzi hely és földrész szerepel az orszagadatok táblában?

**Számítások részcsoportokra – szűrés**

1. Hány európai rekord szerepel az orszagadatok táblában? Ez hány ország adata? Hány országnak van megadva az autójele és ez hány különböző autójelet jelent?
2. Hány országban használnak dollárt?

**Számítások csoportokra – csoportosítás**

Az eddigi feladatokban mindig csak csoportfüggvények szerepeltek a szelektlistában. Ha azonban csoportosítani szeretnénk az adatokat, akkor a csoportosító mezőt is fel szoktuk sorolni a select listában (ez szintaktikailag nem kötelező, logikailag viszont ildomos). Illetve fordítva, amit felsorolunk a szelekt listában, arra mindenképpen kell csoportosítani is! (Ahhoz, hogy egy számított értékhez tartozó egyéb adatokat irassunk ki, majd beágyazott lekérdezés fog kelleni, de azt majd csak a következő héten fogjuk megnézni.)

1. Földrészenként hány ország van a táblában?
2. Földrajzi helyenként hány ország van a táblában?
3. Földrészenként és földrajzi helyenként hány ország van a táblában?
4. Földrészenként hány országban használnak dollárt?

**Csoportok számának korlátozása – HAVING**

Az előző lista utolsó példájában a rekordokra volt feltételünk. Ha azonban olyan feltételünk van, ami az egész csoportra vonatkozik, akkor azt nem adhatjuk meg a WHERE szó után, csak a HAVING után. A feltétel vonatkozhat bármilyen csoportokra számított értékre, nem csak azokra, amelyek szerepelnek a selectlistában.

1. Listázzuk ki azokat a földrajzi helyeket (a földrészeket is megadva), amelyek több mint 4 országból állnak!

**Csoportok csoportosítása**

Csoportfüggvényeket egymásba is ágyazhatunk maximum 2 szinten. A group by a belső függvényre fog vonatkozni, a külső függvény pedig a csoportokra kapott értékekre. Ebben az esetben semmi más nem szerepelhet a selectlistában.

1. Mekkora a legkisebb átlagországméret kontinensenként?

select foldresz, round(avg(terulet),2) átlagterület

from orsz

group by foldresz

order by átlagterület desc; --itt kilistázzuk a földrészeket sorrendben, látjuk, hogy Európában a legkisebbek az országok.

select min(round(avg(terulet),2)) legkisebb\_átlagterület

from orsz

group by foldresz; --itt csak a legkisebb értéket listázzuk ki, de nem tudjuk mellé írni, hogy ez Európa!

**Csoportfüggvények összekapcsolt táblákból**

1. Kérdezzük le a dvd kölcsönző azon ügyfeleit, akik több mint 5-ször kölcsönöztek már ki filmet. Mindegyik különböző film volt?

select kid,nev, count(\*) kölcsönzés\_szám, count(distinct film) film\_szám, count(\*)-count(distinct film) különbség

from mszemely s join mkolcsonzes k

on s.kid=k.kolcsonzo

group by kid,nev

having count(\*)>5;

/\*Fontos, hogy a select listában KID is van!

Meg tudjuk nézni a különbséget a \* és a distinct lekérdezés között.\*/

1. Listázzuk ki csak azokat, akiknél volt filmismétlődés!
2. Átlagosan hány évesek az egyes ügyfeleink által kikölcsönzött filmek. Számoljunk először a hiányzó évszámok kihagyásával, aztán úgy, hogy feltétellezük, hogy azok 1960-ban készültek. Csak azokat írassuk ki, ahol az eredmény eltér.
3. Listázzuk ki, hogy hány filmet kölcsönöztek ki a női ügyfeleink. Akik még nem kölcsönöztek, azoknál 0 szerepeljen! Csak azokat írjuk ki, akik 5-nél kevesebb filmet vettek eddig ki!

**Önálló feladatok:**

1. Listázzuk ki, hogy összesen hány szoba szerepel az adatbázisunkban!
2. Listázzuk ki, hogy szálláshelyenként hány szoba és hány férőhely szerepel az adatbázisunkban! Írjuk ki a szállás nevét és címét is!
3. Listázzuk ki azokat az ügyfelelinket, akik foglaltak már le klímás szobát! Hány alkalommal történt ez meg?
4. Ügyfeleink összesen hányszor foglaltak nálunk szállást? Csak azokat listázzuk ki, akik maximum 2-szer!
5. Ügyfeleink összesen hány éjszakára foglaltak nálunk szállást? Csak azokat írjuk ki, akiknél legalább 40 napot számolunk!
6. Mennyi a legtöbb nap, amit valaki a szállásainkon foglalt?
7. Listázzuk ki, hogy összesen hány hallgató szerepel az adatbázisunkban!
8. Listázzuk ki az egyetemi adatbázisunkból a hallgatókat és hogy hány tárgyat vettek fel a 2015/16-os tanév második félévében! Jelenítsük meg a hallgatók nevét és szakját is!
9. Listázzuk ki, hogy hány pénzügy szakos hallgató vette fel az adatbázisban szereplő választható tárgyakat!
10. Szakonként és félévenként számoljuk ki az eredmények átlagát!
11. Mennyi ezeknek az átlagoknak a maximuma illetve minimuma?
12. Listázzuk ki azokat a kötelező tárgyainkat, amit 20-nál kevesebben vettek fel! Azok is jelenjenek meg, amelyeket egy hallgató sem vett fel!
13. Listázzuk ki, hogy átlagosan mekkora a készletünk a termékeinkből!
14. Listázzuk ki, hogy havonta mekkora volt a rendelések összértéke a webshopunkban!
15. Listázzuk ki azokat az ügyfeleinket, akik rendeltek már nálunk kabátot! Fejenként összesen hány darabot?
16. Listázzuk ki, hogy a termékeinkből összesen mekkora összegben rendeltek! Csak azokat jelenítsük meg, ahol az összeg 1000 Ft-nál kevesebb. Amiből még nem rendeltek, azok is jelenjenek meg!
17. Számoljuk ki, hogy termékkategóriánként mekkora volt a bevételünk! Mennyi ezeknek a bevételeknek az átlaga?