

## Relationele Gegevensbanken: SQL 1

---

Stel SELECT-queries op voor het opvragen van data uit de voetbal databank die beantwoorden aan onderstaande vraagstellingen. Voor de volledigheid is het relationele schema van deze databank gegeven in de appendix van dit bestand. Alle praktische informatie met betrekking tot de evaluatieoefeningen kan je terugvinden in het bestand `sql_introductie.pdf` op Ufora. Lees dit bestand zeer goed na vooraleer in te dienen zodat je zeker niks over het hoofd ziet. De deadline voor deze vragenreeks is **15 november 2023, 22u00**.

### Opgave

1. Druk de marktwaarde van elke speler die voorkomt in de speler tabel en van wie de geboortedatum gekend is uit in miljoenen (1 miljoen = 1000000), afgerond tot het dichtste gehele getal (bv. 1.2 miljoen wordt 1 miljoen en 1.5 miljoen wordt 2 miljoen). Deze nieuwe waarde moet weergegeven worden als een geheel getal, gevolgd door een spatie en het woord 'miljoen' (zie voorbeeld). Wanneer de marktwaarde van een speler niet gekend is, moet er een NULL-waarde teruggegeven worden. Naast de marktwaarde van de spelers zijn we ook geïnteresseerd in de leeftijd (in een geheel aantal jaren) van de spelers (dit ten opzichte van de huidige datum). Let goed op dat je deze leeftijd op een juiste manier berekent en controleer (zeker voor de randgevallen) of je berekening effectief klopt. Enkel spelers van wie de geboortedatum gekend is mogen opgenomen worden in het eindresultaat. In de resultatentabel verwachten we vier kolommen met bijhorend datatype: `id` (integer), `familienaam` (varchar), `marktwaarde_miljoen` (varchar) en `leeftijd` (integer). Sorteert de resultaten in deze tabel op basis van aflopende marktwaarde van spelers. Zorg ervoor dat, wanneer de marktwaarde van een speler niet gekend is, deze speler na

alle andere spelers (van wie de marktwaardes wel gekend zijn) wordt gesorteerd. Wanneer twee rijen dezelfde marktwaarde hebben, sorteer dan ook op basis van de leeftijd van de speler, in oplopende volgorde. In het geval dat zowel de marktwaarde alsook de leeftijd van meerdere spelers gelijk zijn aan elkaar, sorteer dan op basis van het id van de speler, in oplopende volgorde.

**Example:** Tabel 1 geeft, als voorbeeld, een overzicht van alle spelers. Een correcte oplossing, gegeven de data in Tabel 1, wordt voorgesteld in Tabel 2. Je mag veronderstellen in dit voorbeeld dat de huidige datum 1 november 2023 is.

id	familienaam	marktwaarde	geboortedatum
1	Witsel	856000	1989-01-12
2	Lukaku	65000000	1993-05-13
3	De Bruyne	65000000	1991-06-28
4	Hazard	15200000	1991-01-07
5	Courtois	NULL	1992-05-11
6	Vertonghen	8000000	1987-04-24
7	Batshuayi	11000000	NULL

Tabel 1: Voorbeelddata oefening 1.

- Geef, voor elke wedstrijd waarvan (i) het aantal toeschouwers gekend is en (ii) de thuisclub of de uitclub (of beide clubs) een aantal goals heeft gescoord dat een veelvoud is van 2 (0, 2, 4, 6, 8...), aan of de match geëindigd is op een gelijkspel. In het eindresultaat verwachten we 6 kolommen met bijhorend datatype: thuisclub (varchar), uitclub (varchar), datum (date), thuisdoelpunten (integer), uitdoelpunten (integer) en gelijkspel (boolean). De kolom gelijkspel bevat, voor elke wedstrijd, de waarde true indien de wedstrijd geëindigd is op een gelijkspel, en false indien dit niet geval is.

**Voorbeeld:** Tabel 3 geeft, als voorbeeld, een overzicht van alle wedstrijden. Een correcte oplossing, gegeven de data in Tabel 3, wordt voorgesteld in Ta-

id	familienaam	marktwaarde_miljoen	leeftijd
2	Lukaku	65 miljoen	30
3	De Bruyne	65 miljoen	32
4	Hazard	15 miljoen	32
6	Vertonghen	8 miljoen	36
1	Witsel	1 miljoen	34
5	Courtois	NULL	31

Tabel 2: Oplossing gebaseerd op voorbeelddata oefening 1.

bel 4.

thuisclub	uitclub	datum	thuisdoelpunten	uitdoelpunten	toeschouwers
Charleroi	RWDM	2023-10-06	2	2	NULL
Cercle Brugge	OHL	2023-10-07	3	3	7850
Westerlo	Kortrijk	2023-10-07	1	0	6300
Anderlecht	Mechelen	2023-10-07	3	1	NULL
Standard	Club Brugge	2023-10-08	2	2	12900
Antwerp	Eupen	2023-10-08	4	1	10450
Gent	Genk	2023-10-08	1	1	14000
STVV	Union	2023-10-08	0	4	8500

Tabel 3: Voorbeelddata oefening 2.

thuisclub	uitclub	datum	thuisdoelpunten	uitdoelpunten	gelijkspel
Westerlo	Kortrijk	2023-10-07	1	0	false
Standard	Club Brugge	2023-10-08	2	2	true
Antwerp	Eupen	2023-10-08	4	1	false
STVV	Union	2023-10-08	0	4	false

Tabel 4: Oplossing gebaseerd op voorbeelddata oefening 2.

3. Tel, voor elk stadion waarvan de naam uit minstens 15 karakters bestaat, het aantal *unieke* klinkers (dit zijn de karakters 'a', 'e', 'i', 'o' en 'u') in deze naam. Dit betekent dat, wanneer er een bepaalde klinker (bv. 'a') meer dan 1 keer voorkomt (bv. in 'Allianz Arena'), deze klinker slechts 1 keer meegeteld mag worden. Werk op een hoofdletterongevoelige manier (bv. 'A', 'E', 'I', 'O' en 'U' zijn ook klinkers, en 'A' is dezelfde klinker als 'a'...). Ook moet je de klinkers met een speciaal karakter (bv. ã, à...) *niet* meetellen. In het eindresultaat verwachten we 2 kolommen met bijhorend datatype: naam (varchar) en aantal (integer). De kolom naam bevat steeds de volledige naam van het stadion en de kolom aantal bevat het resultaat van de telling. Belangrijk, voor deze oefening mag jouw oplossingsquery enkel concepten (functies, operatoren, sleutelwoorden...) bevatten die in de introductiepresentatie van reeks 1 worden opgelijst (bv. de `regexp_replace` functie is *niet* opgelijst in de introductiepresentatie).

**Voorbeeld:** Tabel 5 geeft, als voorbeeld, een overzicht van alle stadions. Een correcte oplossing, gegeven de data in Tabel 5, wordt voorgesteld in Tabel 6. De letters in het vet zijn enkel ter referentie, je hoeft dit niet te implementeren.

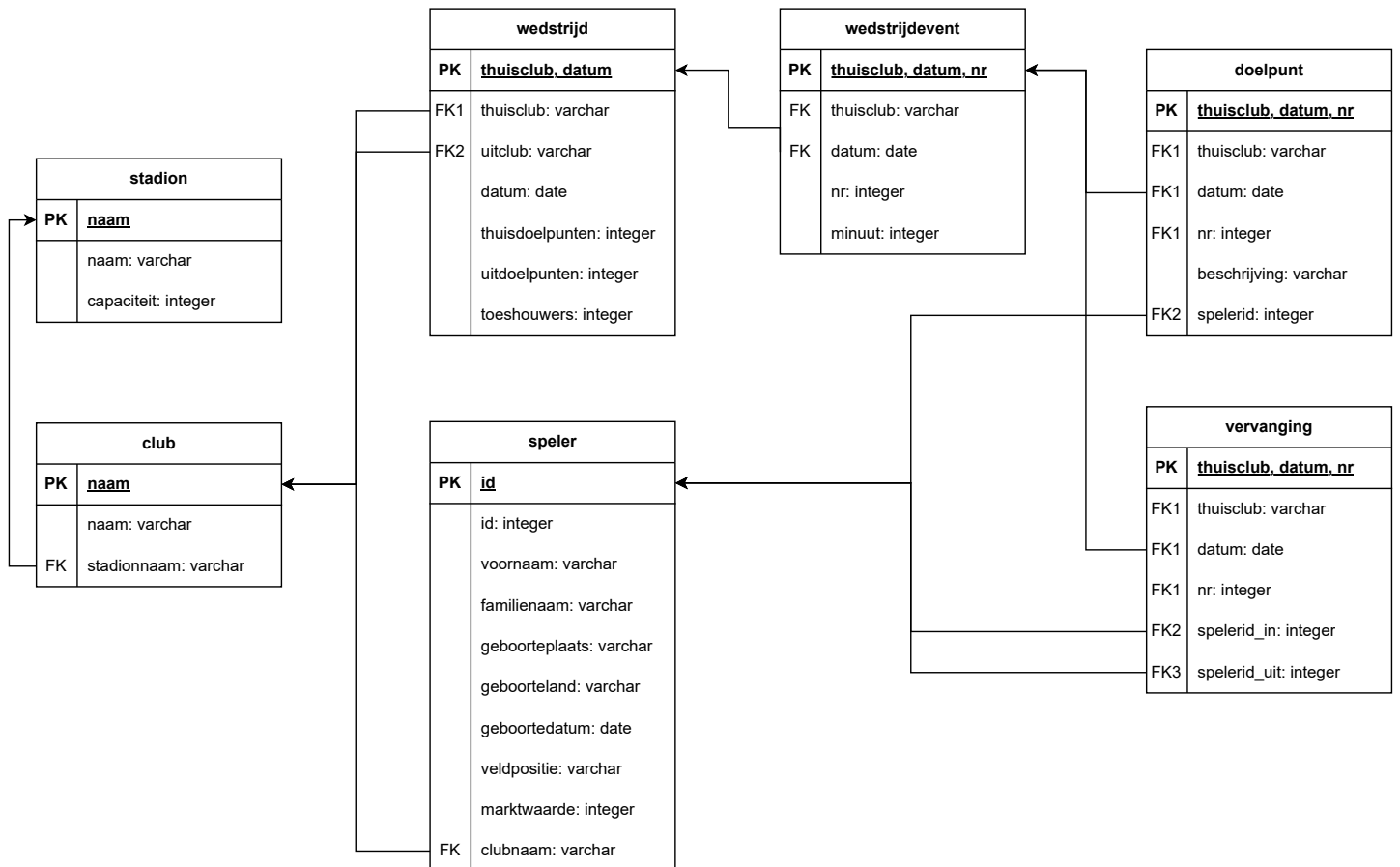
<b>naam</b>	<b>aantal karakters</b>
Jan Breydelstadion	18
Freethielstadion	16
Koning Boudewijnstadion	23
Ghelamco Arena	14
Stayen	6
Stadion 'Lotto Park Stadion'	28
Olympisch Stádion	17

Tabel 5: Voorbeelddata oefening 3.

<b>naam</b>	<b>aantal</b>
<b>Jan Breydelstadion</b>	4
<b>Freethielstadion</b>	4
<b>Koning Boudewijnstadion</b>	5
<b>Stadion 'Lotto Park Stadion'</b>	3
<b>Olympisch Stádion</b>	2

Tabel 6: Oplossing gebaseerd op voorbeelddata oefening 3.

## Appendix: Relationeel schema voetbal databank



In bovenstaande figuur vind je het relationeel databankschema van de voetbal databank. Hierbij wordt iedere basisrelatie weergegeven door een rechthoek, die bovendien een olijsting van alle attributen met bijhorende datatypes bevat. Daarnaast worden de attributen die behoren tot de primaire sleutel (PK) bovenaan weergegeven, en worden vreemde sleutels (FK) voorgesteld door een pijl tussen de betreffende attribuutverzamelingen. Alle extra beperkingen die niet kunnen worden weergegeven in dit schema, worden hieronder opgelijst.

## Extra beperkingen

- stadion:
  - check: capaciteit  $> 0$
- speler:
  - optioneel: voornaam, geboorteplaats, geboorteland, geboortedatum, veldpositie, marktwaarde
  - check: veldpositie  $\in \{ \text{'Goalkeeper'}, \text{'Defender'}, \text{'Midfield'}, \text{'Attack'} \}$ , marktwaarde  $\geq 0$
- wedstrijd:
  - optioneel: toeschouwers
  - uniek: {uitclub, datum}
  - check: thuisdoelpunten  $\geq 0$ , uitdoelpunten  $\geq 0$ , toeschouwers  $\geq 0$ , thuisclub  $\neq$  uitclub
  - controleer bij toevoeging dat het aantal toeschouwers niet groter is dan de capaciteit van het stadion van de thuisclub
  - controleer bij toevoeging dat een club slechts 1 wedstrijd per datum speelt
- wedstrijdevent:
  - check: nr  $\geq 1$ , minuut  $\geq 0$ , minuut  $\leq 120$
- doelpunt:
  - controleer bij toevoeging dat het totaal aantal doelpunten dat gelieerd is aan deze wedstrijd niet groter is dan de som van de scores van de thuis- en uitclub op het einde van deze wedstrijd
- vervanging:
  - check: speler\_in  $\neq$  speler\_uit