

# MySQL: GROUP BY

# Noch einmal der Merksatz:

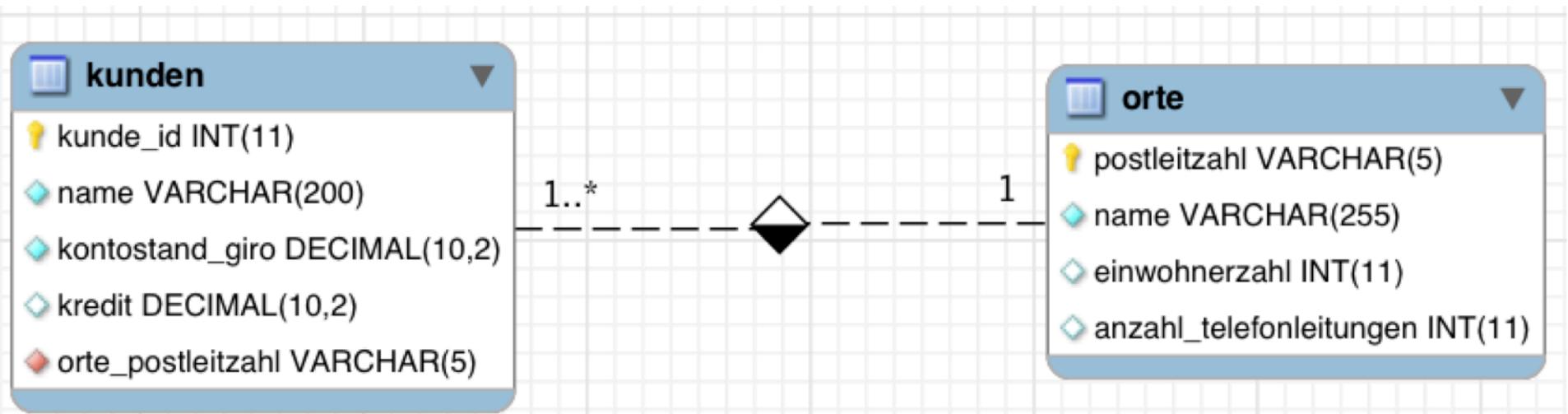
Warum geht Herbert oft laufen?

```
SELECT ...  
FROM ...  
WHERE ...  
GROUP BY ...  
HAVING ...  
ORDER BY ...  
LIMIT
```

# Beispieldatenbank "Kunden"

kunden (kunde\_id, name, ↑ort\_postleitzahl,  
kontostand\_giro, kredit)

orte (postleitzahl, name, einwohnerzahl,  
anzahl\_telefonleitungen)



# Beispieldatenbank "Kunden"

kunden (kunde\_id, name, ↑ort\_postleitzahl, kontostand\_giro, kredit)  
orte (postleitzahl, name, einwohnerzahl, anzahl\_telefonleitungen)

postleitzahl	name	einwohnerzahl	anzahl_telefonleitungen
80995	München	1000000	385
79312	Emmendingen	40000	12
79111	Freiburg	280000	195
20095	Hamburg	2000000	1004

	kunde_id	name	ort_postleitzahl	kontostand_giro	kredit
▶	1	John	79111	182.00	-430320.22
	2	Herbert	79312	10291.32	-10000.00
	3	Sabina	79312	-253.21	-3205.32
	4	Mary	79111	-832.01	NULL
	5	Heinrich	79111	15302.85	0.00
	6	Usal	80995	23012.21	NULL
	7	Johannes	80995	159.31	0.00
	8	Carla	79312	503.06	-15302.68
	9	Ludowika	79111	25201.07	-82213.99
	10	Niemand	99999	-5021.30	-3024.21



# GROUP BY

Bundestag: Wie viele Abgeordnete sitzen hier?



# GROUP BY

Bundestag: Wie viele Abgeordnete sitzen hier?

```
SELECT COUNT (*)  
FROM abgeordnete
```

-- Ergebnis: 16



# GROUP BY

Und wie viele Abgeordnete der einzelnen Parteien sitzen hier?

Farbe "rot" : ...

Farbe "grün" : ...

Farbe "orange" : ...

etc.

```
SELECT COUNT(*)  
FROM abgeordnete
```

```
-- Ergebnis: 16
```



# GROUP BY

Und wie viele Abgeordnete der einzelnen Parteien sitzen hier?

Farbe "rot" : 3

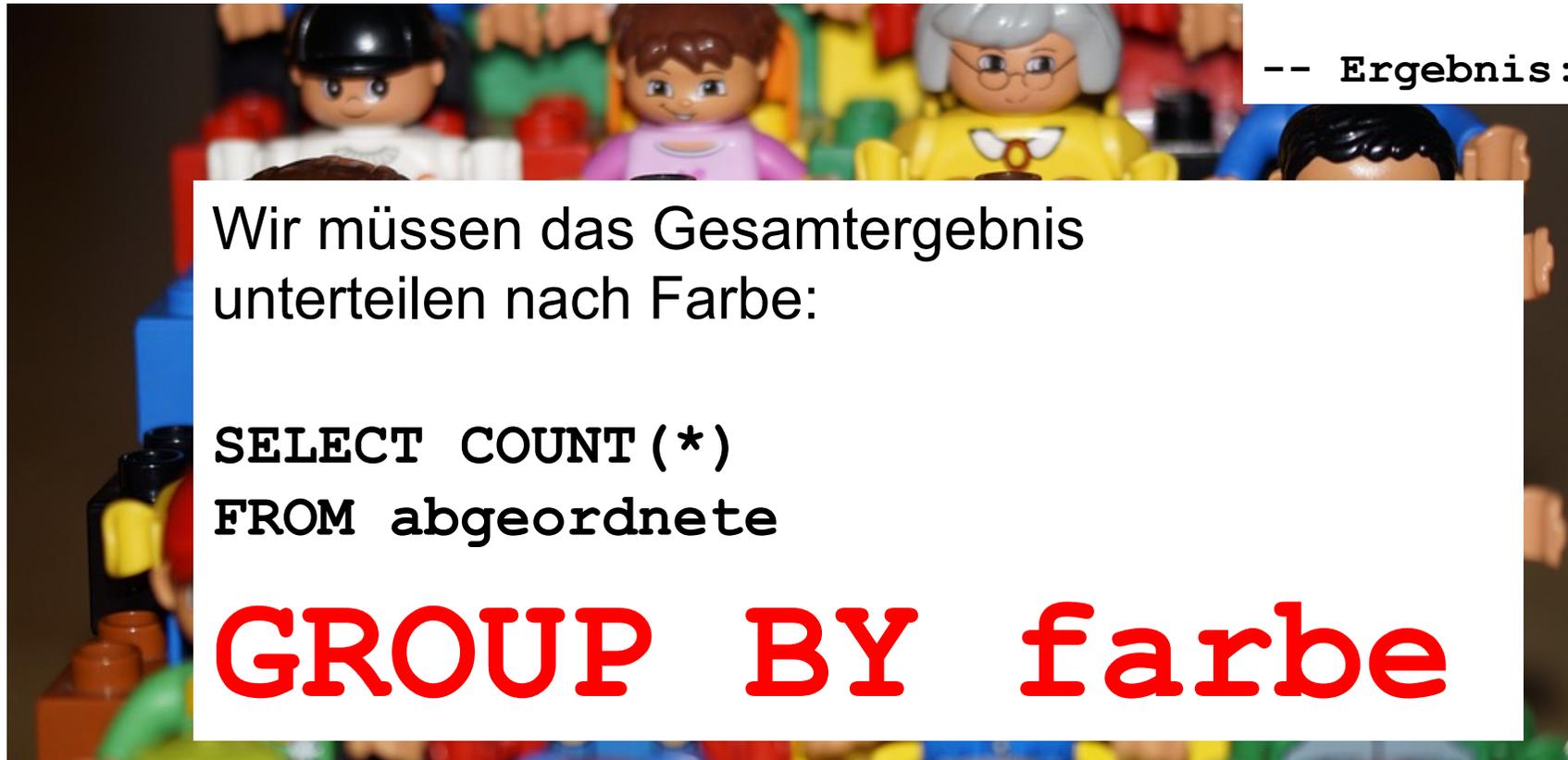
Farbe "grün" : 2

Farbe "orange" : 2

etc.

```
SELECT COUNT(*)  
FROM abgeordnete
```

```
-- Ergebnis: 16
```



Wir müssen das Gesamtergebnis unterteilen nach Farbe:

```
SELECT COUNT(*)  
FROM abgeordnete
```

**GROUP BY farbe**

# GROUP BY

Jetzt wieder unsere Beispieldatenbank:  
Wie viele Personen wohnen in dein einzelnen Orten?

Wir müssen also die Personen nach Ort aufteilen:

**GROUP BY orte.postleitzahl**

postleitzahl					
80995					
79312					
79111	Freiburg	280000	195		
20095	Hamburg	2000000	1004		
1	John	79111	182.00	-430320.22	
2	Herbert	79312	10291.32	-10000.00	
3	Sabina	79312	-253.21	-3205.32	
4	Mary	79111	-832.01	NULL	
5	Heinrich	79111	15302.85	NULL	
6	Usal	80995	23012.21	NULL	
7	Johannes	80995	159.31	NULL	
8	Carla	79312	503.06	-15302.68	
9	Ludowika	79111	25201.07	-82213.99	
10	Niemand	99999	-5021.30	-3024.21	

# GROUP BY

**Achtung: Fehler!**

```
SELECT
    orte.name, COUNT(*) as anzahlKundenProOrt
FROM
    kunden, orte
WHERE
    orte.postleitzahl = kunden.ort_postleitzahl
```

9 Kunden, die einem Ort zugeordnet sind

name	anzahlKundenProOrt
München	9

**Aber: name "München" ist nicht korrekt  
(da Ergebnis eine Zusammenfassung ist, wird hier erster Ort angezeigt)**

**Was wir wollen:  
Das Gesamtergebnis nach Orten gruppieren (München, Freiburg ...)**

# GROUP BY

```
SELECT
    orte.name, COUNT(*) as anzahlKundenProOrt
FROM
    kunden, orte
WHERE
    orte.postleitzahl = kunden.ort_postleitzahl
GROUP BY
    orte.postleitzahl;
```

9 Kunden, die einem Ort zugeordnet sind,  
gruppiert nach Ort

```
-- gleiches Ergebnis:
-- GROUP BY
-- kunden.orte_postleitzahl
```

name	anzahlKundenProO
Emmendingen	3
Freiburg	4
München	2

# GROUP BY

```
SELECT
    orte.name, SUM(kredit)
        AS 'Kredite in den einzelnen Orten'
FROM
    kunden, orte
WHERE
    orte.postleitzahl = kunden.ort_postleitzahl
GROUP BY
    kunden.ort_postleitzahl;
```

	name	Kredite in den einzelnen Orten
▶	Freiburg	-512534.21
	Emmendingen	-28508.00
	München	0.00

# GROUP BY

gruppiert Ergebnisse nach dem Kriterium, z.B.

- Kunden, die im gleichen Ort wohnen

```
GROUP BY orte.postleitzahl
```

-- oder

```
GROUP BY kunden.postleitzahl
```

- Orte, die die gleiche Postleitzahl haben

```
GROUP BY orte.postleitzahl
```

- Kunden, die bei der gleichen Bank sind

```
GROUP BY kunden.bank_id
```