MASARYKOVA UNIVERZITA
INSTITUT BIOSTATISTIKY A ANALÝZ

Databázové systémy a SQL

Lekce 9

Daniel Klimeš

Cvičení

- Tabulka LPZ
- řádek = úmrtí

- Pohlaví
  - 1 = Muž
  - 2 = Žena
- Rodinný stav

  kod="1" popis="svobodný(á)"
  kod="2" popis="ženatý/vdaná"
  kod="3" popis="rozvedený(á)"
  kod="4" popis="ovdovělý(á)"
  kod="5" popis="partnerství"
  kod="6" popis="zaniklé partnerství rozhodnutím"
  kod="7" popis="zaniklé partnerství smrtí"
  kod="8" popis="nezjištěno"

Cvičení

- Zjistěte průměrný věk úmrtí v jednotlivých letech 1987 – 2022
- Průměrný věk úmrtí v roce 2022 po okresech
- Průměrný věk úmrtí v roce 2022 podle pohlaví
- Průměrný věk úmrtí v roce 2022 podle pohlaví a rodinného stavu
- Průměrný věk úmrtí v jednotlivých letech 1987 – 2022
  a změna oproti předchozímu roku
- Průměrný věk úmrtí po měsících v období 2020-2022
- Počet zemřelých po měsících v období 2019 – 2022, počet zemřelých do 50 let věku
- Srovnejte počty zemřelých v pátek 13. se sobotou 14.
  - snadnější varianta pod sebou
  - složitější vedle sebe s uvedením rozdílu

Řešení

```
SELECT
to_char(datumumrti, 'yyyy'),
avg(age(datumumrti,datumnarozeni)) FROM lpz
GROUP BY to_char(datumumrti, 'yyyy')
ORDER BY 1
```

```
SELECT
OkresBydliste,
avg(age(datumumrti,datumnarozeni)) FROM lpz
WHERE Datumumrti BETWEEN '2022-01-01' AND '2022-12-31'
GROUP BY OkresBydliste
ORDER BY 2
```

```
SELECT
Pohlavi,
avg(age(datumumrti,datumnarozeni)) FROM lpz
WHERE Datumumrti BETWEEN '2022-01-01' AND '2022-12-31'
GROUP BY Pohlavi
```

```
SELECT
rodinnystav,
Pohlavi,
avg(age(datumumrti,datumnarozeni)) FROM lpz
WHERE Datumumrti BETWEEN '2022-01-01' AND '2022-12-31'
GROUP BY rodinnystav,Pohlavi
```

Řešení

```
SELECT rok, vek, vek - lag(vek,1,NULL) OVER (ORDER BY rok) zmena
FROM (
SELECT
to_char(datumumrti, 'yyyy') ROK,
avg(age(datumumrti,datumnarozeni)) vek
FROM lpz
GROUP BY to_char(datumumrti, 'yyyy')
) x
ORDER BY 1
```

```
SELECT
to_char(datumumrti, 'yyyy-mm') Mesic,
avg(age(datumumrti,datumnarozeni)) vek
FROM lpz
WHERE  Datumumrti BETWEEN '2020-01-01' AND '2022-12-31'
GROUP BY to_char(datumumrti, 'yyyy-mm')
```

```
SELECT
to_char(datumumrti, 'yyyy-mm') Mesic,
count(*) Pocet,
sum(CASE WHEN age(datumumrti,datumnarozeni) < INTERVAL'50 years' THEN 1 ELSE 0 END) Do50let
FROM lpz
WHERE  Datumumrti BETWEEN '2019-01-01' AND '2022-12-31'
GROUP BY to_char(datumumrti, 'yyyy-mm')
ORDER BY 1
```

Řešení

```
SELECT datumumrti, count(*) Pocet FROM lpz
WHERE to_char(datumumrti, 'dd-Day') IN
('13-Friday   ', '14-Saturday ')
GROUP BY datumumrti
ORDER BY 1

SELECT datumumrti , pocet_patek, pocet_sobota, pocet_patek
- pocet_sobota rozdil
FROM (
SELECT datumumrti, count(*) Pocet_patek,
lead(count(*),1,'0')
OVER (PARTITION BY to_char(datumumrti, 'yyyymm') ORDER BY
datumumrti) pocet_sobota
FROM lpz
WHERE to_char(datumumrti, 'dd-Day') IN ('13-Friday   ',
'14-Saturday ')
GROUP BY datumumrti
) x WHERE pocet_sobota > 0
ORDER BY 1
```