

Databázové systémy a SQL

Lekce 6

Daniel Klimeš, Monika Kratochvílová

- Nakopírovat složku data do dokumentů
- Nastavit oprávnění na složku pro NETWORK_SERVICE
- COPY sites FROM 'c:/Users/student/Documents/data/sites.txt'
NULL " ENCODING 'UTF8';
- COPY patients FROM
'c:/Users/student/Documents/data/patients.txt' NULL " ENCODING
'UTF8';
- COPY studies FROM 'c:/Users/student/Documents/data/studies.txt'
NULL " ENCODING 'UTF8';
- COPY patient_study FROM
'c:/Users/student/Documents/data/patient_study.txt' NULL "
ENCODING 'UTF8';

- CASE WHEN podmínka THEN vysledek
WHEN podmínka2 THEN vysledek 2
ELSE vysledek 3 END
- Až 127 WHEN,
- ELSE nepovinné,
- Vyhodnocování končí na první splněné podmínce
- Všechny výsledky musí být stejného datového typu

Příklad:

```
SELECT vek,
       CASE WHEN vek IS NULL THEN 'neznamo'
            WHEN vek < 30 THEN 'kat < 30'
            WHEN vek < 50 THEN 'kat 30-49'
            WHEN vek < 65 THEN 'kat 50-64'
            ELSE 'kat 65 a starsi' END kategorie
FROM
(SELECT EXTRACT (YEAR FROM AGE(CURRENT_DATE,date_of_birth)) vek
FROM patients) jmeno_vnoreneho /*POSTGRES SQL*/
```

```
SELECT COUNT(*) FROM student
```

```
SELECT studium, COUNT(*) FROM student
GROUP BY studium
```

```
SELECT studium, COUNT(*) pocet , COUNT(*) *
100.0/(SELECT COUNT(*) FROM student) procento FROM
student
GROUP BY studium
```

```
SELECT studium, COUNT(*) pocet, COUNT(*) *100.0 /
SUM(COUNT(*) OVER ()) procento FROM student
GROUP BY studium
```

Procentické zastoupení – standardní SQL:

```
SELECT study_id, COUNT(*),
COUNT(*) * 100.0 / (SELECT COUNT(*) FROM patient_study) procento
FROM patient_study
GROUP BY study_id
```

Analytická funkce

```
SELECT study_id, COUNT(*),
COUNT(*) / SUM(COUNT(*) OVER ()) * 100 procento
FROM patient_study
GROUP BY study_id
```

```
UPDATE student SET ukonceni = 'Z'
WHERE mod(uco,2) = 1 – Rozdělení datového souboru
```

```
SELECT ukonceni, studium, count(*) pocet, COUNT(*) *100.0 /
SUM(COUNT(*)) OVER () procento FROM student
GROUP BY ukonceni, studium
ORDER BY ukonceni
```

```
SELECT ukonceni, studium, count(*) pocet, COUNT(*) *100.0 /
SUM(COUNT(*)) OVER () procento ,
COUNT(*) *100.0 / SUM(COUNT(*)) OVER (PARTITION BY ukonceni)
proc_podskupiny
FROM student
GROUP BY ukonceni, studium
ORDER BY ukonceni
```

```
SELECT study_id, study_site, COUNT(*),
COUNT(*) / SUM(COUNT(*)) OVER (PARTITION BY study_id) * 100 procento
FROM patient_study
GROUP BY study_id, study_site
```

```
SELECT studium, COUNT(*) pocet FROM
student
GROUP BY studium
```

```
SELECT studium, COUNT(*) pocet,
SUM(COUNT(*)) OVER (ORDER BY studium)
FROM student
GROUP BY studium
```

```
SELECT ukonceni, studium, COUNT(*) pocet,
SUM(COUNT(*)) OVER (PARTITION BY ukonceni ORDER BY studium)
kumulace_skupina,
SUM(COUNT(*)) OVER (ORDER BY ukonceni, studium) kumulace_celkem
FROM student
GROUP BY ukonceni, studium
ORDER BY ukonceni, studium
```



```
SELECT study_id, study_site, COUNT(*),
SUM(COUNT(*)) OVER (PARTITION BY study_id ORDER BY study_site)
kumulace
FROM patient_study
GROUP BY study_id, study_site
```

AVG(sloupec) OVER
(ORDER BY sloupec ROWS BETWEEN x PRECEDING AND CURRENT ROW)

• **ROWS BETWEEN**



- **UNBOUNDED PRECEDING**
- **UNBOUNDED FOLLOWING**
- **CURRENT ROW**
- **počet řádků PRECEDING**
- **počet řádků FOLLOWING**

```
CREATE TABLE pocet_pacientu as
SELECT TO_CHAR(date_of_enrollment, 'yyyy-mm') mesic, COUNT(*)
pocet FROM patient_study
WHERE date_of_enrollment >= '2004-01-01'
GROUP BY TO_CHAR(date_of_enrollment, 'yyyy-mm')
ORDER BY TO_CHAR(date_of_enrollment, 'yyyy-mm')
```

```
SELECT * FROM pocet_pacientu
ORDER BY mesic
```

```
SELECT AVG(pocet) FROM
pocet_pacientu
```

```
SELECT mesic, pocet,
ROUND(AVG(pocet) OVER (ORDER BY mesic ROWS BETWEEN 3
PRECEDING AND CURRENT ROW),1) klouzavy_prumer
FROM pocet_pacientu
```

1) Spočítejte v tabulce pocet_pacientu kumulativní počet pacientů

```
SELECT mesic, pocet
FROM pocet_pacientu
ORDER BY mesic
```

```
SELECT mesic, pocet
FROM pocet_pacientu
ORDER BY mesic
```

```
SELECT mesic, pocet,
SUM(pocet) OVER (ORDER BY mesic)
FROM pocet_pacientu
ORDER BY mesic
```

- Přidejte ke kumulativnímu počtu kumulativní procento

- Přidejte ke kumulativnímu počtu kumulativní procento

```
SELECT mesic, pocet, SUM(pocet) OVER (ORDER BY mesic),
SUM(pocet) OVER(), SUM(pocet) OVER (ORDER BY mesic) * 100 /
SUM(pocet) OVER() FROM pocet_pacientu
ORDER BY mesic
```


- Přepište dotaz na původní tabulku patient_study

```
SELECT mesic, pocet, SUM(pocet) OVER (ORDER BY mesic),
SUM(pocet) OVER(), SUM(pocet) OVER (ORDER BY mesic) * 100 /
SUM(pocet) OVER() FROM pocet_pacientu
ORDER BY mesic
```

```
SELECT TO_CHAR(date_of_enrollment, 'yyyy-mm') mesic, COUNT(*) pocet
FROM patient_study
WHERE date_of_enrollment >= '2004-01-01'
GROUP BY TO_CHAR(date_of_enrollment, 'yyyy-mm')
ORDER BY TO_CHAR(date_of_enrollment, 'yyyy-mm')
```

- Přepište dotaz na původní tabulku patient_study

```
SELECT mesic, pocet, SUM(pocet) OVER (ORDER BY mesic),
SUM(pocet) OVER(), SUM(pocet) OVER (ORDER BY mesic) * 100 /
SUM(pocet) OVER() FROM pocet_pacientu
ORDER BY mesic
```

```
SELECT TO_CHAR(date_of_enrollment, 'yyyy-mm') mesic, COUNT(*)
pocet,
SUM(COUNT(*)) OVER (),
SUM(COUNT(*)) OVER (ORDER BY TO_CHAR(date_of_enrollment, 'yyyy-
mm')) * 100 / SUM(COUNT(*)) OVER() kum_procento
FROM patient_study
WHERE date_of_enrollment >= '2004-01-01'
GROUP BY TO_CHAR(date_of_enrollment, 'yyyy-mm')
ORDER BY TO_CHAR(date_of_enrollment, 'yyyy-mm')
```

- Zobrazte kumulativní procentické zastoupení pacientů podle věku
 - Věk, počet pacientů, kumulativní procento

- Zobrazte kumulativní procentické zastoupení pacientů podle věku
 - Věk, počet pacientů, kumulativní procento

```
SELECT EXTRACT (YEAR FROM AGE(date_of_birth))
FROM patients limit 100
```

```
SELECT vek, COUNT(*) FROM (
  SELECT EXTRACT (YEAR FROM AGE(date_of_birth)) vek
  FROM patients) a
WHERE vek > 0 and vek < 100
GROUP BY vek
```

```

SELECT vek, pocet, kum_pocet * 100 / pocet_celkem kum_procento
FROM (
  SELECT vek, COUNT(*) pocet, SUM(COUNT(*)) OVER (ORDER BY VEK)
    kum_pocet, SUM(COUNT(*)) OVER () pocet_celkem
  FROM (
    SELECT EXTRACT (YEAR FROM AGE(date_of_birth)) vek
    FROM patients) a
  WHERE vek > 0 and vek < 100
  GROUP BY vek
  ORDER BY vek
) b

```

Napište dotaz, který vypíše nejmladšího a nejstaršího pacienta.

Napište dotaz, který vypíše nejmladšího a nejstaršího pacienta.

```
SELECT * FROM patients
```

```
WHERE
```

```
date_of_birth = (SELECT MAX(date_of_birth) FROM patients)
```

```
OR
```

```
date_of_birth = (SELECT MIN(date_of_birth) FROM patients);
```