

Zápočet Databáze v biomedicině 2017

Formou analytické zprávy ve wordu nebo powerpointu prezentujte předaná data. Zpráva by měla obsahovat minimálně informace o:

- Celkové počty záznamů, počet center
- Obsahu (vyplněnost, rozložení, rozsah) jednotlivých sloupců
- Kontrolu vazby mezi dvěma soubory (kontrola závislosti dle cizích klíčů)
- Počty záznamů dle center v čase dle měsíců, dle typů vyšetření
- Průměrná velikost léze (X * Y) dle center v čase

V rámci předpřípravy dat připravte SQL, které:

- Odstraní záznamy před rokem 2004
- Odstraní léze s velikostí (lesion_x * lesion_y) > 5000
- Odstraní záznamy center, které nemají záznamy po roce 2015

Ke zpracování použijte databázi (import, čištění, vytěžení). Použité SQL vložte jako přílohu zprávy. Formální úpravu zprávy dle svého nejlepšího svědomí.

Popis dat:

Soubor exam.txt

Provedená vyšetření v jednotlivých centrech (pracovištích). Sloupce v pořadí:

Sloupec	Popis
center_id	Identifikátor centra
examination_id	Identifikátor vyšetření v rámci centra
method	Typ provedeného vyšetření
month_exam	Rok a měsíc, kdy bylo vyšetření provedeno
result	Kód výsledku vyšetření
type	Typ vyšetřované tkáně

Soubor lesion.txt

V rámci vyšetření nalezená patologická ložiska, vazba 1:n k souboru exam.txt, propojitelné přes kombinaci center_id a examination_id

Sloupec	Popis
center_id	Identifikátor centra
lesion_id	Identifikátor ložiska v rámci centra
examination_id	Identifikátor vyšetření v rámci centra
strana	Strana těla, vlevo, vpravo
x	Rozměr léze (délka)
y	Rozměr léze (šířka)

