

Popis datového souboru

Tři datové soubory:

1. Patient.csv (obsahuje základní informace o pacientce)
 - a. **ID pacientky**
 - b. Datum narození
 - c. Pojišťovna
2. Examination.csv (obsahuje informace o provedených screeningových vyšetřeních)
 - a. **ID pacientky**
 - b. **ID vyšetření**
 - c. Plátce daného vyšetření

KOD	ZPŮSOB ÚHRADY VYŠETŘENÍ
1	vyšetření hradí zdravotní pojišťovna
2	vyšetření hradí pacientka (samoplátkyně)
3	vyšetření hrazeno z jiných zdrojů (např. akce AVON)
4	centrální zvaní

- d. Datum vyšetření
- e. Výsledek vyšetření

KOD	VÝSLEDEK VYŠETŘENÍ
1	nejednoznačný nález - potřeba dalšího vyšetření
2	negativní
3	benigní
4	pravděpodobně benigní
5	podezřelý nález
6	pravděpodobně maligní nález
7	bez známek recidivy
8	suspektní recidiva
9	recidiva
10	známá malignita (BIRADS 6)
51	pravděpodobnost malignity nízká (BIRADS 4a)
52	pravděpodobnost malignity je střední (BIRADS 4b)
53	pravděpodobnost malignity je vysoká (BIRADS 4c)

3. Indicators.csv (obsahuje dodatečné informace ke screeningovému vyšetření)
 - a. **ID vyšetření**
 - b. Doplnující vyšetření u pacientky (1 – ANO / 0 – NE)
 - c. Ultrasonografie jako doplňující vyšetření (1 – ANO / 0 – NE)
 - d. Doplnující mamografie jako další vyšetření (1 – ANO / 0 – NE)
 - e. Biopsie jako doplňující vyšetření (1 – ANO / 0 – NE)
 - f. Přítomnost malignity (1 – ANO / 0 – NE)

- g. Pořadí screeningového vyšetření u pacientky (1 – další screeningové vyšetření / 0 – první screeningové vyšetření)

Úkoly

Práce v PostgreSQL:

4. Import dat
 - a. Vytvořit tabulky
 - b. Načíst data
5. Předpříprava dat
 - a. Odstraňte řádky, kde
 - i. Je budoucí datum narození nebo vyšetření
 - ii. Kódy parametrů Plátce daného vyšetření nebo výsledku jsou mimo uvedený rozsah
 - b. U indikátorů prázdné hodnoty nastavte na 0.

Vytvoření reportu:

6. Celkové počty provedených vyšetření v letech.
7. Podíly vyšetření podle jednotlivých indikátorů (procento vyšetření s doplňujícím vyšetření, s USG, atd.) v jednotlivých letech.
8. Počty vyšetření v pětiletých věkových kategoriích – rozdíl v letech 2014 a 2015.
9. Počty žen s jedním, dvěma a více vyšetřeními
10. Vytvořte minimálně jeden graf podle svého vlastního uvážení, co by mohlo být zajímavé v datovém souboru vizualizovat. Iniciativa a originalita se cení. Nezapomeňte pak grafy pořádně popsat. Držte se zásady, že každý slaid vytržený z kontextu by měl mít sám o sobě vypovídající hodnotu.

Vytvoření reportu – Powerpointové prezentace:

1. Vytvořte powerpointovou prezentaci, do které vložíte všechny výstupy úkolů. Jaké grafy vytvoříte je na vás a na všem citu pro vizualizaci. Je také na vás, v jakém softwarovém prostředí budete grafy vytvářet (Rko, Excel, ...).

Vložte do odevzdárny:

2. Do odevzdárny vložte sql script a powerpointovou prezentaci. Obě pojmenujte svým příjmením.