

# Prázdné řádky či sloupce

Prázdné řádky či sloupce vznikají nejčastěji převodem formátované tabulky z tabulkového editoru (sloučené buňky, vzorce) přímo do formátu CSV. Takový obsah CSV sice může být syntakticky validní, ale znesnadňuje práci s reprezentovanými daty. Je totiž třeba nejprve zjišťovat, proč tam mezery jsou, co znamenají, zda například neoddělují součty a podobně, a pak je třeba data upravit předtím, než je možné s nimi pracovat.

Příklad: strojově nečitelné CSV - špatný oddělovač (středník), prázdné řádky, formátování pro tisk

```
;;;;;;;;;;;;;
Back to TOC;;;;;;;;;;;;;
r2 : R2. Máte v Brně trvalé bydliště (hlášené na úřadě)?;;;;;;;;;;;;;
;%;počet;;;;;;;;;;;;;
Ano;89,1%;1385;;;;;;;;;;;;;
Ne;10,9%;169;;;;;;;;;;;;;
CELKEM;100,0%;1554;;;;;;;;;;;;;
"Total sample; Weight: Weight; base n = 1554";;;;;;;;;;;;;
;;;;;;;;;;;;;
Back to TOC;;;;;;;;;;;;;
r3 : R3. Jak dlouho bydlíte v Brně? ;;;;;;;;;;;;;;
;%;počet;;;;;;;;;;;;;
```

## Řešení

Data do CSV je třeba převádět tak, aby vznikla tabulka bez děr, snadno zpracovatelná, tak, aby každý řádek obsahoval kompletní informaci o reprezentované entitě či záznamu. Pokud je zdrojem tabulka formátovaná pro tisk či pro čtení lidmi, je třeba nejprve bez újmy na obsahu

- Odstranit formátování
- Odstranit sloučené buňky například duplikací jejich obsahu na všechny sloučené řádky či sloupce
- Odstranit vzorce například materializací jejich hodnot
- Odstranit prázdné řádky či sloupce před začátkem dat

From:

<https://opendata.gov.cz/> - Otevřená data

Permanent link:

<https://opendata.gov.cz/standardy:csv:pr%C3%A1zdn%C3%A9-%C5%99%C3%A1dky-%C4%8Di-sloupce>

Last update: 2020/11/06 16:50

