

NPFL142.C4DHI – domácí úkol #3

Příklad 3.1

Data: `library(hcandersenr)`

Kód: `andersen.h.R` vytvořte na clusteru ve svém adresáři `3/andersen/`

- Ve které anglicky napsané knížce je nejvíce podstatných jmen (UPOS = noun). Uveďte její název a četnost podstatných jmen.
- Pracujte s knížkami, které jsou k dispozici ve všech pěti jazycích. Pro každý jazyk zvlášť vizualizujte četnosti slovních druhů ve všech knížkách (viz příznak `upos`). Tj. vytvoříte celkem 5 grafů.

Příklad 3.2

Příprava na workshop věnovaný datově orientovaným projektům.

Uveďte alespoň jednu výzkumnou otázku, kterou si kladete. Proč vás zajímá? Hledali jste již na ni odpověď?

- Pokud ano, popište, jaká data jste použil/a a jakou jste zvolil/a metodologii. Narazil/a jste na nějaké problémy? Máte-li prezentaci/článek ke sdílení, uveďte prosím link.
- Pokud ne, máte již data, ve kterých budete hledat odpověď? Pokud ano, popište je - počet příkladů, počet atributů, typy atributů. Pokud ne, prozkoumejte repozitář [LINDAT](#), jestli v něm jsou uložena data, která byste mohl/a použít.

Pokud nemáte žádnou výzkumnou otázku, zformulujte jednu otázku nad datasetem *Titanic*, jednu nad *Migrant Stories* and jednu nad *Andersen*.

Termín odevzdání: 18.4.2024 15:00