

Programming for Data Science

Winter term 2023/2024

Assignment 5

Issued: 2023-10-23

1 Replacer Tool

Write a program that uses the *argparse* module and takes as command line arguments:

- a filename as positional argument
- optional argument (`-p/--pattern`) with default `,`
- optional argument (`-r/--replace`) with default `,`
- optional argument (`-o/--out`)
- flag (`-fo/--force`)

Your program should read the provided input file and write an output file with all occurrences of string passed to `--pattern` option replaced with string passed to `--replace` option. If option `--out` is provided this must be used as output file. Otherwise for input file `/some/path/filename.ext` use `/some/path/filename_mod.ext` as output file. If the output file already exists terminate the program with an informative error message unless the `--force` flag is set - in this case overwrite the existing file.

2 Regular Expressions

Use regular expressions to:

- convert dates in format `yyyy-mm-dd` to format `dd-mm-yyyy`.
- return all integer numbers in a string together with their positions.
- return all float numbers in a string together with their positions.
- convert all `snake_case` words to `camelCase`.

- “compress” strings by replacing a run of $n > 2$ identical characters c by $c[n]$, e.g.
`acccgtaaaaatc` \rightarrow `ac[3]gta[5]tc`

Hint: especially for the last two - the replacement argument of the `sub` function/method can also be a (lambda) function!