

List Cleaning and Statistics

Warning: Windows will always convert the byte `\x0a` to two bytes. So don't be astonished if cleaned files with less entries will be larger than the original.

For a better overview and handling of the bytes it is recommended to use the scripts on Unix systems.

- Remove multiple lines from lists in a directory of average size (about 2MB): [removeMultiLinesInDir.py](#)
- Remove multiple lines from lists in a directory of huge size: [removeMultiLinesInHugeDir.py](#)
- Extract multiple lines from lists in a directory of average size: [extractMultiLinesFromDir.py](#)
- Extract multiple lines from lists in a directory of huge size: [extractMultiLinesFromHugeDir.py](#)
- Initial Character Distribution in a directory of average or huge size: [initialCharDistribution.py](#)

Usage:

1. Install Python27
2. Change CUSTOM DEFINED PARAMETERS in chosen script
3. Run 'cd path/to/folder/with/scripts' in cmd
4. Run 'python nameOfScript.py'