

*** MGsort.exe - Sort a text file

(c) Markus Gnam 2020. Version 1.0.4.5 20200103

USAGE: MGsort [sort field] ... [option] ... <input file> <output file>
SIMPLEST CALL, SORTING THE WHOLE LINE: MGsort <input file> <output file>
Default sort order is alfabetic sorting using ASCII collating sequence.

SORT FIELD: can be sorted either as String (/S) or as Field (/F).

1. **String (/S):** /S=P,L => string from position P with length L
(L can be ommited, this means sorting until the end of the line from P.
If no /S is submitted, the whole line will be sorted.

/S=P,L can be repeated to submit a concatenated sort string)

2. **Field (/F):** [/T<Separator>] /F<field number>=P,L

Field F with given field number (separated with sign after /T)

If no /T is submitted, this means default separator is "\t" (tabulator)

(L can be ommited, this means sorting until the end of the field from P.

=P,L can be ommited, this means sorting just the whole field.

/F<field number>=P,L can be repeated to submit a concatenated sort string)

Instead of a field number a field name can be given, e.g. /Fname=last_name

Further specifications of the sort field, f.e. /S=12,19:n or /F1:b:i

Specification appended with ":" directly after the sort field.

Append **:n** after the sort field in order to sort numerically

Append **:b** after the sort field in order to ignore leading/trailing blanks

Append **:c** after the sort field in order to ignore case (letters [A-Za-z])

Append **:c=u** after the sort field to ignore case and umlauts (ANSI)

Append **:c=uDOS** after the sort field to ignore case and umlauts (DOS)

Append **:i** after the sort field in order to ignore case and accents (ANSI)

Append **:i=d** after the sort field to ignore case and accents (DOS CP 850)

Append **:i=d437** after the sort field to ignore case and accents (DOS CP 437)

Append **:l** after the sort field to sort letters [A-Za-z] after other letters

GENERAL OPTIONS: can be appended before (or after) the sort fields.

The general options are valid for all sort fields.

/N Sort all sort fields numerically

/B Ignore leading/trailing blanks in all sort fields

/C Case insensitive (Ignore case for the letters [A-Za-z])

/C=u Case insensitive: Ignore case (letters [A-Za-z] and umlauts) (ANSI)

/C=uDOS Case insensitive: Ignore case (letters [A-Za-z] and umlauts) (DOS)

/I Ignore case and diacritic marks, break up umlauts and sharp S (ANSI)

/I=d Ignore case and diacritics, break up umlauts and sharp S (DOS CP 850)

/I=d437 Ignore case, diacritics, break up umlauts and sharp S (DOS CP 437)

/L Lines starting with letters [A-Za-z] will be sorted after other letters

/H=n Don't sort n header lines (default: n=0)

/D Ignore records with duplicate keys (according to sort fields)

/D=noSort Ignore records with duplicate keys, keep original order

/D=<file> Ignore records with duplicate keys, write duplicates to <file>

/M=n Use n megabytes as maximal size for temporary files (default: n=250)

/W=n Use Windows temp path for temp files (1, default) or current dir (0)

/V=n Verbosity level of screen output (default yes: n=1 or silent no: n=0)

EXAMPLES:

:: **example1.cmd:** Sort tab sep. file with one header row by the 1st field:

MGsort /H=1 /F1 names.txt names.sor

:: **example2.cmd:** Sort this file with header by the 6th, then the 2nd field:

MGsort /H=1 /F6 /F2 names.txt names.sor

:: **example3.cmd:** Sort this file with header by the 10th field numerically:

MGsort /H=1 /F10:n names.txt names.sor

:: **example4.cmd:** Sort this file with header by field last_name:

MGsort /H=1 /Fname=last_name names.txt names.sor

VERSION HISTORY:

Version 1.0.4.5 20200103

- Instead of a field number also a field name can be given, e.g. /Fname=last_name.
- All duplicates are written to the file rather than only second and further ones.

Version 1.0.3.0 20171019

- Option /M=n Use n megabytes as maximal size for temporary files (default n=250).
- The old option /M=n used lines: Use n lines as memory buffer (default n=100000).

Version 1.0.2.2 20171006

- Option /D=noSort added: Ignore records with duplicate keys, keep original order.
- Option /V=n added: Verbosity level of screen output (default yes n=1 or no n=0).

Version 1.0.1.0 20170514

- Use the Windows PID number for creating temporary files.
- For temporary files use the Windows TEMP path as default.
- This behaviour can be controlled with the new /W option.
- Added some examples.

Version 1.0.0.0 20170421

- Made the sort stable without any issues when merging.
- Increased the default sort buffer from 30.000 to 100.000 for use with modern PCs.
- Fixed minor issues related to spaces treated the same as no spaces in some cases.
- Simplified the merging process.

Version 0.7.5.0 20051111

- This version has worked for many years and is no beta version.