Title

stata.com

string — String manipulation functions

Contents Description Remarks and examples Also see

Contents

Editing, *continued*

strupper( )	`strupper()`	convert to uppercase
	`strlower()`	convert to lowercase
	`strproper()`	convert to proper case
ustrupper( )	`ustrupper()`	convert Unicode characters to uppercase
	`ustrlower()`	convert Unicode characters to lowercase
	`ustrtitle()`	convert Unicode characters to titlecase
strtrim( )	`stritrim()`	replace multiple, consecutive internal blanks with one blank
	`strltrim()`	remove leading blanks
	`strrtrim()`	remove trailing blanks
	`strtrim()`	remove leading and trailing blanks
ustrtrim( )	`ustrtrim()`	remove leading and trailing Unicode whitespace characters and blanks
	`ustrltrim()`	remove leading Unicode whitespace characters and blanks
	`ustrrtrim()`	remove trailing Unicode whitespace characters and blanks
subinstr( )	`subinstr()`	substitute text
	`subinword()`	substitute word
usubinstr( )	`usubinstr()`	replace Unicode substring
_substr( )	`_substr()`	substitute into string
_usubstr( )	`_usubstr()`	substitute into Unicode string
strdup( )	`*`	duplicate string
strreverse( )	`strreverse()`	reverse string in bytes
ustrreverse( )	`ustrreverse()`	reverse string in Unicode characters
soundex( )	`soundex()`	convert to soundex code
	`soundex_nara()`	convert to U.S. Census soundex code

Stata

abbrev( )	`abbrev()`	abbreviate Unicode strings to display columns
strtoname( )	`strtoname()`	translate strings to Stata 13 compatible names
ustrtoname( )	`ustrtoname()`	translate Unicode strings to Stata names

Text translation

strofreal( )	`strofreal()`	convert real to string
strtoreal( )	`strtoreal()`	convert string to real
ustrto( )	`ustrto()`	convert a Unicode string to a string in another encoding
	`ustrfrom()`	convert a string in one encoding to a Unicode string
ustrunescape( )	`ustrunescape()`	convert the escaped hex sequences to Unicode
	`ustrtohex()`	convert a Unicode sequence to hex sequences
ascii( )	`ascii()`	obtain ASCII or byte codes of string
	`char()`	make string from ASCII or byte codes
uchar( )	`uchar()`	make Unicode character from Unicode code-point value

Unicode utilities

ustrcompare( )	`ustrcompare()`	compare or sort Unicode strings
	`ustrsortkey()`	obtain sort key of Unicode string
ustrfix( )	`ustrfix()`	replace invalid sequences in Unicode string
ustrnormalize( )	`ustrnormalize()`	normalize Unicode string

Description

The above functions are for manipulating strings. Strings in Mata are strings of Unicode characters in UTF-8 encoding, usually the printable characters, but Mata enforces no such restriction. In particular, strings may contain binary 0.

Remarks and examples

stata.com

In addition to the above functions, two operators are especially useful for dealing with strings.

The first is +. Addition is how you concatenate strings:

```
: "abc" + "def"
abcdef
: "Café " + "de Flore"
Café de Flore
: command = "list"
: args = "mpg weight"
: result = command + " " + args
: result
list mpg weight
```

The second is *. Multiplication is how you duplicate strings:

```
: 5*"a"
aaaaa
: "Allô"*2
AllôAllô
: indent = 20
: title = indent*" " + "My Title"
: title
                     My Title
```

Also see

[M-4] **intro** — Index and guide to functions