

ustrpos() — Find substring in Unicode string

[Description](#)
[Diagnostics](#)

[Syntax](#)
[Also see](#)

[Remarks and examples](#)

[Conformability](#)

Description

`ustrpos(s, sf [, n])` returns the character position in *s* at which *sf* is first found; otherwise, it returns 0. If *n* is specified and is larger than zero, the search starts at the *n*th Unicode character of *s*.

`ustrrpos(s, sf [, n])` returns the position in *s* at which *sf* is last found; otherwise, it returns 0. If *n* is specified and is larger than zero, the part between the first Unicode character and the *n*th Unicode character of *s* is searched.

When *s* is not a scalar, these functions return element-by-element results.

Syntax

real matrix `ustrpos(string matrix s, string scalar sf [, real scalar n])`

real matrix `ustrrpos(string matrix s, string scalar sf [, real scalar n])`

Remarks and examples

[stata.com](#)

An invalid UTF-8 sequence in *s* or *sf* is replaced with a Unicode replacement character `\ufffd` before the search is performed.

Use `strpos()` or `strrpos()` to find the byte-based location of a substring in a string.

Conformability

`ustrpos(s, sf [, n])`, `ustrrpos(s, sf [, n])`:

<i>s</i> :	$r \times c$
<i>sf</i> :	1×1
<i>n</i> :	1×1
<i>result</i> :	$r \times c$

Diagnostics

None.

Also see

[\[M-5\] strpos\(\)](#) — Find substring in string

[\[M-4\] String](#) — String manipulation functions

[U] 12.4.2 Handling Unicode strings

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow and NetCourseNow are trademarks of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2023 StataCorp LLC, College Station, TX, USA. All rights reserved.



For suggested citations, see the FAQ on [citing Stata documentation](#).