Title

stata.com

ascii( ) — Manipulate ASCII and byte codes

Description Syntax Conformability Diagnostics Also see

Description

`ascii(`*s*`)` returns a row vector containing the ASCII codes (0–127) and byte codes (128–255) corresponding to *s*. For instance, `ascii("abc")` returns `(97, 98, 99)`; `ascii("café")` returns `(99, 97, 102, 195, 169)`. Note that the Unicode character "é" is beyond ASCII range. Its UTF-8 encoding requires 2 bytes and their byte values are 195 and 169.

`char(`*c*`)` returns a UTF-8 encoded string consisting of the specified ASCII and byte codes. For instance, `char((97, 98, 99))` returns `"abc"`, and `char((99, 97, 102, 195, 169))` returns `"café"`.

Syntax

real rowvector `ascii(`*string scalar s*`)`

string scalar `char(`*real rowvector c*`)`

Conformability

`ascii(`*s*`)`:
s: 1×1
result: $1 \times$ `strlen(`*s*`)`

`char(`*c*`)`:
c: $1 \times n, \quad n \geq 0$
result: 1×1

Diagnostics

`ascii(`*s*`)` returns `J(1,0,.)` if `strlen(`*s*`)==0`.

In `char(`*c*`)`, if any element of *c* is outside the range 0 to 255, the returned string is terminated at that point. For instance, `char((97,98,99,1000,97,98,99))=="abc"`.

`char(J(1,0,.))` returns `""`.

Also see

[M-5] **isascii( )** — Whether string scalar contains only ASCII codes

[M-5] **uchar( )** — Convert code point to Unicode character

[M-4] **String** — String manipulation functions