

⁺This command is part of [StataNow](#).

[Description](#)

[Remarks and examples](#)

[Quick start](#)

[Also see](#)

[Syntax](#)

[Options](#)

Description

docx2txt converts a Word (.docx) document to a plain text file.

Quick start

Convert Word document myfile.docx to a plain text file saved as myfile.txt

```
docx2txt myfile
```

Same as above, but save the plain text file as mytext.txt

```
docx2txt myfile, saving(mytext)
```

Same as above, and overwrite the existing mytext.txt

```
docx2txt myfile, saving(mytext) replace
```

Syntax

```
docx2txt srcfile [ , options ]
```

srcfile is a .docx file. If *srcfile* is specified without an extension, .docx is assumed. If *srcfile* contains embedded spaces or other special characters, enclose it in double quotes.

options

Description

saving(*targetfile*)

specify the target plain text file to be saved

replace

replace the target plain text file if it already exists

nomsg

suppress message with link to *targetfile*

Options

saving(*targetfile*) specifies the target plain text file to be saved. If *targetfile* is specified without an extension, .txt is assumed. If *targetfile* contains embedded spaces or other special characters, enclose it in double quotes. If saving() is not specified, the target filename is constructed using the source filename (*srcfile*) with the .txt extension.

replace specifies that the target plain text file be replaced if it already exists.

nomsg suppresses the message that contains a link to the target file.

Remarks and examples

docx2txt converts Word (.docx) documents to plain text files. The graphs in the Word document are ignored.

► Example 1: Convert Word document to a plain text file

To demonstrate how to convert a Word (.docx) document to a plain text file, we begin by creating a .docx file with Stata results. We use the putdocx suite of commands to create a file with a table of descriptive statistics and a table of regression results. The `dtable` command creates a table of descriptive statistics, and `putdocx collect` exports the table to the .docx file. Then we fit a regression model with `regress` and export the table with `putdocx table` by using the `etable` syntax. You can learn more about the putdocx suite of commands in [RPT] [putdocx intro](#).

```
. use https://www.stata.press.com/data/r19/nhanes21
(Second National Health and Nutrition Examination Survey)

. putdocx begin
. putdocx paragraph
. putdocx text ("We analyze data from the Second National Health and")
. putdocx text (" Nutrition Examination Survey."), linebreak(1)
. quietly: dtable bpsystol weight bmi, by(diabetes)
> title("Table 1. Descriptive statistics")
> column(by(, halign(right)) total(, halign(right)))
. putdocx collect
(collection DTable posted to putdocx)
. regress bpsystol age weight
```

Source	SS	df	MS	Number of obs	=	10,351
Model	1707705.53	2	853852.766	F(2, 10348)	=	2250.00
Residual	3926964.49	10,348	379.490191	Prob > F	=	0.0000
Total	5634670.03	10,350	544.412563	R-squared	=	0.3031
				Adj R-squared	=	0.3029
				Root MSE	=	19.481

bpsystol	Coefficient	Std. err.	t	P> t	[95% conf. interval]
age	.6379892	.0111315	57.31	0.000	.6161692 .6598091
weight	.4069041	.0124786	32.61	0.000	.3824435 .4313646
_cons	71.27096	1.041742	68.42	0.000	69.22894 73.31297

```
. putdocx table bweight = etable,
> title("Table 2. Linear regression of systolic blood pressure")
. putdocx save bpreport, replace
successfully created "C:/mypath/bpreport.docx"
```

To convert this file to a plain text file, we type

```
. docx2txt bpreport.docx
```

This creates bpreport.txt, shown below:

```
Do-file Editor - bpreport.txt
File Edit View Language Project Tools

bpreport.txt x
1 We analyze data from the Second National Health and Nutrition Examination Survey.
2
3 Table 1. Descriptive statistics
4
5
6
7
8
9
10
11
12
13
14
15
16
17
```

	Diabetes status		
	Not diabetic	Diabetic	Total
N	9,890 (95.2%)	499 (4.8%)	10,349 (100.0%)
Systolic blood pressure	130.008 (22.759)	146.651 (28.387)	130.887 (23.332)
Weight (kg)	71.658 (15.228)	76.670 (17.175)	71.900 (15.357)
Body mass index (BMI)	25.413 (4.832)	27.996 (5.818)	25.538 (4.915)


```
Table 2. Linear regression of systolic blood pressure
bbsystol Coefficient Std. err. t P>|t| [95% conf. interval]
age .6379892 .0111315 57.31 0.000 .6161692 .6598091
weight .4069041 .0124786 32.61 0.000 .3824435 .4313646
_cons 71.27096 1.041742 68.42 0.000 69.22894 73.31297
```



Also see

[RPT] [docx2epub](#) — Convert a Word (.docx) document to an EPUB file⁺

[RPT] [docx2html](#) — Convert a Word (.docx) document to an HTML file⁺

[RPT] [docx2markdown](#) — Convert a Word (.docx) document to a Markdown file⁺

[RPT] [docx2pdf](#) — Convert a Word (.docx) document to a PDF file

Stata, Stata Press, Mata, NetCourse, and NetCourseNow are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow is a trademark of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2025 StataCorp LLC, College Station, TX, USA. All rights reserved.



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