

2.8 Importing OpenClinica Data Into R

R is widely used open source statistics software. You can obtain the R software for free at http://www.r-project.org/.

There are a few ways to import your OpenClinica data into R. This document provides instructions for importing data into R in three different ways:

- using a Tab-delimited file,
- using a CSV file, and
- using an Excel file.

Importing Tab-delimited data into R

Importing a Tab-delimited file into R is the easiest file type to import. To import our data, your must first change the directory so that R knows where your file is located. To do this, go to File in the menu bar and select Change dir and open the folder where your Tab-delimited file is saved.



You can now import your data using the read.delim function. To do this, type the command: read.delim(file.tsv)

into the R console, where file.tsv is the name of your Tab-delimited file. The example below uses

read.delim(Tab_Example.tsv)



Now the import will run. After your data has been imported the last line may say reached getOption("max.print") indicating how many rows were omitted. This only means that your data was too large to display in the console--but your data was all imported successfully.

Red (32-bit)	
File Edit View Misc Packages Windows Help	
	_
R Console	
33311 \$	^
33312 \$	
33313 \$	
33314 \$	
\$33315	
33316 Hospital General Universitario V\$	
33317 194\$	
33318	
33319	
33320	
33321	
33322	
33324	
33325	
33326 201\$	
33327 \$	
33328 \$	
33329	
\$3330	
33331 \$	
33332 \$	
\$3333	
[reached getOption("max.print") omitted 1284 rows]	
	Ŧ
< III >	.H

Importing CSV data into R

To import a CSV file into R you must first convert your OpenClinica Excel file to a CSV file. First, open your Excel file. When doing this you may see an error message pop-up (as shown below). Click Yes if you see this message.

(· 9 · (· -) ·	100	Microsoft Excel no	-commercial use	-
Home Insert	Page Layout Formulas Data Review V	view Load Test Team		
Paste J Format Painter	・ A A A E = = (※)・ B I 型・(①・A・) 目 = = (※)・ Foot 5 Abo	₩Vap Text Wap Ex Center ~ S ~ % , %	Conditional Format Cell Formatting - as Table - Styles - Styles	Insert Delete Format Σ AutoSum Cells Q Clear
-		printers - Horizon	- J Styles	
Microsoft Office Exce The file yo want to op Was this in	are trying to open, 'Example.x6', is in a different format than s n the file now? ormation helpfu?	peofied by the file extension. Verify that the file is n	ot compted and is from a trusted source l	efore opening the file. Do you

Next, save the file as a CSV file by going to Save As and clicking Other Formats. In the Save as type: dropdown box select CSV (Comma delimited) (*.csv).



You may get a pop-up message when you save your file. If so, click Yes.

	9	(° •) •	;						(Annual loss		CSV_Exan	nple.csv - Mic	rosoft
9	Home	Insert	Page L	ayout For	mulas D	ata Re	eview V	ew Loa	d Test Te	am			
B	🔏 Cut		Calibri	* 11	* A* *	= =	= »-	📑 Wrap	p Text	General	٠	5	
Paste	I Form	at Painter	BI	<u>n</u> - 🖽 -	👌 - <u>A</u> -	≣≣	≡ 详 ()	Merg	ge & Center 😁	\$ - %	· · · · · · · · · · · · · · · · · · ·	Conditional Formatting	Fori as Ta
	Clipboard	5		Font	6		Alig	nment	6	Nu	mber 🕞		
		-		<i>f</i> _x Datas	et Name:								
	А	В	С	D	E	F	G	н	1	J	K	L	М
1 D	Microsoft	Office Exc	el	1000	-	-	-					X	
2 D				_	_	_	_						
4 51		CSV_Exa	mple.csv m	ay contain feat	ures that are r	not compatib	e with CSV (Comma delimi	ited). Do you v	vant to keep	the workbook in	this format?	
5 Pr		To keep To press	p this forma	t, which leaves	out any incom	patible feat	ures, dick Ye	ormat					
6 D.		 To see 	what might	be lost, click He	slp.	copy in the	AND A CAUCIT	on this is a					
7 St					→ 🔽	es	No	Help					
8 St	L		_					_					J

Now to import your CSV file, you must first change the directory so that R knows where your file is located. To do this, go to File in the menu bar, select Change dir and open the folder where your CSV file is saved.



You can now import your data using Rs read.csv function. To do this, type:

read.csv(file.csv)

into the R console, where file.csv is youre the name of your CSV file. The example below uses:

read.csv(CSV_Example.csv)

```
RGui (32-bit)
File Edit View Misc Packages Windows Help
🛎 🖆 🖬 🖻 😳 🐵 🎒
 R Console
                                                                        - • ×
 R version 2.15.3 (2013-03-01) -- "Security Blanket"
 Copyright (C) 2013 The R Foundation for Statistical Computing
 ISBN 3-900051-07-0
 Platform: i386-w64-mingw32/i386 (32-bit)
 R is free software and comes with ABSOLUTELY NO WARRANTY.
 You are welcome to redistribute it under certain conditions.
 Type 'license()' or 'licence()' for distribution details.
  Natural language support but running in an English locale
 R is a collaborative project with many contributors.
 Type 'contributors()' for more information and
 'citation()' on how to cite R or R packages in publications.
 Type 'demo()' for some demos, 'help()' for on-line help, or
 'help.start()' for an HTML browser interface to help.
 Type 'q()' to quit R.
 > read.csv("CSV_Example.csv")
```

After your data has been imported the last line may say reached getOption("max.print") indicating how many rows were omitted. Dont worry, this only means that your data was too large to display in the console--your data was imported successfully.

RGui (32-bit)	
File Edit View Misc Packages Windows Help	
R Console	- • •
265 NA	^
266 NA	
267 NA	
268 NA	
269 NA	
270 NA	
271 NA	
272 NA	
273 NA	
274 NA	
275 NA	
270 NA	
277 NA 279 NA	
270 NA 279 NA	
280 NA	
281 NA	
282 NA	
283 NA	
284 NA	
285 NA	
286 NA	
287 NA	
[reached getOption("max.print") omitted 547 rows]	
>	
4	
	·

Importing Excel data into R

Step 1: Installing the RODBC package

To import your Excel data into R you need to download the RODBC package through R (this step is only necessary the first time you use R). The example below illustrates this with the 32-bit version of R (the 64-bit version requires a different process). For more detailed information about using the RODBC package, see http://cran.r-project.org/web/packages/RODBC/RODBC.pdf. Open R and go to the Packages tab of the menu bar. Click Install package(s)



Select a location to download the package from. The example below selected USA (PA 1) since it represented the closest available location. Then, select RODBC from the list of packages that you can download.



After the download completes, you should see the following message in the R console:



Step 2: Loading the RODBC package

Now that you have installed the RODBC package, it is time to load it into R. Note: this step is necessary every time that you use R. To install the RODBC package, go back to the Packages tab in the menu bar and click Load package



A list of packages should then pop up and you can select RODBC. You should not receive any other messages after you load the package.



Step 3: Importing your Excel data into R

Now that you have R installed, as well as the RODBC package installed and loaded, you are almost ready to import our data into R. First open your Excel file. When you do this you will see an error message pop up. Click Yes when you see this message.

(· · · · []	-	1 1000	Microsoft Excel non-commercial use	-
Home Insert	Page Layout Formulas Da	ta Review View Load Tes	t Team	
Paste J Format Painter	· · A · x · B · Z · U · O · A ·	■ = = ≫··	Center - S - % , S - % Conditional Format Cell Formatting - as Table - Styles -	Insert Delete Format ↓ Clear
Clipboard 9	Font 5	Alignment	6 Number 6 Styles	Cells Ed
Microsoft Office Excel	u are trying to open, "Example.xid", is in a en the file now? formation helpful?	different format than specified by the file of Yes	extension. Verify that the file is not corrupted and is from a trusted source	before opening the file. Do you
		1		

Next save the file as an Excel 97-2003 Workbook. If you overlook this step you will see an error when trying to import your data to R that says "External table is not in the expected format.

<u>()</u> 🖬 🔍 🔊	•) •			Examp	olexis - N
New	Save a copy of the document	Wrap Text	General		
Dpen	Save the file as an Excel Workbook.	律律 Merge & Co	enter • \$ • %	• • • • • • • • • • • • • • • • • • •	Condi Forma
Save	Excel Macro-Enabled Workbook Save the workbook in the XML-based and macro-enabled file format.	Angrimerit		iibei 🦻	<u> </u>
Save As 🔸	Excel Binary Workbook Save the workbook in a binary file format optimized for fast loading and saving.	G H lete_08FEB2013 :te_08FEB2013	I J	К	L
Print >	Excel 27-2003 Workbook Save a copy of the workbook that is fully compatible with Excel 97-2003.	SHEMMES			
Prepare >	OpenDocument Spreadsheet Save the workbook in the Open Document Format.				
Sen <u>d</u> •	PDF or XPS Publish a copy of the workbook as a PDF or XPS file.				
Close	Open the Save As dialog box to select from all possible file types.				
15 Study Sub Prot	Excel Options X Exit Excel	RAND GR EPID ANE EPI	D ANE ANTIB PE	INDCT THE	NDCT T

Now that your data is formatted correctly you are ready to import it into R. First, change the directory so that R knows where the file is located. To do this, go to File in the menu bar, select Change dir, and open the folder where your Excel file is saved.

R R	Gui (32-bit)						
File	Edit View Misc Pack	ages Windows Help					
	Source R code	6					
	New script						
	Open script						
	Display file(s)	^					
	Load Workspace	port but running in an English locale					
	Save Workspace	roject with many contributors.					
	Load History	for more information and					
	Save History	order of a packages in publications.					
	Save History	demos, 'help()' for on-line help, or					
	Change dir 🧲 🗕	HTML browser interface to help.					
	Print						
	Save to File	kgs ()					
		AN mirror for use in this session					
	Exit	ion/zip' length 759673 bytes (741 Kb)					
op	ened URL	E					
do	wnloaded 741 Kb						
pa	package 'RODBC' successfully unpacked and MD5 sums checked						
-	-						
Th	e downloaded binary	<pre>/ packages are in // ackages are in // acka</pre>					
> :	local({pkg <- selec	t.list(sort(.packages(all.available = TRUE)), graphics=TRUE)					
+	if(nchar(pkg)) libr	<pre>ary(pkg, character.only=TRUE)))</pre>					
>		•					
•		m →					

You can now import your data using the odbcConnectExcel function provided to us by the RODBC package. To do this, type:

odbcConnectExcel(file.xls)

into the R console and where file.xls represents the name of your Excel file. The example below uses:

odbcConnectExcel(Excel Example.xls).

After you enter this function you should see a message like the one displayed in the picture below. If you receive this message you have successfully imported your Excel file into R!



Approved for publication by Ben Baumann. Signed on 2014-03-24 8:25AM

Not valid unless obtained from the OpenClinica document management system on the day of use.