4.8 Importing OpenClinica Data Into R

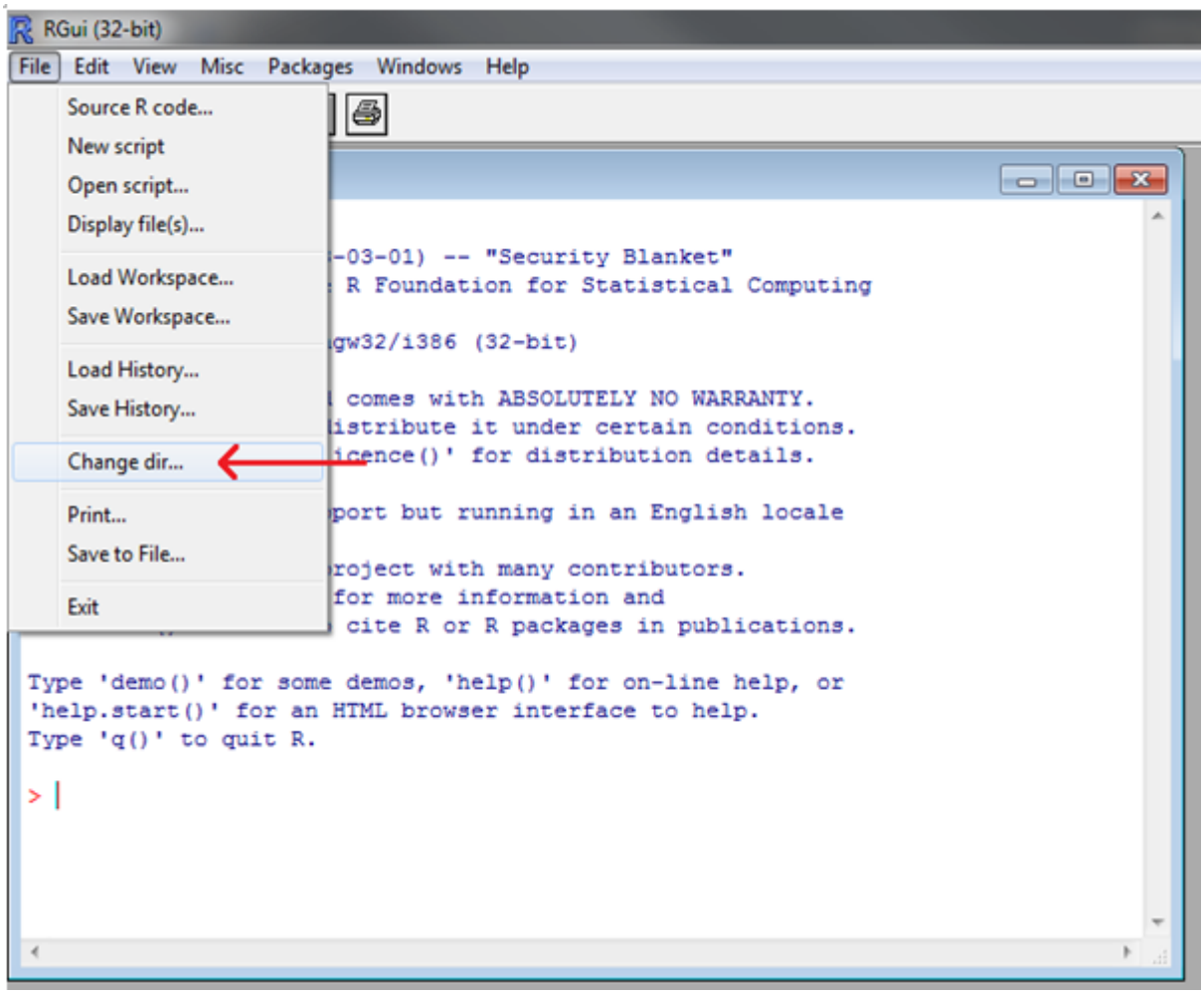
R is widely used open source statistics software. You can obtain the R software for free at http://www.r-project.org/.

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.

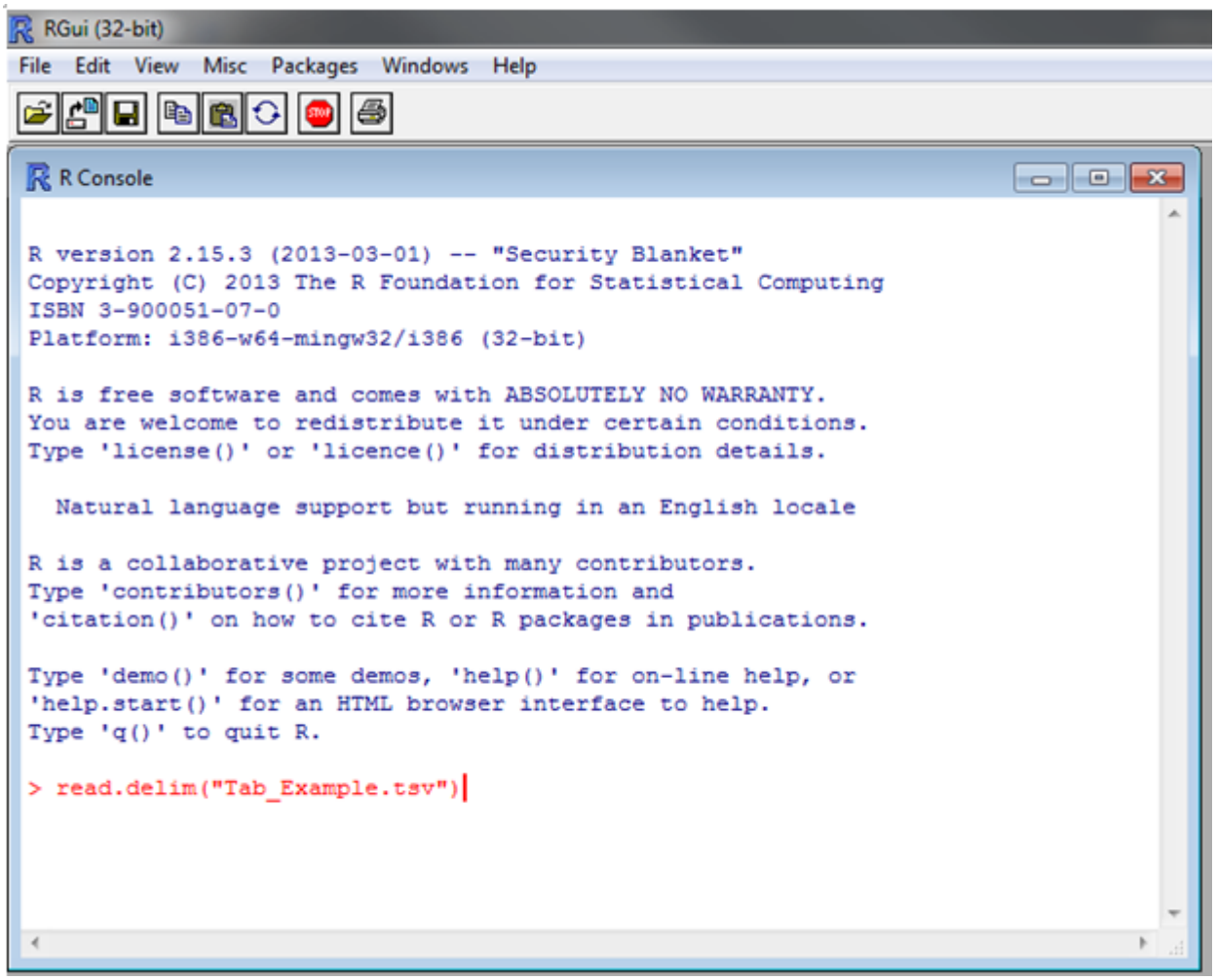
Importing Tab-delimited data into R

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.

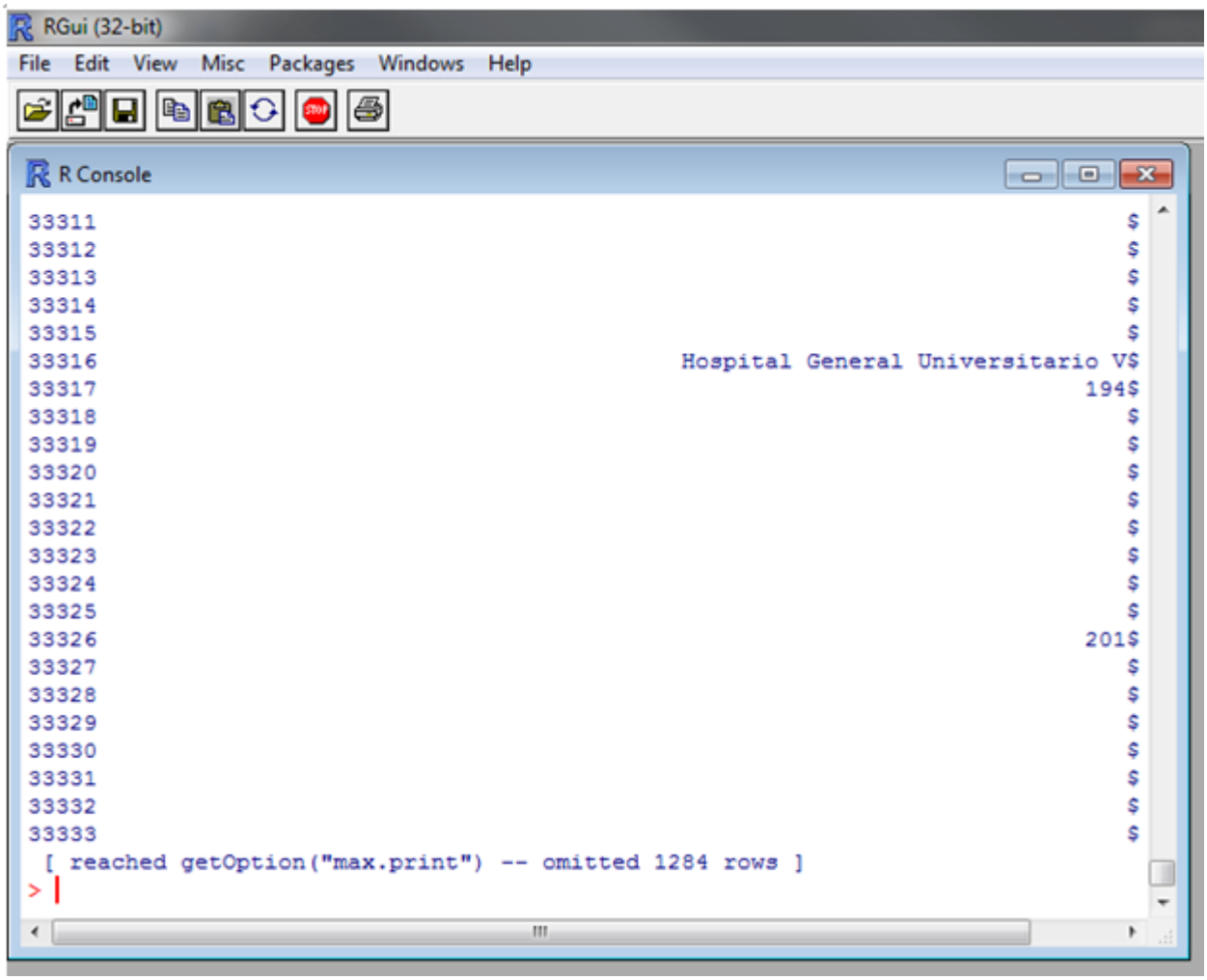


You can now import your data using the read.delim function. To do this, type the command:
read.delim(file.tsv)
into the R console, where file.tsv is the name of your Tab-delimited file. The example below uses

read.delim(Tab_Example.tsv)

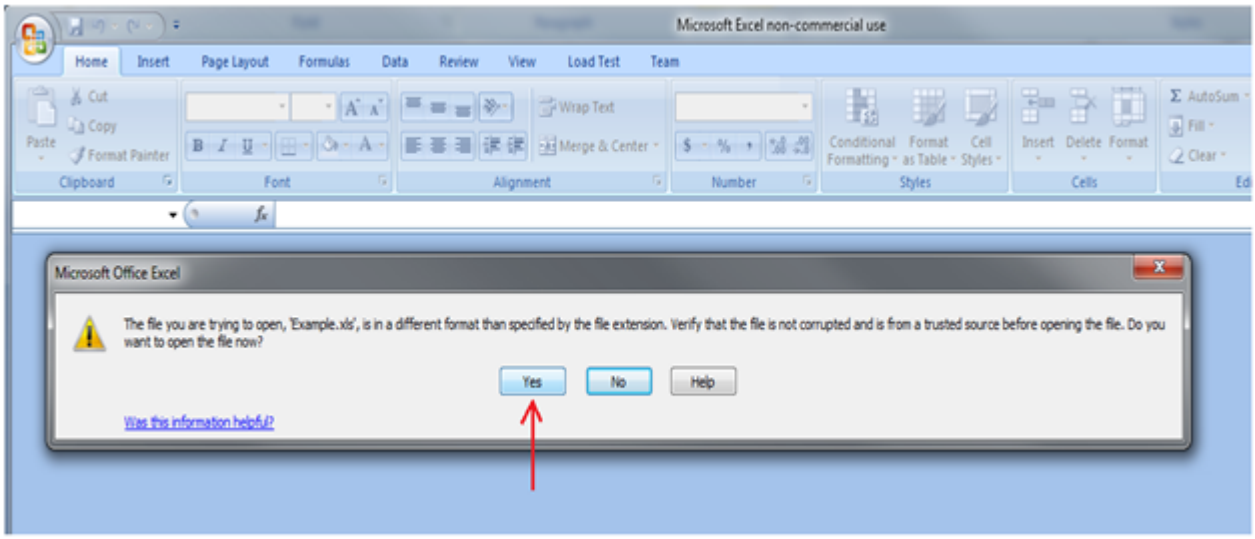


Now the import will run. After your data has been imported the last line may say reached getOption("max.print") indicating how many rows were omitted. This only means that your data was too large to display in the console--but your data was all imported successfully.

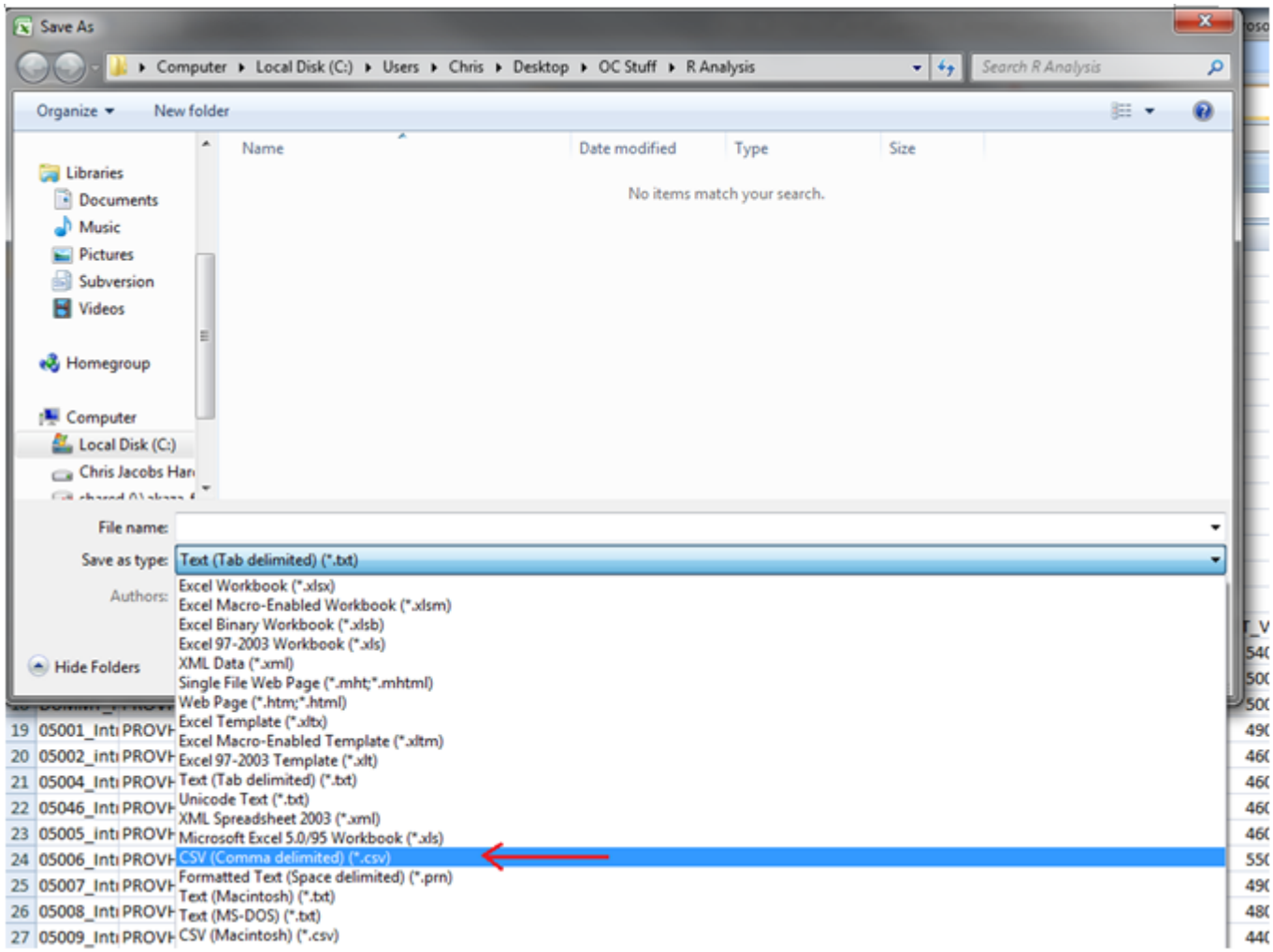


Importing CSV data into R

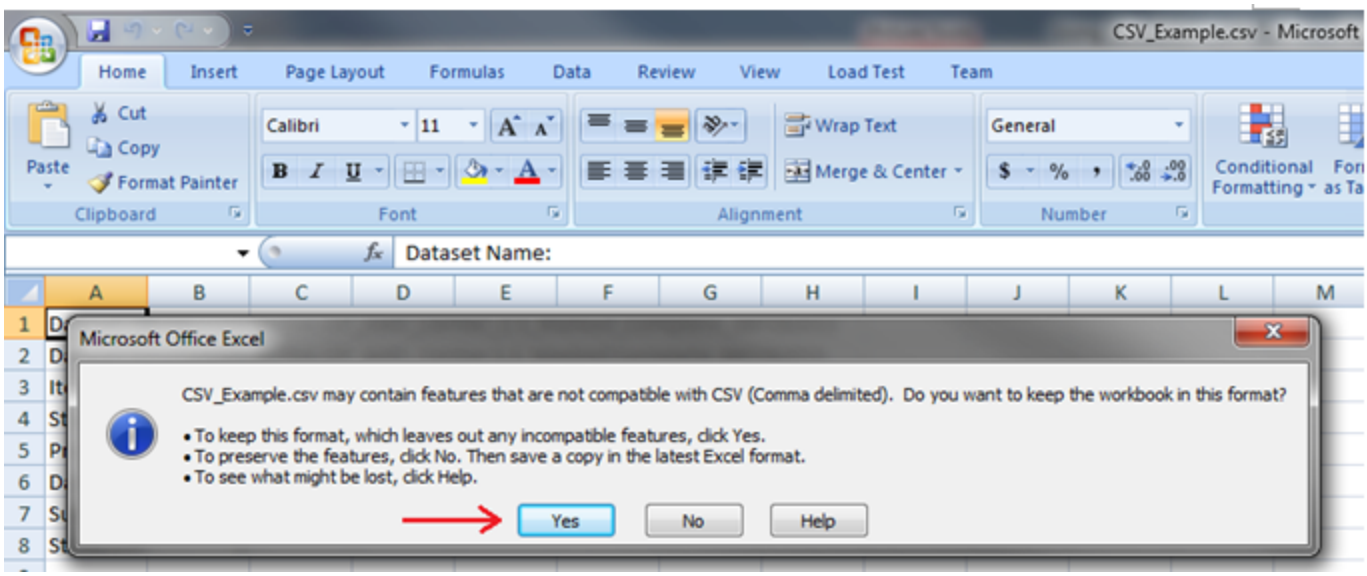
To import a CSV file into R you must first convert your OpenClinica Excel file to a CSV file. First, open your Excel file. When doing this you may see an error message pop-up (as shown below). Click Yes if you see this message.



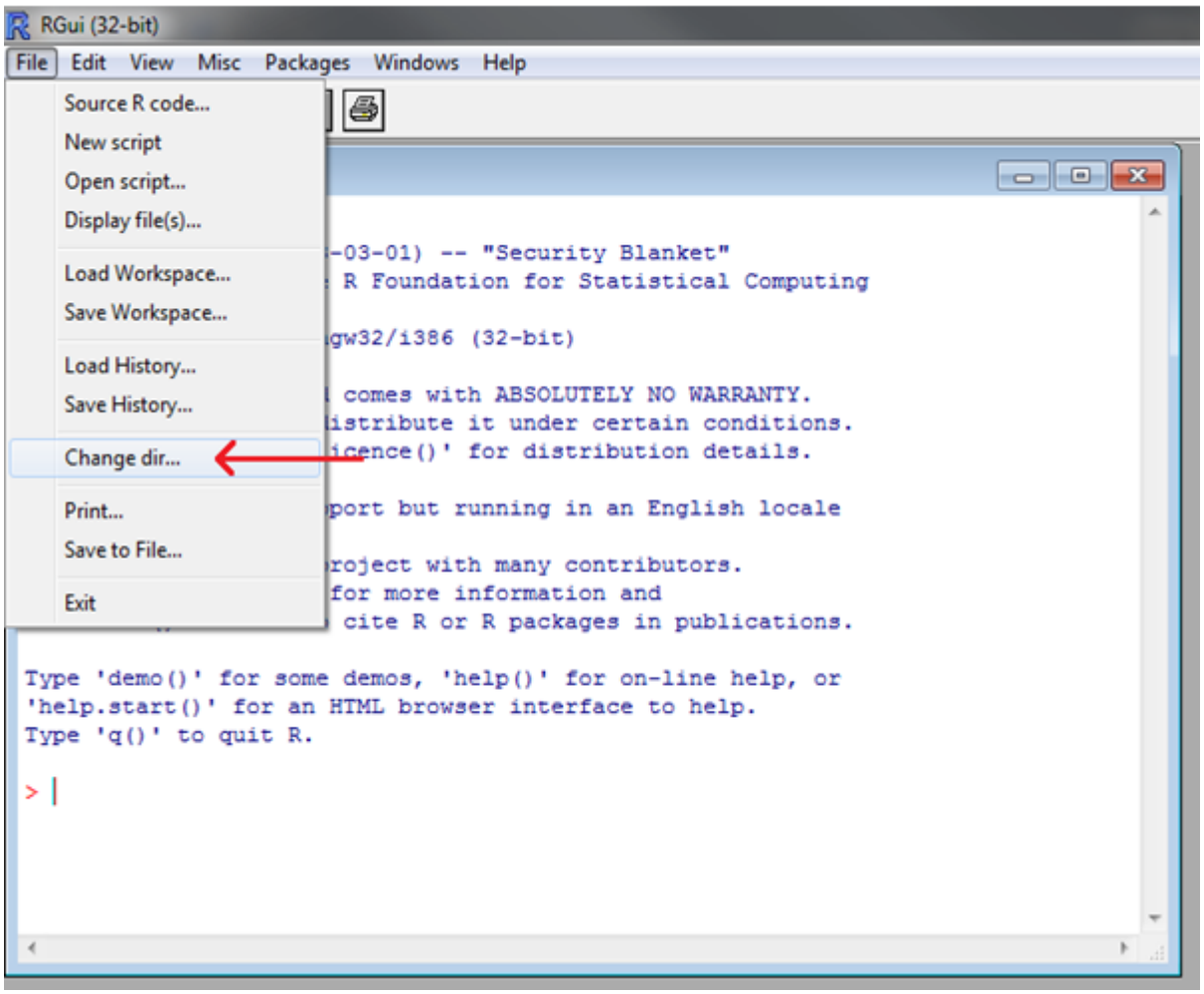
Next, save the file as a CSV file by going to Save As and clicking Other Formats. In the Save as type: dropdown box select CSV (Comma delimited) (*.csv).



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



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.

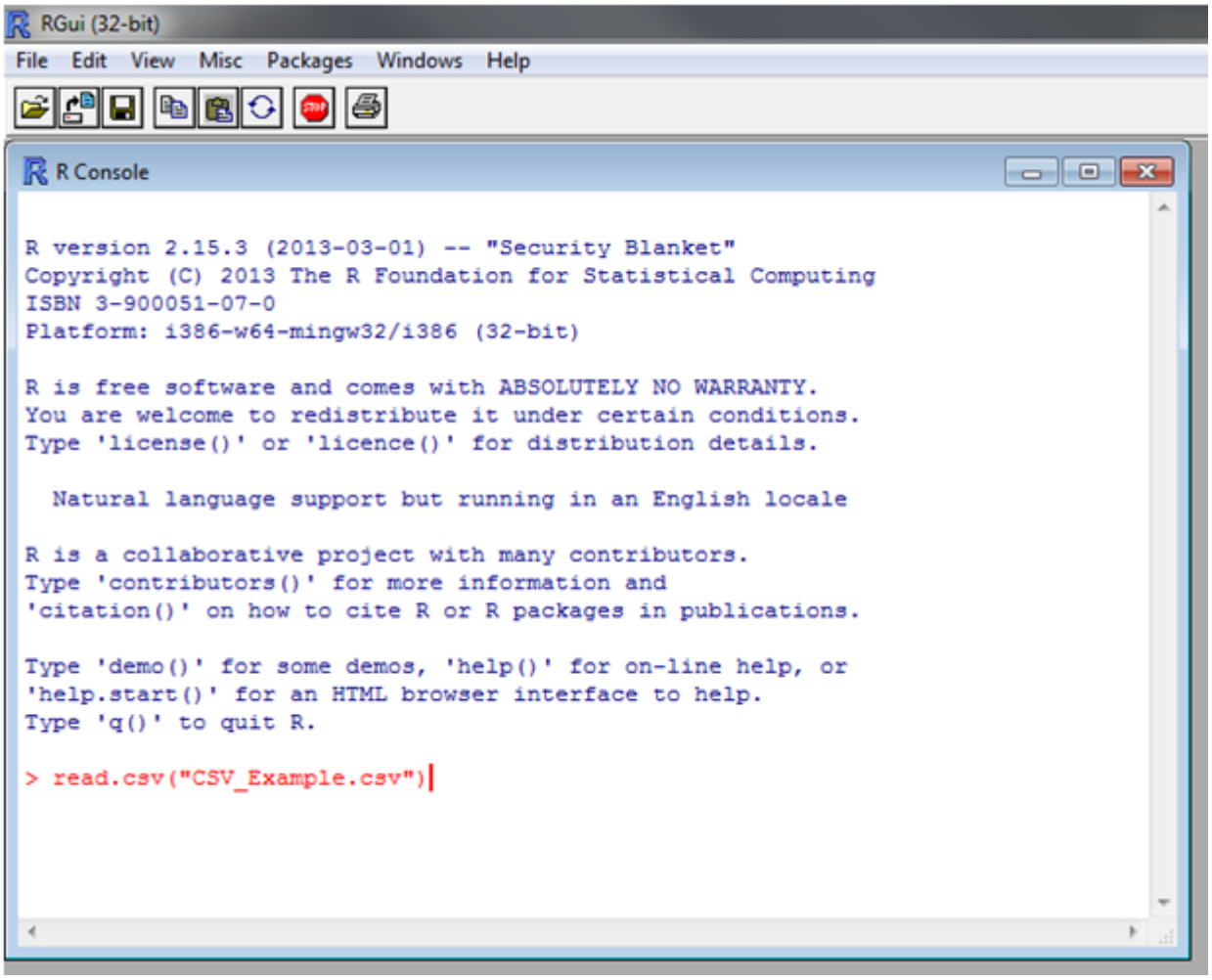


You can now import your data using Rs read.csv function. To do this, type:

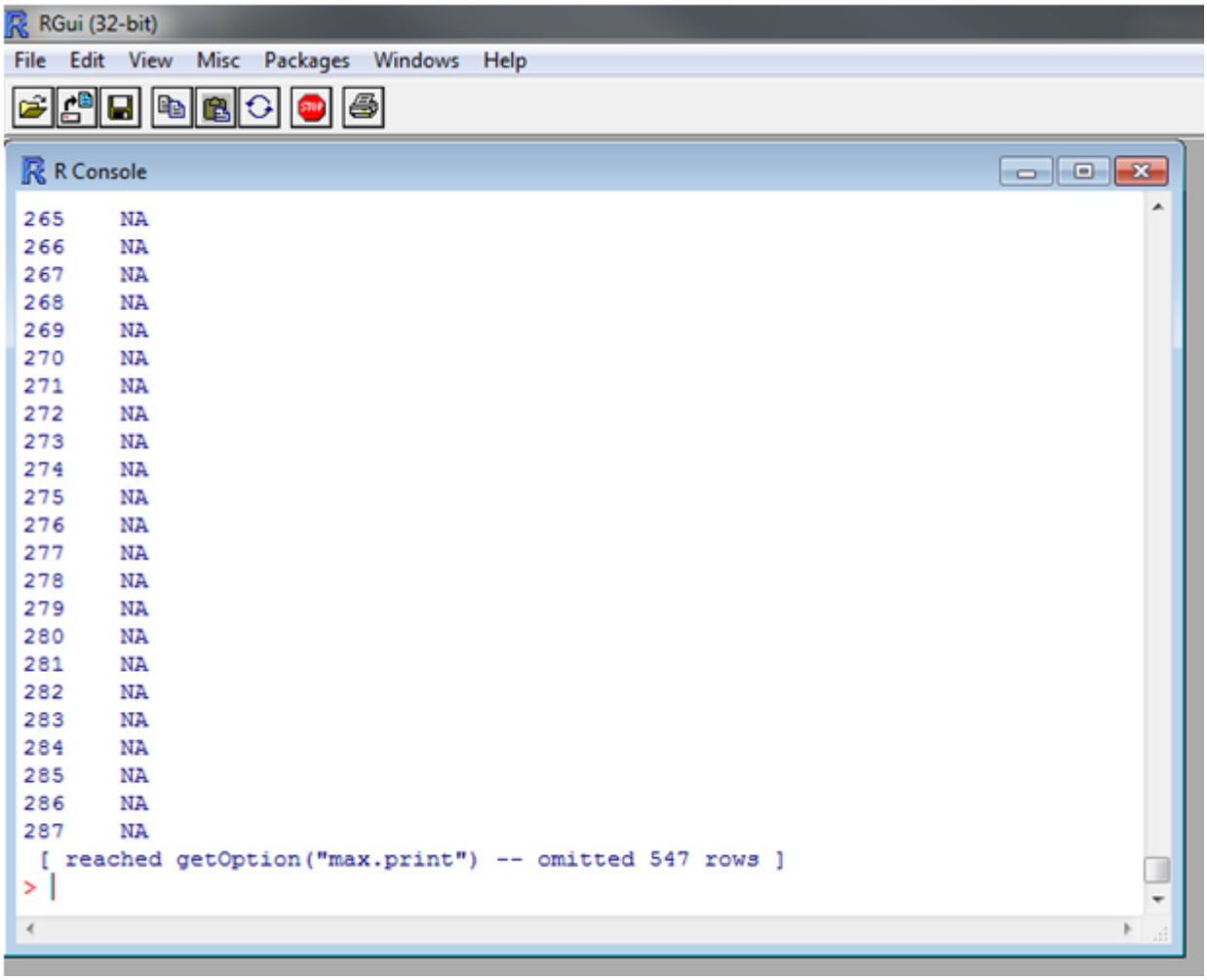
read.csv(file.csv)

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

read.csv(CSV_Example.csv)



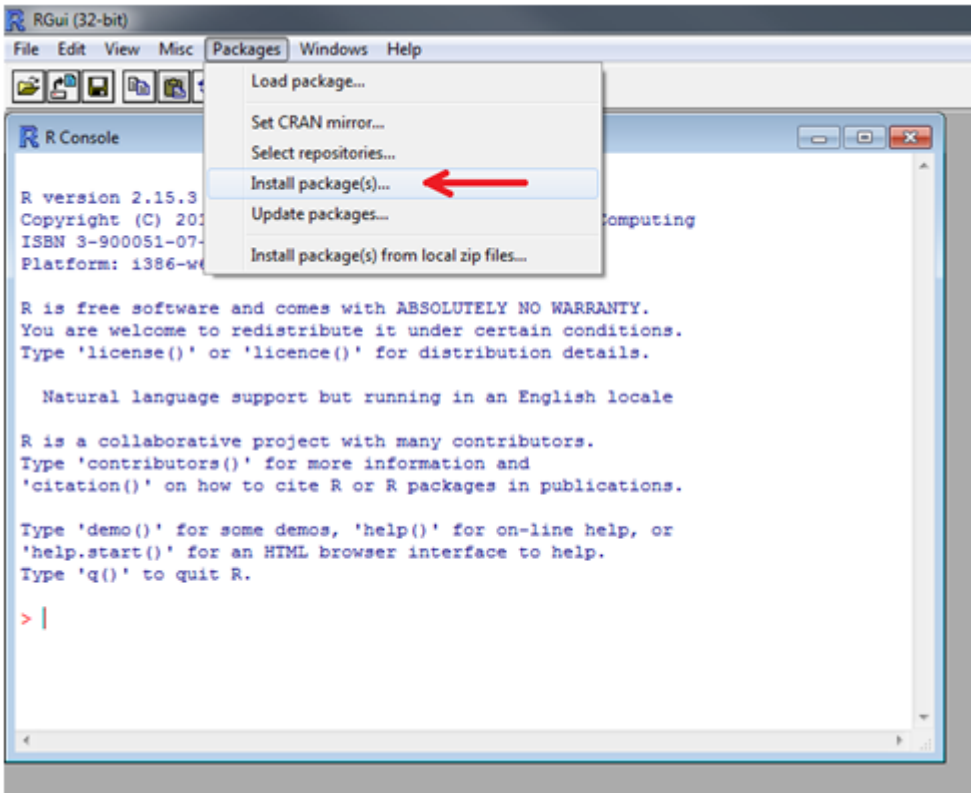
After your data has been imported the last line may say reached getOption("max.print") indicating how many rows were omitted. Dont worry, this only means that your data was too large to display in the console--your data was imported successfully.



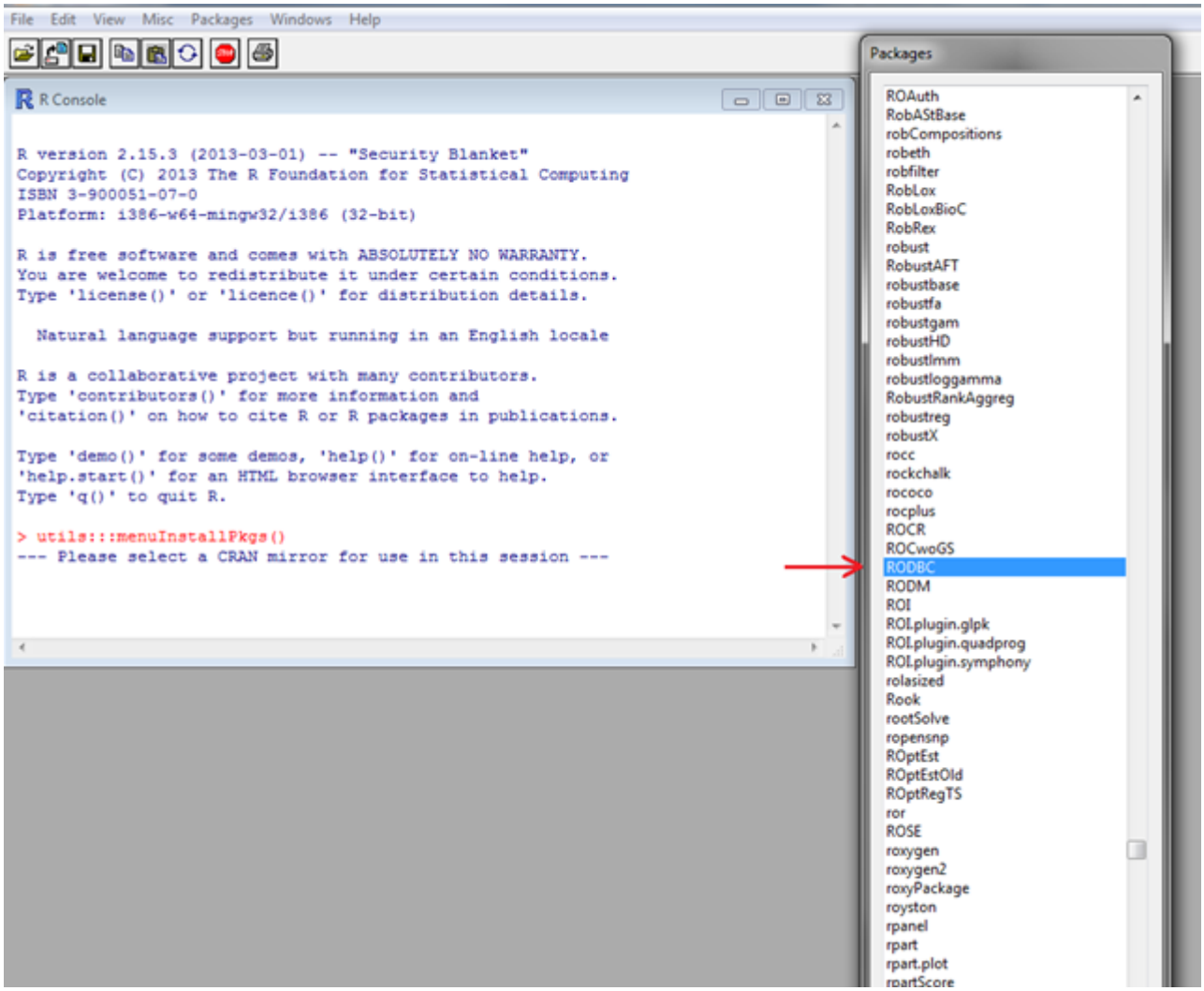
Importing Excel data into R

Step 1: Installing the RODBC package

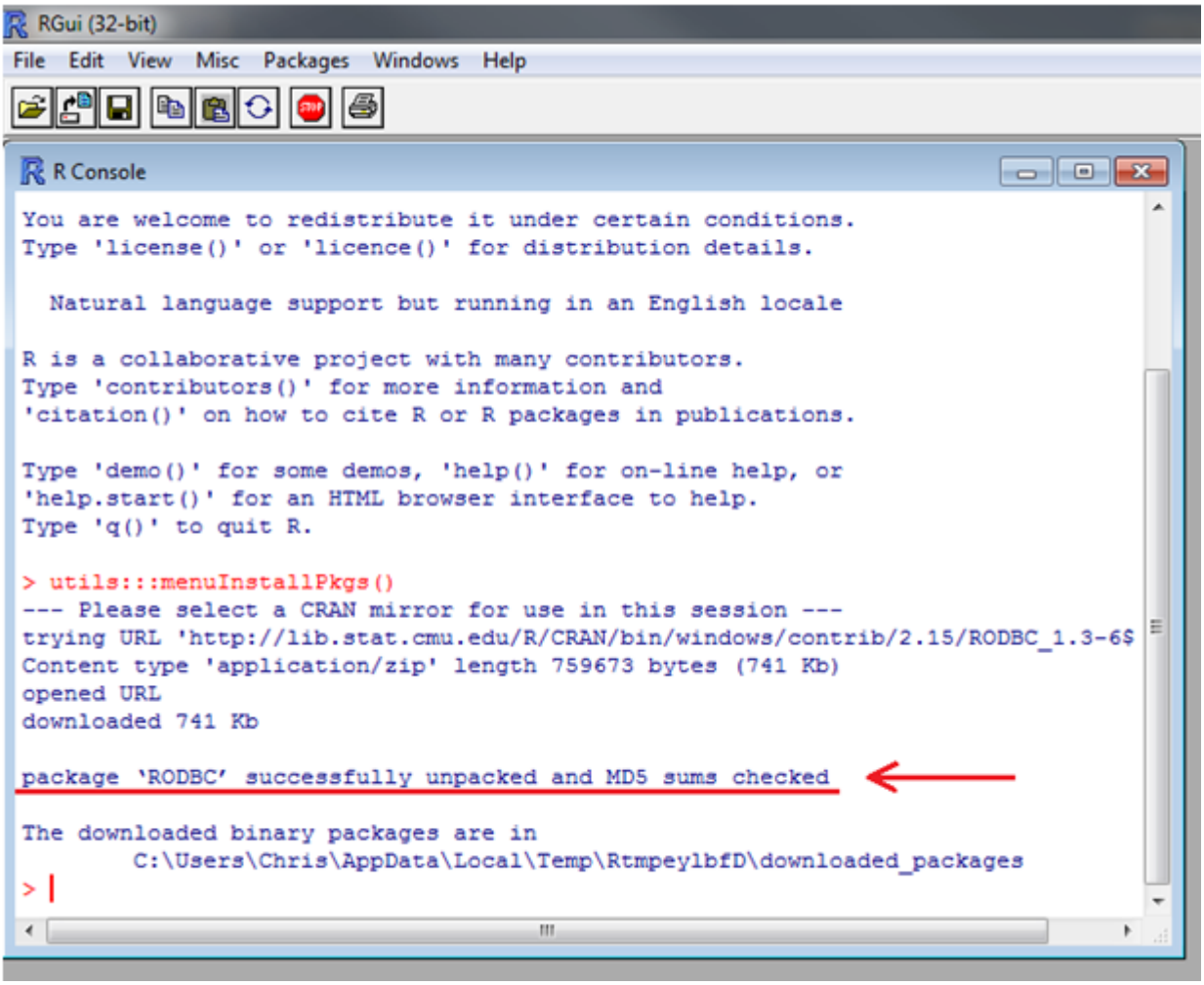
To import your Excel data into R you need to download the RODBC package through R (this step is only necessary the first time you use R). The example below illustrates this with the 32-bit version of R (the 64-bit version requires a different process). For more detailed information about using the RODBC 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.

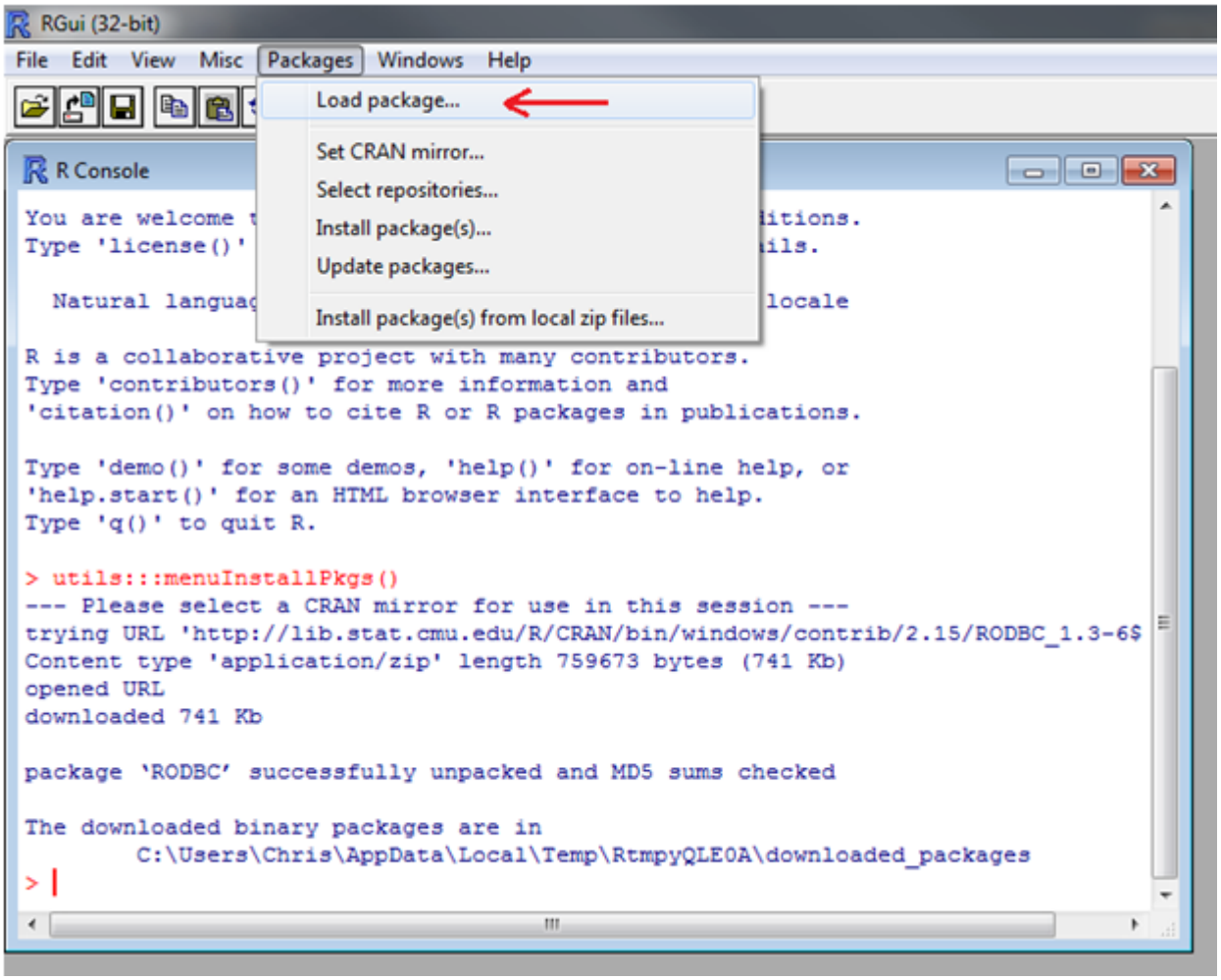


After the download completes, you should see the following message in the R console:

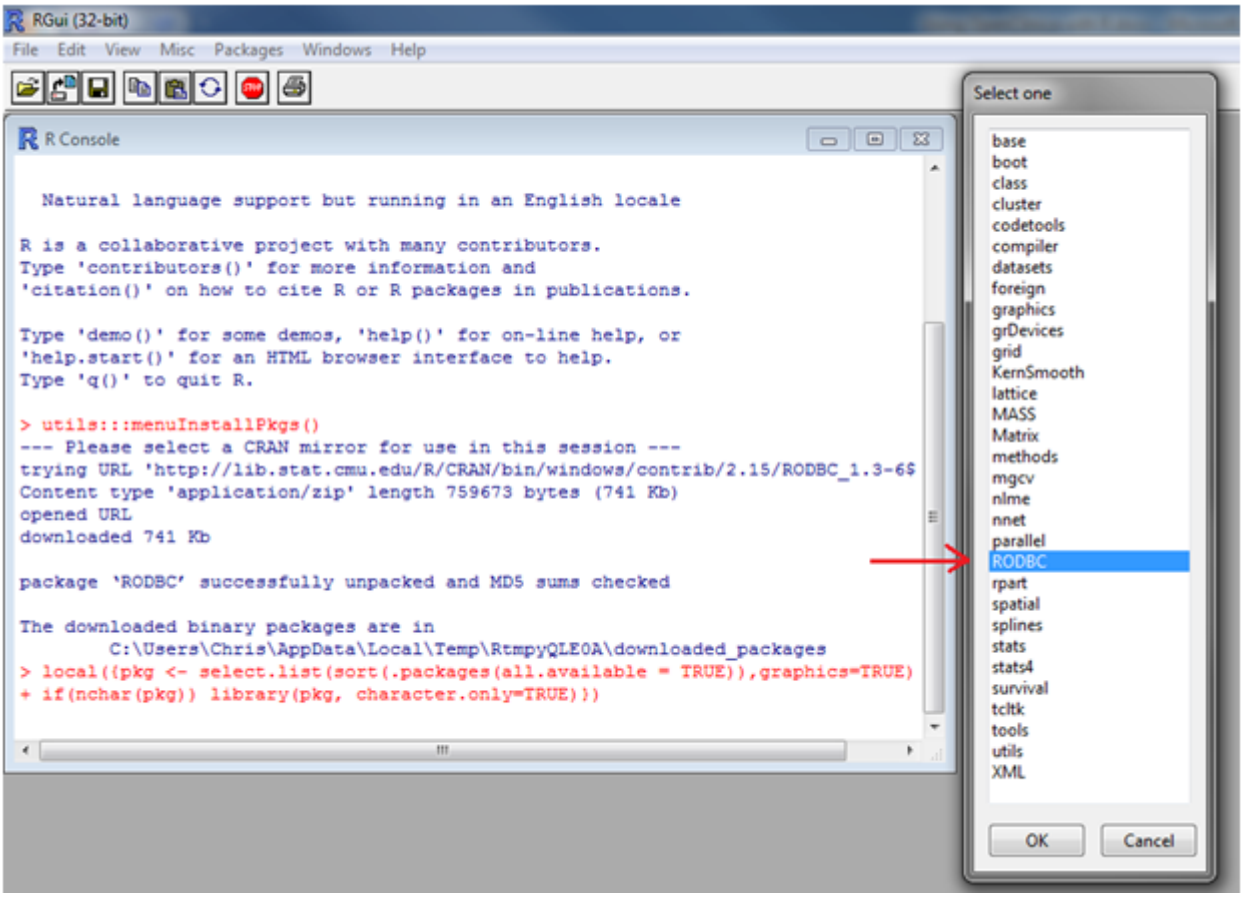


Step 2: Loading the RODBC package

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

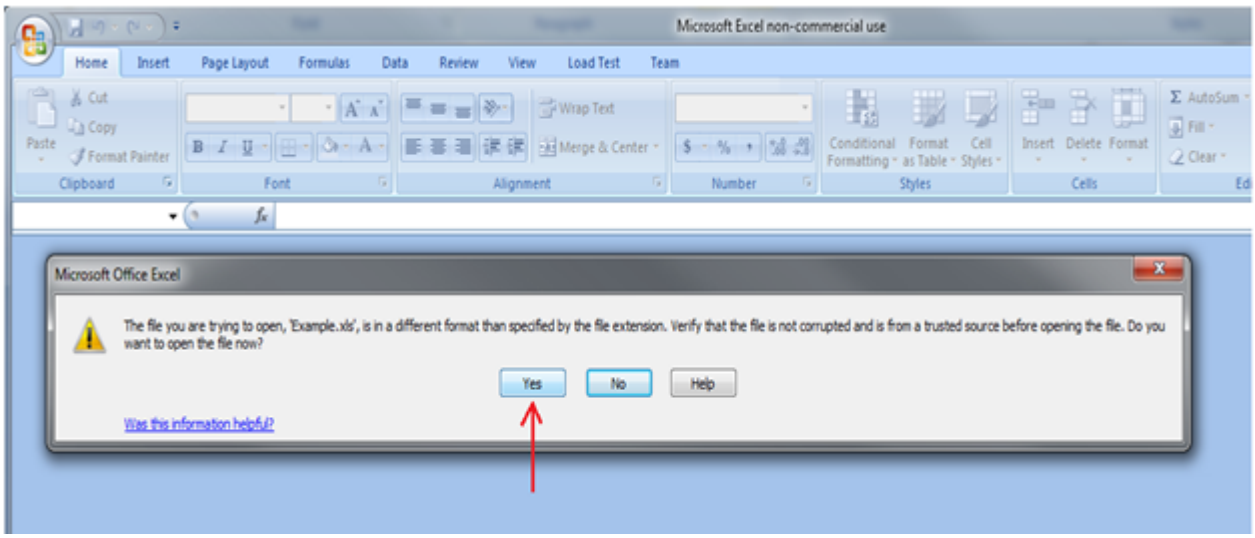


A list of packages should then pop up and you can select RODBC. You should not receive any other messages after you load the package.

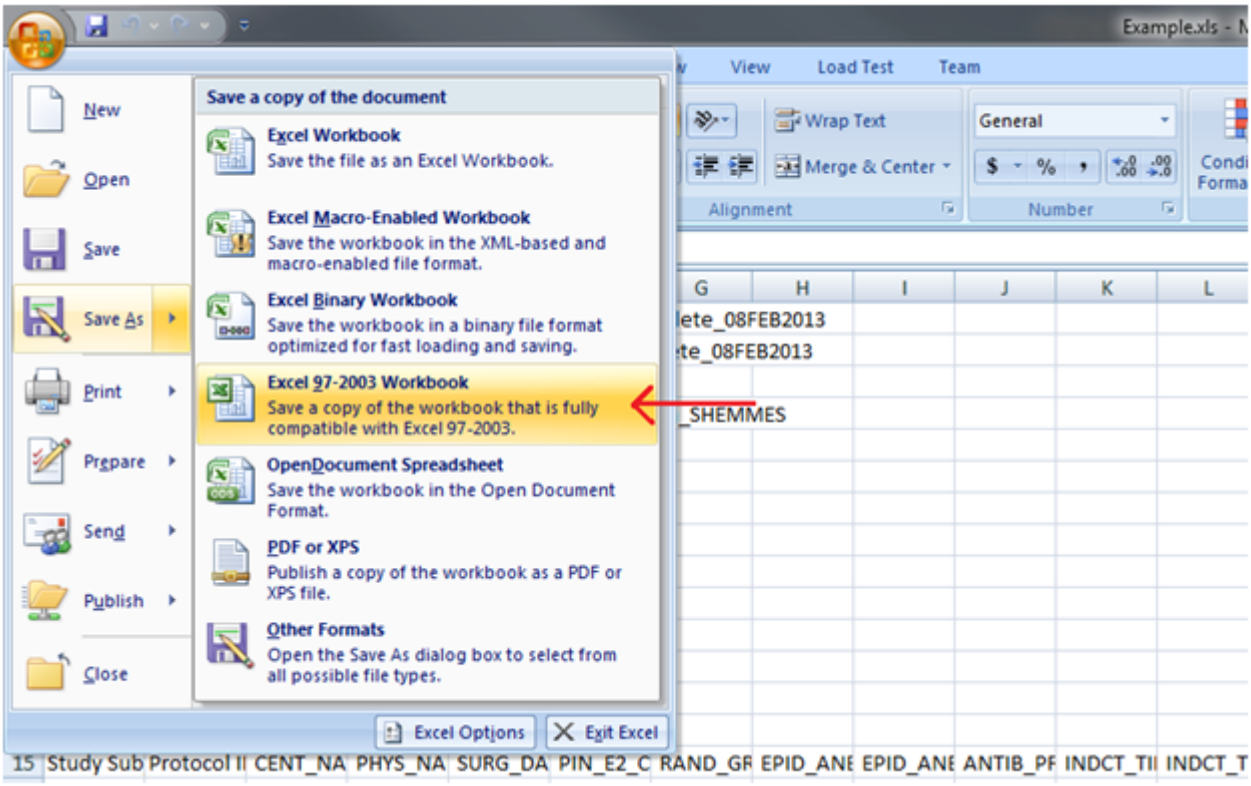


Step 3: Importing your Excel data into R

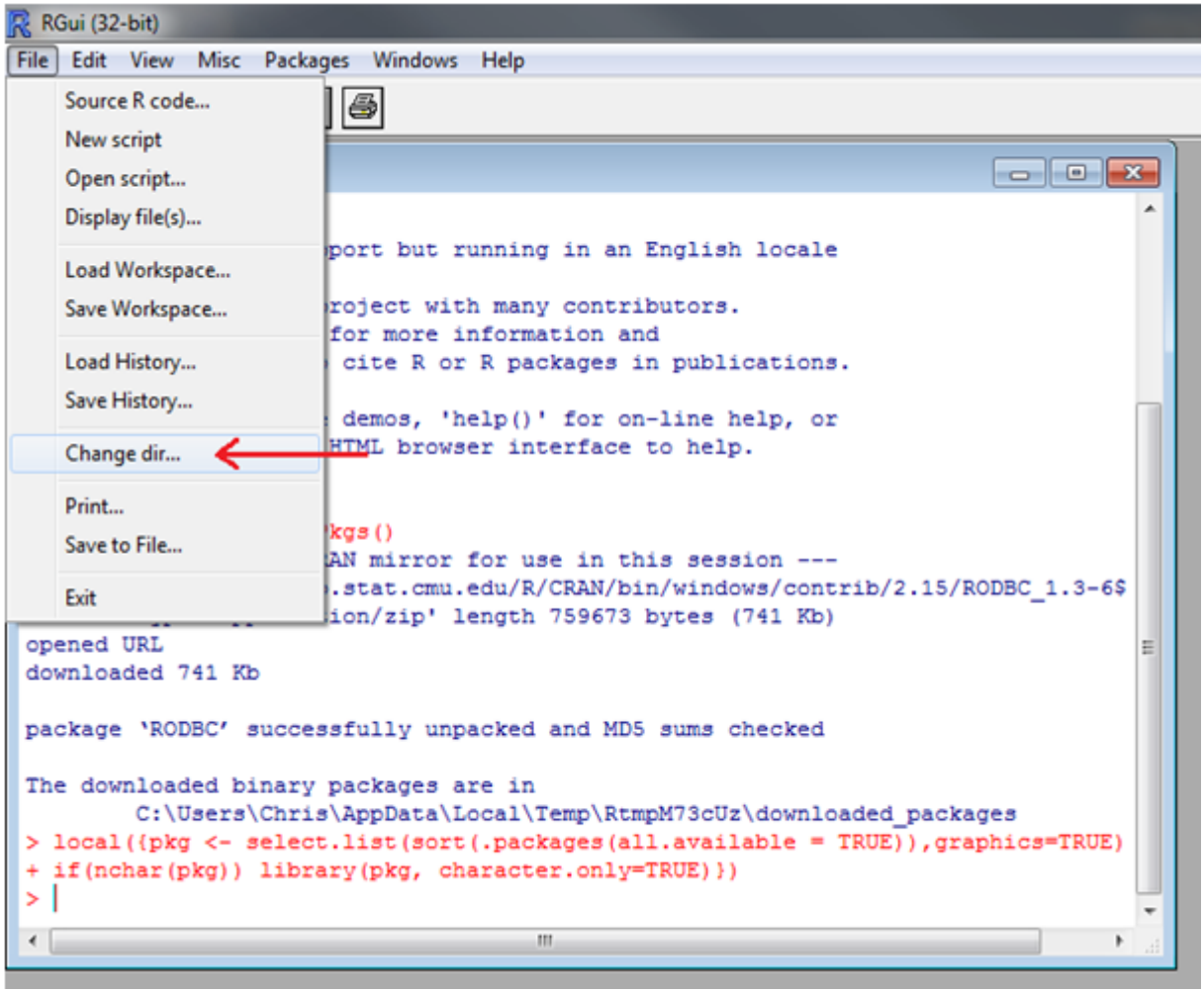
Now that you have R installed, as well as the RODBC package installed and loaded, you are almost ready to import our data into R. First open your Excel file. When you do this you will see an error message pop up. Click Yes when you see this message.



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.



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.



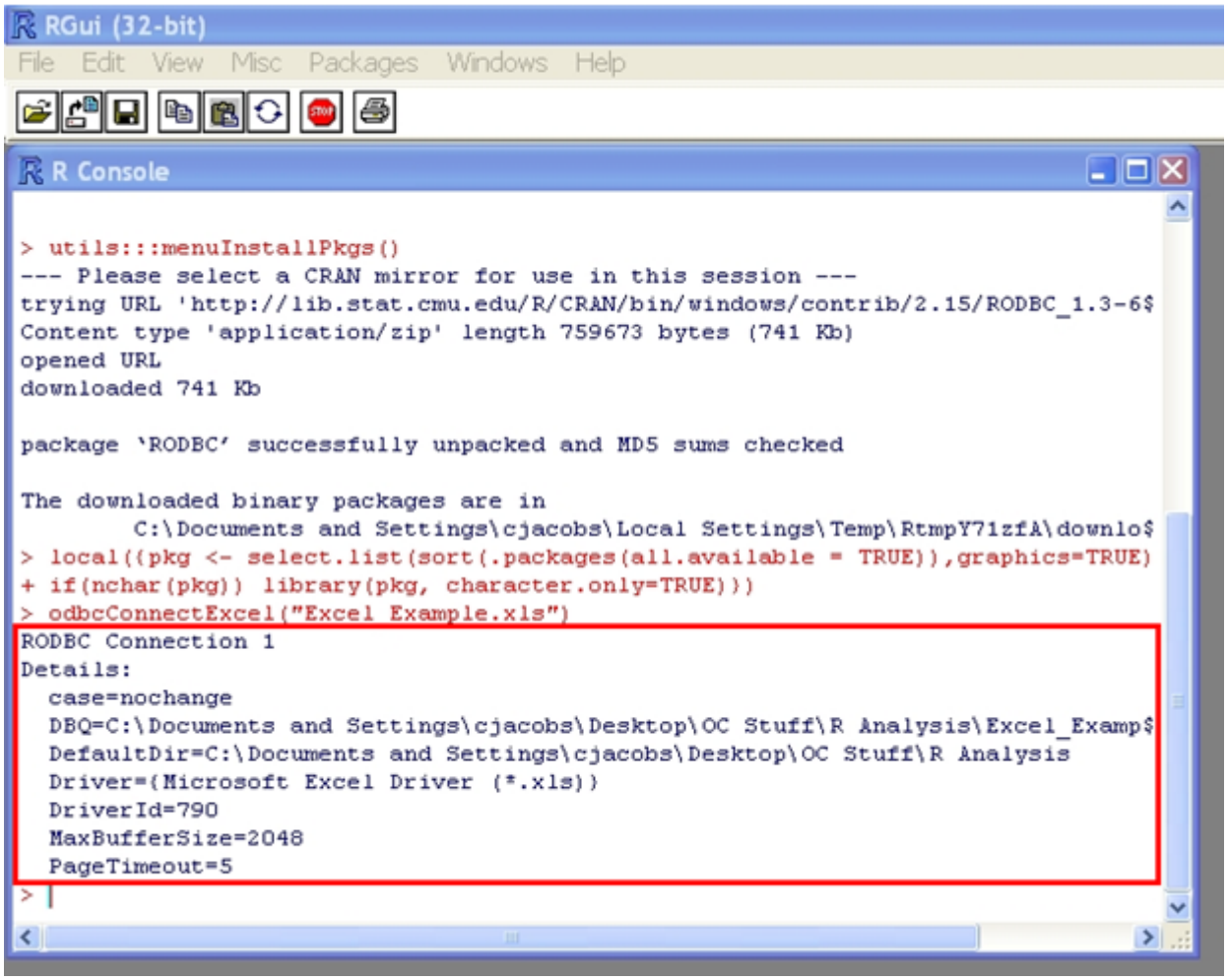
You can now import your data using the odbcConnectExcel function provided to us by the RODBC package. To do this, type:

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.