

Reading Data From Excel Files (xls|xlsx) into R

Previously, we described the [essentials of R programming](#) and some [best practices for preparing your data](#). We also provided quick start guides for [reading](#) and [writing](#) txt and csv files using **R** base functions as well as using a most modern R package named [readr](#), which is faster (X10) than R base functions.

In this article, you'll learn how to **read data** from **Excel xls** or **xlsx** file formats into **R**. This can be done either by:

- **copying data** from Excel
- using **readxl** package
- or using **xlsx** package

Copying data from Excel and import into R

On Windows system

1. **Open** the Excel file containing your data: **select** and **copy the data** (ctrl + c)
2. Type the R code below to import the copied data from the **clipboard** into R and store the data in a data frame (`my_data`):

```
my_data <- read.table(file = "clipboard",  
                      sep = "\t",  
                      header=TRUE)
```

On Mac OSX system

1. **Select** and **copy the data** (Cmd + c)
2. Use the function **pipe(pbpaste)** to import the data you've copied (with Cmd + c):

```
my_data <- read.table(pipe("pbpaste"), sep="\t", header =  
TRUE)
```

Importing Excel files into R using readxl package

The **readxl** package, developed by Hadley Wickham, can be used to easily import Excel files (xls|xlsx) into R without any external dependencies.

Installing and loading readxl package

- Install

```
install.packages("readxl")
```

- Load

```
library("readxl")
```

Using readxl package

The **readxl** package comes with the function **read_excel()** to read xls and xlsx files

1. Read both xls and xlsx files

```
library("readxl")
```

```
my_data <- read_excel("my_file.xls")
```

```
my_data <- read_excel("my_file.xlsx")
```

The above R code, assumes that the file “my_file.xls” and “my_file.xlsx” is in your current [working directory](#). To know your current working directory, type the function **getwd()** in R console.

- It's also possible to choose a file interactively using the function **file.choose()**, which I recommend if you're a beginner in R programming:

```
my_data <- read_excel(file.choose())
```

If you use the R code above in RStudio, you will be asked to choose a file.

2. Specify sheet with a number or name

```
my_data <- read_excel("my_file.xlsx", sheet =  
"data")
```

```
my_data <- read_excel("my_file.xlsx", sheet = 2)
```

- ### 3. Case of missing values: NA (not available).
- If NAs are represented by something (example: “—”) other than blank cells, set the na argument:

```
my_data <- read_excel("my_file.xlsx", na = "---  
")
```

Importing Excel files using xlsx package

The **xlsx** package, a java-based solution, is one of the powerful R packages to **read**, **write** and **format Excel files**.

Installing and loading xlsx package

- Install

```
install.packages("xlsx")
```

- Load

```
library("xlsx")
```

Using xlsx package

There are two main functions in **xlsx** package for reading both xls and xlsx Excel files: **read.xlsx()** and **read.xlsx2()** [faster on big files compared to read.xlsx function].

The simplified formats are:

```
read.xlsx(file, sheetIndex, header=TRUE)
read.xlsx2(file, sheetIndex,
header=TRUE)
```

- **file**: file path
- **sheetIndex**: the index of the sheet to be read
- **header**: a logical value. If TRUE, the first row is used as column names.

Example of usage:

```
library("xlsx")
my_data <- read.xlsx(file.choose(), 1)
```

Read more

Read more about for reading, writing and formatting Excel files:

- [R xlsx package : A quick start guide to manipulate Excel files in R](#)
- [r2excel package: Read, write and format easily Excel files using R software](#)

Summary

- Read Excel files using **readxl** package: **read_excel**(file.choose(), sheet = 1)
- Read Excel files using **xlsx** package: **read.xlsx**(file.choose(), sheetIndex = 1)

