

R Read XLS, XLSX – Excel Files

In this tutorial, we will learn how to read an XLSX file in R programming.

R Read Excel Files

To read Excel (XLS and XLSX) files in R, we will use the package `readxl`.

Install `readxl` package by running the command `install.packages("readxl")`. You should see some information echoed to the screen as shown in the below code snippet. The command installs all the dependencies.

```
> install.packages("readxl")
Warning in install.packages("readxl") :
  'lib = "C:/Program Files/R/R-3.5.2/library"' is not writable
--- Please select a CRAN mirror for use in this session ---
also installing the
dependencies 'magrittr', 'assertthat', 'utf8', 'rematch', 'hms', 'prettyunits', 'R6', 'crayon', 'cli', 'fansi', 'pillar', 'pkgconfig', 'rlang', 'c
trying URL 'https://cloud.r-project.org/bin/windows/contrib/3.5/magrittr_1.5.zip'
Content type 'application/zip' length 155601 bytes (151 KB)
downloaded 151 KB

..

package 'tibble' successfully unpacked and MD5 sums checked
package 'readxl' successfully unpacked and MD5 sums checked

The downloaded binary packages are in
  C:\Users\TutorialKart\AppData\Local\Temp\Rtmpq4O6ls\downloaded_packages
```

Once downloaded you can start using the package in R Console.

To use the package `readxl`, run the following command.

```
> library(readxl)
```

Now you can start using the functions of `readxl` and read XLSX, XLS files.

To read an xlsx file, we can use `read_xlsx()` function or `read_excel()` function. We will learn to use these functions with examples.

R Read XLSX file using read_xlsx() funtion

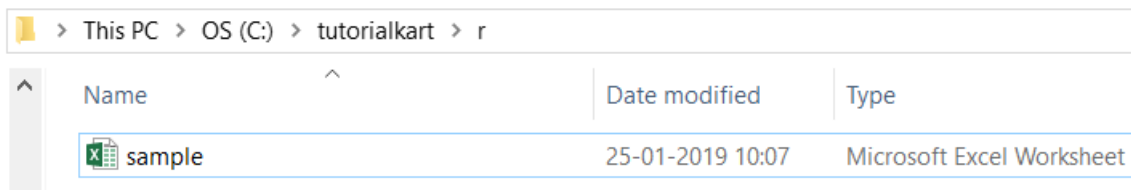
The syntax of `read_xlsx()` function is

```
read_xlsx(path, sheet
```

```
range = NULL, col_names = TRUE,  
col_types = NULL, na = "", trim_ws = TRUE, skip = 0,  
n_max = Inf, guess_max = min(1000, n_max),  
progress = readxl_progress(), .name_repair = "unique")
```

Except for the `path` argument (the first argument) the rest are optional.

In this example, we will consider `sample.xlsx` file containing 3 columns and five rows stored in a local drive.



The contents of the XLSX file are as shown below.

The screenshot shows the Microsoft Excel interface. The title bar reads 'sample'. The menu bar includes 'File', 'Edit', 'View', 'Insert', 'Format', 'Data', and 'Tools'. The ribbon shows 'fx' in the formula bar. The main grid has columns A, B, and C, and rows 1 through 7. The data is as follows:

| | A | B | C |
|---|----|----------|--------|
| 1 | ID | Name | Salary |
| 2 | 22 | John | 25000 |
| 3 | 41 | Samantha | 30000 |
| 4 | 15 | Ron | 37000 |
| 5 | 63 | Rick | 15000 |
| 6 | 87 | Gary | 56000 |
| 7 | | | |

Now we shall run the `read_xlsx()` function with the path to xlsx file as argument.

```
> library(readxl)
```

```
> library(readxl)
> read_xlsx("C:\\tutorialkart\\r\\sample.xlsx")
# A tibble: 5 x 3
  ID Name    Salary
<dbl> <chr> <dbl>
1 22 John 25000
2 41 Samantha 30000
3 15 Ron 37000
4 63 Rick 15000
5 87 Gary 56000
>
```

Now we will go through what `read_xlsx()` function has read from the `xlsx` file.

1. It found that the size of the data is 5×3.
2. It picked the first row as Header i.e., column names for the columns.
3. It interpreted the datatypes of the columns.
 - a. First column `<dbl>` for Double.
 - b. Second column `<chr>` for Char Array.
 - c. Third column `<dbl>` for Double.

R Read XLSX file using `read_excel()` function

The syntax of `read_excel()` function is

```
read_excel(path, sheet
```

```
read_excel(path, sheet = NULL, range = NULL, col_names = TRUE,
col_types = NULL, na = "", trim_ws = TRUE, skip = 0,
n_max = Inf, guess_max = min(1000, n_max),
progress = readxl_progress(), .name_repair = "unique")
```

Except for the `path` argument (the first argument) the rest are optional.

Now we shall run the `read_excel()` function with the path to `xlsx` file as argument.

```
> library(readxl)
```

```

> library(readxl)
> read_excel("C:\\tutorialkart\\r\\sample.xlsx")
# A tibble: 5 x 3
  ID Name    Salary
<dbl> <chr> <dbl>
1  22 John  25000
2  41 Samantha 30000
3  15 Ron   37000
4  63 Rick  15000
5  87 Gary  56000
>

```

Difference between read_xlsx() and read_excel() functions

The result for reading the .xlsx file with the two functions is same. The only difference is that when read_excel() is used, excel_format() is called internally by the read_excel() function to determine if the path is xls or xlsx file from the file extension.

R Read XLX file

Similar to XLSX file, we can use read_excel() function to read an XLS file.

```
> library(readxl)
```

```

> library(readxl)
> read_excel("C:\\tutorialkart\\r\\sample.xls")

```

Or you can use read_xls() function if you know the extension of the excel file to XLS for sure.

The syntax of read_xls() function is

```
read_xls(path, sheet =
```

```

read_xls(path, sheet = NULL, range = NULL, col_names = TRUE,
col_types = NULL, na = "", trim_ws = TRUE, skip = 0,
n_max = Inf, guess_max = min(1000, n_max),
progress = readxl_progress(), .name_repair = "unique")

```

Following is an example to use read_xls() function to read XLS Excel file.

```
> library(readxl)
```

```
> library(readxl)
> read_xls("C:\\tutorialkart\\r\\sample.xls")
# A tibble: 5 x 3
  ID Name    Salary
<dbl> <chr> <dbl>
1  22 John  25000
2  41 Samantha 30000
3  15 Ron   37000
4  63 Rick  15000
5  87 Gary  56000
>
```

Concluding this [R Tutorial](#), we have learned how to read Excel XLS and XLSX files in R programming using readxl package.

Home - Get Started

- [R Tutorial](#)
- [R Script File](#)
- [R Working Directory](#)
- [R Data Types](#)
- [R Variables](#)
- [R Operators](#)
- [R Vectors](#)
- [R Matrix](#)

Decision Making

- [R Decision Making](#)
- [R if](#)
- [R if..else](#)
- [R if..else if...else](#)
- [R switch](#)

Loops

- [R Loops](#)
- [R repeat loop](#)
- [R while loop](#)
- [R for loop](#)
- [R break](#)

Strings

- R Strings
- Find length of String in R
- Extract Substring from a String in R
- Concatenate two or more Strings in R

Functions

- R Functions

DataFrame

- R Data Frame
- Sort R Data Frame by Column
- For each row in an R Data Frame
- Import Excel Data into R Dataframe
- Convert R Dataframe to Matrix
- R Dataframe - Delete Rows
- R Dataframe - Drop Columns
- R Dataframe - Add Column
- R Dataframe - Change Column Name
- R Dataframe - Remove Duplicate Rows
- R Dataframe - Replace NA with 0
- Convert Matrix to R Dataframe

Handling Data from Files

- R CSV Files - Read, Filter, Write
- R Read Excel XLS XLSX files

Charts & Graphs

- R Pie Charts
- R Line Graphs

Statistical Analysis

- R Mean of a Vector
- R Median of a Vector