

Reading sheets

INTRODUCTION TO IMPORTING DATA IN R



Filip Schouwenaars
Instructor, DataCamp



XLConnect

- Martin Studer
- Work with Excel through R
- Bridge between Excel and R
- XLS and XLSX
- Easy-to-use functionality

Installation

```
install.packages("XLConnect")
```

```
also installing the dependencies 'XLConnectJars', 'rJava'  
...
```

- Problems?
 - Install Oracle's Java Development Kit (JDK)
 - Google your error!

loadWorkbook()

```
library("XLConnect")  
book <- loadWorkbook("cities.xlsx")  
str(book)
```

```
Formal class 'workbook' [package "XLConnect"] with 2 slots  
  ..@ filename: chr "cities.xlsx"  
  ..@ jobj      : ...
```

getSheets()

```
getSheets(book)
```

```
"year_1990" "year_2000"
```

```
library(readxl)  
excel_sheets("cities.xlsx")
```

```
"year_1990" "year_2000"
```

readWorksheet()

```
readWorksheet(book, sheet = "year_2000")
```

```
   Capital Population
1 New York  17800000
2   Berlin   3382169
3   Madrid   2938723
4 Stockholm  1942362
```

readWorksheet()

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_2000 col 2

row 3
row 4

```
readWorksheet(book, sheet = "year_2000",  
             startRow = 3,  
             endRow = 4,  
             startCol = 2,  
             header = FALSE)
```

```
Col1  
1 3382169  
2 2938723
```


Let's practice!

INTRODUCTION TO IMPORTING DATA IN R

Adapting sheets

INTRODUCTION TO IMPORTING DATA IN R



Filip Schouwenaars
Instructor, DataCamp

New data!

```
pop_2010 <- data.frame(Capital = c("New York", "Berlin", "Madrid", "Stockholm"),  
  Population = c(8191900, 3460725, 3273000, 1372565))
```

pop_2010

	Capital	Population
1	New York	8191900
2	Berlin	3460725
3	Madrid	3273000
4	Stockholm	1372565

createSheet()

```
pop_2010 <- ... # truncated  
library(XLConnect)  
book <- loadWorkbook("cities.xlsx")
```

Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_1990

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_2000



createSheet()

```
pop_2010 <- ... # truncated
library(XLConnect)
book <- loadWorkbook("cities.xlsx")
createSheet(book, name = "year_2010")
```

The diagram shows a transformation of a single table into a wide table format. On the left, a single table with 5 rows and 2 columns is shown:

Capital	Population
New York	16044000
Berlin	3500000
Madrid	3200000
Stockholm	1200000

Below this table, the text "year_1990" is displayed. An arrow points from this table to a second table on the right. Above the second table is a green Excel icon. The second table has 5 rows and 2 columns:

Capital	Population
New York	17800000
Berlin	3500000
Madrid	2900000
Stockholm	1200000

Below this table, the text "year_2000" is displayed. A third arrow points from the second table to a third table on the right. The third table has 5 rows and 2 columns:

Below this table, the text "year_2010" is displayed. The diagram illustrates how a single table can be transformed into a wide table format, where each row represents a different time point (year) for the same set of categories (Capital).

writeWorksheet()

```
pop_2010 <- ... # truncated
library(XLConnect)
book <- loadWorkbook("cities.xlsx")
createSheet(book, name = "year_2010")
writeWorksheet(book, pop_2010, sheet = "year_2010")
```

The diagram shows a transformation of a single table into a wide table format. On the left, a single table with 2 columns (Capital, Population) and 5 rows (New York, Berlin, Madrid, Stockholm, and an unlabeled row) is shown. This table is transformed into a wide table with 3 columns (Capital, Population, and an unlabeled column) and 5 rows. The transformation is labeled 'year_1990' and 'year_2000' at the bottom. An Excel icon is present in the top right corner.

Capital	Population
New York	16044000
Berlin	3500000
Madrid	3000000
Stockholm	1000000

year_1990


Capital	Population	
New York	17800000	
Berlin	3500000	
Madrid	2800000	
Stockholm	1000000	

year_2000

year_2010

saveWorkbook()

```
pop_2010 <- ... # truncated
library(XLConnect)
book <- loadWorkbook("cities.xlsx")
createSheet(book, name = "year_2010")
writeWorksheet(book, pop_2010, sheet = "year_2010")
```



Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_1990

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_2000

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362


year_2010

saveWorkbook()

```
pop_2010 <- ... # truncated
library(XLConnect)
book <- loadWorkbook("cities.xlsx")
createSheet(book, name = "year_2010")
writeWorksheet(book, pop_2010, sheet = "year_2010")

saveWorkbook(book, file = "cities2.xlsx")
```

Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
year_1990	
Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
year_2000	
Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
year_2010	

 cities2.xlsx

renameSheet()

```
renameSheet(book, "year_1990", "Y1990")
renameSheet(book, "year_2000", "Y2000")
renameSheet(book, "year_2010", "Y2010")
```

Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_1990

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

year_2000

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362


year_2010



renameSheet()


```
renameSheet(book, "year_1990", "Y1990")
renameSheet(book, "year_2000", "Y2000")
renameSheet(book, "year_2010", "Y2010")
saveWorkbook(book, file = "cities3.xlsx")
```

Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
Y1990	
Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
Y2000	
Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
Y2010	

 cities3.xlsx

removeSheet()

```
removeSheet(book, sheet = "Y2010")
```



Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

Y1990

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

Y2000

Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362

Y2010

removeSheet()

```
removeSheet(book, sheet = "Y2010")  
saveWorkbook(book, file = "cities4.xlsx")
```

Capital	Population
New York	16044000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
Y1990	
Capital	Population
New York	17800000
Berlin	3382169
Madrid	2938723
Stockholm	1942362
Y2000	



Wrap-up

- Basic operations
- Reproducibility is the key!
- More functionality
 - Styling cells
 - Working with formulas
 - Arranging cells
 - ...

Let's practice!

INTRODUCTION TO IMPORTING DATA IN R