



# Basic loops in R/RStudio

*(work in progress, 2.0)*

*Oscar Torres-Reyna*  
*otorres@princeton.edu*



November 2018

<http://dss.princeton.edu/training/>



# Basic Loops in R/RStudio

(v. 2.0)

***Oscar Torres-Reyna***

*otorres@princeton.edu*

November 2018

<http://www.princeton.edu/~otorres/>

# **APPLYING SAME R CODE TO MULTIPLE FILES**

# Note

The following code applies the same program to multiple \*.csv files and produces one file containing all files by appending them.

All files must have the variables with the same spelling and same order.

For the example, download the \*.csv files available in this link:

<https://www.princeton.edu/~otorres/Rfiles/>

Make sure to save all files in the same folder.

Source of data is *World Development Indicators* from the *World Bank*.

```
# Set working directory to where the *.csv files are saved
# Go to Session -> Set Working directory -> Choose directory
# or type and run:
```

```
setwd("write path here")
```

```
# Make a list of all *.csv files in the folder,
# make sure only the ones you need are in the working directory
```

```
files <- list.files(path = ".", pattern = ".csv")
```

```
files
```

```
[1] "Canada.csv" "Mexico.csv" "United States.csv"
```

```
# Create a reference data frame with the structure of the final
# version of your data. Read one file first, then leave the
# first row as placeholder.
# If the final version is different from the original
# individual files, then create an empty data frame with
# the expected final structure of the data
```

```
mydata = read.csv(files[1], header = TRUE, stringsAsFactors = FALSE)
```

```
mydata = data.frame(mydata[1,]) # Keep first row as placeholder
```

```
mydata
```

```
  year country      gdppc unempf unempm unemp      exports      imports polity polity2
1 1980  Canada 23,070.40   0.08   0.07 7.50% 1.15284e+11 1.03395e+11     10     10
```

```
# Process each datafile, then append all files into one  
# This loop uses smartbind() from library -gtools
```

```
library(gtools)
```

```
for (record in files) {
```

```
  # Mandatory line
```

```
  temp <- read.csv(record, header = TRUE, stringsAsFactors = FALSE)
```

```
  temp$trade = temp$exports + temp$imports # Your code from here
```

```
  mydata = smartbind(mydata, temp) # Mandatory last line
```

```
}
```

```
# Make sure to change 'temp' with the last version in the last  
mandatory line, only if you created different versions.
```

```
# Make sure to drop the first line used as placeholder
```

```
mydata <- mydata[-1, ]
```

```
View(mydata)
```

# **SAME REGRESSION ACROSS SUBGROUPS**



## **# Get the sample data**

```
mydata <- read.csv("http://www.princeton.edu/~otorres/mydataR.csv",  
                  header=TRUE, stringsAsFactors = FALSE)
```

## **# Create a unique id per group, in this example per country**

```
library(plyr)  
mydata$id <- id(mydata[c("country")], drop = TRUE)
```

## **# Running same regression per group**

```
b <- lapply(1:max(unique(mydata$id)), function(i) {  
  reg <- with(subset(mydata, mydata$id==i), lm(unemp ~ gdppc + polity2))  
  reg$coefficients  
})
```

## **# Getting the betas in data frame form**

```
betas <- as.data.frame(do.call(rbind,b))
```

```
# Adding id to the betas data frame
```

```
betas$id = rownames(betas)
```

Source of data: WDI-World Bank and Polity IV project