

Εισαγωγή στην R

Εύη Γκέλη Ευγενία Μάρκο Αντώνης Μπαλλής
Αγγελική Παναγοπούλου Έλενα Τσίπου

Πανεπιστήμιο Πατρών
Τμήμα Οικονομικών Επιστημών
Εφαρμοσμένη Οικονομική και Ανάλυση Δεδομένων
Διδάσκων: Κωνσταντίνος Κουνετάς

29 Νοεμβρίου 2015

Εισαγωγικά στοιχεία

- Η R είναι μια γλώσσα προγραμματισμού στην οποία οι εντολές διαβάζονται και κατόπιν εκτελούνται αμέσως. Στην εργασία αυτή θα παρουσιάσουμε τις βασικές έννοιες, την εισαγωγή και επεξεργασία δεδομένων, την εξαγωγή αποτελεσμάτων και τέλος την εισαγωγή βιβλιοθηκών που υλοποιούνται με τη γλώσσα προγραμματισμού R .

Εισαγωγή *dataset* στην *R*

Για να εισάγουμε ένα *dataset* στην *R* θα πρέπει πρώτα να δηλώσουμε ποια θα είναι η τοποθεσία του *dataset*. Η εντολή για αυτό είναι :

`setwd (" <location of your Dataset >")` ή

Manually : *Session* - > *Set Working Directory* - > *Choose Directory*

- *Import csv File*

```
csv.data <- read.csv("BETACAR.csv")
```

- *Import xls File*

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

```
library(xlsx)
```

```
excel.data <- read.xlsx("BLOOD.DAT.xls", sheetIndex = 1)
```

- *Import Stata File*

```
library(foreign)
```

```
dta.data <- read.dta("headache.dta")
```

Για παράδειγμα:

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for importing various data formats:


```

1 #importing a csv file dataset
2 csv.data <- read.csv("BETACAR.csv")
3 #alternative command where you choose manually your file
4 csv.data <- read.csv(file.choose())
5
6 #importing an xls file dataset
7 library(xlsx)
8 excel.data <- read.xlsx("BLOOD.DAT.xls", sheetIndex = 1)
9
10 #importing a STATA form dataset
11 library(foreign)
12 dta.data <- read.dta("headache.dta")
13
14
15
16

```
- Environment:** Shows the objects created in the Global Environment:

Object Name	Details
csv.data	23 obs. of 8 variables
dta.data	846 obs. of 6 variables
excel.data	510 obs. of 9 variables
- Console:** Shows the execution output:


```

> #importing a csv file dataset
> csv.data <- read.csv("BETACAR.csv")
> #alternative command where you choose manually your file
> csv.data <- read.csv(file.choose())
> #importing an xls file dataset
> library(xlsx)
Loading required package: r3java
Loading required package: xlsxjars
> excel.data <- read.xlsx("BLOOD.DAT.xls", sheetIndex = 1)
> #importing a STATA form dataset
> library(foreign)
> dta.data <- read.dta("headache.dta")
> |

```
- Packages:** Lists installed and available packages:

Name	Description	Version
<input checked="" type="checkbox"/> foreign	Read Data Stored by Minitab, S, SAS, SPSS, Stata, Systat, Weka, dBase ...	0.8-66
<input type="checkbox"/> lme4	Testing Linear Regression Models	0.9-34
<input type="checkbox"/> Rcpp	Seamless R and C++ Integration	0.12.2
<input type="checkbox"/> readxl	Read Excel Files	1.0.0
<input checked="" type="checkbox"/> rjava	Low-Level R to Java Interface	0.9-7
<input checked="" type="checkbox"/> xlsx	Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files	0.5.7
<input checked="" type="checkbox"/> xlsxjars	Package required for the xlsx package	0.6.1
<input type="checkbox"/> XML	Tools for Parsing and Generating XML Within R and C/Obj	3.98-1.3

Αποθήκευση του *dataset* και των εντολών στην *R*

- Για να μπορέσουμε να αποθηκεύσουμε ο,τι εντολές έχουμε κάνει στην *R* πάμε στην γραμμή εργαλείων και επιλέγουμε

File -- > *Save Workspace* -- > *file name* : το όνομα που θέλουμε να δώσουμε , *save as type* : (**.RData*) -- > *Save*.

- Για να μπορέσουμε να ανακαλέσουμε ο,τι εντολές έχουμε κάνει στην *R* πάμε στην γραμμή εργαλείων και επιλέγουμε

File -- > *Load Workspace* -- > βρίσκουμε το αρχείο που αποθηκεύσαμε το επιλέγουμε -- > *Load*

Αποθήκευση *dataset* σε *xls* αρχείο

Για να μπορέσουμε να αποθηκεύσουμε το *dataset* μας σε ένα αρχείο *xls* ακολουθούμε την παρακάτω διαδικασία:

```
library(xlsx)
```

```
write.xlsx(x,file,sheetname = "Sheet1",col.names =  
TRUE,row.names = TRUE,append = FALSE)
```

Για παράδειγμα:

```
library(xlsx)
```

```
write.xlsx(dta.data, file="mydatafile.xlsx", sheetName="Country GDP")
```

Εγκατάσταση βιβλιοθηκών στην R

- Μέσω της αρχικής εγκατάστασης της R, έχουμε αυτομάτως στη διάθεση μας τις βασικές βιβλιοθήκες (*base packages*). Επιπροσθέτως, υπάρχουν βιβλιοθήκες τις οποίες μπορούμε να εγκαταστήσουμε. Έτσι, αρχικά μπορούμε να πληροφορηθούμε αναφορικά με το πόσες και ποιές βασικές βιβλιοθήκες διαθέτουμε μέσω της αρχικής εγκατάστασης της R.
- Για την εγκατάσταση νέων βιβλιοθηκών απαιτείται η χρήση της εντολής *install.packages* (όνομα βιβλιοθήκης). Έτσι, εκτελώντας την εντολή, *install.packages()* έχουμε τη δυνατότητα για επιλογή της βιβλιοθήκης της αρεσκείας μας με τελικό στόχο την εγκατάσταση της.
 - *Σημείωση : Αποθήκευση *Dataset* σε *Library File* - > *New Project* - > *New directory* - > *R Package* - > *Create Project*