R-Day Teaching R workshop

Ralf Becker

30 October 2018

WORK-SHOP

You: “work”

I: “shop” around for good ideas!

Data and files available from:
https://github.com/datasquad/Rteaching

My context

- Teaching “large” course units to economics/business students
- Core material is econometrics (Stats applied to economic data - the cult of causal inference)
- R is taught as an “add-on”

Issues that arise

- “Why do I have to do this?”/“Is this on the exam?”
- Too little resources to run frequent labs
- Many students with no programming exposure

Aim for today

Develop ideas on how to help students to become resilient “coders”

- Gain motivation
- Don’t be thrown off the rails by problems
- Develop self-help skills

Let's assume the following

The statistical techniques we want students to be able to implement are:

- Importing data
- Cleaning data
- Merging data
- Summary stats
- Plotting data
- Regressions analysis
- Hypothesis testing

The traditional way

Here is how we tend to deliver the R-component

- Make code and data for analysis used in lectures available
- Set extra weekly worksheets (see "Week3practice.pdf")
- Have website with material to learn R (Google "ECLR R" to find it)
- Have "smallish" assessment items which assume that students have done some data work (see "CW3 201516.pdf")
- Drop-in help sessions
- Perhaps one introductory lab
- Link to Datacamp

But self-learning has limitations

The things that trip students up

- RStudio architecture
- Having to write a code file (and not click)

The things that trip students up

- RStudio architecture
- Having to write a code file (and not click)
- Error messages

```
You are recommended to install the tinytex package to build PDF.FALSE
This is pdfTeX, Version 3.14159265-2.6-1.40.17 (MiKTeX 2.9.6210 64-bit)
entering extended mode
! I can't write on file `RDayworkshop.pdf'.
Please type another file name for output
! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

Error: Failed to compile RDayworkshop.tex. See RDayworkshop.log for more info.
Execution halted
```

The things that trip students up

- RStudio architecture
- Having to write a code file (and not click)
- Error messages

```
You are recommended to install the tinytex package to build PDF.FALSE
This is pdfTeX, Version 3.14159265-2.6-1.40.17 (MiKTeX 2.9.6210 64-bit)
entering extended mode
! I can't write on file `RDayworkshop.pdf'.
Please type another file name for output
! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

Error: Failed to compile RDayworkshop.tex. See RDayworkshop.log for more info.
Execution halted
```

- They think they need to know everything - they Google everything else, but not how to solve their problem

The things that trip students up

- RStudio architecture
- Having to write a code file (and not click)
- Error messages

```
You are recommended to install the tinytex package to build PDF.FALSE
This is pdfTeX, Version 3.14159265-2.6-1.40.17 (MiKTeX 2.9.6210 64-bit)
entering extended mode
! I can't write on file `RDayworkshop.pdf'.
Please type another file name for output
! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

Error: Failed to compile RDayworkshop.tex. See RDayworkshop.log for more info.
Execution halted
```

- They think they need to know everything - they Google everything else, but not how to solve their problem
- The fear of plagiarism

The things that trip students up

- RStudio architecture
- Having to write a code file (and not click)
- Error messages

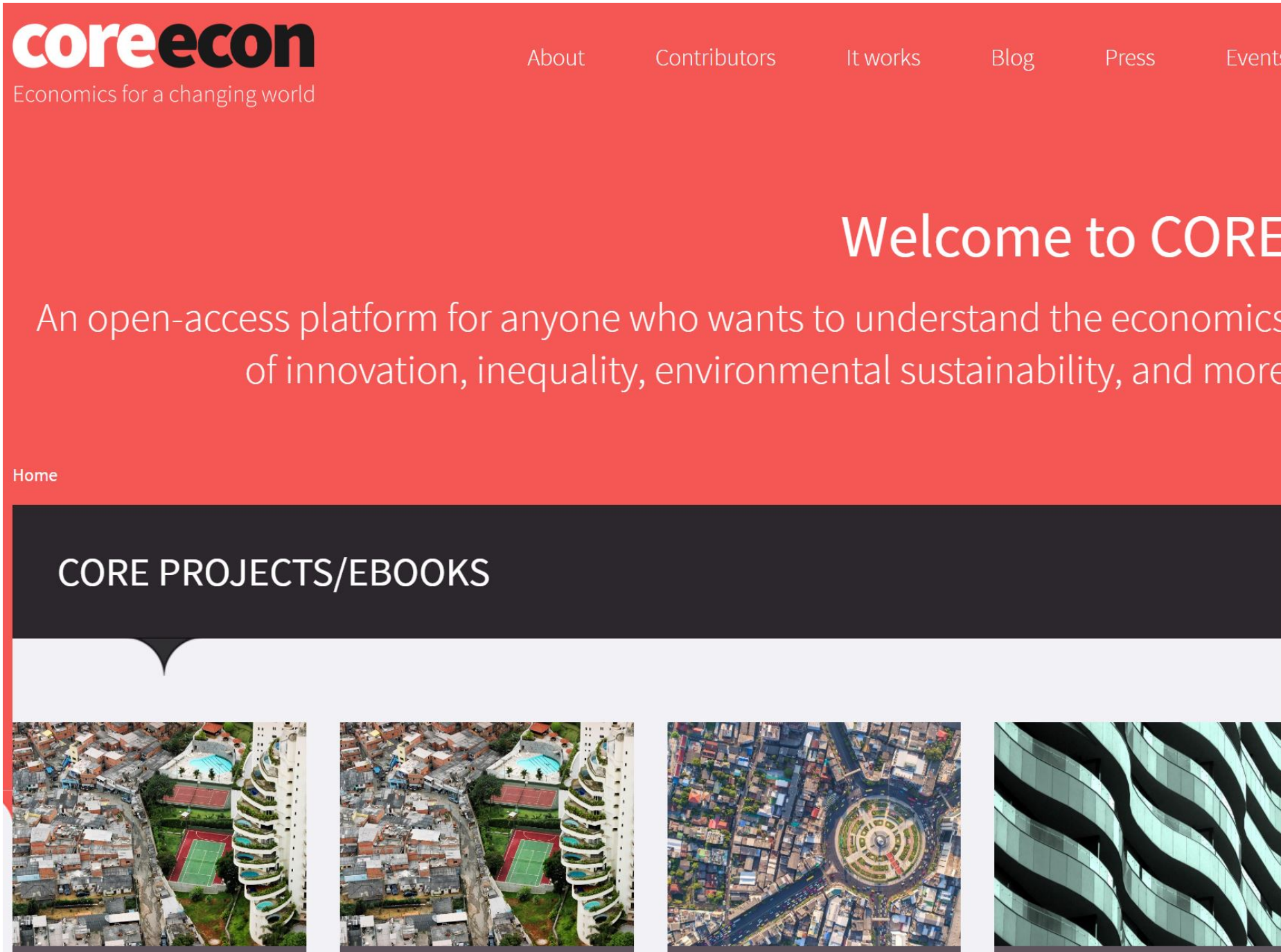
```
You are recommended to install the tinytex package to build PDF.FALSE
This is pdfTeX, Version 3.14159265-2.6-1.40.17 (MiKTeX 2.9.6210 64-bit)
entering extended mode
! I can't write on file `RDayworkshop.pdf'.
Please type another file name for output
! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

! Emergency stop.
\pgfsys@defineimage ...astform }\else {\pdfximage
                                                \pgf@imageheight \pgf@imag...
l.15 ...ight=12pt]{beamericonbook}{beamericonbook}

Error: Failed to compile RDayworkshop.tex. See RDayworkshop.log for more info.
Execution halted
```

- They think they need to know everything - they Google everything else, but not how to solve their problem
- The fear of plagiarism
- They think they will break the computer if they make a mistake

Plan for today

Think about how we can help students to develop these vital coding skills

Our testbed


coreecon
Economics for a changing world
About
Contributors
It works
Blog
Press
Event
Welcome to CORE
An open-access platform for anyone who wants to understand the economics
of innovation, inequality, environmental sustainability, and more
Home
CORE PROJECTS/EBOOKS

The projects

- Measuring climate change
- Collecting and analysing data from experiments
- Measuring the effect of a sugar tax
- Measuring wellbeing
- Measuring inequality: Lorenz curves and Gini coefficients
- Measuring management practices
- Supply and demand
- Measuring the non-monetary cost of unemployment
- Credit-excluded households in a developing country
- Characteristics of banking systems around the world
- Measuring willingness to pay for climate change mitigation
- Government policies and popularity: Hong Kong cash handout

Measuring the effect of a sugar tax

```
# loads dat_c (datafile) and var_info
load("SugarTaxData.RData")
str(dat_c)
```

```
## Classes 'tbl_df', 'tbl' and 'data.frame':    939 obs. of
##  $ store_id      : Factor w/ 26 levels "1","2","3","4",
##  $ type          : Factor w/ 13 levels "ENERGY","ENERGY-
##  $ store_type    : Factor w/ 4 levels "Large Supermarket
##  $ type2         : Factor w/ 4 levels "COCONUT WATER",.
##  $ size          : num  33.8 33.8 64 64 20 67.6 67.6 67
##  $ price         : num  1.69 1.69 2.79 2.79 1.89 2.49 2
##  $ price_per_oz  : num  0.05 0.05 0.0436 0.0436 0.0945
##  $ price_per_oz_c: num  5 5 4.36 4.36 9.45 ...
##  $ taxed         : Factor w/ 2 levels "not taxed","taxed
##  $ supp          : Factor w/ 2 levels "Standard","Supple
##  $ time          : Factor w/ 3 levels "DEC2014","JUN2015
##  $ product_id    : Factor w/ 247 levels "1","2","3","4",
##  $ period_test   : logi  TRUE TRUE TRUE TRUE TRUE TRUE
```

Basic data structure

`dat_c` contains observations for products (`product_id`) for which prices are observed in the same store (`store_id`) at three points in time (`time` - DEC2014, JUN2015, MAR2016).

A sugar tax was introduced on some products (`dat_c$taxed == "taxed"`), sometime between DEC2014 and MAR2016.

Variable Name	Type	Description
price	Num	Purchase Price
price_per_oz	Num	Price per ounce
price_per_oz_c	Num	Price per ounce cents
product_id	Num	Unique product identifier
size	Num	Total package size
store_id	Num	Unique store identifier
store_type	Num	Store Type: 1(Large Supermarket), 2(Small Supermarket), ...
supp	Num	Supplemental(1) or standard(0) item in beverage panel
taxed	Num	Tax status
time	Text	Data collection month and year
type	Text	Product Type
type2	Text	Specifies milk type or coconut water

Things we do in the project

- Find out how many products and stores there are
- Frequency Tables
- Calculating the
- Column/Bar chats
- Testing for statistical significance in Price changes

Your task

Think how you could use the available data to introduce students to one of the vital, generic programming skills

- Understand the RStudio architecture
- The need to write a code file (.r or .Rmd)
- Understand and learn from error messages
- How to google effectively for help
- How to pick someone else's code and adapt it
- Trial and error

Email finished products to ralffbecker@gmail.com (yes, two fs)