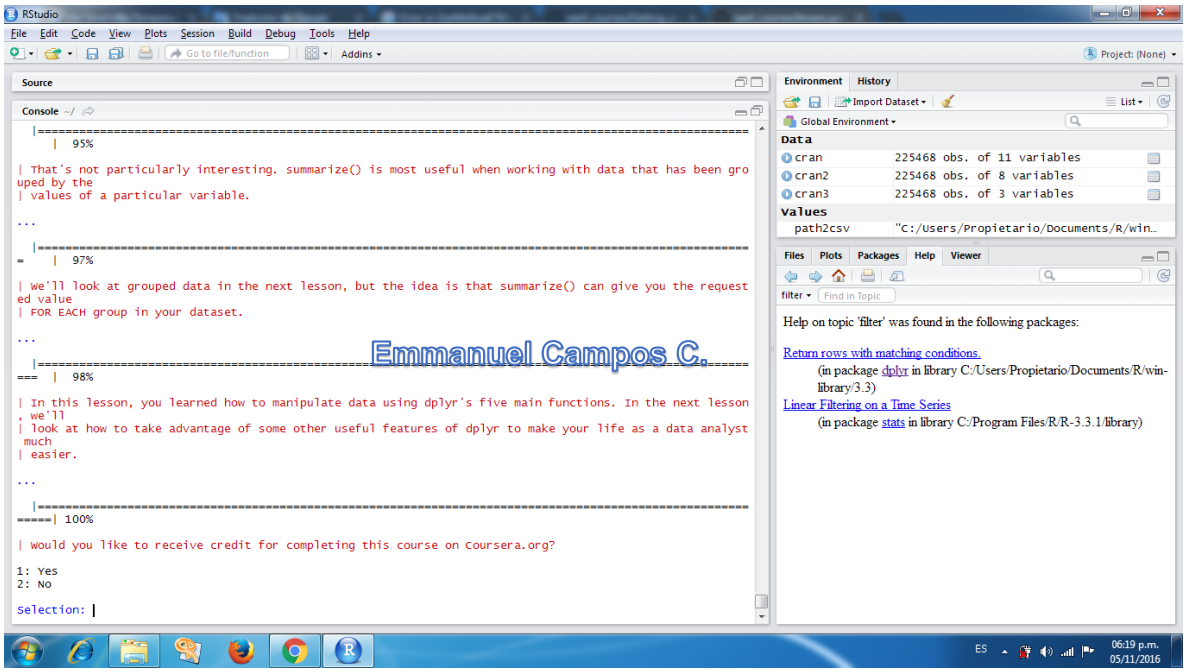
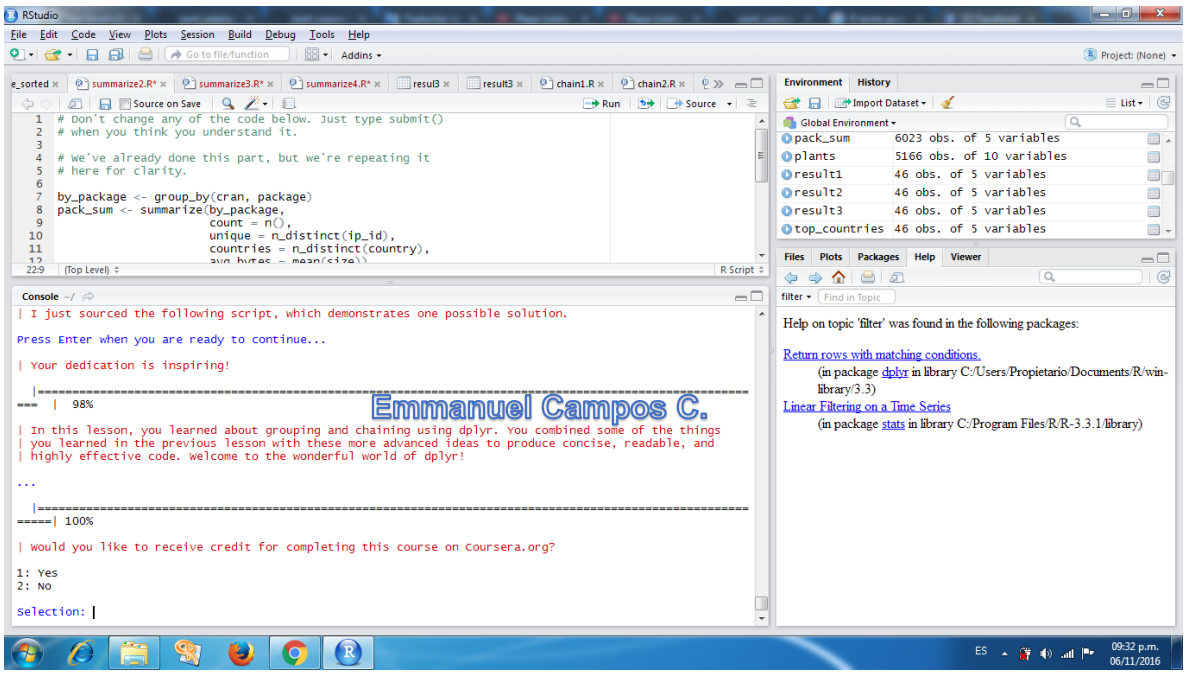
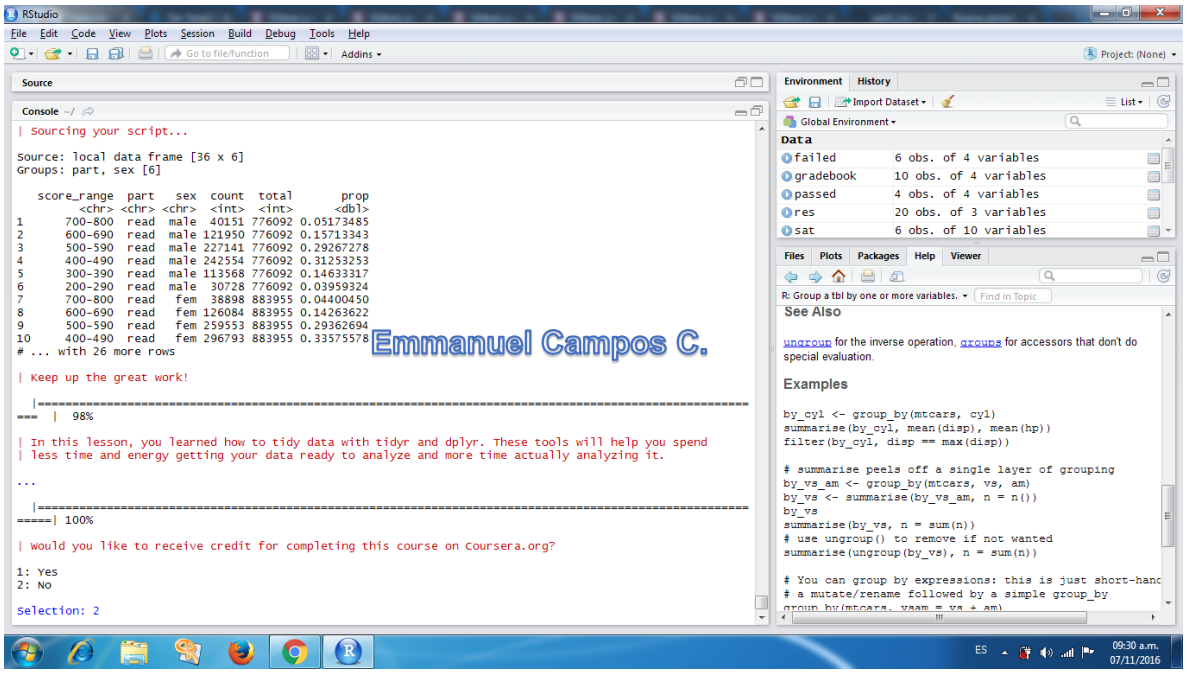
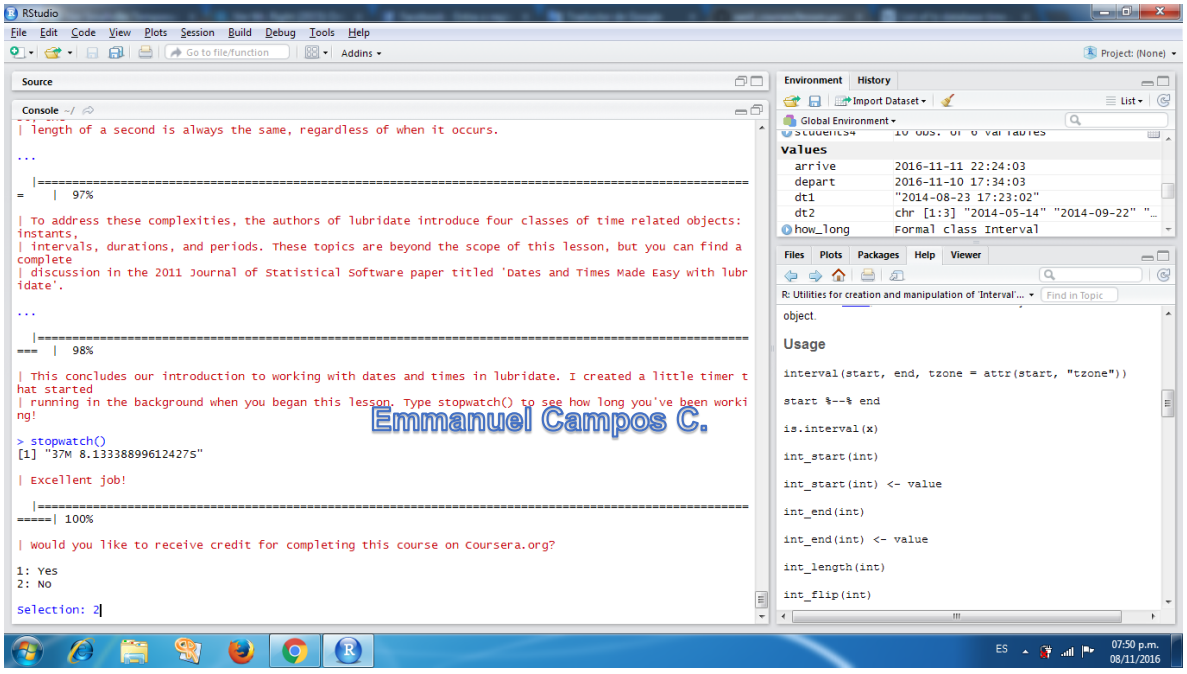

RStudio
File Edit Code View Plots Session Build Debug Tools Help
Source
Console ~/
| 95%
| That's not particularly interesting. summarize() is most useful when working with data that has been grouped by the
| values of a particular variable.
...
= | 97%
| We'll look at grouped data in the next lesson, but the idea is that summarize() can give you the requested value
| FOR EACH group in your dataset.
...
Emmanuel Campos C.
=== | 98%
| In this lesson, you learned how to manipulate data using dplyr's five main functions. In the next lesson, we'll
| look at how to take advantage of some other useful features of dplyr to make your life as a data analyst much
| easier.
...
=====| 100%
| Would you like to receive credit for completing this course on Coursera.org?
1: Yes
2: No
Selection:
Environment History
Import Dataset
Global Environment
Data
cran 225468 obs. of 11 variables
cran2 225468 obs. of 8 variables
cran3 225468 obs. of 3 variables
values
path2csv "C:/Users/Propietario/Documents/R/win...
Files Plots Packages Help Viewer
Help on topic 'filter' was found in the following packages:
Return rows with matching conditions.
(in package dplyr in library C:/Users/Propietario/Documents/R/win-library/3.3)
Linear Filtering on a Time Series
(in package stats in library C:/Program Files/R/R-3.3.1/library)
ES
06:19 p.m.
05/11/2016


RStudio
File Edit Code View Plots Session Build Debug Tools Help
summarize2.R* summarize3.R* summarize4.R* result3 result3 chain1.R chain2.R
Source on Save
Run
Source
1 # Don't change any of the code below. Just type submit()
2 # when you think you understand it.
3
4 # We've already done this part, but we're repeating it
5 # here for clarity.
6
7 by_package <- group_by(cran, package)
8 pack_sum <- summarize(by_package,
9 count = n(),
10 unique = n_distinct(ip_id),
11 countries = n_distinct(country),
22:9 (Top Level)
R Script
Console ~/
| I just sourced the following script, which demonstrates one possible solution.
Press Enter when you are ready to continue...
| Your dedication is inspiring!
=== | 98%
Emmanuel Campos C.
| In this lesson, you learned about grouping and chaining using dplyr. You combined some of the things
| you learned in the previous lesson with these more advanced ideas to produce concise, readable, and
| highly effective code. Welcome to the wonderful world of dplyr!
...
=====| 100%
| Would you like to receive credit for completing this course on Coursera.org?
1: Yes
2: No
Selection:
Environment History
Import Dataset
Global Environment
pack_sum 6023 obs. of 5 variables
plants 5166 obs. of 10 variables
result1 46 obs. of 5 variables
result2 46 obs. of 5 variables
result3 46 obs. of 5 variables
top_countries 46 obs. of 5 variables
Files Plots Packages Help Viewer
Help on topic 'filter' was found in the following packages:
Return rows with matching conditions.
(in package dplyr in library C:/Users/Propietario/Documents/R/win-library/3.3)
Linear Filtering on a Time Series
(in package stats in library C:/Program Files/R/R-3.3.1/library)
ES
09:32 p.m.
06/11/2016


RStudio
File Edit Code View Plots Session Build Debug Tools Help
Source
Console ~/
| Sourcing your script...
Source: local data frame [36 x 6]
Groups: part, sex [6]
score_range part sex count total prop
<chr> <chr> <chr> <int> <int> <dbl>
1 700-800 read male 40151 776092 0.05173485
2 600-690 read male 121950 776092 0.15713343
3 500-590 read male 227141 776092 0.29267278
4 400-490 read male 242554 776092 0.31253253
5 300-390 read male 113568 776092 0.14633317
6 200-290 read male 30728 776092 0.03959324
7 700-800 read fem 38898 883955 0.04400450
8 600-690 read fem 126084 883955 0.14263622
9 500-590 read fem 259553 883955 0.29362694
10 400-490 read fem 296793 883955 0.33575578
... with 26 more rows
| Keep up the great work!
98%
| In this lesson, you learned how to tidy data with tidyr and dplyr. These tools will help you spend
| less time and energy getting your data ready to analyze and more time actually analyzing it.
100%
| Would you like to receive credit for completing this course on Coursera.org?
1: Yes
2: No
Selection: 2
Emmanuel Campos C.
Environment History
Import Dataset
Global Environment
Data
failed 6 obs. of 4 variables
gradebook 10 obs. of 4 variables
passed 4 obs. of 4 variables
res 20 obs. of 3 variables
sat 6 obs. of 10 variables
Files Plots Packages Help Viewer
R: Group a tbl by one or more variables.
See Also
ungroup for the inverse operation, groups for accessors that don't do special evaluation.
Examples
by_cyl <- group_by(mtcars, cyl)
summarise(by_cyl, mean(disp), mean(hp))
filter(by_cyl, disp == max(disp))
# summarise peels off a single layer of grouping
by_vs_am <- group_by(mtcars, vs, am)
by_vs <- summarise(by_vs_am, n = n())
by_vs
summarise(by_vs, n = sum(n))
# use ungroup() to remove if not wanted
summarise(ungroup(by_vs), n = sum(n))
# You can group by expressions: this is just short-hand
# a mutate/rename followed by a simple group_by
ES
09:30 a.m.
07/11/2016


RStudio
File Edit Code View Plots Session Build Debug Tools Help
Source
Console ~/
| length of a second is always the same, regardless of when it occurs.
97%
| To address these complexities, the authors of lubridate introduce four classes of time related objects:
instants,
| intervals, durations, and periods. These topics are beyond the scope of this lesson, but you can find a
complete
| discussion in the 2011 Journal of Statistical Software paper titled 'Dates and Times Made Easy with lubr
idate'.
98%
| This concludes our introduction to working with dates and times in lubridate. I created a little timer t
hat started
| running in the background when you began this lesson. Type stopwatch() to see how long you've been worki
ng!
> stopwatch()
[1] "37M 8.13338899612427S"
| Excellent job!
100%
| Would you like to receive credit for completing this course on Coursera.org?
1: Yes
2: No
Selection: 2
Emmanuel Campos C.
Environment History
Import Dataset
Global Environment
Values
arrive 2016-11-11 22:24:03
depart 2016-11-10 17:34:03
dt1 "2014-08-23 17:23:02"
dt2 chr [1:3] "2014-05-14" "2014-09-22" "...
how_long Formal class Interval
Files Plots Packages Help Viewer
R: Utilities for creation and manipulation of 'Interval'...
object.
Usage
interval(start, end, tzone = attr(start, "tzone"))
start %--% end
is.interval(x)
int_start(int)
int_start(int) <- value
int_end(int)
int_end(int) <- value
int_length(int)
int_flip(int)
ES
07:50 p.m.
08/11/2016