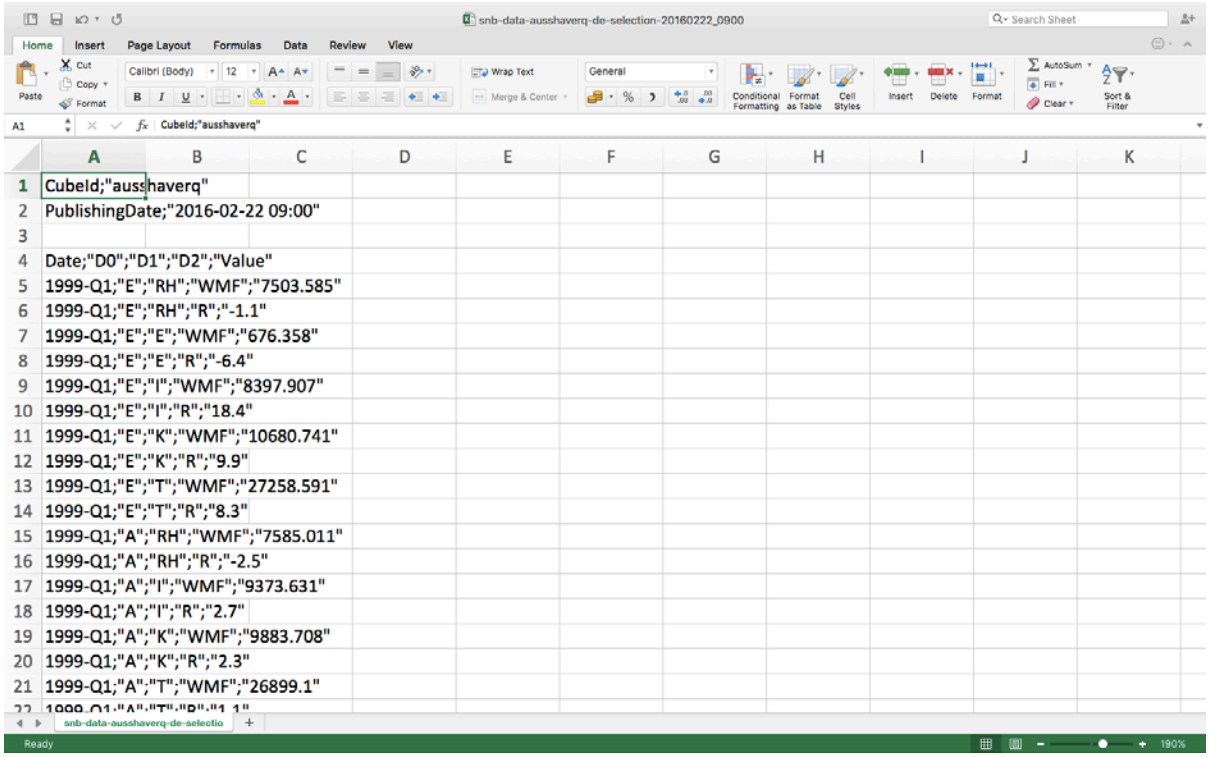
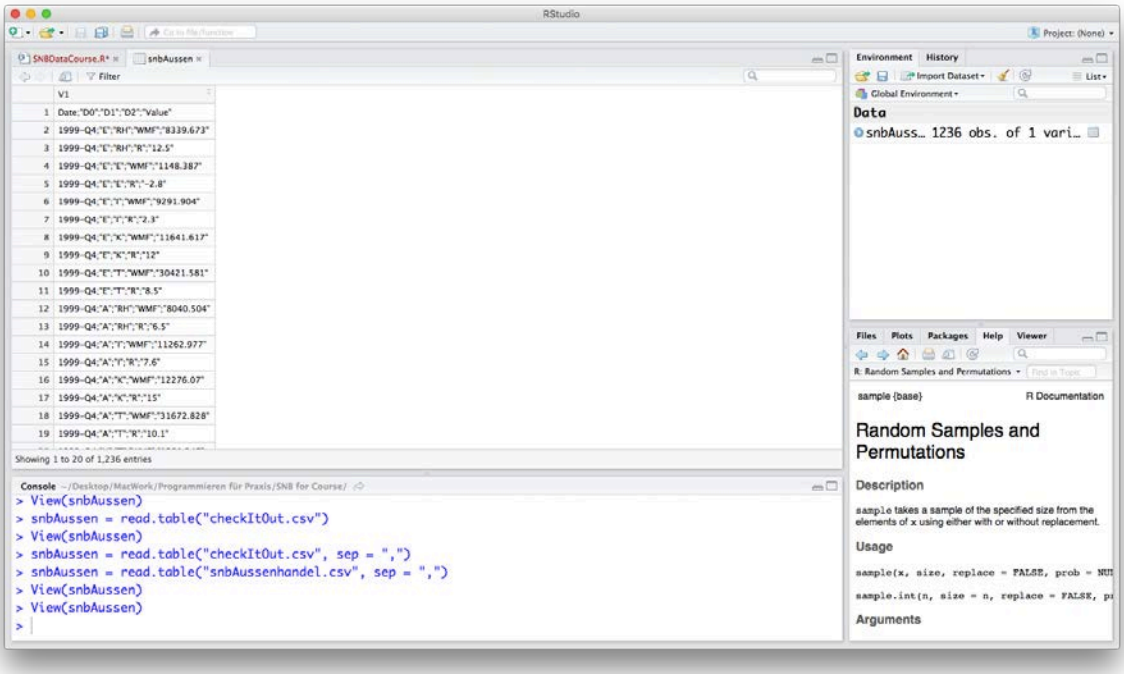
Getting messy-looking csv files in order

Data distribution platforms like the OECD or the Swiss National Bank put their files in a particular “format”. If the regional setting of your device corresponds to a different format, then the downloaded csv files may look pretty ugly, like so (or even uglier):

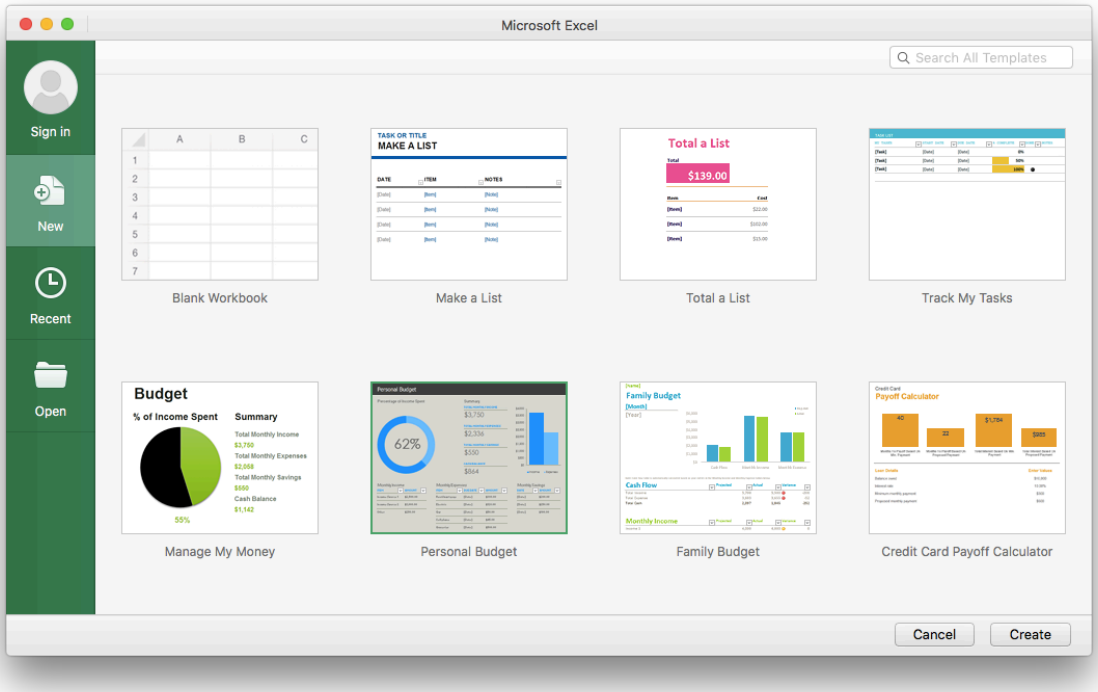


And when you try to read them into R, it looks like this:

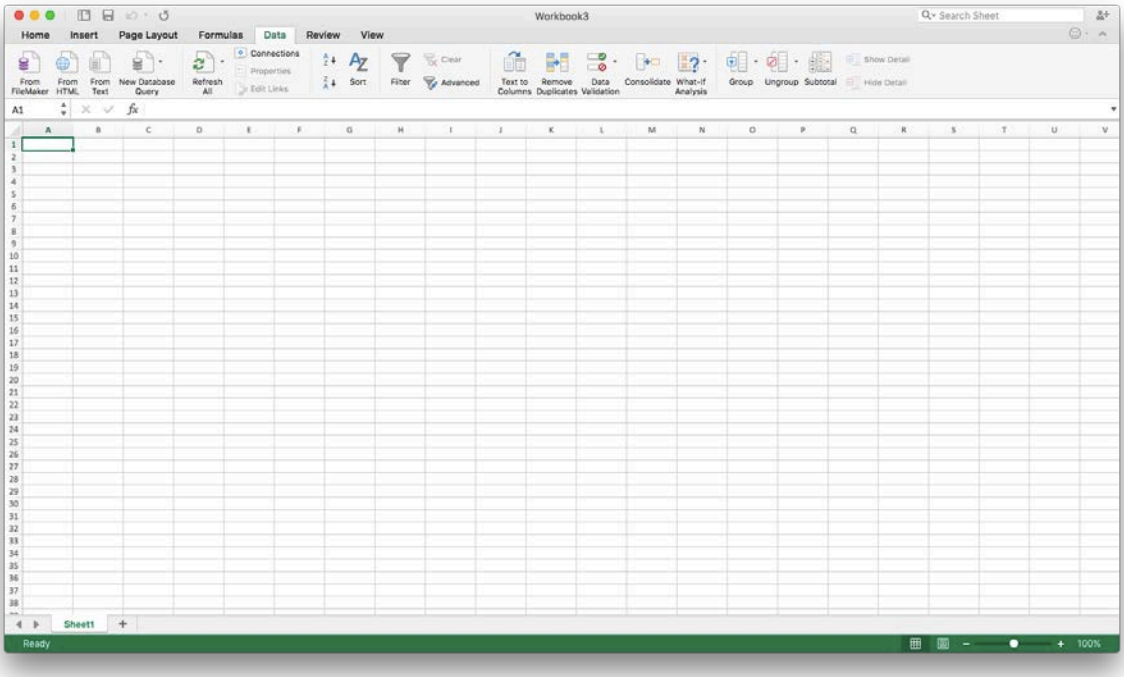


Sometimes, it helps to adjust the regional settings of your device (as in the case of my Windows 7 machine), but sometimes not (as in the case of my Mac). So what can you do? Below you find some screenshots that give you instructions how to get a clean-looking and workable csv file that you can read into R.
I first show this on a Mac (it works the same on Windows), and then I show you a different trick that sometimes works on a Windows machine

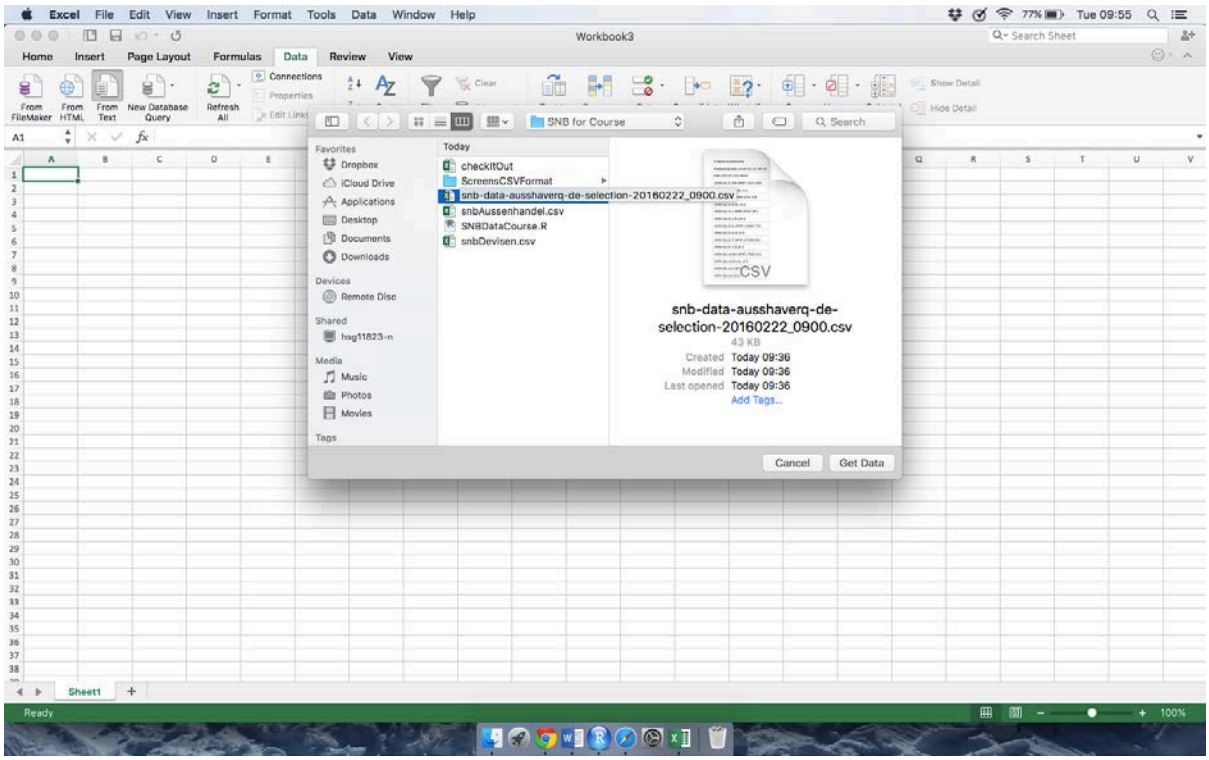
Open Excel



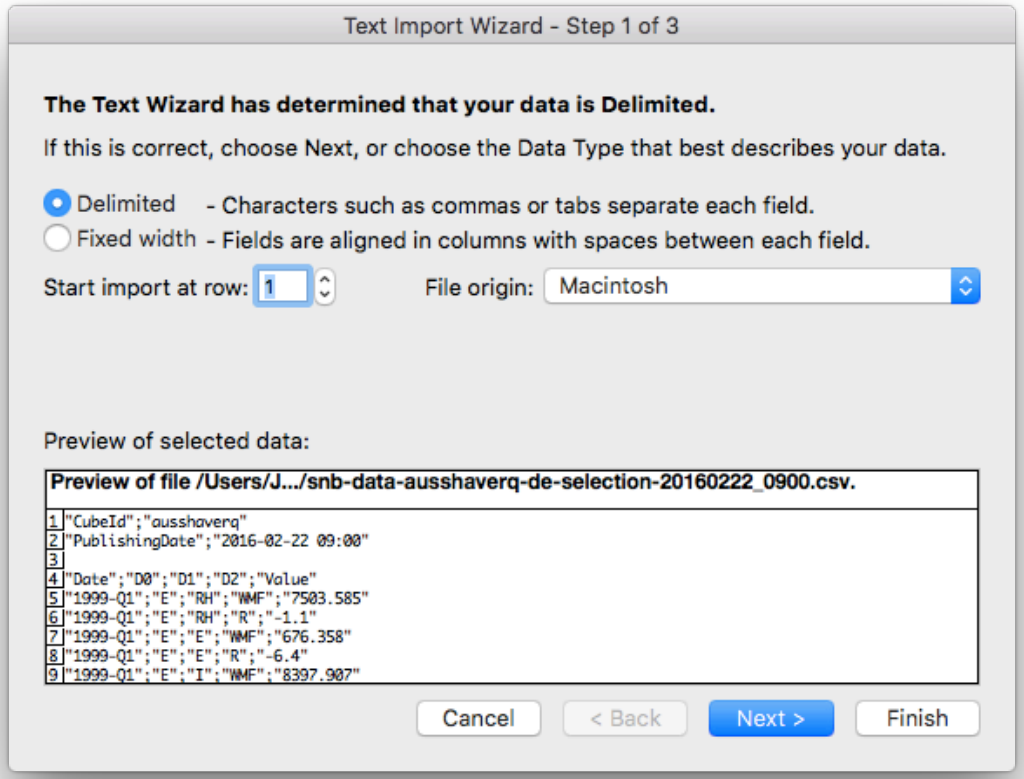
Open a new empty Workbook and go to the Data tab.



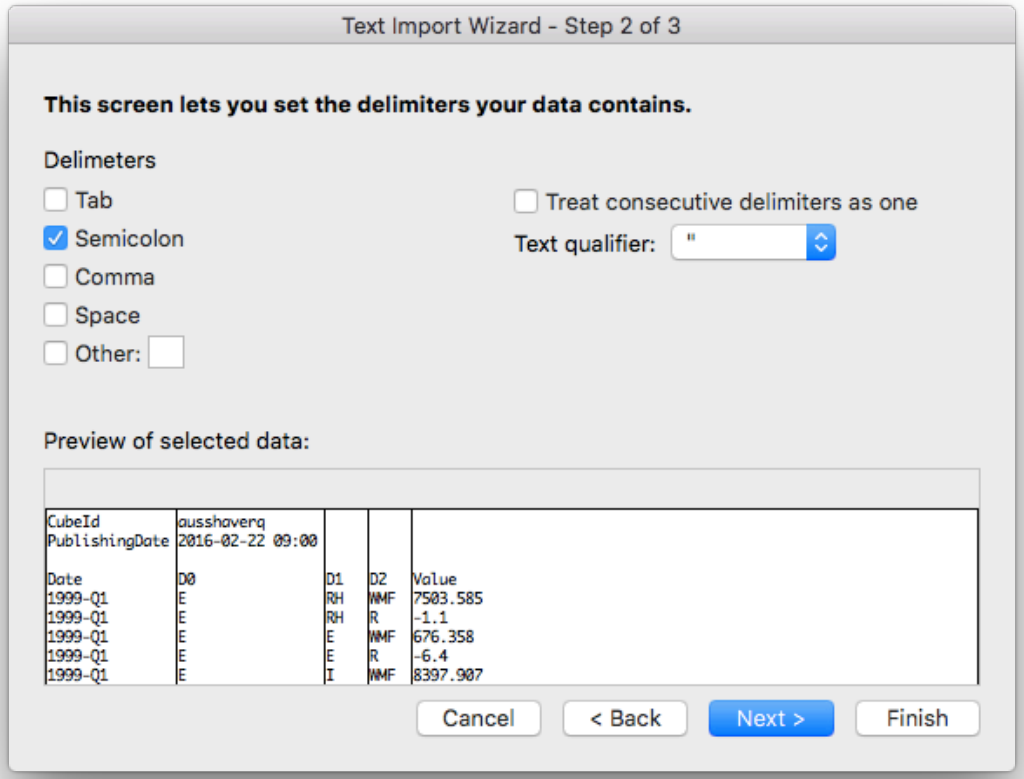
Click “From Text”



This opens the important wizard… Then just follow the steps you see on the below screens.


Text Import Wizard - Step 1 of 3
The Text Wizard has determined that your data is Delimited.
If this is correct, choose Next, or choose the Data Type that best describes your data.
Delimited - Characters such as commas or tabs separate each field.
Fixed width - Fields are aligned in columns with spaces between each field.
Start import at row: 1
File origin: Macintosh
Preview of selected data:
Preview of file /Users/J.../snb-data-ausshaverq-de-selection-20160222_0900.csv.
1 "CubeId";"ausshaverq"
2 "PublishingDate";"2016-02-22 09:00"
3
4 "Date";"D0";"D1";"D2";"Value"
5 "1999-Q1";"E";"RH";"WMF";"7503.585"
6 "1999-Q1";"E";"RH";"R";"-1.1"
7 "1999-Q1";"E";"E";"WMF";"676.358"
8 "1999-Q1";"E";"E";"R";"-6.4"
9 "1999-Q1";"E";"I";"WMF";"8397.907"
Cancel
< Back
Next >
Finish

In the case of the Swiss National Bank data, the solution is most likely “semicolon”. I sometimes encounter the situation that I additionally need to check “Other” and insert double quotes in the field next to other (this leads to a lot of empty columns, but it is easy to delete them in R). When downloading data from another source (say from the OECD), you may need a different trick. Some willingness to play around is required here… Just check it out what works.



Text Import Wizard - Step 3 of 3

This screen lets you select each column and set the Data Format.

Column data format

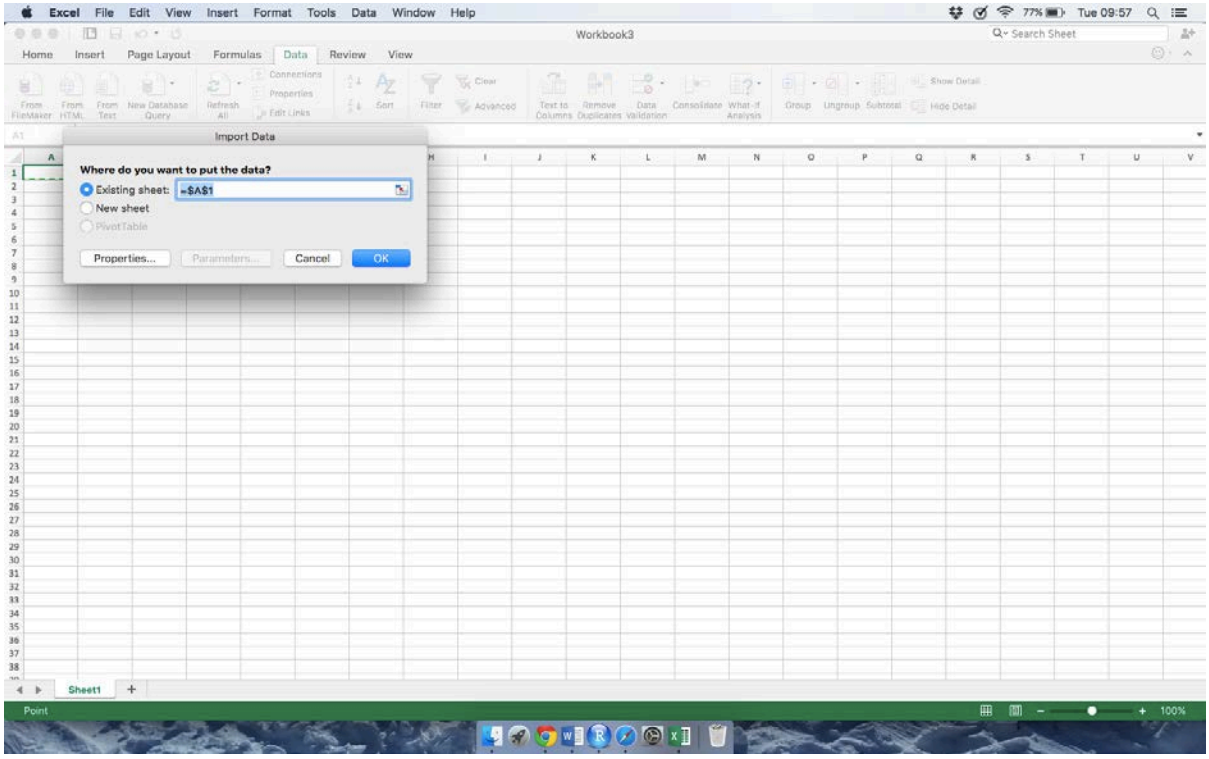
- (•) General
- ( ) Text
- ( ) Date: DMY
- ( ) Do not import column (Skip)

Advanced...

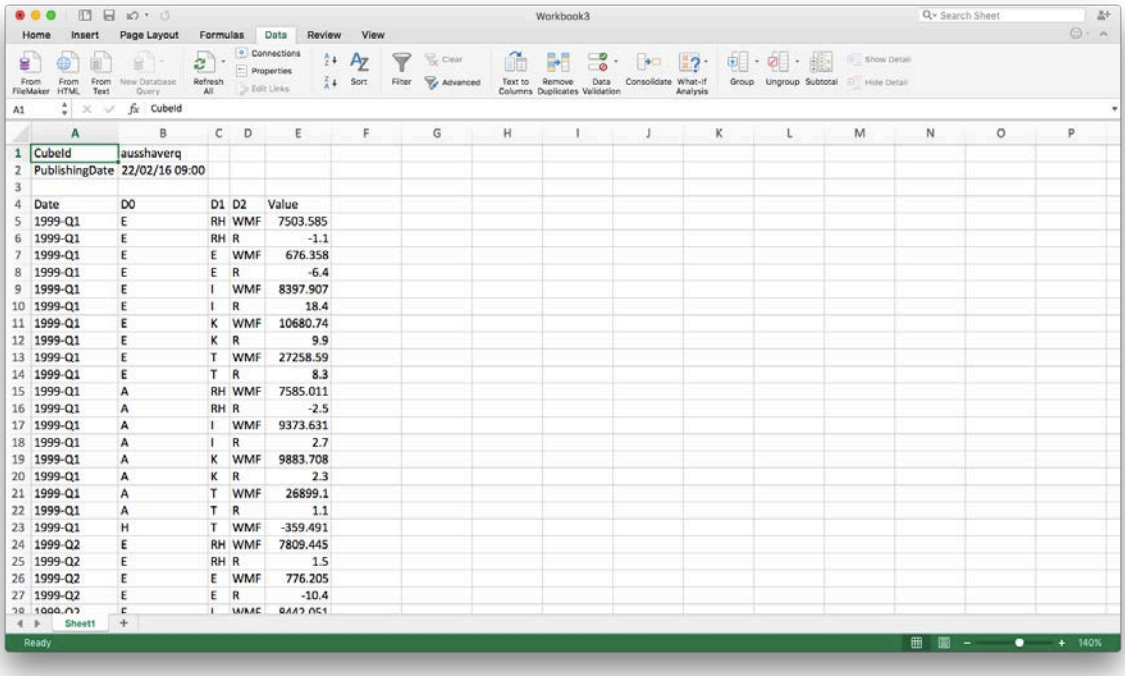
Preview of selected data:

General	General	Gener	Gener	General
CubeId	ausshaverq			
PublishingDate	2016-02-22 09:00			
Date	D0	D1	D2	Value
1999-Q1	E	RH	MMF	7503.585
1999-Q1	E	RH	R	-1.1
1999-Q1	E	E	MMF	676.358
1999-Q1	E	E	R	-6.4
1999-Q1	E	I	MMF	8397.907

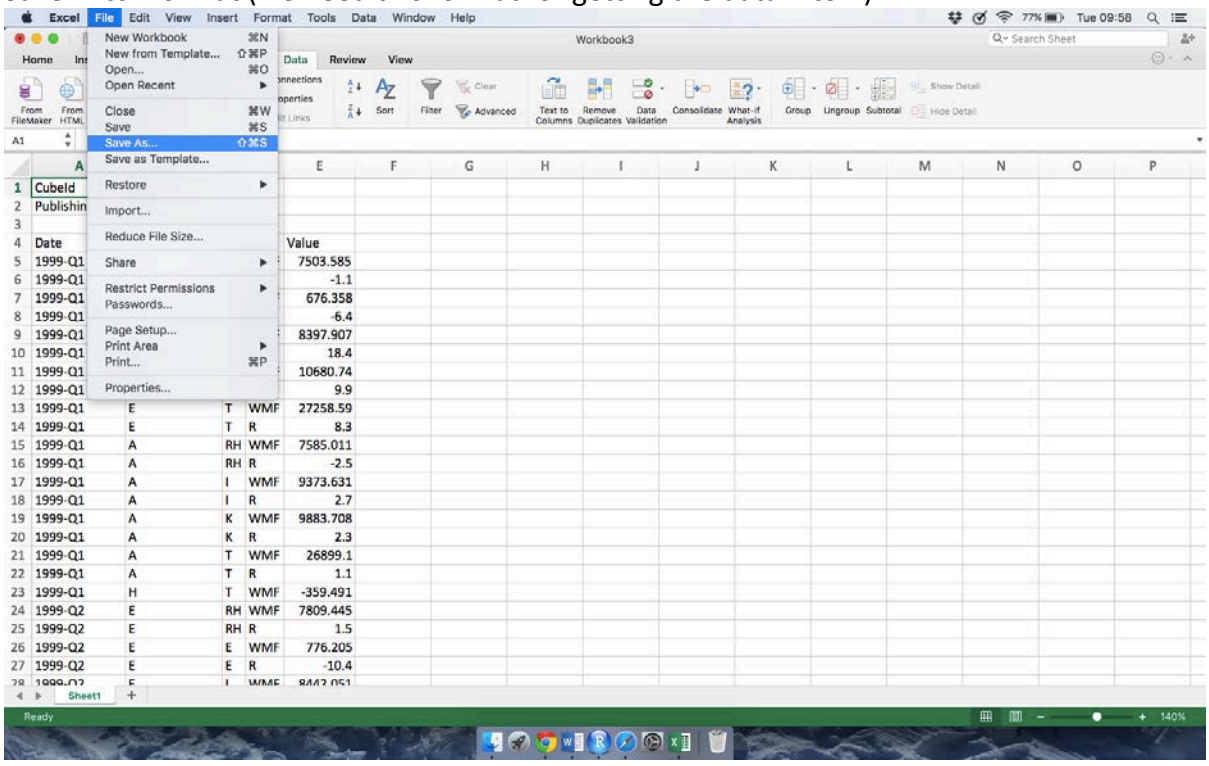
Cancel | < Back | Next > | Finish

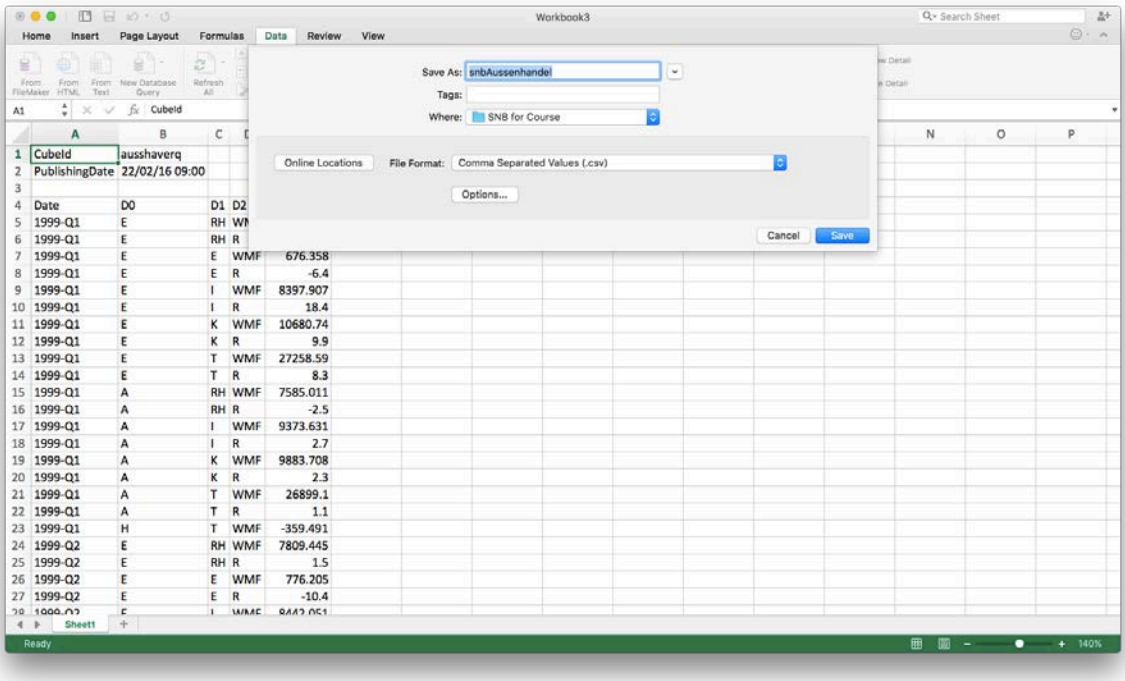


Now it looks neat…

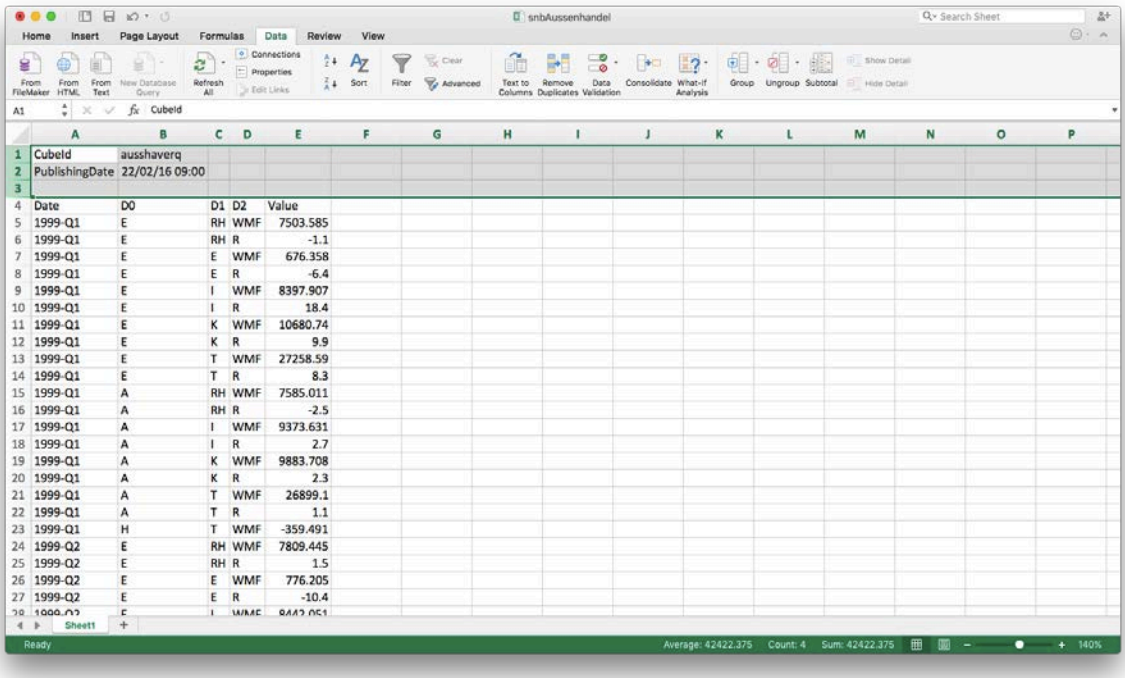


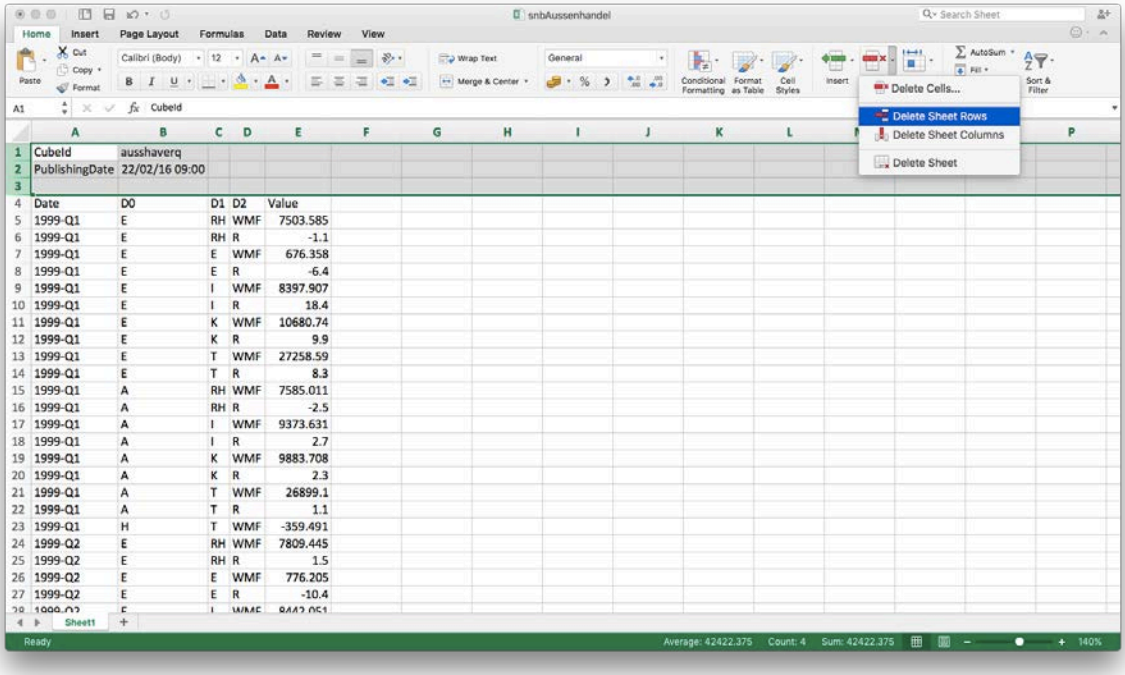
Save in csv format (we need this format for getting the data into R)



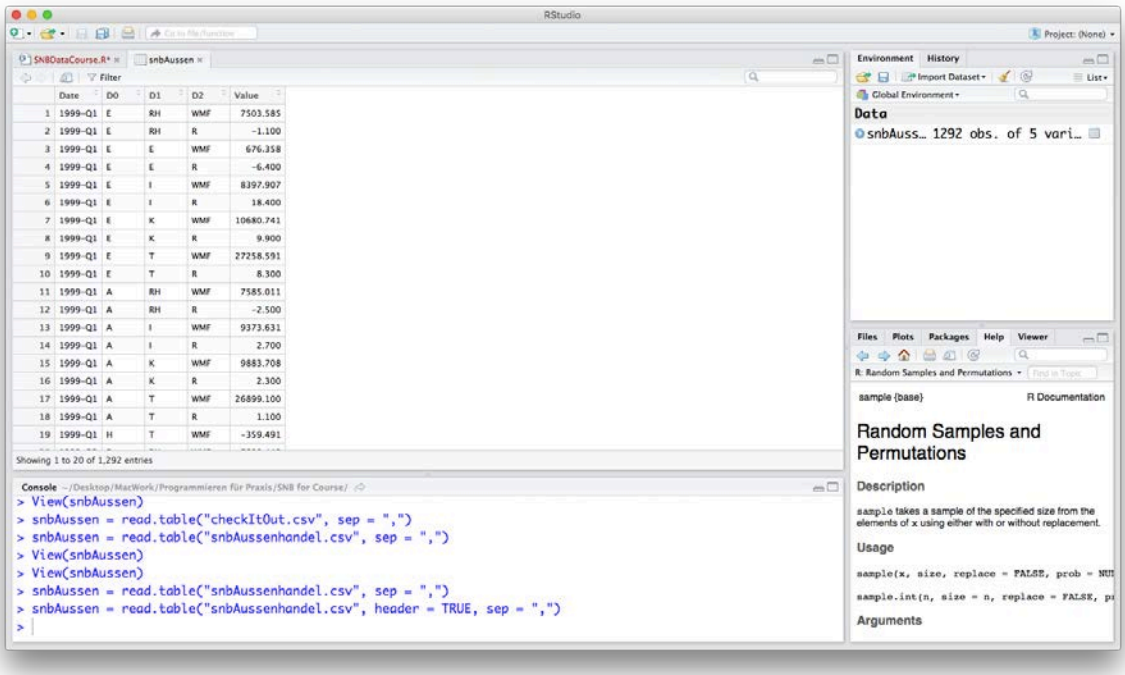


For reading the data into R, we need to delete the first three rows, as they would not make sense for R.



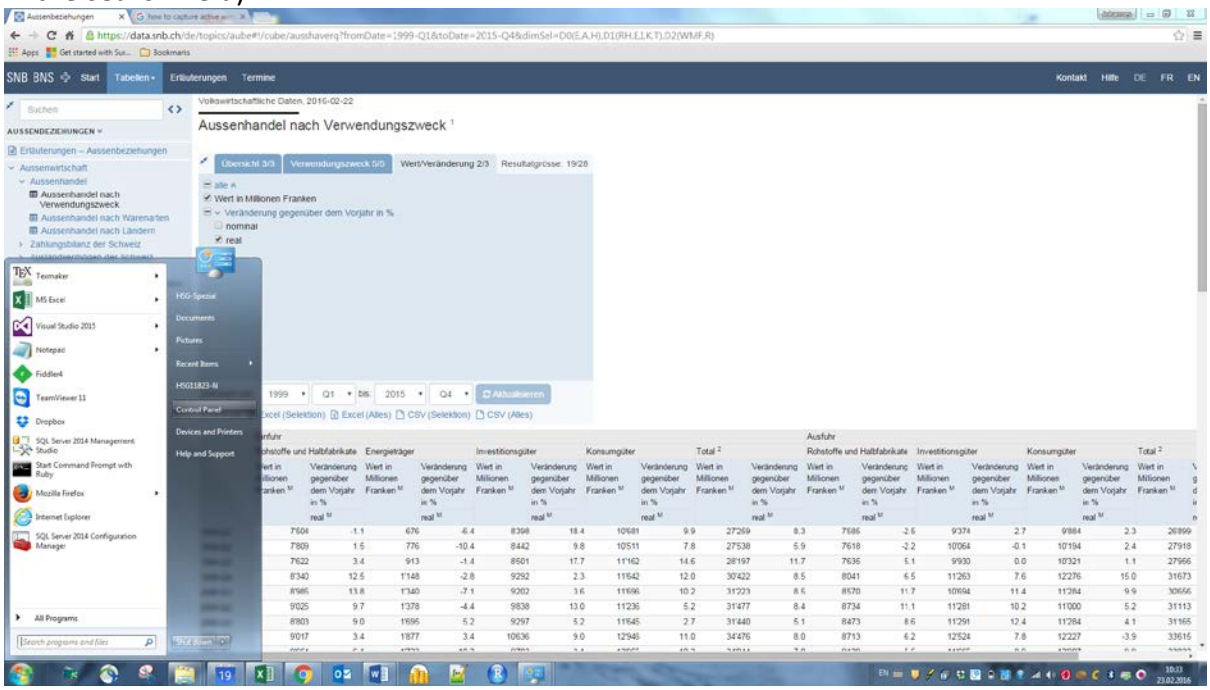


We'll read the data into R in class, I'll show you the details. But this is how it looks like. Finally, it's neat!

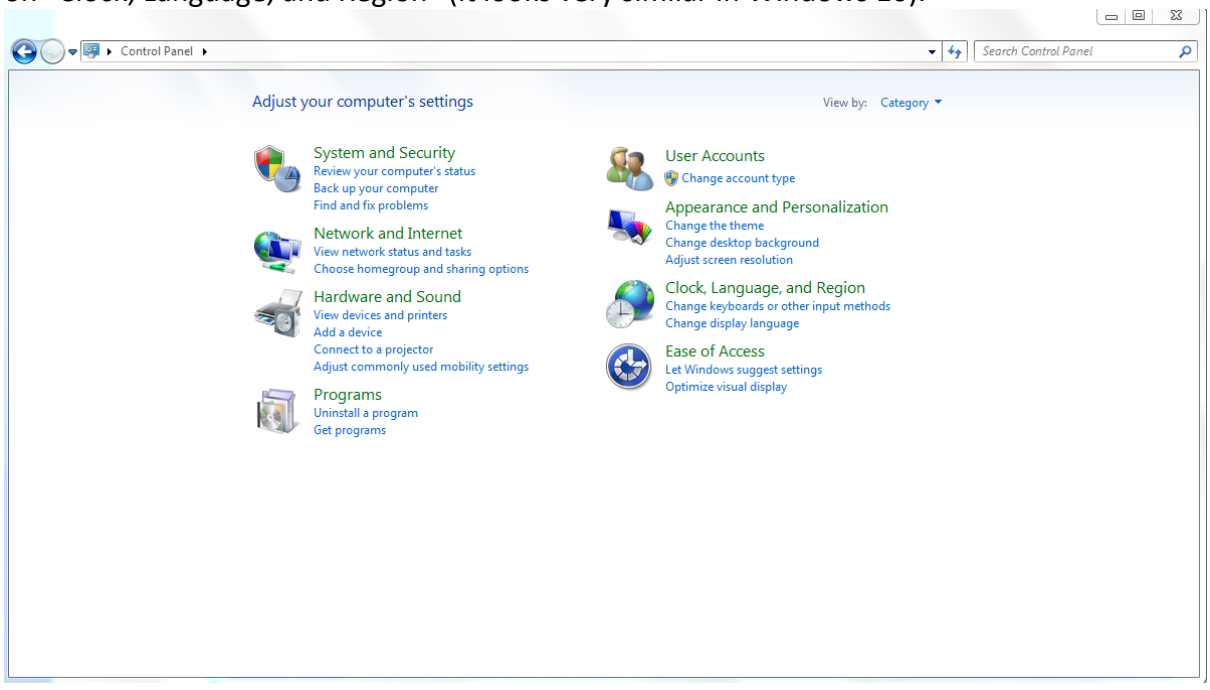


A different trick that works on my windows machine

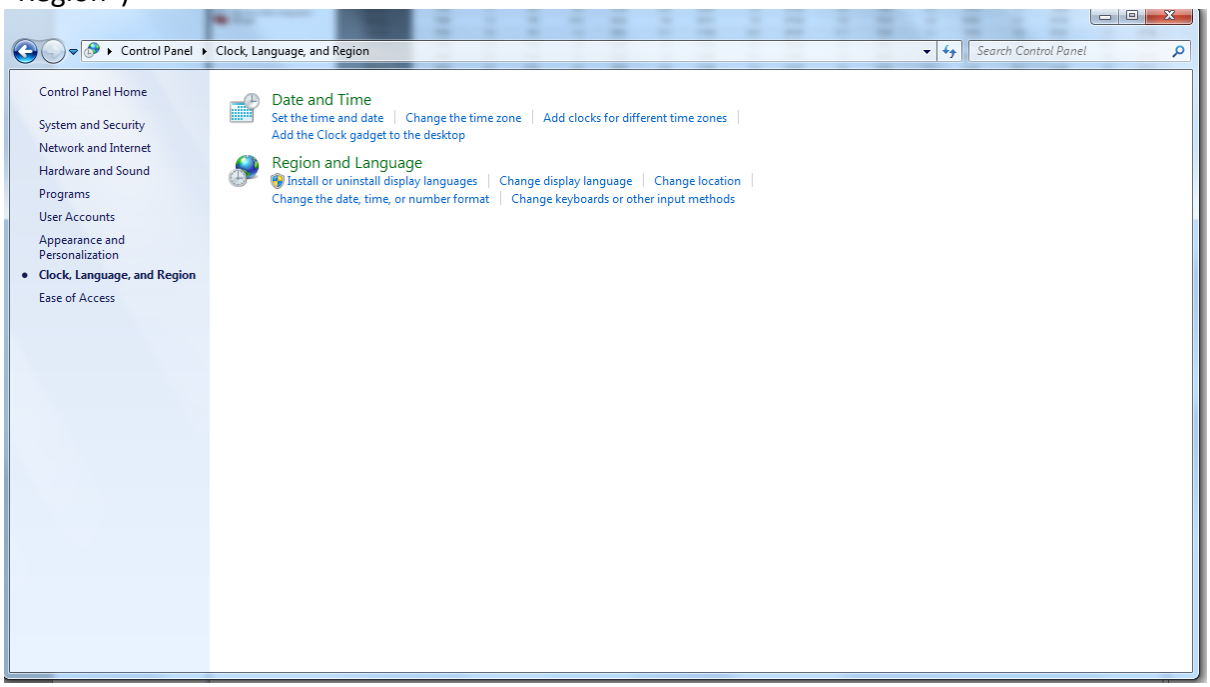
Go to the start button and click on Control Panel (in Windows 10, just type “control panel” in the search field)



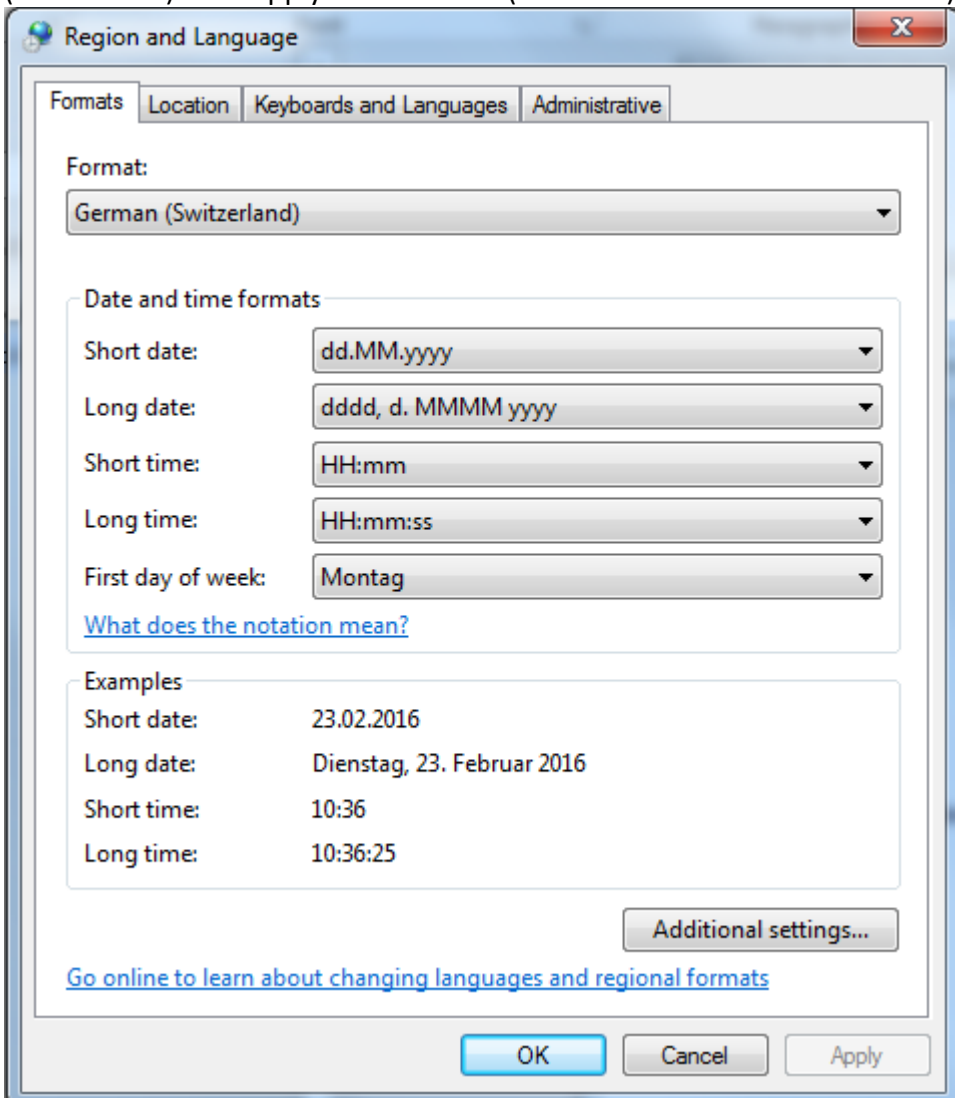
You get the next screen (probably); otherwise, navigate there using the navigation bar. Click on “Clock, Language, and Region” (it looks very similar in Windows 10).



You'll get the following screen. Click on "Region and Language" (in Windows 10, click on "Region")



For a successful download of data from the Swiss National Bank, choose “German (Switzerland)”. Hit Apply and then OK. (It looks the same in Windows 10.)



On my Windows 7 or 10 machines, that’s all I have to do get neat csv files from the SNB.

If you want to download data from the OECD and you do so with Swiss regional settings, it will again look ugly. Then you will have to set your regional settings to, for instance, “English (United States)”.