TWINNING CONTRACT

BA 15 IPA SR 01 17

Support to the reform of the statistics system in Bosnia and Herzegovina

MISSION REPORT

Component 3: Balance of Payment and IIPs

Activity 3.11: Establishing dashboard for CBBH and training of staff

Mission carried out by

Morten-Bo Paulsen, Danmarks Nationalbank

4 - 15 November 2019

Version: Final

STATISTICS DENMARK

Insee Institut national de la statistique et des études économiques Mesurer pour comprendre

Statistics Finland

CROATIAN BUREAU OF STATISTICS

Expert contact information

Morten-Bo Paulsen
Danmarks Nationalbank
Copenhagen, Denmark
Email: Morten-Bo.Paulsen@ecb.europa.eu
Phone: +45 60 59 82 86

Table of contents

List of Abbreviations

BHAS	Agency for Statistics of Bosnia and Herzegovina
BiH	Bosnia and Herzegovina
CBBH	Central Bank of Bosnia and Herzegovina
EC	European Commission
EU	European Union
FBiH	Federation of Bosnia and Herzegovina
FIS	Institute for Statistics of Federation of Bosnia and Herzegovina
MS	EU Member State
RSIS	Institute for Statistics of Republika Srpska
RTA	Resident Twinning Adviser
ToR	Terms of Reference

1. General comments

This mission report was prepared within the EU Twinning Project ” Support to the reform of the statistics system in Bosnia and Herzegovina”.

The purposes of the mission were:

- **Adjust dashboard for CBBH based on Danish Central Bank system**
- **Teach and advise CBBH staff on maintenance and development going forward**
- **Discussions and recommendations for future work**

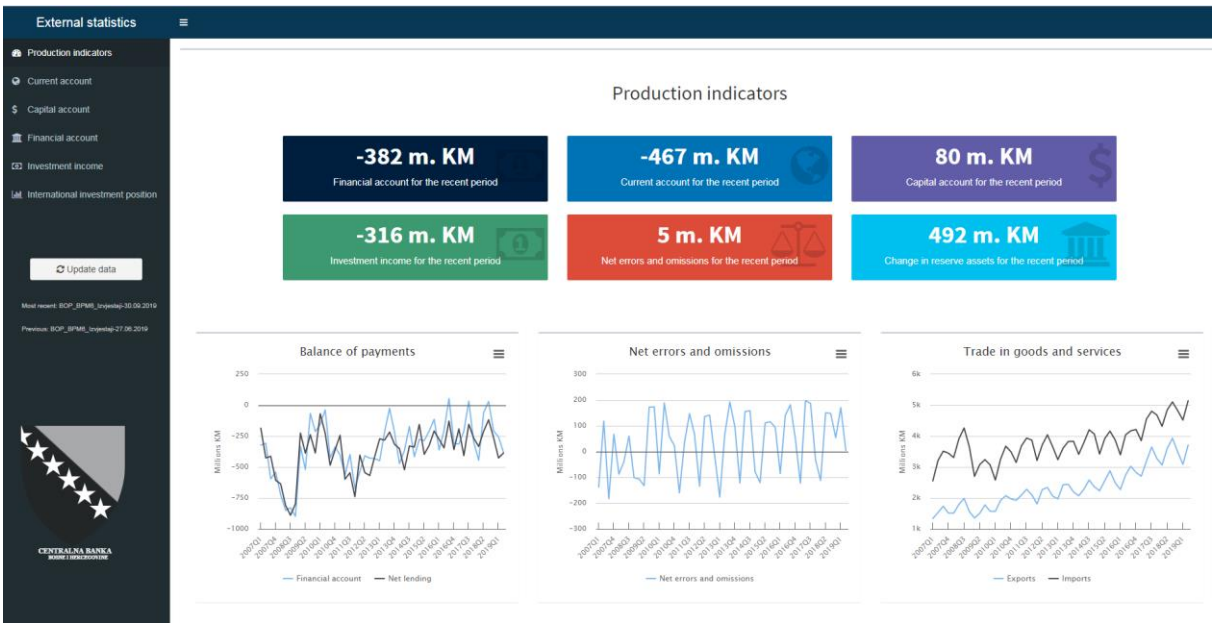
The consultant would like to express his thanks to all officials and individuals met for the kind support and valuable information that was received during the stay in Bosnia-Herzegovina and which highly facilitated the work of the consultant.

This views and observations stated in this report are those of the consultant and do not necessarily correspond to the views of EU, BHAS, FIS, RSIS, CBBH, Statistics Denmark, Danmarks Nationalbank, INSEE, Statistics Finland and Croatian Bureau of Statistics.

2. Assessment and results

Adjust dashboard for CBBH based on Danish Central Bank system:

A dashboard was successfully created based on the codes and interfaces that are used in the Danish Central Bank. In the process, valuable discussion of data compilation and dissemination as well as data structures and IT resources were discussed with staff members. Most work was then focused on transforming the current data sources to a database-like structure that is necessary in order to make the reading of data and subsequent coding of the dashboard convenient.



After achieving a database for balance of payments data as well as data for the international investment position, a dashboard was developed containing interactive charts and tables to display trends in the current account, capital account, financial account, investment income and international investment position. The charts allow the user to select between different instruments, to display series on a credit/debit or asset/liability basis, as well as other customization choices. In addition, the dashboard calculates the differences between two versions of the data, which allow the user to keep track of data revisions.

Teach and advise CBBH staff on maintenance and development going forward:

The staff has been informed about the prerequisites for the dashboard to run. In addition, there has been teaching of how data is transformed and read into the dashboard. There has also been discussion about the version control of packages and IT staff has been involved in order to facilitate that the dashboard can run on a local IP-address. A guide has been drafted so as to assist staff in maintaining the dashboard onwards (see Annex 3).

3. Conclusions and recommendations

In conclusion, the mission successfully achieved its objectives. The results were presented to the section with the participation of the head of the statistics department as well as vice governor of the respective sector. It is expected that the staff will be able to maintain and develop further on the dashboard that has been set up. It is recommended, that the staff ensures that the dashboard can continue to run in its current format from the equipment that is currently available. It is advised that staff become familiar with the way that data is transformed and read into the dashboard as described in the guide (see annex 3). Staff may further seek out further education in R-programming, in particular to become familiar with the "shiny package". Furthermore, it is recommended that staff takes into consideration the possibility of running the dashboard on a server rather than from a local IP-address. As the staff becomes more familiar with R-programming, it can be recommend to seek out assistance in terms of using code collaboration tools such as GitLab.

Actions needed for moving forward as well as for preparing the next mission –add rows as needed.

Action	**Responsible person**
Coordinate with IT department whether dashboard can continue to run on an IP-address	CBBH
Ensure that necessary software and checkpoint of version control are set up (cf. dashboard guide)	CBBH
Ensure that excel sheets maintain same format (same SDMX and instrument codes) and that they are located in correct folders	CBBH
Seek further education in R-programming, mainly how to use the "shiny" package	CBBH

Annex 1. Terms of Reference

Terms of Reference

EU Twinning Project BA 15 IPA ST 01 17

Component 3: Balance of Payment and IIPs
04 - 15 November 2019

Activity 3.11: Establishing dashboard for CBBH and training of staff

Annex 2. Persons met

Central Bank of Bosnia and Herzegovina (CBBH)

- Zelimira Raspudic, Vice Governor
- Amir Hadžiomeragić, Head of Statistics and Publications Department
- Vedran Milisav, Head of BoP statistics section
- Aida Kalco, Economist/Statistician in BoP statistics section
- Masa Kamenica, Economist/Statistician in BoP statistics section

MS Experts

- Morten-Bo Paulsen, Danmarks Nationalbank

Twinning Project Administration

- Katja Møller Hjelvang, RTA
- Đemka Šahinpašić, RTA Assistant

Annex 3. Dashboard guide

Case no.: 190134
Document no.: 1956955

12th. November 2019

CBBH - dashboard guide

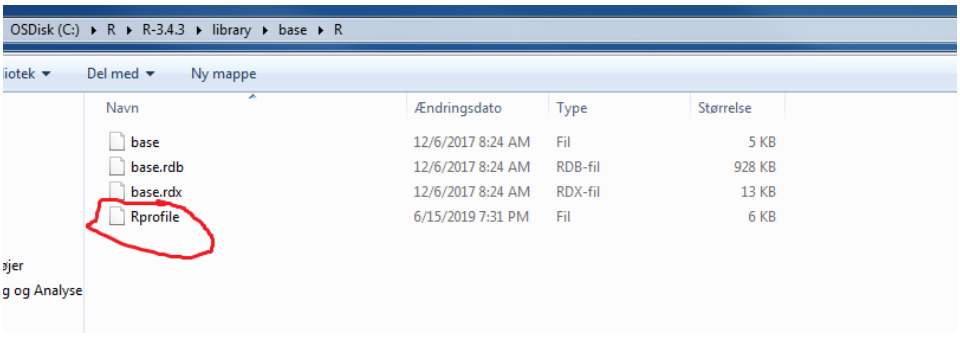
Morten-Bo Paulsen
Financial Statistics
Balance of payments and securities statistics

Prerequisites to run dashboard:

You need Rstudio and R installed (preferably version 3.4.3 due to bugs in newer versions). It is convenient to have both programs installed in the same folder (e.g. Programs folder on your C-drive). After the installation it is convenient to set at checkpoint date in the RProfile-file. This ensures that R always loads the same version of packages (if packages change some codes may not work).

Do the following:

Open the RProfile-file from your installation folder of R (the base-folder in the library folder):



Open the folder in Notepad and add the following to the end of the file:
checkpoint::setSnapshot("2018-03-01"):

```
local({
    br <- Sys.getenv("R_BROWSER", NA_character_)
    if(!is.na(br)) options(browser = br)
    tests_startup <- Sys.getenv("R_TESTS")
    if(nzchar(tests_startup)) source(tests_startup)
    ca_cert <- Sys.getenv("CURL_CA_BUNDLE", NA_character_)
    if(!is.na(ca_cert) &&
       file.exists(ca_path <- file.path(R.home("etc"), "curl-ca-bundle.crt")))
        Sys.setenv(CURL_CA_BUNDLE = ca_path)
})

#Sætter standardbibliotek til pandoc
Sys.setenv(RSTUDIO_PANDOC = "C:/Program Files/RStudio/bin/pandoc")
#Sætter repository til 1. marts 2018
checkpoint::setSnapshot("2018-03-01")
```

Next we need to ensure that all required packages are installed:
The first package to be installed is the one used for the "checkpoint" at it should be installed like this:

```
1
2
3  install.packages("checkpoint", repos = "http://cran.r-project.org")
4
5
6
```

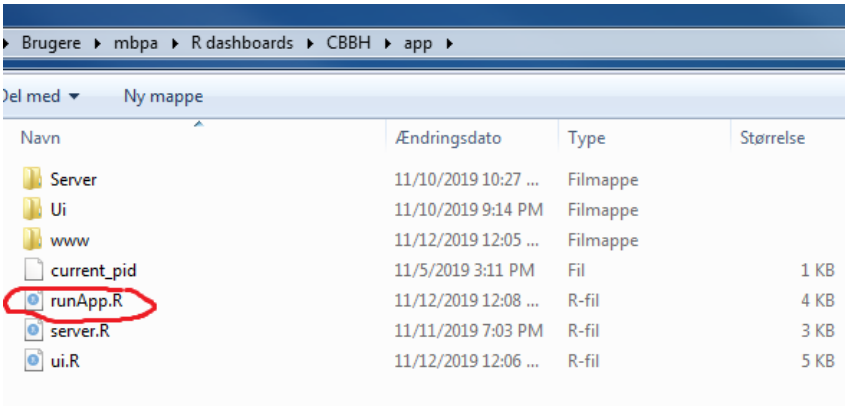
Now we have ensured that all packages that we install will always stick to the version as of 1th March 2018.

We can now install the remaining packages:

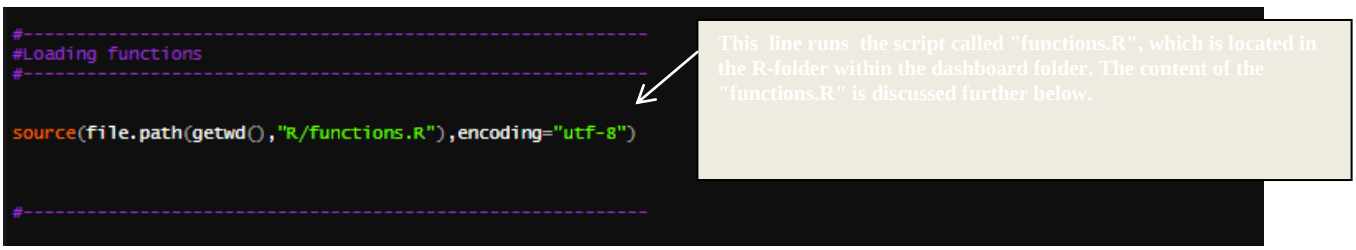
```
2
3  install.packages("highcharter")
4  install.packages("tidyverse")
5  install.packages("rvest")
6  install.packages("stringr")
7  install.packages("forcats")
8  install.packages("janitor")
9  install.packages("readr")
0  install.packages("lubridate")
1  install.packages("magrittr")
2  install.packages("dplyr")
3  install.packages("bizdays")
4  install.packages("tidyr")
5  install.packages("shiny")
6  install.packages("shinydashboard")
7  install.packages("shinyjs")
8  install.packages("shinyWidgets")
9  install.packages("DT")
0  install.packages("data.table")
1  install.packages("DescTools")
2  install.packages("lazyeval")
3  install.packages("readxl")
4  install.packages("countrycode")
5  install.packages("RColorBrewer")
6  install.packages("kableExtra")
7  install.packages("sparkline")
8  install.packages("formattable")
9  install.packages("viridisLite")
0  install.packages("zoo")
1
```

How to launch dashboard:

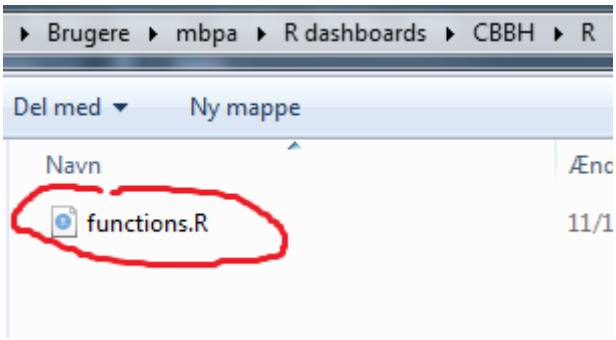
The CBBH dashboard is launched by executing all the lines in the script called "runApp.R", which is located in the "app" folder



The execution of the lines in the "runApp.R" script does the following:



Location of the "functions.R" script:




```
#------------------------------------------------------------
#Retrieve  IP adress from local PC
#------------------------------------------------------------
shell.exec(file.path(getwd(),"findip.bat"))

# Pause for 3 seconds, so that  ip adress can be updated
Sys.sleep(3)

#Read ip adress for dashboard
my_ip <- read.table(file.path(getwd(),"my_ip.txt"))

#------------------------------------------------------------

#------------------------------------------------------------
#START DASHBOARD
#------------------------------------------------------------
runApp(file.path(getwd(),"app"),launch.browser = TRUE,host = as.character(my_ip$V1), port = 5768)
```

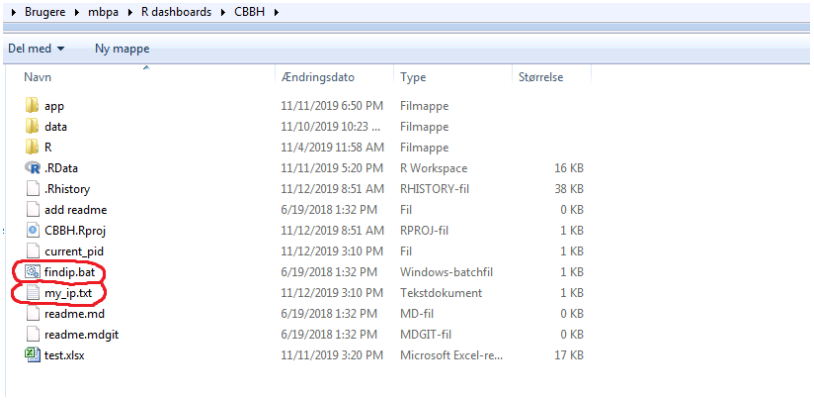
This line executes a bat-file that is located in the dashboard folder. The bat-file retrieves a valid IP-address from the local PC and saves it in a text-file called "my_ip.txt"

Execution of the script is paused so that the IP-address has time to update

The saved IP-address is read in from the local file called "my_ip.txt"

By running the final line, the dashboard is launched on the IP-address that was created in previous steps along with a specified port (should be an open port). The dashboard opens in the default browser (Chrome provides the best user-experience). All users (internal and external) will be able to access the dashboard from the IP-address it has been launched on).

Location of the files "findip.bat" and "my_ip.txt":



Contents of the "functions.R" script:

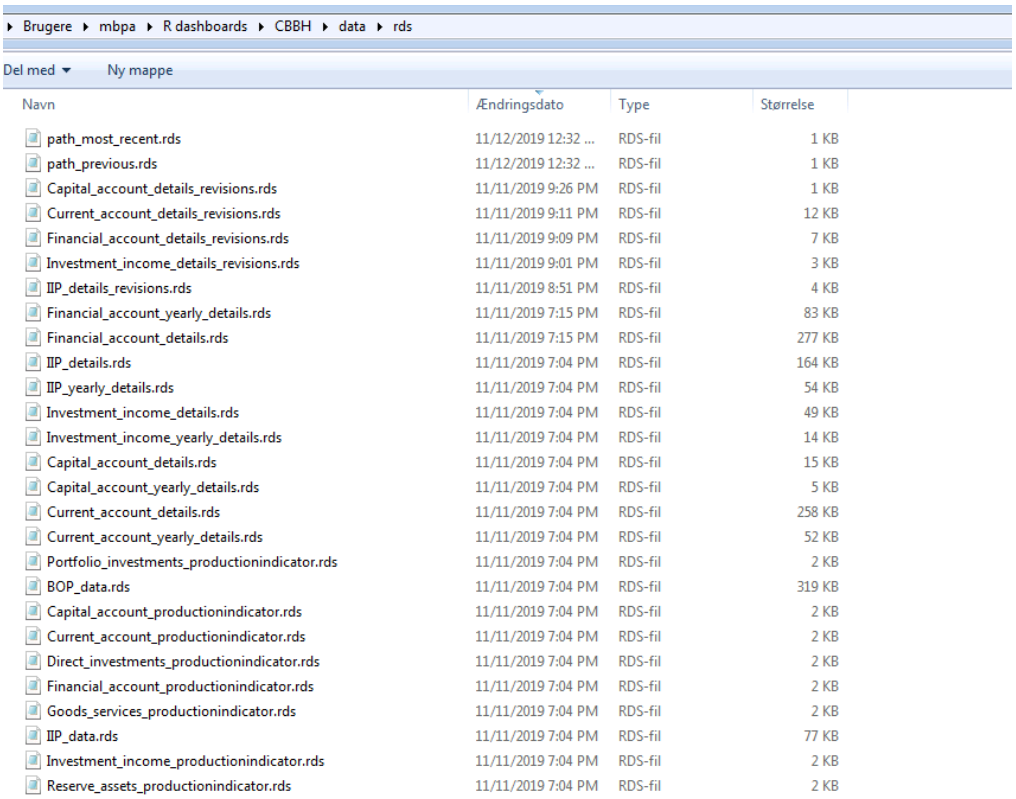
The script contains pre-defined settings, functions and values that are convenient to refer to rather than having to write the same codes multiple times. The contents of the "functions.R" script are as follows:

```
#--------------------
#Functions
#--------------------


##################################################

# Set language settings
 Sys.setlocale("LC_ALL", "English")


# Folder for rds files
save_dir <- file.path(getwd(),"data/rds")

if (!dir.exists(save_dir)){
  dir.create(save_dir)
}
```

Specifies that language options are set to "English" such that e.g. date formats follow the English standard

Specifies the path for the folder in which RDS-files (saved datasets that have been created in R) are stored (if folder not already exists, it is created automatically)

Location of the RDS-files (path is defined by "save_dir"):



```
#-----------------------------------------------------
# Paths for datafiles
#-----------------------------------------------------
path_most_recent<-list.files("C:/Users/mbpa/R dashboards/BPM6 IMF reporting",pattern='*.xls.*')
path_previous<-list.files("C:/Users/mbpa/R dashboards/archives/",pattern='*.xls.*')
```

Specifies the paths for the location of the excel files that all data input is based on. Whichever excel-file (file has to end with .xls. or .xlsx) is located in these paths will be read in, so ideally there should only be one excel-file in each folder. "path most recent" should be the folder where the most recent excel-file is located and "path previous" should be the location of the excel-file that the most recent file is compared against (e.g. revisions will be calculated as the difference between these two files)

Example of how "path_most recent" and "path_previous" points to the relevant excel-files:

```
> path_most_recent
# A tibble: 1 x 1
  file
  <chr>
1 C:/Users/mbpa/R dashboards/BPM6 IMF reporting/BOP_BPM6_Izvjestaji-30.09.2019.xlsx
> path_previous
# A tibble: 1 x 1
  file
  <chr>
1 C:/Users/mbpa/R dashboards/archives/BOP_BPM6_Izvjestaji-27.06.2019.xls
```

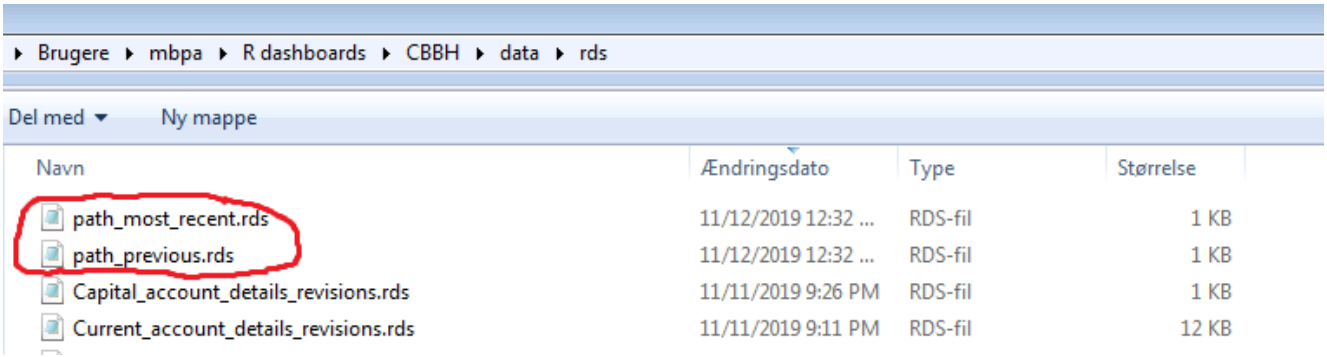
Most recent excel-file

Previous version of excel-file

```
#Saving dataset in rds format for reactive reading in dashboard session
saveRDS(path_most_recent,paste(save_dir,"path_most_recent.rds",sep = "/"))

#Saving dataset in rds format for reactive reading in dashboard session
saveRDS(path_previous,paste(save_dir,"path_previous.rds",sep = "/"))
```

Path-names from previous step are further saved as RDS-files so that the names can be displayed in dashboard sidebar (see description further below)

RDS-files with relevant path names are saved to the RDS-folder (as defined by "save_dir"):



```r
#Basic trim functions

left <- function(s, amount){
  return(substr(s, 1, amount))
}

right <- function(s, amount){
  return(substr(s, nchar(s)-amount+1, nchar(s)))
}

mid <- function(s, offset, amount){
  return(substr(s, offset, offset+amount-1))
}
```

Basic trim-functions are created that can be used to conveniently extract parts of a string variable (se example below)

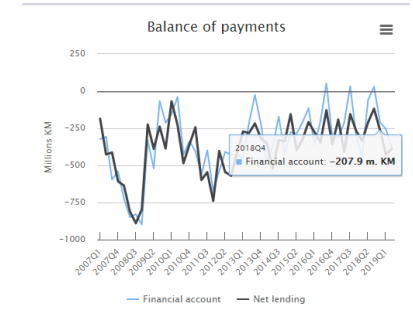
Example of trim-function "left":

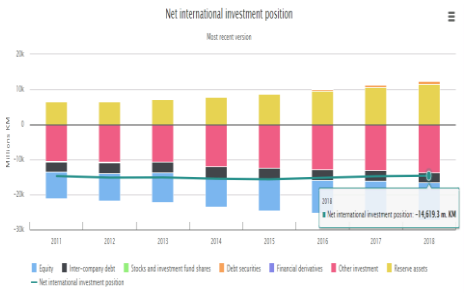
```r
> left("2019Q1",4)
[1] "2019"
>
```

```r
# Defines the interface used for charts

hc <- function(y1.label="Bn. KM",align="high",y2.label="Bn. KM",kolonnetype =NULL) {
  highchart() %>%

    hc_xAxis(tickColor=c("#666666"),tickPosition='inside',lineColor=c("#666666") ,labels = list(style=list(fontFamily="Lucida Sans Unicode",fontsize=6.5,color=c("#666666"))))%>%

    hc_exporting(enabled = TRUE)%>%

    hc_legend(reversed=TRUE,

              symbolRadius=0,itemStyle = list(fontFamily = 'Lucida Sans Unicode',fontsize=6.5,fontWeight= 'Lucida Sans Unicode',color = c("#666666")))%>%

    hc_tooltip(pointFormat="<span style='font-family:Lucida Sans Unicode;color:{point.color}'> &#x25A0;</span> {series.name}: <b>{point.y:,.1f} bn. KM</b><br/>",style=list(fontFamily= "Lucida Sans Unicode",fontsize=6.5),useHTML=TRUE)%>%
    hc_chart(zoomType = "x") %>%
    hc_yAxis_multiples(
      list(gridLineColor=c("#CCCCCC"),startOnTick=TRUE,style=list(fontfamily="Lucida Sans Unicode",fontsize=6.5,color=c("#666666")),plotLines = list(
            list(color = "#666666",
                 width = 1.5,
                 value = 0,
                 zIndex= 5)),

        title = list(
          text = y1.label
        ),
        align = "high"
      ),
      list(
        title = list(
          text = y2.label
        ),
        align = "right",
        opposite = TRUE
      )) %>%
    hc_plotOptions(line = list(marker = list(enabled = FALSE)),
                   area = list(marker = list(enabled = FALSE)))
}
```

"highcharter-function" is defined such that charts in the dashboard have the same visual format.

Example of "highcharts" created with the "hc" function defined in the "functions.R" script as shown above



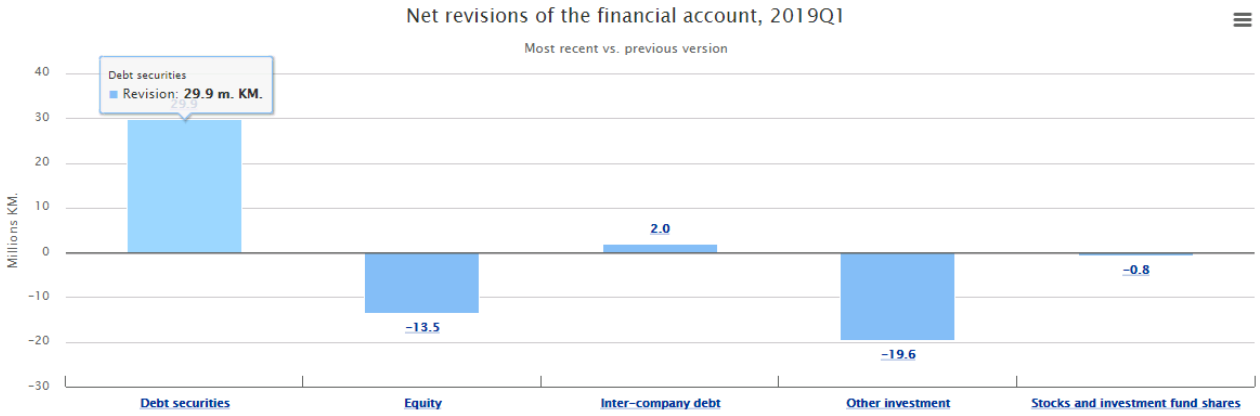


:

```
#Define function to drill-down chart

hc_drill_func <- function(idata,ilevels,ivalue,ititle,inavn,iyakselabel){
  iUpperLevel <- ilevels[1]
  upperdata <-
    idata %>%
    group_by_(interp(as.name(iUpperLevel)))%>%
    summarise_(y = interp(~ sum(x), x = as.name(ivalue))) %>%
    mutate_(drilldown = interp(as.name(iUpperLevel))) %>%
    ungroup() %>%
    rename_(name=as.name(iUpperLevel)) %>%
    select(name,y,drilldown)

  hc_drill <- highchart() %>%
    hc_chart(type = "column") %>%
    hc_title(text = ititle) %>%
    hc_xAxis(type = "category") %>%
    hc_yAxis(title = list (text= iyakselabel)) %>%
    hc_legend(enabled = FALSE) %>%
    hc_plotOptions(
      series = list(
        boderWidth = 0,
        dataLabels = list(enabled = FALSE),
        cropThreshold = 300
      )
    ) %>%
    hc_add_series(
      data = upperdata,
      name = inavn,
      colorByPoint = FALSE
    )
  DrillDownC <- list()
  for (j in 1:(length(ilevels)-1))
  {
    iUpperLevel <- ilevels[j]
    UpperLevel <- unique(idata[[ilevels[j]]])
    UpperLevelCount <- length(UpperLevel)
    if ( j == length(ilevels)){
      iLowerLevel <- ilevels[j]
    } else {
      iLowerLevel <- ilevels[j+1]
    }
    DrillDownList <- list()
    for (i in 1:UpperLevelCount)
    {
      DrillDownList[[i]] = list(name = UpperLevel[i],
                                id = UpperLevel[i],
                                data = list_parse(idata %>%
                                                    filter_(interp(~y == UpperLevel[i] ,y=as.name(iUpperLevel))) %>%
                                                    group_by_(as.name(iLowerLevel)) %>%
                                                    summarise_(y = interp(~ sum(x), x = as.name(ivalue))) %>%
                                                    mutate_(drilldown = interp(as.name(iLowerLevel))) %>%
                                                    ungroup() %>%
                                                    rename_(name=as.name(iLowerLevel)) %>%
                                                    select(name,y,drilldown)))

    }
    DrillDownC = c(DrillDownC,DrillDownList)
  }
  hc_drill <- hc_drill %>%
    hc_drilldown(
      allowPointDrilldown = TRUE,
      series = DrillDownC
    ) %>%
    hc_xAxis(type = "category")
}
```

Another "highcharter-function" is defined that makes it easy to create "drilldown" charts (e.g. as used for the revision charts. See explanation further below).

Example of "drilldown-chart" created with the "hc_drill" function defined in the "functions.R" script as shown above

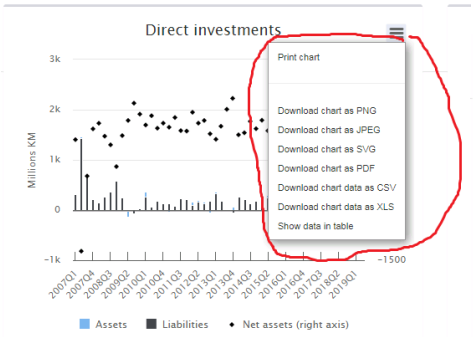


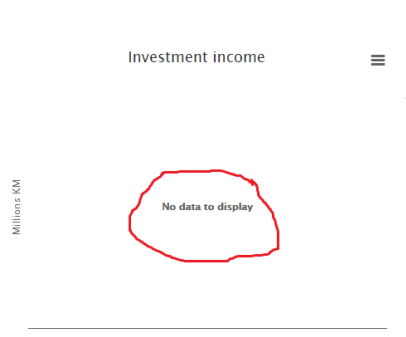
```
#Highcarts settings
hcoptslang <- getOption("highcharter.lang")
hcoptslang$printChart <- "Print chart"
hcoptslang$downloadPNG <- "Download chart as PNG"
hcoptslang$downloadJPEG <- "Download chart as JPEG"
hcoptslang$downloadSVG <- "Download chart as SVG"
hcoptslang$downloadPDF <- "Download chart as PDF"
hcoptslang$downloadCSV <- "Download chart data as CSV"
hcoptslang$downloadXLS <- "Download chart data as XLS"
hcoptslang$viewData    <- "Show data in table"
hcoptslang$noData      <-  "No data to display"
hcoptslang$drillUpText = "Back to {series.name}"
hcoptslang$resetZoom <-"Back"


hcoptslang$decimalPoint="."
hcoptslang$thousandsSep=","
options(highcharter.lang = hcoptslang)
```

Specifies different settings to determine what text should be displayed in charts (e.g. export-options).

Specifies that "." is decimal point and "," is thousand separator in tooltips in charts

Examples of how "hcoptlang" settings display in dashboard:

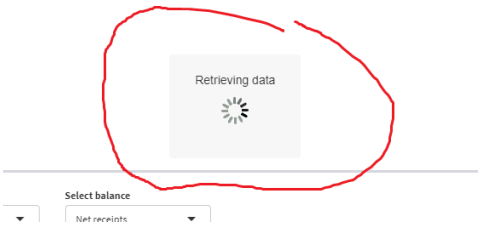


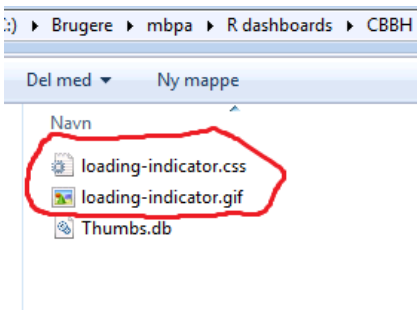


```
#Loading indicator for ui
addResourcePath("bpm6statviz",
                file.path(getwd(),"app/www"))

loading_indicator <- function(wait_text = "Retrieving data",
                              gif = "bpm6statviz/loading-indicator/loading-indicator.gif",
                              wait_time = 1000) {

  tagList(
    singleton(
      tags$head(
        tags$link(rel = "stylesheet", type = "text/css",
                  href = "bpm6statviz/loading-indicator/loading-indicator.css")
      )
    ),
    div(class = "loading-indicator", p(wait_text), img(src = gif)),
    tags$script(
      sprintf("setInterval(function(){
              if ($('html').hasClass('shiny-busy')) {
              setTimeout(function() {
              if ($('html').hasClass('shiny-busy')) {
              $('div.loading-indicator').show()
              }
              }, %d)
              } else {
              $('div.loading-indicator').hide()
              }
              }, 100)", wait_time)
    )
  )
}
```

Defines the "Retrieving data" animation that is shown when data is updating/charts are loaded. The function makes use of a gif-file called "loading-indicator.gif" and a css-file called "loading-indicator.css". Both files are located under "/app/www/loading-indicator" (as shown below)

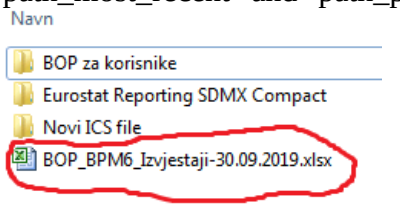
Example of "loading indicator" as defined in "functions.R" as above and location of relevant files that function refers to:



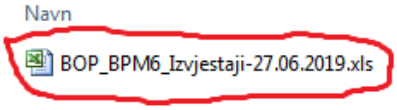


How data is prepared for the dashboard:

All data input comes from the two excel-files that are specified by the paths "path_most_recent" and "path_previous" as previously shown above.

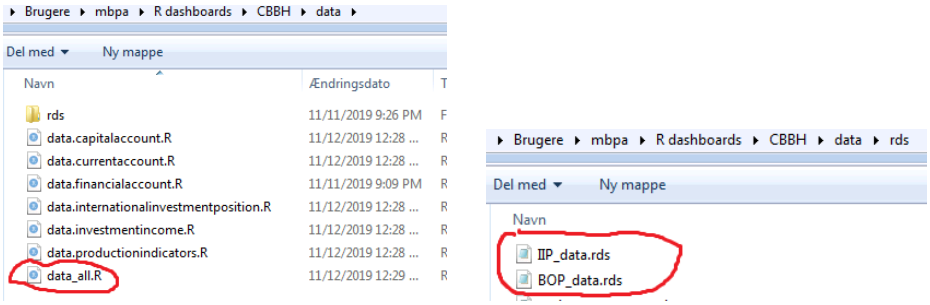


Most recent excel-file



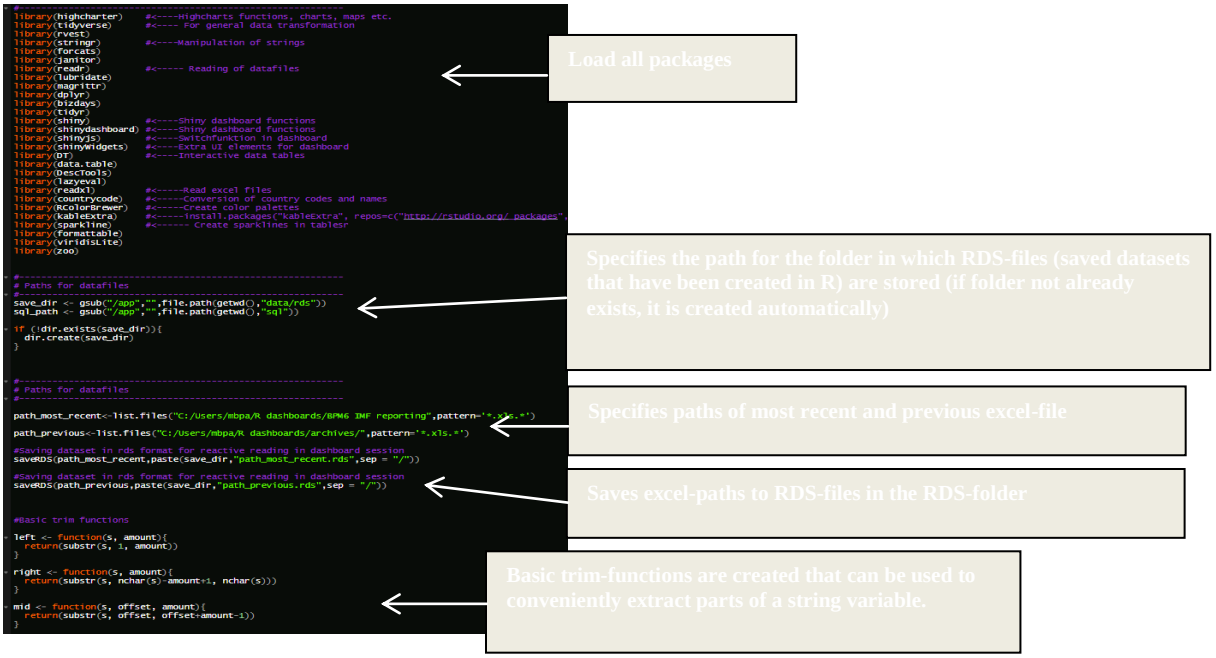
Previous excel-file

The relevant content from the excel-files is read into R via the script called "data_all.R" which is located in the data-folder. The objective of this script is to select the balance of payments data (from the sheet called "BPM6BOPForm") and the international investment positon data (from the sheet called "IIP_Kvartalni") from the excel files and transform the data structure to a database-like structure and save the datasets as RDS-files to be found in the RDS-folder. Details of "data_all.R" follow below:
Location of "data_all.R" script and the RDS-files that the script generates:



Contents of the "data_all.R" script:

The script begins by loading all required packages, specifies the "save_dir" (location of RDS-folder), specifies the paths for the excel-files ("path_most_recent" and "path_previous"), saves the excel-paths as RDS-files in the RDS-folder and defines the basic trim-functions.



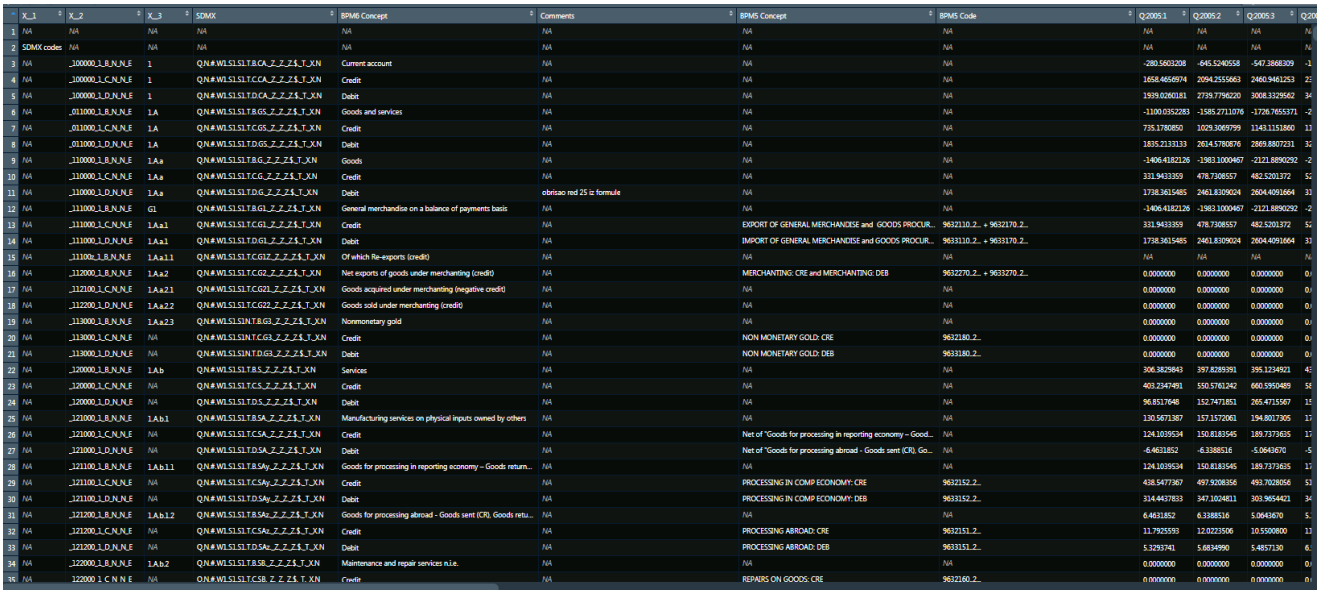
Next, the balance of payments data is read into R from the most recent excel-file (as defined by "path_most_recent") using the sheet "BPM6BOPForm". The first 7 rows are skipped because we want the periods to appear as variable names (see below). The data is saved as a temporary data frame in the global environment called "BPM6BOPForm_df".

```
# BPM6BOPForm
BPM6BOPForm_df <-read_excel(paste0(path_most_recent),  sheet = "BPM6BOPForm",skip=7)
```

By skipping the first 7 rows from the "BPM6BOPForm" sheet, the values on the 7th row will be used as variables names:

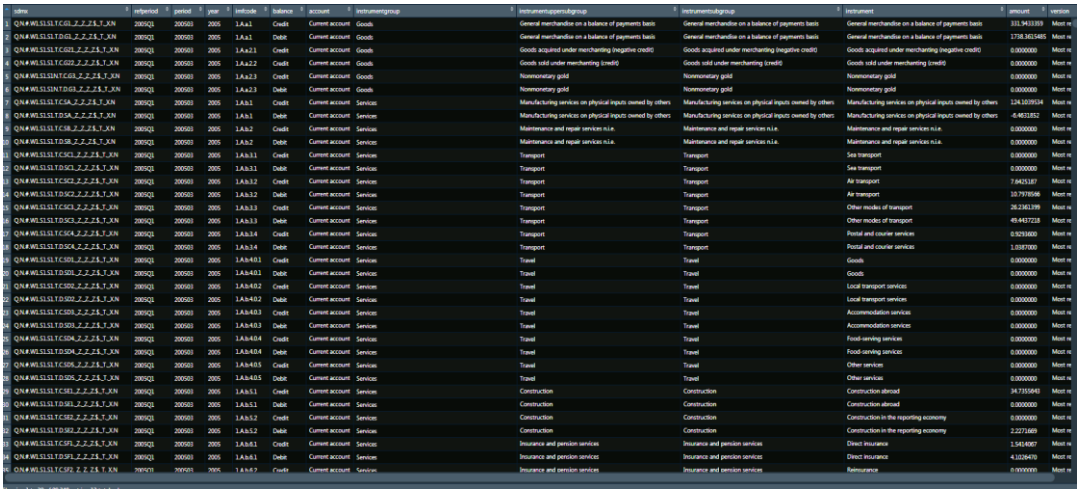
	B	C	D	E	V	W	X	Y	Z	AA	AB	AC	AD	AE
1														
2														
3				Country: Bosnia & Herzegovina										
4				Currency: National Currency										
5				Scale: Millions										
6				Balance of Payments 6 Conversion Table										
7					2008-Q2	2008-Q3	2008-Q4	2009-Q1	2009-Q2	2009-Q3	2009-Q4	2010-Q1	2010-Q2	2010-Q3
8			SDMX	BPM6 Concept	Q:2008:2	Q:2008:3	Q:2008:4	Q:2009:1	Q:2009:2	Q:2009:3	Q:2009:4	Q:2010:1	Q:2010:2	Q:2010:3
9														
11	100000_1_B_N_N_E	1	Q.N.#.W1.S1.S1.T.B.CA._Z._Z._Z.S._T._X.N	Current account	-916.4	-980.5	-889.1	-311.4	-477.8	-328.9	-470.2	-160.6	-328.5	-583.2
12	100000_1_C_N_N_E	1	Q.N.#.W1.S1.S1.T.C.CA._Z._Z._Z.S._T._X.N	Credit	3,238.6	3,529.7	3,027.7	2,562.7	2,725.5	3,015.1	2,742.5	2,588.0	3,092.3	3,297.5
13	100000_1_D_N_N_E	1	Q.N.#.W1.S1.S1.T.D.CA._Z._Z._Z.S._T._X.N	Debit	4,154.9	4,510.1	3,916.8	2,874.0	3,203.3	3,344.0	3,212.7	2,748.6	3,420.8	3,880.7
14	011000_1_B_N_N_E	1.A	Q.N.#.W1.S1.S1.T.B.GS._Z._Z._Z.S._T._X.N	Goods and services	-2,116.4	-2,279.3	-2,092.2	-1,347.8	-1,575.7	-1,464.6	-1,496.3	-1,013.8	-1,334.5	-1,600.3
15	011000_1_C_N_N_E	1.A	Q.N.#.W1.S1.S1.T.C.GS._Z._Z._Z.S._T._X.N	Credit	1,792.9	1,984.2	1,564.9	1,351.8	1,503.7	1,775.7	1,570.8	1,566.6	1,921.7	2,074.4
16	011000_1_D_N_N_E	1.A	Q.N.#.W1.S1.S1.T.D.GS._Z._Z._Z.S._T._X.N	Debit	3,909.3	4,263.4	3,657.1	2,699.6	3,079.3	3,240.2	3,067.2	2,580.5	3,256.2	3,674.6
17	110000_1_B_N_N_E	1.A.a	Q.N.#.W1.S1.S1.T.B.G._Z._Z._Z.S._T._X.N	Goods	-2,632.9	-2,760.8	-2,565.6	-1,749.9	-1,952.1	-1,854.8	-1,879.5	-1,381.1	-1,778.4	-2,013.7
18	110000_1_C_N_N_E	1.A.a	Q.N.#.W1.S1.S1.T.C.G._Z._Z._Z.S._T._X.N	Credit	1,105.4	1,143.2	977.0	807.7	907.0	1,022.9	1,018.6	1,071.4	1,305.7	1,319.5

Resulting data frame will look like below (note that for columns with no value in row 7, the variable names will default to "X__1", "X__2" etc. In earlier versions of the "readxl" package variable names for empty columns may default to "…1", "…2" etc.):



	X__1	X__2	X__3	SDMX	BPM6 Concept	Comments	BPM5 Concept	BPM5 Code	Q:2005:1	Q:2005:2	Q:2005:3	Q:20
1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	N
2	SDMX codes	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	N
3	NA	100000_1_B_N_N_E	1	Q.N.#.W1.S1.S1.T.B.CA._Z._Z._Z.S._T._X.N	Current account	NA	NA	NA	-280.5603208	-645.5240558	-547.3868309	-1
4	NA	100000_1_C_N_N_E	1	Q.N.#.W1.S1.S1.T.C.CA._Z._Z._Z.S._T._X.N	Credit	NA	NA	NA	1658.4659874	2094.2555663	2460.9461253	2
5	NA	100000_1_D_N_N_E	1	Q.N.#.W1.S1.S1.T.D.CA._Z._Z._Z.S._T._X.N	Debit	NA	NA	NA	1939.0260181	2739.7796220	3008.3329562	3
6	NA	011000_1_B_N_N_E	1.A	Q.N.#.W1.S1.S1.T.B.GS._Z._Z._Z.S._T._X.N	Goods and services	NA	NA	NA	-1100.0352283	-1585.2711076	-1726.7655371	-2
7	NA	011000_1_C_N_N_E	1.A	Q.N.#.W1.S1.S1.T.C.GS._Z._Z._Z.S._T._X.N	Credit	NA	NA	NA	735.1780850	1029.3069799	1143.1151860	1
8	NA	011000_1_D_N_N_E	1.A	Q.N.#.W1.S1.S1.T.D.GS._Z._Z._Z.S._T._X.N	Debit	NA	NA	NA	1835.2133133	2614.5780876	2869.8807231	3
9	NA	110000_1_B_N_N_E	1.A.a	Q.N.#.W1.S1.S1.T.B.G._Z._Z._Z.S._T._X.N	Goods	NA	NA	NA	-1406.4182126	-1983.1000467	-2121.8890292	-2
10	NA	110000_1_C_N_N_E	1.A.a	Q.N.#.W1.S1.S1.T.C.G._Z._Z._Z.S._T._X.N	Credit	NA	NA	NA	331.9433359	478.7308557	482.5201372	5
11	NA	110000_1_D_N_N_E	1.A.a	Q.N.#.W1.S1.S1.T.D.G._Z._Z._Z.S._T._X.N	Debit	obrisao red 25 iz formule	NA	NA	1738.3615485	2461.8309024	2604.4091664	3
12	NA	111000_1_B_N_N_E	G1	Q.N.#.W1.S1.S1.T.B.G1._Z._Z._Z.S._T._X.N	General merchandise on a balance of payments basis	NA	NA	NA	-1406.4182126	-1983.1000467	-2121.8890292	-2
13	NA	111000_1_C_N_N_E	1.Aa1	Q.N.#.W1.S1.S1.T.C.G1._Z._Z._Z.S._T._X.N	Credit	NA	EXPORT OF GENERAL MERCHANDISE and GOODS PROCUR...	9632110.2... + 9632170.2...	331.9433359	478.7308557	482.5201372	5
14	NA	111000_1_D_N_N_E	1.Aa1	Q.N.#.W1.S1.S1.T.D.G1._Z._Z._Z.S._T._X.N	Debit	IMPORT OF GENERAL MERCHANDISE and GOODS PROCUR...	9633110.2... + 9633170.2...	1738.3615485	2461.8309024	2604.4091664	3
15	NA	111100_1_B_N_N_E	1.Aa1.1	Q.N.#.W1.S1.S1.T.C.G1Z._Z._Z.S._T._X.N	Of which Re-exports (credit)	NA	NA	NA	NA	NA	NA	N
16	NA	112000_1_B_N_N_E	1.Aa.2	Q.N.#.W1.S1.S1.T.C.G2._Z._Z._Z.S._T._X.N	Net exports of goods under merchanting (credit)	NA	MERCHANTING: CRE and MERCHANTING: DEB	9632270.2... + 9633270.2...	0.0000000	0.0000000	0.0000000	0
17	NA	112100_1_C_N_N_E	1.Aa.21	Q.N.#.W1.S1.S1.T.C.G21._Z._Z._Z.S._T._X.N	Goods acquired under merchanting (negative credit)	NA	NA	NA	0.0000000	0.0000000	0.0000000	0
18	NA	112200_1_D_N_N_E	1.Aa.2.2	Q.N.#.W1.S1.S1.T.C.G22._Z._Z._Z.S._T._X.N	Goods sold under merchanting (credit)	NA	NA	NA	0.0000000	0.0000000	0.0000000	0
19	NA	113000_1_B_N_N_E	1.Aa.2.3	Q.N.#.W1.S1.S1.N.T.B.G3._Z._Z._Z.S._T._X.N	Nonmonetary gold	NA	NA	NA	0.0000000	0.0000000	0.0000000	0
20	NA	113000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.N.T.C.G3._Z._Z._Z.S._T._X.N	Credit	NA	NON MONETARY GOLD: CRE	9632180.2...	0.0000000	0.0000000	0.0000000	0
21	NA	113000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.N.T.D.G3._Z._Z._Z.S._T._X.N	Debit	NA	NON MONETARY GOLD: DEB	9633180.2...	0.0000000	0.0000000	0.0000000	0
22	NA	120000_1_B_N_N_E	1.Ab	Q.N.#.W1.S1.S1.T.B.S._Z._Z._Z.S._T._X.N	Services	NA	NA	NA	306.3829843	397.8289391	395.1234921	4
23	NA	120000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.S._Z._Z._Z.S._T._X.N	Credit	NA	NA	NA	403.2347491	550.5761242	660.5950489	5
24	NA	120000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.S._Z._Z._Z.S._T._X.N	Debit	NA	NA	NA	96.8517648	152.7471851	265.4715567	1
25	NA	121000_1_B_N_N_E	1.Ab.1	Q.N.#.W1.S1.S1.T.B.SA._Z._Z._Z.S._T._X.N	Manufacturing services on physical inputs owned by others	NA	NA	NA	130.5671387	157.1572061	194.8017305	1
26	NA	121000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SA._Z._Z._Z.S._T._X.N	Credit	NA	Net of "Goods for processing in reporting economy – Good...	NA	124.1039534	150.8183545	189.7373635	1
27	NA	121000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SA._Z._Z._Z.S._T._X.N	Debit	NA	Net of "Goods for processing abroad - Goods sent (CR), Go...	NA	-6.4631852	-6.3388516	-5.0643670	-5
28	NA	121100_1_B_N_N_E	1.Ab.1.1	Q.N.#.W1.S1.S1.T.B.SAy._Z._Z._Z.S._T._X.N	Goods for processing in reporting economy – Goods return...	NA	NA	NA	124.1039534	150.8183545	189.7373635	1
29	NA	121100_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SAy._Z._Z._Z.S._T._X.N	Credit	NA	PROCESSING IN COMP ECONOMY: CRE	9632152.2...	438.5477367	497.9208356	493.7028056	5
30	NA	121100_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SAy._Z._Z._Z.S._T._X.N	Debit	NA	PROCESSING IN COMP ECONOMY: DEB	9633152.2...	314.4437833	347.1024811	303.9654421	3
31	NA	121200_1_B_N_N_E	1.Ab.1.2	Q.N.#.W1.S1.S1.T.B.SAz._Z._Z._Z.S._T._X.N	Goods for processing abroad - Goods sent (CR), Goods retu...	NA	NA	NA	6.4631852	6.3388516	5.0643670	5
32	NA	121200_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SAz._Z._Z._Z.S._T._X.N	Credit	NA	PROCESSING ABROAD: CRE	9632151.2...	11.7925593	12.0223506	10.5500800	1
33	NA	121200_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SAz._Z._Z._Z.S._T._X.N	Debit	NA	PROCESSING ABROAD: DEB	9633151.2...	5.3293741	5.6834990	5.4857130	6
34	NA	122000_1_B_N_N_E	1.Ab.2	Q.N.#.W1.S1.S1.T.B.SB._Z._Z._Z.S._T._X.N	Maintenance and repair services n.i.e.	NA	NA	NA	0.0000000	0.0000000	0.0000000	0
35	NA	122000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SB._Z._Z._Z.S._T._X.N	Credit	NA	REPAIRS ON GOODS: CRE	9632160.2...	0.0000000	0.0000000	0.0000000	0

The data frame is not very useful at this stage. The objective of the following lines of code in "data_all.R" is to transform the data frame above so it will have a database-like structure like this:



sdmx	refperiod	period	year	imtrade	balance	account	instrumentgroup	instrumentsuppersubgroup	instrumentsubgroup	instrument	amount	version
Q.N.#.W1.S1.S1.T.C.G1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Aa.1	Credit	Current account	Goods	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	331.9433359	Most re
Q.N.#.W1.S1.S1.T.D.G1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Aa.1	Debit	Current account	Goods	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	1738.3615485	Most re
Q.N.#.W1.S1.S1.T.C.G21._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Aa.2.1	Credit	Current account	Goods	Goods acquired under merchanting (negative credit)	Goods acquired under merchanting (negative credit)	Goods acquired under merchanting (negative credit)	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.G22._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Aa.2.2	Credit	Current account	Goods	Goods sold under merchanting (credit)	Goods sold under merchanting (credit)	Goods sold under merchanting (credit)	0.0000000	Most re
Q.N.#.W1.S1.S1.N.T.C.G3._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Aa.2.3	Credit	Current account	Goods	Nonmonetary gold	Nonmonetary gold	Nonmonetary gold	0.0000000	Most re
Q.N.#.W1.S1.S1.N.T.D.G3._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Aa.2.3	Debit	Current account	Goods	Nonmonetary gold	Nonmonetary gold	Nonmonetary gold	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SA._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.1	Credit	Current account	Services	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	124.1039534	Most re
Q.N.#.W1.S1.S1.T.D.SA._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.1	Debit	Current account	Services	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	-6.4631852	Most re
Q.N.#.W1.S1.S1.T.C.SB._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.2	Credit	Current account	Services	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SB._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.2	Debit	Current account	Services	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SC1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.1	Credit	Current account	Services	Transport	Transport	Sea transport	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SC1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.1	Debit	Current account	Services	Transport	Transport	Sea transport	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SC2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.2	Credit	Current account	Services	Transport	Transport	Air transport	7.6425187	Most re
Q.N.#.W1.S1.S1.T.D.SC2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.2	Debit	Current account	Services	Transport	Transport	Air transport	10.7978598	Most re
Q.N.#.W1.S1.S1.T.C.SC3._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.3	Credit	Current account	Services	Transport	Transport	Other modes of transport	26.2361399	Most re
Q.N.#.W1.S1.S1.T.D.SC3._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.3	Debit	Current account	Services	Transport	Transport	Other modes of transport	49.4437218	Most re
Q.N.#.W1.S1.S1.T.C.SC4._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.4	Credit	Current account	Services	Transport	Transport	Postal and courier services	0.9291600	Most re
Q.N.#.W1.S1.S1.T.D.SC4._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.3.4	Debit	Current account	Services	Transport	Transport	Postal and courier services	1.0387300	Most re
Q.N.#.W1.S1.S1.T.C.SD1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.1	Credit	Current account	Services	Travel	Travel	Goods	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SD1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.1	Debit	Current account	Services	Travel	Travel	Goods	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SD2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.2	Credit	Current account	Services	Travel	Travel	Local transport services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SD2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.2	Debit	Current account	Services	Travel	Travel	Local transport services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SD3._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.3	Credit	Current account	Services	Travel	Travel	Accommodation services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SD3._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.3	Debit	Current account	Services	Travel	Travel	Accommodation services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SD4._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.4	Credit	Current account	Services	Travel	Travel	Food-serving services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SD4._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.4	Debit	Current account	Services	Travel	Travel	Food-serving services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SD5._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.5	Credit	Current account	Services	Travel	Travel	Other services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SD5._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.4.0.5	Debit	Current account	Services	Travel	Travel	Other services	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SE1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.5.1	Credit	Current account	Services	Construction	Construction	Construction abroad	34.7355843	Most re
Q.N.#.W1.S1.S1.T.D.SE1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.5.1	Debit	Current account	Services	Construction	Construction	Construction abroad	0.0000000	Most re
Q.N.#.W1.S1.S1.T.C.SE2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.5.2	Credit	Current account	Services	Construction	Construction	Construction in the reporting economy	0.0000000	Most re
Q.N.#.W1.S1.S1.T.D.SE2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.5.2	Debit	Current account	Services	Construction	Construction	Construction in the reporting economy	2.2271669	Most re
Q.N.#.W1.S1.S1.T.C.SF1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.6.1	Credit	Current account	Services	Insurance and pension services	Insurance and pension services	Direct insurance	1.5414007	Most re
Q.N.#.W1.S1.S1.T.D.SF1._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.6.1	Debit	Current account	Services	Insurance and pension services	Insurance and pension services	Direct insurance	4.1224470	Most re
Q.N.#.W1.S1.S1.T.C.SF2._Z._Z._Z.S._T._X.N	2005Q1	2005Q1	2005	1.Ab.6.2	Credit	Current account	Services	Insurance and pension services	Insurance and pension services	Reinsurance	0.0000000	Most re

To create this structure, we do the following:
We remove empty columns if there should be any (i.e. remove any column for which all rows are *NA*):

```
#Empty columns are removed

BPM6BOPForm_df<-BPM6BOPForm_df[, colSums(is.na(BPM6BOPForm_df)) != nrow(BPM6BOPForm_df)]
```

We now transform the structure from wide format to a panel data format (long format). More specifically, we want the periods (2007Q1, 2007Q2..., 2019Q1 etc.) to appear as values in a time series variable rather than having them as variable names and also to have the associated amounts listed as values in a common variable. By using the "gather" function, we can specify that variables for every time period (starting from column 9 to the last column) should be listed in a time series variable (which we call "refperiod") and the amounts for each time period should be listed in a new variable called "amount" :

```
#Data is transformed from wide format to long panel structure

BPM6BOPForm_df<-gather(BPM6BOPForm_df,refperiod,amount,names(BPM6BOPForm_df[,9:ncol(BPM6BOPForm_df)]))
```

Note in which column the first time period appears. In this data frame, the first period ("2005Q1") is in column 9. "ncol" is the total number of columns so we are telling the "gather" function to apply from column 9 and to the last column

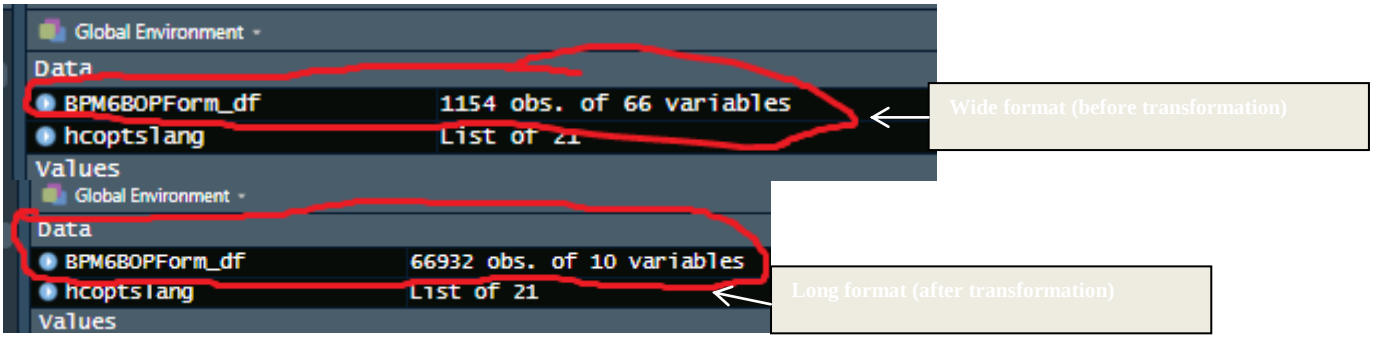
The data frame "BPM6BOPForm_df" now look like this:

	X_1	X_2	X_3	SDMX	BPM6 Concept	Comments	BPM5 Concept	BPM5 Code	refperiod	amount
1	NA	NA	NA	NA	NA	NA	NA	NA	Q:2005:1	NA
2	SDMX codes	NA	NA	NA	NA	NA	NA	NA	Q:2005:1	NA
3	NA	_100000_1_B_N_N_E	1	Q.N.#.W1.S1.S1.T.B.CA_Z_Z_Z.$_T_XN	Current account	NA	NA	NA	Q:2005:1	-280.5603208
4	NA	_100000_1_C_N_N_E	1	Q.N.#.W1.S1.S1.T.C.CA_Z_Z_Z.$_T_XN	Credit	NA	NA	NA	Q:2005:1	1658.4656974
5	NA	_100000_1_D_N_N_E	1	Q.N.#.W1.S1.S1.T.D.CA_Z_Z_Z.$_T_XN	Debit	NA	NA	NA	Q:2005:1	1939.0260181
6	NA	_011000_1_B_N_N_E	1.A	Q.N.#.W1.S1.S1.T.B.GS_Z_Z_Z.$_T_XN	Goods and services	NA	NA	NA	Q:2005:1	-1100.0352283
7	NA	_011000_1_C_N_N_E	1.A	Q.N.#.W1.S1.S1.T.C.GS_Z_Z_Z.$_T_XN	Credit	NA	NA	NA	Q:2005:1	735.1780850
8	NA	_011000_1_D_N_N_E	1.A	Q.N.#.W1.S1.S1.T.D.GS_Z_Z_Z.$_T_XN	Debit	NA	NA	NA	Q:2005:1	1835.2133133
9	NA	_110000_1_B_N_N_E	1.Aa	Q.N.#.W1.S1.S1.T.B.G_Z_Z_Z.$_T_XN	Goods	NA	NA	NA	Q:2005:1	-1406.4182126
10	NA	_110000_1_C_N_N_E	1.Aa	Q.N.#.W1.S1.S1.T.C.G_Z_Z_Z.$_T_XN	Credit	NA	NA	NA	Q:2005:1	331.9433359
11	NA	_110000_1_D_N_N_E	1.Aa	Q.N.#.W1.S1.S1.T.D.G_Z_Z_Z.$_T_XN	Debit	obrisao red 25 iz formule	NA	NA	Q:2005:1	1738.3615485
12	NA	_111000_1_B_N_N_E	G1	Q.N.#.W1.S1.S1.T.B.G1_Z_Z_Z.$_T_XN	General merchandise on a balance of payments basis	NA	NA	NA	Q:2005:1	-1406.4182126
13	NA	_111000_1_C_N_N_E	1.Aa1	Q.N.#.W1.S1.S1.T.C.G1_Z_Z_Z.$_T_XN	Credit	NA	EXPORT OF GENERAL MERCHANDISE and GOODS PROCUR...	9632110.2... + 9632170.2...	Q:2005:1	331.9433359
14	NA	_111000_1_D_N_N_E	1.Aa1	Q.N.#.W1.S1.S1.T.D.G1_Z_Z_Z.$_T_XN	Debit	NA	IMPORT OF GENERAL MERCHANDISE and GOODS PROCUR...	9633110.2... + 9633170.2...	Q:2005:1	1738.3615485
15	NA	_11100c_1_B_N_N_E	1.Aa1.1	Q.N.#.W1.S1.S1.T.C.G1Z_Z_Z_Z.$_T_XN	Of which Re-exports (credit)	NA	NA	NA	Q:2005:1	NA
16	NA	_112000_1_B_N_N_E	1.Aa2	Q.N.#.W1.S1.S1.T.C.G2_Z_Z_Z.$_T_XN	Net exports of goods under merchanting (credit)	NA	MERCHANTING: CRE and MERCHANTING: DEB	9632270.2... + 9633270.2...	Q:2005:1	0.0000000
17	NA	_112100_1_C_N_N_E	1.Aa2.1	Q.N.#.W1.S1.S1.T.C.G21_Z_Z_Z.$_T_XN	Goods acquired under merchanting (negative credit)	NA	NA	NA	Q:2005:1	0.0000000
18	NA	_112200_1_D_N_N_E	1.Aa2.2	Q.N.#.W1.S1.S1.T.C.G22_Z_Z_Z.$_T_XN	Goods sold under merchanting (credit)	NA	NA	NA	Q:2005:1	0.0000000
19	NA	_113000_1_B_N_N_E	1.Aa2.3	Q.N.#.W1.S1.S1N.T.B.G3_Z_Z_Z.$_T_XN	Nonmonetary gold	NA	NA	NA	Q:2005:1	0.0000000
20	NA	_113000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1N.T.C.G3_Z_Z_Z.$_T_XN	Credit	NA	NON MONETARY GOLD: CRE	9632180.2...	Q:2005:1	0.0000000
21	NA	_113000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1N.T.D.G3_Z_Z_Z.$_T_XN	Debit	NA	NON MONETARY GOLD: DEB	9633180.2...	Q:2005:1	0.0000000
22	NA	_120000_1_B_N_N_E	1.Ab	Q.N.#.W1.S1.S1.T.B.S_Z_Z_Z.$_T_XN	Services	NA	NA	NA	Q:2005:1	306.3829843
23	NA	_120000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.S_Z_Z_Z.$_T_XN	Credit	NA	NA	NA	Q:2005:1	403.2347491
24	NA	_120000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.S_Z_Z_Z.$_T_XN	Debit	NA	NA	NA	Q:2005:1	96.8517648
25	NA	_121000_1_B_N_N_E	1.Ab.1	Q.N.#.W1.S1.S1.T.B.SA_Z_Z_Z.$_T_XN	Manufacturing services on physical inputs owned by others	NA	NA	NA	Q:2005:1	130.5671387
26	NA	_121000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SA_Z_Z_Z.$_T_XN	Credit	NA	Net of "Goods for processing in reporting economy – Good...	NA	Q:2005:1	124.1039534
27	NA	_121000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SA_Z_Z_Z.$_T_XN	Debit	NA	Net of "Goods for processing abroad - Goods sent (CR), Go...	NA	Q:2005:1	-6.46311852
28	NA	_121100_1_B_N_N_E	1.Ab.1.1	Q.N.#.W1.S1.S1.T.B.SAy_Z_Z_Z.$_T_XN	Goods for processing in reporting economy – Goods return...	NA	NA	NA	Q:2005:1	124.1039534
29	NA	_121100_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SAy_Z_Z_Z.$_T_XN	Credit	NA	PROCESSING IN COMP ECONOMY: CRE	9632152.2...	Q:2005:1	438.5477367
30	NA	_121100_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SAy_Z_Z_Z.$_T_XN	Debit	NA	PROCESSING IN COMP ECONOMY: DEB	9633152.2...	Q:2005:1	314.4437833
31	NA	_121200_1_B_N_N_E	1.Ab.1.2	Q.N.#.W1.S1.S1.T.B.SAe_Z_Z_Z.$_T_XN	Goods for processing abroad - Goods sent (CR), Goods retu...	NA	NA	NA	Q:2005:1	6.4631852
32	NA	_121200_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SAe_Z_Z_Z.$_T_XN	Credit	NA	PROCESSING ABRDAD: CRE	9632151.2...	Q:2005:1	11.7925593
33	NA	_121200_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SAe_Z_Z_Z.$_T_XN	Debit	NA	PROCESSING ABROAD: DEB	9633151.2...	Q:2005:1	5.3293741
34	NA	_122000_1_B_N_N_E	1.Ab.2	Q.N.#.W1.S1.S1.T.B.SR_Z_Z_Z.$_T_XN	Maintenance and repair services n.i.e.	NA	NA	NA	Q:2005:1	0.0000000
35	NA	_122000_1_C_N_N_E	NA	Q.N.#.W1.S1.S1.T.C.SR_Z_Z_Z.$_T_XN	Credit	NA	REPAIRS ON GOODS: CRE	9632160.2...	Q:2005:1	0.0000000
36	NA	_122000_1_D_N_N_E	NA	Q.N.#.W1.S1.S1.T.D.SR_Z_Z_Z.$_T_XN	Debit	NA	REPAIRS ON GOODS: DEB	9633160.2...	Q:2005:1	0.0000000

As we can see, we now have all periods and the associated amounts listed in the new variables called "refperiod" and "amount", respectively.

We can also see that the number of observations has increased from 1154 observations to 66932 observations. This is because we have 58 periods in this example (2005Q1-2019Q2) which implies that the number of observations in the long format will be N x T = 1154 x 58 =

66932. Also, we see that the number of variables decrease by 56 from 66 variables to 10 variables (we remove 58 variables and create two new variables "refperiod" and "amount").



Next, we run a couple of commands to clean up some irrelevant rows (rows with invalid periods and with no SDMX code) and making sure that all columns have proper data formats and that any "white spaces" are trimmed:

```
#Remove irrelvant periods

BPM6BOPForm_df<-filter(BPM6BOPForm_df,!(right(BPM6BOPForm_df$refperiod,3) %in% c("__1")))

#Remove white spaces from all characters and cast all amounts as numeric

BPM6BOPForm_df[,1:9] <- lapply(BPM6BOPForm_df[,1:9], str_squish)
BPM6BOPForm_df$amount<-as.numeric(BPM6BOPForm_df$amount)

#Remove rows with no SDMX code
BPM6BOPForm_df<-filter(BPM6BOPForm_df,!is.na(SDMX))
```

We now change the format of the "refperiod" variable (e.g. so "2018:Q1" becomes "2018Q1") and we also create an extra time variable called "period" where the time values from "refperiod" are converted from strings to numeric values (e.g. "2018Q1" is a string variable whereas 201803 is numeric and it is sometimes convenient to be able to refer to periods in numeric values). We also create a variable called "year" to make it easy to change the dataset from quarterly to yearly observations if needed:

```
#Convert to proper formats for periods
BPM6BOPForm_df<-mutate(BPM6BOPForm_df,period=as.numeric(gsub(":","0",substr(refperiod,3,nchar(refperiod)))))%>%
  mutate(refperiod=paste0(substr(refperiod,3,6), #<-----Extracting year
   ifelse(as.numeric(right(refperiod,1))==1,paste0("Q1") #<------ Paste 1th Quarter if applicable
  ,ifelse(as.numeric(right(refperiod,1))==2,paste0("Q2") #<------ Paste 2th Quarter if applicable
  ,ifelse(as.numeric(right(refperiod,1))==3,paste0("Q3") #<------ Paste 3th Quarter if applicable
  ,paste0("Q4"))))                                       #<------ Paste 4th Quarter if applicable
)) %>%
    mutate(period=ifelse(right(period,1)==1,as.numeric(paste0(left(period,5),"3"))
           ,ifelse(right(period,1)==2,as.numeric(paste0(left(period,5),"6"))
           ,ifelse(right(period,1)==3,as.numeric(paste0(left(period,5),"9"))
           ,as.numeric(paste0(left(period,4),"12"))))))%>%mutate(year=as.numeric(left(period,4)))
```

The next few lines of code are just changing the variable names to something that will be easier to refer to and also only selects the variables that will be used for the remaining transformation (the final two lines correct some errors in the instrument codes, but since we will generally be relying on the SDMX codes to identify the relevant data, this step is not that important):

```
#Renaming variables
BPM6BOPForm_df<-BPM6BOPForm_df%>%setnames("SDMX","sdmx")%>% setnames("X__3","imfcode")%>%setnames("BPM6 Concept","bpm6")

#Select relevant variables

BPM6BOPForm_df<-BPM6BOPForm_df%>%select(sdmx,imfcode,bpm6,refperiod,period,year,amount)

#Correcting wrong instrument codes

BPM6BOPForm_df<-BPM6BOPForm_df%>%mutate(imfcode=ifelse(sdmx %in% c("Q.N.#.W1.S1.S1.T.C.SE1._Z._Z._Z.$._T._X.N","Q.N.#.W1.S1.S1.T.D.SE1._Z._Z._Z.$._T._X.N"),"1.A.b.5.1",imfcode))
BPM6BOPForm_df<-BPM6BOPForm_df%>%mutate(imfcode=ifelse(sdmx %in% c("Q.N.#.W1.S1.S1.T.C.SE2._Z._Z._Z.$._T._X.N","Q.N.#.W1.S1.S1.T.D.SE2._Z._Z._Z.$._T._X.N"),"1.A.b.5.2",imfcode))
```

Next, the objective is to ensure that relevant rows of data have meaningful names. For the current account and capital account data, the relevant rows will be those that identify the debits and credits on the most detailed instrument level.

For instance, if we look at the data frame "BPM6BOPForm_df" at this stage, relevant data would be rows 11 and 12, for example. Rows 11 and 12 correspond to the debit and credit values for "General merchandise on the balance of payments basis". All preceding rows (rows 1-10) are all aggregate values that will already contain the values from rows 11 and 12, among others.

	sdmx	imfcode	bpm6	refperiod	period	year	amount
1	Q.N.#.W1.S1.S1.T.B.CA._Z._Z._Z.$._T._X.N	1	Current account	2005Q1	200503	2005	-280.5603208
2	Q.N.#.W1.S1.S1.T.C.CA._Z._Z._Z.$._T._X.N	1	Credit	2005Q1	200503	2005	1658.4656974
3	Q.N.#.W1.S1.S1.T.D.CA._Z._Z._Z.$._T._X.N	1	Debit	2005Q1	200503	2005	1939.0260181
4	Q.N.#.W1.S1.S1.T.B.GS._Z._Z._Z.$._T._X.N	1.A	Goods and services	2005Q1	200503	2005	-1100.0352283
5	Q.N.#.W1.S1.S1.T.C.GS._Z._Z._Z.$._T._X.N	1.A	Credit	2005Q1	200503	2005	735.1780850
6	Q.N.#.W1.S1.S1.T.D.GS._Z._Z._Z.$._T._X.N	1.A	Debit	2005Q1	200503	2005	1835.2133133
7	Q.N.#.W1.S1.S1.T.B.G._Z._Z._Z.$._T._X.N	1.Aa	Goods	2005Q1	200503	2005	-1406.4182126
8	Q.N.#.W1.S1.S1.T.C.G._Z._Z._Z.$._T._X.N	1.Aa	Credit	2005Q1	200503	2005	331.9433359
9	Q.N.#.W1.S1.S1.T.D.G._Z._Z._Z.$._T._X.N	1.Aa	Debit	2005Q1	200503	2005	1738.3615485
10	Q.N.#.W1.S1.S1.T.B.G1._Z._Z._Z.$._T._X.N	G1	General merchandise on a balance of payments basis	2005Q1	200503	2005	-1406.4182126
11	Q.N.#.W1.S1.S1.T.C.G1._Z._Z._Z.$._T._X.N	1.Aa1	Credit	2005Q1	200503	2005	331.9433359
12	Q.N.#.W1.S1.S1.T.D.G1._Z._Z._Z.$._T._X.N	1.Aa1	Debit	2005Q1	200503	2005	1738.3615485
13	Q.N.#.W1.S1.S1.T.C.G1Z._Z._Z._Z.$._T._X.N	1.Aa1.1	Of which Re-exports (credit)	2005Q1	200503	2005	*NA*
14	Q.N.#.W1.S1.S1.T.C.G2._Z._Z._Z.$._T._X.N	1.Aa2	Net exports of goods under merchanting (credit)	2005Q1	200503	2005	0.0000000
15	Q.N.#.W1.S1.S1.T.C.G21._Z._Z._Z.$._T._X.N	1.Aa2.1	Goods acquired under merchanting (negative credit)	2005Q1	200503	2005	0.0000000
16	Q.N.#.W1.S1.S1.T.C.G22._Z._Z._Z.$._T._X.N	1.Aa2.2	Goods sold under merchanting (credit)	2005Q1	200503	2005	0.0000000
17	Q.N.#.W1.S1.S1N.T.B.G3._Z._Z._Z.$._T._X.N	1.Aa2.3	Nonmonetary gold	2005Q1	200503	2005	0.0000000
18	Q.N.#.W1.S1.S1N.T.C.G3._Z._Z._Z.$._T._X.N	*NA*	Credit	2005Q1	200503	2005	0.0000000
19	Q.N.#.W1.S1.S1N.T.D.G3._Z._Z._Z.$._T._X.N	*NA*	Debit	2005Q1	200503	2005	0.0000000

Rather than having the aggregate values, we would like to be able to get the aggregates by summing the rows containing the most detailed representation of the data. In order to be able to sum the rows in a meaningful way, we should also need a separate variable that identifies whether a row is a credit or a debit value. For example, we would like rows 11 and 12 from above to appear like this:

	sdmx	imfcode	bpm6	refperiod	period	year	amount	balance	instrument
1	Q.N.#.W1.S1.S1.T.C.G1._Z._Z._Z.$._T._X.N	1.Aa1	*NA*	2005Q1	200503	2005	331.9433359	Credit	General merchandise on a balance of payments basis
2	Q.N.#.W1.S1.S1.T.D.G1._Z._Z._Z.$._T._X.N	1.Aa1	*NA*	2005Q1	200503	2005	1738.3615485	Debit	General merchandise on a balance of payments basis

Such that we have a variable ("instrument") that describes which instrument category the amount belongs to and another variable ("balance") that tells us whether the amount is a credit or a debit.

To achieve the desired structure as discussed above, we do the following:
We create a separate data frame (called "imf_fact) in which all rows under the "bpm6" variable that contains the values "Debit" and "Credit" are replace with *NA (*we select only the "sdmx" and "bpm6" variables for this data frame):

```
#Create fact table with instrument names
imf_fact<-BPM6BOPForm_df%>%mutate(bpm6=ifelse(bpm6 %in% c("Debit","Credit"),factor(NA),bpm6))
imf_fact<-imf_fact%>%select(sdmx,bpm6)
```

The data frame "imf_fact" will now look like this:

	sdmx	bpm6
1	Q.N.#.W1.S1.S1.T.B.CA._Z._Z._Z.$._T._X.N	Current account
2	Q.N.#.W1.S1.S1.T.C.CA._Z._Z._Z.$._T._X.N	*NA*
3	Q.N.#.W1.S1.S1.T.D.CA._Z._Z._Z.$._T._X.N	*NA*
4	Q.N.#.W1.S1.S1.T.B.GS._Z._Z._Z.$._T._X.N	Goods and services
5	Q.N.#.W1.S1.S1.T.C.GS._Z._Z._Z.$._T._X.N	*NA*
6	Q.N.#.W1.S1.S1.T.D.GS._Z._Z._Z.$._T._X.N	*NA*
7	Q.N.#.W1.S1.S1.T.B.G._Z._Z._Z.$._T._X.N	Goods
8	Q.N.#.W1.S1.S1.T.C.G._Z._Z._Z.$._T._X.N	*NA*
9	Q.N.#.W1.S1.S1.T.D.G._Z._Z._Z.$._T._X.N	*NA*
10	Q.N.#.W1.S1.S1.T.B.G1._Z._Z._Z.$._T._X.N	General merchandise on a balance of payments basis
11	Q.N.#.W1.S1.S1.T.C.G1._Z._Z._Z.$._T._X.N	*NA*
12	Q.N.#.W1.S1.S1.T.D.G1._Z._Z._Z.$._T._X.N	*NA*
13	Q.N.#.W1.S1.S1.T.C.G1Z._Z._Z._Z.$._T._X.N	Of which Re-exports (credit)
14	Q.N.#.W1.S1.S1.T.C.G2._Z._Z._Z.$._T._X.N	Net exports of goods under merchanting (credit)
15	Q.N.#.W1.S1.S1.T.C.G21._Z._Z._Z.$._T._X.N	Goods acquired under merchanting (negative credit)
16	Q.N.#.W1.S1.S1.T.C.G22._Z._Z._Z.$._T._X.N	Goods sold under merchanting (credit)
17	Q.N.#.W1.S1.S1N.T.B.G3._Z._Z._Z.$._T._X.N	Nonmonetary gold
18	Q.N.#.W1.S1.S1N.T.C.G3._Z._Z._Z.$._T._X.N	*NA*
19	Q.N.#.W1.S1.S1N.T.D.G3._Z._Z._Z.$._T._X.N	*NA*
20	Q.N.#.W1.S1.S1.T.B.S._Z._Z._Z.$._T._X.N	Services
21	Q.N.#.W1.S1.S1.T.C.S._Z._Z._Z.$._T._X.N	*NA*
22	Q.N.#.W1.S1.S1.T.D.S._Z._Z._Z.$._T._X.N	*NA*
23	Q.N.#.W1.S1.S1.T.B.SA._Z._Z._Z.$._T._X.N	Manufacturing services on physical inputs owned by others
24	Q.N.#.W1.S1.S1.T.C.SA._Z._Z._Z.$._T._X.N	*NA*
25	Q.N.#.W1.S1.S1.T.D.SA._Z._Z._Z.$._T._X.N	*NA*

The idea now is to replace all the *NA* values, with the relevant instrument names. The relevant instrument names should correspond to the names preceding every *NA* value under the "bpm6" variable. For example, we would like the *NA* values in rows 11 and 12 to be replace with the preceding instrument name in row 10 ("General merchandise on a balance of payments basis").

The function "na.locf" does exactly this:

```
imf_fact$bpm6<-na.locf(imf_fact$bpm6)
```

Note how all NA values under the "bpm6" variable have now been replaced with the preceding names. For example, rows 11 and 12 (that previously were called "Debit" and "Credit") now has the value "General merchandise on balance of payments basis".

	sdmx	bpm6
1	Q.N.#.W1.S1.S1.T.B.CA_Z_Z_Z$_T_X.N	Current account
2	Q.N.#.W1.S1.S1.T.C.CA_Z_Z_Z$_T_X.N	Current account
3	Q.N.#.W1.S1.S1.T.D.CA_Z_Z_Z$_T_X.N	Current account
4	Q.N.#.W1.S1.S1.T.B.GS_Z_Z_Z$_T_X.N	Goods and services
5	Q.N.#.W1.S1.S1.T.C.GS_Z_Z_Z$_T_X.N	Goods and services
6	Q.N.#.W1.S1.S1.T.D.GS_Z_Z_Z$_T_X.N	Goods and services
7	Q.N.#.W1.S1.S1.T.B.G_Z_Z_Z$_T_X.N	Goods
8	Q.N.#.W1.S1.S1.T.C.G_Z_Z_Z$_T_X.N	Goods
9	Q.N.#.W1.S1.S1.T.D.G_Z_Z_Z$_T_X.N	Goods
10	Q.N.#.W1.S1.S1.T.B.G1_Z_Z_Z$_T_X.N	General merchandise on a balance of payments basis
11	Q.N.#.W1.S1.S1.T.C.G1_Z_Z_Z$_T_X.N	General merchandise on a balance of payments basis
12	Q.N.#.W1.S1.S1.T.D.G1_Z_Z_Z$_T_X.N	General merchandise on a balance of payments basis
13	Q.N.#.W1.S1.S1.T.C.G1Z_Z_Z_Z$_T_X.N	Of which Re-exports (credit)
14	Q.N.#.W1.S1.S1.T.C.G2_Z_Z_Z$_T_X.N	Net exports of goods under merchanting (credit)
15	Q.N.#.W1.S1.S1.T.C.G21_Z_Z_Z$_T_X.N	Goods acquired under merchanting (negative credit)
16	Q.N.#.W1.S1.S1.T.C.G22_Z_Z_Z$_T_X.N	Goods sold under merchanting (credit)
17	Q.N.#.W1.S1.S1N.T.B.G3_Z_Z_Z$_T_X.N	Nonmonetary gold
18	Q.N.#.W1.S1.S1N.T.C.G3_Z_Z_Z$_T_X.N	Nonmonetary gold
19	Q.N.#.W1.S1.S1N.T.D.G3_Z_Z_Z$_T_X.N	Nonmonetary gold
20	Q.N.#.W1.S1.S1.T.B.S_Z_Z_Z$_T_X.N	Services
21	Q.N.#.W1.S1.S1.T.C.S_Z_Z_Z$_T_X.N	Services
22	Q.N.#.W1.S1.S1.T.D.S_Z_Z_Z$_T_X.N	Services
23	Q.N.#.W1.S1.S1.T.B.SA_Z_Z_Z$_T_X.N	Manufacturing services on physical inputs owned by others
24	Q.N.#.W1.S1.S1.T.C.SA_Z_Z_Z$_T_X.N	Manufacturing services on physical inputs owned by others
25	Q.N.#.W1.S1.S1.T.D.SA_Z_Z_Z$_T_X.N	Manufacturing services on physical inputs owned by others

We also just change the variable name "bpm6" to "instrument" and make sure that the data frame "imf_fact" only has distinct SDMX codes (i.e. no duplicates).

```
imf_fact<-imf_fact%>%select(sdmx,bpm6)%>%distinct()%>%setnames("bpm6","instrument")
```

Next, we return to our main data frame "BPM6BOPForm_df". We now want to create the variable "balance" to indicate whether a row is a "debit" or "credit" (for the current account and capital account) or whether it is an "asset" or a "liability" (for the financial account). All net amounts will just be assigned the value "balance" (we will be removing all net amounts later anyway).
To assign the correct balance indicator, we can just rely on the fact that the SDMX codes will indicate this from the pattern "T.A" ("Asset"), "T.L" ("Liability"), "T.D" ("Debit") and "T.C" ("Credit"):

```
#Add indicator of direction

BPM6BOPForm_df<-mutate(BPM6BOPForm_df,balance=ifelse(sdmx %like% c("%T.A%"),"Assets",
                                                     ifelse(sdmx %like% c("%T.L%"),"Liabilities",
                                                     ifelse(sdmx %like% c("%T.D%"),"Debit",
                                                     ifelse(sdmx %like% c("%T.C%"),"Credit","Balance")))))
```

We can see by comparing the SDMX code with the newly create variable "balance" that the latter identifies correctly whether a row is a credit, debit, asset or liability (or net amount):

	sdmx	imfcode	bpm6	refperiod	period	year	amount	balance
1	Q.N.#.W1.S1.S1.T.B.CA_Z_Z_Z$_T_X.N	1	Current account	2005Q1	200503	2005	-280.5603208	Balance
2	Q.N.#.W1.S1.S1.T.C.CA_Z_Z_Z$_T_X.N	1	Credit	2005Q1	200503	2005	1658.4656974	Credit
3	Q.N.#.W1.S1.S1.T.D.CA_Z_Z_Z$_T_X.N	1	Debit	2005Q1	200503	2005	1939.0260181	Debit
4	Q.N.#.W1.S1.S1.T.B.GS_Z_Z_Z$_T_X.N	1.A	Goods and services	2005Q1	200503	2005	-1100.0352283	Balance
5	Q.N.#.W1.S1.S1.T.C.GS_Z_Z_Z$_T_X.N	1.A	Credit	2005Q1	200503	2005	735.1780850	Credit
6	Q.N.#.W1.S1.S1.T.D.GS_Z_Z_Z$_T_X.N	1.A	Debit	2005Q1	200503	2005	1835.2133133	Debit
7	Q.N.#.W1.S1.S1.T.B.G_Z_Z_Z$_T_X.N	1.A.a	Goods	2005Q1	200503	2005	-1406.4182126	Balance
8	Q.N.#.W1.S1.S1.T.C.G_Z_Z_Z$_T_X.N	1.A.a	Credit	2005Q1	200503	2005	331.9433359	Credit
9	Q.N.#.W1.S1.S1.T.D.G_Z_Z_Z$_T_X.N	1.A.a	Debit	2005Q1	200503	2005	1738.3615485	Debit
10	Q.N.#.W1.S1.S1.T.B.G1_Z_Z_Z$_T_X.N	G1	General merchandise on a balance of payments basis	2005Q1	200503	2005	-1406.4182126	Balance
11	Q.N.#.W1.S1.S1.T.C.G1_Z_Z_Z$_T_X.N	1.A.a.1	Credit	2005Q1	200503	2005	331.9433359	Credit
12	Q.N.#.W1.S1.S1.T.D.G1_Z_Z_Z$_T_X.N	1.A.a.1	Debit	2005Q1	200503	2005	1738.3615485	Debit
13	Q.N.#.W1.S1.S1.T.C.G1Z_Z_Z_Z$_T_X.N	1.A.a.1.1	Of which Re-exports (credit)	2005Q1	200503	2005	*NA*	Credit
14	Q.N.#.W1.S1.S1.T.C.G2_Z_Z_Z$_T_X.N	1.A.a.2	Net exports of goods under merchanting (credit)	2005Q1	200503	2005	0.0000000	Credit
15	Q.N.#.W1.S1.S1.T.C.G21_Z_Z_Z$_T_X.N	1.A.a.2.1	Goods acquired under merchanting (negative credit)	2005Q1	200503	2005	0.0000000	Credit
16	Q.N.#.W1.S1.S1.T.C.G22_Z_Z_Z$_T_X.N	1.A.a.2.2	Goods sold under merchanting (credit)	2005Q1	200503	2005	0.0000000	Credit
17	Q.N.#.W1.S1.S1N.T.B.G3_Z_Z_Z$_T_X.N	1.A.a.2.3	Nonmonetary gold	2005Q1	200503	2005	0.0000000	Balance
18	Q.N.#.W1.S1.S1N.T.C.G3_Z_Z_Z$_T_X.N	*NA*	Credit	2005Q1	200503	2005	0.0000000	Credit
19	Q.N.#.W1.S1.S1N.T.D.G3_Z_Z_Z$_T_X.N	*NA*	Debit	2005Q1	200503	2005	0.0000000	Debit

We now create a new temporary data frame called "CA_KA_df" for the current account and capital account (the financial account has a different structure i.e. the credit/debit representation is different from the asset/liability representation).
We first replace any *NA* values in the "imf_code" variable with the preceding codes, just to avoid having any *NA*s in this column (recall the "na.locf" function):

```
##################################################
# Start of current account and capital account
##################################################

# Fill balance items with relevant instrument codes - nessessry to complete structure for current account and capital account
CA_KA_df<-BPM6BOPForm_df

CA_KA_df$imfcode<-na.locf(CA_KA_df$imfcode)
```

We then replace the "Credit" and "Debit" values in the "bpm6" column with *NA*s (as we just did before in the "imf_fact " data frame):

```
#Replace debit/credit with NA i bpm6 concept
CA_KA_df<-CA_KA_df%>%mutate(bpm6=ifelse(balance %in% c("Credit","Debit"),factor(NA),bpm6))
```

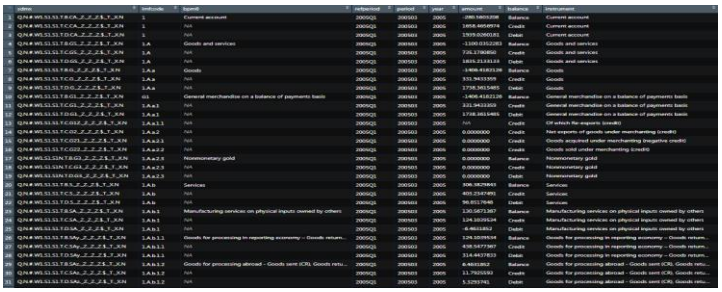
The idea now is that we can replace all these *NA*s, whith the relevant instrument names that we created in "imf_fact".

To enrich the "CA_KA_df" data frame with the instrument variable from the "imf_fact" data frame, we simply join (merge) the two data frames.

We do this as a "left_join" by the SDMX codes, such that instrument names from "imf_fact" are only transferred to those rows in "CA_KA_df that are matched by having identical SDMX codes (they should always identical, however, since "imf_fact" was created based on the same data, so an "inner-join" would have yielded the same result):

```
# Replace NAs with relevant instrument names
CA_KA_df<-left_join(CA_KA_df,imf_fact,by=c("sdmx"))
```

We can see that the variable "instrument" is now present in the "CA_KA_df" data frame (the "bpm6" column is now redundant since it still contains all the *NAs*).



The data frame "CA_KA_df" is now complete with instrument names, but the data frame still contains many redundant rows, such as all the aggregates, the net amounts etc.
The following lines of code are meant to remove these redundant rows and create a meaningful instrument hierarchy:

We can remove all net amounts very simply (filter out all rows for which the variable "balance" contains the value "Balance". Recall that these rows effectively correspond to having the pattern "T.N" in the SDMX code):

```
#Remove net amounts
CA_KA_df<-filter(CA_KA_df,!(balance %in% c("Balance")))
```

Next, we can remove some of the main aggregates. We rely on the fact that aggregates for "Current account", "Capital account", "Financial account", "Goods and services" and "Primary income" can be easily identified by having the SDMX pattern ".CA.", ".KA.", ".FA.", ".GS." and ".IN1.", respectively.

```
#Remove account aggregates

CA_KA_df<-filter(CA_KA_df,!(sdmx %like% c("%.CA.%")))%>%filter(!(sdmx %like% c("%.KA.%")))%>%filter(!(sdmx %like% c("%.FA.%")))%>%
filter(!(sdmx %like% c("%.GS.%")))%>%filter(!(sdmx %like% c( "%.IN1.%")))
```

Since we will focus on the current account and capital account at first, it is convenient to be able to distinguish which to rows that belong to each account. We create a new indicator variable (called "account") and we rely on the fact that all items that has an instrument code (values of the variable "imfcode") starting with "1","2" and "3" will correspond to the current account, capital account and financial account, respectively (Note that this is the only time we will be relying on the instrument codes in the variable "imfcode". We will rely on SDMX codes for all subsequent filtering)

```
# Add indicator of account

CA_KA_df<-CA_KA_df%>%mutate(account=ifelse(imfcode %like% c("1%"),"Current account",
                                           ifelse(imfcode %like% c("2%"),"Capital account",
                                           ifelse(imfcode %like% c("3%"),"Financial account",factor(NA)))))
```

Now, since we removed the aggregates in the previous step, we still would like to have an indicator of which instrument group our remaining instrument names belong to. For instance,

we would like to know that "General merchandise on balance of payments basis" belongs to "Goods", that "Local transport services" belong to "Services" and so forth.The next step creates this indicator variable ("instrumentgroup"), to indicate which groups each item in the "instrument" variable belong to (again we rely on the common patterns in the SDMX codes):

```
#Add indicator of instrument group

CA_KA_df<-CA_KA_df%>%mutate(instrumentgroup=ifelse(account %in% c("Current account") & !(sdmx %like% c("%SG1%")) & sdmx %like% c("%.G1%") |
                                                     sdmx %like% c("%.G2%") & !(sdmx %like% c("%SG2%")) | sdmx %like% c("%.G3%") &
                                                     !(sdmx %like% c("%.D1.%")) &  !(sdmx %like% c("%.D2.%")) &  !(sdmx %like% c("%.D3.%")) &
                                                     !(sdmx %like% c("%.D4.%")) & !(sdmx %like% c("%.IN1.%")),"Goods",
                                                   ifelse(account %in% c("Current account") & sdmx %like% c("%.SA%")  | sdmx %like% c("%.SB%")
                                                          | sdmx %like% c("%.SC%") | sdmx %like% c("%.SD%") |sdmx %like% c("%.SE%") |
                                                          sdmx %like% c("%.SF%") | sdmx %like% c("%.SG%")| sdmx %like% c("%.SH%") | sdmx %like% c("%.SI%") |
                                                     sdmx %like% c("%.SJ%") | sdmx %like% c("%.SK%") | sdmx %like% c("%.SL%")      & !(sdmx %like% c("%.D1.%")) &
                                                     !(sdmx %like% c("%.D2.%")) &  !(sdmx %like% c("%.D3.%")) &  !(sdmx %like% c("%.D4.%")) & !(sdmx %like% c("%.IN1.%")),"Services",
                                                   ifelse(account %in% c("Current account") & sdmx %like% c("%.IN1%") |  sdmx %like% c("%.D1.%") | sdmx %like% c("%.D2.%")| sdmx %like% c("%.D3.%") |
                                                            sdmx %like% c("%.D4.%") ,"Primary income",
                                                          ifelse(account %in% c("Current account") & sdmx %like% c("%.IN2%") | account %in% c("Current account") & sdmx %like% c("%.D5.%") |
                                                                   account %in% c("Current account")  & sdmx %like% c("%.D6.%")| account %in% c("Current account") & sdmx %like% c("%.D7.%") |
                                                                 account %in% c("Current account") & sdmx %like% c("%.D8.%") ,"Secondary income",
                                                                 ifelse(account %in% c("Capital account") & sdmx %like% c("%.NP.%") ,"Gross acquisitions (DR.) / disposals (CR.) of nonproduced nonfinancial assets",
                                                                        ifelse(account %in% c("Capital account") & sdmx %like% c("%.D9.%"),"Capital transfers",factor(NA))))))))
```

We can see in the resulting "CA_KA_df" data frame that items belonging to the service account, for instance, have now been assigned the value "Services" in the "instrumentgroup" variable and items belonging to "Primary income" is identified by having the value "Primary income" in the "instrumentgroup" variable:

sdmx	imfcode	bpm6	refperiod	period	year	amount	balance	instrument	account	instrumentgroup
Q.N.#.W1.S1.S1.T.D.SI3._Z._Z._Z.$._T._XN	1.Ab.9.3	NA	2005Q1	200503	2005	0.0000000	Debit	Information services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SI._Z._Z._Z.$._T._XN	1.Ab.10	NA	2005Q1	200503	2005	24.4383593	Credit	Other business services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SI._Z._Z._Z.$._T._XN	1.Ab.10	NA	2005Q1	200503	2005	0.0000000	Debit	Other business services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SI1._Z._Z._Z.$._T._XN	1.Ab.10.1	NA	2005Q1	200503	2005	0.0000000	Credit	Research and development services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SI1._Z._Z._Z.$._T._XN	1.Ab.10.1	NA	2005Q1	200503	2005	0.0000000	Debit	Research and development services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SI2._Z._Z._Z.$._T._XN	1.Ab.10.2	NA	2005Q1	200503	2005	0.0000000	Credit	Professional and management consulting services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SI2._Z._Z._Z.$._T._XN	1.Ab.10.2	NA	2005Q1	200503	2005	0.0000000	Debit	Professional and management consulting services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SI3._Z._Z._Z.$._T._XN	1.Ab.10.3	NA	2005Q1	200503	2005	24.4383593	Credit	Technical, trade-related, and other business services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SI3._Z._Z._Z.$._T._XN	1.Ab.10.3	NA	2005Q1	200503	2005	0.0000000	Debit	Technical, trade-related, and other business services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SK._Z._Z._Z.$._T._XN	1.Ab.11	NA	2005Q1	200503	2005	0.0000000	Credit	Personal, cultural, and recreational services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SK._Z._Z._Z.$._T._XN	1.Ab.11	NA	2005Q1	200503	2005	0.0000000	Debit	Personal, cultural, and recreational services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SK1._Z._Z._Z.$._T._XN	1.Ab.11.1	NA	2005Q1	200503	2005	0.0000000	Credit	Audiovisual and related services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SK1._Z._Z._Z.$._T._XN	1.Ab.11.1	NA	2005Q1	200503	2005	0.0000000	Debit	Audiovisual and related services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SK2._Z._Z._Z.$._T._XN	1.Ab.11.2	NA	2005Q1	200503	2005	0.0000000	Credit	Other personal, cultural, and recreational services	Current account	Services
Q.N.#.W1.S1.S1.T.D.SK2._Z._Z._Z.$._T._XN	1.Ab.11.2	NA	2005Q1	200503	2005	0.0000000	Debit	Other personal, cultural, and recreational services	Current account	Services
Q.N.#.W1.S1.S1.T.C.SL._Z._Z._Z.$._T._XN	1.Ab.12	NA	2005Q1	200503	2005	0.8559171	Credit	Government goods and services n.i.e.	Current account	Services
Q.N.#.W1.S1.S1.T.D.SL._Z._Z._Z.$._T._XN	1.Ab.12	NA	2005Q1	200503	2005	4.1120000	Debit	Government goods and services n.i.e.	Current account	Services
Q.N.#.W1.S1.S1.T.C.SD._Z._Z._Z.$._T._XN	1.Ab.0.1	NA	2005Q1	200503	2005	NA	Credit	Tourism-related services in travel and passenger transport	Current account	Services
Q.N.#.W1.S1.S1.T.D.SD._Z._Z._Z.$._T._XN	1.Ab.0.1	NA	2005Q1	200503	2005	NA	Debit	Tourism-related services in travel and passenger transport	Current account	Services
Q.N.#.W1.S1.S1.T.C.D1._Z._Z._Z.$._T._XN	1.B.1	NA	2005Q1	200503	2005	189.1161704	Credit	Compensation of employees	Current account	Primary income
Q.N.#.W1.S1.S1.T.D.D1._Z._Z._Z.$._T._XN	1.B.1	NA	2005Q1	200503	2005	4.6965000	Debit	Compensation of employees	Current account	Primary income
Q.N.#.W1.S1.S1.T.C.D4P._T.F._Z.$._T._XN	1.B.2	NA	2005Q1	200503	2005	41.6434351	Credit	Investment income	Current account	Primary income
Q.N.#.W1.S1.S1.T.D.D4P._T.F._Z.$._T._XN	1.B.2	NA	2005Q1	200503	2005	51.6857483	Debit	Investment income	Current account	Primary income
Q.N.#.W1.S1.S1.T.C.D4P.D.F._Z.$._T._XN	1.B.2.1	NA	2005Q1	200503	2005	0.4320000	Credit	Direct investment	Current account	Primary income
Q.N.#.W1.S1.S1.T.D.D4P.D.F._Z.$._T._XN	1.B.2.1	NA	2005Q1	200503	2005	24.7064685	Debit	Direct investment	Current account	Primary income
Q.N.#.W1.S1.S1.T.C.D45.D.FS._Z.$._T._XN	1.B.2.1.1	NA	2005Q1	200503	2005	0.0000000	Credit	Income on equity and investment fund shares	Current account	Primary income
Q.N.#.W1.S1.S1.T.D.D45.D.FS._Z.$._T._XN	1.B.2.1.1	NA	2005Q1	200503	2005	18.1274685	Debit	Income on equity and investment fund shares	Current account	Primary income
Q.N.#.W1.S1.S1.T.C.D425.D.FS._Z.$._T._XN	1.B.2.1.1.1	NA	2005Q1	200503	2005	0.0000000	Credit	Dividends and withdrawals from income of quasi-corporatio...	Current account	Primary income
Q.N.#.W1.S1.S1.T.D.D425.D.FS._Z.$._T._XN	1.B.2.1.1.1	NA	2005Q1	200503	2005	13.3934446	Debit	Dividends and withdrawals from income of quasi-corporatio...	Current account	Primary income
Q.N.#.W1.S1.S1.T.C.D425.D1.FS._Z.$._T._XN	1.B.2.1.1.1.1	NA	2005Q1	200503	2005	0.0000000	Credit	Direct investor in direct investment enterprises	Current account	Primary income
Q.N.#.W1.S1.S1.T.D.D425.D1.FS._Z.$._T._XN	1.B.2.1.1.1.1	NA	2005Q1	200503	2005	13.3934446	Debit	Direct investor in direct investment enterprises	Current account	Primary income
Q.N.#.W1.S1.S1.T.C.D425.D2.FS._Z.$._T._XN	1.B.2.1.1.1.2	NA	2005Q1	200503	2005	0.0000000	Credit	Direct investment enterprises in direct investor (reverse inve...	Current account	Primary income

Next, we can remove all rows belonging to the financial account (we will prepare the financial account data in later steps). We also remove any NA values from the newly-created "instrumentgroup" variable (Note that the assignment of groups in the definition of the "instrumentgroup" variable above should have captured all items belonging to the current account and capital account, so that any remaining *NA*s that we remove should merely be redundant aggregates)

```
#Select only current account and capital account
CA_KA_df<-CA_KA_df%>%filter(!(account %in% c("Financial account")))

#Remove redundant instrument groups
CA_KA_df<- CA_KA_df[!is.na(CA_KA_df$instrumentgroup),]
```

Now, the next many lines of code are meant to remove any redundant aggregates on the current account and capital account while assigning relevant aggregate groups under the "instrumentgroup" variable. At the same time, a new group variable "instrumentsubgroup" is created to add information to the instrument hierarchy. We will not go through all the lines of

code but just show an example of how the instrument hierarchies for the variables "instrumentgroup" and "instrumentsubgroup" are defined:

For example, in this step we are removing redundant aggregates related to transport service" and instead the remaining rows related to transport services are identified by assigning the value "Transport" to the "instrumentgroup" variable.

```
#Remove redundant aggregates
CA_KA_df<-CA_KA_df%>%filter(!(sdmx %in% c("Q.N.#.W1.S1.S1.T.C.SC._Z._Z._Z.$._T._X.N","Q.N.#.W1.S1.S1.T.D.SC._Z._Z._Z.$._T._X.N"))) #<----- Remove "Transport"

CA_KA_df<-CA_KA_df%>%filter(!(sdmx %like% c("%.SCA.%")))%>%filter(!(sdmx %like% c("%.SCB.%")))%>%filter(!(sdmx %like% c("%.SCC.%"))) #<-----Remove "For all modes of transport"

CA_KA_df<-CA_KA_df%>%filter(!(sdmx %like% c("%.SC11%")))%>%filter(!(sdmx %like% c("%.SC12%")))%>%filter(!(sdmx %like% c("%.SC13%"))) #<-----Remove "Passenger, freight and others" from "Sea transport"

CA_KA_df<-CA_KA_df%>%filter(!(sdmx %like% c("%.SC21%")))%>%filter(!(sdmx %like% c("%.SC22%")))%>%filter(!(sdmx %like% c("%.SC23%"))) #<-----Remove "Passenger, freight and others" from "Air transport"

CA_KA_df<-CA_KA_df%>%filter(!(sdmx %like% c("%.SC31%")))%>%filter(!(sdmx %like% c("%.SC32%")))%>%filter(!(sdmx %like% c("%.SC33%"))) #<-----Remove "Passenger, freight and others" from "Other modes of transport"

# Create subgroup "Transport"
CA_KA_df<-CA_KA_df%>%mutate(instrumentsubgroup=ifelse(sdmx %like% c("%.SC1.%") | sdmx %like% c("%.SC2.%")| sdmx %like% c("%.SC3.%")| sdmx %like% c("%.SC4.%") ,
                                                      "Transport",instrumentsubgroup)) #<--- "Sea transport", "Air transport","Other modes of transport" and "Postal courier services"
```

We can see that in the resulting data frame "CA_FA_df", we now have a very complete instrument hierarchy for transport services. Specifically, we have that the variable "instrument" represents the most detailed information related to transport services. For instance, we have "Sea transport", "Air transport" etc. listed under the "instrument" variable. The "instrumentsubgroup" variable then tells us that these items belong to "Transport". The variable "instrumentgroup" tells us that the items belong to "Services" and finally the "account" variable tells us that the items are part of the "Current account".

	sdmx	imfcode	bpm6	refperiod	period	year	amount	balance	instrument	account	instrumentgroup	instrumentsubgroup
237	Q.N.#.W1.S1.S1.T.CSC3_Z_Z_Z.$_T_X.N	1.Ab.3.3	NA	2005Q1	200503	2005	2.623614e+01	Credit	Other modes of transport	Current account	Services	Transport
238	Q.N.#.W1.S1.S1.T.DSC3_Z_Z_Z.$_T_X.N	1.Ab.3.3	NA	2005Q1	200503	2005	4.944372e-01	Debit	Other modes of transport	Current account	Services	Transport
239	Q.N.#.W1.S1.S1.T.CSC4_Z_Z_Z.$_T_X.N	1.Ab.3.4	NA	2005Q1	200503	2005	9.293600e-01	Credit	Postal and courier services	Current account	Services	Transport
240	Q.N.#.W1.S1.S1.T.DSC4_Z_Z_Z.$_T_X.N	1.Ab.3.4	NA	2005Q1	200503	2005	1.038700e+00	Debit	Postal and courier services	Current account	Services	Transport
241	Q.N.#.W1.S1.S1.T.CSC1_Z_Z_Z.$_T_X.N	1.Ab.3.1	NA	2005Q2	200506	2005	0.000000e+00	Credit	Sea transport	Current account	Services	Transport
242	Q.N.#.W1.S1.S1.T.DSC1_Z_Z_Z.$_T_X.N	1.Ab.3.1	NA	2005Q2	200506	2005	0.000000e+00	Debit	Sea transport	Current account	Services	Transport
243	Q.N.#.W1.S1.S1.T.CSC2_Z_Z_Z.$_T_X.N	1.Ab.3.2	NA	2005Q2	200506	2005	8.462451e+00	Credit	Air transport	Current account	Services	Transport
244	Q.N.#.W1.S1.S1.T.DSC2_Z_Z_Z.$_T_X.N	1.Ab.3.2	NA	2005Q2	200506	2005	1.208828e+01	Debit	Air transport	Current account	Services	Transport
245	Q.N.#.W1.S1.S1.T.CSC3_Z_Z_Z.$_T_X.N	1.Ab.3.3	NA	2005Q2	200506	2005	3.273307e+01	Credit	Other modes of transport	Current account	Services	Transport
246	Q.N.#.W1.S1.S1.T.DSC3_Z_Z_Z.$_T_X.N	1.Ab.3.3	NA	2005Q2	200506	2005	6.699947e+01	Debit	Other modes of transport	Current account	Services	Transport
247	Q.N.#.W1.S1.S1.T.CSC4_Z_Z_Z.$_T_X.N	1.Ab.3.4	NA	2005Q2	200506	2005	1.097820e+00	Credit	Postal and courier services	Current account	Services	Transport
248	Q.N.#.W1.S1.S1.T.DSC4_Z_Z_Z.$_T_X.N	1.Ab.3.4	NA	2005Q2	200506	2005	1.211840e+00	Debit	Postal and courier services	Current account	Services	Transport
249	Q.N.#.W1.S1.S1.T.CSC1_Z_Z_Z.$_T_X.N	1.Ab.3.1	NA	2005Q3	200509	2005	0.000000e+00	Credit	Sea transport	Current account	Services	Transport
250	Q.N.#.W1.S1.S1.T.DSC1_Z_Z_Z.$_T_X.N	1.Ab.3.1	NA	2005Q3	200509	2005	0.000000e+00	Debit	Sea transport	Current account	Services	Transport
251	Q.N.#.W1.S1.S1.T.CSC2_Z_Z_Z.$_T_X.N	1.Ab.3.2	NA	2005Q3	200509	2005	1.445634e+01	Credit	Air transport	Current account	Services	Transport
252	Q.N.#.W1.S1.S1.T.DSC2_Z_Z_Z.$_T_X.N	1.Ab.3.2	NA	2005Q3	200509	2005	1.596126e+01	Debit	Air transport	Current account	Services	Transport
253	Q.N.#.W1.S1.S1.T.CSC3_Z_Z_Z.$_T_X.N	1.Ab.3.3	NA	2005Q3	200509	2005	3.298521e+01	Credit	Other modes of transport	Current account	Services	Transport
254	Q.N.#.W1.S1.S1.T.DSC3_Z_Z_Z.$_T_X.N	1.Ab.3.3	NA	2005Q3	200509	2005	6.925672e+01	Debit	Other modes of transport	Current account	Services	Transport

The subsequent lines of code basically aim to create this kind of instrument hierarchy for all relevant items on the current account and capital acount.
When we get to the investment income on primary income, there is a need to add yet another instrument layer to the hierarchy, as the items on investment income has some additional instrument information that is similar to the representation on the financial account (e.g. "equity", "inter-company debt" can be grouped as "direct investment", "portfolio equity" and "debt securities" can be grouped as "portfolio investment" etc.). We add this information under a new variable called "instrumentuppersubgroup":

```
# Add Financial account layer on instrument hierachi  for investment income
CA_KA_df<-CA_KA_df%>%mutate(instrumentuppersubgroup=ifelse(instrumentsubgroup %in% c("Income on equity and investment fund shares","Interest from inter-company debt"),
                                                           "Direct investment",
                                                    ifelse(instrumentsubgroup %in% c("Dividends on equity excluding investment fund shares","Investment income attributable to investment fund shareholders",
                                                                                     "Interest from debt securities"),"Portfolio investment",
                                                     ifelse(instrumentsubgroup %in% c("Withdrawals from income of quasi-corporations","Interest from loans and deposits",
                                                                                      "Investment income attributable to policyholders in insurance, pension schemes, and standardized guarantees"),"Other investment",
                                                     ifelse(instrumentsubgroup %in% c("Interest (Reserve assets)","Income on equity and investment fund shares (Reserve assets)"),"Reserve assets",instrumentsubgroup)))))
```

Once all hierarchies (i.e. "account", "instrumentgroup", "instrumentuppersubgroup", "instrumentsubgroup" and "instrument") have been defined for the current account and the capital account, we can select the relevant variables from the resulting data frame "CA_KA_df":

```
#Select relevant variables
CA_KA_df<-CA_KA_df%>%select(sdmx,refperiod,period,year,imfcode,balance,account,instrumentgroup,instrumentuppersubgroup,instrumentsubgroup,instrument,amount)
```

The data frame "CA_KA_df" should now look like this:

	sdmx	refperiod	period	year	imfcode	balance	account	instrumentgroup	instrumentuppersubgroup	instrumentsubgroup	instrument	amount
1	Q.N.#.W1.S1.S1.T.C.G1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Aa1	Credit	Current account	Goods	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	331.9433359
2	Q.N.#.W1.S1.S1.T.D.G1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Aa1	Debit	Current account	Goods	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	1738.3615485
3	Q.N.#.W1.S1.S1.T.C.G21_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Aa2.1	Credit	Current account	Goods	Goods acquired under merchanting (negative credit)	Goods acquired under merchanting (negative credit)	Goods acquired under merchanting (negative credit)	0.0000000
4	Q.N.#.W1.S1.S1.T.C.G22_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Aa2.2	Credit	Current account	Goods	Goods sold under merchanting (credit)	Goods sold under merchanting (credit)	Goods sold under merchanting (credit)	0.0000000
5	Q.N.#.W1.S1.S1.N.T.C.G3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Aa2.3	Credit	Current account	Goods	Nonmonetary gold	Nonmonetary gold	Nonmonetary gold	0.0000000
6	Q.N.#.W1.S1.S1.N.T.D.G3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Aa2.3	Debit	Current account	Goods	Nonmonetary gold	Nonmonetary gold	Nonmonetary gold	0.0000000
7	Q.N.#.W1.S1.S1.T.C.SA_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab1	Credit	Current account	Services	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	124.1039534
8	Q.N.#.W1.S1.S1.T.D.SA_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab1	Debit	Current account	Services	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	-6.4631852
9	Q.N.#.W1.S1.S1.T.C.SB_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab2	Credit	Current account	Services	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	0.0000000
10	Q.N.#.W1.S1.S1.T.D.SB_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab2	Debit	Current account	Services	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	0.0000000
11	Q.N.#.W1.S1.S1.T.C.SC1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.1	Credit	Current account	Services	Transport	Transport	Sea transport	0.0000000
12	Q.N.#.W1.S1.S1.T.D.SC1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.1	Debit	Current account	Services	Transport	Transport	Sea transport	0.0000000
13	Q.N.#.W1.S1.S1.T.C.SC2_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.2	Credit	Current account	Services	Transport	Transport	Air transport	7.6425187
14	Q.N.#.W1.S1.S1.T.D.SC2_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.2	Debit	Current account	Services	Transport	Transport	Air transport	10.7978566
15	Q.N.#.W1.S1.S1.T.C.SC3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.3	Credit	Current account	Services	Transport	Transport	Other modes of transport	28.2361399
16	Q.N.#.W1.S1.S1.T.D.SC3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.3	Debit	Current account	Services	Transport	Transport	Other modes of transport	49.4437218
17	Q.N.#.W1.S1.S1.T.C.SC4_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.4	Credit	Current account	Services	Transport	Transport	Postal and courier services	0.9293600
18	Q.N.#.W1.S1.S1.T.D.SC4_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab3.4	Debit	Current account	Services	Transport	Transport	Postal and courier services	1.0387000
19	Q.N.#.W1.S1.S1.T.C.SD1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.1	Credit	Current account	Services	Travel	Travel	Goods	0.0000000
20	Q.N.#.W1.S1.S1.T.D.SD1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.1	Debit	Current account	Services	Travel	Travel	Goods	0.0000000
21	Q.N.#.W1.S1.S1.T.C.SD2_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.2	Credit	Current account	Services	Travel	Travel	Local transport services	NA
22	Q.N.#.W1.S1.S1.T.D.SD2_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.2	Debit	Current account	Services	Travel	Travel	Local transport services	NA
23	Q.N.#.W1.S1.S1.T.C.SD3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.3	Credit	Current account	Services	Travel	Travel	Accommodation services	NA
24	Q.N.#.W1.S1.S1.T.D.SD3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.3	Debit	Current account	Services	Travel	Travel	Accommodation services	NA
25	Q.N.#.W1.S1.S1.T.C.SD4_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.4	Credit	Current account	Services	Travel	Travel	Food-serving services	NA
26	Q.N.#.W1.S1.S1.T.D.SD4_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.4	Debit	Current account	Services	Travel	Travel	Food-serving services	NA
27	Q.N.#.W1.S1.S1.T.C.SD5_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.5	Credit	Current account	Services	Travel	Travel	Other services	0.0000000
28	Q.N.#.W1.S1.S1.T.D.SD5_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.Ab4.0.5	Debit	Current account	Services	Travel	Travel	Other services	0.0000000

As we can see, we have kept the SDMX codes in the variable called "sdmx", we have created three time series variables ("refperiod", "period" and "year") and we have kept the instrument codes in the "imfcode" variable. We have an indicator variable to tell us whether a row represents a "credit" or "debit" value (as indicated by the variable "balance"). We then have the instrument hierarchy that was explained above as indicated by the variables "account", "instrumentgroup", "instrumentuppersubgroup", "instrumentsubgroup" and "instrument". And finally, we have the "amount" variable containing the numeric value of each row.
From this data structure, it is easy to add the net amounts while still maintaining the same data structure.

To calculate net amounts, we will be summing all the rows of the "amount" variable grouped by every other variable. If we multiply all "debits" by -1 and sum the rows, we will effectively get the net amounts.

We first have to replace all *NA*s in the "amount" variable with "0" (because we cannot sum *NA*s):

```
#Replace NA amounts with zero

CA_KA_df$amount[is.na(CA_KA_df$amount)] <- c(0)
```

We then do the calculation where we sum the rows (multiplying debits by -1) and store the results in a temporary data frame called "CA_KA_df_net". We also make sure that the

"balance" variable always takes the value "Net receipts" in this data frame (so that we can distinguish the net amounts from the debits and credits):

```
# Create net amounts
CA_KA_df_net<-CA_KA_df%>%select(sdmx,refperiod,period,year,imfcode,balance,account,instrumentgroup,instrumentuppersubgroup,instrumentsubgroup,instrument,amount)%>%
                            group_by(sdmx,refperiod,period,year,imfcode,balance,account,instrumentgroup,instrumentuppersubgroup,instrumentsubgroup,instrument)%>%
                            summarise(amount=sum(ifelse(balance %in% "Debit",amount*-1,amount)))%>%ungroup()%>%
                            mutate(balance="Net receipts")%>%group_by(sdmx,refperiod,period,year,imfcode,balance,account,instrumentgroup,instrumentuppersubgroup,instrumentsubgroup,instrument)%>%
                            summarise(amount=sum(amount))%>%ungroup()%>%arrange(period)
```

The data frame "CA_KA_df_net" will look like this:

	sdmx	refperiod	period	year	imfcode	balance	account	instrumentgroup	instrumentuppersubgroup	instrumentsubgroup	instrument	amount
1	Q.N.#.W1.S1.S1.T.CD1_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.B.1	Net receipts	Current account	Primary income	Compensation of employees	Compensation of employees	Compensation of employees	189.1161704
2	Q.N.#.W1.S1.S1.T.CD2_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.B.3.1	Net receipts	Current account	Primary income	Other primary income	Other primary income	Taxes on production and imports	0.0000000
3	Q.N.#.W1.S1.S1.T.CD3_Z_Z_Z.$_T_XN	2005Q1	200503	2005	1.B.3.2	Net receipts	Current account	Primary income	Other primary income	Other primary income	Subsidies	0.0000000
4	Q.N.#.W1.S1.S1.T.CD41.O.F.A.T.$_T_XN	2005Q1	200503	2005	1.B.2.3.2	Net receipts	Current account	Primary income	Other investment	Interest from loans and deposits	Interest	22.5494351
5	Q.N.#.W1.S1.S1.T.CD41.P.F3.L.$_T_XN	2005Q1	200503	2005	1.B.2.2.2.2	Net receipts	Current account	Primary income	Portfolio investment	Interest from debt securities	Long-term	0.0000000
6	Q.N.#.W1.S1.S1.T.CD41.P.F3.S.$_T_XN	2005Q1	200503	2005	1.B.2.2.2.1	Net receipts	Current account	Primary income	Portfolio investment	Interest from debt securities	Short-term	0.0000000
7	Q.N.#.W1.S1.S1.T.CD41.R.F.T.$.XL_XN	2005Q1	200503	2005	1.B.2.4.2	Net receipts	Current account	Primary income	Reserve assets	Interest (Reserve assets)	Interest	18.6225000
8	Q.N.#.W1.S1.S1.T.CD42.P.F51._Z.$_T_XN	2005Q1	200503	2005	1.B.2.2.1.1	Net receipts	Current account	Primary income	Portfolio investment	Dividends on equity excluding investment fund shares	Dividends on equity excluding investment fund shares	0.0000000
9	Q.N.#.W1.S1.S1.T.CD422.O.F5._Z.$_T_XN	2005Q1	200503	2005	1.B.2.3.1	Net receipts	Current account	Primary income	Other investment	Withdrawals from income of quasi-corporations	Withdrawals from income of quasi-corporations	0.0000000
10	Q.N.#.W1.S1.S1.T.CD42S.D.F5._Z.$_T_XN	2005Q1	200503	2005	1.B.2.1.1.1	Net receipts	Current account	Primary income	Direct investment	Income on equity and investment fund shares	Dividends and withdrawals from income of quasi-corporatio...	0.0000000
11	Q.N.#.W1.S1.S1.T.CD43S.D.F5._Z.$_T_XN	2005Q1	200503	2005	1.B.2.1.1.2	Net receipts	Current account	Primary income	Direct investment	Income on equity and investment fund shares	Reinvested earnings	0.0000000
12	Q.N.#.W1.S1.S1.T.CD431.P.F52._Z.$_T_XN	2005Q1	200503	2005	1.B.2.2.1.2.1	Net receipts	Current account	Primary income	Portfolio investment	Investment income attributable to investment fund shareholders	Dividends	0.0000000
13	Q.N.#.W1.S1.S1.T.CD432.P.F52._Z.$_T_XN	2005Q1	200503	2005	1.B.2.2.1.2.2	Net receipts	Current account	Primary income	Portfolio investment	Investment income attributable to investment fund shareholders	Reinvested earnings	0.0000000
14	Q.N.#.W1.S1.S1.T.CD44P.O.F6._Z.$_T_XN	2005Q1	200503	2005	1.B.2.3.3	Net receipts	Current account	Primary income	Other investment	Investment income attributable to policyholders in insurance...	Investment income attributable to policyholders in insurance...	0.0000000
15	Q.N.#.W1.S1.S1.T.CD45._Z._Z.$_T_XN	2005Q1	200503	2005	1.B.3.3	Net receipts	Current account	Primary income	Other primary income	Other primary income	Rent	0.0000000
16	Q.N.#.W1.S1.S1.T.CD4Q.D.F1._Z.$_T_XN	2005Q1	200503	2005	1.B.2.1.2	Net receipts	Current account	Primary income	Direct investment	Interest from inter-company debt	Interest	0.4320000
17	Q.N.#.W1.S1.S1.T.CD45.R.F5._Z.$.XL_XN	2005Q1	200503	2005	1.B.2.4.1	Net receipts	Current account	Primary income	Reserve assets	Income on equity and investment fund shares (Reserve asse...	Income on equity and investment fund shares	0.0000000
18	Q.N.#.W1.S1.S1.T.CG1._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.a.1	Net receipts	Current account	Goods	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	General merchandise on a balance of payments basis	331.9433359
19	Q.N.#.W1.S1.S1.T.CG21._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.a.2.1	Net receipts	Current account	Goods	Goods acquired under merchanting (negative credit)	Goods acquired under merchanting (negative credit)	Goods acquired under merchanting (negative credit)	0.0000000
20	Q.N.#.W1.S1.S1.T.CG22._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.a.2.2	Net receipts	Current account	Goods	Goods sold under merchanting (credit)	Goods sold under merchanting (credit)	Goods sold under merchanting (credit)	0.0000000
21	Q.N.#.W1.S1.S1.T.CNP._Z._Z.$_T_XN	2005Q1	200503	2005	2.1	Net receipts	Capital account	Gross acquisitions (DR.) / disposals (CR.) of nonproduced no...	Gross acquisitions (DR.) / disposals (CR.) of nonproduced no...	Gross acquisitions (DR.) / disposals (CR.) of nonproduced no...	Gross acquisitions (DR.) / disposals (CR.) of nonproduced no...	0.0000000
22	Q.N.#.W1.S1.S1.T.CSA._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.b.1	Net receipts	Current account	Services	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	Manufacturing services on physical inputs owned by others	124.1039534
23	Q.N.#.W1.S1.S1.T.CSB._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.b.2	Net receipts	Current account	Services	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	Maintenance and repair services n.i.e.	0.0000000
24	Q.N.#.W1.S1.S1.T.CSC1._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.b.3.1	Net receipts	Current account	Services	Transport	Transport	Sea transport	0.0000000
25	Q.N.#.W1.S1.S1.T.CSC2._Z._Z.$_T_XN	2005Q1	200503	2005	1.A.b.3.2	Net receipts	Current account	Services	Transport	Transport	Air transport	7.6425187

We now simply have to union the two data frames "CA_KA_df" and "CA_KA_df_net" to a common data frame (we just call this "CA_KA_df" as well) which is readily done with the "rbind" function:

```
#Combine credits and debits with net receipts

CA_KA_df<-rbind(CA_KA_df,CA_KA_df_net)


############################################################
# End of current account and capital account
############################################################
```

Hence, the final data frame "CA_KA_df", containing all relevant data (on a credit, debit and net basis) for the current account and capital account, will have a database-structure like this:

	sdmx	refperiod	period	year	imfcode	balance	account	instrumentgroup	instrumentuppersubgroup	instrumentsubgroup	instrument	amount
1371	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2016Q4	201612	2016	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	3.384913647
1372	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2016Q4	201612	2016	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.029289730
1373	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2017Q1	201703	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.021638894
1374	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2017Q1	201703	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.057151310
1375	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2017Q2	201706	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.591666596
1376	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2017Q2	201706	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-2.169882078
1377	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2017Q3	201709	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.200103205
1378	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2017Q3	201709	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.181161829
1379	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2017Q4	201712	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.446591305
1380	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2017Q4	201712	2017	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.067744311
1381	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2018Q1	201803	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.123802783
1382	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2018Q1	201803	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.067529572
1383	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2018Q2	201806	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.750833255
1384	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2018Q2	201806	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.105207979
1385	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2018Q3	201809	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.320113526
1386	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2018Q3	201809	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.079325930
1387	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2018Q4	201812	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.591250435
1388	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2018Q4	201812	2018	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.095618610
1389	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2019Q1	201903	2019	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.236183062
1390	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2019Q1	201903	2019	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.367034562
1391	Q.N.#.W1.S1.S1.T.CS1._Z._Z.Z.$_T_XN	2019Q2	201906	2019	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	1.678257258
1392	Q.N.#.W1.S1.S1.T.D.S1._Z._Z.Z.$_T_XN	2019Q2	201906	2019	1.A.b.10.1	Net receipts	Current account	Services	Other business services	Other business services	Research and development services	-0.100946871
1393	Q.N.#.W1.S1.S1.T.CS2._Z._Z.Z.$_T_XN	2005Q1	200503	2005	1.A.b.10.2	Credit	Current account	Services	Other business services	Other business services	Professional and management consulting services	0.000000000
1394	Q.N.#.W1.S1.S1.T.D.S2._Z._Z.Z.$_T_XN	2005Q1	200503	2005	1.A.b.10.2	Debit	Current account	Services	Other business services	Other business services	Professional and management consulting services	0.000000000
1395	Q.N.#.W1.S1.S1.T.CS2._Z._Z.Z.$_T_XN	2005Q2	200506	2005	1.A.b.10.2	Credit	Current account	Services	Other business services	Other business services	Professional and management consulting services	0.000000000
1396	Q.N.#.W1.S1.S1.T.D.S2._Z._Z.Z.$_T_XN	2005Q2	200506	2005	1.A.b.10.2	Debit	Current account	Services	Other business services	Other business services	Professional and management consulting services	0.000000000
1397	Q.N.#.W1.S1.S1.T.CS2._Z._Z.Z.$_T_XN	2005Q3	200509	2005	1.A.b.10.2	Credit	Current account	Services	Other business services	Other business services	Professional and management consulting services	0.000000000
1398	Q.N.#.W1.S1.S1.T.D.S2._Z._Z.Z.$_T_XN	2005Q3	200509	2005	1.A.b.10.2	Debit	Current account	Services	Other business services	Other business services	Professional and management consulting services	0.000000000

The data for the current account and the capital account has now been prepared and the next step now is to create a similar data structure for the financial account.
We start off by filtering the "BPM6BOPForm_df" that we previously created. We rely on the fact that all items belonging to the financial account can easily be identified through the pattern ".FA." in the SDMX codes:

```
#########################################################
# Start of financial account
#########################################################


FA_df<- BPM6BOPForm_df%>%filter(sdmx %like% c("%.FA.%"))
```

The next many lines of codes are more or less the same that were used to prepare the "CA_KA_df" (the current account and capital account), though with a few modifications given that the financial account representation is different from the current and capital account representation.

However, the resulting data structure will be the same as can be seen here:

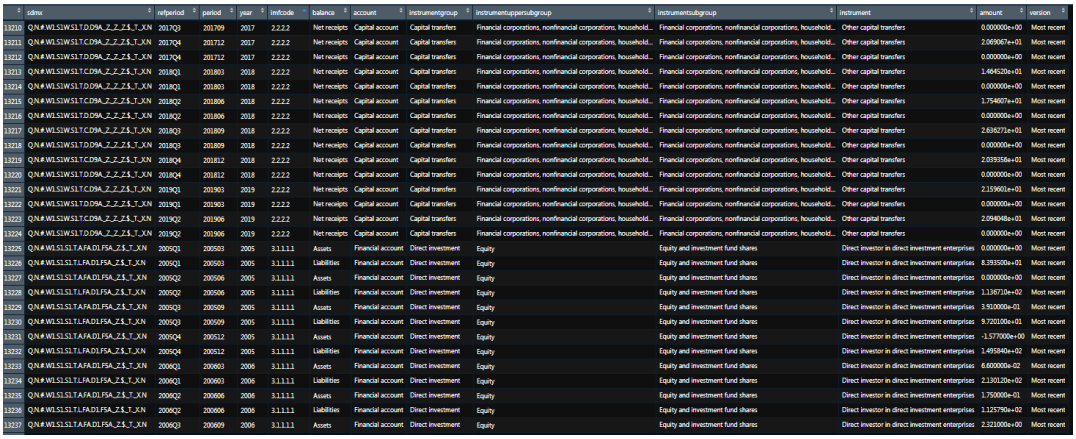
	sdmx	refperiod	period	year	imfcode	balance	account	instrumentgroup	instrumentuppersubgroup	instrumentsubgroup	instrument	amount
6438	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2013Q2	201306	2013	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.130311e+02
6503	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2013Q3	201309	2013	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.161218e+01
6568	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2013Q4	201312	2013	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.867100e+02
6633	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2014Q1	201403	2014	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-4.340156e+00
6698	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2014Q2	201406	2014	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.235079e+02
6763	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2014Q3	201409	2014	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.074882e+02
6828	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2014Q4	201412	2014	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	3.636692e+01
6893	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2015Q1	201503	2015	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.481459e+01
6958	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2015Q2	201506	2015	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	4.605862e+01
7023	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2015Q3	201509	2015	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.090176e+01
7088	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2015Q4	201512	2015	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.293475e+02
7153	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2016Q1	201603	2016	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.532560e+01
7218	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2016Q2	201606	2016	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-3.322675e+01
7283	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2016Q3	201609	2016	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.305218e+02
7348	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2016Q4	201612	2016	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.724642e+02
7413	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2017Q1	201703	2017	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	8.012325e+01
7478	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2017Q2	201706	2017	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	3.288000e+01
7543	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2017Q3	201709	2017	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	6.641241e+01
7608	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2017Q4	201712	2017	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.711369e+02
7673	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2018Q1	201803	2018	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.609333e+01
7738	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2018Q2	201806	2018	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.776437e+02
7803	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2018Q3	201809	2018	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.066459e+02
7868	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2018Q4	201812	2018	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.473225e+02
7933	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2019Q1	201903	2019	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	3.784463e+00
7998	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2019Q2	201906	2019	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	3.010052e+02
8063	Q.N.#.W1.S1.S1.T.AFA.D1.FSA_Z.$_T_XDC_N	2005Q1	200503	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	0.000000e+00
8083	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2005Q1	200503	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-8.393500e+01
8208	Q.N.#.W1.S1.S1.T.AFA.D1.FSA_Z.$_T_XDC_N	2005Q2	200506	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	0.000000e+00
8222	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2005Q2	200506	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-1.136710e+02
8347	Q.N.#.W1.S1.S1.T.AFA.D1.FSA_Z.$_T_XDC_N	2005Q3	200509	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	3.910000e-01
8361	Q.N.#.W1.S1.S1.T.LFA.D1.FSA_Z.$_T_XDC_N	2005Q3	200509	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-9.720100e+01
8488	Q.N.#.W1.S1.S1.T.AFA.D1.FSA_Z.$_T_XDC_N	2005Q4	200512	2005	3.1.1.1.1	Net assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-1.577000e+00

We now have a complete data frame called "FA_df" that contains all relevant data from the financial account, along with the SDMX codes ("sdmx"), instrument codes ("imfcode"), time series variables ("refperiod", "period" and "year"), "balance" indicator (indicating whether a row is an "Asset", "Liability" or "Net asset") , the instrument hierarchy ("account", "instrumentgroup", "instrumentuppersubgroup", "instrumentsubgroup" and "instrument") and the "amount" variable.

We can now combine the final data for the current account and capital account with the data for the financial account. Hence, we combine the data frames "CA_KA_df" and "FA_df" via the "rbind" function and store it in a new data frame called "BOP_most_recent" (we add a new variable "version" to indicate that this is the most recent data):

```
########################################
#Combine all accounts
########################################
BOP_most_recent<-rbind(CA_KA_df,FA_df)%>%mutate(version="Most recent")
########################################################################

#End of most recent data
########################################################################
```

We now have the most recent balance of payments data ready in a database-like structure looking like this:



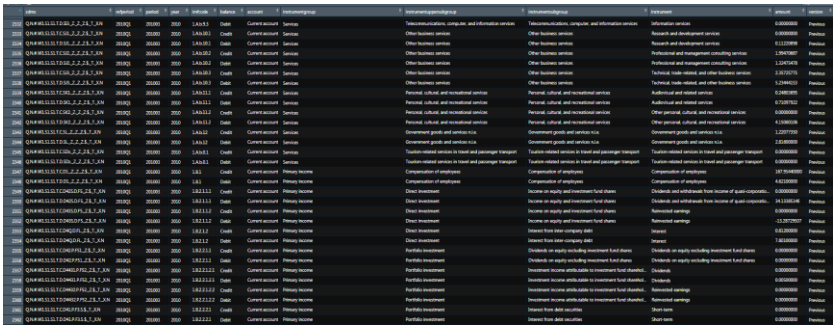
	sdmx	refperiod	period	year	imfcode	balance	account	instrumentgroup	instrumentupperpsubgroup	instrumentsubgroup	instrument	amount	version
13210	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2017Q3	201709	2017	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13211	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2017Q4	201712	2017	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	2.069067e+01	Most recent
13212	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2017Q4	201712	2017	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13213	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2018Q1	201803	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	1.464520e+01	Most recent
13214	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2018Q1	201803	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13215	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2018Q2	201806	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	1.754607e+01	Most recent
13216	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2018Q2	201806	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13217	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2018Q3	201809	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	2.638272e+01	Most recent
13218	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2018Q3	201809	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13219	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2018Q4	201812	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	2.039358e+01	Most recent
13220	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2018Q4	201812	2018	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13221	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2019Q1	201903	2019	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	2.159601e+01	Most recent
13222	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2019Q1	201903	2019	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13223	Q.N.#.W1.S1W.S1.T.C.D9A_Z_Z_Z.$_T_XN	2019Q2	201906	2019	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	2.094048e+01	Most recent
13224	Q.N.#.W1.S1W.S1.T.D.D9A_Z_Z_Z.$_T_XN	2019Q2	201906	2019	2.2.2.2	Net receipts	Capital account	Capital transfers	Financial corporations, nonfinancial corporations, household...	Financial corporations, nonfinancial corporations, household...	Other capital transfers	0.000000e+00	Most recent
13225	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2005Q1	200503	2005	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	0.000000e+00	Most recent
13226	Q.N.#.W1.S1.S1.T.L.FA.D.I.FSA_Z.$_T_XN	2005Q1	200503	2005	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	8.393500e+01	Most recent
13227	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2005Q2	200506	2005	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	0.000000e+00	Most recent
13228	Q.N.#.W1.S1.S1.T.L.FA.D.I.FSA_Z.$_T_XN	2005Q2	200506	2005	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.136710e+02	Most recent
13229	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2005Q3	200509	2005	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	3.910000e-01	Most recent
13230	Q.N.#.W1.S1.S1.T.L.FA.D.I.FSA_Z.$_T_XN	2005Q3	200509	2005	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	9.720100e+01	Most recent
13231	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2005Q4	200512	2005	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	-1.577000e+00	Most recent
13232	Q.N.#.W1.S1.S1.T.L.FA.D.I.FSA_Z.$_T_XN	2005Q4	200512	2005	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.495840e+02	Most recent
13233	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2006Q1	200603	2006	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	6.600000e-02	Most recent
13234	Q.N.#.W1.S1.S1.T.L.FA.D.I.FSA_Z.$_T_XN	2006Q1	200603	2006	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.130120e+02	Most recent
13235	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2006Q2	200606	2006	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.750000e-01	Most recent
13236	Q.N.#.W1.S1.S1.T.L.FA.D.I.FSA_Z.$_T_XN	2006Q2	200606	2006	3.1.1.1.1	Liabilities	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	1.125750e+02	Most recent
13237	Q.N.#.W1.S1.S1.T.A.FA.D.I.FSA_Z.$_T_XN	2006Q3	200609	2006	3.1.1.1.1	Assets	Financial account	Direct investment	Equity	Equity and investment fund shares	Direct investor in direct investment enterprises	2.321000e+00	Most recent

We now have to do the same for the previous data.
We first read in the relevant excel file (the one specified by "path_previous") and store it in a temporary data frame called "BPM6BOPForm_revision":

```
#-------------------------------------

# Create data based on previous data

########################################################################

# BPM6BOPForm
BPM6BOPForm_revision <-read_excel(paste0(path_previous),  sheet = "BPM6BOPForm",skip=7)
```

Based on this data frame, we basically just repeat all the codes that were used to create the data frame "BOP_most_recent".

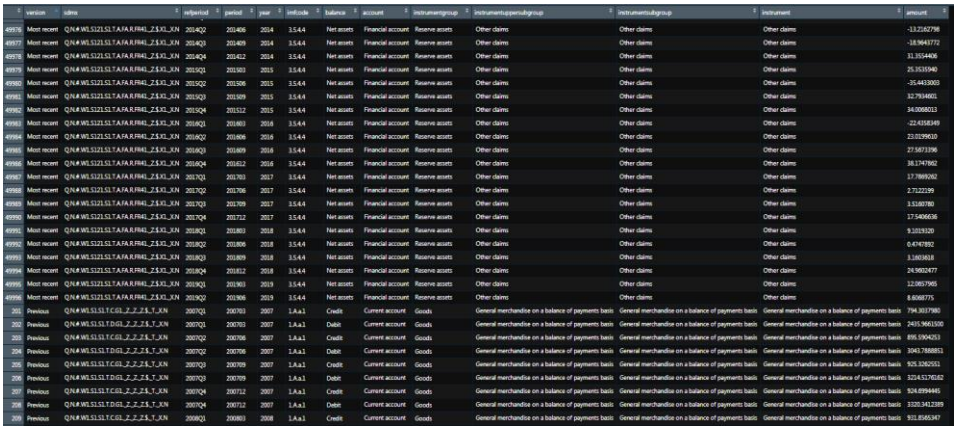
The resulting data frame is called "BOP_revision" and contains all balance of payments data in a database-like structure. We ensure to add the value "Previous" to the "version" variable in order to indicate that all values in the "BOP_revision" data frame are based on the previous excel-file. The resulting data frame "BOP_revision" looks like this:



We can now combine the most recent balance of payments data with the previous version in a common data frame called "BOP_data":

```
############################################################
#Combine most recent data with revision data
############################################################

BOP_data<-rbind(BOP_most_recent,BOP_revision)%>%filter(period>200612)%>%
select(version,sdmx,refperiod,period,year,imfcode,balance,account,instrumentgroup,instrumentuppersubgroup,instrumentsubgroup,instrument,amount)
```

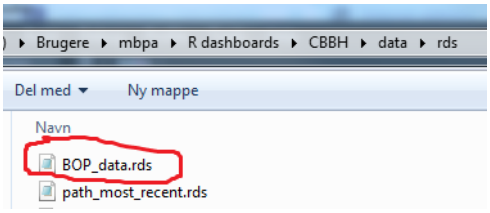
The final data frame "BOP_data" will look like this:



The dataset for the balance of payments data is now complete and we can now save the final data frame as an RDS-file to be located in the RDS-folder:

```
#Saving dataset in rds format for reactive reading in dashboard session
saveRDS(BOP_data,paste(save_dir,"BOP_data.rds",sep = "/"))
```

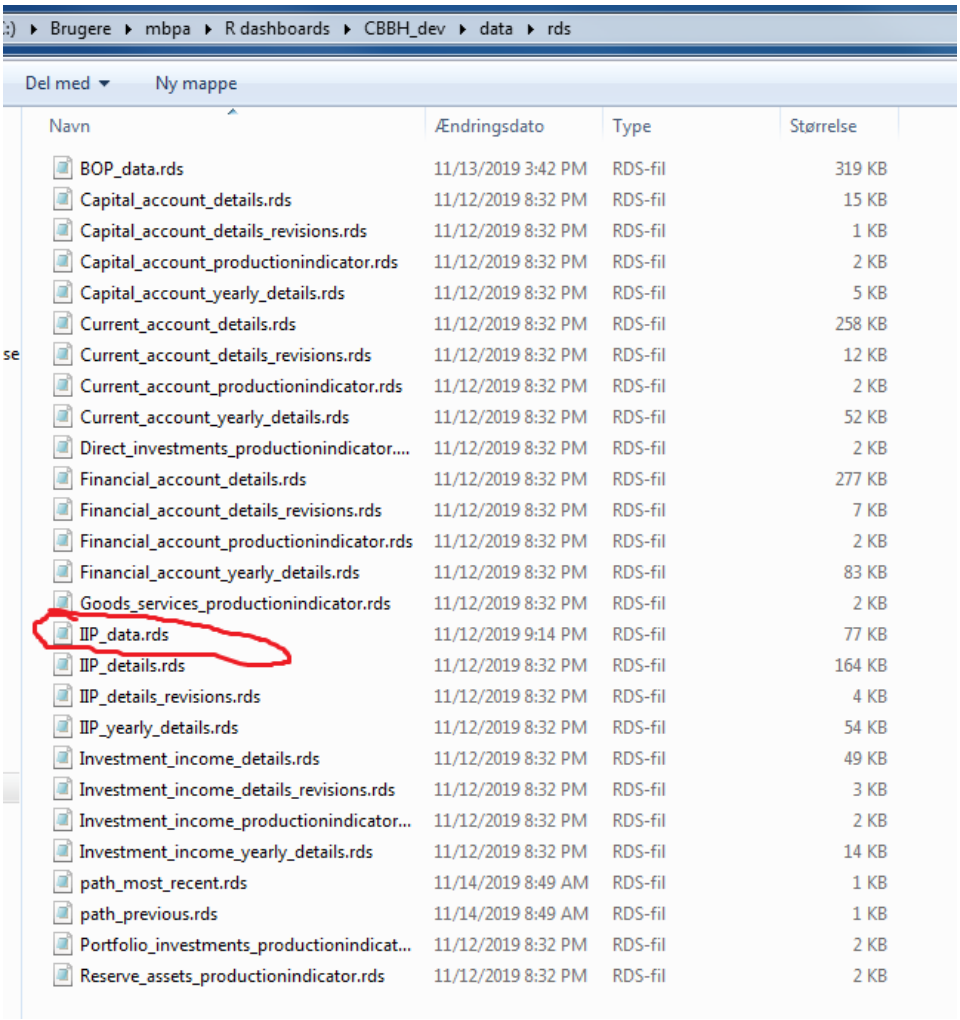
Location of the RDS-file with all the relevant balance of payments data:



We can now remove the temporary data frames that were used to create "BOP_data".

```
#Remove old dataframes
rm(list=c("BOP_most_recent","BOP_revision","BPM6BOPForm_df","BPM6BOPForm_revision","CA_KA_df","CA_KA_df_net","CA_KA_revision","CA_KA_revision_net","FA_df","FA_df_net",
          "FA_revision","FA_revision_net","imf_fact","imf_fact_revision"))

############################################################
```

The final part of the "data_all.R" scripts now aims to create a similar database-like structure for the most recent and previous version of the international investment position data. The resulting data frame called "IIP_Kvartalni_df" is saved as an RDS-file in the RDS-folder:


Brugere ▸ mbpa ▸ R dashboards ▸ CBBH_dev ▸ data ▸ rds
Del med
Ny mappe
Navn
Ændringsdato
Type
Størrelse
BOP_data.rds 11/13/2019 3:42 PM RDS-fil 319 KB
Capital_account_details.rds 11/12/2019 8:32 PM RDS-fil 15 KB
Capital_account_details_revisions.rds 11/12/2019 8:32 PM RDS-fil 1 KB
Capital_account_productionindicator.rds 11/12/2019 8:32 PM RDS-fil 2 KB
Capital_account_yearly_details.rds 11/12/2019 8:32 PM RDS-fil 5 KB
Current_account_details.rds 11/12/2019 8:32 PM RDS-fil 258 KB
Current_account_details_revisions.rds 11/12/2019 8:32 PM RDS-fil 12 KB
Current_account_productionindicator.rds 11/12/2019 8:32 PM RDS-fil 2 KB
Current_account_yearly_details.rds 11/12/2019 8:32 PM RDS-fil 52 KB
Direct_investments_productionindicator.... 11/12/2019 8:32 PM RDS-fil 2 KB
Financial_account_details.rds 11/12/2019 8:32 PM RDS-fil 277 KB
Financial_account_details_revisions.rds 11/12/2019 8:32 PM RDS-fil 7 KB
Financial_account_productionindicator.rds 11/12/2019 8:32 PM RDS-fil 2 KB
Financial_account_yearly_details.rds 11/12/2019 8:32 PM RDS-fil 83 KB
Goods_services_productionindicator.rds 11/12/2019 8:32 PM RDS-fil 2 KB
IIP_data.rds 11/12/2019 9:14 PM RDS-fil 77 KB
IIP_details.rds 11/12/2019 8:32 PM RDS-fil 164 KB
IIP_details_revisions.rds 11/12/2019 8:32 PM RDS-fil 4 KB
IIP_yearly_details.rds 11/12/2019 8:32 PM RDS-fil 54 KB
Investment_income_details.rds 11/12/2019 8:32 PM RDS-fil 49 KB
Investment_income_details_revisions.rds 11/12/2019 8:32 PM RDS-fil 3 KB
Investment_income_productionindicator... 11/12/2019 8:32 PM RDS-fil 2 KB
Investment_income_yearly_details.rds 11/12/2019 8:32 PM RDS-fil 14 KB
path_most_recent.rds 11/14/2019 8:49 AM RDS-fil 1 KB
path_previous.rds 11/14/2019 8:49 AM RDS-fil 1 KB
Portfolio_investments_productionindicat... 11/12/2019 8:32 PM RDS-fil 2 KB
Reserve_assets_productionindicator.rds 11/12/2019 8:32 PM RDS-fil 2 KB