

```
In [3]: library(magrittr)
library(dplyr)
library(stringr)
library(countrycode)
```

Dataset CPIS

Importo dataset.

```
In [1]: full_data <- read.csv("CPIS.csv",
                             sep=";", header=TRUE)
head(full_data[,1:4])
head(full_data[,5:8])
```

A data.frame: 6 × 4

	year	holder_country	holder_country_code	issuer_country
	<int>	<chr>	<int>	<chr>
1	2001	Argentina	213	Afghanistan, Islamic Rep. of
2	2001	Argentina	213	Albania
3	2001	Argentina	213	Algeria
4	2001	Argentina	213	American Samoa
5	2001	Argentina	213	Andorra, Principality of
6	2001	Argentina	213	Angola

A data.frame: 6 × 4

	issuer_country_code	total_assets	five_year	five_yrs_avg_total_assets
	<int>	<dbl>	<int>	<dbl>
1	512	0	2000	0
2	914	0	2000	0
3	612	0	2000	0
4	859	0	2000	0
5	171	0	2000	0
6	614	0	2000	0

Limito il dataset all'anno, al nome del paese creditore e debitore, e all'ammontare del credito in USD.

```
In [5]: new_data <- full_data %>%
select(year, issuer_country, holder_country, total_assets)
```

```
colnames(new_data)<-c("year","origin","destination","value_cpis")

#riordino così le colonne
new_data <- new_data %>%
  select(origin,destination,year,value_cpis)
```

Utilizzo la funzione della library "countrycode" per trasformare i nomi dei paesi in codici ISO3c. Non vengono riconosciuti dei paesi che non sono scritti in modo corretto: li inserisco a mano; altri paesi, invece, van tolti, in quanto si riferiscono a entità geografiche non più esistenti in quei confini.

```
In [6]: paesi_nome <- c(unique(new_data$origin),setdiff(
  unique(new_data$destination),unique(new_data$origin)))
paesi_iso3c <- countrycode(paesi_nome,origin = 'country.name',
  destination = 'iso3c')
paesi <- as.data.frame(cbind(paesi_nome,paesi_iso3c))
paesi$paesi_iso3c[10] <- "ABW" #aruba
paesi$paesi_iso3c[90] <- "GUF" #guyana french
paesi$paesi_iso3c[205] <- "STP" #Sao Tome and Principe
paesi$paesi_iso3c[244] <- "CUW" #l'ha trascritto sbagliato
paesi$paesi_iso3c[245] <- "SXM" #sint maarten
paesi$paesi_iso3c[246] <- NA #qui mette insieme due paesi
```

Warning message:

"Some values were not matched unambiguously: Aruba, Kingdom of the Netherlands, Guiana, French, International Organizations, Kosovo, Rep. of, Netherlands Antilles, Not Specified (including Confidential), SÃ£o TomÃ© and PrÃ­ncipe, Dem. Rep. of, Sint Maarten, Kingdom of the Netherlands, US Pacific Islands, World
"

Warning message:

"Some strings were matched more than once, and therefore set to <NA> in the result: Aruba, Kingdom of the Netherlands,ABW,NLD; Sint Maarten, Kingdom of the Netherlands, NLD,SXM
"

Successivamente, cambio il paese creditore e debitore con il corrispondente ISO3c così determinato.

```
In [7]: cpis <- new_data
cpis$origin <- paesi$paesi_iso3c[
  match(as.factor(cpis$origin),
    as.factor(paesi$paesi_nome))]
cpis$destination <- paesi$paesi_iso3c[
  match(as.factor(cpis$destination),
    as.factor(paesi$paesi_nome))]
#rimuovo le righe con i NA dei paesi non considerati
cpis <- na.omit(cpis)
```

Rimuovo i link interni, ossia che hanno l'importatore uguale all'esportatore. Infine, creo un dataset completo con tutte le possibili combinazioni di *origin*, *destination* e *year* (tra quelli presenti nel dataset originale).

```
In [8]: dati <- which(cpis$origin == cpis$destination)
cpis_ripulito <- cpis[-dati,]

p <- as.character(na.omit(paesi$paesi_iso3c))
y <- unique(cpis$year)
tot <- as.data.frame(cbind(rep(p, each = length(p)*length(y)),
                           rep(rep(p,times=length(p)),each=length(y)),
                           rep(y, times = length(p)^2)))
colnames(tot) <- c("origin","destination","year")
tot <- tot[-c(which(tot$origin == tot$destination)),]
tot <- cbind(tot,rep(0,nrow(tot)))
colnames(tot) <- c("origin","destination","year","value_cpis")
tot$year <- as.numeric(tot$year)
merge <- rbind(cpis_ripulito,tot)
cpis_tot <- merge %>%
  group_by(origin,destination,year) %>%
  summarise(value_cpis = sum(value_cpis))
```

`summarise()` has grouped output by 'origin', 'destination'. You can override using the `.groups` argument.

Valori negativi

```
In [9]: print(paste0("Ci sono ", length(which(cpis$value_cpis < 0)),
                    " valori negativi: li elimino"))
cpis$value_cpis[which(cpis$value_cpis < 0)] <- 0
```

```
[1] "Ci sono 170 valori negativi: li elimino"
```

```
In [10]: cpis_tot <- cpis_tot[order(cpis_tot$year),]
write.csv(cpis_tot,"cpis_tot.csv", row.names = FALSE)
```