

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

Dataset CDIS

Importo il dataset contenente 86 variabili. Le prime riportano il nomi dei paesi, il rispettivo codice ISO3c dei medesimi che dichiarano di inviare o di ricevere l'investimento diretto e l'anno in cui viene registrato l'investimento; le successivi colonne contengono diverse osservazioni: il valore degli strumenti del debito sia lordi che netti, sia intesi come debiti (*liabilities*) sia come crediti (*asset*), sia nei confronti di un'impresa di investimento diretto sia nei confronti di un'impresa "amica" (*fellow enterprise*), di cui non si possiede il 10% del potere di voto; di ognuna di queste variabili c'è il valore "*derived*" ossia il corrispettivo valore del flusso di investimento ottenuto considerando ciò che dichiara il paese partner nello scambio. Ho ritenuto opportuno mantenere, oltre alle colonne che riportano l'ISO3c dei paesi creditori e debitori e l'anno in cui viene registrato lo scambio, solo le colonne contenenti il valore degli investimenti diretti totali,escludendo i debiti/crediti con le *fellow enterprises*, differenziati per direzionalità, ossia considerando il caso *outward* e *inward*, ottenuti dalla somma tra ogni tipo di strumento di debito (asset-liabilities nel caso outward, e liabilities-asset nel caso inward) e il netto dell'equity. E' il dato che più si avvicina anche ai dati del dataset UNCTAD.

Sia nel caso outward che inward mi interessa studiare il collegamento tra due paesi in cui il paese di origin è il paese creditore e il paese di destination è quello debitore, indipendentemente da chi dichiara l'investimento. Per questo, decido di selezionare la colonna che riporta il valore degli investimenti diretti totali outward e la colonna che riporta il valore degli investimento diretti totali outward derived (che sono gli stessi degli investimenti diretti totali inward ponendo come paese di destinazione il paese reporting e quello di origin il paese partner).

```
In [9]: full_data <- read.csv("CDIS1.csv",
                             sep=";", header=TRUE)
#riscrivo meglio le colonne del dataset
colnames(full_data) <- str_split_fixed(colnames(full_data), "..IOW",n=2)[,1]
colnames(full_data) <- str_split_fixed(colnames(full_data), "..IIW",n=2)[,1]
colnames(full_data)[1] <- "Country.Name"
#tolgo le colonne status che sono ogni 2 a partire dalla 7ima colonna
#(che riportano solo se il paese ha voluto rilasciare il dato oppure no)
#rimuovo anche il country code (lascio i paesi scritti con il nome)
# e tolgo anche l'ultima colonna che ha tutti elementi NA
remove <- c(2,4,seq(7,85,2),86)
new_data <- full_data[,-c(remove)]
head(new_data[,1:3])
```

```
head(new_data[,4:8])
colnames(new_data)
```

A data.frame: 6 × 3

	Country.Name	Counterpart.Country.Name	Time.Period
	<chr>	<chr>	<int>
1	Afghanistan, Islamic Rep. of	Türkkiye, Rep of	2009
2	Afghanistan, Islamic Rep. of	Türkkiye, Rep of	2010
3	Afghanistan, Islamic Rep. of	Türkkiye, Rep of	2011
4	Afghanistan, Islamic Rep. of	Türkkiye, Rep of	2012
5	Afghanistan, Islamic Rep. of	Türkkiye, Rep of	2013
6	Afghanistan, Islamic Rep. of	Türkkiye, Rep of	2014

	Inward.Debt.Instruments.Assets.Positions..Gross...Derived..US.Dollars	Inward.Debt.Instrument
		<dbl>
1		0
2		0
3		0
4		0
5		0
6		0

'Country.Name' · 'Counterpart.Country.Name' · 'Time.Period' ·
 'Inward.Debt.Instruments.Assets.Positions..Gross...Derived..US.Dollars' ·
 'Inward.Debt.Instruments.Assets.Positions..Gross...US.Dollars' ·
 'Inward.Debt.Instruments.Liabilities.Positions..Gross...Derived..US.Dollars' ·
 'Inward.Debt.Instruments.Liabilities.Positions..Gross...US.Dollars' ·
 'Inward.Debt.Instruments.Positions..Net...Derived..US.Dollars' ·
 'Inward.Debt.Instruments.Positions..Net...US.Dollars' ·
 'Inward.Debt.Positions..Net...Resident.Enterprises.that.are.not.Financial.Intermediaries..Derived..US.Dollars' ·
 'Inward.Debt.Positions..Net...Resident.Enterprises.that.are.not.Financial.Intermediaries..US.Dollars' ·
 'Inward.Debt.Positions..Net...Resident.Financial.Intermediaries..Derived..US.Dollars' ·
 'Inward.Debt.Positions..Net...Resident.Financial.Intermediaries..US.Dollars' ·
 'Inward.Direct.Investment.Assets.Positions..Gross..with.Fellow.Enterprises..Derived..US.Dollars' ·
 'Inward.Direct.Investment.Assets.Positions..Gross..with.Fellow.Enterprises..US.Dollars' ·
 'Inward.Direct.Investment.Liabilities.Positions..Gross..with.Fellow.Enterprises..Derived..US.Dollars' ·
 'Inward.Direct.Investment.Liabilities.Positions..Gross..with.Fellow.Enterprises..US.Dollars' ·
 'Inward.Direct.Investment.Positions..Net..with.Fellow.Enterprises..Derived..US.Dollars' ·
 'Inward.Direct.Investment.Positions..Net..with.Fellow.Enterprises..US.Dollars' ·
 'Inward.Direct.Investment.Positions..Derived..US.Dollars' ·
 'Inward.Direct.Investment.Positions..US.Dollars' ·
 'Inward.Equity.Positions..Net...Derived..US.Dollars' · 'Inward.Equity.Positions..Net...US.Dollars' ·
 'Outward.Debt.Instruments.Assets.Positions..Gross...Derived..US.Dollars' ·
 'Outward.Debt.Instruments.Assets.Positions..Gross...US.Dollars' ·
 'Outward.Debt.Instruments.Liabilities.Positions..Gross...Derived..US.Dollars' ·
 'Outward.Debt.Instruments.Liabilities.Positions..Gross...US.Dollars' ·
 'Outward.Debt.Instruments.Positions..Net...Derived..US.Dollars' ·
 'Outward.Debt.Instruments.Positions..Net...US.Dollars' ·
 'Outward.Debt.Positions..Net...Resident.Enterprises.that.are.not.Financial.Intermediaries..Derived..US.Dollars' ·
 'Outward.Debt.Positions..Net...Resident.Enterprises.that.are.not.Financial.Intermediaries..US.Dollars' ·
 'Outward.Debt.Positions..Net...Resident.Financial.Intermediaries..Derived..US.Dollars' ·
 'Outward.Debt.Positions..Net...Resident.Financial.Intermediaries..US.Dollars' ·
 'Outward.Direct.Investment.Assets.Positions..Gross..with.Fellow.Enterprises..Derived..US.Dollars' ·
 'Outward.Direct.Investment.Assets.Positions..Gross..with.Fellow.Enterprises..US.Dollars' ·
 'Outward.Direct.Investment.Liabilities.Positions..Gross..with.Fellow.Enterprises..Derived..US.Dollars' ·
 'Outward.Direct.Investment.Liabilities.Positions..Gross..with.Fellow.Enterprises..US.Dollars' ·
 'Outward.Direct.Investment.Positions..Net..with.Fellow.Enterprises..Derived..US.Dollars' ·
 'Outward.Direct.Investment.Positions..Net..with.Fellow.Enterprises..US.Dollars' ·
 'Outward.Direct.Investment.Positions..Derived..US.Dollars' ·
 'Outward.Direct.Investment.Positions..US.Dollars' ·
 'Outward.Equity.Positions..Net...Derived..US.Dollars' · 'Outward.Equity.Positions..Net...US.Dollars'

```

In [4]: fdi_cdis <- new_data %>%
  select(Country.Name,Counterpart.Country.Name,Time.Period,
    Outward.Direct.Investment.Positions..US.Dollars,

```

```

Outward.Direct.Investment.Positions..Derived..US.Dollars)
colnames(fdi_cdis) <- c("origin","destination","year","outward_cdis",
                        "outward_derived_cdis")

```

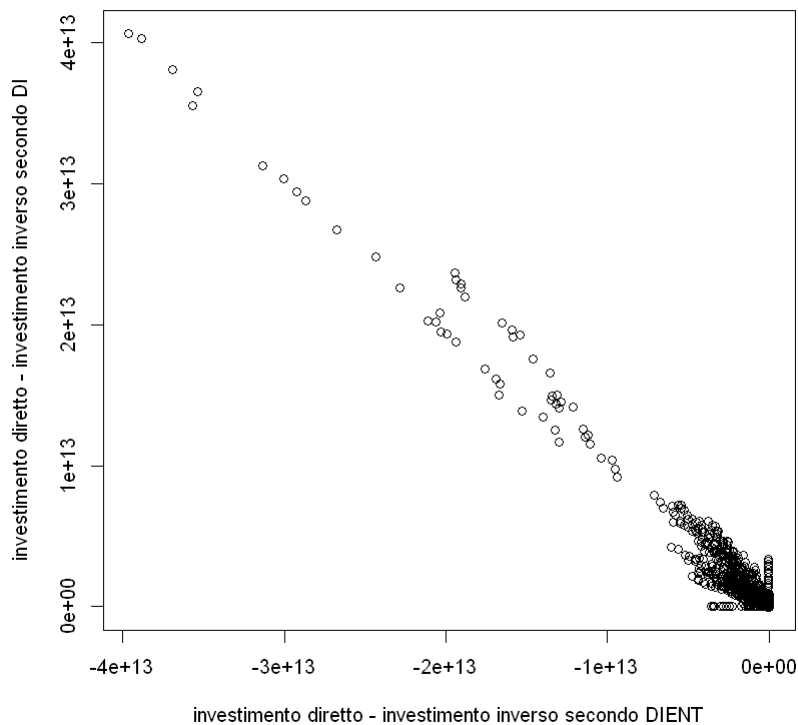
```

In [8]: data_cdis <- data.frame(y = fdi_cdis$outward_cdis,
                                x = (-1)*fdi_cdis$outward_derived_cdis)

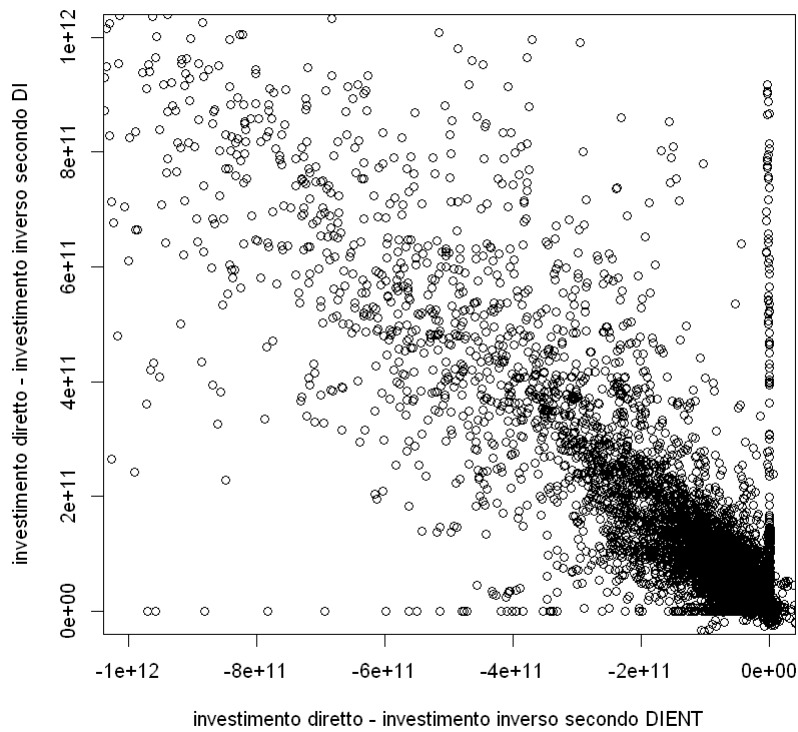
mfrow =c(1,3)
plot(data_cdis$x,data_cdis$y,
     xlab = "investimento diretto - investimento inverso secondo DIENT",
     ylab = "investimento diretto - investimento inverso secondo DI",
     main = "Confronto tra valori di investimento diretto-investimento inverso")
plot(data_cdis$x,data_cdis$y,
     xlab = "investimento diretto - investimento inverso secondo DIENT",
     ylab = "investimento diretto - investimento inverso secondo DI",
     main = "Confronto tra valori di investimento diretto-investimento inverso",
     ylim = c(0,1e+12),xlim = c(-1e+12,0))
plot(data_cdis$x,data_cdis$y,
     xlab = "investimento diretto - investimento inverso secondo DIENT",
     ylab = "investimento diretto - investimento inverso secondo DI",
     main = "Confronto tra valori di investimento diretto-investimento inverso",
     ylim = c(0,4e+11),xlim = c(-4e+11,0))

```

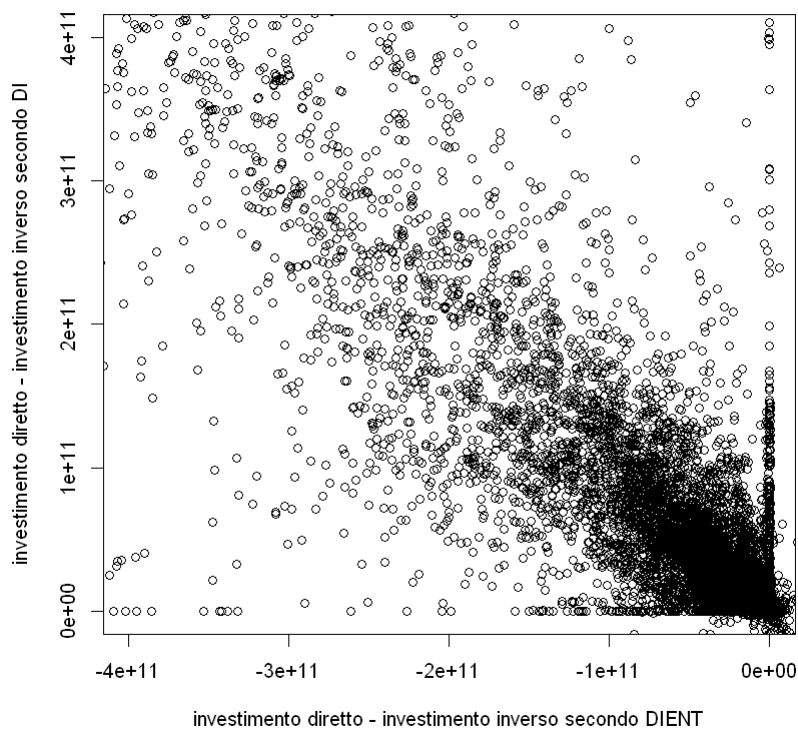
Confronto tra valori di investimento diretto-investimento inverso



Confronto tra valori di investimento diretto-investimento inverso



Confronto tra valori di investimento diretto-investimento inverso



Utilizzo la funzione della library "countrycode" per trasformare i nomi dei paesi in codici ISO3c. Alcuni nomi si riferiscono ad aree geografiche molto generali: lascio il valore NA al posto dell'ISO3c, altri li sostituisco io manualmente perchè i nomi non sono scritti in modo

corretto. Successivamente, sostituisco i nomi dei paesi presenti nel dataset con i rispettivi ISO3c così determinati.

```
In [ ]: paesi_name <- c(unique(fdi_cdis$origin),
                        setdiff(unique(fdi_cdis$destination),
                                unique(fdi_cdis$origin)))
paesi_iso3c <- countrycode(paesi_name, origin = "country.name",
                           destination = "iso3c")
paesi <- as.data.frame(cbind(paesi_name,paesi_iso3c))
paesi$paesi_iso3c[237] <- "TUR"
paesi$paesi_iso3c[11] <- "ABW"
#Economies of Persian Gulf che ha sostituito con IRN:
paesi$paesi_iso3c[68] <- NA
paesi$paesi_iso3c[207] <- "SXM"
paesi$paesi_iso3c[60] <- "CUW"
paesi$paesi_iso3c[200] <- "STP"

fdi_cdis <- fdi_cdis[order(fdi_cdis$year),]

fdi_cdis$origin <- paesi$paesi_iso3c[
  match(as.factor(fdi_cdis$origin),
        as.factor(paesi$paesi_name))]
fdi_cdis$destination <- paesi$paesi_iso3c[
  match(as.factor(fdi_cdis$destination),
        as.factor(paesi$paesi_name))]

#rimuovo le righe con i NA dei paesi non considerati
fdi_cdis <- na.omit(fdi_cdis)
```

Creo un dataset completo con tutte le possibili combinazioni di *origin*, *destination* e *year* (tra quelli presenti nel dataset originale).

```
In [ ]: p <- as.character(na.omit(paesi$paesi_iso3c))
y <- unique(fdi_cdis$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)),rep(0,nrow(tot)))
colnames(tot) <- c("origin","destination","year","outward","outward_derived")
tot$year <- as.numeric(tot$year)
merge <- rbind(fdi_cdis,tot)
fdi_cdis_tot <- merge %>%
  group_by(origin,destination,year) %>%
  summarise(outward = sum(outward),outward_derived=sum(outward_derived))
fdi_cdis_tot <- fdi_cdis_tot[order(fdi_cdis_tot$year),]
colnames(fdi_cdis_tot) <- c("origin","destination","year","outward_cdis",
                          "outward_derived_cdis")
```

```
In [ ]: #cdis_fin <- cdis_confronto %>%
# mutate(dummy_inward = rep(0,times=nrow(cdis_confronto)), dummy_outward = rep(1,
#cdis_fin <- cdis_fin %>%
# select(origin,destination,year,inward_cdis,dummy_inward,outward_cdis,dummy_outwa
```

```
#write.csv(fdi_cdis_tot, "fdi_cdis.csv", row.names = FALSE)
write.csv(fdi_cdis_tot, "cdis_fromcredto debt.csv", row.names = FALSE)
```