

```
In [56]: library(magrittr)
library(dplyr)
library(reshape)
library(countrycode)
library(readxl)
```

UNCTAD

Importo lista di file: due relativi ai flussi di investimento (flow) e uno relativo agli stock di investimento (stock).

```
In [57]: list_of_files <- list.files(path = "C:", recursive = TRUE,
                                     pattern = "\\.", full.names = TRUE)
```

Dataset FLOW

Per i primi due file relativi ai flow sistemo il dataset rinominando le colonne, togliendo le prime righe e le ultime colonne che hanno informazioni non pertinenti, togliendo quelle righe in cui l'origine e la destinazione del flusso sono lo stesso paese. Per questi due dataset i paesi presi in considerazione sono gli stessi.

```
In [58]: flow_1 <- suppressMessages(read_excel(list_of_files[1], col_names = FALSE))
colnames(flow_1) <- flow_1[2,]
flow_1 <- flow_1[-c(1:2),]
flow_1 <- flow_1[which(flow_1$FDI == "Flows"),] #o erano Flows o NA
flow_1 <- flow_1[, -3] #a sto punto tolgo la colonna
flow_2 <- suppressMessages(read_excel(list_of_files[2], col_names = FALSE))
colnames(flow_2) <- flow_2[5,]
flow_2 <- flow_2[-c(1:5),]
flow_2 <- flow_2[which(flow_2$FDI == "Flows"),] # o erano Flows o NA
flow_2 <- flow_2[, -3] #a sto punto tolgo la colonna
flow_2 <- flow_2[-c(which(flow_2[,37] == TRUE)),]
#a quanto pare sono i flowi dove origin e destination son uguali
flow_1 <- flow_1[-c(which(flow_1[,1] == flow_1[,2])),]
#controllo che non ci siano anche in flow_1
flow_2 <- flow_2[, -c(36:37)] #togliamo il gran total e la colonna di veri e falsi
colnames(flow_1) <- colnames(flow_2) <- c("reporting",
                                           "partner", "direction", c(1990:2021))

paesi_1 <- c(unique(flow_1$partner),
             setdiff(unique(flow_1$reporting), unique(flow_1$partner)))
paesi_2 <- c(unique(flow_2$partner),
             setdiff(unique(flow_2$reporting), unique(flow_2$partner)))
head(flow_1)
```

reporting	partner	direction	1990	1991	1992	1993	1994	1995	1996	...	20
<chr>	<chr>	<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	...	<db
Albania	Austria	Inward	NA	NA	NA	NA	NA	NA	NA	...	NA
Albania	Bulgaria	Inward	NA	NA	NA	NA	NA	NA	NA	...	NA
Albania	Cayman Islands	Inward	NA	NA	NA	NA	NA	NA	NA	...	NA
Albania	China	Inward	NA	NA	NA	NA	NA	NA	NA	...	NA
Albania	Croatia	Inward	NA	NA	NA	NA	NA	NA	NA	...	NA
Albania	Cyprus	Inward	NA	NA	NA	NA	NA	NA	NA	...	NA

Del valore dei flussi di investimento distinguo tra la direzionalità del flusso. Prima considero solo quelli di inward (di entrambi i dataset), ossia dove l'origine è il paese che dichiara di ricevere un investimento diretto (DIENT) mentre la destinazione è il paese dove risiede l'investitore diretto (DI). Il valore di questo flusso è la differenza tra l'investimento inverso e l'investimento diretto. Una volta considerati solo i dati inward riscrivo il dataset con il format panel (tramite la funzione melt): ottengo un dataset con 4 colonne (origin, destination, year, value); successivamente raggruppo le righe che hanno le prime tre colonne uguali sommando i valori del flusso.

```
In [59]: ###primo dataset
new_inward_1<-as.data.frame(flow_1[which(flow_1[,3] == "Inward"),-3])
new_inward_1[is.na(new_inward_1)] <- 0
new_inward_1[,3] <- as.numeric(new_inward_1[,3])
resh_1_inward <- melt(new_inward_1,id=c(1,2))
colnames(resh_1_inward) <- c("origin","destination","year","value1")
inward_1_ridotto <- resh_1_inward %>%
  group_by(origin,destination,year) %>%
  summarise(value1 = sum(value1))

###secondo dataset
new_inward_2<-as.data.frame(flow_2[which(flow_2[,3] == "Inward"),-3])
new_inward_2[is.na(new_inward_2)] <- 0
new_inward_2[,3] <- as.numeric(new_inward_2[,3])
resh_2_inward <- melt(new_inward_2,id=c(1,2))
colnames(resh_2_inward) <- c("origin","destination","year","value2")
inward_2_ridotto <- resh_2_inward %>%
  group_by(origin,destination,year) %>%
  summarise(value2 = sum(value2))
```

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

Ora unisco i due dataset in un unico per poi confrontarli. Per ottenere ciò creo un dataset con tutte le possibili combinazioni tra origin e destination (dati i paesi totali presenti tra i

due dataset) e anni, inizialmente con valori nulli che andrò a riempire laddove c'è il corrispettivo valore in uno dei due dataset o in entrambi.

```
In [60]: rm(new_inward_1)
rm(new_inward_2)
rm(resh_1_inward)
rm(resh_2_inward)
```

```
In [61]: p <- paesi_1
y <- c(1990:2021)
flow_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(flow_tot) <- c("origin","destination","year")
flow_tot <- flow_tot[-c(which(flow_tot$origin ==
                             flow_tot$destination)),]
flow_tot <- cbind(flow_tot,rep(0,nrow(flow_tot)))
colnames(flow_tot) <- c("origin","destination","year","value1")

merge1 <- rbind(inward_1_ridotto,flow_tot)
fdi_flow_inward1 <- merge1 %>%
  group_by(origin,destination,year) %>%
  summarise(value1 = sum(value1))

colnames(flow_tot)[4] <- "value2"
merge2 <- rbind(inward_2_ridotto,flow_tot)
fdi_flow_inward2 <- merge2 %>%
  group_by(origin,destination,year) %>%
  summarise(value2 = sum(value2))

#creo un unico dataset
fdi_flow_inward <- cbind(fdi_flow_inward1,fdi_flow_inward2$value2)
colnames(fdi_flow_inward) <- c("origin","destination","year","value1","value2")
```

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

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

New names:

* `` -> `...5`

```
In [62]: rm(inward_1_ridotto)
rm(inward_2_ridotto)
rm(fdi_flow_inward1)
rm(fdi_flow_inward2)
rm(merge1)
rm(merge2)
```

Confronto i valori non nulli di entrambi i dataset.

```
In [63]: sovrapposizione <- intersect(which(fdi_flow_inward$value1 != 0),
                                       which(fdi_flow_inward$value2 != 0))
diversi <- which(fdi_flow_inward$value1[sovrapposizione] !=
                fdi_flow_inward$value2[sovrapposizione])
```

```

print(paste0("Su ", length(sovrapposizione),
" scambi in cui il valore del primo e del secondo dataset",
"sono entrambi non nulli, solo ",
length(diversi), " riportano un valore diverso"))
data_unctad_inward <- data.frame(
  x = fdi_flow_inward$value1[sovrapposizione[diversi]],
  y = fdi_flow_inward$value2[sovrapposizione[diversi]])
print(paste0("Nel primo dataset ci sono ",
length(which(fdi_flow_inward$value1 != 0))
- length(sovrapposizione), " scambi in più del secondo"))
print(paste0("Nel secondo dataset ci sono ",
length(which(fdi_flow_inward$value2 != 0))
- length(sovrapposizione), " scambi in più del primo"))

```

```

[1] "Su 113030 scambi in cui il valore del primo e del secondo dataset sono entrambi
non nulli, solo 3362 riportano un valore diverso"
[1] "Nel primo dataset ci sono 83 scambi in più del secondo"
[1] "Nel secondo dataset ci sono 1852 scambi in più del primo"

```

Vediamo lo scatterplot tra i due dataset quando riportano due valori diversi per lo stesso scambio:

```

In [64]: plot(data_unctad_inward$x,
  data_unctad_inward$y, xlim = c(-1000,1000), ylim = c(-1000,1000),
  xlab = "1 dataset", ylab = "2 dataset",
  main = "Confronto tra i due dataset di flow inward")
regr <- lm(y ~ x, data = data_unctad_inward)
summary(regr)
abline(coefficients(regr)[1],coefficients(regr)[2],col="blue",lwd = 3)

```

Call:

```
lm(formula = y ~ x, data = data_unctad_inward)
```

Residuals:

Min	1Q	Median	3Q	Max
-37179	-15	-15	-9	40922

Coefficients:

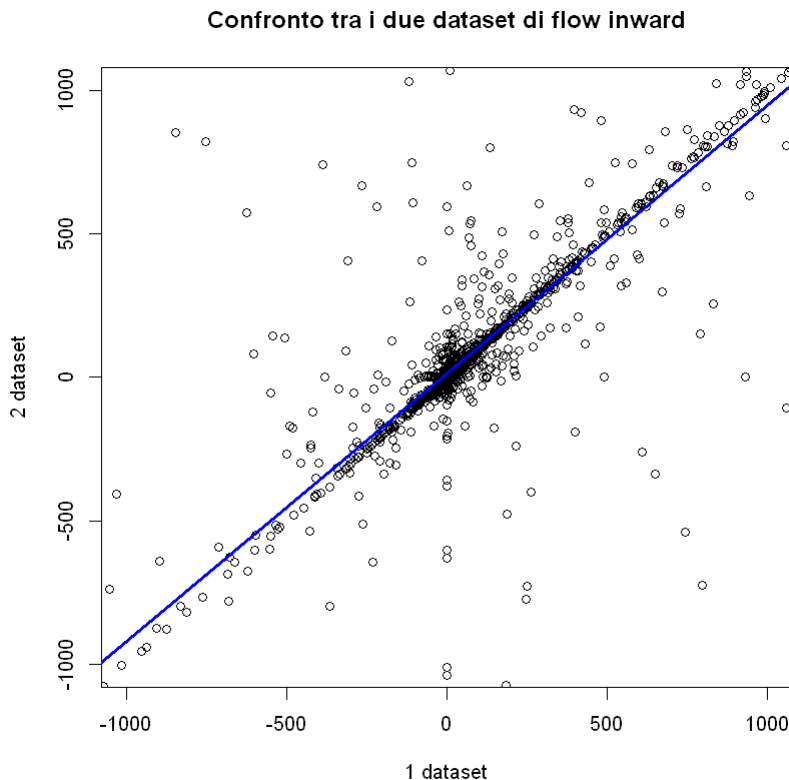
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	15.034360	23.416685	0.642	0.521
x	0.933151	0.007283	128.136	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1352 on 3360 degrees of freedom

Multiple R-squared: 0.8301, Adjusted R-squared: 0.8301

F-statistic: 1.642e+04 on 1 and 3360 DF, p-value: < 2.2e-16



Facciamo lo stesso con i dati di flusso di investimento di tipo outward, ossia dove l'origine è il paese che dichiara dove risiede l'investitore diretto (DI) mentre la destinazione è il paese dove risiede l'impresa di investimento diretto (DIENT). Il valore di questo flusso è la differenza tra l'investimento diretto e l'investimento inverso.

```
In [65]: ##primo dataset
new_outward_1<-as.data.frame(flow_1[which(flow_1[,3] == "Outward"),-3])
new_outward_1[is.na(new_outward_1)] <- 0
new_outward_1[,3] <- as.numeric(new_outward_1[,3])
resh_1_outward <- melt(new_outward_1,id=c(1,2))
colnames(resh_1_outward) <- c("origin","destination","year","value1")
outward_1_ridotto <- resh_1_outward %>%
  group_by(origin,destination,year) %>%
  summarise(value1 = sum(value1))

##secondo dataset
new_outward_2<-as.data.frame(flow_2[which(flow_2[,3] == "Outward"),-3])
new_outward_2[is.na(new_outward_2)] <- 0
new_outward_2[,3] <- as.numeric(new_outward_2[,3])
resh_2_outward <- melt(new_outward_2,id=c(1,2))
colnames(resh_2_outward) <- c("origin","destination","year","value2")
outward_2_ridotto <- resh_2_outward %>%
  group_by(origin,destination,year) %>%
  summarise(value2 = sum(value2))
```

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

```
In [66]: rm(flow_1)
rm(flow_2)
rm(new_outward_1)
rm(new_outward_2)
rm(resh_1_outward)
rm(resh_2_outward)
```

Creo un unico dataset.

```
In [67]: colnames(flow_tot)[4] <- "value1"
merge1 <- rbind(outward_1_ridotto,flow_tot)
fdi_flow_outward1 <- merge1 %>%
  group_by(origin,destination,year) %>%
  summarise(value1 = sum(value1))

colnames(flow_tot)[4] <- "value2"
merge2 <- rbind(outward_2_ridotto,flow_tot)
fdi_flow_outward2 <- merge2 %>%
  group_by(origin,destination,year) %>%
  summarise(value2 = sum(value2))

fdi_flow_outward <- cbind(fdi_flow_outward1,fdi_flow_outward2$value2)
colnames(fdi_flow_outward) <- c("origin","destination",
                                "year","value1","value2")
```

`summarise()` has grouped output by 'origin', 'destination'. You can override using the `.groups` argument.
`summarise()` has grouped output by 'origin', 'destination'. You can override using the `.groups` argument.
New names:
* `` -> `...5`

```
In [68]: rm(flow_tot)
rm(outward_1_ridotto)
rm(outward_2_ridotto)
rm(fdi_flow_outward1)
rm(fdi_flow_outward2)
rm(merge1)
rm(merge2)
```

Confronto i due dataset anche utilizzando i dati outward.

```
In [69]: sovrapposizione <- intersect(which(fdi_flow_outward$value1 != 0),
                                         which(fdi_flow_outward$value2 != 0))
diversi <- which(fdi_flow_outward$value1[sovrapposizione] !=
                fdi_flow_outward$value2[sovrapposizione])
print(paste0("Su ", length(sovrapposizione),
              " scambi in cui il valore del primo e del secondo dataset",
              "sono entrambi non nulli, solo ",
              length(diversi), " riportano un valore diverso"))
data_unctad_outward <- data.frame(
  x = fdi_flow_outward$value1[sovrapposizione[diversi]],
  y = fdi_flow_outward$value2[sovrapposizione[diversi]])
print(paste0("Nel primo dataset ci sono ",
```

```

length(which(fdi_flow_outward$value1 != 0))
- length(sovrapposizione), " scambi in più del secondo"))
print(paste0("Nel secondo dataset ci sono ",
length(which(fdi_flow_outward$value2 != 0))
- length(sovrapposizione), " scambi in più del primo"))

```

[1] "Su 82364 scambi in cui il valore del primo e del secondo dataset sono entrambi non nulli, solo 2370 riportano un valore diverso"

[1] "Nel primo dataset ci sono 44 scambi in più del secondo"

[1] "Nel secondo dataset ci sono 1522 scambi in più del primo"

Creo grafico.

```

In [70]: plot(data_unctad_outward$x,
data_unctad_outward$y, xlim = c(-1000,1000), ylim = c(-1000,1000),
xlab = "1 dataset", ylab = "2 dataset",
main = "confronto tra i due dataset di flow outward")
regr <- lm(y ~ x, data = data_unctad_outward)
summary(regr)
abline(coefficients(regr)[1],coefficients(regr)[2],col="blue", lwd = 3)

```

Call:

lm(formula = y ~ x, data = data_unctad_outward)

Residuals:

Min	1Q	Median	3Q	Max
-15397.2	6.2	10.1	29.0	24952.5

Coefficients:

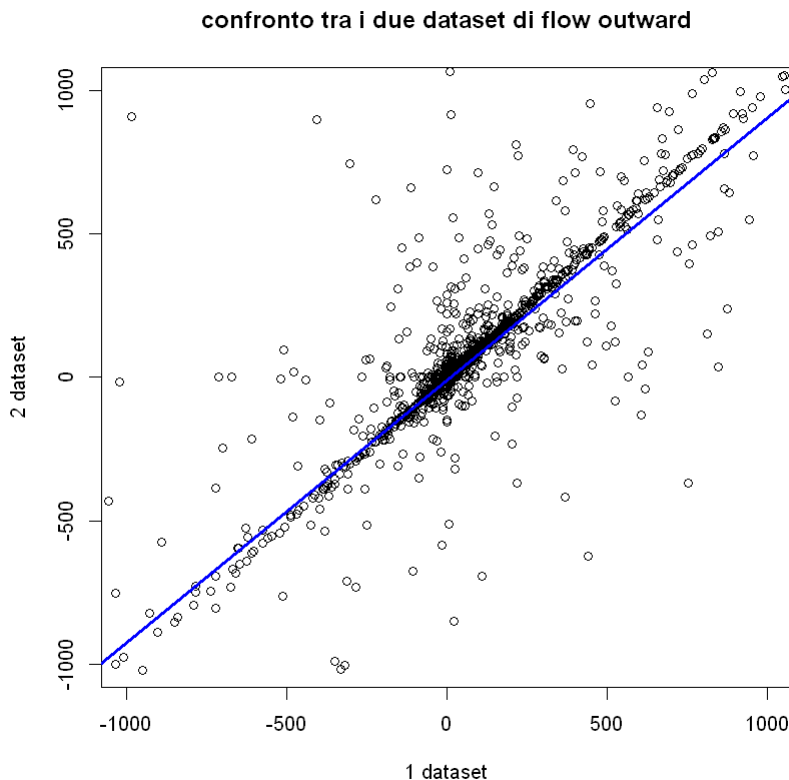
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-9.814678	23.323221	-0.421	0.674
x	0.914669	0.005805	157.568	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1130 on 2368 degrees of freedom

Multiple R-squared: 0.9129, Adjusted R-squared: 0.9129

F-statistic: 2.483e+04 on 1 and 2368 DF, p-value: < 2.2e-16



Decido di prendere il secondo dataset in quanto ha più dati. Creo un dataset flow unico di investimento inward e outward.

```
In [71]: flow_unctad <- as.data.frame(cbind(fdi_flow_inward[,1:3],
      inward = fdi_flow_inward$value2,
      outward = fdi_flow_outward$value2))
```

Converto il nome completo dei paesi nel loro corrispettivo codice ISO3c attraverso la funzione countrycode. Alcuni non li riconosce perchè non sono scritti in modo corretto, altri perchè non esistono più come entità geografiche o non esiste il loro codice iso3c o comprendono aree geografiche troppo generiche.

```
In [72]: paesi_iso3c <- countrycode(paesi_1, origin = "country.name",
      destination = "iso3c")
paesi <- as.data.frame(cbind(paesi_1,paesi_iso3c))
#non ha riconosciuto saint martin: lo scrivo a mano
paesi$paesi_iso3c[241] <- "MAF"
paesi$paesi_iso3c[263] <- NA #soviet union non la conto
```

Warning message:

"Some values were not matched unambiguously: APEC, ASEAN 10, Belgium / Luxembourg, Canary Islands, Channel Islands, Czechoslovakia (former), International Organization s, Kosovo, Melilla, Netherlands Antilles, OPEC, Other Capital, Private buying and selling of property, Reinvested earnings, Saint Martin, Serbia and Montenegro, SPEs, Unspecified Africa, Virgin Islands, West Indies, Yugoslavia (former)
"

C'è un paese il cui codice appare due volte ma si riferisce allo stesso paese che ha cambiato nome.

```
In [73]: p <- na.omit(paesi)
codice_dupl <- p$paesi_iso3c[which(
  duplicated(p$paesi_iso3c) == TRUE)]
print(p$paesi_1[which(
  p$paesi_iso3c == codice_dupl)])
```

```
[1] "Congo, Democratic Republic of" "Zaire"
```

C'è un paese il cui codice appare due volte ma si riferisce allo stesso paese che ha cambiato nome.

```
In [74]: flow_unctad$origin <- paesi$paesi_iso3c[match(
  as.factor(flow_unctad$origin),
  as.factor(paesi$paesi_1))]
flow_unctad$destination <- paesi$paesi_iso3c[match(
  as.factor(flow_unctad$destination),
  as.factor(paesi$paesi_1))]
flow_unctad <- na.omit(flow_unctad)
```

```
In [75]: flow_unctad <- flow_unctad %>%
  group_by(origin, destination, year) %>%
  summarise(inward = sum(inward), outward = sum(outward))
flow_unctad <- flow_unctad[-c(which(flow_unctad$origin ==
  flow_unctad$destination)),]
```

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

```
In [76]: flow_unctad <- flow_unctad[order(flow_unctad$year),]
colnames(flow_unctad) <- c("origin", "destination", "year",
  "inward_flow_unctad", "outward_flow_unctad")
```

```
In [77]: write.csv(flow_unctad, "flow_unctad.csv", row.names = FALSE)
```

Valori negativi

```
In [78]: print(paste0("Nel dataset flow la percentuale dei valori",
  "negativi sul totale dei valori non nulli è del ",
  round(length(which(flow_unctad$inward_flow_unctad < 0))/length(
  which(flow_unctad$inward_flow_unctad != 0))*100,1), "% per inward e del ",
  round(length(which(flow_unctad$outward_flow_unctad < 0))/length(which(
  flow_unctad$outward_flow_unctad != 0))*100,1), "% per outward."))
```

```
[1] "Nel dataset flow la percentuale dei valori negativi sul totale dei valori non nulli è del 23.6% per inward e del 26.2% per outward."
```

Da creditore a debitore

Invertiamo l'origine con la destinazione nel caso inward in modo che l'origin sia il paese dove risiede DI e la destination sia il paese dove risiede DIENT.

```
In [79]: flow_inv <- flow_unctad %>%
  select(destination,origin,year,
         outward_derived_flow_unctad = inward_flow_unctad)
colnames(flow_inv)[1:2] <- colnames(flow_unctad)[1:2]
flow_in <- flow_inv[order(flow_inv$origin),]
flow_confronto <- cbind(flow_in[order(flow_in$year),],
                       flow_unctad$outward_flow_unctad)
#tolgo la quarta colonna che è inward
colnames(flow_confronto) <- c("origin","destination","year",
                             "outward_derived_flow_unctad", "outward_flow_unctad")
```

New names:

```
* `` -> `...5`
```

Confronto graficamente i due valori del flusso di investimento trattandoli entrambi dal punto di vista outward (investimento diretto-investimento inverso) ma in un caso il valore è dichiarato da DI, quindi dal creditore, mentre nell'altro caso è dichiarato da DIENT, quindi dal debitore.

```
In [80]: sovr_unctad <- which(flow_confronto$outward_derived_flow_unctad != 0 &
                             flow_confronto$outward_flow_unctad != 0)
print(paste0("Dal punto di vista inward ci sono ",length(
  which(flow_confronto$outward_derived_flow_unctad != 0)) - length(sovr_unctad),
" valori non nulli per un dato scambio che son nulli,",
      "invece, dal punto di vista outward"))
print(paste0("Dal punto di vista outward ci sono ",length(
  which(flow_confronto$outward_flow_unctad != 0)) - length(sovr_unctad),
" valori non nulli per un dato scambio che son nulli, invece,",
      "dal punto di vista inward"))
```

```
[1] "Dal punto di vista inward ci sono 71583 valori non nulli per un dato scambio che son nulli, invece, dal punto di vista outward"
```

```
[1] "Dal punto di vista outward ci sono 41465 valori non nulli per un dato scambio che son nulli, invece, dal punto di vista inward"
```

Confrontiamo graficamente solo i dati che riportano valori non nulli sia in inward che in outward.

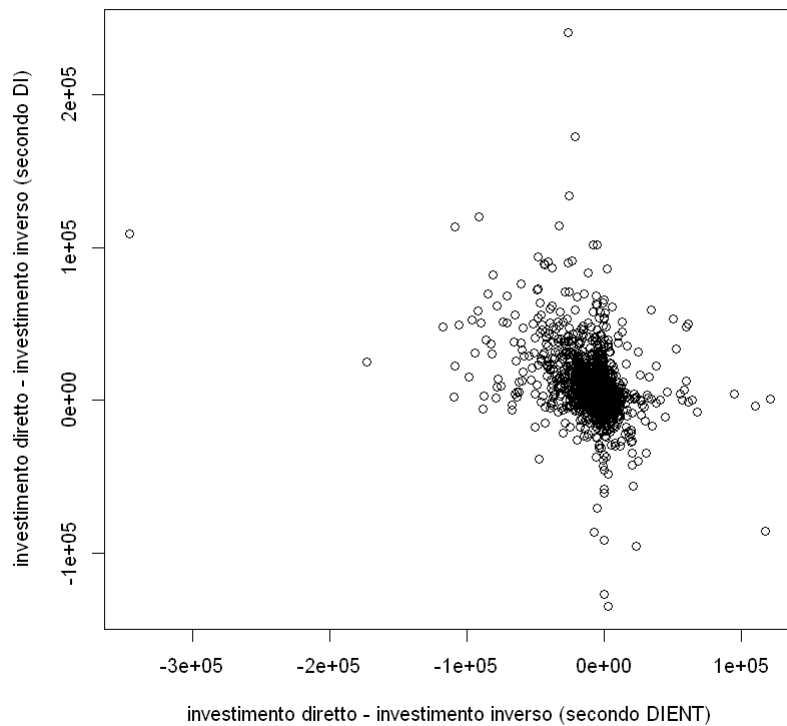
```
In [81]: data_unctad <- data.frame(
  x = (-1)*flow_confronto$outward_derived_flow_unctad[sovr_unctad],
  y = flow_confronto$outward_flow_unctad[sovr_unctad])
plot(data_unctad$x,
     data_unctad$y,
     xlab = "investimento diretto - investimento inverso (secondo DIENT)",
     ylab = "investimento diretto - investimento inverso (secondo DI)",
     main = "Confronto tra flusso di investimento dichiarato da DI e da DIENT")
plot(data_unctad$x,
     data_unctad$y,
     xlab = "investimento diretto - investimento inverso (secondo DIENT)",
     ylab = "investimento diretto - investimento inverso (secondo DI)",
```

```

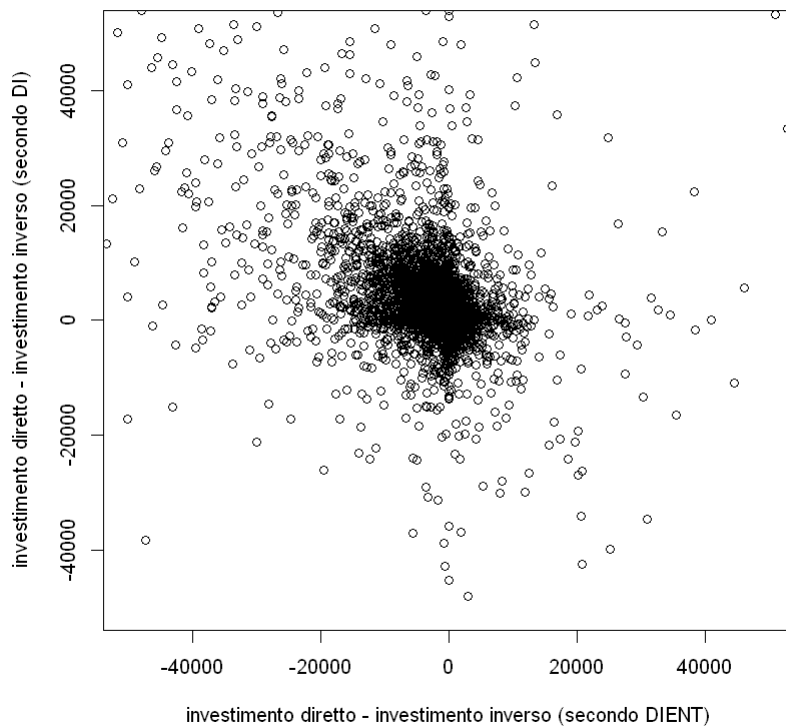
    main = "Confronto tra flusso di investimento dichiarato da DI e da DIENT",
    xlim = c(-5e+4, 5e+04), ylim = c(-5e+04, 5e+04))
plot(data_unctad$x,
     data_unctad$y,
     xlab = "investimento diretto - investimento inverso (secondo DIENT)",
     ylab = "investimento diretto - investimento inverso (secondo DI)",
     main = "Confronto tra flusso di investimento dichiarato da DI e da DIENT",
     xlim = c(-20000, 20000), ylim = c(-20000, 20000))

```

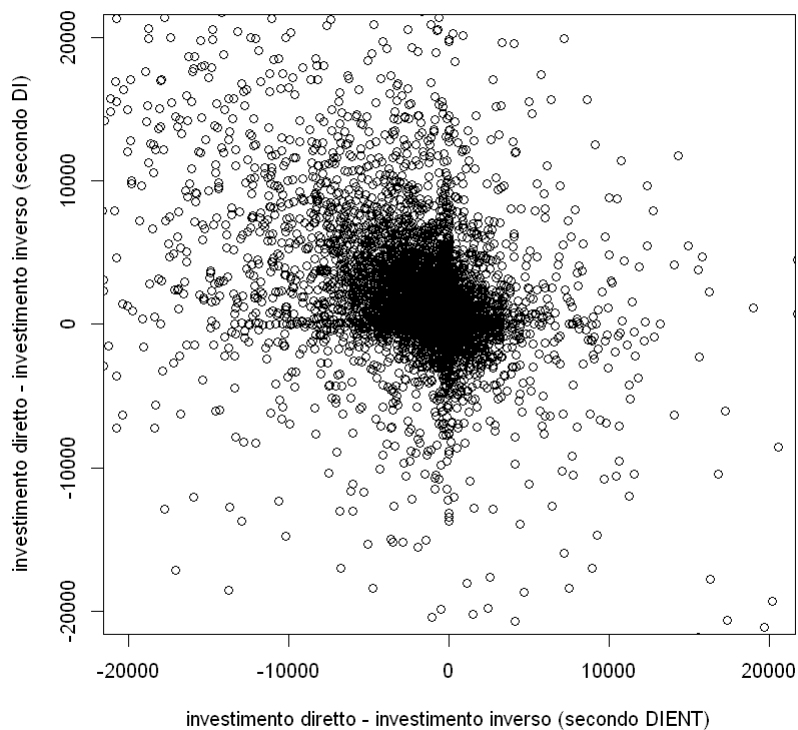
Confronto tra flusso di investimento dichiarato da DI e da DIENT



Confronto tra flusso di investimento dichiarato da DI e da DIENT



Confronto tra flusso di investimento dichiarato da DI e da DIENT



Per la maggior parte dei casi si possono trarre le seguenti osservazioni: se secondo il paese in cui risiede DIENT il suo investimento inverso è maggiore dell'investimento diretto che riceve dal paese dove risiede DI, secondo DI invece il suo investimento diretto in DIENT non eccede l'investimento inverso; in modo analogo, se secondo DIENT, DI non sta investendo

più di quanto DIENT non reinvesta, secondo DI, l'investimento diretto supera di molto l'investimento inverso.

Aggiungo una variabile dummy accanto alle due colonne di flow inward e flow outward dove la riga assume valore 1 se il creditore è anche il paese che dichiara, mentre -1 se il creditore non è il paese che dichiara.

```
In [82]: #flow_unctad_fin <- flow_confronto %>%
# mutate(dummy_outward = rep(1,times=nrow(flow_confronto)), dummy_inward = rep(-1,
#flow_unctad_fin <- flow_unctad_fin %>%
# select(origin,destination,year,outward_derived_flow_unctad = outward_derived_unc
write.csv(flow_confronto,"flow_unctad_fromcredtoebt.csv", row.names = FALSE)
```

Dataset STOCK

Analizziamo ora il dataset sugli stock, quindi sul valore accumulato dell'investimento diretto alla fine del periodo di osservazione. Sistemo il dataset rinominando le colonne, togliendo le prime righe e le ultime colonne che hanno informazioni non pertinenti, togliendo quelle righe in cui l'origine e la destinazione del flusso sono lo stesso paese.

```
In [83]: stock <- suppressMessages(read_excel(list_of_files[3], col_names = FALSE))
colnames(stock)<- stock[5,]
stock <- stock[-c(1:5),]
stock <- stock[which(stock$FDI == "Stock"),-3]
colnames(stock) <- c("reporting", "partner", "direction",c(1990:2021))
stock <- stock[-c(which(stock$reporting ==
stock$partner)),]
```

Divido il dataset nelle due direzionalità: inward e outward.

```
In [84]: new_inward_3<-as.data.frame(stock[which(stock$direction == "Inward"),-3])
new_inward_3[is.na(new_inward_3)] <- 0
new_inward_3[,3] <- as.numeric(new_inward_3[,3])
resh_3_inward <- melt(new_inward_3,id=c(1,2))
colnames(resh_3_inward) <- c("origin","destination","year","value")

ridotto_stock_inward <- resh_3_inward %>%
  group_by(origin,destination,year) %>%
  summarise(value = sum(value))

new_outward_3<-as.data.frame(stock[which(stock[,3] == "Outward"),-3])
new_outward_3[is.na(new_outward_3)] <- 0
new_outward_3[,3] <- as.numeric(new_outward_3[,3])
resh_3_outward <- melt(new_outward_3,id=c(1,2))
colnames(resh_3_outward) <- c("origin","destination","year","value")

ridotto_stock_outward <- resh_3_outward %>%
  group_by(origin,destination,year) %>%
  summarise(value = sum(value))
```

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

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

```
In [85]: rm(new_inward_3)
rm(new_outward_3)
rm(resh_3_inward)
rm(resh_3_outward)
```

Creo un dataset unico.

```
In [86]: p <- c(unique(stock$partner), setdiff(
              unique(stock$reporting),
              unique(stock$partner)))

y <- c(1990:2021)
stock_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(stock_tot) <- c("origin", "destination", "year")
stock_tot <- stock_tot[-c(which(stock_tot$origin ==
                                stock_tot$destination)),]
stock_tot <- cbind(stock_tot, rep(0, nrow(stock_tot)))
colnames(stock_tot) <- c("origin", "destination", "year", "value")

merge1 <- rbind(ridotto_stock_inward, stock_tot)
fdi_stock_inward <- merge1 %>%
  group_by(origin, destination, year) %>%
  summarise(value = sum(value))

merge2 <- rbind(ridotto_stock_outward, stock_tot)
fdi_stock_outward <- merge2 %>%
  group_by(origin, destination, year) %>%
  summarise(value = sum(value))
```

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

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

```
In [87]: rm(merge1)
rm(merge2)
rm(stock_tot)
rm(ridotto_stock_inward)
rm(ridotto_stock_outward)
```

```
In [88]: stock_unctad <- cbind(fdi_stock_inward[,1:3],
                              inward_stock = fdi_stock_inward$value,
                              outward_stock = fdi_stock_outward$value)
```

Cambio il nome dei paesi con il loro rispettivo codice ISO3c. Come nel dataset flow alcuni paesi non sono scritti in modo corretto, altri non esistono più come entità geografiche o non hanno un loro codice iso3c o comprendono aree geografiche troppo generiche. Riporto modifiche manualmente.

```
In [89]: paesi_iso3c <- countrycode(p, origin = "country.name", destination = "iso3c")
paesi <- as.data.frame(cbind(p, paesi_iso3c))

paesi$paesi_iso3c[261] <- "MAF"
paesi$paesi_iso3c[257] <- NA #soviet union non la conto

stock_unctad$origin <-
  paesi$paesi_iso3c[match(
    as.factor(stock_unctad$origin),
    as.factor(paesi$p))]
stock_unctad$destination <-
  paesi$paesi_iso3c[match(
    as.factor(stock_unctad$destination),
    as.factor(paesi$p))]
stock_unctad <- na.omit(stock_unctad)
```

Warning message:

"Some values were not matched unambiguously: APEC, ASEAN 10, Belgium / Luxembourg, Canary Islands, Channel Islands, Czechoslovakia (former), International Organization s, Kosovo, Netherlands Antilles, OPEC, Other Capital, Private buying and selling of property, Saint Martin, Serbia and Montenegro, SPEs, Unspecified Africa, Virgin Islands, West Indies, Yugoslavia (former)"

```
In [90]: stock_unctad <- stock_unctad %>%
  group_by(origin, destination, year) %>%
  summarise(inward_stock = sum(inward_stock), outward_stock = sum(outward_stock))
stock_unctad <- stock_unctad[-c(which(stock_unctad$origin ==
  stock_unctad$destination)),]
```

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

```
In [91]: stock_unctad <- stock_unctad[order(stock_unctad$year),]
colnames(stock_unctad) <- c("origin", "destination", "year",
  "inward_stock_unctad", "outward_stock_unctad")
write.csv(stock_unctad, "stock_unctad.csv", row.names = FALSE)
```

Valori negativi

```
In [92]: print(paste0("Nel dataset stock la percentuale dei valori negativi",
  "sul totale dei valori non nulli è del ",
  round(length(which(stock_unctad$inward_stock_unctad < 0))/length(
  which(stock_unctad$inward_stock_unctad != 0))*100,1), "% per inward e del ",
  round(length(which(stock_unctad$outward_stock_unctad < 0))/length(which(
  stock_unctad$outward_stock_unctad != 0))*100,1), "% per outward."))
```

[1] "Nel dataset stock la percentuale dei valori negativi sul totale dei valori non nulli è del 6.9% per inward e del 3.4% per outward."

Da creditore a debitore

Invertiamo l'origine con la destinazione nel caso inward in modo che l'origin sia il paese dove risiede DI e la destination sia il paese dove risiede DIENT.

```
In [93]: stock_inv <- stock_unctad %>%
  select(destination,origin,year,
         inward_derived_stock_unctad = inward_stock_unctad)
colnames(stock_inv)[1:2] <- colnames(stock_unctad)[1:2]
stock_inv <- stock_inv[order(stock_inv$origin),]
stock_confronto <- cbind(stock_inv[order(stock_inv$year),],
                        stock_unctad$outward_stock_unctad)
colnames(stock_confronto) <- c("origin", "destination", "year",
                              "outward_derived_stock_unctad", "outward_stock_unctad")
```

New names:

```
* `` -> `...5`
```

Confronto graficamente i due valori del flusso di investimento trattandoli entrambi dal punto di vista outward (investimento diretto-investimento inverso) ma in un caso il valore è dichiarato da DI, quindi dal creditore, mentre nell'altro caso è dichiarato da DIENT, quindi dal debitore.

```
In [94]: #stock_confronto$inward_stock <- (-1)*stock_confronto$inward_stock
sovr_stock <- which(stock_confronto$outward_stock_unctad != 0 &
                  stock_confronto$outward_derived_stock_unctad != 0)
print(paste0("Dal punto di vista inward ci sono ",length(which(
  stock_confronto$outward_stock_unctad != 0)) - length(sovr_unctad),
" valori non nulli per un dato scambio che son nulli, invece,",
"dal punto di vista outward"))

print(paste0("Dal punto di vista outward ci sono ",length(which(
  stock_confronto$outward_derived_stock_unctad != 0)) - length(sovr_unctad),
" valori non nulli per un dato scambio che son nulli, invece,",
"dal punto di vista inward"))
```

```
[1] "Dal punto di vista inward ci sono 77755 valori non nulli per un dato scambio che son nulli, invece, dal punto di vista outward"
```

```
[1] "Dal punto di vista outward ci sono 119289 valori non nulli per un dato scambio che son nulli, invece, dal punto di vista inward"
```

Confrontiamo graficamente solo i dati che riportano valori non nulli sia in inward che in outward.

```
In [95]: data_unctad <- data.frame(y = stock_confronto$outward_stock_unctad[sovr_unctad],
                                x = (-1)*stock_confronto$outward_derived_stock_unctad[sovr_unctad])
plot(data_unctad$x,
     data_unctad$y,
     xlab = "investimento diretto - investimento inverso (secondo DIENT)",
     ylab = "investimento diretto - investimento inverso (secondo DI)",
     main = "Confronto tra stock di investimento dichiarato da DI e da DIENT")
plot(data_unctad$x,
     data_unctad$y,
     xlab = "investimento diretto - investimento inverso (secondo DIENT)",
     ylab = "investimento diretto - investimento inverso (secondo DI)",
```

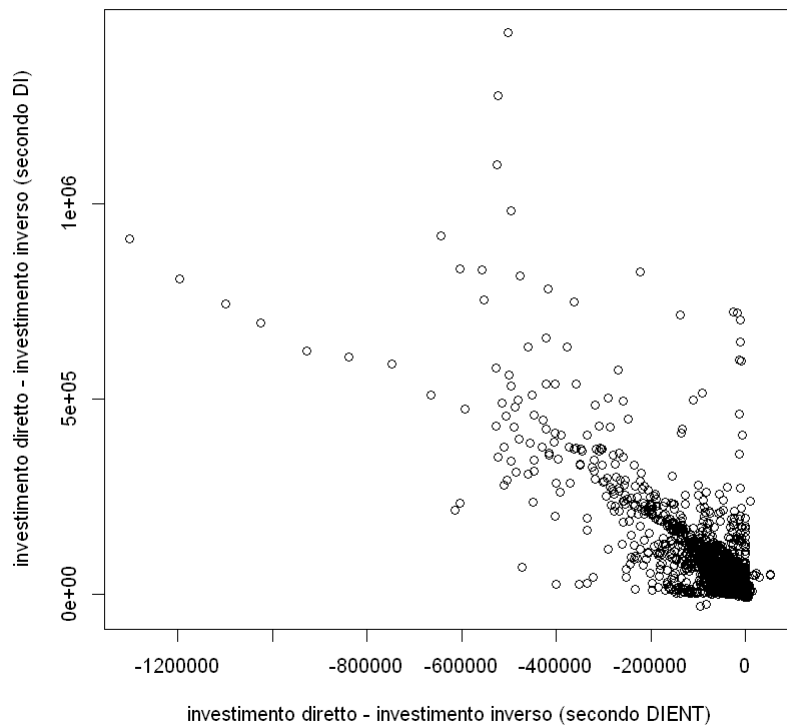


```

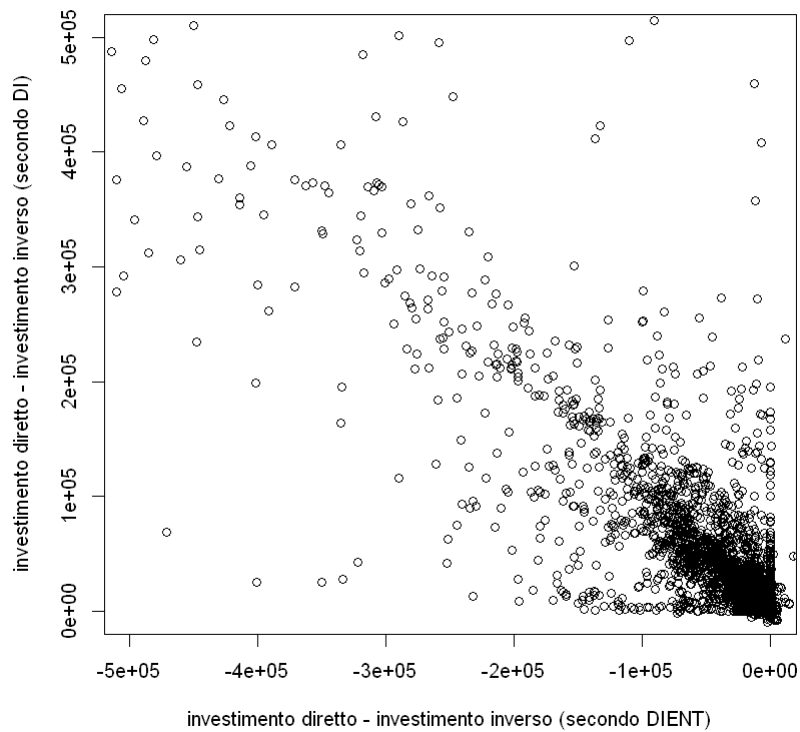
main = "Confronto tra stock di investimento dichiarato da DI e da DIENT",
xlim = c(-5e+05,0),ylim = c(0,5e+05))
plot(data_unctad$x,
      data_unctad$y,
      xlab = "investimento diretto - investimento inverso (secondo DIENT)",
      ylab = "investimento diretto - investimento inverso (secondo DI)",
      main = "Confronto tra stock di investimento dichiarato da DI e da DIENT",
      xlim = c(-1e+05,0),ylim = c(0,1e+05))

```

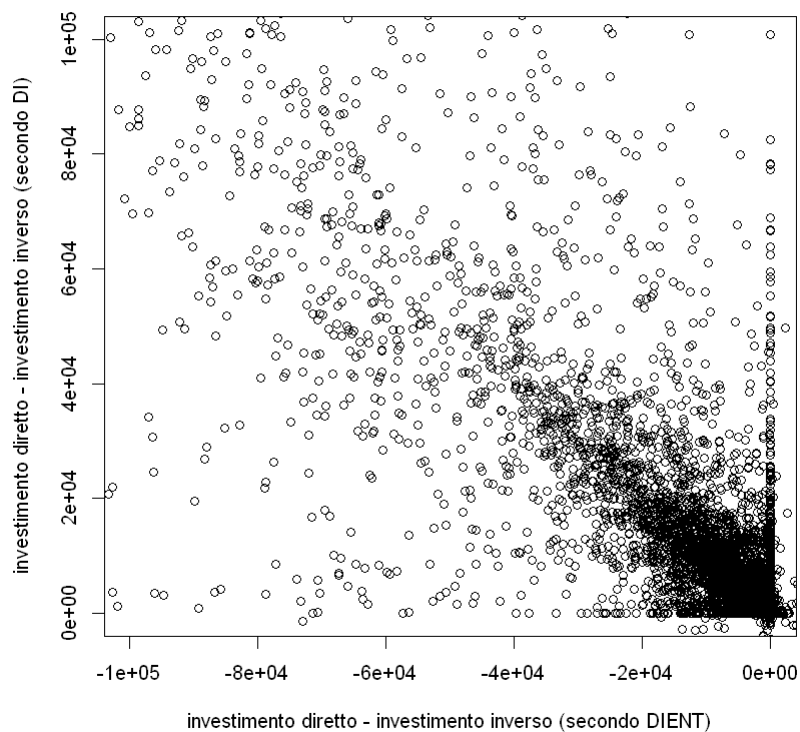
Confronto tra stock di investimento dichiarato da DI e da DIENT



Confronto tra stock di investimento dichiarato da DI e da DIENT



Confronto tra stock di investimento dichiarato da DI e da DIENT



Come per il dataset flow, si traggono le seguente osservazioni: se secondo il paese in cui risiede DIENT il suo investimento accumulato inverso è maggiore dell'investimento accumulato diretto che riceve dal paese dove risiede DI, secondo DI invece il suo investimento accumulato diretto in DIENT non eccede l'investimento accumulato inverso; in

modo analogo, se secondo DIENT, DI non ha investito più di quanto DIENT non abbia reinvestito, secondo DI, l'investimento accumulato diretto supera di molto l'investimento accumulato inverso.

Aggiungo una variabile dummy accanto alle due colonne di flow inward e flow outward dove la riga assume valore 1 se il creditore è anche il paese che dichiara, mentre -1 se il creditore non è il paese che dichiara.

```
In [96]: #stock_unctad_fin <- stock_confronto %>%  
# mutate(dummy_outward = rep(1,times=nrow(stock_confronto)), dummy_inward = rep(-1  
#stock_unctad_fin <- stock_unctad_fin %>%  
# select(origin,destination,year,inward_stock_unctad = inward_stock,dummy_inward_s  
write.csv(stock_confronto,"stock_unctad_fromcredto debt.csv", row.names = FALSE)
```