
Context socioeconòmic

PID_00258216

Xavier Fernández Marín

Índex

1. Indicadors de desenvolupament mundial (<i>World Development Indicators</i>).....	5
2. Desigualtat.....	11
Bibliografia.....	15

1. Indicadors de desenvolupament mundial (*World Development Indicators*)

La base de dades del Banc Mundial amb els indicadors de desenvolupament mundial (WDI per les sigles en anglès) (WDI) és la font de dades sobre característiques socioeconòmiques més rellevant, tant en termes temporals (bona part de les sèries temporals comencen a mitjan segle XX) com d'unitats cobertes (nombre de països) i de variables. En realitat, el Banc Mundial no és qui genera les dades de manera primària, sinó que són cadascun dels països els que s'encarreguen de proporcionar-les i el Banc Mundial les ordena, les sistematitza i les presenta de manera comprensiva.

El valor fonamental de la base de dades dels WDI, més enllà de l'abast que acabem de comentar, és que són indicadors oberts a l'accés i la consulta de qualsevol persona. D'aquesta manera, el primer lloc al qual has de recórrer a l'hora de fer un treball sobre qualsevol país, o per a comparar països i regions, o, en definitiva, per a encarar qualsevol recerca, és aquest. Tenir xifres al cap i poder-les contextualitzar en la geopolítica mundial és clau per a entendre què passa, com passa i per què passa quelcom en relacions internacionals.

Hi ha dues qüestions que cal destacar per a posar en perspectiva l'ús dels WDI com a font de dades. La primera és que les dades més noves sempre estan subjectes a canvis. És a dir, els instituts d'estadística estatals revisen la comptabilitat nacional regularment i no és estrany veure que hi ha canvis entre la primera xifra que es reporta un any i el que queda marcat com a dada històrica cinc anys després. Si bé no seran mai canvis exagerats, no és estrany trobar que la riquesa d'un país s'ha mogut fins a un cinc per cent amunt o avall amb les successives revisions. Per tant, cal posar sempre les dades en perspectiva quan es tracta de dades recents.

La segona qüestió és que la qualitat i la disponibilitat de les dades estan clarament relacionades amb la capacitat organitzativa de l'estat. Dit d'una altra manera: un estat que funciona bé, amb capacitat per a registrar transaccions, amb personal qualificat per a comptabilitzar el que passa dins de les seves fronteres, per a revisar-ho i per a publicar-ho, ha de tenir certes capacitats organitzatives. I no és estrany que siguin els estats més rics els que puguin fer això. Si un estat té una baixa capacitat de funcionament, és més probable que les dades que genera siguin o de més baixa qualitat, o, directament, que sigui incapaç de generar-les.

Si bé la disponibilitat de les dades en funció de la capacitat organitzativa i la fortalesa de l'administració estatal és en certa manera lògica, no és normalment tan lògic considerar que la manca de dades pot alterar la validesa científica de la recerca. Quan analitzem dades esperem no trobar-hi forats, no trobar

⁽¹⁾ Traducció del concepte *missing values* en anglès.

n'hi ha que falten, dades anomenades perdudes¹. Si tenim la dissort que no n'hi ha en alguns casos, el que esperem si més no, des del punt de vista estadístic i matemàtic, és que el patró de dades perdudes sigui aleatori; no tingui a veure amb el que s'està analitzant.

Exemple

Vols recollir dades sobre la riquesa d'un país. En aquest cas observaràs la magnitud de la seva economia tal i com la mesura el producte interior brut. Ara bé, resulta que dels deu països que analitzes n'hi ha dos que no tenen dades. Això, en si mateix, no hauria de ser cap problema si els dos països que no reporten dades no responen a cap patró sistemàtic. Però resulta que en general en relacions internacionals aquest patró és esbiaixat i sistemàtic. És més probable que un país amb un nivell baix de riquesa no reporti el seu PIB que no pas que ho faci un país amb un nivell alt. I és així perquè en bona part el PIB connota també una certa capacitat d'un estat. I això és un problema a l'hora de treure conclusions, perquè vol dir que les podries treure basant-te en un biaix original en les dades, pel fet que països de baixa riquesa són més propicis a no generar-ne. És més fàcil que els Estats Units donin un valor de PIB que no pas que ho faci el Sudan, i això està relacionat amb el fet que el PIB és més elevat en el primer cas que no pas en el segon. Per tant, obviar països que no tenen dades sense haver-se plantejat per què no en tenen no és una bona idea.

Cal anar amb compte, doncs, amb la disponibilitat de les dades socioeconòmiques, ja que tenir o no tenir una dada pot donar-nos molta informació sobre la dada mateixa, i en general aquesta informació sol perjudicar estats amb problemes de funcionament, que en general són aquells dels quals volem informació, precisament.

Les dades dels indicadors del Banc Mundial inclouen aspectes com agricultura, educació, ajut al desenvolupament, canvi climàtic, economia i creixement, mineria i recursos naturals, energia, medi ambient, deute extern, sector financer, gènere, salut, infraestructura, pobresa, sectors privat i públic, desenvolupament social, ciència i tecnologia, treball, protecció social, comerç i desenvolupament urbà.

És molt simple obtenir les dades dels WDI per a analitzar amb el paquet WDI del programa R.

library(WDI)

El primer pas és buscar la variable que t'interessi fent una cerca a la base de dades amb la funció `WDIsearch()` i posar com a primer argument de la funció la paraula clau que vols buscar (en anglès). En aquest cas, buscarem dades que tinguin a veure amb internet:

WDIsearch ("Internet")

```
##      indicator
## [1,] "IT.NET.BBND"
## [2,] "IT.NET.BBND.P2"
## [3,] "IT.NET.CONN.CD"
## [4,] "IT.NET.CONN.CN"
## [5,] "IT.NET.SECR"
## [6,] "IT.NET.SECR.P6"
## [7,] "IT.NET.SUB.CD"
## [8,] "IT.NET.SUB.CN"
## [9,] "IT.NET.USER"
## [10,] "IT.NET.USER.P2"
## [11,] "IT.NET.USER.P3"
##      name
## [1,] "Fixed broadband Internet subscribers"
## [2,] "Fixed broadband Internet subscribers (per 100 people)"
## [3,] "Fixed broadband Internet connection charge (current US$)"
## [4,] "Fixed broadband Internet connection charge (current LCU)"
## [5,] "Secure Internet servers"
## [6,] "Secure Internet servers (per 1 million people)"
## [7,] "Fixed broadband Internet monthly subscription (current US$)"
## [8,] "Fixed broadband Internet monthly subscription (current LCU)"
## [9,] "Internet users"
## [10,] "Internet users (per 100 people)"
## [11,] "Internet users (per 1,000 people)"
```

El resultat ens retorna dues columnes. La primera (*indicator*) és la paraula clau de l'indicador, i la segona (*name*) és el nom complet de l'indicador, amb un petit detall. Per a saber més detalls dels indicadors, ves a la pàgina web del Banc Mundial i hi trobaràs una descripció més detallada de la font originària de les dades.

Per a tenir el percentatge d'usuaris d'internet amb banda ampla només has d'especificar l'indicador a la funció `WDI()` (en aquest cas, `IT.NET.BBND.P2`) i desar-ho en un objecte que creïs amb el nom que vulguis (li direm en aquest cas `int`):

```
int <- WDI(indicator = "IT.NET.BBND.P2")
```

Ara, en memòria, hi tens un objecte anomenat `int` amb el qual pots operar. Per a veure'n l'estructura fes servir la funció `str()`:

```
str(int)

## 'data.frame':    1848 obs. of  4 variables:
## $ iso2c          : chr  "1A" "1A" "1A" "1A" ...
## $ country        : chr  "Arab World" "Arab World" "Arab World"
## "Arab World" ...
## $ IT.NET.BBND.P2: num  2.728 2.079 1.998 1.294 0.886 ...
## $ year           : num  2011 2010 2009 2008 2007 ...
```

La funció WDI

Els arguments de la funció `WDI` també permeten especificar el país (`country =`), l'any d'inici (`start =`) i l'any de finalització (`end =`).

Es tracta d'una matriu de dades (*data frame*) amb 1.848 observacions i 4 variables, la primera de les quals és el codi ISO del país en dos dígits; la segona, el nom del país (o de la regió, en cas que la variable també tingui variants regionals); la tercera, el valor de la variable que busques; i la quarta, l'any d'observació.

Per a tenir una idea dels valors que conté la matriu de dades no cal observar els 1.848 casos, sinó que amb els primers ja te'n pots fer una idea. En aquest cas, la funció `head()` mostra els primers valors.

```
head(int)
```

```
##   iso2c   country IT.NET.BBND.P2 year
## 1    1A Arab World      2.72761 2011
## 2    1A Arab World      2.07884 2010
## 3    1A Arab World      1.99806 2009
## 4    1A Arab World      1.29426 2008
## 5    1A Arab World      0.88585 2007
## 6    1A Arab World      0.49973 2006
```

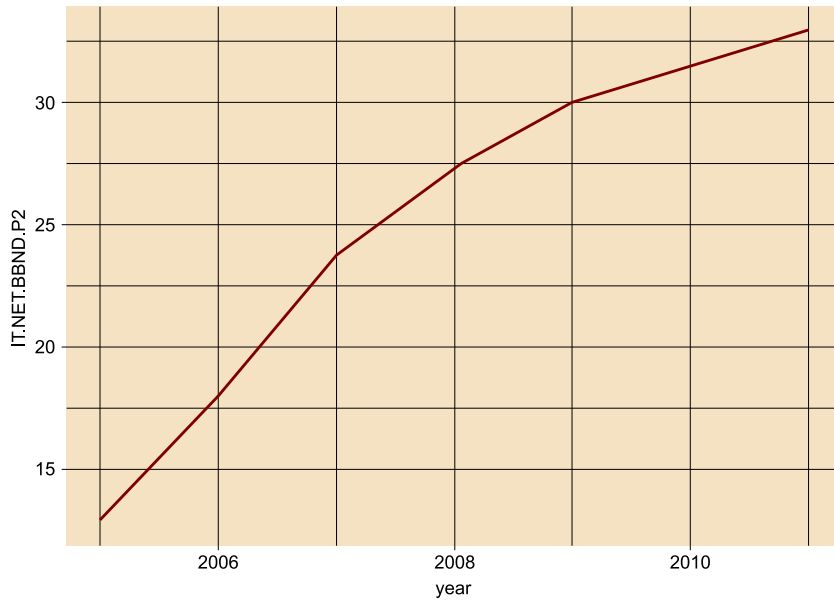
Per a fer una figura d'un estat concret has de filtrar el nom que vulguis a la variable `country` (o, si el saps, fer servir el seu codi ISO). En aquest cas, començant per l'objecte `int`, per mitjà d'un *pipeline* (representada per `%>%`) l'envies a la funció `filter()`, en la qual concretes que vols les observacions el `country` de les quals sigui estrictament igual a «Germany».

```
int.de <- int %>%
  filter(country == "Germany")
```

Ara, l'objecte `int.de`, el tens desat a la memòria i el pots fer servir per a dibuixar una figura amb l'evolució temporal, especificant els elements estètics (`aes`) `x` i `y` corresponents a l'eix horitzontal i vertical, respectivament. Finalment, l'element geomètric que vols que aparegui a la figura és una línia, i l'afegeixes a l'objecte definit per `ggplot()` amb la funció `geom_line()`:

```
ggplot(int.de, aes(x = year, y = IT.NET.BBND.P2)) +
  geom_line()
```

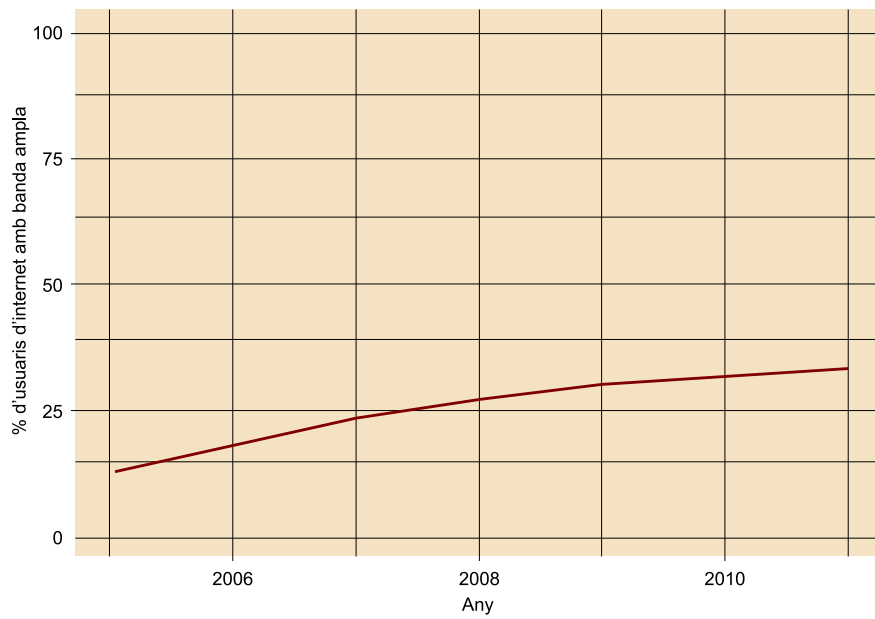

Figura 1. Resultat inicial d'una figura sense interès per aspectes formals



Finalment, cal que la figura feta tingui uns eixos adequats tant en el seu rang com en els noms, de manera que has d'afegir altres elements. En aquest cas, `expand_limit()` et permet dir-li que, com que l'eix vertical es correspon amb un percentatge, necessites que el valor mínim sigui el zero (0), i el valor màxim el cent (100). Només així pots tenir una perspectiva real de fins a quin punt canvia el valor anualment. Per a especificar un text millor per als eixos necessites les funcions `xlab()` i `ylab()`:

```
ggplot(int.de, aes(x = year, y = IT.NET.BBND.P2)) +  
  geom_line() +  
  expand_limits(y = c(0, 100)) +  
  xlab("Any") + ylab("% usuaris d'Internet\namb banda ampla")
```

Figura 2. Evolució temporal del percentatge d'usuaris d'internet amb banda ampla a Alemanya (2005-2011)



Font: *World Development Indicators*.

Hi ha també un altre paquet en \mathbb{R} que permet interactuar de manera similar a com ho fa el `WDI`, però amb les dades de països europeus provinents de l'Eurostat, l'oficina d'informació estadística europea. Es tracta del paquet `eu-rostat`.

2. Desigualtat

A banda de les mesures socioeconòmiques generals que podem trobar als indicadors de desenvolupament mundial, dins mateix del Banc Mundial un dels seus economistes, Branko Milanovic, manté una acurada base de dades sobre desigualtat, juntament amb una referència bibliogràfica imprescindible per a entendre el procés de desenvolupament econòmic i social els últims cent cinquanta anys i, de retruc, les relacions internacionals. Es tracta de Milanovic (2010).

El primer capítol explica què entenem per desigualtat, fa un repàs de les teories per les quals tradicionalment hem entès que tenia un efecte sobre el comportament econòmic, i –especialment rellevant per a fonts d’informació– presenta breument com es mesura la desigualtat.

El coeficient de Gini és la mesura estàndard de la *desigualtat* econòmica dins d’un grup social. El més habitual és considerar els estats com els grups socials de referència, però una de les virtuts del coeficient de Gini és que és aplicable a qualsevol grup social format per més de dues unitats (ja siguin persones o llars, vegeu el capítol). Ho calculem sumant la diferència d’ingressos (a vegades també es fa servir renda) entre totes les combinacions possibles d’unitats d’un grup social. En un grup social de tres unitats (A , B i C), l’operació és igual a $(A - B) + (A - C) + (B - C)$. Aquesta suma de diferències d’ingressos, la dividim pel nombre d’unitats (tres en l’exemple anterior) i la mitjana dels ingressos del grup. Aquesta operació aritmètica ens dona un valor que està en un rang teòric entre 0 i 1. Com que ja hem comentat que és una mesura de desigualtat, el zero correspon a la menor desigualtat, o quan tothom en el grup té exactament el mateix nivell d’ingressos (fixa’t que les diferències d’ingressos serien totes zero, i la seva suma també seria zero, de manera que al numerador només hi quedaria zero, que dividit per qualsevol cosa dona zero). A l’altre extrem hi ha l’1, que correspon a casos de grups amb desigualtat extrema, quan un únic individu té tots els ingressos i la resta no en té cap. Com s’indica a la lectura, el rang efectivament observable del coeficient de Gini, va de 0,25 a 0,60, o més habitualment expressat en una escola de zero a cent: de 25 a 60, que és el que coneixem com a índex de Gini.

El segon capítol se centra fonamentalment a explicar la desigualtat en termes comparats entre països i al llarg del temps. Milanovic explica quines estratègies es fan servir per a comparar ingressos en diferents moments del temps, introduint el concepte de paritat de poder adquisitiu (PPP, per les sigles en anglès). També explica per què és millor que ens refiem de les dades d’ingressos obtingudes a partir de les enquestes, que no pas de les dades obtingudes a partir de registres burocràtics (declaracions d’impostos, bàsicament), tot i que

Lectura obligatòria

Branko Milanovic (2010). *The Haves and the Have-Nots: A Brief and Idiosyncratic History of Global Inequality*. Nova York: Basic Books.

les primeres tenen un recorregut temporal més breu (des de mitjan segle XX per a economies avançades i des de finals del XX per a economies emergents, i encara molt problemàtiques per a economies no desenvolupades).

El nivell de renda dels països presentat en aquest segon capítol no deixa de ser el que hem introduït a la secció anterior (el PIB), tot i que en no en valor absolut (que és una mesura de la *mida d'una economia*), sinó en valor relatiu, que esdevé llavors una mesura de *riquesa d'una societat*.

Per tant, a l'hora de voler obtenir dades de riquesa o benestar material, i sobretot si vols fer comparacions en el temps, has de fer servir sèries temporals del PIB per càpita en PPP²:

⁽²⁾Recorda que necessites la llibreria `WDI`.

```
gdp.pc.ppp <- WDI(indicator = "NY.GDP.PCAP.PP.KD",
                 start = 1990, end = 2015)
str(gdp.pc.ppp)

## 'data.frame':    6864 obs. of  4 variables:
## $ iso2c          : chr  "1A" "1A" "1A" "1A" ...
## $ country        : chr  "Arab World" "Arab World" "Arab World"
## $ NY.GDP.PCAP.PP.KD: num  15323 15144 15040 14852 14462 ...
## $ year           : num  2015 2014 2013 2012 2011 ...
```

Per a calcular com ha anat divergint el PIB per càpita en els estats per als quals hi ha dades, fem servir la desviació estàndard, que és una mesura de diferència respecte a la mitjana en un grup d'observacions. El primer que fem és agafar l'objecte `gdp.pc.ppp` que conté les quatre variables que hem obtingut abans (codi ISO en dos caràcters, nom del país, valor de la variable i any) i enviar-lo a un *pipeline* (`%>%`) per a agrupar-lo per anys, de tal manera l'última línia, que fa un resum de la variable que ens interessa, ens farà el resum anualment. En aquest cas, el resum és l'aplicació de la variable `sd()`, que calcula la desviació estàndard (*standard deviation*) i que obvia els valors perduts (`na.rm = TRUE`).

Desviació estàndard

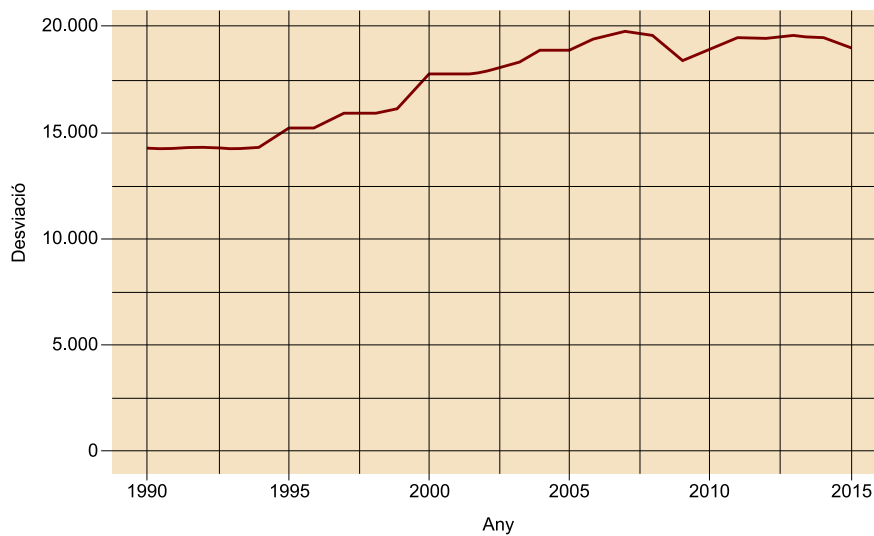
La desviació estàndard s'interpreta com el valor esperable que té la diferència de dues observacions agafades a l'atzar en un grup de valors. Com a mitjana, quan divergiràn dues observacions?

```
gdp.pc.de <- gdp.pc.ppp %>%
  group_by(year) %>%
  summarize(Desviació = sd(NY.GDP.PCAP.PP.KD, na.rm = TRUE))
```

Per a generar la figura, hem de passar l'objecte acabat de crear a la funció `ggplot()` com anteriorment, i especificar tant els valors que aniran a cada eix com el tipus de figura que volem (línies, per a marcar una evolució temporal), i fixar un valor de referència per a l'eix vertical.

```
ggplot(gdp.pc.de, aes(x = year, y = Desviació)) +
  geom_line() +
  expand_limits(y = 0) +
  xlab("Any")
```

Figura 3. Evolució temporal de la desigualtat de renda entre països, mesurada per la desviació estàndard anual de la renda per càpita als països del món



Fixa't que la crisi financera del 2008 en realitat va significar una equiparació entre els països, un fre a la creixent desigualtat entre estats. Font: *World Development Indicators*.

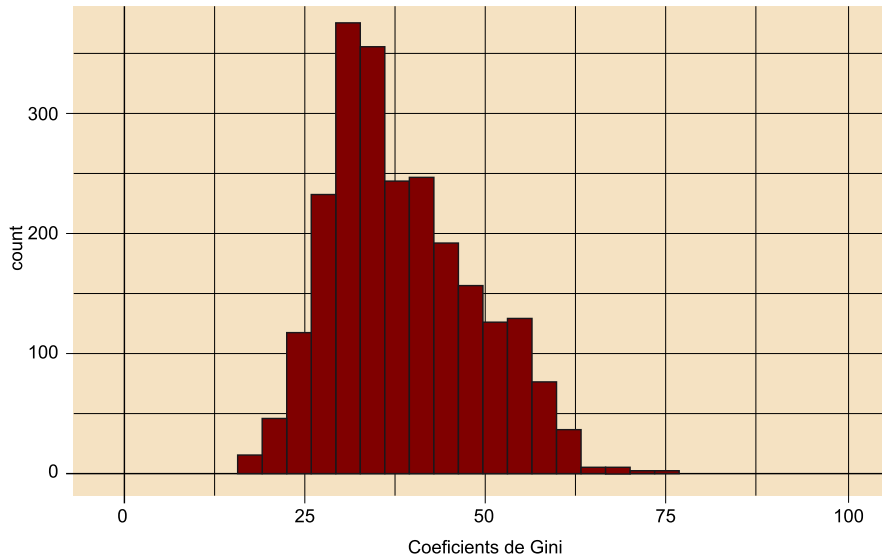
Per a obtenir les dades dels índexs de Gini, preparades per Branko Milanovic, hem de fer servir les seves matrius, que estan en format Stata (extensió `.dta`), però que amb el paquet `foreign` es poden llegir fàcilment en R:

```
library(foreign)
gini <-
read.dta("https://wfs.gc.cuny.edu/njohnson/www/BrankoData/allginis_2016.dta")
```

La distribució dels índexs de Gini (variable `Giniall`), l'obtenim amb un histograma, on marquem els rangs teòrics.

```
ggplot(gini, aes(x = Giniall)) +
  geom_histogram() +
  xlab("Coeficients de Gini") +
  expand_limits(x = c(0, 100))
```

Figura 4. Distribució dels índexs de Gini estandarditzats (1980-2015)



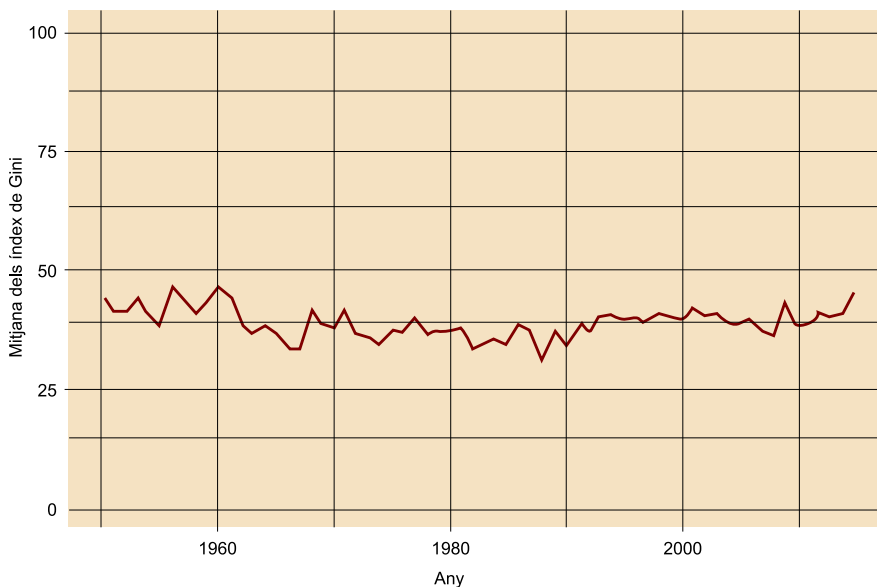
L'evolució de la mitjana dels índexs de Gini es pot obtenir agrupant les observacions anuals i fent-ne la mitjana³.

```
gini.any <- gini %>%
  group_by(year) %>%
  summarize(Mitjana = mean(GiniIall, na.rm = TRUE))
```

```
ggplot(gini.any, aes(x = year, y = Mitjana)) +
  geom_line() +
  xlab("Any") + ylab("Mitjana dels índexs de Gini") +
  expand_limits(y = c(0, 100))
```

⁽³⁾Fixa't, però, que aquesta és la mitjana dels índexs de Gini de cada país, de manera que els països tenen el mateix pes, sense ponderar pel seu nivell de població. Per tant, no es pot llegir com una evolució de la desigualtat al món, sinó només com una evolució de la desigualtat entre estats.

Figura 5. Evolució temporal de la mitjana d'índexs de Gini dels diferents països (1980-2015)



Mesura la desigualtat entre Estats, però no la desigualtat global.

Bibliografia

Milanovic, Branko (2010). *The Haves and the Have-Nots: A Brief and Idiosyncratic History of Global Inequality* (cap. 2 i 3). Nova York: Basic Books.

