

Wprowadzenie do reshape2

Statystyka I z R

Bartosz Maćkiewicz

Wprowadzenie do reshape2

Główną funkcją pakietu **reshape2** jest możliwość przekształcenia danych z formatu “długiego” do “szerokiego” i odwrotnie.

Format “szeroki” - każdej zmiennej odpowiada kolumna. Jest to format, z którym najczęściej się do tej pory spotykaliśmy (choć mieliśmy również styczność z “długim” formatem).

Format “długi” - zawiera kolumny z możliwymi zmiennymi oraz ich wartościami.

Po co dokonywać konwersji pomiędzy formatami? Wszystko zależy od tego, co z nimi chcemy zrobić - czasami potrzebne nam są dane w jednym formacie, czasami w drugim. W rzeczywistości “długi” format potrzebny jest częściej, niż może się Państwu wydawać.

Pakiet **reshape2** instalujemy standardowym poleceniem **install.packages** i ładujemy za pomocą **library**.

```
# install.packages('reshape2')
library(reshape2)
```

Do ćwiczeń załadujemy dane dotyczące śmiertelności dzieci poniżej 5 roku życia.

```
df <- read.csv('unicef-u5mr.csv')
```

Nasze dane, jak możemy zobaczyć, są w formacie “długim” - każdemu rokowi odpowiada jedna kolumna.

```
head(df)
```

```
##      CountryName U5MR.1950 U5MR.1951 U5MR.1952 U5MR.1953 U5MR.1954 U5MR.1955
## 1      Afghanistan      NA      NA      NA      NA      NA      NA
## 2      Albania      NA      NA      NA      NA      NA      NA
## 3      Algeria      NA      NA      NA      NA      251      249.9
## 4      Andorra      NA      NA      NA      NA      NA      NA
## 5      Angola      NA      NA      NA      NA      NA      NA
## 6 Antigua & Barbuda      NA      NA      NA      NA      NA      NA
##      U5MR.1956 U5MR.1957 U5MR.1958 U5MR.1959 U5MR.1960 U5MR.1961 U5MR.1962
## 1      NA      NA      NA      NA      NA      356.5      350.6
## 2      NA      NA      NA      NA      NA      NA      NA
## 3      249      248      247.5      246.7      246.3      246.1      246.2
## 4      NA      NA      NA      NA      NA      NA      NA
## 5      NA      NA      NA      NA      NA      NA      NA
## 6      NA      NA      NA      NA      NA      NA      NA
##      U5MR.1963 U5MR.1964 U5MR.1965 U5MR.1966 U5MR.1967 U5MR.1968 U5MR.1969
## 1      345.0      339.7      334.1      328.7      323.3      318.1      313.0
## 2      NA      NA      NA      NA      NA      NA      NA
## 3      246.8      247.4      248.2      248.7      248.4      247.4      245.3
## 4      NA      NA      NA      NA      NA      NA      NA
```

## 5	NA	NA	NA	NA	NA	NA	NA
## 6	NA	NA	NA	NA	NA	NA	NA
##	U5MR.1970	U5MR.1971	U5MR.1972	U5MR.1973	U5MR.1974	U5MR.1975	U5MR.1976
## 1	307.8	302.1	296.4	290.8	284.9	279.4	273.6
## 2	NA	NA	NA	NA	NA	NA	NA
## 3	241.7	236.5	230.0	222.5	214.2	205.0	195.2
## 4	NA	NA	NA	NA	NA	NA	NA
## 5	NA	NA	NA	NA	NA	NA	NA
## 6	NA	NA	NA	NA	NA	NA	NA
##	U5MR.1977	U5MR.1978	U5MR.1979	U5MR.1980	U5MR.1981	U5MR.1982	U5MR.1983
## 1	267.8	261.6	255.5	249.1	242.7	236.2	229.7
## 2	NA	91.1	84.7	78.6	73.0	67.8	62.8
## 3	184.9	173.8	161.8	148.1	132.5	115.8	99.2
## 4	NA	NA	NA	NA	NA	NA	NA
## 5	NA	NA	NA	234.1	232.8	231.5	230.2
## 6	NA	NA	NA	NA	NA	NA	NA
##	U5MR.1984	U5MR.1985	U5MR.1986	U5MR.1987	U5MR.1988	U5MR.1989	U5MR.1990
## 1	222.9	216.0	209.2	202.1	195.0	187.8	181.0
## 2	58.3	54.3	50.7	47.6	44.9	42.5	40.6
## 3	83.8	71.2	61.9	55.4	51.2	48.5	46.8
## 4	NA	NA	NA	NA	NA	NA	8.5
## 5	229.1	228.3	227.5	226.9	226.5	226.2	226.0
## 6	NA	NA	NA	NA	NA	NA	25.5
##	U5MR.1991	U5MR.1992	U5MR.1993	U5MR.1994	U5MR.1995	U5MR.1996	U5MR.1997
## 1	174.2	167.8	162.0	156.8	152.3	148.6	145.5
## 2	38.8	37.3	36.0	34.6	33.2	31.8	30.3
## 3	45.7	44.9	44.1	43.3	42.5	41.8	41.1
## 4	7.9	7.4	6.9	6.4	6.0	5.7	5.3
## 5	225.9	226.0	225.8	225.5	224.8	224.0	222.6
## 6	24.2	23.1	21.9	20.8	19.7	18.8	17.9
##	U5MR.1998	U5MR.1999	U5MR.2000	U5MR.2001	U5MR.2002	U5MR.2003	U5MR.2004
## 1	142.6	139.9	137.0	133.8	130.3	126.8	123.2
## 2	28.9	27.5	26.2	24.9	23.6	22.5	21.5
## 3	40.6	40.2	39.7	38.9	37.8	36.5	35.1
## 4	5.0	4.8	4.6	4.4	4.2	4.1	4.0
## 5	220.8	218.9	216.7	214.1	211.7	209.2	206.7
## 6	17.0	16.2	15.5	14.8	14.1	13.5	12.9
##	U5MR.2005	U5MR.2006	U5MR.2007	U5MR.2008	U5MR.2009	U5MR.2010	U5MR.2011
## 1	119.6	116.3	113.2	110.4	107.6	105.0	102.3
## 2	20.5	19.5	18.7	17.9	17.3	16.6	16.0
## 3	33.6	32.1	30.7	29.4	28.3	27.3	26.6
## 4	3.9	3.7	3.6	3.5	3.4	3.3	3.2
## 5	203.9	200.5	196.4	192.0	187.3	182.5	177.3
## 6	12.4	11.8	11.3	10.9	10.4	9.9	9.5
##	U5MR.2012	U5MR.2013	U5MR.2014	U5MR.2015			
## 1	99.5	96.7	93.9	91.1			
## 2	15.5	14.9	14.4	14.0			
## 3	26.1	25.8	25.6	25.5			
## 4	3.1	3.0	2.9	2.8			
## 5	172.2	167.1	162.2	156.9			
## 6	9.1	8.7	8.4	8.1			

Spróbujmy trochę uporządkować nazwy kolumn, tak by odpowiadały im liczby. Po wczytaniu danych nazwy kolumn wyglądają tak:

```
colnames(df)[2:ncol(df)]
```

```
## [1] "U5MR.1950" "U5MR.1951" "U5MR.1952" "U5MR.1953" "U5MR.1954" "U5MR.1955"
## [7] "U5MR.1956" "U5MR.1957" "U5MR.1958" "U5MR.1959" "U5MR.1960" "U5MR.1961"
## [13] "U5MR.1962" "U5MR.1963" "U5MR.1964" "U5MR.1965" "U5MR.1966" "U5MR.1967"
## [19] "U5MR.1968" "U5MR.1969" "U5MR.1970" "U5MR.1971" "U5MR.1972" "U5MR.1973"
## [25] "U5MR.1974" "U5MR.1975" "U5MR.1976" "U5MR.1977" "U5MR.1978" "U5MR.1979"
## [31] "U5MR.1980" "U5MR.1981" "U5MR.1982" "U5MR.1983" "U5MR.1984" "U5MR.1985"
## [37] "U5MR.1986" "U5MR.1987" "U5MR.1988" "U5MR.1989" "U5MR.1990" "U5MR.1991"
## [43] "U5MR.1992" "U5MR.1993" "U5MR.1994" "U5MR.1995" "U5MR.1996" "U5MR.1997"
## [49] "U5MR.1998" "U5MR.1999" "U5MR.2000" "U5MR.2001" "U5MR.2002" "U5MR.2003"
## [55] "U5MR.2004" "U5MR.2005" "U5MR.2006" "U5MR.2007" "U5MR.2008" "U5MR.2009"
## [61] "U5MR.2010" "U5MR.2011" "U5MR.2012" "U5MR.2013" "U5MR.2014" "U5MR.2015"
```

Za pomocą polecenia `strsplit` podzielmy nazwy kolumn na dwa ciągi znaków: `U5MR` i `rok`. Funkcja `strsplit` przyjmuje jako pierwszy argument wektor typu `character` a jako drugi wyrażenie regularne mówiące, jak chcemy podzielić nasze napisy. W tym przypadku użyłem wzorca `\\.`, który dopasowuje znak podziału do kropki.

```
str = c('123y21321s3432', '12z3213y213213', '21321512y1321z321')
```

```
strsplit(str, '[a-z]')
```

```
## [[1]]
## [1] "123"      "21321"    "3432"
##
## [[2]]
## [1] "12"       "3213"     "213213"
##
## [[3]]
## [1] "21321512" "1321"     "321"
```

```
head(strsplit(colnames(df)[2:ncol(df)], '\\.'))
```

```
## [[1]]
## [1] "U5MR" "1950"
##
## [[2]]
## [1] "U5MR" "1951"
##
## [[3]]
## [1] "U5MR" "1952"
##
## [[4]]
## [1] "U5MR" "1953"
##
## [[5]]
## [1] "U5MR" "1954"
##
## [[6]]
## [1] "U5MR" "1955"
```

```
head(strsplit(colnames(df)[2:ncol(df)], 'R'))
```

```
## [[1]]
## [1] "U5M"      ".1950"
```

```
##
## [[2]]
## [1] "U5M" ".1951"
##
## [[3]]
## [1] "U5M" ".1952"
##
## [[4]]
## [1] "U5M" ".1953"
##
## [[5]]
## [1] "U5M" ".1954"
##
## [[6]]
## [1] "U5M" ".1955"
```

```
head(strsplit(colnames(df)[2:ncol(df)], '\\.'))
```

```
## [[1]]
## [1] "" "" "" "" "" "" "" "" ""
##
## [[2]]
## [1] "" "" "" "" "" "" "" "" ""
##
## [[3]]
## [1] "" "" "" "" "" "" "" "" ""
##
## [[4]]
## [1] "" "" "" "" "" "" "" "" ""
##
## [[5]]
## [1] "" "" "" "" "" "" "" "" ""
##
## [[6]]
## [1] "" "" "" "" "" "" "" "" ""
```

Funkcja `strsplit` zwraca listę, z której za pomocą polecenia `unlist` możemy stworzyć wektor. Problemem jest to, że otrzymamy wówczas wektor typu `character` o postaci:

```
unlist(strsplit(colnames(df)[2:ncol(df)], '\\.'))
```

```
## [1] "U5MR" "1950" "U5MR" "1951" "U5MR" "1952" "U5MR" "1953" "U5MR" "1954"
## [11] "U5MR" "1955" "U5MR" "1956" "U5MR" "1957" "U5MR" "1958" "U5MR" "1959"
## [21] "U5MR" "1960" "U5MR" "1961" "U5MR" "1962" "U5MR" "1963" "U5MR" "1964"
## [31] "U5MR" "1965" "U5MR" "1966" "U5MR" "1967" "U5MR" "1968" "U5MR" "1969"
## [41] "U5MR" "1970" "U5MR" "1971" "U5MR" "1972" "U5MR" "1973" "U5MR" "1974"
## [51] "U5MR" "1975" "U5MR" "1976" "U5MR" "1977" "U5MR" "1978" "U5MR" "1979"
## [61] "U5MR" "1980" "U5MR" "1981" "U5MR" "1982" "U5MR" "1983" "U5MR" "1984"
## [71] "U5MR" "1985" "U5MR" "1986" "U5MR" "1987" "U5MR" "1988" "U5MR" "1989"
## [81] "U5MR" "1990" "U5MR" "1991" "U5MR" "1992" "U5MR" "1993" "U5MR" "1994"
## [91] "U5MR" "1995" "U5MR" "1996" "U5MR" "1997" "U5MR" "1998" "U5MR" "1999"
## [101] "U5MR" "2000" "U5MR" "2001" "U5MR" "2002" "U5MR" "2003" "U5MR" "2004"
## [111] "U5MR" "2005" "U5MR" "2006" "U5MR" "2007" "U5MR" "2008" "U5MR" "2009"
## [121] "U5MR" "2010" "U5MR" "2011" "U5MR" "2012" "U5MR" "2013" "U5MR" "2014"
## [131] "U5MR" "2015"
```

Musimy więc wybrać co drugą wartość z tego wektora.

```
lata = unlist(strsplit(colnames(df)[2:ncol(df)], '\\. '))[seq(2*(ncol(df)-1),from = 2, by = 2)]
lata
```

```
## [1] "1950" "1951" "1952" "1953" "1954" "1955" "1956" "1957" "1958" "1959"
## [11] "1960" "1961" "1962" "1963" "1964" "1965" "1966" "1967" "1968" "1969"
## [21] "1970" "1971" "1972" "1973" "1974" "1975" "1976" "1977" "1978" "1979"
## [31] "1980" "1981" "1982" "1983" "1984" "1985" "1986" "1987" "1988" "1989"
## [41] "1990" "1991" "1992" "1993" "1994" "1995" "1996" "1997" "1998" "1999"
## [51] "2000" "2001" "2002" "2003" "2004" "2005" "2006" "2007" "2008" "2009"
## [61] "2010" "2011" "2012" "2013" "2014" "2015"
```

Następnie przypisujemy ten wektor do nazw kolumn.

```
colnames(df)[2:ncol(df)] = unlist(strsplit(colnames(df)[2:ncol(df)], '\\. '))[seq(2*(ncol(df)-1),from = 2, by = 2)]
```

Widzimy, że uporządkowaliśmy sobie kwestię nazw kolumn. Teraz są nimi kolejne lata.

```
head(df)
```

```
##      CountryName 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960
## 1      Afghanistan    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 2           Albania    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 3           Algeria    NA    NA    NA    NA   251 249.9 249 248 247.5 246.7 246.3
## 4           Andorra    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 5           Angola    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 6 Antigua & Barbuda    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
##      1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973
## 1 356.5 350.6 345.0 339.7 334.1 328.7 323.3 318.1 313.0 307.8 302.1 296.4 290.8
## 2    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 3 246.1 246.2 246.8 247.4 248.2 248.7 248.4 247.4 245.3 241.7 236.5 230.0 222.5
## 4    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 5    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 6    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
##      1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986
## 1 284.9 279.4 273.6 267.8 261.6 255.5 249.1 242.7 236.2 229.7 222.9 216.0 209.2
## 2    NA    NA    NA    NA   91.1  84.7  78.6  73.0  67.8  62.8  58.3  54.3  50.7
## 3 214.2 205.0 195.2 184.9 173.8 161.8 148.1 132.5 115.8  99.2  83.8  71.2  61.9
## 4    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 5    NA    NA    NA    NA    NA    NA 234.1 232.8 231.5 230.2 229.1 228.3 227.5
## 6    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
##      1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999
## 1 202.1 195.0 187.8 181.0 174.2 167.8 162.0 156.8 152.3 148.6 145.5 142.6 139.9
## 2  47.6  44.9  42.5  40.6  38.8  37.3  36.0  34.6  33.2  31.8  30.3  28.9  27.5
## 3  55.4  51.2  48.5  46.8  45.7  44.9  44.1  43.3  42.5  41.8  41.1  40.6  40.2
## 4    NA    NA    NA   8.5   7.9   7.4   6.9   6.4   6.0   5.7   5.3   5.0   4.8
## 5 226.9 226.5 226.2 226.0 225.9 226.0 225.8 225.5 224.8 224.0 222.6 220.8 218.9
## 6    NA    NA    NA  25.5  24.2  23.1  21.9  20.8  19.7  18.8  17.9  17.0  16.2
##      2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
## 1 137.0 133.8 130.3 126.8 123.2 119.6 116.3 113.2 110.4 107.6 105.0 102.3  99.5
## 2  26.2  24.9  23.6  22.5  21.5  20.5  19.5  18.7  17.9  17.3  16.6  16.0  15.5
## 3  39.7  38.9  37.8  36.5  35.1  33.6  32.1  30.7  29.4  28.3  27.3  26.6  26.1
## 4   4.6   4.4   4.2   4.1   4.0   3.9   3.7   3.6   3.5   3.4   3.3   3.2   3.1
## 5 216.7 214.1 211.7 209.2 206.7 203.9 200.5 196.4 192.0 187.3 182.5 177.3 172.2
## 6  15.5  14.8  14.1  13.5  12.9  12.4  11.8  11.3  10.9  10.4   9.9   9.5   9.1
##      2013 2014 2015
```

```
## 1  96.7  93.9  91.1
## 2  14.9  14.4  14.0
## 3  25.8  25.6  25.5
## 4   3.0   2.9   2.8
## 5 167.1 162.2 156.9
## 6   8.7   8.4   8.1
```

melt

Funkcja `melt` pozwala nam na przekształcenie naszej ramki danych z formatu “szerokiego” na format “długi”. Jako pierwszy argument przyjmuje ramkę danych, kolejne argumenty określają jak ma wyglądać wynikowa ramka danych.

```
tail(melt(df))
```

```
## Using CountryName as id variables
##      CountryName variable value
## 12931    Uruguay      2015  10.1
## 12932 Uzbekistan      2015  39.1
## 12933  Venezuela      2015  14.9
## 12934     Samoa      2015  17.5
## 12935     Yemen      2015  41.9
## 12936     Zambia      2015  64.0
```

Możemy od razu określić jak nazywać się będzie kolumna z nazwą zmiennej. W naszym przypadku zupełnie sensownie jest nadać jej zamiast nazwy `variable` nazwę `year`, bo faktycznie znajdują się w niej daty.

```
df_dlugie = melt(df, variable.name = 'year', value.name = 'U5MR')
```

```
## Using CountryName as id variables
tail(df_dlugie)
```

```
##      CountryName year U5MR
## 12931    Uruguay 2015  10.1
## 12932 Uzbekistan 2015  39.1
## 12933  Venezuela 2015  14.9
## 12934     Samoa 2015  17.5
## 12935     Yemen 2015  41.9
## 12936     Zambia 2015  64.0
```

dcast

Za pomocą funkcji `dcast` możemy dokonać przekształcenia w drugą stronę - z formatu “długiego” do formatu “szerokiego”. `dcast` korzysta ze standardowego interfejsu formuł w R. W naszym przypadku przekazaliśmy następującą formułę:

```
CountryName ~ year
```

`CountryName` ma być naszą kolumną wyznaczającą wiersze, `year` jest zaś kolumną, w której znajdują się nazwy zmiennych.

Argument `value` mówi, w której kolumnie znajdują się wartości wierszy.

Funkcja `dcast` umożliwia również bardziej zaawansowane przekształcanie ramek danych w “długim” formacie. Nam jednak na początek starczy najbardziej podstawowe jej użycie.

```
df_krotkie = dcast(CountryName ~ year, data = df_dlugie, value = 'U5MR')
```

```
## Using U5MR as value column: use value.var to override.
```

```
tail(df_krotkie)
```

```
##      CountryName 1950  1951  1952  1953  1954  1955  1956  1957  1958  1959
## 191    Vanuatu    NA    NA 182.6 179.4 175.9 172.4 169.0 165.8 162.7 158.9
## 192  Venezuela    NA 121.3 115.5 110.1 105.0 100.2  95.8  91.6  87.7  84.2
## 193    Vietnam    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 194      Yemen    NA    NA    NA    NA    NA    NA    NA    NA    NA    NA
## 195      Zambia    NA    NA    NA 238.0 233.3 229.3 224.6 220.5 216.2 211.7
## 196  Zimbabwe    NA    NA    NA 181.8 177.2 172.6 168.0 163.8 159.5 155.1
##      1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971
## 191 155.2 151.0 147.0 142.4 137.7 132.9 128.1 123.0 118.1 113.2 108.6 104.2
## 192  81.1  78.3  75.9  73.8  72.0  70.4  68.9  67.5  66.0  64.3  62.6  60.6
## 193    NA    NA    NA    NA  95.0  93.3  91.8  90.3  88.8  87.3  85.7  84.2
## 194    NA    NA    NA 406.1 397.7 388.2 377.9 366.5 354.6 341.8 329.3 317.0
## 195 207.3 203.0 199.0 195.4 192.3 190.3 189.0 188.2 187.1 185.0 181.6 176.7
## 196 150.7 146.1 141.6 137.2 132.8 128.6 124.7 121.1 118.1 115.8 114.0 112.6
##      1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983
## 191  99.9  95.9  92.1  88.5  84.8  81.1  77.1  73.0  68.7  64.3  59.9  55.6
## 192  58.8  56.8  55.0  53.2  51.4  49.6  47.7  45.8  43.9  42.1  40.3  38.8
## 193  82.6  81.0  79.3  77.6  75.9  74.1  72.3  70.5  68.7  66.9  65.3  63.6
## 194 305.2 294.2 283.0 271.6 259.5 246.9 234.0 221.2 208.4 196.0 184.1 173.0
## 195 171.1 165.4 160.7 157.2 155.4 155.0 155.4 156.0 156.8 158.5 160.9 164.5
## 196 111.7 111.1 110.7 110.4 110.0 109.4 108.4 106.4 103.4  99.4  94.7  89.8
##      1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995
## 191  51.5  47.8  44.6  41.7  39.3  37.3  35.6  34.1  32.9  31.9  31.0  30.4
## 192  37.4  36.1  34.8  33.5  32.1  30.8  29.6  28.6  27.8  27.3  26.8  26.3
## 193  62.0  60.3  58.5  56.7  54.8  52.8  50.8  48.8  46.8  44.8  43.0  41.2
## 194 163.1 154.4 147.0 140.7 135.2 130.4 126.3 122.8 119.6 116.8 114.4 112.2
## 195 169.2 174.7 180.1 184.7 188.0 190.0 190.6 190.3 189.4 187.5 184.5 181.1
## 196  84.9  80.6  77.1  74.6  73.5  73.9  75.8  78.8  82.7  87.0  91.2  95.5
##      1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008
## 191  29.9  34.1  29.1  28.8  28.5  28.2  27.9  27.9  27.9  27.9  28.0  28.2  28.2
## 192  25.5  24.5  23.5  27.7  21.7  21.0  20.4  19.8  19.2  18.5  17.9  17.4  17.1
## 193  39.5  38.0  36.5  35.1  33.8  32.5  31.4  30.4  29.4  28.5  27.7  26.9  26.2
## 194 109.8 107.2 103.8  99.9  95.4  90.7  85.9  81.3  76.8  72.6  68.5  64.7  61.0
## 195 178.1 175.7 172.9 169.1 163.1 154.2 142.9 131.2 120.6 111.7 104.5  98.0  93.1
## 196  99.2 102.0 103.9 105.2 105.8 105.6 105.1 104.0 103.0 101.9 100.0  98.0  95.4
##      2009 2010 2011 2012 2013 2014 2015
## 191  28.4  28.5  28.6  28.6  28.4  28.1  27.5
## 192  16.8  16.6  16.4  16.1  15.7  15.3  14.9
## 193  25.5  24.8  24.2  23.5  22.9  22.3  21.7
## 194  57.6  54.2  51.1  48.4  46.0  43.8  41.9
## 195  87.4  82.1  78.6  74.4  70.2  66.6  64.0
## 196  92.9  89.5  85.6  78.5  74.5  72.3  70.7
```

Mając dane w formacie “długim” możemy z łatwością stworzyć wykres obrazujący, jak śmiertelność dzieci poniżej 5 roku życia malała w Polsce na przestrzeni lat.

```
df_polska = df_dlugie[df_dlugie$CountryName == 'Poland',]
plot(as.numeric(as.character(df_polska$year)), df_polska$U5MR, type = 'l', col = 'red')
```

