

Proyecto Final: Principales Causas de Mortalidad en la Población Mexicana 2012-2017

José Luis Zarate Cortés (000183347)

15/7/2019

IMPORTAR DATOS

Dado que estamos utilizando la impagen RStudio palmoreck, es necesario introducir los archivos *.csv en esta ruta, modificar el directorio y descargar los datos.

```
getwd()

## [1] "/home/rstudio"

#Datos del INEGI, 2012 a 2017. Se analiza para los periodos 2015-2017
data2017 <- read.csv("conjunto_de_datos_defunciones_generales_2017.csv", sep = ",", header = TRUE)
datacap <- read.csv("decapitulo.csv", sep = ",", header = TRUE)
dataocu <- read.csv("deocupa.csv", sep = ",", header = TRUE)

print("El número de fila de las tablas importadas es (en orden 2012-2017):")

## [1] "El número de fila de las tablas importadas es (en orden 2012-2017):"

numfil<-c(nrow(data2017))#,nrow(data2016),nrow(data2017))
numcol<-c(ncol(data2017))#,ncol(data2016),ncol(data2017))
print(numfil)
```

```
## [1] 703047
```

```
print(numcol)
```

```
## [1] 59
```

Consolidamos los datos en una sola tabla y validamos la dimension

Ahora eliminaremos datos inecesarios, por ejemplo analizaremos si los datos de año de fallecimiento están en el rango 2015-2017, en caso de existir datos anteriores a 2014 se eliminarán.

```
print("año menor de ocurrencia:")
```

```
## [1] "año menor de ocurrencia:"
```

```
min(data2017$anio_ocur)
```

```
## [1] 1943
```

```
print("número de registro anteriores a 2015:")
```

```
## [1] "número de registro anteriores a 2015:"
```

```
sum(data2017$anio_ocur < 2014)
```

```
## [1] 1259
```

```
# conservamos únicamente edad, capítulo, sexo,entidad residencia, ocupación, capítulo, edad agrupada
```

```
data2017<-data.frame(data2017$ent_resid,data2017$sexo,data2017$edad-4000,data2017$anio_ocur,data2017$ocupa)
```

Realizamos el cruce de la tabla de categorías de enfermedades y realizamos un pequeño resumen

```
#consdata[consdata$consdat2.anio_ocur > 2014, ]
dim(data2017)

## [1] 703047      6

names(data2017)

## [1] "data2017.ent_resid" "data2017.sexo"      "data2017.edad...4000"
## [4] "data2017.anio_ocur" "data2017.ocupacion" "data2017.capitulo"

datamx1<-merge(data2017,datacap, by.x="data2017.capitulo", by.y="CVE")
datamx1<-merge(datamx1,dataocu, by.x="data2017.ocupacion", by.y="CLAVE")
summary(datamx1)

## data2017.ocupacion data2017.capitulo data2017.ent_resid data2017.sexo
## Min. : 1.00      Min. : 1.00      Min. : 1.00      Min. :1.000
## 1st Qu.: 7.00      1st Qu.: 4.00      1st Qu.:10.00     1st Qu.:1.000
## Median :11.00      Median : 9.00      Median :15.00     Median :1.000
## Mean :20.54      Mean : 9.04      Mean :17.53      Mean :1.444
## 3rd Qu.:11.00      3rd Qu.:11.00     3rd Qu.:23.00     3rd Qu.:2.000
## Max. :99.00      Max. :20.00      Max. :99.00      Max. :9.000
##
## data2017.edad...4000 data2017.anio_ocur
## Min. : -2999.000      Min. :1943
## 1st Qu.: 51.000      1st Qu.:2017
## Median : 68.000      Median :2017
## Mean : 2.296      Mean :2018
## 3rd Qu.: 81.000      3rd Qu.:2017
## Max. : 998.000      Max. :9999
##
##
## DESCRIP.x
## Enfermedades del sistema circulatorio :180551
## Enfermedades endocrinas, nutricionales y metabólicas:123151
## Tumores (neoplasias) : 90572
## Causas externas de morbilidad y de mortalidad : 80400
## Enfermedades del sistema digestivo : 67542
## Enfermedades del sistema respiratorio : 59254
## (Other) :101577
##
## DESCRIP.y
## No trabaja :369804
## Trabajadores en actividades agrícolas, ganaderas, forestales, caza y pesca: 75037
## Trabajadores artesanales : 44411
## Comerciantes, empleados en ventas y agentes de ventas : 36493
## Ocupaciones insuficientemente especificadas : 32643
## No aplica a menores de 5 años : 30368
## (Other) :114291

causasenf<-data.frame(summary(datamx1$DESCRIP.x))
propcaumxprop <- prop.table(summary(datamx1$DESCRIP.x))

par(mfrow=c(1,2))
par(cex=0.4)
hist(causasenf$summary.datamx1.DESCRIP.x.)
library(MASS)
truehist(causasenf$summary.datamx1.DESCRIP.x., col='red')
```

```

par(mfrow=c(1,2))
par(cex=0.4)
hist(causasenf$summary.datamx1.DESCRIP.x.)
library(MASS)
truehist(causasenf$summary.datamx1.DESCRIP.x., col='red')

```

