



EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA



TECNOLÓGICO
NACIONAL DE MÉXICO®



Tecnológico Nacional de México
Instituto Tecnológico de Tijuana

Departamento de Sistemas y Computación
Ingeniería en Sistemas Computacionales

Semestre:

Febrero - Junio 2022

Materia:

Minería de datos

Docente:

JOSE CHRISTIAN ROMERO HERNANDEZ

Unidad:

U1

Título del trabajo:

Práctica Evaluativa - Unidad 1

alumno:

Anahi Del Carmen Hernandez Pablo 18210486

Pérez Mora Ana Ivonne

18212074



En esta parte escribimos la información que necesitábamos más la información que nos brindó el maestro

```
Countries_2013_Dataset <- c("Aruba","Afghanistan","Angola","Albania","United Arab Emirates","Argentina","Armenia","Antigua and Barbuda","Country_Code <- c("ABW","AFG","AGO","ALB","ARE","ARG","ARM","ATG","AUS","AUT","AZE","BDI","BEL","BEN","BFA","BGD","BGR","BHR","BHS","Regions_2013_Dataset <- c("The Americas","Asia","Africa","Europe","Middle East","The Americas","Asia","The Americas","Oceania","Life_Expectancy_At_Birth_1960 <- c(65.5693658536586,32.328512195122,32.9848292682927,62.2543658536585,52.2432195121951,65.215536Life_Expectancy_At_Birth_2013 <- c(75.3286585365854,60.0282682926829,51.8661707317073,77.537243902439,77.1956341463415,75.9860975
```

Aquí usamos esta función para cargar el documento

```
Informacion<- read.csv(file.choose())
```

Aquí filtró la información por expectativa de vida para cada año

```
Datos1960 <-data.frame(Country_Code, Life_Expectancy_At_Birth_1960)  
Datos2013<- data.frame(Country_Code, Life_Expectancy_At_Birth_2013)
```

Aquí nos muestra los nombres de las columnas que hay y filtra la información por año 1960-2013

```
names(Informacion)|  
  
info2013 <- filter(Informacion, Year == 2013)  
info1960 <- filter(Informacion, Year == 1960)  
  
DatosC <- data.frame(Codigo= Countries_2013_Dataset,Region= Regions_2013_Dataset)  
head(DatosC)
```

Aquí filtramos la información por código y región

```
DatosC <- data.frame(Codigo= Countries_2013_Dataset,Region= Regions_2013_Dataset)|  
head(DatosC)  
> head(DatosC)  
  
      Codigo      Region  
1      Aruba The Americas  
2 Afghanistan      Asia  
3      Angola      Africa  
4      Albania      Europe  
5 United Arab Emirates Middle East  
6      Argentina The Americas  
> |
```

En esta parte se agrupan por código y expectativa de vida según su año 1960-2013

```
Evida1960 <- data.frame(Code = Country_Code,  
                        Life_Expectancy1960 = Life_Expectancy_At_Birth_1960)  
  
head(Evida1960, 4)  
  
TFert2013<- data.frame(Code = Country_Code,  
                      Life_Expectancy2013 = Life_Expectancy_At_Birth_2013)  
  
head(TFert2013, 4)
```



```
> Evida1960 <- data.frame(Code = Country_Code,
+                          Life_Expectancy1960 = Life_Expectancy_At_Birth_1960)
>
> head(Evida1960, 4)
  Code Life_Expectancy1960
1 ABW                65.56937
2 AFG                32.32851
3 AGO                32.98483
4 ALB                62.25437
> TFert2013<- data.frame(Code = Country_Code,
+                         Life_Expectancy2013 = Life_Expectancy_At_Birth_2013)
>
> head(TFert2013, 4)
  Code Life_Expectancy2013
1 ABW                75.32866
2 AFG                60.02827
3 AGO                51.86617
4 ALB                77.53724
> |
```

Aquí usamos ese código para filtrar la información por región y año

```
MCountriesData1960 <- merge(Informacion, Evida1960,
+                           by.x = "Country.Code",
+                           by.y = "Code")
```

```
|
head(MCountriesData1960, 4)
> head(MCountriesData1960, 4)
  Country.Code Country.Name      Region Year Fertility.Rate Life_Expectancy1960
1          ABW      Aruba The Americas 1960         4.820             65.56937
2          ABW      Aruba The Americas 2013         1.669             65.56937
3          AFG Afghanistan      Asia 2013         5.050             32.32851
4          AFG Afghanistan      Asia 1960         7.450             32.32851
> |
```

Para esta práctica utilizamos las librerías de ggplot y ds labs

```
|
#install.packages("dslabs")
library(dslabs)

install.packages("ggplot2")
library(ggplot2)
```

En esta parte creamos la gráfica de regiones y en su eje X esta tasa de fertilidad y en su eje Y expectativa de vida cada una con su año este caso fue del 2013

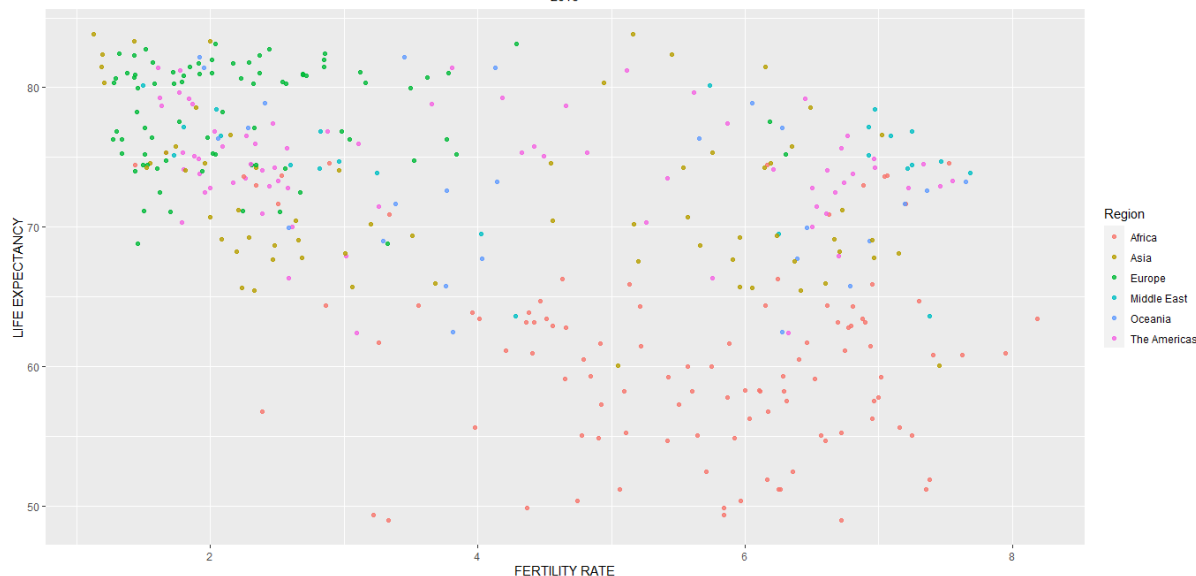
```
MCountriesData2013$Region <- as.factor(MCountriesData2013$Region)

RInfo2013 <- ggplot(data = MCountriesData2013,
+                  mapping = aes(x = Fertility.Rate,
+                                y = Life_Expectancy2013,
+                                color = Region))+
  geom_point(alpha = 0.8) +
  labs(title = "FERTILITY RATE VS LIFE EXPECTANCY",
+       subtitle = "2013",
+       x = "FERTILITY RATE",
+       y = "LIFE EXPECTANCY") +
  theme(plot.title=element_text(hjust = 0.5)) +
  theme(plot.subtitle=element_text(hjust = 0.5))

RInfo2013
```

Plot Zoom

FERTILITY RATE vs LIFE EXPECTANCY
2013



como se mencionó anteriormente pero con la gráfica del año 1960.

```
MCountriesData1960$Region <- as.factor(MCountriesData1960$Region)
RInfo1960 <- ggplot(data = MCountriesData1960,
  mapping = aes(x = Fertility.Rate,
    y = Life_Expectancy1960,
    color = Region))+
  geom_point(alpha = 0.8) +
  labs(title = "FERTILITY RATE vs LIFE EXPECTANCY",
    subtitle = "1960",
    x = "FERTILITY RATE",
    y = "LIFE EXPECTANCY") +
  theme(plot.title=element_text(hjust = 0.5)) +
  theme(plot.subtitle=element_text(hjust = 0.5))
RInfo1960
```

Plot Zoom

FERTILITY RATE vs LIFE EXPECTANCY
1960

