Proyecto-final

Estephanie Gomez

2025-02-10

```r
install.packages("arules", repos = "https://cloud.r-project.org/")
```

```
## Installing package into 'C:/Users/estep/AppData/Local/R/win-library/4.4'
## (as 'lib' is unspecified)
```

```
## package 'arules' successfully unpacked and MD5 sums checked
```

```
## Warning: cannot remove prior installation of package 'arules'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE): problem copying
## C:\Users\estep\AppData\Local\R\win-library\4.4\00LOCK\arules\libs\x64\arules.dll
## to C:\Users\estep\AppData\Local\R\win-library\4.4\arules\libs\x64\arules.dll:
## Permission denied
```

```
## Warning: restored 'arules'
```

```
##
## The downloaded binary packages are in
##   C:\Users\estep\AppData\Local\Temp\Rtmp2b7GRl\downloaded_packages
```

```r
# Verificar si el paquete 'arules' está instalado
if (!requireNamespace("arules", quietly = TRUE)) {
  install.packages("arules")
}

# Cargar el paquete 'arules'
library(arules)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'arules'
```

```
## The following objects are masked from 'package:base':
##
##     abbreviate, write
```

R Markdown

Objetivo:

Este script en R de algoritmo apriori va a obtener la posibilidad de tener diabetes dato las condiciones de alta presion arterial (HighBP), colesterol alto (HighChol), chequeo de colesterol (CholCheck) y fumadora (Smoker).

Eventos (A,B,C,D) -> Evento (diabetes)

Métricas:

- Soporte: mide la frecuencia con la que ocurre una combinación de elementos (reglas) en el conjunto de datos.
- Confianza: mide la probabilidad de que el lado derecho (rhs) de la regla ocurra dado que el lado izquierdo (lhs) de la regla es cierto
- Lift: fuerza de la asociación entre los elementos del lhs y rhs en la regla.

Librerias

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.4     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------- tidyverse_conflicts() --
## x tidyr::expand() masks Matrix::expand()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## x tidyr::pack()   masks Matrix::pack()
## x dplyr::recode() masks arules::recode()
## x tidyr::unpack() masks Matrix::unpack()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
datos <- read_csv(file = "Diabetes_data_modified.csv", col_names = TRUE)
```

```
## Rows: 70692 Columns: 6
## -- Column specification -------------------------------------------------------
## Delimiter: ","
## dbl (6): SessionID, HighBP, HighChol, CholCheck, Smoker, Diabetes_binary
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(datos)
```

```
## # A tibble: 6 x 6
##   SessionID HighBP HighChol CholCheck Smoker Diabetes_binary
##       <dbl>  <dbl>    <dbl>     <dbl>  <dbl>           <dbl>
## 1         1      1        1         0      1               0
## 2         2      1        1         1      1               0
## 3         3      0        0         1      0               0
## 4         4      1        1         1      1               0
## 5         5      0        0         1      1               0
## 6         6      0        0         1      0               0
```

EDA

Este bloque saca un EDA de los datos, en realidad todos son binarios expecto el Session ID, segun lo visto en clases a priori funciona con datos que vayan siendo transacciones

```
head(datos)
```

```
## # A tibble: 6 x 6
```

```
##   SessionID HighBP HighChol CholCheck Smoker Diabetes_binary
##       <dbl>  <dbl>    <dbl>     <dbl>  <dbl>           <dbl>
## 1        1      1        0         1      0               0
## 2        2      1        1         1      1               0
## 3        3      0        0         1      0               0
## 4        4      1        1         1      1               0
## 5        5      0        0         1      1               0
## 6        6      0        0         1      0               0
```

str(datos)

```
## spc_tbl_ [70,692 x 6] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
##  $ SessionID      : num [1:70692] 1 2 3 4 5 6 7 8 9 10 ...
##  $ HighBP         : num [1:70692] 1 1 0 1 0 0 0 0 0 0 ...
##  $ HighChol       : num [1:70692] 0 1 0 1 0 0 1 0 0 0 ...
##  $ CholCheck      : num [1:70692] 1 1 1 1 1 1 1 1 1 1 ...
##  $ Smoker         : num [1:70692] 0 1 0 1 1 0 1 1 0 1 ...
##  $ Diabetes_binary: num [1:70692] 0 0 0 0 0 0 0 0 0 0 ...
##  - attr(*, "spec")=
##   .. cols(
##   ..   SessionID = col_double(),
##   ..   HighBP = col_double(),
##   ..   HighChol = col_double(),
##   ..   CholCheck = col_double(),
##   ..   Smoker = col_double(),
##   ..   Diabetes_binary = col_double()
##   .. )
##  - attr(*, "problems")=<externalptr>
```

colSums(**is.na**(datos))

```
##       SessionID          HighBP        HighChol       CholCheck          Smoker
##               0               0               0               0               0
## Diabetes_binary
##               0
```

summary(datos)

```
##    SessionID          HighBP          HighChol         CholCheck
##  Min.   :    1   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
##  1st Qu.:17674   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:1.0000
##  Median :35347   Median :1.0000   Median :1.0000   Median :1.0000
##  Mean   :35347   Mean   :0.5635   Mean   :0.5257   Mean   :0.9753
##  3rd Qu.:53019   3rd Qu.:1.0000   3rd Qu.:1.0000   3rd Qu.:1.0000
##  Max.   :70692   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000
##      Smoker       Diabetes_binary
##  Min.   :0.0000   Min.   :0.0
##  1st Qu.:0.0000   1st Qu.:0.0
##  Median :0.0000   Median :0.5
##  Mean   :0.4753   Mean   :0.5
##  3rd Qu.:1.0000   3rd Qu.:1.0
##  Max.   :1.0000   Max.   :1.0
```

Transformacion de datos

Filtrar por una condición, en este caso diabetes binary es la condicion y seleccionar las columnas de interés, trabajando el data set de diabetesFiltrar por una condición y seleccionar las columnas de interés, trabajando el data set de diabetes.

```r
data_trans <- data.frame(
  HighBP = as.factor(datos$HighBP),          # Cambiar 'data' por 'datos'
  HighChol = as.factor(datos$HighChol),      # Cambiar 'data' por 'datos'
  CholCheck = as.factor(datos$CholCheck),    # Cambiar 'data' por 'datos'
  Smoker = as.factor(datos$Smoker),          # Cambiar 'data' por 'datos'
  Diabetes_binary = as.factor(datos$Diabetes_binary)  # Cambiar 'data' por 'datos'
)

# Convertir el dataframe a transacciones
transactions <- as(data_trans, "transactions")

# Ver los primeros datos
head(datos)
```

```
## # A tibble: 6 x 6
##   SessionID HighBP HighChol CholCheck Smoker Diabetes_binary
##       <dbl>  <dbl>    <dbl>     <dbl>  <dbl>           <dbl>
## 1         1      1        0         1      0               0
## 2         2      1        1         1      1               0
## 3         3      0        0         1      0               0
## 4         4      1        1         1      1               0
## 5         5      0        0         1      1               0
## 6         6      0        0         1      0               0
```

Algoritmo Apriori

Aplicar el algoritmo Apriori para encontrar reglas de asociación y Definir los parámetros de support y confidence

```r
reglas <- apriori(transactions,
                  parameter = list(support = 0.1, confidence = 0.8))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.8    0.1    1 none FALSE            TRUE       5     0.1      1
##  maxlen target   ext
##      10  rules  TRUE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 7069
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[10 item(s), 70692 transaction(s)] done [0.01s].
## sorting and recoding items ... [9 item(s)] done [0.00s].
```

```
## creating transaction tree ... done [0.02s].
## checking subsets of size 1 2 3 4 5 done [0.00s].
## writing ... [60 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```r
# Ver las reglas generadas
inspect(reglas)
```

```
##       lhs                      rhs              support confidence  coverage      lift count
## [1]  {}                    => {CholCheck=1}   0.9752589  0.9752589 1.0000000 1.0000000 68943
## [2]  {HighBP=0}            => {CholCheck=1}   0.4177842  0.9570318 0.4365416 0.9813105 29534
## [3]  {HighChol=0}          => {CholCheck=1}   0.4558932  0.9611978 0.4742970 0.9855822 32228
## [4]  {Smoker=1}            => {CholCheck=1}   0.4631783  0.9745521 0.4752730 0.9992753 32743
## [5]  {Diabetes_binary=0}   => {CholCheck=1}   0.4786680  0.9573360 0.5000000 0.9816225 33838
## [6]  {Diabetes_binary=1}   => {CholCheck=1}   0.4965908  0.9931817 0.5000000 1.0183775 35105
## [7]  {Smoker=0}            => {CholCheck=1}   0.5120806  0.9758991 0.5247270 1.0006564 36200
## [8]  {HighChol=1}          => {CholCheck=1}   0.5193657  0.9879450 0.5257030 1.0130080 36715
## [9]  {HighBP=1}            => {CholCheck=1}   0.5574747  0.9893804 0.5634584 1.0144798 39409
## [10] {HighBP=0,
##       HighChol=0}          => {CholCheck=1}   0.2700588  0.9461294 0.2854354 0.9701316 19091
## [11] {HighBP=0,
##       Smoker=1}            => {CholCheck=1}   0.1771346  0.9532582 0.1858202 0.9774412 12522
## [12] {HighBP=0,
##       Diabetes_binary=0}   => {CholCheck=1}   0.2955073  0.9444796 0.3128784 0.9684399 20890
## [13] {HighBP=0,
##       Diabetes_binary=1}   => {CholCheck=1}   0.1222769  0.9887898 0.1236632 1.0138741  8644
## [14] {HighBP=0,
##       Smoker=0}            => {CholCheck=1}   0.2406496  0.9598285 0.2507214 0.9841782 17012
## [15] {HighBP=0,
##       HighChol=1}          => {CholCheck=1}   0.1477253  0.9776259 0.1511062 1.0024271 10443
## [16] {HighChol=0,
##       Smoker=1}            => {CholCheck=1}   0.1936287  0.9579397 0.2021304 0.9822414 13688
## [17] {HighChol=0,
##       Diabetes_binary=0}   => {CholCheck=1}   0.2924518  0.9453564 0.3093561 0.9693390 20674
## [18] {HighChol=0,
##       Diabetes_binary=1}   => {CholCheck=1}   0.1634414  0.9909091 0.1649409 1.0160472 11554
## [19] {HighChol=0,
##       Smoker=0}            => {CholCheck=1}   0.2622645  0.9636175 0.2721666 0.9880633 18540
## [20] {HighBP=1,
##       HighChol=0}          => {CholCheck=1}   0.1858343  0.9839712 0.1888615 1.0089334 13137
## [21] {Smoker=1,
##       Diabetes_binary=0}   => {CholCheck=1}   0.2061195  0.9535371 0.2161631 0.9777271 14571
## [22] {Smoker=1,
##       Diabetes_binary=1}   => {CholCheck=1}   0.2570588  0.9920839 0.2591099 1.0172518 18172
## [23] {HighChol=1,
##       Smoker=1}            => {CholCheck=1}   0.2695496  0.9868455 0.2731426 1.0118806 19055
## [24] {HighBP=1,
##       Smoker=1}            => {CholCheck=1}   0.2860437  0.9882221 0.2894528 1.0132921 20221
## [25] {Smoker=0,
##       Diabetes_binary=0}   => {CholCheck=1}   0.2725485  0.9602293 0.2838369 0.9845891 19267
## [26] {HighChol=1,
##       Diabetes_binary=0}   => {CholCheck=1}   0.1862163  0.9767752 0.1906439 1.0015548 13164
## [27] {HighBP=1,
##       Diabetes_binary=0}   => {CholCheck=1}   0.1831608  0.9788328 0.1871216 1.0036646 12948
## [28] {Smoker=0,
```

```
##          Diabetes_binary=1} => {CholCheck=1}         0.2395321 0.9943626 0.2408901 1.0195883 16933
## [29] {HighChol=1,
##          Diabetes_binary=1} => {HighBP=1}             0.2717846 0.8111543 0.3350591 1.4395993 19213
## [30] {HighChol=1,
##          Diabetes_binary=1} => {CholCheck=1}         0.3331494 0.9943004 0.3350591 1.0195246 23551
## [31] {HighBP=1,
##          Diabetes_binary=1} => {CholCheck=1}         0.3743139 0.9946249 0.3763368 1.0198573 26461
## [32] {HighChol=1,
##          Smoker=0}          => {CholCheck=1}         0.2498161 0.9891341 0.2525604 1.0142272 17660
## [33] {HighBP=1,
##          Smoker=0}          => {CholCheck=1}         0.2714310 0.9906040 0.2740055 1.0157344 19188
## [34] {HighBP=1,
##          HighChol=1}        => {CholCheck=1}         0.3716404 0.9921075 0.3745968 1.0172761 26272
## [35] {HighBP=0,
##          HighChol=0,
##          Smoker=1}          => {CholCheck=1}         0.1065326 0.9392617 0.1134216 0.9630896  7531
## [36] {HighBP=0,
##          HighChol=0,
##          Smoker=0}          => {Diabetes_binary=0} 0.1398461 0.8129934 0.1720138 1.6259868  9886
## [37] {HighBP=0,
##          HighChol=0,
##          Diabetes_binary=0} => {CholCheck=1}         0.2104906 0.9353196 0.2250467 0.9590476 14880
## [38] {HighBP=0,
##          HighChol=0,
##          Smoker=0}          => {CholCheck=1}         0.1635263 0.9506579 0.1720138 0.9747749 11560
## [39] {HighBP=0,
##          Smoker=1,
##          Diabetes_binary=0} => {CholCheck=1}         0.1161518 0.9369010 0.1239744 0.9606690  8211
## [40] {HighBP=0,
##          Smoker=0,
##          Diabetes_binary=0} => {CholCheck=1}         0.1793555 0.9494533 0.1889040 0.9735398 12679
## [41] {HighChol=0,
##          Smoker=1,
##          Diabetes_binary=0} => {CholCheck=1}         0.1162791 0.9389993 0.1238330 0.9628206  8220
## [42] {HighChol=0,
##          Smoker=0,
##          Diabetes_binary=0} => {CholCheck=1}         0.1761727 0.9495997 0.1855231 0.9736899 12454
## [43] {HighBP=1,
##          HighChol=0,
##          Diabetes_binary=1} => {CholCheck=1}         0.1038731 0.9935056 0.1045521 1.0187096  7343
## [44] {HighChol=1,
##          Smoker=1,
##          Diabetes_binary=1} => {HighBP=1}             0.1471878 0.8140354 0.1808125 1.4447125 10405
## [45] {HighChol=1,
##          Smoker=1,
##          Diabetes_binary=1} => {CholCheck=1}         0.1797092 0.9938977 0.1808125 1.0191116 12704
## [46] {HighBP=1,
##          Smoker=1,
##          Diabetes_binary=1} => {CholCheck=1}         0.1960759 0.9939763 0.1972642 1.0191923 13861
## [47] {HighBP=1,
##          HighChol=1,
##          Smoker=1}          => {CholCheck=1}         0.1989475 0.9910507 0.2007441 1.0161924 14064
## [48] {HighBP=1,
##          HighChol=1,
```

```
##          Diabetes_binary=0} => {CholCheck=1}        0.1011996  0.9843148 0.1028122 1.0092857  7154
## [49] {HighChol=1,
##         Smoker=0,
##         Diabetes_binary=1} => {HighBP=1}            0.1245968  0.8077770 0.1542466 1.4336054  8808
## [50] {HighChol=1,
##         Smoker=0,
##         Diabetes_binary=1} => {CholCheck=1}         0.1534403  0.9947726 0.1542466 1.0200087 10847
## [51] {HighBP=1,
##         Smoker=0,
##         Diabetes_binary=1} => {CholCheck=1}         0.1782380  0.9953393 0.1790726 1.0205898 12600
## [52] {HighBP=1,
##         HighChol=1,
##         Diabetes_binary=1} => {CholCheck=1}         0.2704408  0.9950554 0.2717846 1.0202988 19118
## [53] {HighChol=1,
##         CholCheck=1,
##         Diabetes_binary=1} => {HighBP=1}            0.2704408  0.8117702 0.3331494 1.4406924 19118
## [54] {HighBP=1,
##         HighChol=1,
##         Smoker=0}          => {CholCheck=1}         0.1726928  0.9933279 0.1738528 1.0185274 12208
## [55] {HighBP=0,
##         HighChol=0,
##         Smoker=0,
##         Diabetes_binary=0} => {CholCheck=1}         0.1316556  0.9414323 0.1398461 0.9653153  9307
## [56] {HighBP=0,
##         HighChol=0,
##         CholCheck=1,
##         Smoker=0}          => {Diabetes_binary=0}   0.1316556  0.8051038 0.1635263 1.6102076  9307
## [57] {HighBP=1,
##         HighChol=1,
##         Smoker=1,
##         Diabetes_binary=1} => {CholCheck=1}         0.1464239  0.9948102 0.1471878 1.0200473 10351
## [58] {HighChol=1,
##         CholCheck=1,
##         Smoker=1,
##         Diabetes_binary=1} => {HighBP=1}            0.1464239  0.8147827 0.1797092 1.4460389 10351
## [59] {HighBP=1,
##         HighChol=1,
##         Smoker=0,
##         Diabetes_binary=1} => {CholCheck=1}         0.1240169  0.9953451 0.1245968 1.0205958  8767
## [60] {HighChol=1,
##         CholCheck=1,
##         Smoker=0,
##         Diabetes_binary=1} => {HighBP=1}            0.1240169  0.8082419 0.1534403 1.4344305  8767
```

Responder que reglas hacen que un paciente tenga diabetes

Filtrar reglas que tengan 'Diabetes_binary = 1' en el lado derecho

```
# Ver las reglas filtradas
inspect(reglas[1:10])
```

```
##      lhs                  rhs             support   confidence coverage
## [1] {}                => {CholCheck=1} 0.9752589 0.9752589  1.0000000
## [2] {HighBP=0}        => {CholCheck=1} 0.4177842 0.9570318  0.4365416
```

```
## [3]  {HighChol=0}           => {CholCheck=1} 0.4558932 0.9611978  0.4742970
## [4]  {Smoker=1}             => {CholCheck=1} 0.4631783 0.9745521  0.4752730
## [5]  {Diabetes_binary=0}    => {CholCheck=1} 0.4786680 0.9573360  0.5000000
## [6]  {Diabetes_binary=1}    => {CholCheck=1} 0.4965908 0.9931817  0.5000000
## [7]  {Smoker=0}             => {CholCheck=1} 0.5120806 0.9758991  0.5247270
## [8]  {HighChol=1}           => {CholCheck=1} 0.5193657 0.9879450  0.5257030
## [9]  {HighBP=1}             => {CholCheck=1} 0.5574747 0.9893804  0.5634584
## [10] {HighBP=0, HighChol=0} => {CholCheck=1} 0.2700588 0.9461294  0.2854354
##      lift      count
## [1]  1.0000000 68943
## [2]  0.9813105 29534
## [3]  0.9855822 32228
## [4]  0.9992753 32743
## [5]  0.9816225 33838
## [6]  1.0183775 35105
## [7]  1.0006564 36200
## [8]  1.0130080 36715
## [9]  1.0144798 39409
## [10] 0.9701316 19091
```

```r
# Filtrar reglas donde 'Diabetes_binary=1' esté en el lado izquierdo (lhs)
reglas_diabetes <- subset(reglas, lhs %in% "Diabetes_binary=1")

# Ver las reglas filtradas
inspect(reglas_diabetes)
```

```
##      lhs                       rhs              support confidence coverage     lift count
## [1]  {Diabetes_binary=1} => {CholCheck=1} 0.4965908   0.9931817 0.5000000 1.018378 35105
## [2]  {HighBP=0,
##       Diabetes_binary=1} => {CholCheck=1} 0.1222769   0.9887898 0.1236632 1.013874  8644
## [3]  {HighChol=0,
##       Diabetes_binary=1} => {CholCheck=1} 0.1634414   0.9909091 0.1649409 1.016047 11554
## [4]  {Smoker=1,
##       Diabetes_binary=1} => {CholCheck=1} 0.2570588   0.9920839 0.2591099 1.017252 18172
## [5]  {Smoker=0,
##       Diabetes_binary=1} => {CholCheck=1} 0.2395321   0.9943626 0.2408901 1.019588 16933
## [6]  {HighChol=1,
##       Diabetes_binary=1} => {HighBP=1}    0.2717846   0.8111543 0.3350591 1.439599 19213
## [7]  {HighChol=1,
##       Diabetes_binary=1} => {CholCheck=1} 0.3331494   0.9943004 0.3350591 1.019525 23551
## [8]  {HighBP=1,
##       Diabetes_binary=1} => {CholCheck=1} 0.3743139   0.9946249 0.3763368 1.019857 26461
## [9]  {HighBP=1,
##       HighChol=0,
##       Diabetes_binary=1} => {CholCheck=1} 0.1038731   0.9935056 0.1045521 1.018710  7343
## [10] {HighChol=1,
##       Smoker=1,
##       Diabetes_binary=1} => {HighBP=1}    0.1471878   0.8140354 0.1808125 1.444712 10405
## [11] {HighChol=1,
##       Smoker=1,
##       Diabetes_binary=1} => {CholCheck=1} 0.1797092   0.9938977 0.1808125 1.019112 12704
## [12] {HighBP=1,
##       Smoker=1,
##       Diabetes_binary=1} => {CholCheck=1} 0.1960759   0.9939763 0.1972642 1.019192 13861
## [13] {HighChol=1,
```

```
##         Smoker=0,
##          Diabetes_binary=1} => {HighBP=1}      0.1245968  0.8077770 0.1542466 1.433605  8808
## [14] {HighChol=1,
##          Smoker=0,
##          Diabetes_binary=1} => {CholCheck=1} 0.1534403  0.9947726 0.1542466 1.020009 10847
## [15] {HighBP=1,
##          Smoker=0,
##          Diabetes_binary=1} => {CholCheck=1} 0.1782380  0.9953393 0.1790726 1.020590 12600
## [16] {HighBP=1,
##          HighChol=1,
##          Diabetes_binary=1} => {CholCheck=1} 0.2704408  0.9950554 0.2717846 1.020299 19118
## [17] {HighChol=1,
##          CholCheck=1,
##          Diabetes_binary=1} => {HighBP=1}      0.2704408  0.8117702 0.3331494 1.440692 19118
## [18] {HighBP=1,
##          HighChol=1,
##          Smoker=1,
##          Diabetes_binary=1} => {CholCheck=1} 0.1464239  0.9948102 0.1471878 1.020047 10351
## [19] {HighChol=1,
##          CholCheck=1,
##          Smoker=1,
##          Diabetes_binary=1} => {HighBP=1}      0.1464239  0.8147827 0.1797092 1.446039 10351
## [20] {HighBP=1,
##          HighChol=1,
##          Smoker=0,
##          Diabetes_binary=1} => {CholCheck=1} 0.1240169  0.9953451 0.1245968 1.020596  8767
## [21] {HighChol=1,
##          CholCheck=1,
##          Smoker=0,
##          Diabetes_binary=1} => {HighBP=1}      0.1240169  0.8082419 0.1534403 1.434431  8767
```

Data frame con tablas de asociacion

Mejora visual, para ver las reglas.

```r
# Convertir las reglas filtradas a un data.frame
reglas_df <- as(reglas_diabetes, "data.frame")

# Guardar en csv

write.csv(reglas_df, "reglas_diabetes.csv", row.names = FALSE)

# Ver las primeras filas del data.frame
head(reglas_df)
```

```
##                                        rules   support confidence
## 6             {Diabetes_binary=1} => {CholCheck=1} 0.4965908  0.9931817
## 13   {HighBP=0,Diabetes_binary=1} => {CholCheck=1} 0.1222769  0.9887898
## 18 {HighChol=0,Diabetes_binary=1} => {CholCheck=1} 0.1634414  0.9909091
## 22   {Smoker=1,Diabetes_binary=1} => {CholCheck=1} 0.2570588  0.9920839
## 28   {Smoker=0,Diabetes_binary=1} => {CholCheck=1} 0.2395321  0.9943626
## 29     {HighChol=1,Diabetes_binary=1} => {HighBP=1} 0.2717846  0.8111543
##     coverage     lift count
## 6  0.5000000 1.018378 35105
```

```
## 13 0.1236632 1.013874  8644
## 18 0.1649409 1.016047 11554
## 22 0.2591099 1.017252 18172
## 28 0.2408901 1.019588 16933
## 29 0.3350591 1.439599 19213
```

Conclusiones

Vamos a analizar una de las reglas, recordar que R da las reglas con asociacion mas fuerte.

Regla: {Smoker=1, Diabetes_binary=1} => {CholCheck=1} Esta regla sugiere que, aunque el fumar es un factor de riesgo, los fumadores con diabetes tienen una probabilidad extremadamente alta (99.21%) de hacerse un chequeo de colesterol.

El soporte es de 25.715 es decir es la regla que cumple con las condiciones de persona fumadora, con diabetes que tiene chequeos de colesterol, la confianza de que esta regla sea cierta cuando se cumple las 3 condiciones es de un 99,21%. El lift de 1.071, lo que quiere decir es que hay una relacion positiva entre los dos eventos.