

INFORME TÉCNICO

SATISFACCIÓN DE NIÑOS Y ABUELOS

Por:

Daniel Espinal Mosquera

Juan Sebastián Falcón

Juan F. Peña Tamayo

Brayan M. Ortiz Fajardo

Thalea Marina Hesse



Universidad Nacional de Colombia
Sede Medellín

Introducción

La satisfacción de la vida y la felicidad varía entre países [7] y juegan un papel importante en el desarrollo de un país. Sin embargo, no se logró determinar documentación o estudios relacionados sobre cómo se pueden predecir estos factores en Colombia dados unos parámetros. Por tal motivo, se propone el planteamiento, desarrollo, análisis y posterior productización de un modelo con el cual se busca predecir la satisfacción de niños y abuelos. Para lograr esto, se toma como base una encuesta realizada por el DANE en el 2020: Colombia - Encuesta Nacional de Calidad de Vida - ECV 2020. Esta investigación, según el DANE “Busca cuantificar y caracterizar las condiciones socioeconómicas de los hogares colombianos, con el fin de obtener la información necesaria para la actualización de indicadores sociales a nivel de viviendas, hogares y personas, y para la definición de políticas que permitan diseñar y ejecutar planes sociales.” (Metodología ECV, 2009, p.17)

La estructura del estudio se planteó de la siguiente manera: se hizo una búsqueda exhaustiva sobre documentación para determinar cuáles de las variables que se tienen afectan de manera significativa la satisfacción. Después se planteó un modelo general, sin embargo al revisar las correlaciones entre las variables predictoras se decidió partir ese modelo general en tres sub-modelos: satisfacción de salud, seguridad y trabajo. Para cada uno de estos también se realizó la búsqueda de documentación al respecto. Adicionalmente, se creó una página web para poder interactuar con los modelos. Finalmente, se obtuvieron los resultados y plantearon las conclusiones.

Planteamiento del Problema

El Instituto Colombiano de Bienestar Familiar es una entidad que trabaja por la prevención y protección integral de la primera infancia, la niñez, la adolescencia y el bienestar general de las familias en Colombia, llegando a millones de colombianos mediante sus programas, estrategias y servicios de atención. En el marco de los objetivos de esta institución se encontró que el ICBF actualmente no cuenta con una herramienta para conocer en prospectiva, y de forma adecuada y efectiva la satisfacción general de vida tanto de niños como de adultos en la tercera edad. Es para ellos de vital importancia conocer esta información pues es un indicador fundamental a tener en cuenta a la hora de crear programas preventivos y de protección que tienen como objetivo el mejoramiento de vida de la población destinataria. Por esto se busca implementar en el ICBF tanto los 3 sub-modelos como el modelo de satisfacción general, para que sea usado por la institución en pro de mejorar futuros planeamientos en todo proyecto social que involucre niños y adultos de la tercera edad como población objetivo.

Justificación

Niños

En primera instancia se planteó tomar a los niños en dos grupos, uno como aquellos pertenecientes a la primera infancia (0 a 5 años) y otro con aquellos niños con edad entre 6 y 12 años. Sin embargo, luego se decidió que tomaríamos como niño la definición integrada en el código de infancia y adolescencia, donde se expone que “Para todos los efectos de esta ley son sujetos titulares de derechos todas las personas menores de 18 años. Sin perjuicio de lo establecido en el artículo 34 del Código Civil, se entiende por niño o niña las personas entre los 0 y los 12 años, y por adolescente las personas entre 12 y 18 años de edad”(Artículo 3).

Una vez adoptada esta definición se analizó cuántas observaciones de la ECV cumplían esta condición, resultando en un total de 56128 niños.

Abuelos

Para los abuelos, al igual que con los niños, se pensó inicialmente en tomar un rango de edad que tomara la definición popular de este colectivo, los adultos de la tercera edad (mayores de 60 años). Sin embargo, luego planteamos que debíamos tener en cuenta cuál es la definición literal de abuelo, y mediante un sistema de

grafos logramos determinar el numero de hombres y mujeres que tenian un nieto. Filtrandolos por su rango edad.

Mayor a 60 años	1467
Menor a 60 años	1049

Se observa que con la definicion inicial estabamos omitiendo un total de 1049 observaciones, ademas, se observa que los abuelos resgitrados en la base de datos son relativamente pocos pues solo representan el x % del total de personas.

Ante esta situación tomamos la desicion de ...

Seleccion de Variables

Para predecir la satisfacción se tomaron en cuenta las evidencias halladas en la bibliografía, donde se encontró que las variables relacionadas con el nivel de satisfacción de una persona son las siguientes:

- NIVEL_DE_EDUCACION
- SEXO
- SALUD_AUTOPERCIBIDA
- ESTADO_CIVIL
- ETNIA
- INGRESO_AUTOPERCIBIDO
- SEGURIDAD_AUTOPERCIBIDA
- TRABAJO_AUTOPERCIBIDO
- SATISFACCION
- I_HOGAR
- PERCAPITA
- COND_VIDA_DEL_HOGAR

El primer paso para verificar si las variables anteriores servirán para ajustar algún modelo es hacer un análisis descriptivo de cada una agrupando los niños y abuelos.

Modelos Predictivos

Inicialmente, se intentó englobar en un modelo a los abuelos y niños con el fin de predecir la satisfacción. Sin embargo, como lo ilustra la tabla 1, las variables objetivos que se seleccionaron no fueron respondidas, en su mayoría, por niños. Este comportamiento se asemeja con los resultados encontrados en [7], donde se puede observar que los abuelos y niños tienen diferentes definiciones de satisfacción y, por ende, diferentes factores que la influyen. Por esta razón, se decidió trabajar de forma independiente los modelos para los niños y abuelos.

Modelos Predictivos en Abuelos

```
## Loading required package: lattice
```

Satisfacción de la Vida en General

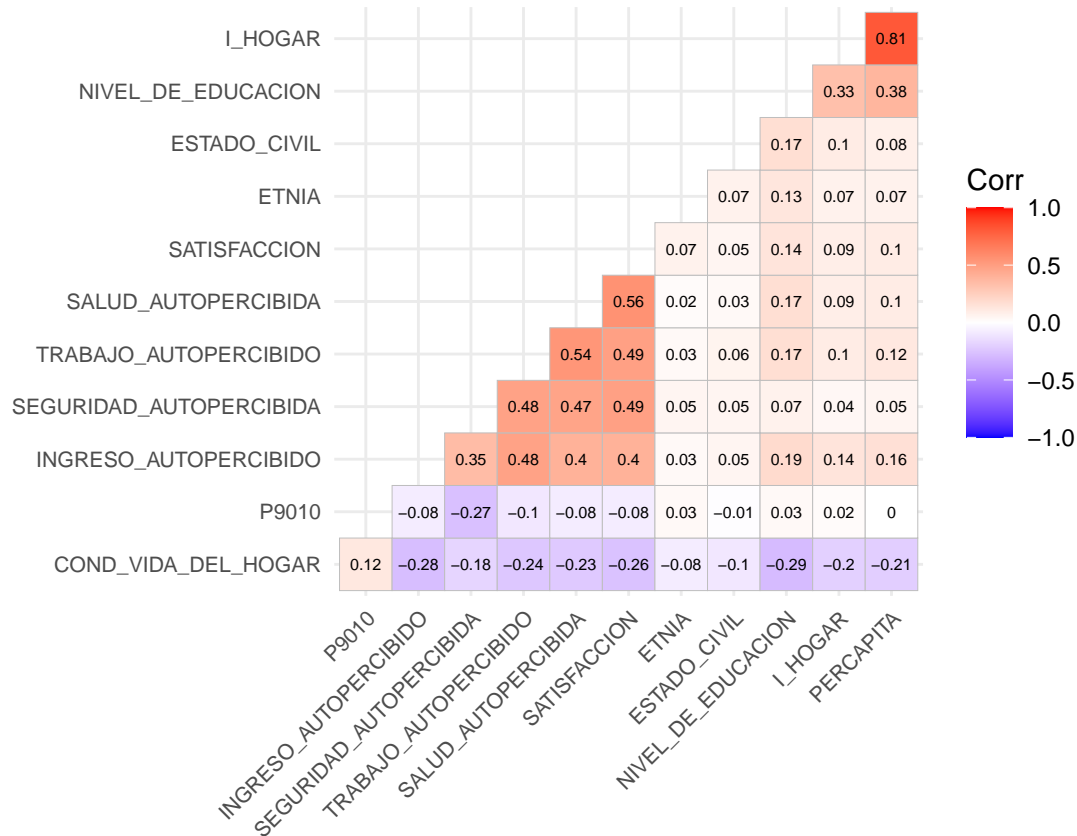
Análisis Descriptivo

Clasificación	ID_Persona	NIVEL_DE_EDUCACION	SEXO	SALUD_AUTOPERCIBIDA	ESTADO_CIVIL	ETNIA	INGRESO_AUTOPERCIBIDO
3ra_edad	37721	37721	37721	37721	37721	37721	37721
niño	32576	32576	32576	0	14464	32576	0

SEGURIDAD_AUTOPERCIBIDA	TRABAJO_AUTOPERCIBIDO	SATISFACCION	I_HOGAR	PERCAPITA	COND_VIDA_DEL_HOGAR	P9010
37721	37721	37721	37721	37721	37721	37721
0	0	0	32576	32576	32576	32576

Conteo de respuestas Según la tabla anterior las variables que preguntan directamente por la satisfacción no son respondidas por los niños, por lo tanto se decide abordar y desarrollar modelos por separado para niños y abuelos.

Matriz de Correlaciones Para empezar la selección de posibles variables objetivo y predictoras de las escogidas anteriormente se desarrolla el siguiente gráfico de correlación:



Se evalúa la posibilidad de establecer como posible variable objetivo la satisfacción, esta tiene una buena correlación con las variables de satisfacción pero no con el resto, por lo tanto es posible que no tenga mucho sentido predecir la satisfacción general de un abuelo a partir de todas estas variables. Con base en lo anterior se evalúa la implementación de varios modelos, uno por cada variable de satisfacción.

KNN

```
##      Y KNN_Predict
## 2    8           10
## 4    4           10
## 11   6           10
## 17   9           10
## 18   7           10
## 20   8           10
```

Regresion con lm y glm

```
##
```

```
## Call:
## lm(formula = Y ~ ., data = train)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.7221 -0.7421  0.0334  0.7687  7.2361
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.882886   0.059204  65.585 < 2e-16 ***
## X1           0.291916   0.005082  57.447 < 2e-16 ***
## X2           0.206558   0.005097  40.528 < 2e-16 ***
## X3           0.015651   0.005008   3.125 0.00178 **
## X4          -0.291487   0.014676 -19.862 < 2e-16 ***
## X5           0.137483   0.004334  31.720 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.437 on 28279 degrees of freedom
## Multiple R-squared:  0.4074, Adjusted R-squared:  0.4073
## F-statistic: 3889 on 5 and 28279 DF, p-value: < 2.2e-16

##      Y Y_Predict
## 2  8  7.576154
## 4  4  7.471508
## 11 6  8.276218
## 17 9  8.620259
## 18 7  6.114202
## 20 8  7.915893

##      1
## 4.96325
```

Satisfacción en la Salud

Según (PONER BIBLIOGRAFÍA) se seleccionan las siguientes variables como posibles predictores de la satisfacción de la Salud de los abuelos:

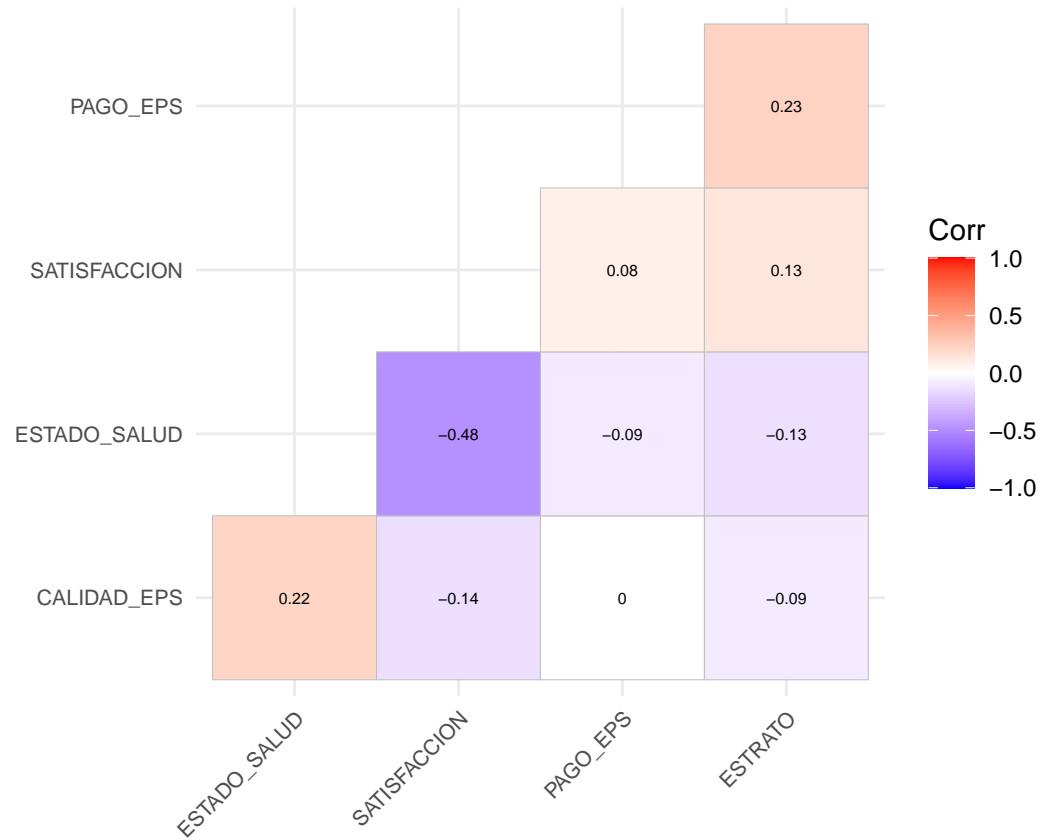
- AFILIADO (p6090)
- PAGO_EPS (p8551)
- CALIDAD_PRESTADOR (p6181)
- ESTADO_SALUD (p6127)
- TIPO_PAGO (p6115)

Análisis Descriptivo

```
##      DIRECTORIO P8520S1A1 ID_Persona SECUENCIA_P ID_Hogar P6020 P1897 P5502 P6080
## 1      7247300          3 724730011          1 72473001      2      8      2      5
## 2      7247300          3 724730021          1 72473001      1      7      2      5
## 3      7247300          3 724730031          1 72473001      1     10      4      5
## 4      7247301          3 724730111          1 72473011      1      7      6      6
## 5      7247301          3 724730121          1 72473011      2      7      6      6
## 6      7247302          3 724730211          1 72473021      1      7      2      6
##      Clasificación P1896 P1898 P1899 P1895 P6081S1 P6083S1 P6090 P8551 P6181 P6115
## 1      3ra_edad      5      8      6      7      NA      NA      1      NA      2      5
## 2      3ra_edad      5      8      6      8      NA      NA      1      NA      2      5
```

```
## 3      3ra_edad      6      10      4      9      NA      NA      1      NA      2      5
## 4      3ra_edad      7      10      0      4      NA      NA      1      NA      2      5
## 5      3ra_edad     10      10      0     10      NA      NA      1      NA      2      5
## 6          otro      NA       8       7       6      NA      NA      1      NA     NA     NA
##      P6127
## 1       2
## 2       2
## 3       2
## 4       3
## 5       3
## 6       2

##      SATISFACCION PAGO_EPS CALIDAD_EPS ESTADO_SALUD ESTRATO
## 89              10   110000              2              2        2
## 199             3   120000              2              2        1
## 209             4    74500              2              3        1
## 242             6   145000              2              3        2
## 641             6    87780              2              2        2
## 664             9   144000              2              2        2
```



Matriz de Correlaciones

Regresion con lm

```
##      SATISFACCION PAGO_EPS CALIDAD_EPS ESTADO_SALUD ESTRATO
## 89              10   110000              2              2        2
## 209             4    74500              2              3        1
## 641             6    87780              2              2        2
## 664             9   144000              2              2        2
```

```
## 752          10    98000          2          2          1
## 767           7   160000          2          2          3

##
## Call:
## lm(formula = SATISFACCION ~ ., data = train_salud)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.7624 -0.8855  0.1418  1.1809  5.2537
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.081e+01  1.352e-01  79.920 < 2e-16 ***
## PAGO_EPS      4.279e-07  1.550e-07   2.761  0.00579 **
## CALIDAD_EPS   -9.211e-02  4.342e-02  -2.121  0.03395 *
## ESTADO_SALUD -1.517e+00  4.318e-02 -35.132 < 2e-16 ***
## ESTRATO       7.972e-02  1.969e-02   4.049 5.23e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.727 on 4636 degrees of freedom
## Multiple R-squared:  0.2385, Adjusted R-squared:  0.2379
## F-statistic: 363 on 4 and 4636 DF, p-value: < 2.2e-16

##      SATISFACCION Y_Predict
## 199           3  7.721832
## 242           6  6.295304
## 666           8  7.795272
## 1233          7  7.672508
## 1234          8  7.612173
## 1535          6  7.840058

## 'data.frame':  4641 obs. of  5 variables:
## $ SATISFACCION: int  10 4 6 9 10 7 10 6 10 6 ...
## $ PAGO_EPS : int  110000 74500 87780 144000 98000 160000 75000 170000 400000 90000 ...
## $ CALIDAD_EPS : num  2 2 2 2 2 2 1 4 2 2 ...
## $ ESTADO_SALUD: int  2 3 2 2 2 2 1 3 2 3 ...
## $ ESTRATO : int  2 1 2 2 1 3 1 2 1 3 ...
## - attr(*, "na.action")= 'omit' Named int [1:30652] 1 2 3 4 5 6 7 8 9 10 ...
## ..- attr(*, "names")= chr [1:30652] "1" "2" "3" "4" ...
```

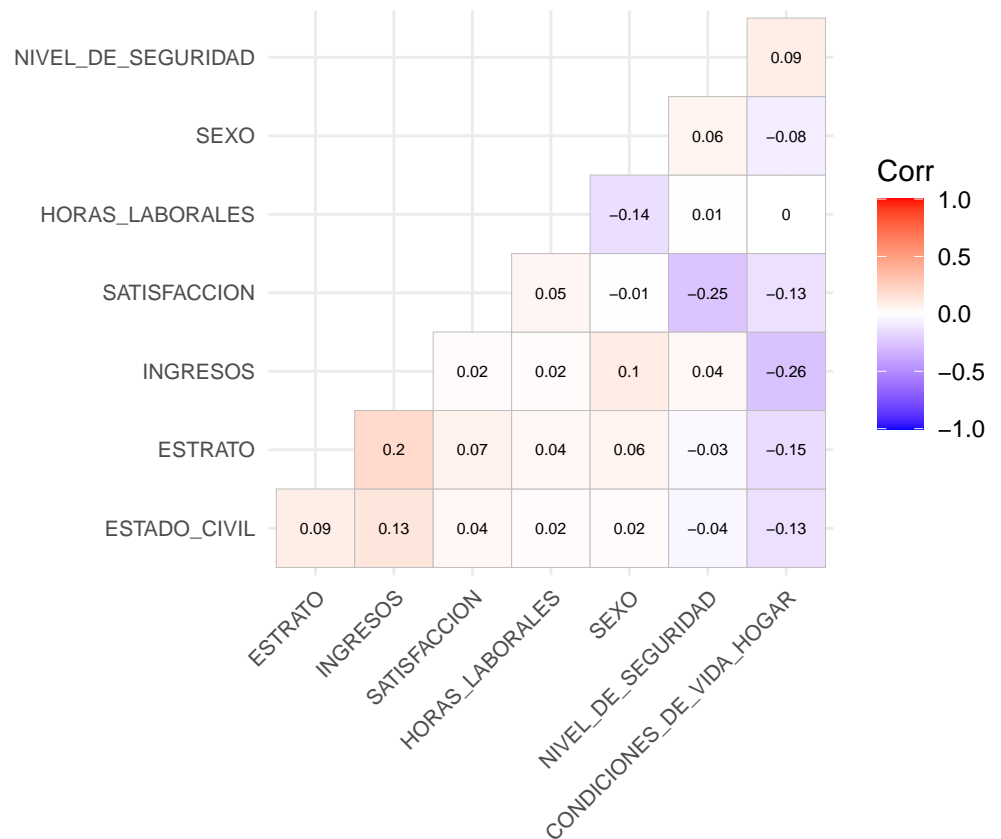
Satisfacción sobre el Nivel de Seguridad

Según (PONER BIBLIOGRAFÍA) se seleccionan las siguientes variables como posibles predictores de la satisfacción de la seguridad de los abuelos:

- ESTADO_CIVIL (P5502)
- SEXO (P6020')
- ESTRATO (P8520S1A1)
- INGRESOS (P8624)
- HORAS_LABORALES (P415)
- NIVEL_DE_SEGURIDAD (P9010)
- CONDICIONES_DE_VIDA_HOGAR (P9030)

Análisis Descriptivo

##	DIRECTORIO	ID_Hogar	ID_Persona	SECUENCIA_P	P6020	P1897	P5502	P6080		
## 1	7247300	72473001	724730011	1	2	8	2	5		
## 2	7247300	72473001	724730031	1	1	10	4	5		
## 3	7247300	72473001	724730021	1	1	7	2	5		
## 4	7247301	72473011	724730111	1	1	7	6	6		
## 5	7247301	72473011	724730121	1	2	7	6	6		
## 6	7247302	72473021	724730231	1	1	8	5	6		
##	Clasificación	P1896	P1898	P1899	P1895	P6081S1	P6083S1	P6435	P6440	P6460
## 1	3ra_edad	5	8	6	7	NA	NA	NA	NA	NA
## 2	3ra_edad	6	10	4	9	NA	NA	NA	NA	NA
## 3	3ra_edad	5	8	6	8	NA	NA	NA	NA	NA
## 4	3ra_edad	7	10	0	4	NA	NA	NA	NA	NA
## 5	3ra_edad	10	10	0	10	NA	NA	NA	NA	NA
## 6	otro	NA	6	8	6	1	2	1	NA	1
##	P8624	P8631	P8642	P415	P9030	P9010	P8520S1A1			
## 1	NA	NA	2	NA	3	1	3			
## 2	NA	NA	2	NA	3	1	3			
## 3	NA	NA	2	NA	3	1	3			
## 4	NA	NA	2	NA	2	1	3			
## 5	NA	NA	2	NA	2	1	3			
## 6	1000000	2	2	48	2	1	3			
##	SATISFACCION	ESTADO_CIVIL	SEXO	ESTRATO	INGRESOS	HORAS_LABORALES				
## 58	5		4	1	3	20000				
## 166	6		6	1	1	99				
## 193	8		2	1	2	99				
## 203	7		5	1	1	400000				
## 524	7		5	2	2	500000				
## 594	10		2	1	2	872000				
##	NIVEL_DE_SEGURIDAD	CONDICIONES_DE_VIDA_HOGAR								
## 58	1	1								
## 166	1	2								
## 193	1	2								
## 203	1	2								
## 524	1	2								
## 594	1	2								



Matriz de Correlaciones

Regresion con lm

```
##
## Call:
## lm(formula = SATISFACCION ~ ., data = train_seguridad)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.1082 -0.9855  0.1060  1.4853  3.9804
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   9.638e+00  3.107e-01  31.017 < 2e-16 ***
## ESTADO_CIVIL   6.777e-04  2.662e-02   0.025  0.97969
## SEXO           2.707e-02  9.843e-02   0.275  0.78332
## ESTRATO        6.282e-02  2.591e-02   2.425  0.01540 *
## INGRESOS      -9.806e-09  2.401e-08  -0.408  0.68300
## HORAS_LABORALES 7.428e-03  2.754e-03   2.697  0.00707 **
## NIVEL_DE_SEGURIDAD -1.474e+00  1.309e-01 -11.259 < 2e-16 ***
## CONDICIONES_DE_VIDA_HOGAR -3.429e-01  7.786e-02 -4.404 1.12e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.838 on 1881 degrees of freedom
## Multiple R-squared:  0.08628,    Adjusted R-squared:  0.08287
## F-statistic: 25.37 on 7 and 1881 DF,  p-value: < 2.2e-16
##
##      SATISFACCION Y_Predict
## 166              6  8.300956
```

```
## 203          7  7.776413
## 693          5  7.606514
## 986         10  7.798966
## 1046         7  7.908798
## 1245         9  7.773315
```

Satisfacción en el Trabajo

Según (PONER BIBLIOGRAFÍA) se seleccionan las siguientes variables como posibles predictores de la satisfacción en el trabajo de los abuelos:

- SEXO (P6020')
- CARGO (P6435)
- TIENE_CONTRATO (P6440)
- TIPO_CONTRATO (P6460)
- SALARIO (P8624)
- HORAS_LABORALES (P415)
- RECIBIO_PRIMAS (P8631)
- RECIBIO_PENSIONES (P8642)

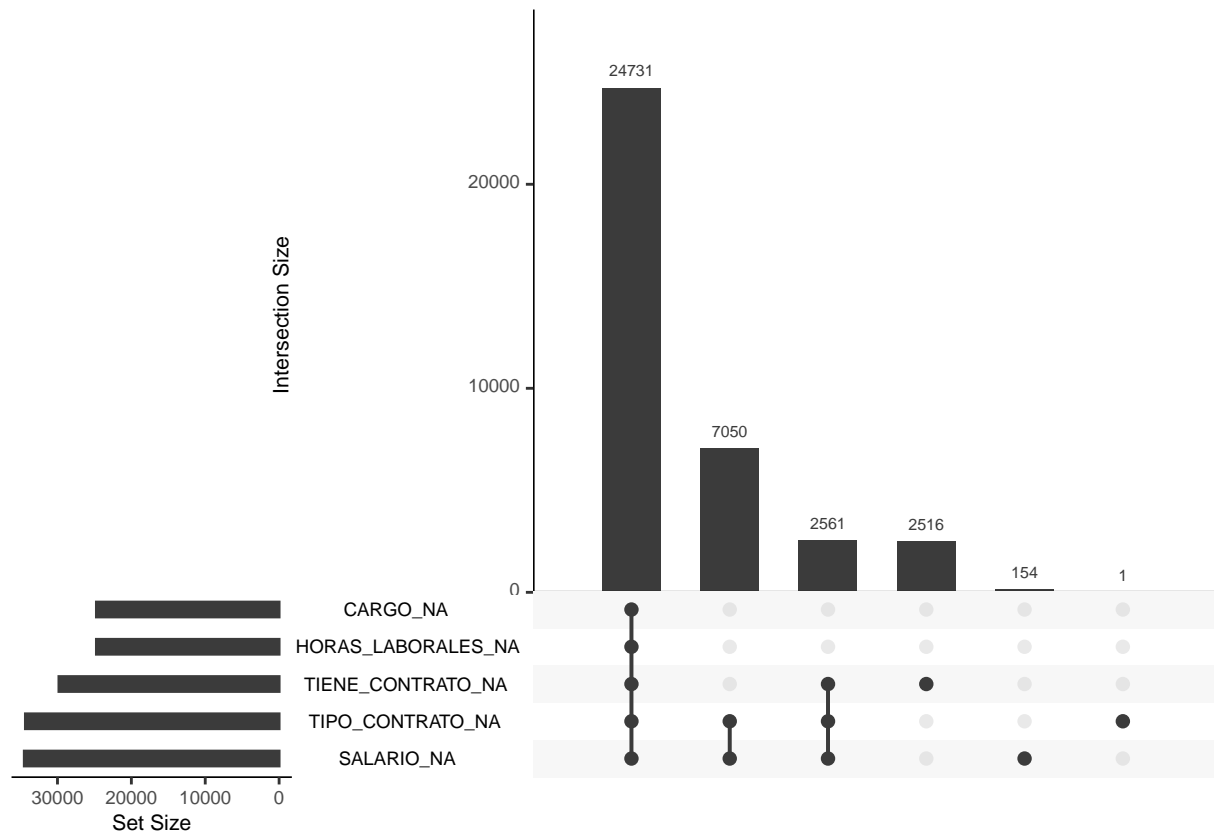
Análisis Descriptivo

```
## SATISFACCION SEXO CARGO TIENE_CONTRATO TIPO_CONTRATO SALARIO HORAS_LABORALES
## 1          6    2   NA              NA              NA      NA      NA
## 2          4    1   NA              NA              NA      NA      NA
## 3          6    1   NA              NA              NA      NA      NA
## 4          0    1   NA              NA              NA      NA      NA
## 5          0    2   NA              NA              NA      NA      NA
## 10         4    2   NA              NA              NA      NA      NA
## RECIBIO_PRIMAS RECIBIO_PENSIONES
## 1          NA              2
## 2          NA              2
## 3          NA              2
## 4          NA              2
## 5          NA              2
## 10         NA              2
```

Datos Faltantes Para este dataframe se tiene la siguiente cantidad de abuelos:

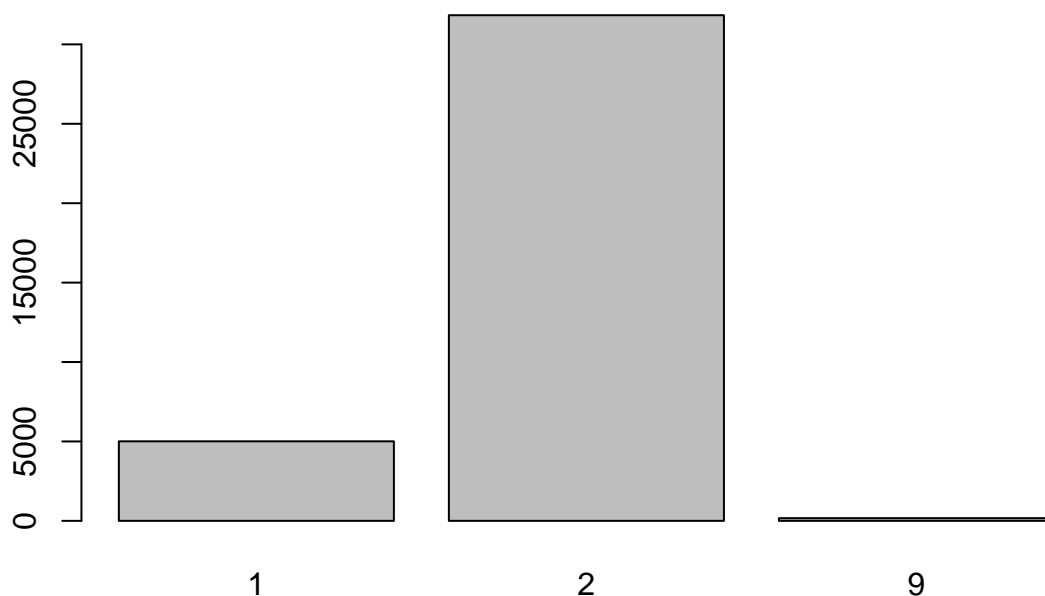
```
## Cantidad de Abuelos
## 1          37013
```

Sin embargo, si se observa la cantidad de abuelos que respondieron a las preguntas seleccionadas:



se puede determinar que la mayoría de estos no respondieron a las preguntas que se les hicieron sobre el trabajo. Este mismo procedimiento se repitió con variables diferentes, pero no se obtuvieron resultados distintos a los presentados. Por tal motivo, se decide no realizar un modelo de predicción para la satisfacción del trabajo en los abuelos. En primera instancia se pensó que este comportamiento se debía a que la mayoría de los abuelos estaban pensionados, pero si se observan los abuelos pensionados:

¿Recibe pensión?



la gran mayoría de estos respondieron que no recibían algún tipo de pensión (2). Por tanto, se puede inferir que la mayoría de los abuelos no trabajan y tampoco reciben pensión, es decir, viven dependientes de sus familiares.

Modelos Predictivos en Niños

```
# los datos:
ninos <- read.csv2("Datos/datos_ninos.csv", header=TRUE, dec=".", encoding="UTF-8")
```

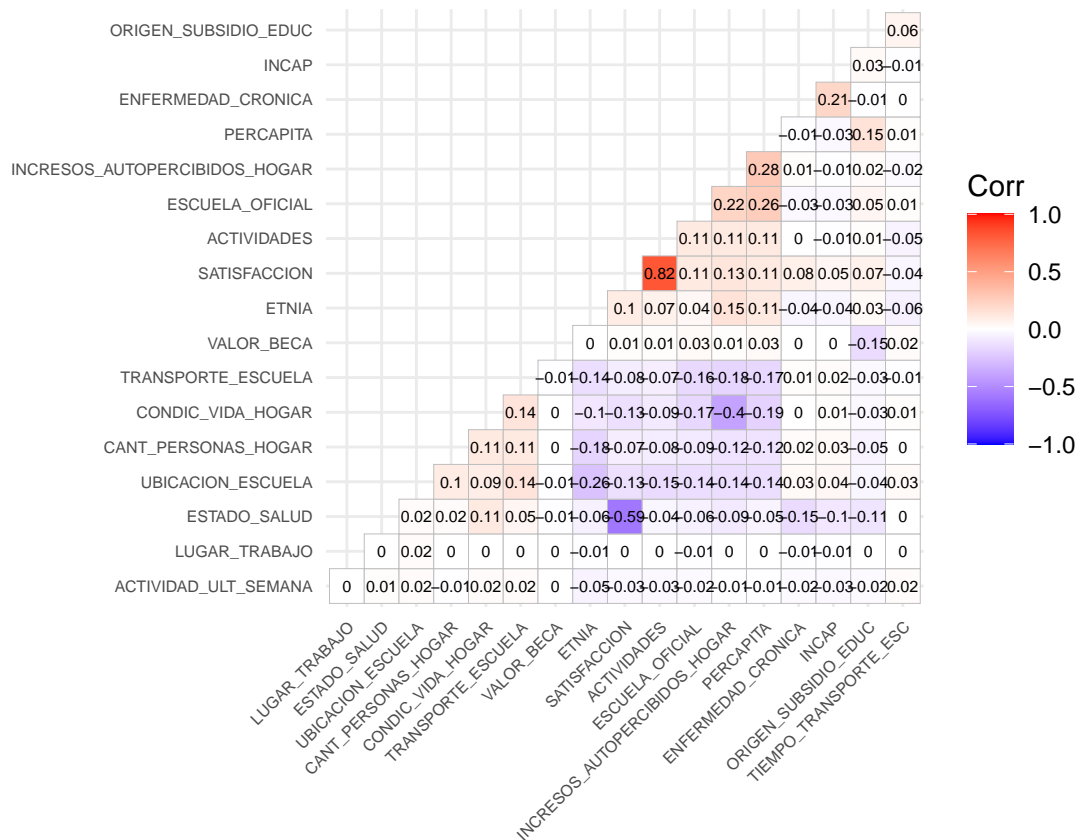
Correlaciones

```
corr_ninos <- ninos %>% select(SATISFACCION, ESTADO_SALUD, ACTIVIDADES, ETNIA,
                              CONDIC_VIDA_HOGAR, INCRESOS_AUTOPERCIBIDOS_HOGAR,
                              PERCAPITA, CANT_PERSONAS_HOGAR, ESCUELA_OFICIAL,
                              VALOR_BECA, ORIGEN_SUBSIDIO_EDUC, UBICACION_ESCUELA,
                              TRANSPORTE_ESCUELA, TIEMPO_TRANSPORTE_ESC,
                              ACTIVIDAD_ULT_SEMANA, LUGAR_TRABAJO,
                              ENFERMEDAD_CRONICA, INCAP)

M = cor(corr_ninos, use = "pairwise.complete.obs")

M[is.na(M)]=0

ggcorrplot(M, hc.order = TRUE, type = "lower", lab = TRUE, tl.cex = 6,
           pch.col = "red", lab_size = 2)
```



Los datos para niños:

#KNN Minutos para escuela, ValorBeca, PERCAPITA

```
df_ninos <- ninos %>% select(EDAD, INCAP, SATISFACCION,
                             PERCAPITA, INGRESOS_AUTOPERCIBIDOS_HOGAR, UBICACION_ESCUELA, CONDIC_VIDA_HOGAR)

df_ninos['ETNIA'] = as.factor(ninos[, 'ETNIA'])

df_ninos['VIVE_CON_PADRE'] = factor(ninos[, 'VIVE_CON_PADRE'], labels = c(TRUE,
                                                                           FALSE,
                                                                           'Muerto'))

df_ninos['VIVE_CON_MADRE'] = factor(ninos[, 'VIVE_CON_MADRE'], labels = c(TRUE,
                                                                           FALSE,
                                                                           'Muerto'))

df_ninos['ESCUELA_OFICIAL'] = factor(ninos[, 'ESCUELA_OFICIAL'], labels = c('Oficial',
                                                                           'conSubstituto',
                                                                           'SinSubstituto'))

df_ninos['TRANSPORTE_ESCUELA'] = addNA(factor(ninos[, 'TRANSPORTE_ESCUELA'],
                                              labels = c('Carro', 'escolar', 'público',
                                                         'pie', 'Bicicleta', 'Caballo',
                                                         'canoa', 'Otro'))))

df_ninos['ENFERMEDAD_CRONICA'] = addNA(factor(ninos[, 'ENFERMEDAD_CRONICA'],
                                              labels = c(TRUE, FALSE)))

df_ninos['ACTIVIDAD_ULT_SEMANA'] = addNA(factor(ninos[, 'ACTIVIDAD_ULT_SEMANA'],
                                              labels = c('Trabajando', 'Buscando',
                                                         'Estudiando', 'Oficios_hogar',
                                                         'Incapacitado trabajar', 'Otra'))))

df_ninos['LUGAR_TRABAJO'] = addNA(factor(ninos[, 'LUGAR_TRABAJO'],
                                          labels = c('no trabajo', 'la vivienda',
                                                         'otra vivienda', 'Puerta',
                                                         'calle', 'oficina', 'campo',
                                                         'obra'))))

head(df_ninos)
```

##	EDAD	INCAP	SATISFACCION	PERCAPITA	INGRESOS_AUTOPERCIBIDOS_HOGAR
## 1	7	32	31	694500.0	2
## 2	10	32	25	265083.3	2
## 3	8	32	34	300000.0	2
## 4	7	32	25	759078.5	1
## 5	6	32	31	954500.0	2
## 6	7	32	31	600000.0	2

##	UBICACION_ESCUELA	CONDIC_VIDA_HOGAR	ETNIA	VIVE_CON_PADRE	VIVE_CON_MADRE	
## 1		1	2	6	FALSE	FALSE
## 2		1	2	6	FALSE	TRUE
## 3		1	2	6	TRUE	TRUE
## 4		1	2	6	FALSE	TRUE
## 5		1	3	6	TRUE	TRUE
## 6		1	2	6	TRUE	TRUE

```
## ESCUELA_OFICIAL TRANSPORTE_ESCUELA ENFERMEDAD_CRONICA ACTIVIDAD_ULT_SEMANA
## 1      Oficial          pie          FALSE      Estudiando
## 2      Oficial          Carro         FALSE      Estudiando
## 3      SinSubstituto    escolar      FALSE      Estudiando
## 4      SinSubstituto    Carro         FALSE      Estudiando
## 5      SinSubstituto    pie          FALSE      Estudiando
## 6      SinSubstituto    pie          FALSE      Estudiando
## LUGAR_TRABAJO
## 1      no trabajo
## 2      no trabajo
## 3      no trabajo
## 4      no trabajo
## 5      no trabajo
## 6      no trabajo
```

Resultados del entrenamiento

```
# Separación de los datos en entrenamiento y prueba
```

```
datos1 <- sample(2, nrow(df_ninos),
                replace = TRUE,
                prob = c(0.75, 0.25))
```

```
train <- df_ninos[datos1 == 1,]
```

```
test <- df_ninos[datos1 == 2,]
```

```
# Modelo de arbol de decicion
```

```
modelo <- ctree(formula = SATISFACCION ~ ., data=train, controls = ctree_control(mincriterion = 0.7))
```

```
#t <- predict(modelo2, newdata = test)
```

```
#colnames(t) <- c('pred_Satisfaccion_3')
```

```
#test_pred <- cbind(test_pred, t)
```

```
png(file = "decision_tree.png", width = 3000, height = 1500, )
```

```
plot(modelo)
```

```
dev.off()
```

```
## pdf
```

```
## 2
```

```
modelo
```

```
##
```

```
## Conditional inference tree with 63 terminal nodes
```

```
##
```

```
## Response: SATISFACCION
```

```
## Inputs: EDAD, INCAP, PERCAPITA, INCRESOS_AUTOPERCIBIDOS_HOGAR, UBICACION_ESCUELA, CONDIC_VIDA_HOGAR
```

```
## Number of observations: 23461
```

```
##
```

```
## 1) INCRESOS_AUTOPERCIBIDOS_HOGAR <= 1; criterion = 1, statistic = 434.179
```

```
## 2) UBICACION_ESCUELA <= 2; criterion = 1, statistic = 155.675
```

```
## 3) ENFERMEDAD_CRONICA == {FALSE}; criterion = 1, statistic = 96.629
```

```
## 4) TRANSPORTE_ESCUELA == {Carro, pie, Bicicleta}; criterion = 1, statistic = 75.579
```

```
## 5) PERCAPITA <= 33333.33; criterion = 1, statistic = 34.371
```

```
## 6) CONDIC_VIDA_HOGAR <= 2; criterion = 1, statistic = 23.587
```

```
## 7) VIVE_CON_PADRE == {FALSE, Muerto}; criterion = 0.956, statistic = 11.483
```

```
## 8) ACTIVIDAD_ULT_SEMANA == {NA}; criterion = 0.826, statistic = 10.684
```

```

##          9)* weights = 8
##          8) ACTIVIDAD_ULT_SEMANA == {Estudiando, Oficios_hogar, Otra}
##          10)* weights = 48
##          7) VIVE_CON_PADRE == {TRUE}
##          11)* weights = 78
##          6) CONDIC_VIDA_HOGAR > 2
##          12) ETNIA == {6}; criterion = 0.999, statistic = 19.798
##          13)* weights = 60
##          12) ETNIA == {1, 5}
##          14)* weights = 107
##          5) PERCAPITA > 33333.33
##          15) ESCUELA_OFICIAL == {Oficial, SinSubstituto}; criterion = 1, statistic = 26.436
##          16) PERCAPITA <= 196000; criterion = 1, statistic = 18.893
##          17) UBICACION_ESCUELA <= 1; criterion = 0.901, statistic = 11.8
##          18) ETNIA == {3, 6}; criterion = 0.958, statistic = 13.888
##          19)* weights = 1397
##          18) ETNIA == {1, 5}
##          20) LUGAR_TRABAJO == {NA}; criterion = 0.918, statistic = 7.81
##          21) CONDIC_VIDA_HOGAR <= 2; criterion = 0.827, statistic = 7.385
##          22)* weights = 25
##          21) CONDIC_VIDA_HOGAR > 2
##          23)* weights = 42
##          20) LUGAR_TRABAJO == {no trabajo}
##          24) TRANSPORTE_ESCUELA == {pie, Bicicleta}; criterion = 0.971, statistic = 12.3
##          25)* weights = 340
##          24) TRANSPORTE_ESCUELA == {Carro}
##          26)* weights = 21
##          17) UBICACION_ESCUELA > 1
##          27) VIVE_CON_PADRE == {TRUE}; criterion = 0.971, statistic = 12.324
##          28)* weights = 532
##          27) VIVE_CON_PADRE == {FALSE, Muerto}
##          29)* weights = 387
##          16) PERCAPITA > 196000
##          30) UBICACION_ESCUELA <= 1; criterion = 0.969, statistic = 9.307
##          31) VIVE_CON_PADRE == {TRUE, Muerto}; criterion = 0.915, statistic = 10.13
##          32)* weights = 1308
##          31) VIVE_CON_PADRE == {FALSE}
##          33)* weights = 1121
##          30) UBICACION_ESCUELA > 1
##          34)* weights = 555
##          15) ESCUELA_OFICIAL == {conSubstituto}
##          35)* weights = 59
##          4) TRANSPORTE_ESCUELA == {escolar, público, Caballo, canoa, Otro, NA}
##          36) ETNIA == {2, 4, 5, 6}; criterion = 1, statistic = 41.566
##          37) CONDIC_VIDA_HOGAR <= 1; criterion = 1, statistic = 20.885
##          38)* weights = 72
##          37) CONDIC_VIDA_HOGAR > 1
##          39) TRANSPORTE_ESCUELA == {escolar, público, Otro, NA}; criterion = 0.994, statistic = 2
##          40) INCAP <= 30; criterion = 0.991, statistic = 11.557
##          41)* weights = 40
##          40) INCAP > 30
##          42) UBICACION_ESCUELA <= 1; criterion = 0.989, statistic = 11.193
##          43)* weights = 1990
##          42) UBICACION_ESCUELA > 1

```

```

##          44)* weights = 440
## 39) TRANSPORTE_ESCUELA == {Caballo, canoa}
##          45) ETNIA == {5}; criterion = 0.994, statistic = 12.306
##          46) CONDIC_VIDA_HOGAR <= 2; criterion = 0.855, statistic = 6.445
##          47)* weights = 9
##          46) CONDIC_VIDA_HOGAR > 2
##          48)* weights = 31
##          45) ETNIA == {6}
##          49) TRANSPORTE_ESCUELA == {Caballo}; criterion = 0.749, statistic = 6.543
##          50)* weights = 9
##          49) TRANSPORTE_ESCUELA == {canoa}
##          51) UBICACION_ESCUELA <= 1; criterion = 0.919, statistic = 9.782
##          52)* weights = 16
##          51) UBICACION_ESCUELA > 1
##          53)* weights = 13
## 36) ETNIA == {1, 3}
##          54) CONDIC_VIDA_HOGAR <= 2; criterion = 0.926, statistic = 8.881
##          55)* weights = 210
##          54) CONDIC_VIDA_HOGAR > 2
##          56) PERCAPITA <= 8958.333; criterion = 0.987, statistic = 10.983
##          57)* weights = 14
##          56) PERCAPITA > 8958.333
##          58) PERCAPITA <= 320000; criterion = 0.935, statistic = 7.967
##          59)* weights = 151
##          58) PERCAPITA > 320000
##          60)* weights = 12
## 3) ENFERMEDAD_CRONICA == {TRUE}
##          61) EDAD <= 10; criterion = 0.912, statistic = 13.208
##          62)* weights = 175
##          61) EDAD > 10
##          63)* weights = 87
## 2) UBICACION_ESCUELA > 2
##          64) ENFERMEDAD_CRONICA == {FALSE}; criterion = 1, statistic = 71.53
##          65) ETNIA == {3, 6}; criterion = 1, statistic = 40.045
##          66)* weights = 1885
##          65) ETNIA == {1, 2, 4, 5}
##          67) TRANSPORTE_ESCUELA == {Bicicleta, Otro}; criterion = 1, statistic = 56.279
##          68)* weights = 30
##          67) TRANSPORTE_ESCUELA == {Carro, escolar, público, pie, Caballo, canoa, NA}
##          69) ESCUELA_OFICIAL == {SinSubstituto}; criterion = 1, statistic = 30.945
##          70)* weights = 19
##          69) ESCUELA_OFICIAL == {Oficial, conSubstituto}
##          71) ACTIVIDAD_ULT_SEMANA == {Estudiando, NA}; criterion = 0.989, statistic = 16.714
##          72) VIVE_CON_MADRE == {TRUE}; criterion = 0.78, statistic = 8.076
##          73)* weights = 1238
##          72) VIVE_CON_MADRE == {FALSE, Muerto}
##          74) VIVE_CON_PADRE == {FALSE}; criterion = 0.991, statistic = 14.672
##          75) TRANSPORTE_ESCUELA == {Carro, escolar, NA}; criterion = 0.952, statistic = 17.
##          76)* weights = 20
##          75) TRANSPORTE_ESCUELA == {pie, Caballo, canoa}
##          77)* weights = 72
##          74) VIVE_CON_PADRE == {TRUE, Muerto}
##          78) PERCAPITA <= 60722.22; criterion = 0.77, statistic = 5.546
##          79)* weights = 12

```



```

##          78) PERCAPITA > 60722.22
##          80)* weights = 40
##      71) ACTIVIDAD_ULT_SEMANA == {Oficios_hogar, Otra}
##          81) PERCAPITA <= 175333.3; criterion = 0.97, statistic = 21.895
##          82)* weights = 73
##          81) PERCAPITA > 175333.3
##          83)* weights = 12
##      64) ENFERMEDAD_CRONICA == {TRUE}
##          84) PERCAPITA <= 352500; criterion = 0.808, statistic = 8.654
##          85)* weights = 59
##          84) PERCAPITA > 352500
##          86)* weights = 10
##      1) INCRESOS_AUTOPERCIBIDOS_HOGAR > 1
##          87) CONDIC_VIDA_HOGAR <= 1; criterion = 1, statistic = 342.804
##          88) INCRESOS_AUTOPERCIBIDOS_HOGAR <= 2; criterion = 0.993, statistic = 20.676
##          89) ENFERMEDAD_CRONICA == {FALSE}; criterion = 0.987, statistic = 11.016
##          90)* weights = 669
##          89) ENFERMEDAD_CRONICA == {TRUE}
##          91)* weights = 24
##          88) INCRESOS_AUTOPERCIBIDOS_HOGAR > 2
##          92) ETNIA == {6}; criterion = 0.953, statistic = 14.549
##          93)* weights = 215
##          92) ETNIA == {1, 3, 5}
##          94)* weights = 16
##          87) CONDIC_VIDA_HOGAR > 1
##          95) UBICACION_ESCUELA <= 1; criterion = 1, statistic = 138.86
##          96) ENFERMEDAD_CRONICA == {FALSE}; criterion = 1, statistic = 40.837
##          97) ESCUELA_OFICIAL == {SinSubstituto}; criterion = 1, statistic = 41.446
##          98) TRANSPORTE_ESCUELA == {Carro, escolar, público, pie, Otro, NA}; criterion = 0.996, sta
##          99)* weights = 773
##          98) TRANSPORTE_ESCUELA == {Bicicleta}
##          100)* weights = 13
##          97) ESCUELA_OFICIAL == {Oficial, conSubstituto}
##          101) INCRESOS_AUTOPERCIBIDOS_HOGAR <= 2; criterion = 1, statistic = 24.089
##          102) PERCAPITA <= 332857.1; criterion = 1, statistic = 25.923
##          103) TRANSPORTE_ESCUELA == {Carro, público, pie, NA}; criterion = 1, statistic = 35.98
##          104) CONDIC_VIDA_HOGAR <= 2; criterion = 0.731, statistic = 9.546
##          105)* weights = 1793
##          104) CONDIC_VIDA_HOGAR > 2
##          106)* weights = 479
##          103) TRANSPORTE_ESCUELA == {escolar, Bicicleta, Caballo, canoa, Otro}
##          107)* weights = 268
##          102) PERCAPITA > 332857.1
##          108) INCAP <= 30; criterion = 0.955, statistic = 15.731
##          109)* weights = 55
##          108) INCAP > 30
##          110)* weights = 2740
##          101) INCRESOS_AUTOPERCIBIDOS_HOGAR > 2
##          111)* weights = 240
##          96) ENFERMEDAD_CRONICA == {TRUE}
##          112)* weights = 173
##          95) UBICACION_ESCUELA > 1
##          113) UBICACION_ESCUELA <= 2; criterion = 1, statistic = 22.464
##          114)* weights = 1210

```

```
##      113) UBICACION_ESCUELA > 2
##      115) ENFERMEDAD_CRONICA == {FALSE}; criterion = 1, statistic = 23.79
##      116) ESCUELA_OFICIAL == {SinSubstituto}; criterion = 0.997, statistic = 17.06
##      117) PERCAPITA <= 545000; criterion = 0.997, statistic = 20.032
##      118)* weights = 84
##      117) PERCAPITA > 545000
##      119)* weights = 8
##      116) ESCUELA_OFICIAL == {Oficial, conSubstituto}
##      120) INCRESOS_AUTOPERCIBIDOS_HOGAR <= 2; criterion = 0.987, statistic = 18.459
##      121) TRANSPORTE_ESCUELA == {Caballo, NA}; criterion = 0.809, statistic = 18.965
##      122)* weights = 338
##      121) TRANSPORTE_ESCUELA == {Carro, escolar, público, pie, Bicicleta, canoa, Otro}
##      123)* weights = 1469
##      120) INCRESOS_AUTOPERCIBIDOS_HOGAR > 2
##      124)* weights = 29
##      115) ENFERMEDAD_CRONICA == {TRUE}
##      125)* weights = 38

# used variables: P9090,P6223,P9030, PERCAPITA, P1930, ETNIA, P6167, incap, CANT_PERSONAS_HOGAR, P5673,
# variables inútiles: P8610S1, P6238, P420
```

```
test_pred <- predict(modelo, newdata = test)
colnames(test_pred) <- c('pred_Satisfaccion')
test_pred <- cbind(test_pred, test['SATISFACCION'])

cat_test_pred = lapply(test_pred/5.6, as.integer)
table(cat_test_pred)
```

```
##          SATISFACCION
## pred_Satisfaccion  0   1   2   3   4   5   6   7   8   9
##          3   0   1   4   8  15   2   2   1   0   0
##          4   1  17  41 347 1298 437 343 100 41   1
##          5   0   6  33 569 1865 1046 880 367 147   8
##          6   0   1   0  18  60  20  71  54  32   2
```

```
train_pred <- predict(modelo, newdata = train)
colnames(train_pred) <- c('pred_Satisfaccion')
train_pred <- cbind(train_pred, train['SATISFACCION'])

cat_train_pred = lapply(train_pred/5.6, as.integer)
table(cat_train_pred)
```

```
##          SATISFACCION
## pred_Satisfaccion  0   1   2   3   4   5   6   7   8   9
##          3   1  11  16  54  35   8   2   2   0   0
##          4   3  51 148 1135 3781 1159 986 251  50   4
##          5   1  35  89 1634 5575 3227 2673 1150 460  28
##          6   0   1   1  46  174  71  269  194 127   9
```

calculate MSE

```
print('MSE:' )
```

```
## [1] "MSE:"
```

```
print(mean((train_pred$SATISFACCION - train_pred$pred_Satisfaccion)^2))
```

```
## [1] 43.87954
```

```

print('MS:' )

## [1] "MS:"
print(mean(train_pred$SATISFACCION)^2)

## [1] 811.6294
print('MSE test:' )

## [1] "MSE test:"
print(mean((test_pred$SATISFACCION - test_pred$pred_Satisfaccion)^2))

## [1] 44.97206
print('Para las 10 categorias: MSE:' )

## [1] "Para las 10 categorias: MSE:"
print(mean((cat_train_pred$SATISFACCION - cat_train_pred$pred_Satisfaccion)^2))

## [1] 1.621627
print('MS:' )

## [1] "MS:"
print(mean(cat_train_pred$SATISFACCION)^2)

## [1] 21.99207
print('MSE test:' )

## [1] "MSE test:"
print(mean((cat_test_pred$SATISFACCION - cat_test_pred$pred_Satisfaccion)^2))

## [1] 1.673258

```

Conclusiones

Recomendaciones

Referencias

- [1] Ramírez Pérez, Mauricio; Lee Maturana, Sau-Lyn (2012). Factores asociados a la satisfacción vital en adultos mayores de 60 años. Polis (Santiago), 11(33), 407–428. doi:10.4067/s0718-65682012000300020
- [2] Kutubaeva RZh (2019) Analysis of life satisfaction of the elderly population on the example of Sweden, Austria and Germany. Population and Economics 3(3): 102-116. <https://doi.org/10.3897/popecon.3.e47192>
- [3] Palmore, E., Luikart, C. (1972). Health and Social Factors Related to Life Satisfaction. Journal of Health and Social Behavior, 13(1), 68–80. doi: 10.2307/2136974
- [4] Naidu, Aditi (2009). Factors affecting patient satisfaction and healthcare quality. International Journal of Health Care Quality Assurance, 22(4), 366–381. doi:10.1108/09526860910964834
- [5] ROBLES-GARCIA, Monica et al. Variables relacionadas con la satisfaccion laboral: un estudio transversal a partir del modelo EFQM. Gac Sanit [online]. 2005, vol.19, n.2, pp.127-134. ISSN 0213-9111

- [6] Booth, Jaime; Ayers, Stephanie L.; and Marsiglia, Flavio F. (2012) "Perceived Neighborhood Safety and Psychological Distress: Exploring Protective Factors," *The Journal of Sociology & Social Welfare*: Vol. 39 : Iss. 4 , Article 8. Available at: <https://scholarworks.wmich.edu/jssw/vol39/iss4/>
- [7] <https://ourworldindata.org/happiness-and-life-satisfaction>