

## CAPÍTULO VIII

### Anexos

#### *Anexo 1. Depuración de base de datos*

2019			2021		
Observaciones	Descripción	Porcentaje	Observaciones	Descripción	Porcentaje
139.144	Total de datos	100,0%	101.881	Total de datos	100,0%
104.166	Datos depurados	74,9%	100.417	Datos depurados	98,6%
34.978	Datos faltantes	25,1%	1.464	Datos faltantes	1,4%

*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

#### *Anexo 2. Tratamiento variable condición de actividad laboral.*

Categorías Originales	Categorías Agrupadas
Empleo adecuado (1)	Empleo adecuado
Subempleo por insuficiencia de tiempo (2)	Subempleo
Subempleo por insuficiencia de ingresos (3)	
Otro empleo inadecuado (4)	Empleo inadecuado: categoría de referencia
Empleo no remunerado (5)	
Empleo no clasificado (6)	
Desempleo abierto (7)	Desempleo
Desempleo oculto (8)	
Población económicamente inactiva (9)	Población económicamente inactiva

*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

<b>Categorías Originales</b>	<b>Categorías Agrupadas</b>
Ninguna instrucción (1)	Ningún nivel de instrucción: categoría de referencia
Centro de alfabetización (2)	
Jardín de infantes (3)	Instrucción primaria
Primaria (4)	
Educación básica (5)	Instrucción secundaria
Secundaria (6)	
Educación media (7)	
Superior no universitario (8)	Instrucción universitaria o superior
Superior universitario (9)	
Posgrado (10)	

*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

#### *Anexo 4. Tratamiento variable Seguridad Social en Salud*

<b>Categorías Originales</b>	<b>Categorías Agrupadas</b>
IESS seguro general (1)	
IESS seguro general, Voluntario (2)	
IESS seguro campesino (3)	
ISSFA, ISSPOL (4)	
Seguro privado con hospitalización (5)	Recibe Seguridad Social en Salud (1)
Seguro privado sin hospitalización (6)	
AUS (7)	
Seguros municipales (8)	
Seguro ministerio salud pública (9)	
Ninguno (10)	No recibe Seguridad Social en Salud: categoría de referencia (0)

*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

#### *Anexo 5. Tratamiento variable tipo de vivienda.*

<b>Categorías Originales</b>	<b>Categorías Agrupadas</b>
Casa o villa (1)	Casa o departamento
Departamento (2)	
Cuartos de inquilinato (3)	Cuartos o mediaguas
Mediagua (4)	
Rancho o covacha (5)	
Choza (6)	Otro tipo de vivienda: categoría de referencia
Otro (7)	

*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

*Anexo 6. Tratamiento de variable nivel de ingresos per-cápita.*

<b>Categorías de la Variable</b>	<b>Referencia de la Categoría</b>
Ingresos per-cápita bajos (1)	Ingresos menores a 186,7 \$
Ingresos per-cápita medios (2)	Ingresos entre 186,8 \$ a 350\$
Ingresos per-cápita altos (3)	Ingresos mayores a 350 \$

*Fuente: ENEMDU (2019) y ENEMDU (2021)*  
*Elaboración: Autores*

*Anexo 7. Test de Breusch Pagan*

<b>Ho: Homocedasticidad</b>		
	<b>2019</b>	<b>2021</b>
Chi2	144,2	11873,0
Prob > chi2	0,0	0,0
df	38	38

*Fuente: ENEMDU (2019) y*  
*ENEMDU (2021)*  
*Elaboración: Autores*

*Anexo 8. Test Ramsey Reset*

<b>Ho: El modelo no tiene variables omitidas</b>	
<b>2019</b>	<b>2021</b>
F(3, 104.135) = 91,63	F(3, 100.386) = 76,00
Prob > F = 0,00	Prob > F = 0,00

*Fuente: ENEMDU (2019) y ENEMDU (2021)*  
*Elaboración: Autores*

Anexo 9. Factor de inflación de la varianza generalizado (GVIF), 2019 y 2021

Modelo		2019			2021		
Variables Demográficas		VIF	GVIF	1/VIF	VIF	GVIF	1/VIF
Área (categoría referencia: Rural)							
	Urbano	1,35	1,35	0,74	1,27	1,27	0,79
Etnia		1,15	1,16	0,86	1,12	1,13	0,88
Variables Individuales							
Nivel de instrucción		1,67	1,70	0,59	1,63	1,67	0,60
Estado Civil		1,13	1,14	0,88	1,13	1,14	0,88
Edad		1,83	1,83	0,55	1,86	1,86	0,54
Seguridad social en salud (categoría referencia: No posee ninguna seguridad social en salud)							
	Posee alguna seguridad social en salud	1,37	1,37	0,73	1,46	1,46	0,68
Variables Económicas							
Nivel socioeconómico (categoría referencia: Nivel socioeconómico Bajo)							
	Nivel socioeconómico Alto	1,60	1,60	0,63	1,42	1,42	0,70
	Nivel socioeconómico Medio	1,54	1,54	0,65	1,38	1,38	0,72
Pobreza (categoría referencia: No se encuentra en situación de pobreza)							
	Se encuentra en situación de pobreza	1,74	1,74	0,57	1,79	1,79	0,56
Extrema pobreza (categoría referencia: No se encuentra en situación de extrema pobreza)							
	Se encuentra en situación de extrema pobreza	1,42	1,42	0,70	1,44	1,44	0,69
Condición de actividad (categoría referencia: empleo inadecuado)							
	Subempleo	1,38	1,38	0,72	1,49	1,49	0,67
	Empleo adecuado	2,28	2,28	0,44	2,23	2,23	0,45
	Desempleo	1,08	1,08	0,93	1,11	1,11	0,90
	Población económicamente inactiva	1,77	1,77	0,56	1,76	1,76	0,57
Variables de Ingresos externos							
Remesas (categoría referencia: no recibe dinero del extranjero)							
	Recibe dinero del extranjero	1,01	1,01	0,99	1,01	1,01	0,99
Bono discapacidad (categoría referencia: no recibe el bono de discapacidad)							
	Recibe el bono de discapacidad	1,01	1,01	0,99	1,01	1,01	0,99
Bono desarrollo humano (categoría referencia: no recibe el bono de desarrollo humano)							
	Recibe el bono de desarrollo humano	1,20	1,20	0,83	1,18	1,18	0,85
Variables de Hogar							
Miembros del hogar (categoría referencia: tamaño familiar del hogar grande)							
	Tamaño familiar del hogar pequeño	3,03	3,03	0,33	2,95	2,95	0,34
	Tamaño familiar del hogar mediano	2,67	2,67	0,37	2,62	2,62	0,38
Tipo de vivienda (categoría referencia: otro tipo de vivienda)							
	Casa o departamento	2,47	2,47	0,40	2,22	2,22	0,45
	Cuartos o mediagua	2,37	2,37	0,42	2,13	2,13	0,47
Mean VIF - GVIF		1,68	1,68		1,64	1,64	

Fuente: ENEMDU (2019) y ENEMDU (2021)

Elaboración: Autores

Anexo 10. Valores influyentes, Distancias de Cook

Di > 1				
2019		2021		
VERDADERO	FALSO	VERDADERO	FALSO	
0	104.166	0	100.417	

Fuente: ENEMDU (2019) y ENEMDU (2021)

Elaboración: Autores

*Anexo 11. Residuos Chi-cuadrado de Pearson*

Ho: $b_1=b_2=b_3=\dots=b_n=0$		
	2019	2021
Chi2	1018,1	1018,1
Prob > chi2	0,0	0,0

*Fuente: ENEMDU (2019) y ENEMDU (2021)  
Elaboración: Autores*

*Anexo 12. Test de Wald*

Ho: $b_j = 0$		
	2019	2021
Chi2	30.204,1	28.464,2
Prob > chi2	0,0	0,0

*Fuente: ENEMDU (2019) y ENEMDU (2021)  
Elaboración: Autores*

*Anexo 13. Sensibilidad, especificidad y correctamente predichos.*

	2019	2021
Sensibilidad Pr ( 1   DR )	77,15%	76,15%
Especificidad Pr ( 0   DR )	88,55%	87,84%
Falsos positivos Pr ( DR   1 )	74,28%	74,88%
Falsos negativos Pr ( DR   0 )	90,04%	88,55%
Correctamente clasificado	85,13%	84,07%

*Fuente: ENEMDU (2019) y ENEMDU (2021)  
Elaboración: Autores*

*Anexo 14. Área bajo la curva ROC*

<b>Logistic model for mujer</b>	<b>2019</b>	<b>2021</b>
<b>Number of observations</b>	104.166	100.417
<b>Area under ROC curve</b>	0,9009	0,8944

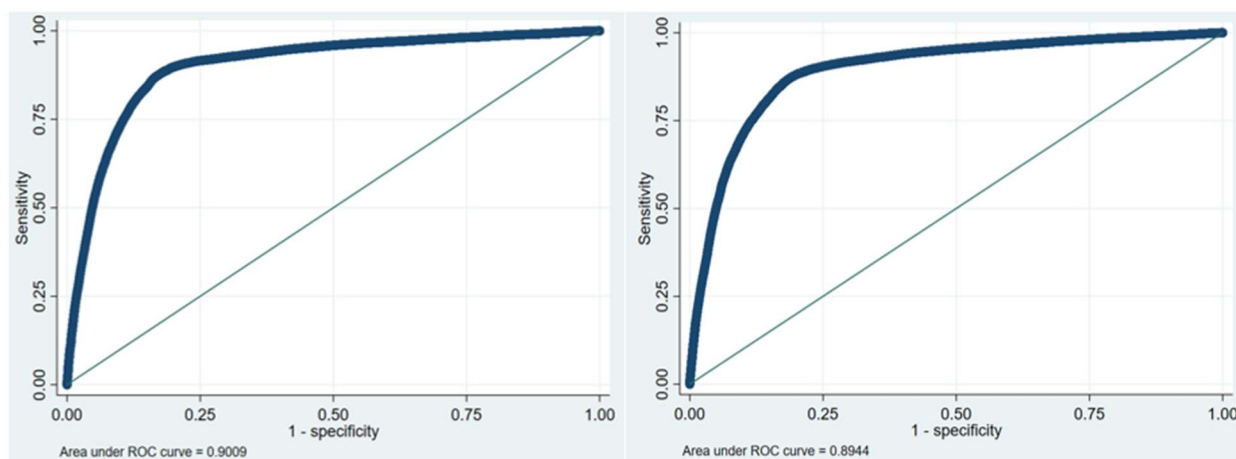
*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

*Anexo 15. Curvas ROC*

2019

2021



*Fuente: ENEMDU (2019) y ENEMDU (2021)*

*Elaboración: Autores*

```

## CÓDIGO EN R
library(dplyr)
library(outliers)
library(haven)
library(readxl)
library(openxlsx)
library(DataExplorer)
library(PerformanceAnalytics)
library(stringi)
library(tidyverse)
library(e1071)
library(pastecs)
library(Hmisc)
# 2019
df1<-read_sav("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/01_ENEMDU_Base datos_anual_2019_spss/BDDenemdu_personas_2019_anual.sav")
df<-select(df1, "p04", "area", "estrato", "p02", "p03", "p06", "p10a", "p15",
          "pobreza", "epobreza", "upm", "fexp", "conduct", "id_vivienda", "id_hogar", "id_persona",
          "p05a", "p74a", "p75", "ingpc", "p72a", "p77")
df$nivelsocio<- stri_sub(df$estrato,-1)
df<-filter(df, df$nivelsocio!=9)
df<-filter(df, df$p04==1)
col<- c("p04", "estrato", "p02", "conduct", "nivelsocio", "id_hogar", "id_vivienda",
        "id_persona", "p05a", "p74a", "p75")
df[col] <- lapply(df[col], factor)
df<-na.omit(df)
df<-mutate(df, pension_jubilacion= ifelse(p72a==2,1,0))
df<-mutate(df, bono_discapacidad= ifelse(p77==2,1,0))
df<-mutate(df, genero_mujer= ifelse(p02==2,1,0))
df<-mutate(df, jh_mujer= ifelse(genero_mujer==1 & p04==1,1,0))
df<-mutate(df, ingresos_bajos= ifelse(ingpc< 187,1,0))
df<-mutate(df, ingresos_medios= ifelse(ingpc>=187 & ingpc<350,1,0))
df<-mutate(df, ingresos_altos= ifelse(ingpc>=350,1,0))
df<-mutate(df, nivelsocio_bajo= ifelse(nivelsocio==3,1,0))
df<-mutate(df, nivelsocio_medio= ifelse(nivelsocio==2,1,0))
df<-mutate(df, nivelsocio_alto= ifelse(nivelsocio==1,1,0))
df<-mutate(df, subempleo= ifelse(conduct%in% c(2,3),1,0))
df<-mutate(df, empleo_inadecuado= ifelse(conduct %in% c(4,5,6),1,0))
df<-mutate(df, empleo_adequado= ifelse(conduct==1,1,0))
df<-mutate(df, desempleo= ifelse(conduct%in% c(7,8),1,0))
df<-mutate(df, poblacion_economicamente_inactiva= ifelse(conduct%in% c(9,0),1,0))
df<-mutate(df, indigena= ifelse(p15==1,1,0))
df<-mutate(df, Negro= ifelse(p15%in% c(2,3),1,0))
df<-mutate(df, Mulato= ifelse(p15%in% c(4,8),1,0))

```

```

df<-mutate(df, Montubio= ifelse(p15==5,1,0))
df<-mutate(df, Mestizo= ifelse(p15==6,1,0))
df<-mutate(df, Blanco= ifelse(p15==7,1,0))
df<-mutate(df, ninguna_instruccion= ifelse(p10a==1,1,0))
df<-mutate(df, educacion_primaria= ifelse(p10a %in% c(2,3,4,5),1,0))
df<-mutate(df, educacion_secundaria= ifelse(p10a %in% c(6,7),1,0))
df<-mutate(df, educacion_universitaria= ifelse(p10a %in% c(8,9,10),1,0))
df<-mutate(df, casado= ifelse(p06==1,1,0))
df<-mutate(df, separado= ifelse(p06==2,1,0))
df<-mutate(df, divorciado= ifelse(p06==3,1,0))
df<-mutate(df, viudo= ifelse(p06==4,1,0))
df<-mutate(df, union_libre= ifelse(p06==5,1,0))
df<-mutate(df, soltero= ifelse(p06==6,1,0))
df<-mutate(df, area_urbana= ifelse(area==1,1,0))
df<-mutate(df, seguro_social= ifelse(p05a==10,1,0))
df<-mutate(df, dinero_del_extranjero= ifelse(p74a==2,1,0))
df<-mutate(df, bono_desarrollo_humano= ifelse(p75==2,1,0))
df<-mutate(df, pobreza2= ifelse(pobreza==1,0,1))
df<-mutate(df, epobreza2= ifelse(epobreza==1,0,1))
vivienda<-read_sav("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/01_ENEMDU_Base datos_anual_2019_spss/BDDenemdu_vivienda_2019_anual.sav")
names(vivienda) #cuartos y dormitorios: "vi06", "vi07"
vivienda<-select(vivienda, "vi14", "id_hogar", "vi01", "vi02", "vi12")
vivienda<-mutate(vivienda, vivienda_propia= ifelse(vi14 %in% c(3,4,5),1,0))
vivienda<-mutate(vivienda, carretera_pavimento= ifelse(vi01==1,1,0))
vivienda<-mutate(vivienda, casa_departamento= ifelse(vi02%in% c(1,2),1,0))
vivienda<-mutate(vivienda, cuarto_mediagua= ifelse(vi02%in% c(3,4),1,0))
vivienda<-mutate(vivienda, otra_vivienda= ifelse(vi02%in% c(5,6,7),1,0))
vivienda<-mutate(vivienda, alumbrado_publico= ifelse(vi12==1,1,0))
df<-merge(df,vivienda, by="id_hogar")
num_unique <- . %>% unique %>% length
num_unique(df1$id_hogar)
df2<- aggregate(df1[,c("id_hogar")],by = list(id_hogar = df1$id_hogar), FUN =length)
colnames(df2)[2]<-"miembros_hogar"
df2$miembros_hogar <-as.factor(df2$miembros_hogar)
df3<-merge(df,df2, by="id_hogar")
df3$miembros_hogar<-as.double(df3$miembros_hogar)
df3<-mutate(df3, tamanoH_bajo= ifelse(miembros_hogar<=2,1,0))
df3<-mutate(df3, tamanoH_medio= ifelse(miembros_hogar>2 & miembros_hogar<=5,1,0))
df3<-mutate(df3, tamanoH_alto= ifelse(miembros_hogar>5,1,0)) #categoria de referencia
df3<-drop_columns(df3, "estrato")
df3<-drop_columns(df3, "p02")
df3<-drop_columns(df3, "conduct")
df3<-drop_columns(df3, "nivelesocio")
df3<-drop_columns(df3, "vi14")
df3<-drop_columns(df3, "p74a")

```



```

df3<-drop_columns(df3, "p75")
df3<-drop_columns(df3, "p05a")
df3<-drop_columns(df3, "upm")
df3<-drop_columns(df3, "id_hogar")
df3<-drop_columns(df3, "id_vivienda")
df3<-drop_columns(df3, "id_persona")
df3<-drop_columns(df3, "fexp")
df3<-drop_columns(df3, "p72a")
df3<-drop_columns(df3, "p77")
df3<-drop_columns(df3, "genero_mujer")
df3<-drop_columns(df3, "vi01")
df3<-drop_columns(df3, "vi02")
df3<-drop_columns(df3, "vi12")
df3<-drop_columns(df3, "pobreza")
df3<-drop_columns(df3, "epobreza")
writexl::write_xlsx(df3,"C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/Base final jefatura de hogar 2019.xlsx" )
# 2021
df1<-read_sav("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de datos/01
ENEMDU_Base datos_anual_2021_spss/BDDenemdu_personas_2021_anual.sav")
df<-select(df1, "p04", "area", "estrato", "p02", "p03", "p06", "p10a", "p15",
"pobreza", "epobreza", "upm", "fexp", "conduct", "id_vivienda", "id_hogar", "id_persona", "p05a",
"p74a", "p75", "ingpc", "p72a", "p77")
df$nivelsocio<- stri_sub(df$estrato,-1)
df<-filter(df, df$nivelsocio!=9)
df<-filter(df, df$p04==1)
col<- c("p04", "estrato", "p02", "upm", "conduct", "nivelsocio", "id_hogar", "id_vivienda",
"id_persona", "p05a", "p74a", "p75")
df[col] <- lapply(df[col], factor)
df<-na.omit(df)
df<-mutate(df, pension_jubilacion= ifelse(p72a==2,1,0))
df<-mutate(df, bono_discapacidad= ifelse(p77==2,1,0))
df<-mutate(df, genero_mujer= ifelse(p02==2,1,0))
df<-mutate(df, jh_mujer= ifelse(genero_mujer==1 & p04==1,1,0))
df<-mutate(df, ingresos_bajos= ifelse(ingpc< 187,1,0))
df<-mutate(df, ingresos_medios= ifelse(ingpc>=187 & ingpc<350,1,0))
df<-mutate(df, ingresos_altos= ifelse(ingpc>=350,1,0))
df<-mutate(df, nivelsocio_bajo= ifelse(nivelsocio==3,1,0))
df<-mutate(df, nivelsocio_medio= ifelse(nivelsocio==2,1,0))
df<-mutate(df, nivelsocio_alto= ifelse(nivelsocio==1,1,0))
df<-mutate(df, subempleo= ifelse(conduct%in% c(2,3),1,0))
df<-mutate(df, empleo_inadecuado= ifelse(conduct %in% c(4,5,6),1,0))
df<-mutate(df, empleo_adequado= ifelse(conduct==1,1,0))
df<-mutate(df, desempleo= ifelse(conduct%in% c(7,8),1,0))
df<-mutate(df, poblacion_economicamente_inactiva= ifelse(conduct%in% c(9,0),1,0))
df<-mutate(df, indigena= ifelse(p15==1,1,0))

```

```

df<-mutate(df, Negro= ifelse(p15%in% c(2,3),1,0))
df<-mutate(df, Mulato= ifelse(p15%in% c(4,8),1,0))
df<-mutate(df, Montubio= ifelse(p15==5,1,0))
df<-mutate(df, Mestizo= ifelse(p15==6,1,0))
df<-mutate(df, Blanco= ifelse(p15==7,1,0))
df<-mutate(df, ninguna_intrucion= ifelse(p10a==1,1,0))
df<-mutate(df, educacion_primaria= ifelse(p10a %in% c(2,3,4,5),1,0))
df<-mutate(df, educacion_secundaria= ifelse(p10a %in% c(6,7),1,0))
df<-mutate(df, educacion_universitaria= ifelse(p10a %in% c(8,9,10),1,0))
df<-mutate(df, casado= ifelse(p06==1,1,0))
df<-mutate(df, separado= ifelse(p06==2,1,0))
df<-mutate(df, divorciado= ifelse(p06==3,1,0))
df<-mutate(df, viudo= ifelse(p06==4,1,0))
df<-mutate(df, union_libre= ifelse(p06==5,1,0))
df<-mutate(df, soltero= ifelse(p06==6,1,0))
df<-mutate(df, area_urbana= ifelse(area==1,1,0))
df<-mutate(df, seguro_social= ifelse(p05a==10,1,0))
df<-mutate(df, dinero_del_extranjero= ifelse(p74a==2,1,0))
df<-mutate(df, bono_desarrollo_humano= ifelse(p75==2,1,0))
df<-mutate(df, pobreza2= ifelse(pobreza==1,0,1))
df<-mutate(df, epobreza2= ifelse(epobreza==1,0,1))
vivienda<-read_sav("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de datos/01
ENEMDU_Base datos_anual_2021_spss/BDDenemdu_vivienda_2021_anual.sav")
names(vivienda) #cuartos y dormitorios: "vi06", "vi07"
vivienda<-select(vivienda, "vi14", "id_hogar", "vi01", "vi02", "vi12")
vivienda<-mutate(vivienda, vivienda_propia= ifelse(vi14 %in% c(3,4,5),1,0))
vivienda<-mutate(vivienda, carretera_pavimento= ifelse(vi01==1,1,0))
vivienda<-mutate(vivienda, casa_departamento= ifelse(vi02%in% c(1,2),1,0))
vivienda<-mutate(vivienda, cuarto_mediaagua= ifelse(vi02%in% c(3,4),1,0))
vivienda<-mutate(vivienda, otra_vivienda= ifelse(vi02%in% c(5,6,7),1,0))
vivienda<-mutate(vivienda, alumbrado_publico= ifelse(vi12==1,1,0))
df<-merge(df,vivienda, by="id_hogar")
num_unique <- . %>% unique %>% length
num_unique(df1$id_hogar)
df2<- aggregate(df1[,c("id_hogar")],by = list(id_hogar = df1$id_hogar), FUN =length)
colnames(df2)[2]<-"miembros_hogar"
df2$miembros_hogar <-as.factor(df2$miembros_hogar)
df3<-merge(df,df2, by="id_hogar")
df3$miembros_hogar<-as.double(df3$miembros_hogar)
df3<-mutate(df3, tamanoH_bajo= ifelse(miembros_hogar<=2,1,0))
df3<-mutate(df3, tamanoH_medio= ifelse(miembros_hogar>2 & miembros_hogar<=5,1,0))
df3<-mutate(df3, tamanoH_alto= ifelse(miembros_hogar>5,1,0)) #categoria de referencia
df3<-drop_columns(df3, "estrato")
df3<-drop_columns(df3, "p02")
df3<-drop_columns(df3, "conduct")
df3<-drop_columns(df3, "nivelsocio")

```

```

df3<-drop_columns(df3, "vi14")
df3<-drop_columns(df3, "p74a")
df3<-drop_columns(df3, "p75")
df3<-drop_columns(df3, "p05a")
df3<-drop_columns(df3, "upm")
df3<-drop_columns(df3, "id_hogar")
df3<-drop_columns(df3, "id_vivienda")
df3<-drop_columns(df3, "id_persona")
df3<-drop_columns(df3, "fexp")
df3<-drop_columns(df3, "p72a")
df3<-drop_columns(df3, "p77")
df3<-drop_columns(df3, "genero_mujer")
df3<-drop_columns(df3, "vi01")
df3<-drop_columns(df3, "vi02")
df3<-drop_columns(df3, "vi12")
df3<-drop_columns(df3, "pobreza")
df3<-drop_columns(df3, "epobreza")
writexl::write_xlsx(df3,"C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/Base final jefatura de hogar 2021.xlsx" )
#2019
df <- read.xlsx("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de datos/Base
final jefatura de hogar 2019.xlsx")
#2021
df1 <- read.xlsx("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de datos/Base
final jefatura de hogar 2021.xlsx")
df<-mutate(df, año= 2019)
df1<-mutate(df1, año= 2021)
col<- c("jh_mujer", "pobreza2","area", "epobreza2", "pension_jubilacion", "bono_discapacidad",
"ingresos_bajos", "ingresos_medios", "ingresos_altos", "nivelesocio_bajo", "nivelesocio_medio", "niv
elsocio_alto", "subempleo", "empleo_inadecuado", "empleo_adequado", "desempleo", "poblacion_e
conomicamente_inactiva", "indigena", "Negro", "Mulato", "Montubio", "Mestizo", "Blanco", "ningu
na_intrusion", "educacion_primaria", "educacion_secundaria", "educacion_universitaria",
"casado", "separado", "divorciado", "viudo", "union_libre", "soltero", "area_urbana", "seguro_social",
"dinero_del_extranjero", "bono_desarrollo_humano", "vivienda_propia", "carretera_pavimento",
"casa_departamento", "cuarto_mediagua", "otra_vivienda", "alumbrado_publico",
"miembros_hogar", "tamanoH_bajo", "tamanoH_medio", "tamanoH_alto")
df[col] <- lapply(df[col], as.double)
df1[col] <- lapply(df1[col], as.double)
df2<-full_join(df,df1)
df2 <- df2 %>% mutate(etniagraf = case_when( Negro==1 ~ "Afroecuatoriano", indigena==1
~ "Indígena", Mulato==1 ~ "Mulato", Montubio==1 ~ "Montubio", Mestizo== 1 ~ "Mestizo",
Blanco== 1 ~ "Blanco"))
df2 <- df2 %>% mutate(areagraf = case_when( area==1 ~ "Urbano", area==0 ~ "Rural"))
df2 <- df2 %>% mutate(jefehogargraf = case_when( jh_mujer==1 ~ "Jefe de hogar mujer",
jh_mujer==0 ~ "Jefe de hogar hombre"))
df2 <- df2 %>% mutate(pobrezagraf = case_when( pobreza2==0 ~ "Pobreza",

```

```

pobreza2 ==1 ~ "No pobre"))
df2 <- df2 %>% mutate(pobrezaextremagraf = case_when( epobreza2==0 ~ "Pobreza extrema",
  epobreza2==1 ~ "No pobreza extrema"))
df2 <- df2 %>% mutate(pension_jubilaciongraf = case_when( pension_jubilacion==1 ~ "No
recibe pensión o jubilación", pension_jubilacion==0 ~ "Si recibe pensión o jubilación"))
df2 <- df2 %>% mutate(bono_discapacidadgraf = case_when( bono_discapacidad==1 ~ "No
recibe bono por discapacidad", bono_discapacidad==0 ~ "Si recibe bono por discapacidad"))
df2 <- df2 %>% mutate(ingresosgraf = case_when( ingresos_bajos==1 ~ "Ingresos bajos",
ingresos_medios==1 ~ "Ingresos medios", ingresos_altos==1 ~ "Ingresos altos"))
df2 <- df2 %>% mutate(nivelsocioeconomicograf = case_when( nivelsocio_bajo==1 ~ "Nivel
socioeconómico bajo", nivelsocio_medio==1 ~ "Nivel socioeconómico medio",
  nivelsocio_alto==1 ~ "Nivel socioeconómico alto"))
df2 <- df2 %>% mutate(situacion_laboralgraf = case_when( subempleo==1 ~ "Subempleo",
empleo_inadecuado==1 ~ "Empleo inadecuado", empleo_adequado==1 ~ "Empleo adecuado",
desempleo==1 ~ "Desempleo", poblacion_economicamente_inactiva== 1 ~ "Población
economicamente inactiva"))
df2 <- df2 %>% mutate(intrucciongraf = case_when( ninguna_intrucion==1 ~ "Ninguna
instrucción", educacion_primaria==1 ~ "Educación primaria", educacion_secundaria==1 ~
"Educación secundaria", educacion_universitaria==1 ~ "Educacion universitaria"))
df2 <- df2 %>% mutate(estado_civilgraf = case_when( casado==1 ~ "Casado/a", separado==1
~ "Separado/a", divorciado==1 ~ "Divorciado/a", viudo==1 ~ "Viudo/a", union_libre== 1 ~
"Union libre", soltero== 1 ~ "Soltero/a"))
df2 <- df2 %>% mutate(seguro_socialgraf = case_when( seguro_social==0 ~ "Otro caso",
  seguro_social==1 ~ "No tiene seguro social"))
df2 <- df2 %>% mutate(dinero_extranjerograf = case_when( dinero_del_extranjero==1 ~ "No
recibe dinero del extranjero", dinero_del_extranjero==0 ~ "Si recibe dinero del extranjero"))
df2 <- df2 %>% mutate(bono_desarrollograf = case_when( bono_desarrollo_humano==1 ~
"No recibe el bono de desarrollo humano", bono_desarrollo_humano==0 ~ "Si recibe el bono de
desarrollo humano"))
df2 <- df2 %>% mutate(vivienda_propiagraf = case_when( vivienda_propia==1 ~ "Otro caso",
  vivienda_propia==0 ~ "Tiene vivienda propia"))
df2 <- df2 %>% mutate(carretera_pavimentograf = case_when( carretera_pavimento==1 ~ "Via
de acceso: carretera o calle pavimentada", carretera_pavimento==0 ~ "Otro caso"))
df2 <- df2 %>% mutate(tipo_viviendagraf = case_when( casa_departamento==1 ~ "Tipo de
vivienda: casa o departamento", cuarto_mediagua==1 ~ "Tipo de vivienda: cuarto o mediagua",
  otra_vivienda==1 ~ "Otro tipo de vivienda"))
df2 <- df2 %>% mutate(alumbradograf = case_when( alumbrado_publico==1 ~ "Posee
alumbrado público", alumbrado_publico==0 ~ "Otro caso"))
df2 <- df2 %>% mutate(tamano_hogargraf = case_when( tamanoH_bajo==1 ~ "Tamaño del
hogar: pequeño", tamanoH_medio==1 ~ "Tamaño del hogar: mediano",tamanoH_alto==1 ~
"Tamaño del hogar: grande"))
df2 <- df2 %>% mutate(edadgraf = case_when( p03>=13 & p03<=29 ~ "De 13 a 29 años de
edad", p03>29 & p03<=59 ~ "De 30 a 59 años de edad", p03>59 ~ "De 60 años o más"))
df2$p03<-as.double(df2$p03)
df3<-dplyr::select(df2,"edadgraf","tamano_hogargraf","alumbradograf","tipo_viviendagraf",
"carretera_pavimentograf","vivienda_propiagraf","bono_desarrollograf",

```

```

"dinero_extranjero", "seguro_social", "estado_civil", "instruccion", "situacion_labor",
"nivel_socioeconomico", "ingresos", "bono_discapacidad", "pension_jubilacion",
"pobreza_extrema", "pobreza", "jefe_hogar", "area", "etnia", "año", "p03")
write_xlsx(df3, "C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/Base para graficas 2019 y 2021.xlsx")
df <- read_xlsx("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de datos/Base
para graficas 2019 y 2021.xlsx", sheet = "Sheet1")
df <- mutate(df, edad_cuadrado = p03^2)
df1 <- split(df, df$año)
glimpse(df1$`2019`)
describe(df1$`2019`)
describe(df1$`2021`)
stat.desc(df1$`2019`)
stat.desc(df1$`2021`)
library(lmtest)
library(DescTools)
library("DMwR2")
library(openxlsx)
library(gmodels)
library(modeest)
library(GGally)
require(MPV)
library("MVN")
library("performance")
# 2019
df <- read_xlsx("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/Base final jefatura de hogar 2019.xlsx")
df <- DataExplorer::drop_columns(df, "area")
df <- DataExplorer::drop_columns(df, "p04")
df <- DataExplorer::drop_columns(df, "p06")
df <- DataExplorer::drop_columns(df, "p10a")
df <- DataExplorer::drop_columns(df, "p20")
df <- DataExplorer::drop_columns(df, "ingpc")
df <- mutate(df, edad2 = p03^2)
df <- DataExplorer::drop_columns(df, "menor_26")
df <- DataExplorer::drop_columns(df, "de26_a40")
df <- DataExplorer::drop_columns(df, "de41_a55")
df <- DataExplorer::drop_columns(df, "mayor_55")
df <- DataExplorer::drop_columns(df, "NoTrabajó_semanaP")
df <- DataExplorer::drop_columns(df, "este_lugar")
df <- DataExplorer::drop_columns(df, "p15")
df <- DataExplorer::drop_columns(df, "epobreza")
df <- DataExplorer::drop_columns(df, "miembros_hogar")
df <- DataExplorer::drop_columns(df, "tamanoH_alto")
df <- DataExplorer::drop_columns(df, "ingresos_bajos")
df <- DataExplorer::drop_columns(df, "nivel_socio_bajo")

```

```

df<-DataExplorer::drop_columns(df,"empleo_adequado")
df<-DataExplorer::drop_columns(df,"divorciado")
df<-DataExplorer::drop_columns(df,"ninguna_intruccion")
df<-DataExplorer::drop_columns(df,"Negro")
df<-DataExplorer::drop_columns(df,"otra_vivienda")
df$jh_mujer <- as.factor(df$jh_mujer)
logitrose <- glm(jh_mujer~., data = df, family = binomial(link = "logit"))
bptest(logitrose) # H0: homocedasticidad
pearprob <-sum ( residuals ( logitrose , type ="pearson" ) ^2)
pvaluepers<-1 - pchisq ( pearprob , logitrose$df.null - logitrose$df.residual )
pvaluepers
# Multicolinealidad de los predictores
car::vif (logitrose)
# Distancias de cook
cook <-cooks.distance (logitrose)
# valores influyentes
influyentes <-cook >1
table (influyentes)
df<-readxl::read_xlsx("C:/Users/Sergio P. Medina/OneDrive/Escritorio/TESIS 3/base de
datos/Base final jefatura de hogar 2021.xlsx")
df<-DataExplorer::drop_columns(df,"area")
df<-DataExplorer::drop_columns(df,"p04")
df<-DataExplorer::drop_columns(df,"p06")
df<-DataExplorer::drop_columns(df,"p10a")
df<-DataExplorer::drop_columns(df,"ingpc")
df<-mutate(df, edad2=p03^2)
df<-DataExplorer::drop_columns(df,"NoTrabajó_semanaP")
df<-DataExplorer::drop_columns(df,"este_lugar")
df<-DataExplorer::drop_columns(df,"p15")
df<-DataExplorer::drop_columns(df,"epobreza")
df<-DataExplorer::drop_columns(df,"miembros_hogar")
df<-DataExplorer::drop_columns(df,"tamanoH_alto")
df<-DataExplorer::drop_columns(df,"ingresos_bajos")
df<-DataExplorer::drop_columns(df,"nivelesocio_bajo")
df<-DataExplorer::drop_columns(df,"empleo_adequado")
df<-DataExplorer::drop_columns(df,"divorciado")
df<-DataExplorer::drop_columns(df,"ninguna_intruccion")
df<-DataExplorer::drop_columns(df,"Negro")
df<-DataExplorer::drop_columns(df,"otra_vivienda")
df$jh_mujer <- as.factor(df$jh_mujer)
prop.table(table(df$jh_mujer))
logitrose <- glm(jh_mujer~., data = df, family = binomial(link = "logit"))
# homocedasticidad
bptest(logitrose) # H0: homocedasticidad
# Multicolinealidad de los predictores
car::vif (logitrose)

```

```

# Distancias de cook
cook <-cooks.distance (logitrose)
# valores influyentes
influyentes <-cook >1
table (influyentes)
## CÓDIGO EN STATA
//modelo 2019
import excel "C:\Users\Sergio P. Medina\OneDrive\Escritorio\TESIS 3\base de datos\Base final
jefatura de hogar 2019.xlsx", sheet("Sheet1") firstrow
gen edad2=p03^2
**pruebas de multicolinealidad
collin jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco p06 p10a
area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano vivienda_propia
carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico tamanoH_medio
tamanoH_bajo p03
reg jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv p15 p06 p10a area_urbana seguro_social
dinero_del_extranjero bono_desarrollo_humano vivienda_propia carretera_pavimento
casa_departamento cuarto_mediagua alumbrado_publico tamanoH_medio tamanoH_bajo p03
vif
** variable omitida
estat ovtest
logistic jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, robust
logit jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, robust
mfx
predict jefe_prob
sum jefe_prob jh_mujer
** correctamente predichos y matriz de confusion
estat class
lsens
**Test de wald

```

```

test pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios ingresos_altos
nivelesocio_medio nivelesocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03
**valor de pearson (valor del estadistico Chi Cuadrado de Pearson)
binreg jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelesocio_medio nivelesocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, or
display "pvalueChi2=" chiprob(1, 1018.074413) // p.value chi2=0 buen modelo.
logit jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelesocio_medio nivelesocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, robust
** curva roc
lroc
clear all
//modelo 2021
import excel "C:\Users\Sergio P. Medina\OneDrive\Escritorio\TESIS 3\base de datos\Base final
jefatura de hogar 2021.xlsx", sheet("Sheet1") firstrow
gen edad2=p03^2
**pruebas de multicolinealidad
collin jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelesocio_medio nivelesocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco p06 p10a
area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano vivienda_propia
carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico tamanoH_medio
tamanoH_bajo p03
reg jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelesocio_medio nivelesocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv p15 p06 p10a area_urbana seguro_social
dinero_del_extranjero bono_desarrollo_humano vivienda_propia carretera_pavimento
casa_departamento cuarto_mediagua alumbrado_publico tamanoH_medio tamanoH_bajo p03
vif
** variable omitida
estat ovtest

```



```

logistic jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, robust
logit jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, robust
mfx
predict jefe_prob
sum jefe_prob jh_mujer
** correctamente predichos y matriz de confusion
estat class
lsens
**Test de wald
test pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios ingresos_altos
nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03
**valor de pearson (valor del estadistico Chi Cuadrado de Pearson)
binreg jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, or
display "pvalueChi2=" chiprob(1, 1018.074413) // p.value chi2=0 buen modelo.
logit jh_mujer pobreza2 pension_jubilacion bono_discapacidad epobreza2 ingresos_medios
ingresos_altos nivelsocio_medio nivelsocio_alto subempleo empleo_inadecuado desempleo
poblacion_economicamente_inactiv indigena Mulato Montubio Mestizo Blanco
educacion_primaria educacion_secundaria educacion_universitaria casado separado viudo
union_libre soltero area_urbana seguro_social dinero_del_extranjero bono_desarrollo_humano
vivienda_propia carretera_pavimento casa_departamento cuarto_mediagua alumbrado_publico
tamanoH_medio tamanoH_bajo edad2 p03, robust
** curva roc
lroc

```