

```

#~~~~~
#####Validaciones CHC#####
#~~~~~

rm(list=ls())

#Definir rutas
setwd("d:/Users/Usuario/Documents/Jurgen/SDP/CHC")

#Cargar paquetes
library(sf)
library(dplyr)
library(openxlsx)
library(readxl)
library(stringr)

wb <- createWorkbook()

# 1. Cargar la base de datos -----
-----

#shapename <- read_sf('BD CHC 04 de
septiembre/Habitante_de_Calle_Censo_20240812.shp')
shapename <- st_read('Input/20241205_base_CHC.gdb')

Base_prueba_CHC<-shapename
Base_prueba_CHC<-as.data.frame(Base_prueba_CHC)

###Eliminar variables operativo

Base_prueba_CHC<-subset(Base_prueba_CHC,select=-
c(inicio_encuesta_hcalle,TIP_FOR,COD_CUAD,COD_CUAD_MAPA,nom_encuestador,
cod_encuestador,nom_coordinador,cod_coordinador,nom_supervisor,
cod_supervisor,nom_lider_cuad,cod_lider_cuad,cod_apoyo_infor,
nom_apoyo_infor,SHAPE,P39_4,P39_5,P39_6,P39_7,hora_final_entrevista_o,act
iva_anulada,
fin_encuesta1,fin_encuesta_hcalle,CreationDate,Creator,EditDate,Editor,ho
ra_inicio_entrevista_o,
fecha_entrevista_o,autorizacion_tratam_datos1,P7S1)) #Incluir esta en el
listado

#Eliminar variables sensibles

Base_prueba_CHC<-subset(Base_prueba_CHC,select=-
c(direccion_concatenada,observacion_direccion,C_XY,primer_nombre_8_1c,

```

```
segundo_nombre_8_2c,primer_apellido_8_3c,segundo_apellido_8_4c,apodo_8_5c  
,
```

```
P9_1,P12S1,UBIC_2,barrio_sector_catastral_ciud,    upl_escogida,  
    upz_escogida1,    LOCALIDAD_MAPA,    localidad_escogida1_codigo))  
#incluir
```

```
#Variables que se deben ajustar para su utilización
```

```
Base_prueba_CHC<-subset(Base_prueba_CHC,select=-c(P5S1, P11S2,  
                                                    no_sabe_edad,  
            P17_1,      P17_2_2,    P17_3_2,  
                                                    P19S2, P20S1,  
            P21S1,      P22S7_1,    P23S10_1,  
                                                    P26S1, P28_1S11_1,  
            P30S9_1,    P30_1A,      P32_1,  
                                                    P33S9_1,    P38S5_1_1,  
            P38S5_1_2, P42_1,  
                                                    P45_1, P50))
```

```
#Base_prueba_CHC<-Base_prueba_CHC[Base_prueba_CHC$globalid=="{E742EF50-  
9553-4214-8B46-3AC12B2C6E62}",,]
```

```
##Recodificación de variables
```

```
Base_prueba_CHC$visita_municipio<-ifelse(!is.na(Base_prueba_CHC$P24S1D) |  
!is.na(Base_prueba_CHC$P24S2A) | !is.na(Base_prueba_CHC$P24S3A),1,2)  
Base_prueba_CHC$visita_municipio<-  
ifelse(Base_prueba_CHC$CTL1==1,NA,Base_prueba_CHC$visita_municipio)  
Base_prueba_CHC$visita_municipio<-  
ifelse(is.na(Base_prueba_CHC$ninguna_ciudad_f24) &  
Base_prueba_CHC$visita_municipio==2 &  
is.na(Base_prueba_CHC$P24S4),NA,Base_prueba_CHC$visita_municipio)
```

```
Base_prueba_CHC$visita_pais<-ifelse(!is.na(Base_prueba_CHC$P24S4_1),1,2)  
Base_prueba_CHC$visita_pais<-  
ifelse(Base_prueba_CHC$CTL1==1,NA,Base_prueba_CHC$visita_pais)  
Base_prueba_CHC$visita_pais<-  
ifelse(is.na(Base_prueba_CHC$ninguna_ciudad_f24) &  
Base_prueba_CHC$visita_pais==2 &  
is.na(Base_prueba_CHC$P24S1A),NA,Base_prueba_CHC$visita_pais)
```

```
#Base_prueba_CHC$P9<-ifelse(Base_prueba_CHC$P9==4,3,Base_prueba_CHC$P9)  
Base_prueba_CHC$P12<-ifelse(Base_prueba_CHC$P12>=75,"75 años o  
más",ifelse(Base_prueba_CHC$P12<15,"Menos de 15  
años",as.character(Base_prueba_CHC$P12)))  
Base_prueba_CHC$P44<-ifelse(Base_prueba_CHC$P44>=75,"75 años o  
más",ifelse(Base_prueba_CHC$P44<15,"Menos de 15  
años",as.character(Base_prueba_CHC$P44)))
```

```

#Base_prueba_CHC$P16<-
ifelse(Base_prueba_CHC$P16==8,7,Base_prueba_CHC$P16)
Base_prueba_CHC$P18S1<-ifelse(Base_prueba_CHC$P18S1>=60,"60 años o
más",as.character(Base_prueba_CHC$P18S1))

##Supresión de variables sensibles
Base_prueba_CHC<-subset(Base_prueba_CHC,select=-
c(UBIC_4,P7S2A_1,P7S2A,P7S2B,

P7S2B_1,P11S2B,P11S2A,P11S1A ,P11S1B,

P19S2A,P19S2B,

P24S1A, P24S1D, P24S1C, P24S2B,

P24S3A, P24S3D, P24S3C, P24S4_1, P24S4,P24S2C))

cols <- names(Base_prueba_CHC)
new_order <- append(cols[!cols %in% "visita_municipio"],
"visita_municipio", after = match("P23S10", cols))

Base_prueba_CHC <- Base_prueba_CHC[, new_order]

cols <- names(Base_prueba_CHC)

new_order <- append(cols[!cols %in% "visita_pais"], "visita_pais", after
= match("visita_municipio", cols))
Base_prueba_CHC <- Base_prueba_CHC[, new_order]

library(haven)

write.xlsx(Base_prueba_CHC,"Output/20241218_base_anonimizada.xlsx")

write.csv(Base_prueba_CHC, file = "Output/Base
anonimizada/20241218_base_anonimizada.csv", row.names =
FALSE,sep=";",fileEncoding = "latin1")
write_sas(Base_prueba_CHC, path = "Output/Base
anonimizada/20241218_base_anonimizada.sas7bdat")
write_dta(Base_prueba_CHC, path = "Output/Base
anonimizada/20241218_base_anonimizada.dta")
saveRDS(Base_prueba_CHC, file = "Output/Base
anonimizada/20241218_anonimizada.rds")

```