

ANEXOS

Anexo 1: Script 1

```
1 # Nombre: Script de manipulación y unión de La ENDES
2 # Versión: 1.0 (https://github.com/horaciochacon/Analisis-Endes-Peru/blob/master/scripts/Script1.R)
3 # Autores: Akram Hernández-Vásquez, Horacio Chacón-Torrico
4 # Objetivos: Cargar Las bases de datos de La ENDES, unir las en una sola base y
5 #            construir las variables para el análisis (obesidad, hipertensión arterial y dieta)
6 # Fecha: 28/02/19
7
8 library(tidyverse)
9 library(haven)
10 library(survey)
11
12 # Cargamos Las bases de datos de La ENDES 2017 con La función read_sav()
13
14 SALUD <- read_sav("data/CSALUD01.sav", encoding = 'UTF-8')
15 MUJER_OBS <- read_sav("data/RE223132.SAV", encoding = 'UTF-8')
16 MUJER_LAC <- read_sav("data/REC42.SAV", encoding = 'UTF-8')
17 MUJER_ANT <- read_sav("data/RECH5.SAV", encoding = 'UTF-8')
18 PERSONA <- read_sav("data/RECH1.SAV", encoding = 'UTF-8')
19 VIVIENDA <- read_sav("data/RECH23.SAV", encoding = 'UTF-8')
20 HOGAR <- read_sav("data/RECH0.SAV", encoding = 'UTF-8')
21
22 # Transformamos Las variables de unión (DESCOMPOSICIÓN DEL CASEID EN HHID Y QSNUMERO):
23 # 1. Descartamos últimos 3 caracteres de CASEID y Lo convertimos a HHID
24 # 2. Extraemos Los últimos 2 caracteres de CASEID y Lo convertimos a QSNUMERO
25 # 3. Convertimos La variable de texto a numérica
26
27 MUJER_OBS$HHID <- str_sub(MUJER_OBS$CASEID,1,(str_length(MUJER_OBS$CASEID)-3))
28 MUJER_OBS$QSNUMERO <- str_sub(MUJER_OBS$CASEID,-2,-1)
29 MUJER_OBS$QSNUMERO <- as.numeric(MUJER_OBS$QSNUMERO)
30 MUJER_LAC$HHID <- str_sub(MUJER_LAC$CASEID,1,(str_length(MUJER_LAC$CASEID)-3))
31 MUJER_LAC$QSNUMERO <- str_sub(MUJER_LAC$CASEID,-2,-1)
32 MUJER_LAC$QSNUMERO <- as.numeric(MUJER_LAC$QSNUMERO)
33 PERSONA$QSNUMERO <- PERSONA$HVIDX
34 MUJER_ANT$QSNUMERO <- MUJER_ANT$HA0
35
36 # Realizamos La unión mediante Los identificadores o Llaves HHID y QSNUMERO
37
38 BASE1 <- left_join(PERSONA, SALUD, by = c("HHID", "QSNUMERO"))
39 BASE2 <- left_join(MUJER_OBS, MUJER_LAC, by = c("HHID", "QSNUMERO"))
40 BASE2 <- left_join(BASE2, MUJER_ANT, by = c("HHID", "QSNUMERO"))
41 BASE3 <- left_join(HOGAR, VIVIENDA, by = "HHID")
42 BASE4 <- left_join(BASE1, BASE2, by = c("HHID", "QSNUMERO"))
43 BASE5 <- left_join(BASE3, BASE4, by = "HHID")
44
45 # Filtramos La base de datos final (BASE5)
46
47 BASE_FINAL <- filter(BASE5, V213 != 1 | is.na(V213)) # Descartamos a Las gestantes
48 BASE_FINAL <- filter(BASE_FINAL, QSRESINF == 1) # Filtramos entrevistas completas del CSALUD (n = 32 514)
49
50 # Descartamos Las bases de datos que ya no son necesarias para ahorrar memoria
51
52 rm(SALUD, BASE1, BASE2, BASE3, BASE4, BASE5, PERSONA, HOGAR, VIVIENDA, MUJER_LAC,
53 MUJER_OBS, MUJER_ANT)
54
55 # Calculamos La variable Hipertensión Arterial
56
57 BASE_FINAL$PAMS <- (BASE_FINAL$QS905S + BASE_FINAL$QS903S) / 2 # Obtenemos La PAM sistólica de dos mediciones
58 BASE_FINAL$PAMD <- (BASE_FINAL$QS905D + BASE_FINAL$QS903D) / 2 # Obtenemos La PAM diastólica de dos mediciones
59 BASE_FINAL$HIPERTENSION <- (BASE_FINAL$PAMS >= 140) | (BASE_FINAL$PAMD >= 90) # Definimos el criterio de HTA
60
61 # Calculamos La variable Obesidad
62
63 BASE_FINAL <- BASE_FINAL %>%
64 mutate(PESO = ifelse(QS902 == 1 & (QS900 >= 1 & QS900 < 999), QS900,
65 ifelse(QS902 == 4, HA2/10, NA)),
66 TALLA = ifelse(QS902 == 1 & (QS901 >= 1 & QS901 < 999), QS901,
67 ifelse(QS902 == 4, HA3/10, NA))) %>%
68 mutate(IMC = (PESO / (TALLA^2)) * 10000) %>%
69 mutate(OBESIDAD = IMC >= 30)
70
71 # Calculamos La variable Dieta
72
73 # Fruta
74 BASE_FINAL <- BASE_FINAL %>%
75 mutate(fruta_dxs = ifelse(QS213U == 1, QS213C, NA), fruta_uxd = QS214C,
76 fruta_semana = fruta_dxs * fruta_uxd)
77
```

```

78 # Jugo de Fruta
79 BASE_FINAL <- BASE_FINAL %>%
80   mutate(jugof_dxs = ifelse(QS215U == 1, QS215C, NA), jugof_vxd = QS216C,
81     jugof_semana = jugof_dxs * jugof_vxd * 2)
82
83 # Ensalada de Fruta
84 BASE_FINAL <- BASE_FINAL %>%
85   mutate(ensaladaf_dxs = ifelse(QS217U == 1, QS217C, NA), ensaladaf_pxd = QS218C,
86     ensaladaf_semana = ensaladaf_dxs * ensaladaf_pxd * 2)
87
88 # Ensalada de Verdura
89 BASE_FINAL <- BASE_FINAL %>%
90   mutate(ensaladav_dxs = ifelse(QS219U == 1, QS219C, NA), ensaladav_u = QS220U,
91     ensaladav_uxd = coalesce(as.numeric(QS220CV), as.numeric(QS220CC)),
92     ensaladav_porcion = ifelse(ensaladav_u == 2, ensaladav_uxd / 4,
93       ifelse(ensaladav_u == 1, ensaladav_uxd, NA)),
94     ensaladav_semana = ensaladav_dxs * ensaladav_porcion)
95
96 # Porciones de alimentos sanos
97 BASE_FINAL <- BASE_FINAL %>%
98   mutate(frutaensalada_pxs = rowSums(cbind(ensaladav_semana,ensaladaf_semana,
99     jugof_semana, fruta_semana),
100     na.rm = TRUE)) %>%
101   mutate(DIETA = ifelse(frutaensalada_pxs == 0, NA,
102     as.numeric(frutaensalada_pxs >= 31.5)))
103
104 # Generamos la base de datos final solo con las variables necesarias
105
106 ENDES <- select(BASE_FINAL, HHID, QSNUMERO, CONGLOMERADO = HV001, ESTRATO = HV022,
107   PONDERACION = PESO15_AMAS, SEXO = HV104, AREA_RESIDENCIA = HV025,
108   QUINTIL_BIENESTAR = HV270, REGION_NATURAL = SHREGION,HIPERTENSION, PESO,
109   TALLA, IMC, OBESIDAD,DIETA)
110
111 # Especificamos el diseño muestral de la encuesta
112
113 diseño <- svydesign(id =~ CONGLOMERADO, strata =~ ESTRATO, weights=~ PONDERACION, data=ENDES)

```