

Manejo y análisis del Cuestionario de Salud de la ENDES

Méd. Akram Hernández-Vásquez

Universidad San Ignacio de Loyola

Noviembre 2019

Curso Pre-Congreso

Metodología, redacción y ética en publicaciones científicas

Declaración de potenciales conflictos de intereses

- No he recibido remuneración, honorarios o financiamiento de cualquier tipo, para la realización de esta presentación.
- Tengo relación laboral con instituciones que realizan investigaciones usando ENDES: USIL, CRONICAS, UCSUR.
- Tengo relación laboral con la RPMESP que publica estudios usando ENDES.

ENDES 2018

OBJETIVO: Proveer información actualizada sobre la dinámica demográfica, el estado de salud de las madres y niños menores de cinco años, así como brindar información sobre el estado y factores asociados a las enfermedades no transmisibles y transmisibles.

34291

Encuestad@s - Salud
Feb-Dic



36760
viviendas

248
estratos

3254
conglomerados

97,6%
tasa de
respuesta



Cuestionarios
C del hogar
C individual
C de salud



Nivel de representatividad:

- a) Nacional
- b) Área de residencia: Urbana y Rural
- c) Región natural: LM, Resto Costa, Sierra y Selva
- d) Departamentos

Marco de estudio

Infancia

Madre y RN

Atención prenatal,
atención del parto,
peso al nacer,
mortalidad neonatal.

Menores 5

Desnutrición y anemia,
lactancia, IRA y EDA,
vacunación y CRED,
mortalidad infantil.

MEF

15-49 años

Reproducción y
anticoncepción;
embarazo, parto y
perperio; nupcialidad;
preferencias de
fecundidad; sida e ITS;
violencia.

Grupos especiales

6-11 años

50 y más

Salud ocular, bucal y
mental.

15-49 años

VIH/sida.

15-59 años

Prevención y control
del cancer.

15+ años

HT, DM, factores de
riesgo ET y ENT, TB,
salud mental,
antropometría, PA.

Bases disponibles

Bases		
▪ CSALUD01	▪ ps_WAWAWASI	• REC93DVdiscipli
▪ CSALUD08	▪ RE223132	na
▪ DIT	▪ RE516171	• REC94
▪ Programas	▪ RE758081	• REC95
Sociales x	▪ REC21	• REC0111_MOD
▪ Hogar	▪ REC41	ULO66.SAV
▪ ps_beca18	▪ REC42	• RECH0
▪ ps_COMEDOR	▪ REC43	• RECH1
▪ ps_PENSION65	▪ REC44	• RECH4
▪ ps_QALIWARM	▪ REC82	• RECH5
A	▪ REC83	• RECH6
▪ ps_TRABAJA	▪ REC84DV	• RECH23
▪ ps_VL	▪ REC91	• RECHM

«Llaves» para fusión de archivos

HHID



CASEID



HHID: Código del hogar

HHID
(9 DÍGITOS)



Conglomerado (HV001)

+

Vivienda (HV002)

+

Hogar (HV002A)

Composición del HHID

1 (HV001)

+

7 (HV002)

+

1 (HV002A)



000100701

CASEID: Código de la persona

CASEID
(12 dígitos)



Conglomerado (HV001)

+

Vivienda (HV002)


+

Hogar (HV002A)

+

HHID
(9 DÍGITOS)

Composición del CASEID

000100701 (HHID) 

+

_ (espacio)

+

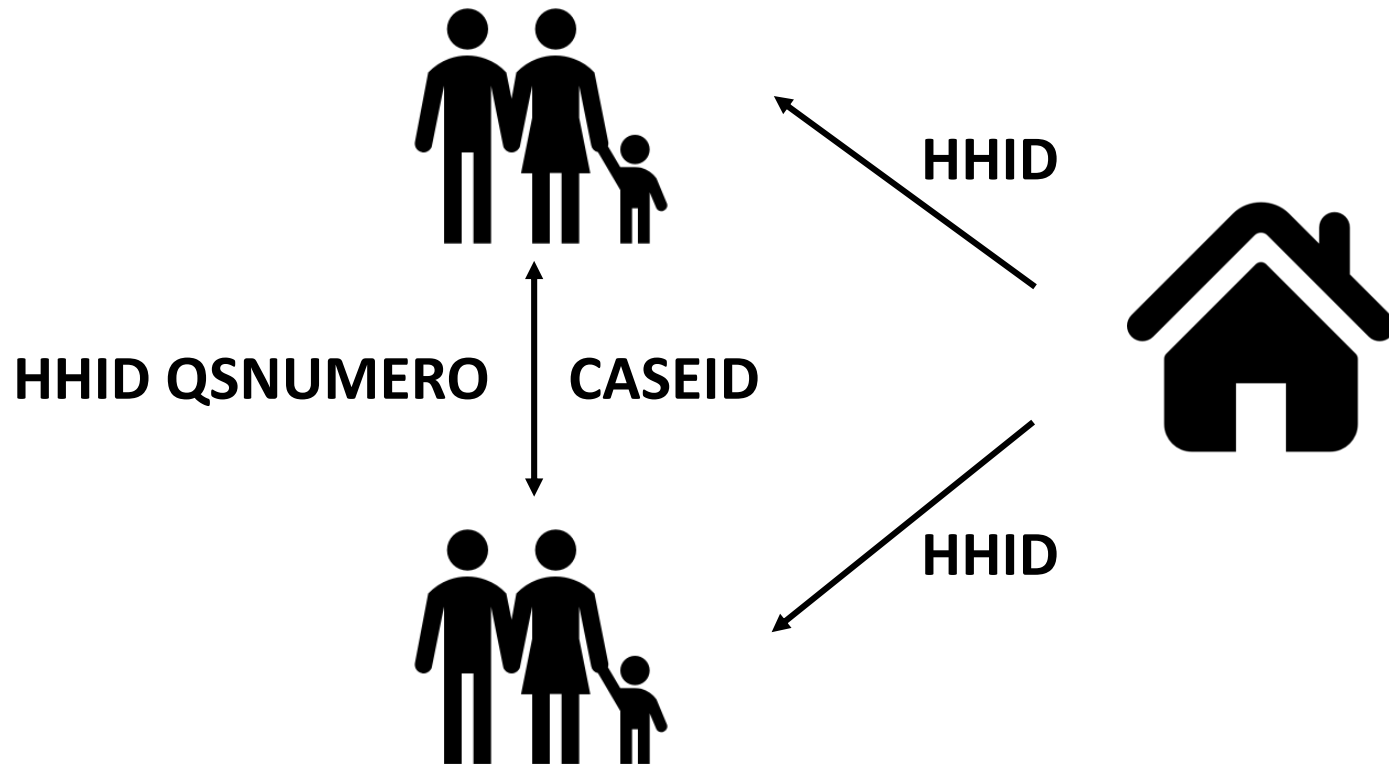
2 (QSNUMERO) 



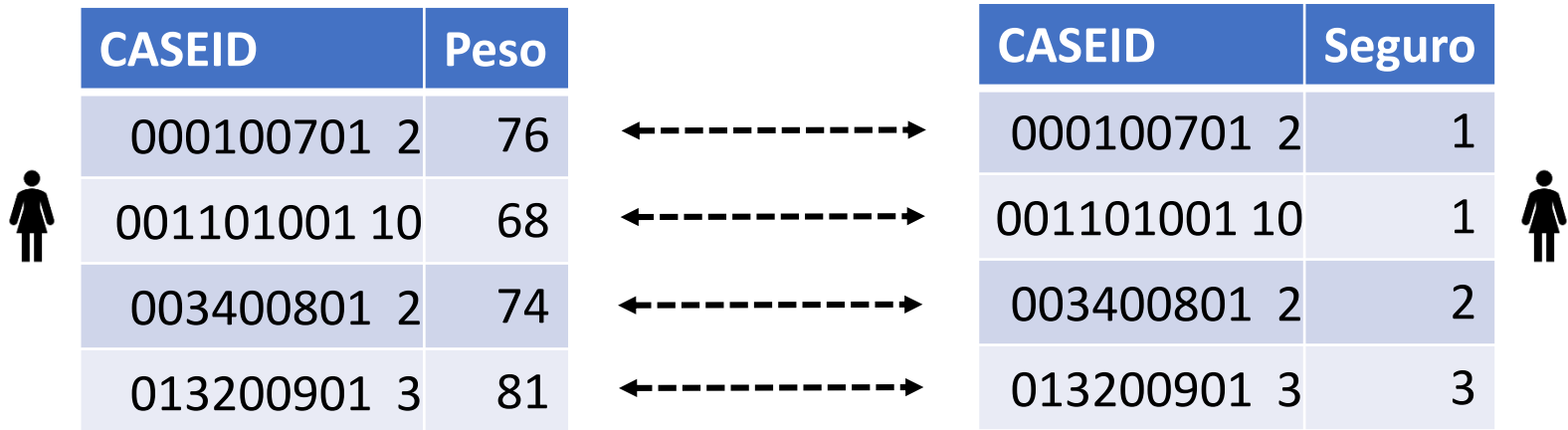
000100701 2

Nº de orden entrevistado (QSNUMERO)

Relación entre las «llaves»

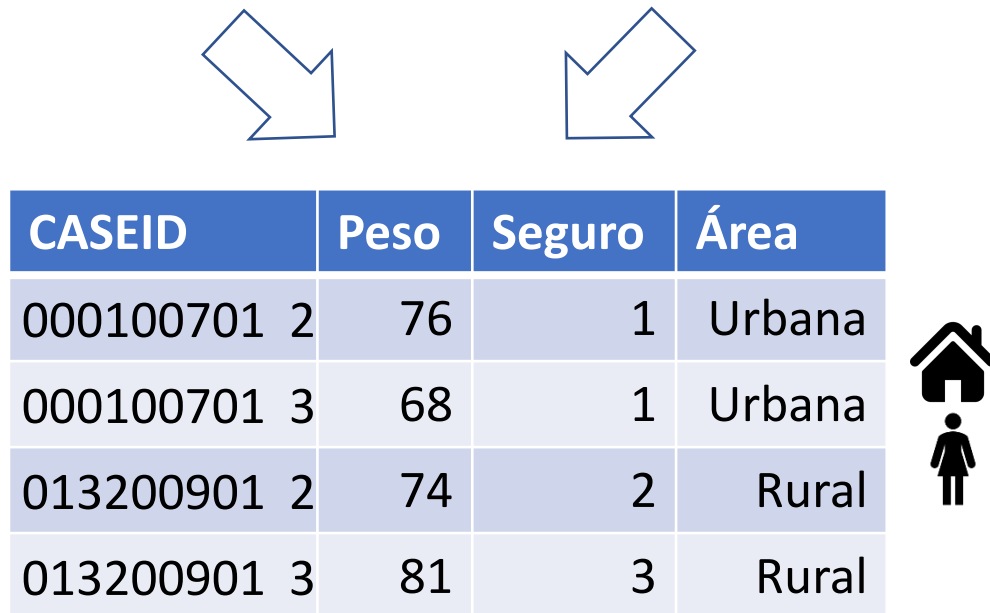
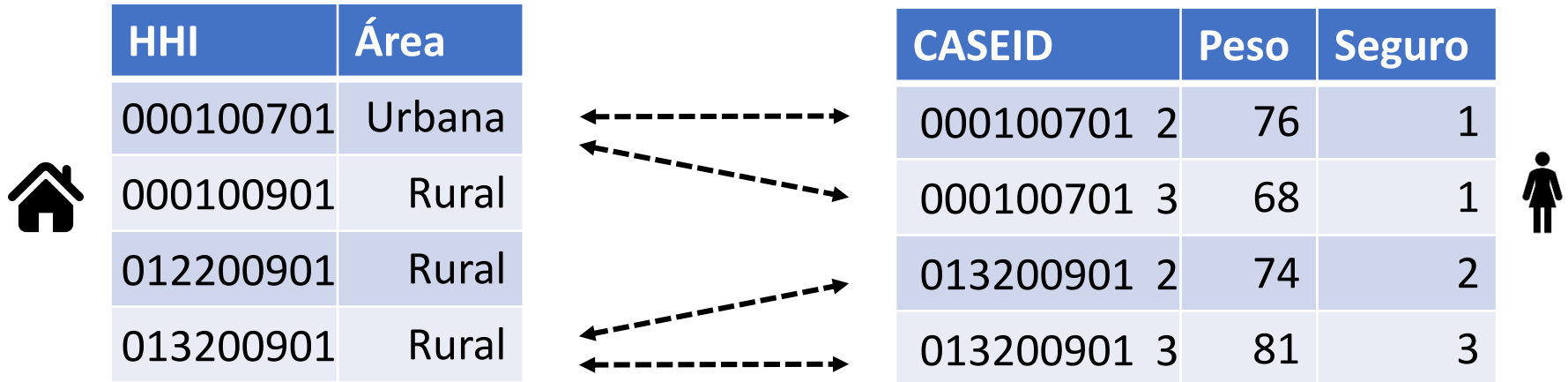


1:1 merge



CASEID	Peso	Seguro
000100701 2	76	1
001101001 10	68	1
003400801 2	74	2
013200901 3	81	3

1:m merge



Unión de individuos



CSALUD01

n=35388 | HHID, QSNUMBER



RECH1

n=147629 | HHID, HVIDX

Unir mediante:

- CASEID (no disponible)
- Hogar + orden (nombres ≠)

¿Creo un CASEID?

¿Renombro la variable HVIDX?

```
usesps "D:\RECH1.SAV"  
gen QSNUMERO=HVIDX  
save "D:\RECH1.dta"
```

CSALUD01

	QHCLUSTER	QNUMBER	QHHOME	QSNUMERO	HHID
1	1	7	1	3	000100701
2	1	14	1	1	000101401
3	1	35	1	1	000103501
4	1	53	1	1	000105301
5	1	71	1	2	000107101
6	1	84	1	1	000108401

RECH1

	HHID	HVIDX	HV101
1	000100701	1	Jefe
2	000100701	2	Esposa o esposo
3	000100701	3	Hijo/hija
4	000100701	4	Hijo/hija
5	000100701	5	Hijo/hija

Unión de individuos *merge 1:1*

 **CSALUD01**

n=35388 | HHID, QSNUMERO



 **RECH1**

n=147629 | HHID, QSNUMERO



 **RECH4**

n=147629 | HHID, IDXH4

```
. use "D:\CSALUD01.dta" //cargar la base
. merge 1:1 HHID QSNUMERO using "D:\RECH1.dta"
Result # of obs.
-----
not matched 112,241
  from master 0 (_merge==1)
  from using 112,241 (_merge==2)
matched 35,388 (_merge==3)
-----
. keep if _merge==3 //mantener los "matcheados"
(112,241 observations deleted)
. drop _merge
. save "D:\CSALUD01_RECH1.dta"

. usespss "D:\RECH4.SAV"
. gen QSNUMERO=IDXH4
. sort HHID QSNUMERO //ordenamos las variables
. save "D:\RECH4.dta"

. use "D:\CSALUD01_RECH1.dta"
. merge 1:1 HHID QSNUMERO using "D:\RECH4.dta"
. keep if _merge==3
(112,241 observations deleted)
. drop _merge
. save "D:\CSALUD01_RECH1.dta"
```

Unión de hogares *merge 1:1*

 **RECH0**
n=37486 | HHID

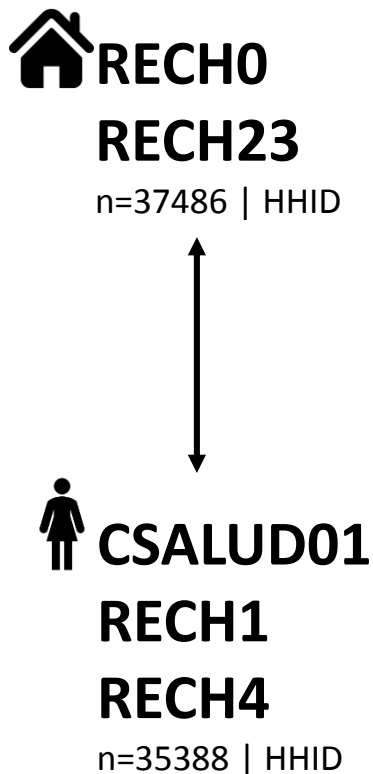


 **RECH23**
n=37486 | HHID

```
. use "D:\RECH0.dta" //cargar la base
. merge 1:1 HHID using "D:\RECH23.dta"
      Result                                     # of obs.
-----
not matched                                     0
matched                                       37,486  (_merge==3)
-----
. drop _merge
. save "D:\hogares.dta"
```

Unión de hogares e individuos

merge 1:m



```
. use "D:\hogares.dta" //cargar la base
. merge 1:m HHID using "D:\CSALUD01_RECH1_RECH4.dta"
Result # of obs.
-----
not matched 2,098
  from master 2,098 (_merge==1)
  from using 0 (_merge==2)

matched 35,388 (_merge==3)
-----
. keep if _merge==3 //mantener los "matcheados"
(2,098 observations deleted)
. drop _merge
. save "D:\hogares_salud.dta"
```

Presión arterial alta

CUADRO N° 1.1
PERÚ: PERSONAS DE 15 Y MÁS AÑOS DE EDAD CON PRESIÓN ARTERIAL ALTA DE ACUERDO A MEDICIÓN EFECTUADA,
SEGÚN CARACTERÍSTICA SELECCIONADA, 2014 - 2018
(Porcentaje)

Característica seleccionada	2014	2015	2016	2017	2018			Total de casos sin ponderar	Diferencias		Test de diferencia de medias		
	Valor estimado	Valor estimado	Valor estimado	Valor estimado	Valor estimado	Intervalo de confianza al 95%			2018/2014	2018/2017	2018/2014	2018/2017	
						Inferior	Superior						
Total	14,8	12,3	12,7	13,6	14,8	14,2	15,5	2,2	33 794	0,0	1,2	***	
Sexo													
Hombre	18,5	14,4	15,6	16,6	18,5	17,4	19,5	2,9	14 599	-0,0	1,8	**	
Mujer	11,6	10,2	9,9	10,7	11,4	10,6	12,1	3,4	19 195	-0,2	0,7	***	


Nota: Excluye a mujeres gestantes


*Filtramos los casos

```
. use "D:\hogares_salud.dta"
. tab QSRESINF
. keep if QSRESINF==1 & QS23>=15
(25,236 observations deleted)
. tab QSRESINF
```

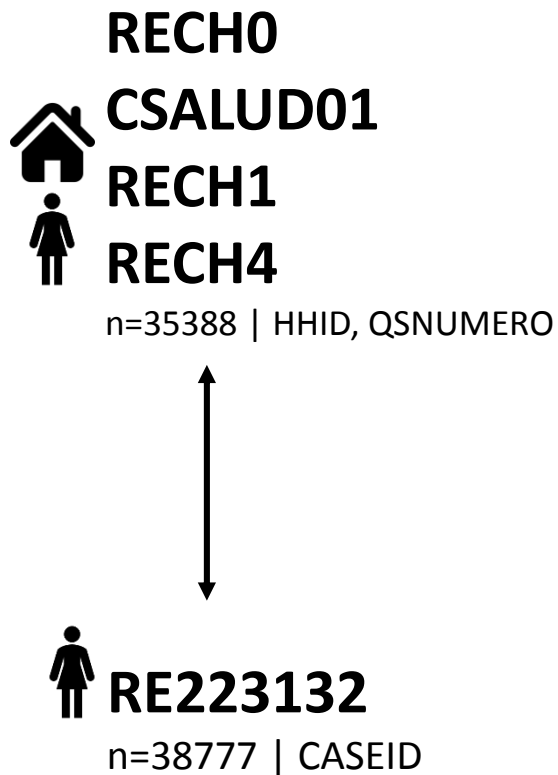
Resultado informante	Freq.	Percent	Cum.
Iniciar entrevista	34,291	96.90	96.90
Ausente	485	1.37	98.27
Rechazada	299	0.84	99.12
Incompleta	185	0.52	99.64
Discapacitada(o)	125	0.35	99.99
Otra	3	0.01	100.00
Total	35,388	100.00	

n=34291

14,7 (14,1-15,4)  **18,5 (17,5-19,5)**

11,2 (10,4-11,9) 

Unión de individuos



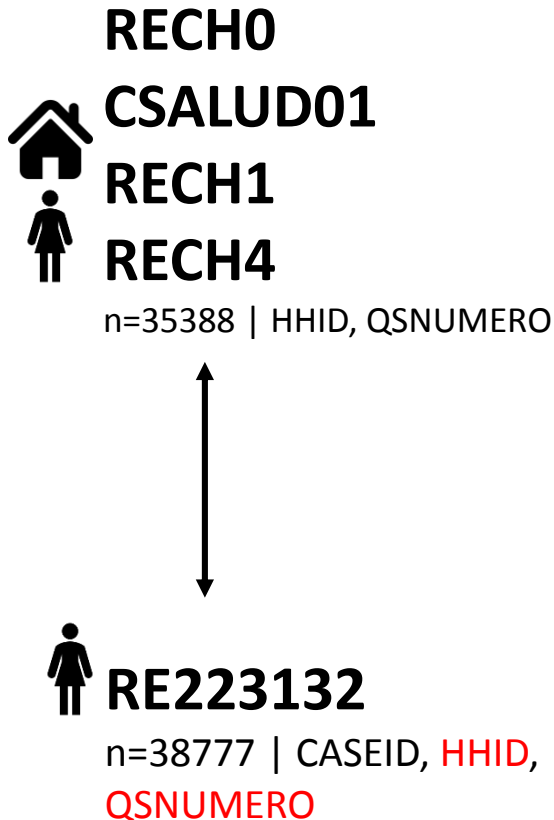
Unir mediante:

- CASEID (no disponible)
- Hogar + orden (no disponible)

¿Descompongo el CASEID?

```
usespsss "D:\RE223132.SAV"  
generate numero=length(CASEID)  
generate num_adj=numero-3  
generate HHID=substr(CASEID,1,num_adj)  
generate QSNUMERO=substr(CASEID,-2,2)  
destring QSNUMERO, replace  
sort HHID QSNUMERO  
save "D:\RE223132.dta"
```

Unión de individuos



Unir mediante:

- CASEID (no disponible)
- Hogar + orden (no disponible)

```
. use "D:\hogares_salud.dta" //cargar la base
. merge 1:1 HHID QSNUMERO using "D:\RE223132.dta"
Result # of obs.
-----
not matched 44,889
  from master 20,750 (_merge==1)
  from using 24,139 (_merge==2)

matched 14,638 (_merge==3)
-----

. tab V213

Actualmente
embarazada Freq. Percent Cum.

No, insegura 37,619 97.01 97.01
Si 1,158 2.99 100.00

. drop if V213==1
(1,158 observations deleted)

. save "D:\hogares_salud_mujeres.dta"
```

Unión de individuos

RECH0

CSALUD01

RECH1

RECH4

RE223132

n=58369 | CASEID, HHID,
QSNUMERO

Tabulamos y filtramos:

```
. use "D:\hogares_salud_mujeres.dta" //cargar la base
. tab QSREINF, miss
Iniciar entrevista |      33,794      57.90      57.90
      Ausente |         485       0.83      58.73
      Rechazada |         299       0.51      59.24
      Incompleta |         185       0.32      59.56
      Discapacitada(o) |         125       0.21      59.77
      Otra |              3       0.01      59.78
      . |      23,478      40.22     100.00
-----+-----
Total |      58,369     100.00
```

```
. keep if QSRESINF==1 & QS23>=15
(24,575 observations deleted)
. tab QSRESINF, miss
```

```
Iniciar entrevista |      33,794     100.00     100.00
-----+-----
Total |      33,794     100.00
```

2018				
Valor estimado	Intervalo de confianza al 95%		Coeficiente de variación	Total de casos sin ponderar
	Inferior	Superior		
14,8	14,2	15,5	2,2	33 794

Diseño muestral y ponderación

- Diseño muestral: estrato (HV022) y conglomerado (HV001)
- Factor de ponderación: factor (PESO15_AMAS)

```
**generando peso
. generate peso=PESO15_AMAS/1000000
**seteando la encuesta
. svyset HV001 [pw= peso], strata(HV022)
      pweight: peso
          VCE: linearized
Single unit: missing
Strata 1: HV022
  SU 1: HV001
  FPC 1: <zero>
```

- Uso: *svy: tabulate var*

Creación de una submuestra

- Usar *subpop* dentro de las estimaciones
- Crear la submuestra de interés:

```
**submuestra de adultos mayores  
. gen AM=.  
. replace AM=1 if (QS23>=60)  
. replace AM=0 if (QS23<60)
```

- Uso: *svy, subpop (if AM==1): tabulate var*
- NO: *svy: tabulate var if AM==1*

Análisis de presión arterial alta

- Total de la población (n=33719)

```

**creamos la variables PAMS y PAMD
. gen PAMS=((QS905S+QS903S)/2) //PA media sistolica
. gen PAMD=((QS905D+QS903D)/2) //PA media diastolica

**creamos la variables presión arterial alta (PA)
. gen PA=.
. replace PA=0 if PAMS<140 | PAMD<90
. replace PA=1 if PAMS>=140 | PAMD>=90
    
```

- *svy: tabulate PA, percent ci col cv format(%7.1f)*

```

Number of strata = 248
Number of PSUs = 3,254
Number of obs = 33,794
Population size = 35,239.822
Design df = 3,006
    
```

PA	column	cv	lb	ub
0	85.2	0.4	84.5	85.8
1	14.8	2.2	14.2	15.5
Total	100.0			

Key: column = column percentage
 cv = coefficients of variation of column percentage
 lb = lower 95% confidence bound for column percentage
 ub = upper 95% confidence bound for column percentage

2018				
Valor estimado	Intervalo de confianza al 95%		Coeficiente de variación	Total de casos sin ponderar
	Inferior	Superior		
14,8	14,2	15,5	2,2	33 794

Análisis de presión arterial alta

- Adultos mayores (n=4874)

```
. svy, subpop(if AM==1): tabulate PA, percent ci col cv format(%7.1f)
```

```
Number of PSUs      =      3,246      Population size      = 35,163.668
Subpop. no. obs     =      4,874      Subpop. size        = 6,062.5559
Design df           =      3,001
```

PA	column	cv	lb	ub
0	61.3	1.8	59.1	63.4
1	38.7	2.8	36.6	40.9
Total	100.0			

Característica seleccionada	2018				
	Valor estimado	Intervalo de confianza al 95%		Coeficiente de variación	Total de casos sin ponderar
		Inferior	Superior		
Grupo de edad					
15 - 19	(2,4)	1,4	3,4	20,5	2.899
20 - 29	5,4	4,6	6,3	7,8	7.836
30 - 39	7,5	6,6	8,4	6,0	9.195
40 - 49	14,4	13,0	15,8	5,1	5.365
50 - 59	20,8	18,8	22,7	4,8	3.625
60 y más	38,7	36,6	40,8	2,8	4.874

Análisis de presión arterial alta

- Regresión logística entre PA y edad (H_1 : >edad, >PA)

```
. svy: logistic PA QS23, cformat(%7.3f)
```

Survey: Logistic regression

```
Number of strata   =          248           Number of obs       =       33,794
Number of PSUs    =       3,254           Population size     =  35,239.822
                                                Design df          =         3,006
                                                F( 1, 3006)       =       1179.36
                                                Prob > F          =         0.0000
```

PA	Odds Ratio	Linearized Std. Err.	t	P> t	[95% Conf. Interval]
QS23	1.054	0.002	34.34	0.000	1.051 1.057
_cons	0.015	0.001	-51.36	0.000	0.013 0.017

Mayores detalles...

Email: ahernandez@usil.edu.pe

Hernández-Vásquez A, Chacón-Torrico H. Manipulación, análisis y visualización de datos de la Encuesta Demográfica y de Salud Familiar con el programa R. Rev Peru Med Exp Salud Publica. 2019;36(1):128-33.doi:10.17843/rpmesp.2019.361.4062.

Códigos R - GitHub: <https://github.com/horaciochacon/Analisis-Endes-Peru/blob/master/scripts/Script1.R>

Códigos R - OSF: <https://osf.io/ahjd6/?branch=master>

Diccionario de variables:

https://webinei.inei.gov.pe/anda_inei/index.php/catalog/671/data_dictionary

Perú: Enfermedades No Transmisibles y Transmisibles, 2018

https://www.inei.gov.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1657/libro.pdf

¡Gr@ci@s por su @tención!