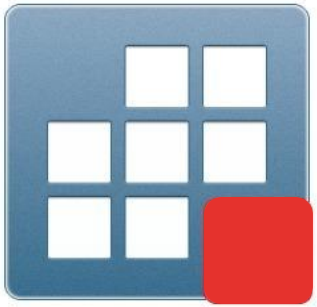
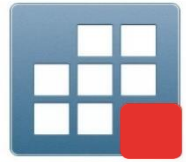
Programación en STATA®

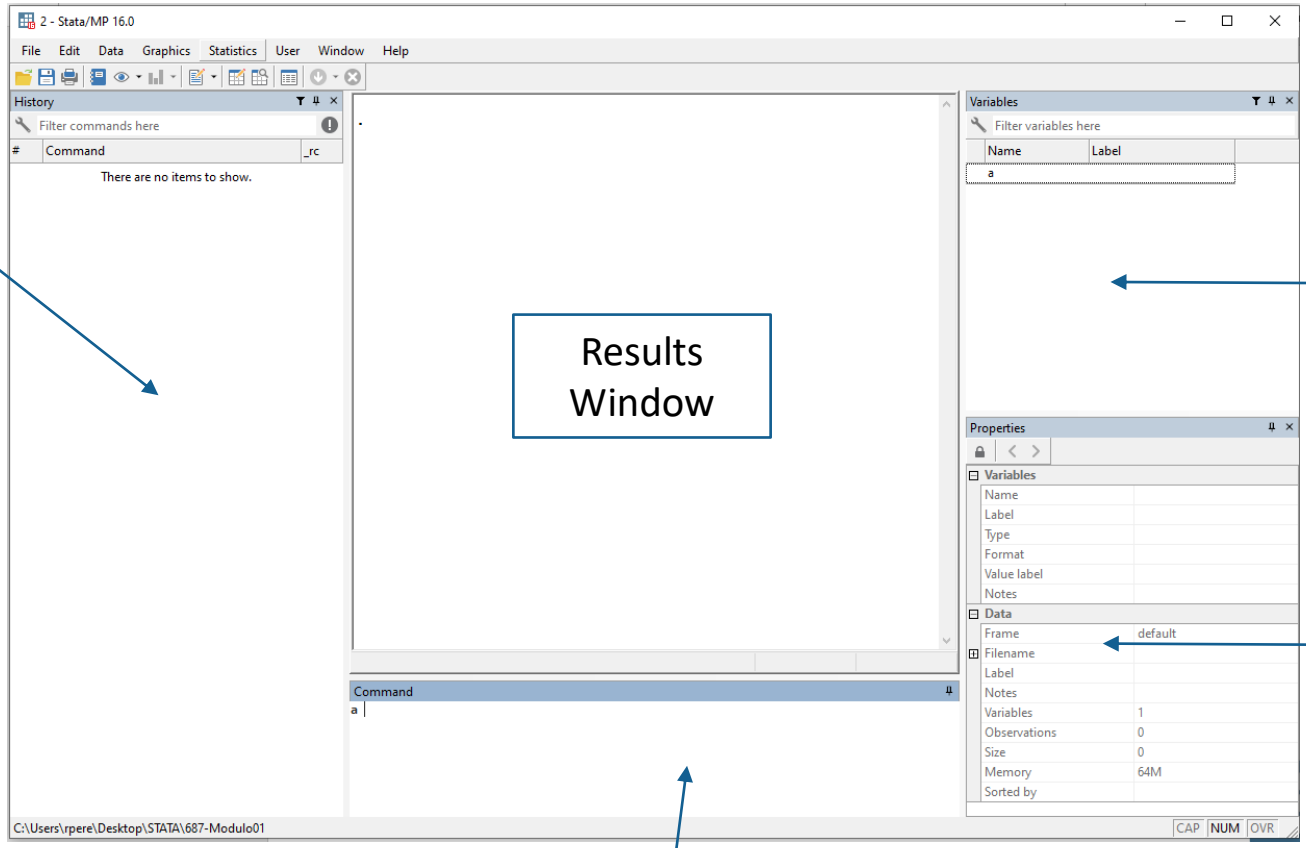
Richard P. Pérez Palma Ponce



Parte 1: Introducción



Interfaz Gráfica de Usuarios (GUI) de Stata



Review Window

Results Window

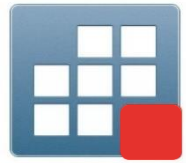
Variables Window

Properties Window

Commands Window

Métodos Abreviados (Shortcuts)

F2	Describe data
Ctrl + 8	Abrir el data editor
Ctr + 9	Abrir nuevo do-file
Ctr + D	Ejecuta texto seleccionado



Establecer los parámetros iniciales

cd

Cambiar el directorio actual

dir (.*dta)

Mostrar nombre de archivos en la ruta de trabajo

capture log close

Cierra el log existente de cualquier do-file

log using ".txt", replace

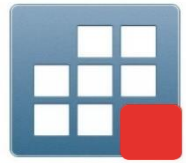
Crea un nuevo log file para guardar resultados y códigos

search (findit)

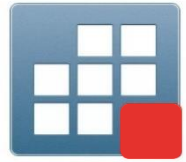
Encontrar el paquete a instalar

ssc install

Instalar paquete



Parte 2: Elementos de Programación



Creación de objetos: Scalar y Matrix

Scalar

Scalar A = 3

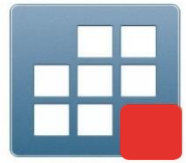
Matrix

Llenar Matrices

(nx1) : matrix A = $(a_1 \backslash a_2, \backslash \cdots \backslash a_n)$
(1xn) : matrix A = $(a_1, a_2, , \ldots, a_n)$

Operaciones con matrices:
Matriz Inversa: invsym
Determinante: det

Operator	Symbol
Unary operators	
negation	-
transposition	'
Binary operators	
(lowest precedence)	
row join	\
column join	,
addition	+
subtraction	-
multiplication	*
division by scalar	/
Kronecker product	#
(highest precedence)	



Bucles y Condicionales

Loops

- **Textos y listas**

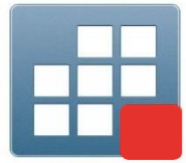
```
foreach x of varlist mpg weight foreign {
        summarize `x'
}

foreach x in "Richard" "Manuel" "Gabriela" {
        display length("`x'")
}
```

- **Numérico**

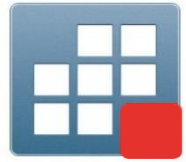
```
forvalues i = 10(10)50 {
        generate d`i' = 0
}
```

Condicionales

```
local m = 1
forvalues i =  10(10)40 {
        if `i' == 30  {
                display "Se ha elegido el número `m'"
        }
        local m = `++m'
}
```



Parte 3: Importación de datos



Importación de datos: .dta y bases externas.

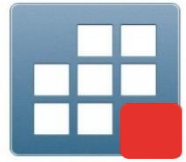
use

```
. use sumaria-2019.dta, replace

. describe

Contains data from sumaria-2019.dta
  obs:        34,565
 vars:           158                          10 Sep 2020 02:23
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
aÑo             str4    %4s                   año de la encuesta
mes             str2    %2s                   mes de ejecución de la encuesta
nconglome       str6    %6s                   número de conglomerado (proveniente del marco)
conglome        str6    %6s                   número de conglomerado
vivienda        str3    %3s                   número de selección de vivienda
hogar           str2    %2s                   número secuencial del hogar
ubigeo          str6    %6s                   ubicación geográfica
dominio         byte    %8.0g      dominio    dominio geográfico
estrato         byte    %46.0g     estrato    estrato geográfico
percepho        byte    %8.0g                 total de perceptores de ingresos
```

import

```
. import delimited "Ubigeo.csv", stringcols(1) clear
(7 vars, 1,874 obs)

. save Ubigeo.dta, replace
file Ubigeo.dta saved
```

Parte 4: Manejo de Base de Datos



Análisis de Base de Datos

describe

```
describe

Contains data from https://stats.idre.ucla.edu/stat/data/patient_pt1_stata_dm.dta
  obs:           120
 vars:            25                          6 May 2017 19:05
 size:        11,400
-------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
-------------------------------------------------------------------------------
hospital        str14   %14s
hospid          byte    %8.0g                 hospid
docid           str5    %9s
```

list

```
. list dominio estrato mieperho percepho in 100/120

     +-------------------------------------------------------------------+
     | dominio                                estrato   mieperho   percepho |
     |-------------------------------------------------------------------|
100. |   selva   Área de empadronamiento rural (aer) simple          2          1 |
101. |   selva   Área de empadronamiento rural (aer) simple          5          2 |
102. |   selva   Área de empadronamiento rural (aer) simple          3          2 |
103. |   selva   Área de empadronamiento rural (aer) simple          5          3 |
104. |   selva   Área de empadronamiento rural (aer) simple          3          2 |
     +-------------------------------------------------------------------+
```

Análisis de Base de Datos

codebook

```
dominio                                                    dominio geográfico
-------------------------------------------------------------------------------

                type:  numeric (byte)
               label:  dominio

               range:  [1,8]                        units:  1
       unique values:  8                    missing .:  0/34,565

          tabulation:  Freq.   Numeric  Label
                       4,782         1  costa norte
                       3,258         2  costa centro
                       2,196         3  costa sur
                       2,362         4  sierra norte
                       6,049         5  sierra centro
                       4,898         6  sierra sur
                       6,990         7  selva
                       4,030         8  lima metropolitana
```

tabulate

```
. tabulate dominio
```

dominio geográfico	Freq.	Percent	Cum.
costa norte	4,782	13.83	13.83
costa centro	3,258	9.43	23.26
costa sur	2,196	6.35	29.61
sierra norte	2,362	6.83	36.45
sierra centro	6,049	17.50	53.95
sierra sur	4,898	14.17	68.12
selva	6,990	20.22	88.34
lima metropolitana	4,030	11.66	100.00
Total	34,565	100.00	

Creación de Variables

generate

```
. gen d_gf  = 1 if mieperho >= 5
(25,464 missing values generated)
```

replace

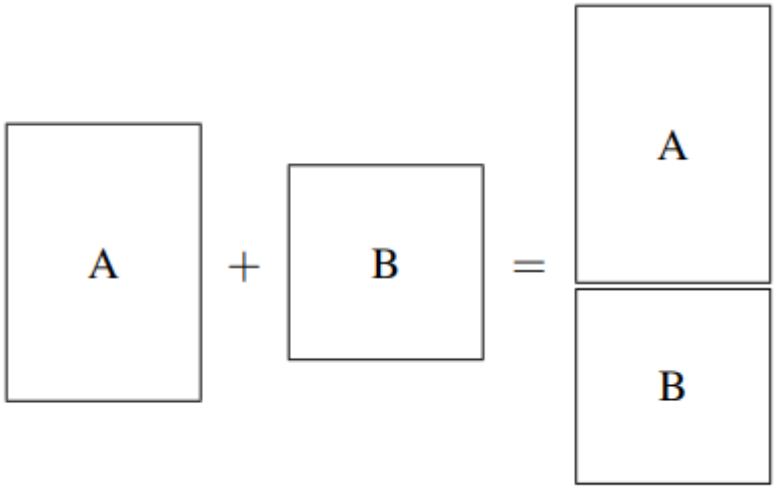
```
. gen d_gf  = 1 if mieperho >= 5
(25,464 missing values generated)
```

values

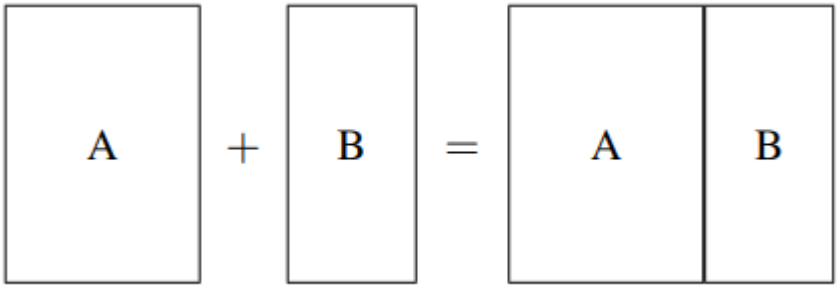
```
label values d_gf etiquetas


label variable d_gf  "Hogares con 5 o más miembros"


label define etiquetas 0 "Hogares con menos de 5 miembros"  1 "Hogares con 5 o más miembros"
```

Combinar Base de Datos

Append

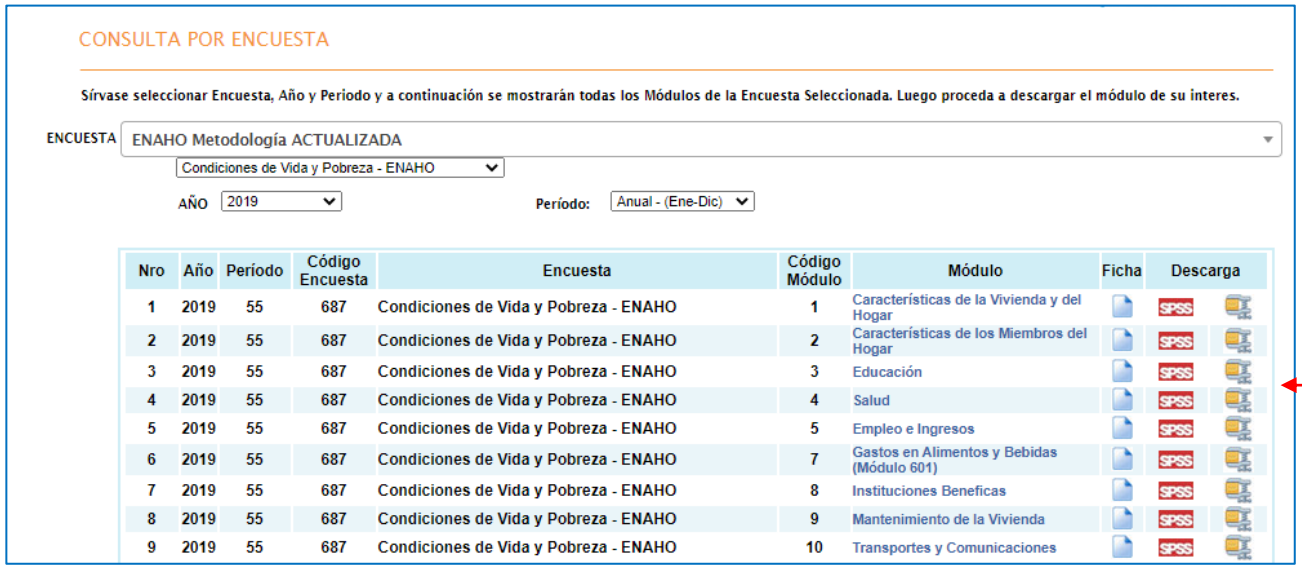


Merge



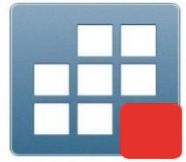
Descargar datos de ENAHO: Encuesta Nacional de Hogares

https://www.inei.gob.pe/bases-de-datos/

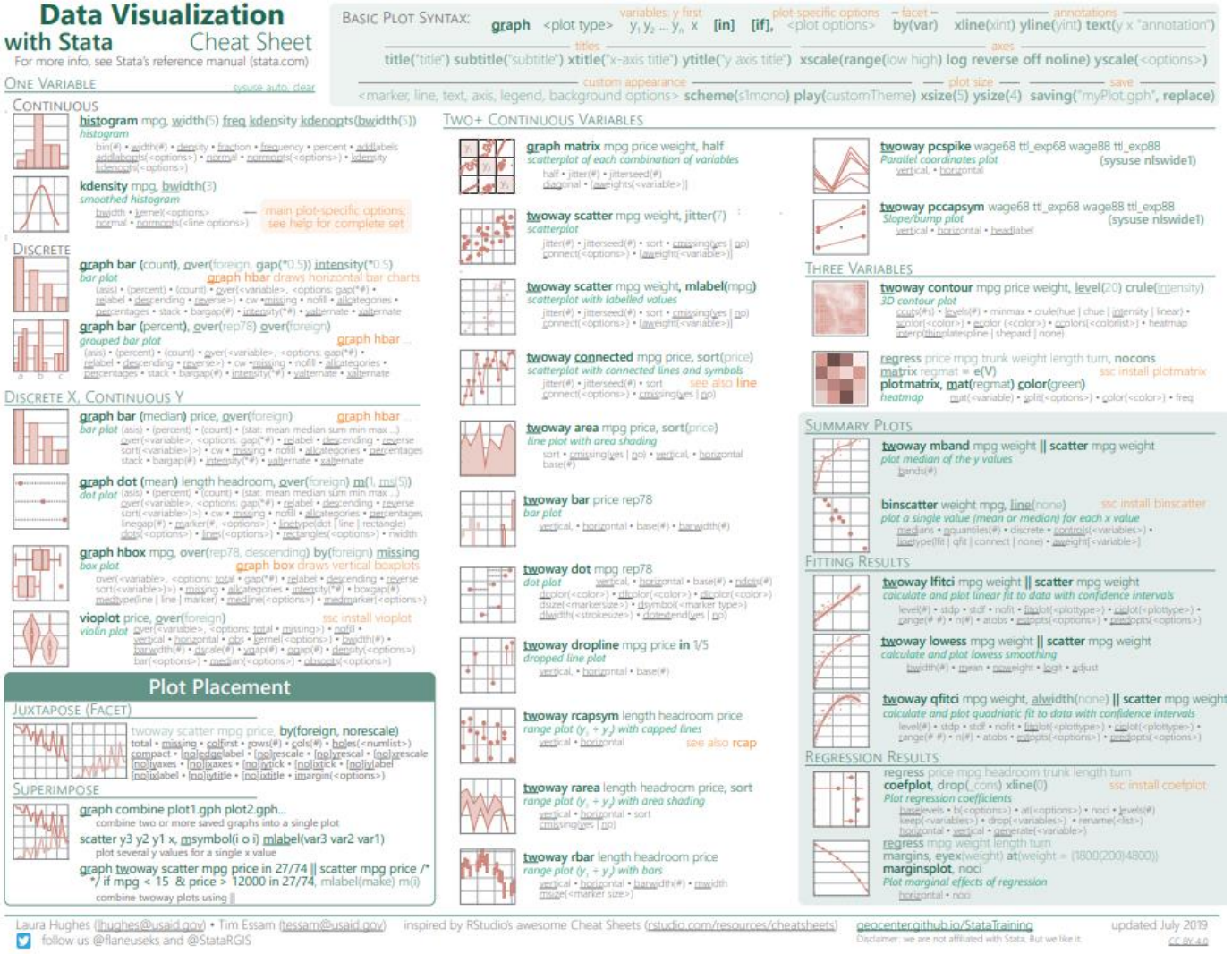


Nro	Año	Período	Código Encuesta	Encuesta	Código Módulo	Módulo	Ficha	Descarga
1	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	1	Características de la Vivienda y del Hogar		
2	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	2	Características de los Miembros del Hogar		
3	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	3	Educación		
4	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	4	Salud		
5	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	5	Empleo e Ingresos		
6	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	7	Gastos en Alimentos y Bebidas (Módulo 601)		
7	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	8	Instituciones Beneficas		
8	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	9	Mantenimiento de la Vivienda		
9	2019	55	687	Condiciones de Vida y Pobreza - ENAHO	10	Transportes y Comunicaciones		

Parte 5: Personalización de Gráficos

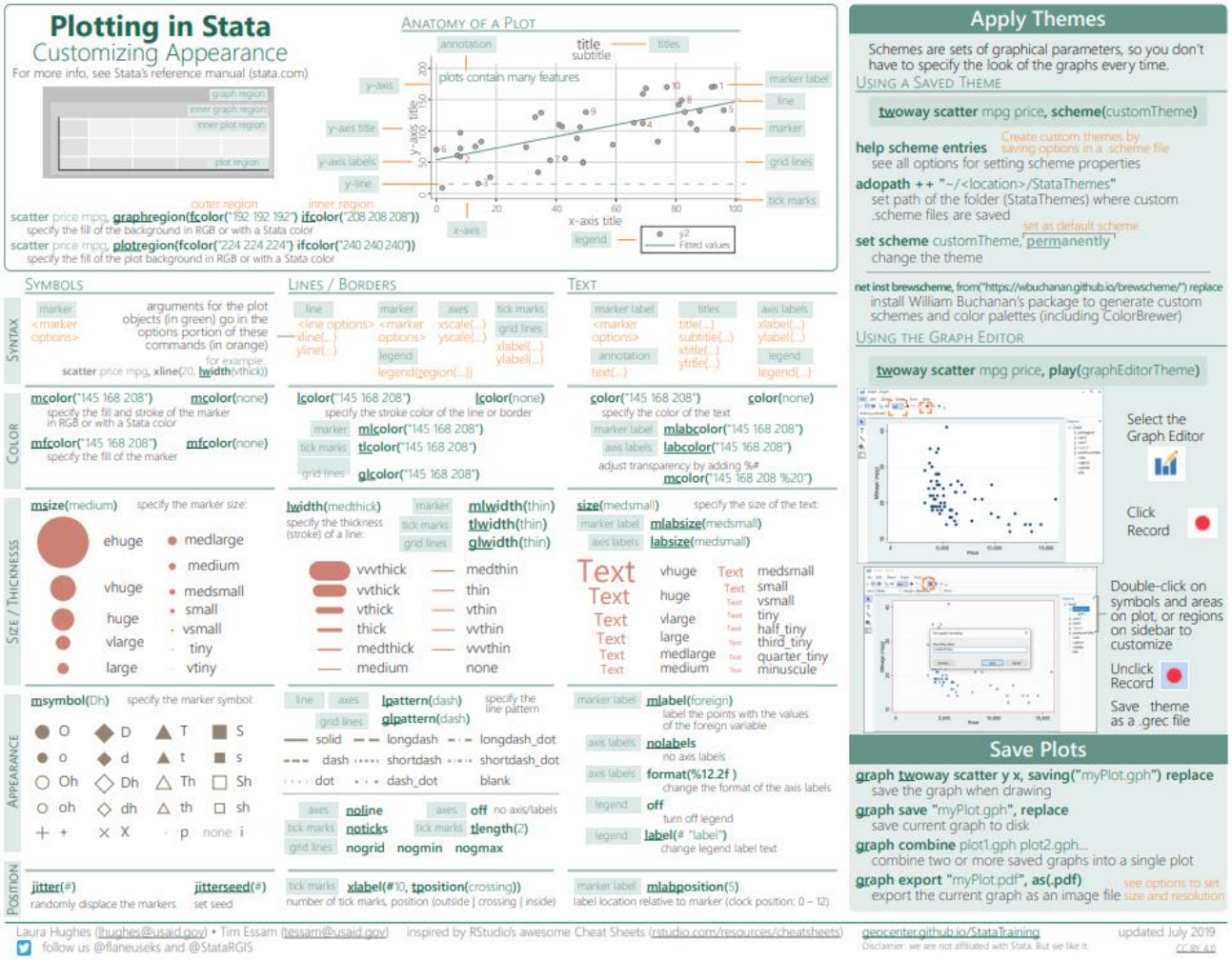


Cheat Sheet: Visualización de información



https://www.stata.com/bookstore/statacheatsheets.pdf

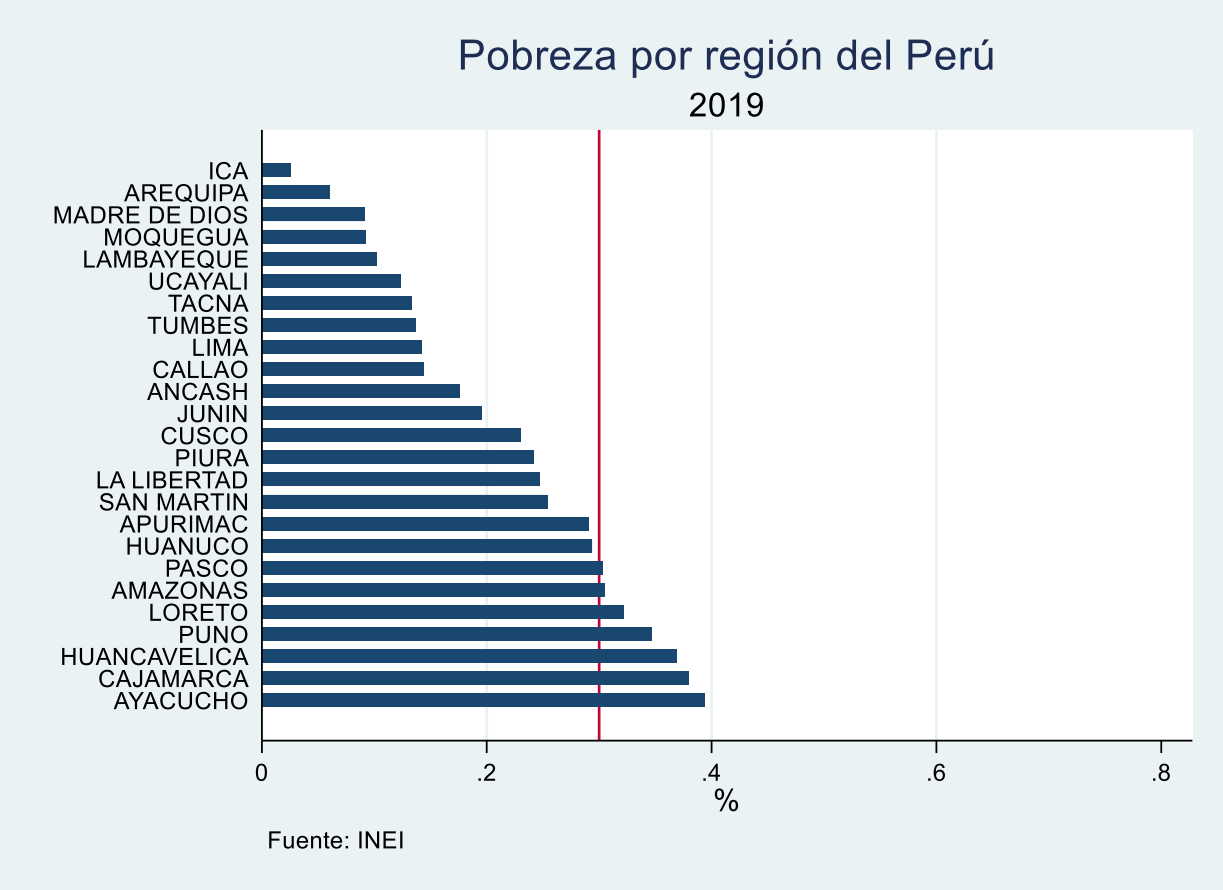
Cheat Sheet: Personalización de Gráficos



https://www.stata.com/bookstore/statacheatsheets.pdf

Ejemplo 1: Gráfico de Barra + Parámetros

```
graph hbar pobreza_n [w = fp], over(departamento, sort((mean) pobreza_n) label(labsize(small)))  ytitle("%") title("Pobreza por región del Perú") note("Fuente: INEI") subtitle("2019") ylabel(0(0.2)0.8, labsize(small)) yline(0.3)
```



Ejemplo 2: Dos Gráficos en 1 sólo + Parámetros

```
graph twoway (scatter mieperho inghog2d[w = factor07],
mcolor("0 155 255") msize(tiny)) (lfit mieperho inghog2d [w = factor07],
lwidth(vthick)), title("Relación: Miembros del hogar e Ingreso") xtitle("Ingresos")
ytitle("Miembros del hogar")
```