# Creating customized tables in Stata

# (Creando tablas de resultados personalizadas en Stata)

Gabriela Ortiz
StataCorp LLC

June 3, 2021

# Learn to create tables such as these
## (Aprenderemos a crear tablas como estas)

reporte.html

File | C:/Users/Stata/tables/reporte.html

| | Región | | | | |
|---|---|---|---|---|---|
| | NE | MO | S | O | Total |
| Diabetes | | | | | |
| No | 95.32% | 95.49% | 94.36% | 95.62% | 95.18% |
| Sí | 4.68% | 4.51% | 5.64% | 4.38% | 4.82% |
| Presión alta | | | | | |
| No | 55.92% | 59.99% | 57.34% | 57.19% | 57.72% |
| Sí | 44.08% | 40.01% | 42.66% | 42.81% | 42.28% |
| Tuvo ataque al corazón | | | | | |
| No | 96.32% | 95.64% | 95.41% | 94.41% | 95.40% |
| Sí | 3.68% | 4.36% | 4.59% | 5.59% | 4.60% |
| Estado de salud | | | | | |
| Excelente | 26.95% | 26.33% | 19.14% | 21.68% | 23.29% |
| Muy buena | 26.76% | 26.01% | 22.82% | 25.18% | 25.07% |
| Regular | 30.26% | 26.52% | 28.29% | 29.14% | 28.43% |
| No muy buena | 12.33% | 15.12% | 18.65% | 17.60% | 16.16% |
| Terrible | 3.69% | 6.02% | 11.11% | 6.40% | 7.05% |

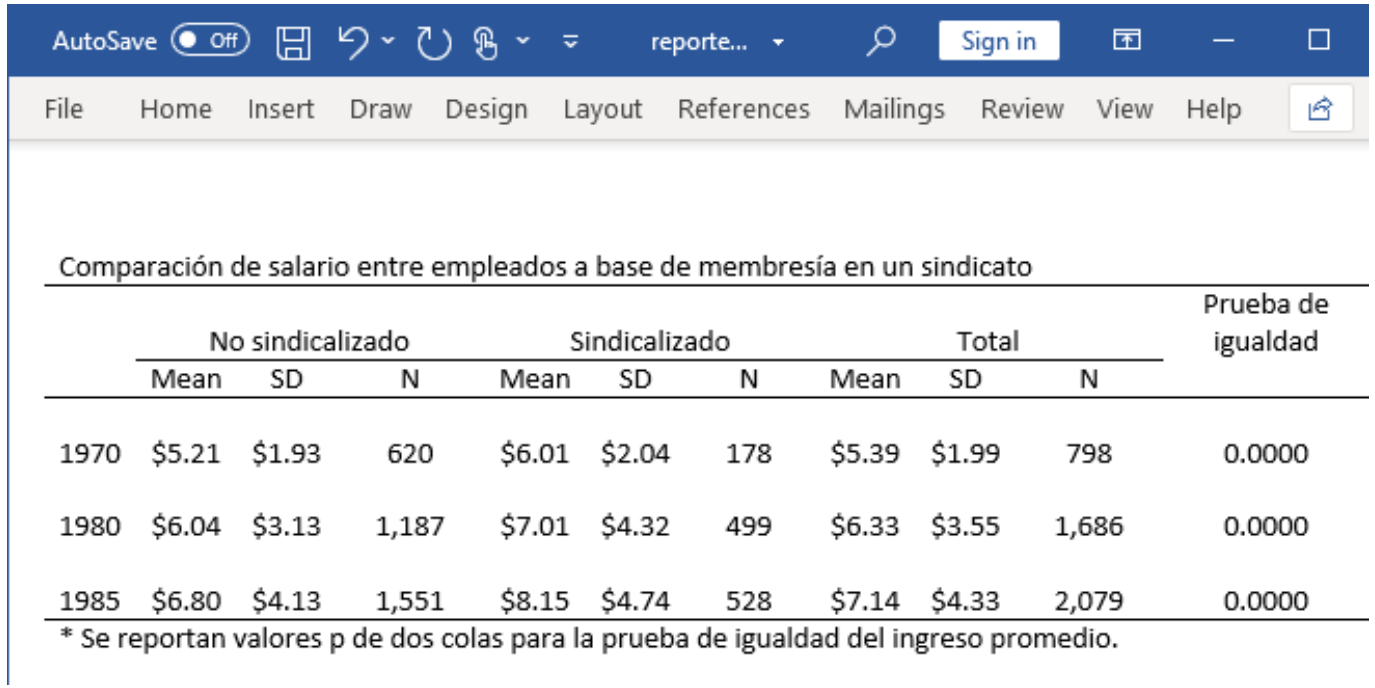
AutoSave Off » report2

File Home Insert Draw Page Layou Formulas Data Review View

H20

| | A | B | C | D |
|---|---|---|---|---|
| 1 | | % of the population | Probability of developing: | |
| 2 | | | High BP | Diabetes |
| 3 | Age group | | | |
| 4 | 20–29 | 22.41% | 0.199 | 0.007 |
| 5 | 30–39 | 15.67% | 0.287 | 0.013 |
| 6 | 40–49 | 12.29% | 0.384 | 0.032 |
| 7 | 50–59 | 12.47% | 0.519 | 0.052 |
| 8 | **60–69** | **27.63%** | **0.562** | **0.080** |
| 9 | 70+ | 9.53% | 0.657 | 0.111 |
| 10 | BMI | | | |
| 11 | <18.5 | 3.18% | 0.227 | 0.032 |
| 12 | **18.5-24.99** | **48.49%** | **0.321** | **0.035** |
| 13 | 25-29.99 | 32.81% | 0.472 | 0.044 |
| 14 | 30+ | 15.53% | 0.656 | 0.093 |

# Learn to create tables such as these
# (Aprenderemos a crear tablas como estas)




Comparación de salario entre empleados a base de membresía en un sindicato

| | No sindicalizado | | | Sindicalizado | | | Total | | | Prueba de igualdad |
|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | N | Mean | SD | N | Mean | SD | N | |
| 1970 | $5.21 | $1.93 | 620 | $6.01 | $2.04 | 178 | $5.39 | $1.99 | 798 | 0.0000 |
| 1980 | $6.04 | $3.13 | 1,187 | $7.01 | $4.32 | 499 | $6.33 | $3.55 | 1,686 | 0.0000 |
| 1985 | $6.80 | $4.13 | 1,551 | $8.15 | $4.74 | 528 | $7.14 | $4.33 | 2,079 | 0.0000 |

* Se reportan valores p de dos colas para la prueba de igualdad del ingreso promedio.

# Learn to create tables such as these

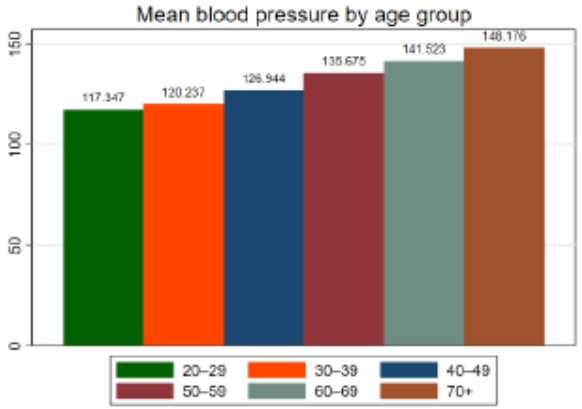
(Aprenderemos a crear tablas como estas)

report.pdf - Adobe Acrobat Reader DC (32-bit)

File Edit View Sign Window Help

Home Tools report.pdf

1 / 1 65%




Probability of developing high blood pressure

| | % Population | Prob. | 95% CI |
|---|---|---|---|
| Sex | | | |
| Male | 47.49% | 0.471 | [0.458, 0.484] |
| Female | 52.51% | 0.379 | [0.367, 0.392] |
| Diabetes status | | | |
| Not diabetic | 95.18% | 0.418 | [0.409, 0.427] |
| Diabetic | 4.82% | 0.512 | [0.471, 0.553] |
| Age group | | | |
| 20-29 | 22.42% | 0.172 | [0.156, 0.187] |
| 30-39 | 15.66% | 0.284 | [0.262, 0.306] |
| 40-49 | 12.28% | 0.396 | [0.369, 0.423] |
| 50-59 | 12.47% | 0.543 | [0.516, 0.570] |
| 60-69 | 27.64% | 0.579 | [0.561, 0.597] |
| 70+ | 9.53% | 0.665 | [0.635, 0.694] |

# What are Stata's capabilities for creating tables? (¿Que capacidades tiene Stata para crear tablas?)

- Create tables of summary statistics, results of hypothesis tests, and regression results. Create tables with a combination of these results, or with any other results returned from a Stata command.
- Customize your table by adding stars for significance, formatting table headers and statistics, modifying borders, and more.
- Export your finalized table to Excel, Word, HTML, Markdown, LaTeX, PDF, and SMCL.
- Create a standard look with style files
- Work interactively in the Tables Builder or by typing commands

# How does it work?
# (¿Cómo funciona?)

1. Collect results with the **`collect`** prefix, **`collect get`** command, or **`table`** command.

2. Lay out the table by telling Stata what it should place on the rows and columns.

3. Customize your table to obtain your preferred style. Format your statistics, add percent signs, enclose numbers in parenthesis, and more.

4. Export your table to another format. Create a document with just one table, or incorporate your table in more extensive reports.

# Creating tables with the new table command

# (Creando tablas con el nuevo comando table)

# Table with percentages (Tabla con porcentajes)

```
. use nhanessp2, clear
(Second National Health and Nutrition Examination Survey)

. table región, statistic(percent)
```

| | Percent |
|---|---|
| Región | |
| NE | 20.25 |
| MO | 26.80 |
| S | 27.56 |
| O | 25.39 |
| Total | 100.00 |

# Table with percentages across categories (Tabla con porcentajes en cada categoría)

```
. table salud región, statistic(percent, across(salud))
```

| | Región | | | | |
|---|---|---|---|---|---|
| | NE | MO | S | O | Total |
| Estado de salud | | | | | |
| Excelente | 26.95 | 26.33 | 19.14 | 21.68 | 23.29 |
| Muy buena | 26.76 | 26.01 | 22.82 | 25.18 | 25.07 |
| Regular | 30.26 | 26.52 | 28.29 | 29.14 | 28.43 |
| No muy buena | 12.33 | 15.12 | 18.65 | 17.60 | 16.16 |
| Terrible | 3.69 | 6.02 | 11.11 | 6.40 | 7.05 |
| Total | 100.00 | 100.00 | 100.00 | 100.00 | 100.00 |

# Table with percentages across multiple variables (Tabla con porcentajes para varias variables)

```
. table (var) (región), ///
>  statistic(fvpercent diabetes alta ataque) ///
>  style(table-1) sformat("%s%%")
```

| | Región | | | | |
|---|---|---|---|---|---|
| | NE | MO | S | O | Total |
| Diabetes | | | | | |
| No | 95.32% | 95.49% | 94.36% | 95.62% | 95.18% |
| Sí | 4.68% | 4.51% | 5.64% | 4.38% | 4.82% |
| Presión alta | | | | | |
| No | 55.92% | 59.99% | 57.34% | 57.19% | 57.72% |
| Sí | 44.08% | 40.01% | 42.66% | 42.81% | 42.28% |
| Tuvo ataque al corazón | | | | | |
| No | 96.32% | 95.64% | 95.41% | 94.41% | 95.40% |
| Sí | 3.68% | 4.36% | 4.59% | 5.59% | 4.60% |

# Table with percentages and means (Tabla con porcentajes y medias)

```
. table (var) (región), statistic(mean presión imc edad) ///
>  statistic(fvpercent ataque alta diabetes) ///
>  style(table-1) nformat(%6.2f mean) sformat("%s%%" fvpercent)
```

| | Región | | | | |
|---|---|---|---|---|---|
| | NE | MO | S | O | Total |
| Presión arterial sistólica | 131.38 | 130.49 | 131.16 | 130.59 | 130.88 |
| Índice de masa corporal | 25.58 | 25.52 | 25.63 | 25.42 | 25.54 |
| Edad | 3.39 | 3.25 | 3.41 | 3.38 | 3.36 |
| Tuvo ataque al corazón | | | | | |
| No | 96.32% | 95.64% | 95.41% | 94.41% | 95.40% |
| Sí | 3.68% | 4.36% | 4.59% | 5.59% | 4.60% |
| Presión alta | | | | | |
| No | 55.92% | 59.99% | 57.34% | 57.19% | 57.72% |
| Sí | 44.08% | 40.01% | 42.66% | 42.81% | 42.28% |
| Diabetes | | | | | |
| No | 95.32% | 95.49% | 94.36% | 95.62% | 95.18% |
| Sí | 4.68% | 4.51% | 5.64% | 4.38% | 4.82% |

# Comparison of blood pressure across sex (Comparar la presión arterial basado en el sexo)

```
. ttest presión, by(sexo)

Two-sample t test with equal variances
```

| Group | Obs | Mean | Std. err. | Std. dev. | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| Hombre | 4,915 | 132.8877 | .2994383 | 20.99274 | 132.3007 | 133.4747 |
| Mujer | 5,436 | 129.0679 | .3407989 | 25.12684 | 128.3998 | 129.736 |
| Combined | 10,351 | 130.8817 | .2293364 | 23.33265 | 130.4321 | 131.3312 |
| diff | | 3.81981 | .4577395 | | 2.922552 | 4.717068 |

```
    diff = mean(Hombre) - mean(Mujer)                             t =   8.3449
H0: diff = 0                                     Degrees of freedom =    10349

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

# Returned results (Resultados)

```
. return list

scalars:
                r(level) =  95
                   r(sd) =  23.33265014752308
                 r(sd_2) =  25.12684271524318
                 r(sd_1) =  20.99274059696966
                   r(se) =  .4577395495253661
                  r(p_u) =  4.01647153672e-17
                  r(p_l) =  1
                    r(p) =  8.03294307343e-17
                    r(t) =  8.344941903935108
                 r(df_t) =  10349
                 r(mu_2) =  129.067880794702
                  r(N_2) =  5436
                 r(mu_1) =  132.8876907426246
                  r(N_1) =  4915
```

# Table with results from other Stata commands (Tabla con resultados de otros comandos de Stata)

```
table (command) (result), nformat(%6.2f) nformat(%05.4f p) ///
 command(Hombres=r(mu_1) Mujeres=r(mu_2) Diferencia=r(mu_2)-r(mu_1) r(p): ttest presión, by(sexo)) ///
 command(Hombres=r(mu_1) Mujeres=r(mu_2) Diferencia=r(mu_2)-r(mu_1) r(p): ttest colesterol, by(sexo))
```

| | Hombres | Mujeres | Diferencia | Valor p |
|---|---|---|---|---|
| Presión arterial sistólica | 132.89 | 129.07 | -3.82 | 0.0000 |
| Colesterol | 213.17 | 221.74 | 8.56 | 0.0000 |

# Table with results from other Stata commands (Tabla con resultados de otros comandos de Stata)

```
table (command) (result), nformat(%6.2f) nformat(%05.4f p) ///
 command(Hombres=r(mu_1) Mujeres=r(mu_2) Diferencia=r(mu_2)-r(mu_1) r(p): ttest presión, by(sexo)) ///
 command(Hombres=r(mu_1) Mujeres=r(mu_2) Diferencia=r(mu_2)-r(mu_1) r(p): ttest colesterol, by(sexo))
```

| | Hombres | Mujeres | Diferencia | Valor p |
|---|---|---|---|---|
| Presión arterial sistólica | 132.89 | 129.07 | -3.82 | 0.0000 |
| Colesterol | 213.17 | 221.74 | 8.56 | 0.0000 |

```
collect label levels command 1 "Presión arterial sistólica" 2 "Colesterol", modify
collect label levels result p "Valor p", modify
```

# The table command (El comando table)

**table** (*rows*) (*columns*) (*tables*), **statistic**(*stats*) **command**(*command*) **totals**(*totals*)

| | |
|---|---|
| mean | mean |
| semean | standard error of the mean |
| sebinomial | standard error of the mean, binomial |
| sepoisson | standard error of the mean, Poisson |
| variance | variance |
| sd | standard deviation |
| skewness | skewness |
| kurtosis | kurtosis |
| cv | coefficient of variation |
| count | number of nonmissing values |
| median | median |
| p# | #th percentile |
| q1 | first quartile |
| q2 | second quartile |
| q3 | third quartile |
| iqr | interquartile range |
| min | minimum value |
| max | maximum value |
| range | range |

| | |
|---|---|
| first | first value |
| last | last value |
| firstnm | first nonmissing value |
| lastnm | last nonmissing value |
| total | total |
| rawtotal | unweighted total |
| fvfrequency | frequency of each factor-variable level |
| fvrawfrequency | unweighted frequency of each factor-variable level |
| fvproportion | proportion within each factor-variable level |
| fvrawproportion | unweighted proportion within each factor-variable level |
| fvpercent | percentage within each factor-variable level |
| fvrawpercent | unweighted percentage within each factor-variable level |
| proportion | proportion |
| percent | percentage |
| rawproportion | proportion ignoring optionally specified weights |
| rawpercent | percentage ignoring optionally specified weights |

# Table with means for each level of a categorical variable (Tabla con medias por categorías)

```
. table () región, statistic(mean presión imc)
```

| | Región | | | | |
|---|---|---|---|---|---|
| | NE | MO | S | O | Total |
| Presión arterial sistólica | 131.3836 | 130.4863 | 131.1626 | 130.5936 | 130.8817 |
| Índice de masa corporal | 25.57535 | 25.51936 | 25.63317 | 25.42299 | 25.5376 |

# Table without variable labels in rows
# (Tabla sin etiquetas en los encabezados de las filas)

| | Región | | | | |
|---|---|---|---|---|---|
| | NE | MO | S | O | Total |
| presión | 131.3836 | 130.4863 | 131.1626 | 130.5936 | 130.8817 |
| imc | 25.57535 | 25.51936 | 25.63317 | 25.42299 | 25.5376 |

# Table without variable labels in rows
# (Tabla sin etiquetas en los encabezados de las filas)

| | Región | |
|---|---|---|
| | NE | MO |
| presión | 131.3836 | 130.4863 |
| imc | 25.57535 | 25.51936 |

| Value | Tags | | |
|---|---|---|---|
| 131.383 | var[presión] | región[NE] | result[mean] |
| 130.486 | var[presión] | región[MO] | result[mean] |
| 25.575 | var[imc] | región[NE] | result[mean] |
| 25.519 | var[imc] | región[MO] | result[mean] |

# Dimensions in our collection
# (Dimensiones en nuestra colección)

```
. collect dims

Collection dimensions
Collection: Table
─────────────────────────────────────────
                   Dimension   No. levels
─────────────────────────────────────────
Layout, style, header, label
                      cmdset   1
                     colname   2
                     command   1
                     región   5
                      result   1
                     statcmd   1
                         var   2

Style only
                 border_block   4
                    cell_type   4
─────────────────────────────────────────
```

# Table of regression results

(Tabla de resultados de análisis de regresión)

# Goal: Table of regression results
# (Meta: Tabla de resultados de análisis de regresión)

| | 1 | 2 |
|---|---|---|
| Peso (kg) | 0.42 | 0.43 |
| | (0.01) | (0.02) |
| Diabetes | 14.46 | 12.57 |
| | (1.02) | (1.54) |
| Mujer | | 0.95 |
| | | (0.48) |
| Diabetes x Mujer | | 3.15 |
| | | (2.05) |
| Intercept | 99.97 | 98.52 |
| | (1.04) | (1.24) |
| R-squared | 0.10 | 0.10 |

# Workflow (Flujo de trabajo)

1. Collect results using the collect prefix

   **`collect :`** *`command`*

2. Lay out the table

   **`collect layout`** *`(rows) (columns)`*

3. Customize the table

   **`collect style cell`** …
   **`collect style header`** …
   **`collect style column`** …
   **`collect style row`** …

4. Export the table

   **`collect export`** …

# Exploring our collection (Explorando la colección)

- List the dimensions in our collection

  **`collect dims`**

- List the levels of a dimension

  **`collect levelsof`** …

- List the labels for a dimension

  **`collect label list`** …

- Preview the table after we've customized it

  **`collect preview`**

# Collecting regression results
# (Recolectar resultados de análisis de regresión)

```
. collect create reg1
(current collection is reg1)

. use nhanessp, clear
(Second National Health and Nutrition Examination Survey)

. collect: regress presión peso i.diabetes

      Source |       SS           df       MS      Number of obs   =    10,349
-------------+----------------------------------   F(2, 10346)     =    570.36
       Model |  559427.954         2  279713.977   Prob > F        =    0.0000
    Residual |  5073852.09    10,346  490.416788   R-squared       =    0.0993
-------------+----------------------------------   Adj R-squared   =    0.0991
       Total |  5633280.05    10,348   544.38346   Root MSE        =    22.145

------------------------------------------------------------------------------
     presión | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        peso |   .4203622   .0142106    29.58   0.000     .3925068    .4482176
             |
    diabetes |
   Diabetes  |   14.45636   1.018656    14.19   0.000     12.45959    16.45312
       _cons |   99.96567   1.042461    95.89   0.000     97.92224    102.0091
------------------------------------------------------------------------------
```

# Returned results (Resultados)

```
. return list

scalars:
              r(level) =  95

matrices:
              r(table) :  9 x 4

. matrix list r(table)

r(table)[9,4]
                          0b.         1.
             peso   diabetes   diabetes      _cons
        b  .42036218         0  14.456357  99.965668
       se  .01421055         .  1.0186564  1.0424612
        t  29.580993         .  14.191594    95.8939
   pvalue  1.11e-184         .  2.744e-45          0
       ll  .39250675         .  12.459594  97.922242
       ul   .4482176         .  16.453121  102.00909
       df      10346     10346      10346      10346
     crit  1.9601933  1.9601933  1.9601933  1.9601933
    eform          0         0          0          0
```

Covariate names are the column identifiers, or column names.

Row identifiers are results, or statistics. ->

# Dimensions in our collection
# (Dimensiones en nuestra colección)

```
. collect dims

Collection dimensions
Collection: reg1
─────────────────────────────────────────────
                    Dimension   No. levels
─────────────────────────────────────────────
Layout, style, header, label
                       cmdset   1
                        coleq   1
                      colname   4
                program_class   1
                       result   30
                  result_type   3

Header, label
                     diabetes

Style only
                 border_block   4
                    cell_type   4
─────────────────────────────────────────────
```

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname) (result)

Collection: reg1
      Rows: colname
   Columns: result
   Table 1: 4 x 8
```

| | Coefficient | 95% CI | | df | 95% lower bound | p-value | Std. error | 95% upper bound | t |
|---|---|---|---|---|---|---|---|---|---|
| Peso (kg) | .4203622 | .3925068 | .4482176 | 10346 | .3925068 | 0.000 | .0142106 | .4482176 | 29.58 |
| No diabetes | 0 | | | | | | 0 | | |
| Diabetes | 14.45636 | 12.45959 | 16.45312 | 10346 | 12.45959 | 0.000 | 1.018656 | 16.45312 | 14.19 |
| Intercept | 99.96567 | 97.92224 | 102.0091 | 10346 | 97.92224 | 0.000 | 1.042461 | 102.0091 | 95.89 |

# The levels of dimension result
# (Los niveles de la dimensión result)

```
. collect levelsof result

Collection: reg1
 Dimension: result
    Levels: F N _r_b _r_ci _r_df _r_lb _r_p _r_se _r_ub _r_z cmd cmdline depvar df_m df_r estat_cmd ll ll_0
            marginsok model mss predict properties r2 r2_a rank rmse rss title vce
```

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname) (result[_r_b _r_ci])

Collection: reg1
      Rows: colname
   Columns: result[_r_b _r_ci]
   Table 1: 4 x 2
```

| | Coefficient | 95% CI | |
|---|---|---|---|
| Peso (kg) | .4203622 | .3925068 | .4482176 |
| No diabetes | 0 | | |
| Diabetes | 14.45636 | 12.45959 | 16.45312 |
| Intercept | 99.96567 | 97.92224 | 102.0091 |

# Collect results (Recolectar resultados)

```
. collect: regress presión peso diabetes##sexo

      Source |       SS           df       MS      Number of obs   =    10,349
-------------+----------------------------------   F(4, 10344)     =    287.31
       Model |  563294.124         4  140823.531   Prob > F        =    0.0000
    Residual |  5069985.92    10,344   490.13785   R-squared       =    0.1000
-------------+----------------------------------   Adj R-squared   =    0.0996
       Total |  5633280.05    10,348   544.38346   Root MSE        =    22.139

------------------------------------------------------------------------------
        presión | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
----------------+-------------------------------------------------------------
           peso |   .4335342   .0153559    28.23   0.000     .4034335    .4636348
                |
       diabetes |
      Diabetes  |   12.57211   1.538361     8.17   0.000     9.556621    15.58759
                |
           sexo |
         Mujer  |   .9520999   .4817911     1.98   0.048     .0076962    1.896504
                |
  diabetes#sexo |
Diabetes#Mujer  |   3.146466   2.048958     1.54   0.125    -.8698888     7.16282
                |
          _cons |    98.5238   1.237787    79.60   0.000      96.0975    100.9501
------------------------------------------------------------------------------
```

# Dimensions in our collection
# (Dimensiones en nuestra colección)

```
. collect dims

Collection dimensions
Collection: reg1
─────────────────────────────────────────────
                    Dimension   No. levels
─────────────────────────────────────────────
Layout, style, header, label
                       cmdset   2
                        coleq   1
                      colname   10
                program_class   1
                       result   30
                  result_type   3

Header, label
                     diabetes
                         sexo

Style only
                 border_block   4
                    cell_type   4
─────────────────────────────────────────────
```

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname) (cmdset#result[_r_b _r_ci])

Collection: reg1
      Rows: colname
   Columns: cmdset#result[_r_b _r_ci]
   Table 1: 10 x 4
```

| | 1 | 1 | | 2 | 2 | |
|---|---|---|---|---|---|---|
| | Coefficient | 95% CI | | Coefficient | 95% CI | |
| Peso (kg) | .4203622 | .3925068 | .4482176 | .4335342 | .4034335 | .4636348 |
| No diabetes | 0 | | | 0 | | |
| Diabetes | 14.45636 | 12.45959 | 16.45312 | 12.57211 | 9.556621 | 15.58759 |
| Hombre | | | | 0 | | |
| Mujer | | | | .9520999 | .0076962 | 1.896504 |
| No diabetes # Hombre | | | | 0 | | |
| No diabetes # Mujer | | | | 0 | | |
| Diabetes # Hombre | | | | 0 | | |
| Diabetes # Mujer | | | | 3.146466 | -.8698888 | 7.16282 |
| Intercept | 99.96567 | 97.92224 | 102.0091 | 98.5238 | 96.0975 | 100.9501 |

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname#result[_r_b _r_se]) (cmdset)

Collection: reg1
      Rows: colname#result[_r_b _r_se]
   Columns: cmdset
   Table 1: 30 x 2
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | | |
| Coefficient | .4203622 | .4335342 |
| Std. error | .0142106 | .0153559 |
| No diabetes | | |
| Coefficient | 0 | 0 |
| Std. error | 0 | 0 |
| Diabetes | | |
| Coefficient | 14.45636 | 12.57211 |
| Std. error | 1.018656 | 1.538361 |
| Hombre | | |
| Coefficient | | 0 |
| Std. error | | 0 |
| Mujer | | |
| Coefficient | | .9520999 |
| Std. error | | .4817911 |
| No diabetes # Hombre | | |
| Coefficient | | 0 |

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname#result[_r_b _r_se] result[r2]) (cmdset)

Collection: reg1
      Rows: colname#result[_r_b _r_se] result[r2]
   Columns: cmdset
   Table 1: 31 x 2
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | | |
| Coefficient | .4203622 | .4335342 |
| Std. error | .0142106 | .0153559 |
| No diabetes | | |
| Coefficient | 0 | 0 |
| Std. error | 0 | 0 |
| Diabetes | | |
| Coefficient | 14.45636 | 12.57211 |
| Std. error | 1.018656 | 1.538361 |
| (output omitted) | | |
| Intercept | | |
| Coefficient | 99.96567 | 98.5238 |
| Std. error | 1.042461 | 1.237787 |
| R-squared | .0993077 | .099994 |

# Workflow (Flujo de trabajo)

1. Collect results using the collect prefix

   **`collect :`** *`command`*

2. Lay out the table

   **`collect layout`** *`(rows) (columns)`*

3. Customize the table

   **`collect style cell`** …
   **`collect style header`** …
   **`collect style column`** …
   **`collect style row`** …

4. Export the table

   **`collect export`** …

# Step 3: Customizing the table (Paso 3: Personalizar la tabla)

We would like to

- Hide the base levels
- Format the results to display only two digits after the decimal
- Enclose the standard errors in parentheses
- Remove the border on the right side
- Remove the labels for the coefficients and standard errors
- Change the delimiter

# Base levels and numeric format
# (Nivel de referencia y  formato de número)

```
. collect style showbase off

. collect style cell result, nformat(%5.2f)

. collect preview
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | | |
| Coefficient | 0.42 | 0.43 |
| Std. error | 0.01 | 0.02 |
| Diabetes | | |
| Coefficient | 14.46 | 12.57 |
| Std. error | 1.02 | 1.54 |
| Mujer | | |
| Coefficient | | 0.95 |
| Std. error | | 0.48 |
| Diabetes # Mujer | | |
| Coefficient | | 3.15 |
| Std. error | | 2.05 |
| Intercept | | |
| Coefficient | 99.97 | 98.52 |
| Std. error | 1.04 | 1.24 |
| R-squared | 0.10 | 0.10 |

# Enclose standard errors in parentheses (Colocar error estándar entre paréntesis)

```
. collect style cell result[_r_se], sformat("(%s)")

. collect preview
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | | |
| Coefficient | 0.42 | 0.43 |
| Std. error | (0.01) | (0.02) |
| Diabetes | | |
| Coefficient | 14.46 | 12.57 |
| Std. error | (1.02) | (1.54) |
| Mujer | | |
| Coefficient | | 0.95 |
| Std. error | | (0.48) |
| Diabetes # Mujer | | |
| Coefficient | | 3.15 |
| Std. error | | (2.05) |
| Intercept | | |
| Coefficient | 99.97 | 98.52 |
| Std. error | (1.04) | (1.24) |
| R-squared | 0.10 | 0.10 |

* `%s` refers to the numeric value

# Remove borders and horizontally center results (Eliminar el borde y centrar horizontalmente los resultados)

```
. collect levelsof cell_type

Collection: reg1
 Dimension: cell_type
    Levels: column-header corner item row-header

. collect style cell cell_type[item column-header], halign(center)

. collect style cell border_block[corner row-header], border(right, pattern(nil))

. collect preview
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | | |
| Coefficient | 0.42 | 0.43 |
| Std. error | (0.01) | (0.02) |
| Diabetes | | |
| Coefficient | 14.46 | 12.57 |
| Std. error | (1.02) | (1.54) |
| Mujer | | |
| Coefficient | | 0.95 |
| Std. error | | (0.48) |
| Diabetes # Mujer | | |
| Coefficient | | 3.15 |
| Std. error | | (2.05) |
| Intercept | | |
| Coefficient | 99.97 | 98.52 |
| Std. error | (1.04) | (1.24) |
| R-squared | 0.10 | 0.10 |

# Labels for the the levels of dimension result (Etiquetas para los niveles de la dimensión result)

```
. collect label list result

  Collection: reg1
   Dimension: result
       Label: Result
Level labels:
           F  F statistic
           N  Number of observations
        _r_b  Coefficient
       _r_ci  __LEVEL__% CI
       _r_df  df
       _r_lb  __LEVEL__% lower bound
        _r_p  p-value
       _r_se  Std. error
       _r_ub  __LEVEL__% upper bound
        _r_z  t
         cmd  Command
     cmdline  Command line as typed
      depvar  Dependent variable
```

# Hiding labels for levels of dimension result (Ocultar las etiquetas)

```
. collect style header result, level(hide)

. collect style header result[r2], level(label)

. collect preview
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | 0.42 | 0.43 |
| | (0.01) | (0.02) |
| Diabetes | 14.46 | 12.57 |
| | (1.02) | (1.54) |
| Mujer | | 0.95 |
| | | (0.48) |
| Diabetes # Mujer | | 3.15 |
| | | (2.05) |
| Intercept | 99.97 | 98.52 |
| | (1.04) | (1.24) |
| R-squared | 0.10 | 0.10 |

# Changing the delimiter (Cambiar el delimitador)

```
. collect style column, extraspace(1)

. collect style row stack, delimiter(" x ") spacer

. collect preview
```

| | 1 | 2 |
|---|---|---|
| Peso (kg) | 0.42 | 0.43 |
| | (0.01) | (0.02) |
| Diabetes | 14.46 | 12.57 |
| | (1.02) | (1.54) |
| Mujer | | 0.95 |
| | | (0.48) |
| Diabetes x Mujer | | 3.15 |
| | | (2.05) |
| Intercept | 99.97 | 98.52 |
| | (1.04) | (1.24) |
| R-squared | 0.10 | 0.10 |

# Exporting our table (Exportar la tabla)

- **`collect export`** can be used to export the table to Microsoft Word, Microsoft Excel, PDF, HTML, LaTeX, Markdown, SMCL, and plain text files

```
. collect export report1.xlsx
```

# Exporting our table (Exportar la tabla)

- **`collect export`** can be used to export the table to Microsoft Word, Microsoft Excel, PDF, HTML, LaTeX, Markdown, SMCL, and plain text files

```
. collect export report1.xlsx
```

- If you are creating more extensive reports in Microsoft Word or PDF format, you can include the table in those documents: **`putdocx collect`** and **`putpdf collect`**
  - You can also add a title and notes, set the table width, and more:

    **`collect style putdocx`** and **`collect style putpdf`**

# Exporting our table (Exportar la tabla)

- **`collect export`** can be used to export the table to Microsoft Word, Microsoft Excel, PDF, HTML, LaTeX, Markdown, SMCL, and plain text files

```
. collect export report1.xlsx
```

- If you are creating more extensive reports in Microsoft Word or PDF format, you can include the table in those documents: **`putdocx collect`** and **`putpdf collect`**
  - You can also add a title and notes, set the table width, and more:
    **`collect style putdocx`** and **`collect style putpdf`**

- Similarly, you could include this table in an Excel file, along other Stata results, with the **`putexcel`** *`ul_cell`* **`= collect`** syntax

# Managing styles and collections

# (Manejo de estilos y colecciones)

# Save the layout and style (Guardar el diseño y estilo)

- We can save the layout and style so that we can apply it to other tables:

```
. collect style save tablareg
```

# List the collections we have in memory

(Obtener lista de las colecciones en memoria)

```
. collect dir

Collections in memory
Current: reg1
------------------------
   Name    No. items
------------------------
  Table    10
default    0
   reg1    118
------------------------
```

We can have multiple collections in memory. All the style changes we have made only apply to this current collection.

# Creating another collection (Crear otra colección)

```
. collect create reg2
(current collection is reg2)
. collect _r_b _r_se : regress colesterol i.diabetes i.sexo
. collect _r_b _r_se : regress colesterol diabetes##sexo
(output omitted)
```

Now our current (active) collection is **reg2**. Any results we collect will be stored in this collection. Any style changes we make will be applied to this collection, by default.

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname#result) (cmdset)

Collection: reg2
      Rows: colname#result
   Columns: cmdset
   Table 1: 27 x 2
```

| | 1 | 2 |
|---|---|---|
| No diabetes | | |
| Coefficient | 0 | 0 |
| Std. error | 0 | 0 |
| Diabetes | | |
| Coefficient | 9.763889 | 5.748875 |
| Std. error | 2.256087 | 3.4131 |
| Hombre | | |
| Coefficient | 0 | 0 |
| Std. error | 0 | 0 |
| Mujer | | |
| Coefficient | 8.50188 | 8.162971 |
| Std. error | .9678378 | .9916246 |
| No diabetes # Hombre | | |
| Coefficient | | 0 |
| Std. error | | 0 |

# Apply the style file to this collection

# (Aplicar el estilo a esta colección)

```
. collect style use tablareg, replace layout

Collection: reg2
      Rows: colname#result[_r_b _r_se] result[r2]
   Columns: cmdset
   Table 1: 13 x 2

. collect preview
```

| | 1 | 2 |
|---|---|---|
| Diabetes | 9.76 | 5.75 |
| | (2.26) | (3.41) |
| Mujer | 8.50 | 8.16 |
| | (0.97) | (0.99) |
| Diabetes x Mujer | | 7.13 |
| | | (4.55) |
| Intercept | 212.74 | 212.92 |
| | (0.71) | (0.72) |
| R-squared | 0.01 | 0.01 |

# Add stars for significance and export the table
(Agregar estrellas a los resultados significativos y exportar la tabla)

```
. collect stars _r_p "0.01" "***" "0.05" "**" "0.10" "*", attach(_r_b)

. collect export "report2", as(pdf) replace
```

# Table with summary statistics and postestimation results

(Crear una tabla con porcentajes y resultados a base de análisis de regresión)

# Goal

# (Meta)

reporte3 - Compatibility Mode

Insert Draw Design Layout References Mailings Review View

Probabilidad de tener la presión alta

| | % Población | Prob. | 95% CI |
|---|---|---|---|
| Sex | | | |
| Hombre | 47.49% | 0.471 | [0.458, 0.484] |
| Mujer | 52.51% | 0.379 | [0.367, 0.392] |
| Diabetes | | | |
| No diabetes | 95.18% | 0.418 | [0.409, 0.427] |
| Diabetes | 4.82% | 0.512 | [0.471, 0.553] |
| Edad | | | |
| 20–29 | 22.42% | 0.172 | [0.156, 0.187] |
| 30–39 | 15.66% | 0.284 | [0.262, 0.306] |
| 40–49 | 12.28% | 0.396 | [0.369, 0.423] |
| 50–59 | 12.47% | 0.543 | [0.516, 0.570] |
| 60–69 | 27.64% | 0.579 | [0.561, 0.597] |
| 70+ | 9.53% | 0.665 | [0.635, 0.694] |

# Workflow (Flujo de trabajo)

1. Collect probabilities from margins
   **`collect : margins`** …

2. Collect percentages
   **`table`** …

3. Lay out the table
   **`collect layout`** *`(rows) (columns)`*

4. Customize the table
   **`collect style cell`** …
   **`collect style row`** …
   **`collect label levels`** …

5. Export the table
   **`collect export`** …

# Logistic regression (Regresión logística)

```
. collect create logit
(current collection is logit)

. logit alta i.sexo i.diabetes i.edadcat, nolog

Logistic regression                                     Number of obs =  10,349
                                                        LR chi2(7)    = 1563.35
                                                        Prob > chi2   =  0.0000
Log likelihood = -6267.9897                             Pseudo R2     =  0.1109

------------------------------------------------------------------------------
        alta | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
        sexo |
      Mujer  |  -.4349361   .0431725   -10.07   0.000    -.5195526   -.3503196
             |
    diabetes |
   Diabetes  |   .4424432    .100436     4.41   0.000     .2455922    .6392943
             |
     edadcat |
       30-39 |   .6570507   .0785988     8.36   0.000     .5029998    .8111015
       40-49 |    1.16578   .0802232    14.53   0.000     1.008545    1.323014
       50-59 |   1.768663   .0793167    22.30   0.000     1.613206    1.924121
       60-69 |   1.915989   .0678957    28.22   0.000     1.782916    2.049062
         70+ |   2.285936   .0885412    25.82   0.000     2.112399    2.459474
             |
       _cons |  -1.385992    .058852   -23.55   0.000     -1.50134   -1.270644
------------------------------------------------------------------------------
```

# Collect probabilities from margins
# (Recolectar probabilidades del comando margins)

```
. collect: margins sexo diabetes edadcat

Predictive margins                                      Number of obs = 10,349
Model VCE: OIM

Expression: Pr(alta), predict()
```

| | Margin | Delta-method std. err. | z | P>\|z\| | [95% conf. | interval] |
|---|---|---|---|---|---|---|
| sexo | | | | | | |
| Hombre | .4709147 | .0065918 | 71.44 | 0.000 | .457995 | .4838344 |
| Mujer | .3794736 | .0061375 | 61.83 | 0.000 | .3674444 | .3915029 |
| diabetes | | | | | | |
| No diabetes | .4180688 | .0046332 | 90.23 | 0.000 | .408988 | .4271496 |
| Diabetes | .5121547 | .0209538 | 24.44 | 0.000 | .4710859 | .5532235 |
| edadcat | | | | | | |
| 20-29 | .1715306 | .0078717 | 21.79 | 0.000 | .1561024 | .1869587 |
| 30-39 | .2840992 | .0111984 | 25.37 | 0.000 | .2621507 | .3060477 |
| 40-49 | .3960843 | .013634 | 29.05 | 0.000 | .3693621 | .4228066 |
| 50-59 | .5431238 | .0137738 | 39.43 | 0.000 | .5161276 | .5701199 |
| 60-69 | .5789079 | .0092216 | 62.78 | 0.000 | .5608338 | .596982 |
| 70+ | .6645402 | .015055 | 44.14 | 0.000 | .635033 | .6940474 |

# Segments of the population (Grupos de la población)

```
. table var if e(sample), stat(fvpercent sexo diabetes edadcat) ///
>   name(logit) append
```

| | Factor variable percent |
|---|---|
| Sex=Hombre | 47.49 |
| Sex=Mujer | 52.51 |
| Diabetes=No diabetes | 95.18 |
| Diabetes=Diabetes | 4.82 |
| Edad=20-29 | 22.42 |
| Edad=30-39 | 15.66 |
| Edad=40-49 | 12.28 |
| Edad=50-59 | 12.47 |
| Edad=60-69 | 27.64 |
| Edad=70+ | 9.53 |

**`table`** will store these results in a collection called Table, by default. But we told Stata to append them to the collection logit.

# Dimensions in our collection (Dimensiones en nuestra colección)

```
. collect dims

Collection dimensions
Collection: logit
─────────────────────────────────────────────
                   Dimension   No. levels
─────────────────────────────────────────────
Layout, style, header, label
                      cmdset   2
                     colname   13
                     command   1
               program_class   1
                      result   32
                 result_type   3
                     rowname   10
                     statcmd   1
                         var   10

Header, label
                    diabetes
                     edadcat
                        sexo

Style only
                 border_block   4
                    cell_type   4
─────────────────────────────────────────────
```

# Lay out the table (Diseñar la tabla)

```
. collect layout (colname) (result[fvpercent _r_b _r_ci])

Collection: logit
      Rows: colname
   Columns: result[fvpercent _r_b _r_ci]
   Table 1: 10 x 3
```

| | Factor variable percent | Coefficient | 95% CI | |
|---|---|---|---|---|
| Sex=Hombre | 47.49 | .4709147 | .457995 | .4838344 |
| Sex=Mujer | 52.51 | .3794736 | .3674444 | .3915029 |
| Diabetes=No diabetes | 95.18 | .4180688 | .408988 | .4271496 |
| Diabetes=Diabetes | 4.82 | .5121547 | .4710859 | .5532235 |
| Edad=20-29 | 22.42 | .1715306 | .1561024 | .1869587 |
| Edad=30-39 | 15.66 | .2840992 | .2621507 | .3060477 |
| Edad=40-49 | 12.28 | .3960843 | .3693621 | .4228066 |
| Edad=50-59 | 12.47 | .5431238 | .5161276 | .5701199 |
| Edad=60-69 | 27.64 | .5789079 | .5608338 | .596982 |
| Edad=70+ | 9.53 | .6645402 | .635033 | .6940474 |

# Modify labels and stack row elements
## (Modificar etiquetas y apilar los encabezados de las filas)

```
. collect label levels result fvpercent "% Población" _r_b "Prob.", modify

. collect style row stack, nobinder

. collect preview
```

| | % Población | Prob. | 95% CI | |
|---|---|---|---|---|
| Sex | | | | |
| Hombre | 47.49 | .4709147 | .457995 | .4838344 |
| Mujer | 52.51 | .3794736 | .3674444 | .3915029 |
| Diabetes | | | | |
| No diabetes | 95.18 | .4180688 | .408988 | .4271496 |
| Diabetes | 4.82 | .5121547 | .4710859 | .5532235 |
| Edad | | | | |
| 20-29 | 22.42 | .1715306 | .1561024 | .1869587 |
| 30-39 | 15.66 | .2840992 | .2621507 | .3060477 |
| 40-49 | 12.28 | .3960843 | .3693621 | .4228066 |
| 50-59 | 12.47 | .5431238 | .5161276 | .5701199 |
| 60-69 | 27.64 | .5789079 | .5608338 | .596982 |
| 70+ | 9.53 | .6645402 | .635033 | .6940474 |

# Modify numeric formats and borders
# (Modificar el formato de los números y esconder bordes)

```
. collect style cell result[_r_b _r_ci], nformat(%05.3f) halign(center)

. collect style cell result[_r_ci], cidelimiter(", ") sformat("[%s]")

. collect style cell result[fvpercent], nformat(%5.2f) sformat("%s%%") halign(right)

. collect style cell border_block[corner row-header], border(right, pattern(nil))

. collect preview
```

| | % Población | Prob. | 95% CI |
|---|---|---|---|
| Sex | | | |
| Hombre | 47.49% | 0.471 | [0.458, 0.484] |
| Mujer | 52.51% | 0.379 | [0.367, 0.392] |
| Diabetes | | | |
| No diabetes | 95.18% | 0.418 | [0.409, 0.427] |
| Diabetes | 4.82% | 0.512 | [0.471, 0.553] |
| Edad | | | |
| 20-29 | 22.42% | 0.172 | [0.156, 0.187] |
| 30-39 | 15.66% | 0.284 | [0.262, 0.306] |
| 40-49 | 12.28% | 0.396 | [0.369, 0.423] |
| 50-59 | 12.47% | 0.543 | [0.516, 0.570] |
| 60-69 | 27.64% | 0.579 | [0.561, 0.597] |
| 70+ | 9.53% | 0.665 | [0.635, 0.694] |

# Adding customized tables to extensive documents

(Incluir tablas personalizadas en documentos extensos)

# Create a Word document with a customized table and graph
## (Crear un documento de Word con una tabla y un gráfico)

```
* Create a graph
graph bar (mean) presión, over(edadcat) asyvars legend(rows(2)) scheme(s1color) ///
 blabel(bar) ytitle("") title("Presión promedio")
graph export presion.png, replace

* Create an active .docx file and insert the graph
putdocx begin
putdocx paragraph, halign(center)
putdocx image presion.png, width(4.5)

* Specify the style to be used when exporting the table
collect style putdocx, width(80%) title(Probabilidad de tener la presión alta)
putdocx collect
putdocx save report3.docx, replace
```

# Other capabilities

# (Otras funcionalidades)

# Saving and loading collections (Guardar y usar una colección)

- Save a collection
  - `. collect save filename`

- Load the collection into memory to customize it or create other tables with those results
  - `. collect use filename`

# More on managing collections
# (Más información sobre el manejo de colecciones)

- Copy collections
- Rename collections
- Combine collections
- Drop collections
- Reorganize items in a collection by remapping tags and recoding levels of a dimension

# Saving and loading labels (Guardar y usar etiquetas)

- Save a set of labels
  - `. collect label save filename`
- Load the file with labels for use with other collections
  - `. collect label use filename`

# More on styles (Más información sobre estilos)

- Use predefined styles that are shipped with Stata
- Use your own style file as the default for collections and tables created by **`table`**

  set collect_style

  set table_style
- Specify the levels of a dimension to be automatically displayed

  collect style autolevels

# Where to learn more
## (¿En dónde puedo encontrar más información?)

- Overview of customizable tables
- Stata Customizable Tables and Collected Results Reference Manual
- Stata Base Reference Manual

# References (Referencias)


- McDowell, A., A. Engel, J. T. Massey, and K. Maurer. 1981. Plan and operation of the Second National Health andNutrition Examination Survey, 1976–1980.Vital and Health Statistics1(15): 1–144.

# Thank you! Gracias!