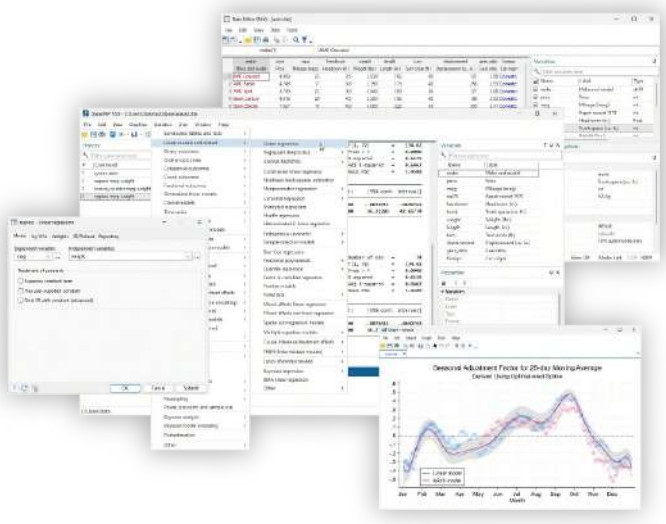
¿Quiénes lo utilizan?

Por más de treinta años, Stata ha sido utilizado por profesionales de todo el mundo en áreas como economía, ingeniería, Ciencia de Datos, finanzas, salud, educación, ciencias sociales, investigación de mercados, análisis de políticas públicas, entre otras, para una amplia gama de tareas y procedimientos específicos que vienen incorporados y validados en el software.

Funcionalidades y Herramientas

Para la mayoría de procedimientos de gestión y análisis de datos, Stata cuenta con menús que le permiten al usuario trabajar de manera gráfica e intuitiva, esta interfaz se encuentra disponible en varios idiomas, entre ellos español, adicionalmente permite la automatización y reproducibilidad de rutinas mediante comandos y la conexión con lenguajes de programación como Python.

Dentro de sus múltiples posibilidades, el software cuenta con una serie de funcionalidades para la gestión y manejo de grandes volúmenes de datos. En general, facilita al usuario el trabajo y la importación de datos desde distintos formatos, entre ellos xls, csv, datos de texto con formato fijo, conexión a fuentes de datos mediante ODBC y JDBC e importación directa de archivos trabajados previamente en SAS y SPSS.



Igualmente, permite la combinación y administración de conjuntos de datos, reportes de duplicados, creación y transformación de variables, agrupación o segmentación de datos mediante características previamente establecidas o bajo condiciones dadas, entre otras tareas. Para conocer las funcionalidades completas haga clic aquí.

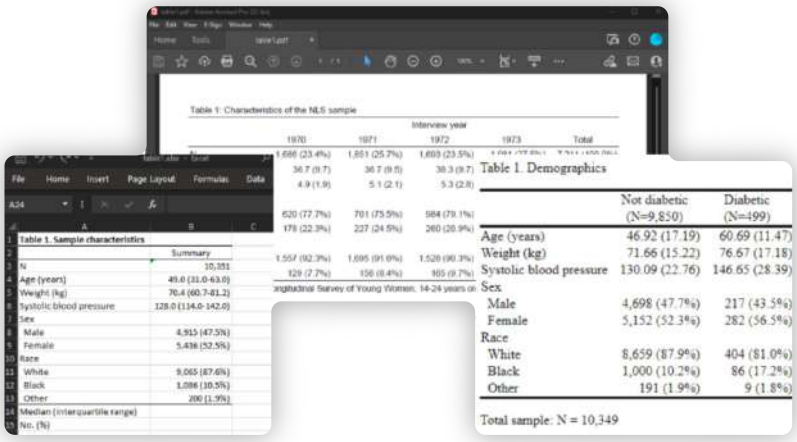
Novedades

En su más reciente versión, presenta un amplio desarrollo de funcionalidades, que abarcan desde herramientas Lasso, procedimientos asociados a la estadística Bayesiana, Machine Learning, análisis causal, mejoras y personalización en la elaboración de informes, creación de tablas, gráficos y exportación de resultados, hasta la integración con lenguajes de programación como Python y Java.

Tablas de estadísticas descriptivas

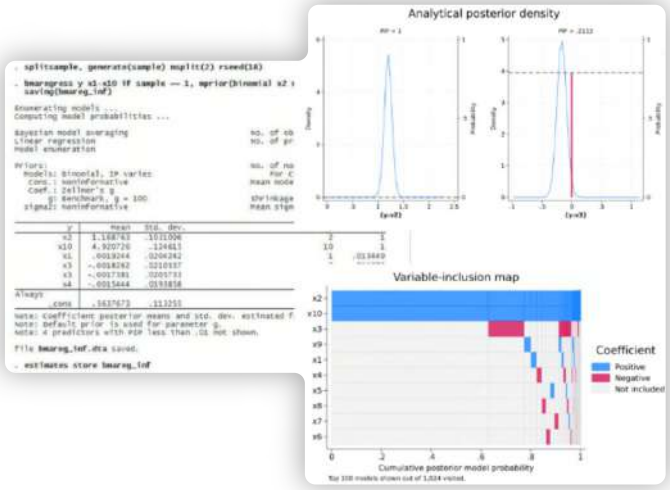
Cree fácilmente su Tabla con el nuevo comando dtable, resuma estadísticas descriptivas, compare información entre grupos y exporte facilmente a Word, Excel, PDF, HTML y más para incluir en informes.

Con dtable , crear una tabla de estadísticas descriptivas puede ser tan fácil como especificar las variables que desea en su tabla, está diseñado para que pueda crear y exportar una tabla a varios formatos en un solo paso. Por supuesto, puede personalizar la tabla incluyendo una variedad de estadísticas y datos relacionados.



Promedio bayesiano del modelo (BMA) para regresión lineal

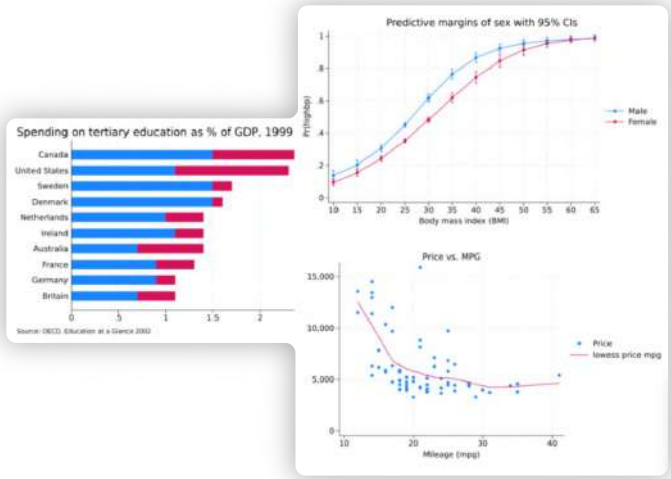
¿Por qué elegir un solo modelo cuando puede tomar información de varios? La nueva suite bma realiza un promedio del modelo bayesiano para tener en cuenta la incertidumbre del modelo en su análisis. ¿No está seguro de qué predictores incluir en su modelo de regresión lineal? con esta nueva opción puede usar bmaregress para averiguar qué predictores son importantes.



¿Por qué elegir un solo modelo cuando puede tomar información de varios? La nueva suite bma realiza un promedio del modelo bayesiano para tener en cuenta la incertidumbre del modelo en su análisis. ¿No está seguro de qué predictores incluir en su modelo de regresión lineal? con esta nueva opción puede usar bmaregress para averiguar qué predictores son importantes.

Gráficos con estilos completamente nuevos

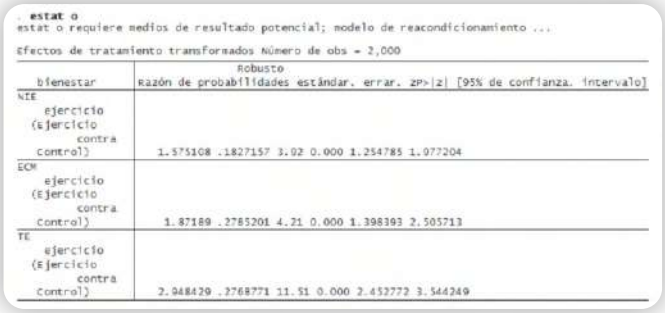
Stata se caracteriza por ofrecer a sus usuarios gráficos y reportes personalizados y de alta calidad para publicación, Stata 18, ofrece lo mejor de ambas opciones, en esta versión se han incluido los nuevos esquemas de gráficos ***stcolor , stcolor_alt , stgcolor y stgcolor_alt.***



Análisis de mediación causal

El nuevo comando mediate amplía la poderosa suite de inferencia causal de Stata para respaldar el análisis de mediación causal. El análisis causal identifica y cuantifica los efectos causales. El análisis de la mediación causal los desenreda. ¿Estos efectos están mediados por otra variable, un mediador?

Elija una de las 23 combinaciones de modelos de resultado y mediador, incluidos lineal, logit y Poisson, para estimar el efecto total y descomponerlo en efectos directos e indirectos (a través del mediador). Calcule los efectos directos controlados y mediados por proporciones. Reformule los efectos en proporciones de probabilidad, riesgo y tasa de incidencia, grafique los efectos estimados, obtenga predicciones y mucho más.



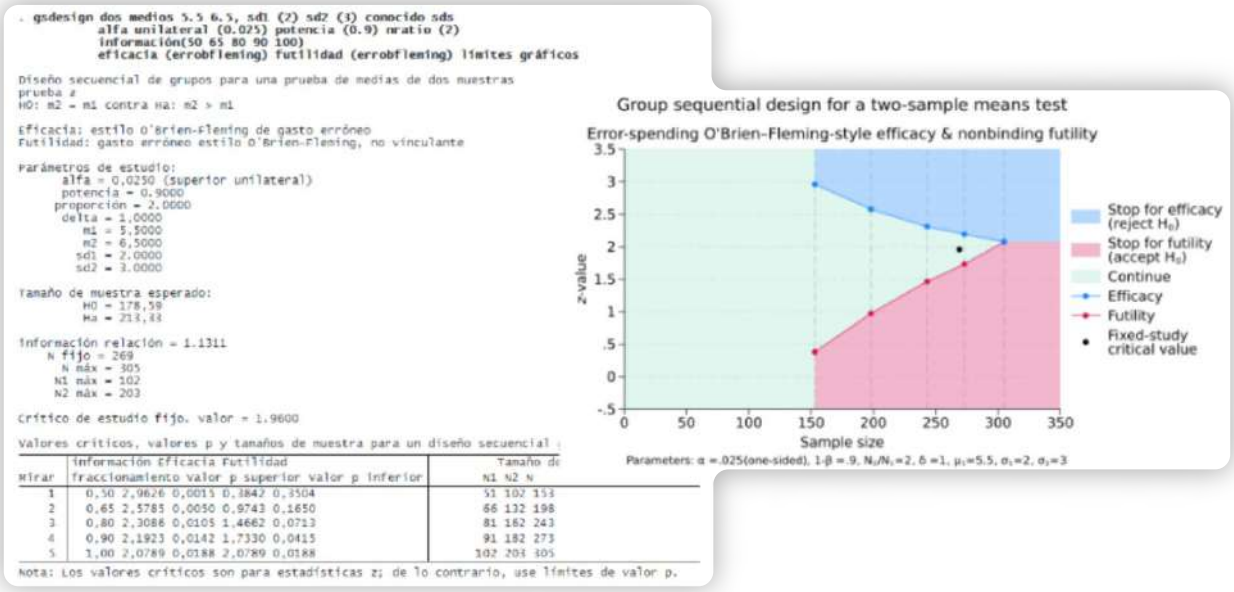
```
. estat o
estat o requiere medios de resultado potencial; modelo de reacondicionamiento ...

Efectos de tratamiento transformados Número de obs = 2,000
```

	Robusto				
bienestar	Razón de probabilidades	estándar. errar.	z	P>|z|	[95% de confianza. intervalo]
NIE					
ejercicio (Ejercicio contra Control)	1.575108	.1827157	3.92	0.000	1.254785 1.977204
ECM					
ejercicio (Ejercicio contra Control)	1.87189	.2785201	4.21	0.000	1.398393 2.505713
TE					
ejercicio (Ejercicio contra Control)	2.948429	.2768771	11.51	0.000	2.452772 3.544249

Diseños secuenciales grupales para ensayos clínicos

¿Diseñando un ensayo clínico? En Stata 18, puede usar los nuevos comandos ***gsbounds*** y ***gsdesign*** para calcular los límites de parada para ensayos secuenciales grupales.

gsdesign calcula límites de parada y tamaños de muestra para análisis intermedios con pruebas de medias, proporciones, funciones de supervivencia e incluso métodos definidos por el usuario, por su parte, ***gsbounds*** calcula los límites de eficacia y futilidad en función de la cantidad de análisis, también llamados miradas, el error de tipo I general deseado y la potencia deseada.



```
. gsdesign dos medios 5.5 6.5, sd1 (2) sd2 (3) conocido sds
          alfa unilateral (0.025) potencia (0.9) oratio (2)
          información(50 65 80 90 100)
          eficacia (errorbfleming) futilidad (errorbfleming) límites gráficos

Diseño secuencial de grupos para una prueba de medias de dos muestras
prueba z
HO: m2 = m1 contra Ha: m2 > m1

Eficacia: estilo O'Brien-Fleming de gasto erróneo
Futilidad: gasto erróneo estilo O'Brien-Fleming, no vinculante

Parámetros de estudio:
      alfa = 0,0250 (superior unilateral)
     potencia = 0,9000
    proporción = 2,0000
     delta = 1,0000
        m1 = 5,5000
        m2 = 6,5000
       sd1 = 2,0000
       sd2 = 3,0000

Tamaño de muestra esperado:
        HO = 178,59
        Ha = 213,83

Información relación = 1,1311
    N fijo = 269
     N máx = 305
    N1 máx = 102
    N2 máx = 203

Crítico de estudio fijo. valor = 1,9600

Valores críticos, valores p y tamaños de muestra para un diseño secuencial
```

	Información	Eficacia	Futilidad		Tamaño de		
Mirar	fraccionamiento	valor p superior	valor p inferior		N1	N2	N
1	0,50	2,9626	0,0015	0,1842	0,3504	51	102 153
2	0,65	2,5785	0,0050	0,9743	0,1650	66	132 198
3	0,80	2,3086	0,0105	1,4662	0,0713	81	162 243
4	0,90	2,1923	0,0142	1,7330	0,0415	91	182 273
5	1,00	2,0789	0,0188	2,0789	0,0188	102	203 305

```
Nota: Los valores críticos son para estadísticas z; de lo contrario, use límites de valor p.
```

Metanálisis multinivel

Stata 18 agrega dos nuevos comandos, ***meta meregress*** y ***meta multilevel.*** Incluya intersecciones aleatorias y coeficientes en diferentes niveles de jerarquía, y suponga diferentes estructuras de covarianza de efectos aleatorios, incluidas las intercambiables y las no estructuradas.

Realice un análisis de sensibilidad colocando varias restricciones en los componentes de varianza, evalúe la heterogeneidad, prediga efectos aleatorios y sus errores estándar comparativos y de diagnóstico.

```
. meta meregress stdmdiff || district: || school: , essevariable(se)

Performing EM optimization ...

Performing gradient-based optimization:
Iteration 0:  Log restricted-likelihood =  -104.8525  (not concave)
Iteration 1:  Log restricted-likelihood = -46.670529  (not concave)
Iteration 2:  Log restricted-likelihood = -22.871266  (not concave)
Iteration 3:  Log restricted-likelihood = -12.977299
Iteration 4:  Log restricted-likelihood = -7.9642885
Iteration 5:  Log restricted-likelihood = -7.9587271
Iteration 6:  Log restricted-likelihood = -7.9587239
Iteration 7:  Log restricted-likelihood = -7.9587239

Computing standard errors ...

Multilevel REML meta-analysis                     Number of obs = 56

Grouping information

------------------------------------------------------------------------------
                |      No. of       Observations per group
 Group variable |     groups    Minimum    Average    Maximum
----------------+-------------------------------------------------------------
       district |         11          3        5.1         11
         school |         56          1        1.0          1
------------------------------------------------------------------------------

                                                  Wald chi2(0)  =      .
Log restricted-likelihood = -7.9587239            Prob > chi2   =      .

------------------------------------------------------------------------------
    stdmdiff |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   .1847132   .0845559     2.18   0.029     .0189866    .3504397
------------------------------------------------------------------------------
Test of homogeneity: Q_M = chi2(55) = 578.86              Prob > Q_M = 0.0000

------------------------------------------------------------------------------
  Random-effects parameters |   Estimate
----------------------------+-------------------------------------------------
district: Identity          |
                  sd(_cons) |   .2550724
----------------------------+-------------------------------------------------
school: Identity            |
                  sd(_cons) |   .1809324
------------------------------------------------------------------------------
```

Exceso de riesgo relativo debido a la interacción (RERI)

¿Cómo interactúan las exposiciones para aumentar el riesgo? ¿Sospecha que la interacción es aditiva? Use reri para averiguarlo. Se proporcionan tres medidas de interacciones bidireccionales: RERI, AP y SI. Se admiten muchos modelos que estiman el RR, incluidos el logístico, el lineal generalizado binomial y el de supervivencia.

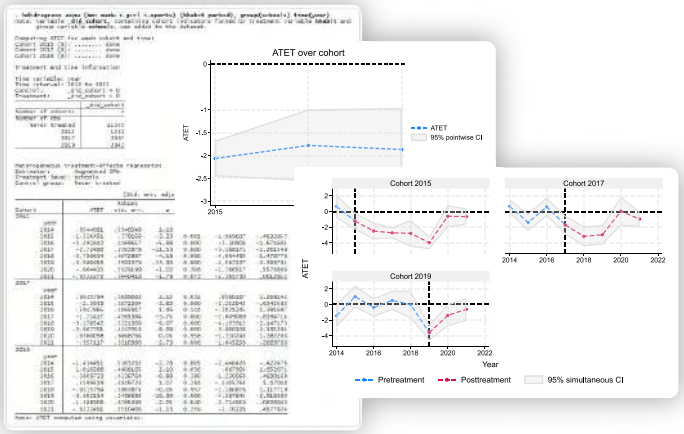
El nuevo comando reri estima interacciones aditivas en modelos lineales generalizados binomiales; regresiones logísticas, de Poisson y binomial negativa; y Cox y otros modelos de supervivencia.

```
. reri binreg tubo drogas nitrato

Montaje binreg...

Interacción de droga y nitrato en una escala aditiva

Modelo: Binomial lineal generalizado Número de observaciones = 1,284

droga#nitrato
        - + No#Alto
        + - Sí#Bajo
        + + Sí#Alto

------------------------------------------------------------------------------
                |                 EIM
                |        ERR      est.      errar.     zP>|z|     [95% de confianza, intervalo]
----------------+-------------------------------------------------------------
droga#nitrato   |
        - +     |    .1161419   .1571024     0.78   0.435    -.1529507    .4707206
        + -     |    .2385481   .2392316     1.11   0.268    -.1517982    .8085336
        + +     |    .7769424    .407042     2.51   0.012     .1342      1.783922
----------------+-------------------------------------------------------------
           RERI |    .4222523   .453286      0.93   0.352    -.466172     1.310677
Atr. apuntalar. |    .2376286   .2187062     1.09   0.277    -.1910276    .6662848
Sinergia ind.   |    2.190483   2.034328     0.84   0.398     .3548285    13.52263
------------------------------------------------------------------------------
Nota: P>|z| para el índice de sinergia (SI) es para la prueba H0: SI = 1.
```

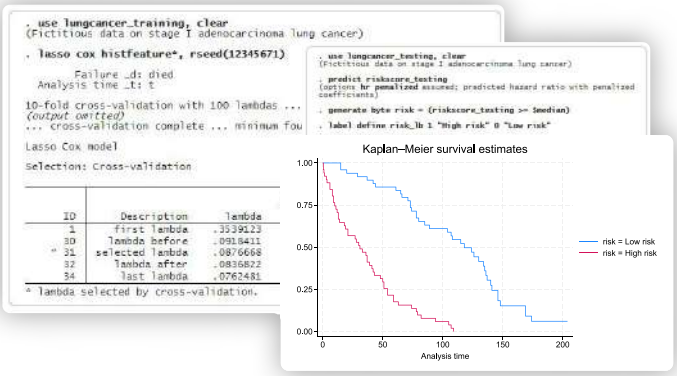
Diferencia heterogénea en diferencias (DID)

Cuando los efectos de tratamiento promedio varían con el tiempo y la cohorte, ahora puede usar los nuevos comandos ***hdidregress*** y ***xthdidregress*** para estimar los efectos de tratamiento promedio heterogéneos en los tratados (ATET). Utilice ***hdidregress*** con datos transversales repetidos y ***xthdidregress*** con datos de panel.

Elija uno de los cuatro estimadores, incluido el ajuste de regresión y la ponderación de probabilidad inversa, trace perfiles de tiempo de ATETs para cada cohorte con ***estat atetplot.*** Agregue los ATETs dentro de la cohorte, el tiempo y la exposición al tratamiento con ***estat aggregation,*** además explore más funciones de postestimación.



Lasso para modelos de riesgos proporcionales de Cox



Si tiene datos de tiempo de supervivencia o tiempo de falla, y muchos predictores, consulte los comandos ***lasso cox*** y ***elasticnet cox*** que amplían la suite ***lasso*** existente para predicción y selección de modelos para incluir un modelo semiparamétrico de riesgos proporcionales de Cox de alta dimensión.

Después de ***lasso cox*** y ***elasticnet cox,*** puede usar stcurve para trazar la función de supervivencia, falla, riesgo o riesgo acumulativo o usar cualquiera de las otras herramientas de postestimación disponibles después de ***lasso*** y ***elasticnet.***

Cluster bootstrap

¿Sus datos tienen un pequeño número de conglomerados o un número impar de observaciones por conglomerado? ¿Quiere hacer inferencias sobre parámetros en un modelo lineal? Con el nuevo comando ***wildbootstrap,*** puede usar wild cluster bootstrap (WCB) para resolver estas situaciones.

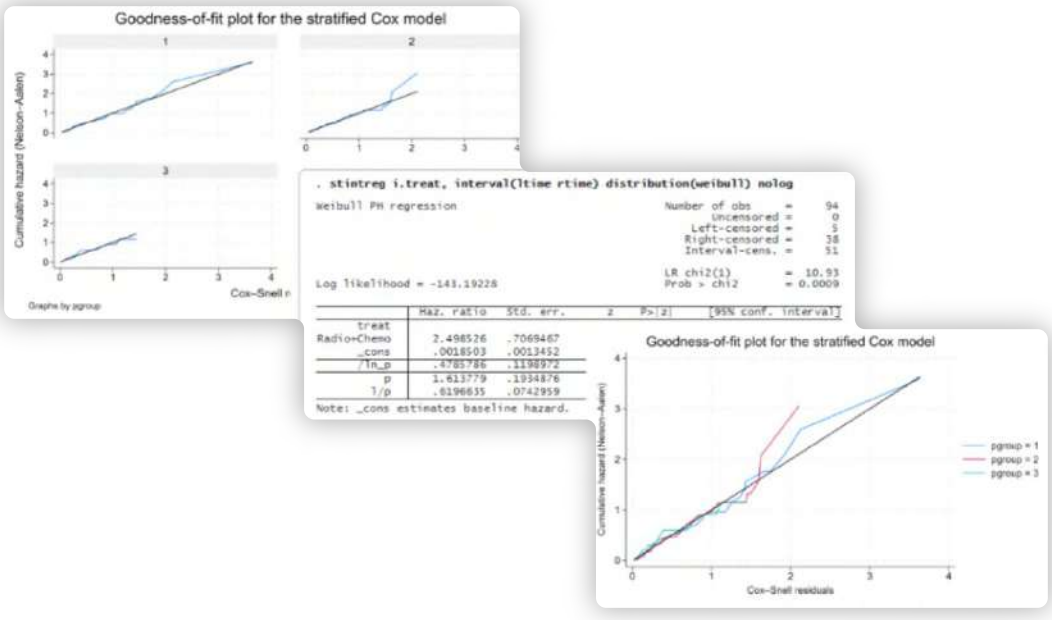
```
. webuse nlsw88
(NLSW, 1988 extract)

. wildbootstrap regress wage tenure, cluster(industry) rseed(12345)

Performing 1,000 replications for p-value for tenure = 0 ...
Computing confidence interval for tenure
  Lower bound: .........10.........20..... done (26)
  Upper bound: .........10.........20.... done (24)

Wild cluster bootstrap                          Number of obs      = 2,217
Linear regression                               Number of clusters =    12
                                                Cluster size:
Cluster variable: industry                                  min =     4
Error weight: Rademacher                                    avg = 184.8
                                                            max =   817
```

wage	Estimate	t	p-value	[95% conf.	interval]
constraint					
tenure = 0	.1830716	6.95	0.000	.1274023	.3258156

Gráficas de bondad de ajuste para modelos de supervivencia

Stata 18 proporciona el nuevo comando ***estat gofplot*** para producir diagramas de bondad de ajuste (GOF) para modelos de supervivencia.

Puede usarlo después de cuatro modelos de supervivencia: Cox censurado por la derecha ***(stcox),*** Cox censurado por intervalos ***(stintcox),*** paramétrico censurado por la derecha ***(streg)*** y paramétrico censurado por intervalos ***(stintreg).*** Verifique el ajuste del modelo después de los modelos estratificados o por separado para cada grupo. Los diagramas GOF proporcionan comprobaciones visuales de qué tan bien se ajusta el modelo a los datos.



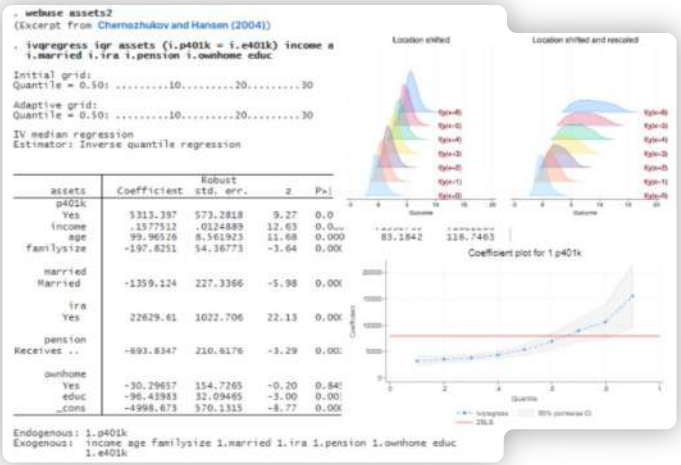
Alias variables across frames

Stata admite múltiples conjuntos de datos en la memoria, cada conjunto de datos reside en un ***marco*** o ***frame.*** En Stata 18, ahora puede trabajar con variables de diferentes marcos como si existieran en uno.

Las variables de alias, creadas por el nuevo comando ***fralias add,*** definen referencias a variables en marcos vinculados. Estas variables ocupan muy poca memoria porque las observaciones en realidad se almacenan en otro marco. Stata trata las variables de alias como cualquier otra variable en su conjunto de datos, con la excepción de que no puede cambiar sus valores.

```
. webuse persons

. describe

Contains data from https://www.stata-press.com/data/r18/persons.dta
 Observations:            20
    Variables:             3                  16 Apr 2022 13:36
                                              (_dta has notes)
-------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------
personid        byte    %9.0g
countyid        byte    %9.0g
income          float   %9.0g
-------------------------------------------------------------------------------
Sorted by:

. frame create txcounty

. frame txcounty: webuse txcounty
(Median income in Texas counties)

. frame txcounty: describe

Contains data from https://www.stata-press.com/data/r18/txcounty.dta
 Observations:             8                  Median income in Texas counties
    Variables:             2                  30 Dec 2022 06:13
                                              (_dta has notes)
-------------------------------------------------------------------------------
Variable      Storage   Display    Value
    name         type    format    label      Variable label
-------------------------------------------------------------------------------
countyid        byte    %9.0g      cty        County ID
median_income   float   %9.0g                 Household median income
-------------------------------------------------------------------------------
Sorted by:
```

Regresión cuantil de variables instrumentales



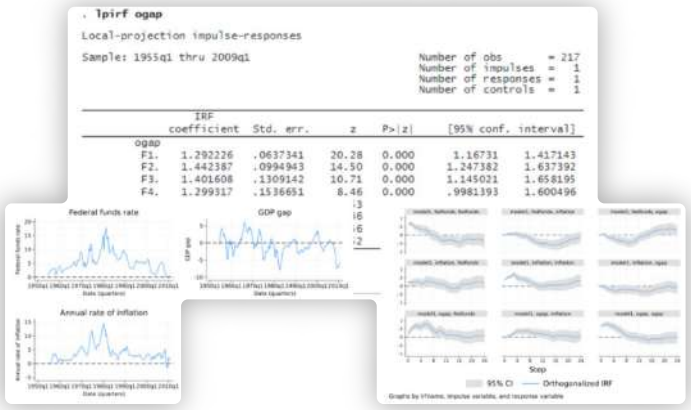
Cuando queremos estudiar los efectos de las covariables en diferentes cuantiles del resultado, usamos la regresión por cuantiles. Pero ¿y si sospechamos que una covariable es endógena? El nuevo comando ***ivqregress*** permite modelar los cuantiles del resultado y, al mismo tiempo, controla los problemas que surgen de la endogeneidad.

Selección de modelos para ARIMA y ARFIMA

¿Quiere encontrar el mejor modelo ARIMA o ARFIMA para sus datos? Compare modelos potenciales utilizando AIC, BIC y HQIC. Utilice los nuevos comandos arimasoc y arfimasoc para seleccionar la mejor cantidad de términos autorregresivos y de promedio móvil.

Proyecciones locales para funciones de impulso-respuesta

Con las funciones de impulso-respuesta, puede averiguar cómo una perturbación en una variable afecta a otras variables a lo largo del tiempo. Con las proyecciones locales, puede estimar funciones de impulso-respuesta directamente mediante regresiones de varios pasos. Utilice el nuevo comando ***lpirf*** para estimar las proyecciones locales y graficarlas o tabularlas con la suite ***irf.***



Encuentre también en Stata 18

- TVC con modelo Cox censurado por intervalos
- Nuevas funciones de spline
- Metanálisis de prevalencia
- Estimación de sistemas de demanda flexible
- Integración numérica vectorizada
- Modelo probit fraccional para variables instrumentales
- Mejoras en el editor de archivos Do
- Mejoras en el editor de datos
- Y mucho más…

Especificaciones técnicas

	Stata/MP* La edición más rápida de Stata para base de datos grandes	Stata/SE La edición de Stata para bases de datos más grandes	Stata/BE La edición de Stata para base de datos medianas
# observaciones	20 Billones	2,14 Billones	2,14 Billones
# de variables	120,000	32,767	2,048
Máximo # de variables independientes	65,532	10,998	798
# de caracteres en un comando	15,480,216	4,227,159	264,408

* Disponible para procesadores de 2, 4, 6, 8, 10, 12, 16, 24, 32, 64 núcleos.

Sistemas operativos compatibles

Stata para Windows®

- Windows 10 u 11 *
- Windows Server 2022, 2019, 2016, 2012R2*

Stata para Mac®

- MacOS 10.13 (High Sierra) o más reciente para Mac con procesadores Intel y macOS 11.0 (Big Sur) o más reciente para Mac con Apple Silicon

Stata para Linux

- Cualquier 64-bit (x86-64 o compatible) ejecutando Linux
- Los requisitos mínimos incluyen la biblioteca GNU C (glibc) 2.17 o superior y libcurl4

STATA 18

Nota:

Cualquier inquietud adicional o problemas con la instalación del software, podrá comunicarse con el siguiente contacto:

Área de Soporte
Software Shop
Teléfono: Según País + Ext. 148
Soporte@Software-Shop.com

Argentina
+54 (11) 5077 9516
Brasil
+55 (21) 9357 1215
Chile
+56 (2) 656 2790
Colombia
+57+60+1 619 4000
México
+52 (555) 351 1755
Perú
+51 (1) 706 8197
USA
+1 (425) 996 0636
Venezuela
+58 (212) 335 0588

Manténgase conectado con nosotros