

NOTA METODOLÓGICA

Tutorial sobre redacción de informes automatizados en Salud Pública con Stata

Francisco Rodríguez-Cabrera^a, Elena Vanessa Martínez Sánchez^{a,b}, Carlos Fernández Escobar^a

^a Unidad Docente de Medicina Preventiva y Salud Pública, Instituto de Salud Carlos III

^b CIBER en Epidemiología y Salud Pública. Instituto de Salud Carlos III

Contacto para correspondencia: fdrodriguez@isciii.es

RESUMEN

Recientemente, salía la versión 15 del programa estadístico Stata, usado con frecuencia por profesionales de salud pública. Como novedad en este software comercial (no así en R, donde ya se podía), es posible generar informes automatizados. Consideramos que su aplicación podría ahorrar una carga relevante de trabajo repetitivo en un sector de profesionales en el campo. En esta nota, explicamos nuestra experiencia en el Centro Nacional de Epidemiología con esta funcionalidad, así como métodos y consejos para replicar el trabajo realizado.

Introducción

La realización de informes epidemiológicos es una actividad fundamental en el campo de la salud pública. Es imprescindible como resumen final de un estudio, de la vigilancia de una o más enfermedades, o incluso como descripción de inicio en una investigación, con la que desarrollar las hipótesis posteriores de estudio. En ellos se resume las características de una población de forma ordenada, lo que ha requerido hasta ahora un gran trabajo, no sólo de análisis e interpretación, sino también de maquetación posterior.

Uno de los software estadísticos más usados en salud pública es Stata®, cuya última versión, la 15, está disponible desde junio de 2017 (1). Una de las novedades más destacables en esta versión parece haber pasado algo desapercibida y es una nueva función que podría ser de gran utilidad: la generación de informes automatizados. Esta posibilidad, que ya existía en el software libre R(2), permite ahorrar una carga importante de trabajo en profesionales que utilizan Stata y realizan informes periódicos similares entre sí.

Para facilitar la comprensión de esta nota, se adjunta una sintaxis para Stata (en un archivo .do).

Exposición del método

Generación de informes con putdocx

El comando oficial *putdocx* crea documentos en Microsoft Word(3). La generación del informe comienza por *putdocx begin*, que abre un documento de texto en la memoria. Una vez generado, se debe crear un párrafo a partir de *putdocx paragraph* (aquí podemos elegir si queremos que el texto o imagen esté centrado o, como está por defecto, alineado a la izquierda, y otras opciones como la tipografía de la fuente y su tamaño). Posteriormente,

se añade el texto, tabla o imagen que queramos a través de *putdocx text("Aquí nuestro texto")*, *putdocx table* y *putdocx image aquinuestraimagen.png*, respectivamente.

Dentro de *putdocx text*, es posible incorporar los cálculos y macros que se ha redactado previamente en Stata. Esto permite diseñar algoritmos “inteligentes” que utilicen la información numérica de las bases de datos para redactar textos con la descripción de los resultados e incluso la interpretación de los mismos. El texto puede personalizarse con diferentes tipografías, tamaños de letra, colores, alineaciones, etc. Por ejemplo, en *putdocx text("Gaceta Sanitaria")*, *font(Arial,12) bold linebreak* escribimos el texto en letra Arial, tamaño 12, negrita, y hacemos un salto de línea entre este texto y el siguiente.

Para finalizar, debe recordarse guardarlo con *putdocx save nombredemiarchivo.docx*.

putdocx table

Mención pausada requiere *putdocx table*. Con este comando podemos exportar una tabla de análisis estadísticos realizada en Stata, o crear tablas nuevas con los datos que queramos. En el primer caso, antes de exportar la tabla es recomendable dar un formato correcto al análisis estadístico en Stata. Una vez hecho, se incorporará al documento a través de *putdocx table nombredemitabla = **etable***

No siempre, eso sí, nos sale una tabla a la primera. Si se quiere modificar, o crear tablas desde cero, es necesario conocer primero cómo sitúa Stata las celdas de una tabla. Se trata de dos coordenadas: la primera corresponde a la fila (“cuánto bajamos en la tabla”), y la segunda a la columna (“cuánto nos desplazamos hacia la derecha”) (ver tabla 1).

Conociendo esto, podemos crear una tabla de 3 filas y 4 columnas a partir de *putdocx table nombredemitabla = (3,4)*. Será posible incluir texto o números en una celda concreta, a partir de *putdocx table nombredemitabla(2,3)=("Holi, Gaceta Sanitaria")*, y eliminar una fila entera a partir de *putdocx table (1,.) drop*. Fíjese en que la colocación de un punto dentro del paréntesis se podría traducir como “en todos los números de la coordenada que tiene mi tabla” (ver tabla 2).

putdocx table incorpora muchas opciones de formato de tablas y celdas. Podemos elegir la alineación vertical y horizontal del texto, el modo de ajuste del tamaño al contenido, el color del sombreado, el aspecto de los bordes y las líneas divisorias, o fusionar celdas de la misma columna o fila. También podemos definir un título y un pie de página, e incluso introducir imágenes dentro de las celdas.

putdocx image

A partir de *putdocx image* podemos incorporar las gráficas que hemos generado en Stata o en cualquier otro programa. Los formatos de imagen compatibles son .jpg y .png. Además, la imagen debe estar en el mismo lugar donde está el directorio de Stata; de lo contrario, se ha de escribir su ubicación (*putdocx image C:\Descargas\miimagen.jpg*, por ejemplo).

Adjuntar los documentos que ya hemos hecho

Además, es posible unir documentos ya existentes (supóngase que se ha redactado introducción y métodos, y se quiere añadir la sección de resultados de forma automatizada).

Esta unión se realiza en orden, debiendo estar el primero de los documentos en nuestro texto (en este caso, *introducción.docx*) al inicio del comando.

```
putdocx append introduccion.docx materialymetodos.docx resultados.docx,  
saving("borrador_automatizado.docx")
```

Para conocer más: Existen en Stata comandos no oficiales como *putwrap*, *markstat*, *textdoc(4)* y *markdown*, que podrían ser complementarios a *putdocx* en la automatización de documentos, y especialmente útiles para aquellas personas que tengan versiones de Stata previas a la 15. También existe la posibilidad de generar PDF, a través de *putpdf*.

Aplicación práctica

Con esta metodología se realizaron parcialmente modelos de informes para el Centro Nacional de Epidemiología (CNE). Entre ellos, el informe de la Red Nacional de Vigilancia Epidemiológica (RENAVE) para la Tuberculosis de 2016 (ver figura 1) (5), o los resultados de vigilancia en los últimos años para diferentes enfermedades. A partir de los datos de casos individualizados y brotes notificados a la RENAVE, una sintaxis de Stata devolvía un informe automático con los siguientes apartados:

- Título, enfermedad y periodo estudiado, imágenes corporativas de la RENAVE y el CNE.
- Resumen de los resultados. Número de casos y brotes notificados, porcentaje de casos confirmados, probables, sospechosos o importados, evolución de la tasa anual, distribución estacional y meses con más y menos casos, distribución geográfica, y grupos de edad y sexo con más casos.
- Tablas con el número de casos y brotes por CCAA y año, y con el alcance y ámbito de los brotes. En las enfermedades de transmisión alimentaria, el alimento implicado.
- Gráficos de líneas de tasa de incidencia anual y estacional, gráficos de barras de casos por grupos de edad y sexo, mapas coropléticos de tasa de incidencia por CCAA y de símbolos proporcionales para el número de brotes.

Para ello, se utilizaron los comandos descritos de Stata, junto con comandos de análisis estadístico y bucles lógicos. Por ejemplo, un fragmento de sintaxis contaba el número de casos en hombres y en mujeres, los comparaba y decidía entre escribir “*habido más casos notificados en hombres (x%) que en mujeres (y%)*”, o lo contrario. Este sistema permite la actualización casi inmediata de los informes después de un cambio en los datos de referencia.

Conclusiones

Dada la situación actual, con un gran déficit de personal en las instituciones que se dedican a la salud pública, toda herramienta que maximice la eficiencia del tiempo invertido en el trabajo es esencial. Esta funcionalidad podrá agilizar los análisis rutinarios, independientemente de la periodicidad necesaria, permitiendo utilizar el tiempo ahorrado en otras actividades que, aunque necesarias, no se realizaban hasta el momento.

Contribuciones de autoría

Todas las personas autoras han participado en la concepción, diseño y elaboración de este artículo. Específicamente, FRC y CFE participaron en la conceptualización y diseño de la sintaxis del programa estadístico que originó la idea del artículo. EVMS realizó la revisión de la idea y sus aplicaciones. FRC y EVMS elaboraron un primer borrador del manuscrito, al que CFE realizó amplios aportes posteriores. Todas las personas autoras participaron en el proceso de revisión del artículo, y han leído y aprobado la versión final.

Financiación

Ninguna.

Conflicto de intereses

Ninguno.

Referencias

1. Gould W, Cox NJ. History of Stata. Stata Official Website. Disponible en <https://www.stata.com/support/faqs/resources/history-of-stata/>
2. RStudio Inc. Shiny package. Disponible en: <https://shiny.rstudio.com/>
3. Manual de Stata. putdocx — generate office open XML (.docx) file. Disponible en: <https://www.stata.com/manuals/pputdocx.pdf>
4. Jann, B. 2016. Creating LaTeX documents from within Stata using texdoc. Stata Journal 16: 245–263.
5. Centro Nacional de Epidemiología. Instituto de Salud Carlos III. Informe epidemiológico sobre la situación de la tuberculosis en España. Año 2016. Disponible en: http://www.isciii.es/ISCIII/es/contenidos/fd-servicios-cientifico-tecnicos/fd-vigilancias-alertas/fd-enfermedades/pdf_2019/RENAVE_TBC_2016.pdf

Tabla 1. Correlación de posiciones de celdas con *putdocx*.

(1,1)	(1,2)	(1,3)	(1,4)
(2,1)	(2,2)	(2,3)	(2,4)

Tabla 2. Esquema de cómo quedaría mi tabla finalmente.

BORRADA	BORRADA	BORRADA	BORRADA
(2,1)	(2,2)	Holi, Gaceta Sanitaria	(2,4)
(3,1)	(3,2)	(3,3)	(3,4)