

RESPUESTAS MÚLTIPLES EN LA INVESTIGACIÓN EDUCATIVA: CODIFICACIÓN, TABULACIÓN Y ANÁLISIS

José Serrano Angulo
Universidad de Málaga

RESUMEN

La codificación y tabulación de preguntas de respuestas múltiples es bien simple y conocida, pero no por ello está exenta de errores y laboriosas tareas de introducción de datos y análisis de los mismos. En este artículo se expone como facilitar la codificación y la tabulación de este tipo de variables, suponiendo un ahorro drástico del tiempo necesario para su tabulación y una reducción considerable de los errores de datos que se pueden producir al introducirlos en el ordenador. Asimismo, se obtiene una variable "viva" que se puede analizar y obtener información adicional. Además, se indica como obtener mayor información del estudio conjunto de más de dos variables dicotómicas a partir de la definición de una nueva variable conjunta en la matriz de datos. Todo ello se aplica a datos de una investigación sobre las transversales en educación primaria.

***Palabras claves:** multi-respuestas, variables dicotómicas, codificación, tabulación, materias transversales.*

MULTIPLE-CHOICE QUESTIONS IN EDUCATIONAL RESEARCH: CODING, TABULATION AND ANALYSIS

ABSTRACT

The process of coding and tabulating multiple-choice questions is simple and well known, yet not free from errors and time-consuming data entry and analysis. This article explains how

Correspondencia:

José Serrano Angulo, Departamento de Métodos de Investigación e Innovación Educativa. Facultad de Ciencias de la Educación. Universidad de Málaga. E-mail: joseserrano@uma.es.

to code and tabulate variables, which will save time and dramatically decrease the number of errors derived from entering data into a computer. An "active" variable is calculated which may be further analyzed. This article also explains how to obtain more information when analyzing more than two dichotomous variables from the creation of a new joint variable in the data matrix. These processes were applied to data of a research study on cross-curricular subjects in primary education.

Keywords: *multiple-choice questions, dichotomous variables, coding, tabulation, cross-curricular subjects.*

I. INTRODUCCIÓN

Una pregunta de respuesta múltiple se plantea de tal modo que el sujeto puede elegir más de una de las opciones que se le ofrece, o todas. Esto hace que estas preguntas se desglosen para su tabulación en tantas variables dicotómicas como opciones se planteen. Tabulando cada variable dicotómicas por cero, si no es elegida la opción correspondiente, o por uno si es elegida. Si el cuestionario tiene varias preguntas de este tipo el trabajo de tabulación aumenta considerablemente, y el número de errores en la matriz de datos puede aumentar fácilmente, por dos motivos uno por el aumento propio de la matriz de datos y otro porque los datos de 0 y 1 pueden fácilmente intercambiarse en el proceso de tabulación generando errores de difícil localización.

Las investigaciones que utilizan encuestas a veces aparecen preguntas de respuestas múltiples, en las que se pueden elegir una, o más respuestas. Así por ejemplo, en una pregunta de respuestas múltiples con siete opciones resulta que para codificar y tabular los datos de esa pregunta se definen siete variables dicotómicas que se tabulan con ceros y unos, 0 si no se elige la opción y 1 si se elige. En muchos casos bastantes ceros, al ser respuestas no elegidas. Esto, además de suponer un trabajo tedioso, se presta a acumular errores de difícil localización entre los datos, al ser todos ceros y unos, con la consiguiente monotonía y ser indistinguibles en un análisis de los mismos.

En otras ocasiones puede haber un número de preguntas con respuestas dicotómicas con las que se suelen presentar el porcentaje de elección de cada una de las opciones en cada uno de las preguntas. Maganto, Bartau y Etxebarria (2003) ofrecen una tabla de frecuencias con las tareas domésticas con las que los hijos participan en casa. Se puede plantear la cuestión, qué porcentaje de hijos participan en un grupo de tareas más laboriosa, o menos. Esto es, los que hacen la cama, ¿hacen la compra, y ponen la mesa?, esta pregunta con el análisis que generalmente se hace no se puede contestar. Aunque es relativamente fácil contestarla si se cruzan las variables de dos en dos, no lo es si se trata de una combinación de tres variables, o más variables. Así, resultaría interesante conocer, no sólo cuántos ayudan en cada una de las tareas domésticas, sino cuántos participan en sólo una tarea, o en dos, o en más de dos, y en que tipos de tareas se dan estas combinaciones.

En el primer caso de variables de respuestas múltiples, se tienen ciertas limitaciones de análisis y relaciones con las demás variables bajo estudio. En la que algunos programas como el SPSS hay procesos para agrupar las variables dicotómicas resultantes en una variable de respuestas múltiples con la que se puede obtener la distribución de

frecuencias de las respuestas en una sola tabla, y representaciones gráficas, y como es obvio tiene limitado el análisis con las demás variables del estudio.

La cuestión es transformar la columna E de X5 en las distintas columnas de X5.1 a X5.8 con 0 y 1 según haya elegido la correspondiente opción, o no. Véase la gráfica 1.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	N°C	X2	X3	X4	X5	X5.1	X5.2	X5.3	X5.4	X5.5	X5.6	X5.7	X5.8
2	1	3	2	5	12	1	1	0	0	0	0	0	0
3	2	4	2	6	23	0	1	1	0	0	0	0	0
4	3	5	3	8	13456	1	0	1	1	1	1	0	0
5	4	7	4	6	156	1	0	0	0	1	1	0	0
6	5	4	5	8	48	0	0	0	1	0	0	0	1
7	6	1	6	5	12	1	1	0	0	0	0	0	0
8	7	1	8	2	23	0	1	1	0	0	0	0	0

GRÁFICA 1
EJEMPLO TABLA DESDOBLE DE DATOS

Si el número de opciones es mayor que nueve se puede utilizar para su codificación las letras, así en el caso de tener 12 opciones se puede utilizar 1,2,3,...,9,a,b,c, obteniendo codificaciones del tipo 25c, 489b, o, ac. La función que se propone utilizar distingue entre mayúscula y minúscula, por lo que de utilizar las letras hay que tener cuidado de ponerlas todas minúsculas, o todas mayúsculas.

Hasta aquí la situación parece simple, pero qué hacer para tener la variable X5 como indican los manuales, desdoblada en las variables de X5.1 a la X5.8 para su posterior análisis. Esto es, hay que obtener tantas variables dicotómicas como opciones tenga la pregunta y en cada una de ellas un 1 si es elegida, o un 0 si no lo es. Para ello hay que desdoblar la variable en tantas opciones como se tenga. Esto no es un problema simple ya que no se tiene un lugar fijo para cada valor, como sería en el olvidado “formato fijo” ni se tiene separadores alguno entre los valores. Para resolver este problema hay que hacer uso de varias funciones que están definidas en el Excel. Obteniendo al final el resultado deseado.

La propuesta que se hace en este artículo es codificar las respuestas en una sola variable para después desdoblarla de forma casi automática en tantas variables dicotómicas como opciones se tenga.

Aunque cada vez más se utilizan los cuestionarios en línea, Serrano (2012), todavía hay situaciones en las que se sigue utilizando el cuestionario en papel, por lo que hay que codificarlos y tabularlos, en aquellos en los que aparecen preguntas de respuestas múltiples, como, como en el trabajo de Montilla y Hernando (2009) en el que consideran las relaciones entre los profesores y el orientador pudiendo ser coordinación, organización, aspectos burocráticos, planificación, tarea tutorial, tarea departamental, otros aspectos. En total siete categorías que analizan con las frecuencias que aparecen cada una. Laorden y Peñafiel (2010) quienes plantean cuestiones sobre la formación recibida por los profesores, o sobre el tipo de recursos humanos, y otros temas, todos ellos de respuestas múltiples. Caballero y García-Lagos (2010), al preguntar sobre hábi-

tos y motivos sobre la lectura. Limón, et al. (2011) al preguntar sobre el conocimiento previos en distintos contenidos de matemáticas 18 categorías y sobre el reconocimiento de los distintos tipos de símbolos, 11 símbolos.

En todos estos estudios los autores presentan las distribuciones de frecuencias de cada una de las opciones. Como es bien sabido, la suma de los correspondientes porcentajes es mayor que 100, ya que cada persona que responde puede elegir más de una opción. Como indican los manuales al respecto, se supone que han codificado y tabulado esas cuestiones como variable dicotómicas para su posterior análisis. Una pregunta interesante en este tipo de cuestiones es ¿qué combinación de respuestas es la más frecuente?, o ¿qué combinación no se da?. Con la codificación y análisis que se hacen de los datos no se puede contestar a estas cuestiones. En cambio, con la codificación propuesta en este artículo sí se puede dar respuestas a esa cuestiones y a otras muchas, al tener una variable que recoge toda esa información se puede analizar junto con las demás variables que se tenga, a la vez que se puede hacer agrupación de la misma, por ejemplo considerar los que eligen una sola opción los que eligen combinaciones de dos, los que eligen combinaciones de tres, y los que eligen combinaciones de 4 ó más. Por supuesto también se puede tener las frecuencias con las que cada opción es elegida.

En primer lugar se expone como codificar, tabular y analizar las variables de respuestas múltiples, y en segundo lugar como unir varias variables dicotómicas en una sola, para estudiar las combinaciones de respuestas que se dan. Esto se hace con ayuda de la hoja de cálculo Excel.

2. CODIFICACIÓN DE LAS RESPUESTAS MÚLTIPLES

La forma que se propone de codificar las respuestas múltiples de una pregunta es asignarle un número según las opciones que se eligen. Esto es, si se elige la primera opción y la quinta, se le asigna el 15, si se elige la segunda opción, la tercera y la cuarta, se le asigna el 234, y así respectivamente.

De este modo la pregunta ocupa una sola columna en la matriz de datos, al igual que cualquier otra pregunta simple y se simplifica enormemente la tabulación de los datos. Siendo ahora más fácil la revisión de los datos y la localización de posibles errores en la introducción de los mismos. Además a la hora de analizar los datos se puede obtener una distribución de frecuencias de esta variable, en la que se observar cuantos han elegido una sola opción y cuantos han elegido más de una y que combinaciones de ellas son elegidas, y cuantos no han elegido ninguna opción. Cuestión esta última que no se obtiene con la codificación de 0 y 1 indicada en los manuales. Esto aporta una información adicional que de otro modo no se tendría.

Para seguir mejor la explicación se utiliza una variable de elección múltiple en un caso real.

En una investigación sobre las transversales que trabajaban los profesores de primaria se tenían preguntas referidas a las transversales, cada una con 8 opciones posibles, por lo que el número de variables a la hora de codificar las respuestas se multiplicaba por 8, este fue el motivo de buscar una solución al problema de tabulación y codificación de este tipo de variables, después de conseguir varias

soluciones utilizando separadores entre las respuestas, la mejor solución es la que se expone en este artículo. Esta solución es fácil de manejar y supone un ahorro de trabajo considerable en la tabulación de los datos, y en el número de errores que se pueden cometer.

En la investigación se eligió una muestra aleatoria de 278 profesores preguntándoles entre otras cosas que transversales han trabajado con los materiales curriculares básicos diseminados en los Centros Escolares.

¿Qué transversales has trabajado utilizando los materiales curriculares básicos?

- 1 Educación ambiental
- 2 Educación para la salud
- 3 Educación para el consumo
- 4 Coeducación
- 5 Educación moral
- 6 Educación vial
- 7 Educación para la paz
- 8 Educación sexual

Las respuestas se codificaron como se indica al inicio de esta sección, para conseguir el desdoblamiento de la variable en tantas variables dicotómicas como respuestas posibles se utiliza la siguiente función

=SI(ESERR(ENCONTRAR(valor a buscar; celdilla donde buscar;1));0;1)

Aunque la expresión puede resultar compleja, el resultado es el deseado y sólo se necesita escribirla una vez, haciendo uso después de la utilidad del Excel de rellenar hacia la derecha y hacia abajo.

Para poder entender que hace esta función compuesta se explica cada una de las funciones que la componen.

Explicación de la función:

ENCONTRAR(expresión; donde; a partir de que lugar)

Es una función de texto con tres argumentos separados por “;”. Esta función busca el texto que hay en el primer argumento –*expresión*– en la celdilla que aparece en el segundo argumento –*donde*–, a partir del lugar que aparece en el tercer argumento –*a partir de que lugar*–. Si encuentra la expresión el resultado es el orden del lugar que ocupa en el texto, o en caso contrario devuelve el error #**¡VALOR!**. Si se omite el tercer argumento de la función, la búsqueda la hace desde el primer lugar, al no necesitar buscar desde un lugar determinado, sino en cualquier lugar de la celdilla, en adelante se escribirá esta función sin el último argumento para simplificar la expresión.

Así **ENCONTRAR**(1;A2) busca si hay algún 1 en la celdilla A2. Obsérvese que da igual el lugar que ocupe en esa celdilla. Es decir que si en A2 hay un 135 devuelve un 1, y si hay 351 devuelve un 3. Si es así devuelve el lugar que ocupa el “texto” en caso contrario devuelve el error #**¡VALOR!**.

A esta función se le aplica la función

ESERR(argumento)

Esta es una función lógica que da como su resultado un VERDADERO, o FALSO, de tal modo que si el argumento es cualquier tipo de error devuelve VERDADERO, y FALSO en caso contrario.

La combinación de estas dos funciones **ESERR(ENCONTRAR(1;A2))** actúa de la siguiente forma: si en la celdilla A2 hay un 1 la primer función dará como resultado un número, que es el orden en el que esté el 1, en caso contrario el resultado es un error, entonces la función **ESERR()** da como resultado FALSO si en A2 hay un 1, y VERDADERO en caso contrario.

Ahora sólo hay que aplicar la función lógica condicional

SI(condición; resultado si verdadero; resultado si falso)

Esta función de tres argumentos opera de la siguiente forma: si el primer argumento *–condición–* es VERDADERO devuelve el segundo argumento, y si el primer argumento es FALSO devuelve el tercer argumento.

Ahora combinado estas funciones adecuadamente se puede obtener el desdoblamiento de la variable en variables dicotómicas.

=SI(ESERR(ENCONTRAR(valor a buscar; celdilla donde buscar));0;1)

De este modo el resultado de **=SI(ESERR(ENCONTRAR(1;A2));0;1)** será

- Si hay un 1 en celdilla A2, la función **ENCONTRAR(1;A2)** daría el lugar que ocupa el 1, esto es 1, ó 2, ó 3, etc. La función **ESERR()** devuelve FALSO, ya que no hay error, y por último la función condicional dará como resultado el tercer argumento, en este caso un 1. Por lo que si hay un 1 en la celdilla A2 el resultado es un 1.
- Si no hay un 1 en la celdilla A2, la función **ENCONTRAR(1;A2)** daría un error concretamente **#¡VALOR!**, la función **ESERR()** devuelve VERDADERO, ya que hay un error, y por último la función condicional dará como resultado el segundo argumento, en este caso un 0. Por lo que si no hay un 1 en la celdilla A2 el resultado es el valor 0.

Cuando se tenga más de 9 alternativas se puede utilizar letras para las opciones 10,11, y sucesivas. En el caso de buscar una letra hay que ponerla entre comillas en la función.

ENCONTRAR("b";A2)

El resultado es un número, si hay una letra "b" en la celdilla A2. Ese número sería el lugar que ocupa la letra dentro del contenido de la celdilla. En caso contrario el resultado es **#¡VALOR!**, que es un error de valor.

Pasos a seguir en la hoja de cálculo,

Una vez que se ha creado el archivo de datos en el ordenador con todos los datos se puede abrir desde la hoja de cálculo Excel, en esta los pasos a seguir son los siguientes:

- 1º Seleccionar la columna posterior a la variable que se quiere desdoblar, para ello hay que hacer clic en la letra de la columna.
- 2º Insertar tantas columnas como opciones distintas tenga la pregunta. Para ello se selecciona en el menú **Insertar** la opción **columnas**, repetir esto tantas veces como opciones distintas tenga la pregunta.
- 3º Escribir en la primera celdillas de cada columna nueva los nombres de las nuevas variables. En el ejemplo AMBIENTAL, SALUD, CONSUMO, etc.

	A	B	C	D	E	F	G	H	I
1	TRANS T	AMBIENTAL	SALUD	CONSUMO	COEDUCA	MORAL	VIAL	PAZ	SEXUAL
2		7							
3		27							
4		1257							
5		23578							
6		123567							
7		1							
8		2							

GRÁFICA 2

HOJA EXCEL CON LA COLUMNA A PARA DESDOBLAR EN LAS COLUMNAS B A LA I

- 4º En la segunda celdilla de la primera columna insertada escribir la expresión con referencia a la celdilla de la columna anterior. En esta ocasión se escribe con el símbolo \$ delante de la letra de la celdilla para que al rellenar las demás columnas con la expresión permanezca fija la columna que se pretende desdoblar. Al darle al *Intro* se tiene la expresión en la primera celdilla de la primera columna que se ha insertado. Ahora sólo queda escribirla en todas las demás, para ello se puede hacer uso de las opciones de rellenar hacia la derecha y la de rellenar hacia abajo.

	A	B	C	D	E	F	G	H	I
1	TRANS T	AMBIENTAL	SALUD	CONSUMO	COEDUCA	MORAL	VIAL	PAZ	SEXUAL
2		7							
3		27							
4		1257							
5		23578							
6		123567							
7		1							
8		2							

GRÁFICA 3

EXPRESIÓN EN LA CELDILLA B2

- 5º Seleccionar la celdillas B2 donde se ha escrito la expresión hasta la celdilla I2 de la última columna que se ha insertado, e ir al menú **Edición** y elegir **Rellenar** y aquí la opción **Hacia la derecha**.
- 6º Modificar en la expresión de cada una de las celdillas desde C2 hasta I2 el valor que corresponda buscar, según en la columna en la que se este. Así, en la primera

buscará el 1, en la segunda el 2 y así sucesivamente. Con esto se ha conseguido tener la expresión adecuadamente en la primera celdilla de cada una de las nuevas columnas.

- 7º Una vez que se tenga cumplimentadas todas las segundas filas de las columnas que se han insertado seleccionar desde estas hacia abajo hasta llegar a la última fila de los datos y elegir el comando rellenar hacia abajo. Para ello hay que rellenar con esta expresión todas estas columnas hacia abajo. Para ello seleccionar todas esas primeras celdillas desde B2 hasta I2 hasta la última fila de datos y en el menú **Edición** elegir **Rellenar** y aquí el comando **Hacia abajo**.

	A	B	C	D	E	F	G	H	I
1	TRANS_T	AMBIENTAL	SALUD	CONSUMO	COEDUCA	MORAL	VIAL	PAZ	SEXUAL
2	7	0	0	0	0	0	0	1	0
3	27	0	1	0	0	0	0	1	0
4	1257	1	1	0	0	1	0	1	0
5	23578	0	1	1	0	1	0	1	1
6	123567	1	1	1	0	1	1	1	0
7	1	1	0	0	0	0	0	0	0
8	2	0	1	0	0	0	0	0	0

GRÁFICA 4

RESULTADO FINAL DEL DESDOBLAMIENTO DE LA VARIABLE

Automáticamente se rellenará las columnas insertadas con los valores de unos y ceros según se haya elegido esa opción, o no.

Esto se repetirá para cada variable de este tipo.

Obsérvese, que el proceso descrito anteriormente se consigue llevarlo a la práctica mucho más rápido que leer los pasos a seguir. Con esto se consigue desdoblar la variable de una forma rápida y exenta de errores, al menos, de confusiones de ceros y unos.

Una forma automática de hacer lo anterior es insertando una fila al principio y después de insertar las columnas necesarias en el encabezado de las mismas escribir 1, 2, y así sucesivamente, hasta el número de respuestas posibles. De este modo sólo hay que escribir el nombre de la primera variable, y la expresión que aparece en la gráfica 3 debajo con referencias a la primera fila, al rellenar hacia la derecha se tiene todas las expresiones automáticamente, como se muestra en la gráfica 5.

	A	B	C	D	E	F	G	H	I
1		1	2	3	4	5	6	7	8
2	TRA CR	=A2&B1							
3	7	=SI(ESERR(ENCONTRAR(B\$1,\$A3));0;1)							
4	27								
5	1257								
6	23578								
7	123567								
8	1								

GRÁFICA 5

NOMBRES Y EXPRESIONES AUTOMÁTICOS

Sólo quedará seleccionar las celdillas B2 y B3 y rellenar a la derecha hasta la columna I8. Después seleccionar desde la celdilla B3 hasta la I3 y rellenar hacia abajo.

Al finalizar se guardará una copia como sólo texto, para poder eliminar la primera fila sin que se altere los resultados, quedando los datos de la forma esperada. Obsérvese, que en la celdilla B2 se ha escrito el \$ delante de A2, para cuando se rellene hacia la derecha se mantenga fija la columna. Asimismo, en la celdilla B3 se ha puesto el \$ en B\$1, para cuando se rellene hacia abajo se mantenga fija la primera fila. De este modo se automatiza el cambiar el número a buscar en cada columna.

3. ANÁLISIS DE LAS VARIABLES: LAS VARIABLES DESDOBLADAS

Ahora se tiene las variables dicotómicas y la primera variable en la que aparece las combinaciones elegidas con esta variable se puede hacer un análisis de frecuencias obteniendo las combinaciones que se dan y en que porcentaje. Así mismo, cuanto han elegido una sola opción. En cualquier caso hay que tener en cuenta que el número de combinaciones posibles son 2^k donde k es el número de opciones, para 8 opciones se puede tener hasta 256 combinaciones distintas, muchas de ellas no aparecerán. En cualquier caso se tiene esa información disponible, y se puede agrupar la variable para estudiar las combinaciones de interés y/o cruzarlas con otras variables, cosa que no se podría hacer si sólo se tiene las variables dicotómicas.

Se puede guardar el archivo en formato Excel, pero para su uso en cualquier programa informático de análisis de datos se recomienda que se guarde una copia en formato texto separado por tabuladores, de este modo el archivo de datos está sólo con los datos, sin expresiones ni formulas intermedias. Para ello ir al menú **Archivo** a la opción **guardar como** y elegir el formato de texto separado por tabulaciones.

También se puede copiar toda la hoja y en otra hoja elegir del menú edición, o con el botón derecho del ratón, pegar especial seleccionando solo contenido de este modo se tiene los datos sin las formulas listo para analizar en cualquier programa de análisis de datos.

Una vez concluido desde el Excel se le da a guardar como y se selecciona la opción de formato sólo texto separados por tabulaciones. De este modo sólo se guarda los datos de sin formulas. Este nuevo archivo se puede leer desde cualquier programa de análisis estadístico en particular desde el spss.

4. UNA APLICACIÓN PRACTICA EN LA EVALUACIÓN DE MATERIALES EDUCATIVOS

A continuación se muestra una aplicación de los anterior con datos de una investigación sobre el uso de los materiales curriculares básicos por parte del profesorado para trabajar las transversales. El tamaños de muestra fue de 278 profesores. Se obtuvo el siguiente resultado en la cuestión, ¿qué transversales has trabajado utilizando los materiales curriculares básicos?

¿Qué transversales has trabajado utilizando los materiales curriculares básicos?

- 1 Educación ambiental
- 2 Educación para la salud
- 3 Educación para el consumo

- 4 Coeducación
- 5 Educación moral
- 6 Educación vial
- 7 Educación para la paz
- 8 Educación sexual

TABLA 1
DISTRIBUCIÓN DE FRECUENCIAS DE LAS MULTIRESPUESTAS

Estadísticos

TRANS

N	Válidos	249
	Perdidos	29

Frecuencias \$TransvTrab

		Respuestas		Porcentaje de casos
		Nº	Porcentaje	
Transversalidades trabajadas ^a	AMBIENTAL	211	17,1%	84,7%
	SALUD	204	16,5%	81,9%
	CONSUMO	132	10,7%	53,0%
	COEDUCA	115	9,3%	46,2%
	MORAL	154	12,5%	61,8%
	VIAL	129	10,4%	51,8%
	PAZ	197	15,9%	79,1%
	SEXUAL	94	7,6%	37,8%
Total		1236	100,0%	496,4%

a. Agrupación de dicotomías. Tabulado el valor 1.

Este es el resultado que se obtiene en el SPSS después de definir la variable de respuesta múltiple. Resultaría interesante conocer si el 84,7% que trabajan la transversal de Ambiental trabajan también la transversal de Salud, y la de Consumo, etc. Esta información se puede obtener analizando la variable original con la que se ha tabulado las respuestas en la tabla 2 se muestra la distribución de frecuencias de todas las combinaciones que se han dado. Así, el 12,4% dicen trabajar con todas las transversales; el 6% dice trabajar con todas menos con Consumo, coeducación y sexual; el 5,6% dicen trabajar con todas menos con coeducación y sexual; el 4,4% dicen trabajar con todas menos con sexual, etc.

TABLA 2
DISTRIBUCIÓN DE LAS COMBINACIONES DE RESPUESTAS

Valor	Frec	%	%vál	Valor	Frec	%	%vál	Valores	Frec	%	%vál
1	6	2,2	2,4	568	1	0,4	0,4	12467	2	0,7	0,8
2	1	0,4	0,4	1234	2	0,7	0,8	12567	15	5,4	6
4	2	0,7	0,8	1236	1	0,4	0,4	12578	2	0,7	0,8
5	1	0,4	0,4	1237	6	2,2	2,4	13467	2	0,7	0,8
7	1	0,4	0,4	1247	4	1,4	1,6	13478	1	0,4	0,4
12	4	1,4	1,6	1248	1	0,4	0,4	14578	2	0,7	0,8
13	1	0,4	0,4	1257	9	3,2	3,6	15678	2	0,7	0,8
17	2	0,7	0,8	1267	6	2,2	2,4	23457	1	0,4	0,4
23	3	1,1	1,2	1278	2	0,7	0,8	23567	1	0,4	0,4
24	2	0,7	0,8	1367	1	0,4	0,4	23578	1	0,4	0,4
28	1	0,4	0,4	1368	2	0,7	0,8	123457	3	1,1	1,2
38	2	0,7	0,8	1378	2	0,7	0,8	123467	1	0,4	0,4
57	1	0,4	0,4	1457	3	1,1	1,2	123478	3	1,1	1,2
67	1	0,4	0,4	1458	1	0,4	0,4	123567	14	5	5,6
123	4	1,4	1,6	1467	1	0,4	0,4	123578	3	1,1	1,2
124	1	0,4	0,4	1567	3	1,1	1,2	123678	1	0,4	0,4
125	1	0,4	0,4	2357	1	0,4	0,4	124578	3	1,1	1,2
126	1	0,4	0,4	2358	1	0,4	0,4	124678	4	1,4	1,6
127	1	0,4	0,4	2367	1	0,4	0,4	125678	3	1,1	1,2
135	1	0,4	0,4	2457	3	1,1	1,2	1234567	11	4	4,4
145	1	0,4	0,4	2578	1	0,4	0,4	1234578	5	1,8	2
146	2	0,7	0,8	4578	1	0,4	0,4	1234678	4	1,4	1,6
236	1	0,4	0,4	12347	3	1,1	1,2	1235678	5	1,8	2
238	2	0,7	0,8	12356	1	0,4	0,4	1245678	3	1,1	1,2
245	1	0,4	0,4	12357	4	1,4	1,6	1345678	1	0,4	0,4
248	1	0,4	0,4	12367	4	1,4	1,6	12345678	31	11,2	12,4
257	4	1,4	1,6	12378	1	0,4	0,4	Total	249	89,6	100
267	1	0,4	0,4	12456	2	0,7	0,8	Perdidos	Sistema	29	10,4
458	1	0,4	0,4	12457	6	2,2	2,4	Total		278	100

A partir de aquí se puede agrupar la variable inicial de respuestas según los intereses de información, a continuación se muestra la distribución de la variable recodificada agrupando los valores con dos cifras en el nuevo valor DOS, los de tres cifras en TRES, y así sucesivamente, se obtiene la distribución de las transversales que se trabajan y cuantas trabajan un mismo profesor.

TABLA 3
DISTRIBUCIÓN DEL NÚMERO DE TRANSVERSALES TRABAJADAS POR EL
PROFESORADO

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
AMBIENTAL	6	2,2	2,4	2,4
SALUD	1	0,4	0,4	2,8
COEDUCA	2	0,7	0,8	3,6
MORAL	1	0,4	0,4	4
PAZ	1	0,4	0,4	4,4
DOS	17	6,1	6,8	11,2
TRES	24	8,6	9,6	20,9
CUATRO	52	18,7	20,9	41,8
CINCO	50	18,0	20,1	61,8
SEIS	35	12,6	14,1	75,9
SIETE	29	10,4	11,6	87,6
OCHO	31	11,2	12,4	100
Total	249	89,6	100	
Sistema	29	10,4		
	278	100,0		

Hay 19 profesores que no trabajan con las transversales, 11 que trabajan con solo una y 31 que trabajan con todas.

Además se tiene información de cuantos no han elegido ninguna opción, en este caso 29 el 10,4%.

5. CONCATENACIÓN DE VARIABLES DICOTÓMICAS

El caso contrario resulta más interesante desde un punto de vista analítico, ya que se puede obtener información relevante sobre las combinaciones de respuesta que se dan a las distintas preguntas del cuestionario.

Cuando se tiene la tabulación de variables dicotómicas se puede construir la variable de las combinaciones de respuestas con la función **CONCATENAR**(arg1;arg2;...). Esta función de texto permite unir cadenas de alfanuméricas, por lo que se puede construir la variable fácilmente, sólo hay que eliminar los ceros, para ello se puede utilizar la función condicional como se indica a continuación.

=**CONCATENAR**(SI(A2=1;1;"");SI(B2=1;2;"");SI(C2=1;3;""))

Si se tiene la tabulación de las respuestas de una pregunta de opción múltiple, fácilmente se puede obtener la variable con las respuestas numeradas como al inicio de este artículo, bastara con concatenar los códigos previamente multiplicados por su opción y evitando los ceros

$$SI(A2=1;1;''')\&SI(B1=1;2;''')$$

Con esta función se consigue el proceso inverso, esto es a partir de variables dicotómicas se obtiene una variable en la que aparecen las combinaciones elegidas entre las variables dicotómicas. Esto se puede utilizar también en aquellas ocasiones donde se tiene variables dicotómicas y se pretenda obtener información de las combinaciones de las respuestas. En muchas investigaciones de sondeos de opinión se presentan las distribuciones de las respuestas a las distintas cuestiones planteadas, un resultado interesante es conocer cuantos de los que han contestado una opción en una pregunta lo han hecho en otra opción de otra pregunta, construyendo esta nueva variable se puede obtener esa información detalladamente.

La cuestión que se puede hacer es, ¿son los mismos los que responden mayoritariamente a una cuestión y a otra? Si se quiere comparar dos a dos de estas variables la cuestión es simple pero si se quiere compara más de dos simultáneamente la cuestión se complica. Con el procedimiento descrito anteriormente se puede construir la variable de las combinaciones elegidas. De este modo se puede hacer un estudio mas detallado de los porcentajes que eligen sólo una opción o las combinaciones de ellas.

También se puede combinar las opciones de interés de distintas variables seleccionando el valor de interese en cada una de las variables a combinar.

Con esta nueva variable es fácil obtener el porcentaje de los que eligen sólo una opción, los que eligen dos opciones, los que eligen tres, y así sucesivamente. Basta con recodificar la variable en unidades, decenas, centenas, etc.

6. CONCLUSIÓN

Con esta tabulación en cuestionarios con preguntas de multirespuestas se simplifica el número el trabajo de introducción de datos con un ahorro importante, y se reduce considerablemente los posibles errores a la hora de tabular, a la vez se dispone de una variable en la que se tiene información detallada de las combinaciones que se dan en las respuestas múltiples, permitiendo hacer un mejor análisis de los datos pudiéndose obtener información detallada de las elecciones y pudiendo cruzar esta variable con otras, como por ejemplo las de carácter descriptivo de la muestra, como puede ser la edad o el género, o la titularidad. Además se tiene información de cuantos no han elegido ninguna opción. Todo esto es complicado obtenerlo de las variables dicotómicas que resultan del desdoble.

De otro lado el proceso inverso de unir variables dicotómicas, como las de respuestas sí, no, en una variable de respuesta múltiple permite obtener información adicional sobre los porcentajes de lección de cada una de las combinaciones posibles.

REFERENCIAS

Caballero, P. A. y García-Lagos, V. (2010). La lectura como factor determinante del desarrollo de la competencia emocional: Un estudio hecho con población universitaria. *Revista de Investigación Educativa*, 28(2), 345-359.

- Laorden, C. y Peñafiel, E. (2010). Proyectos bilingües en los centros de la comunidad de Madrid: Percepción de los equipos directivos. *Revista de Investigación Educativa*, 28(2), 325-344.
- Limón, L.; Vilchez, M. L.; Congregado, E.; Gessa, A.; García, J. J.; Velasco, F. Chamizo, C.; Naranjo, C. y García M. E. (2011) Propuesta de programa de contenidos para la asignatura “matemáticas” de los nuevos grados de ADE y FICO. *Revista de Investigación Educativa*, 29(2), 363-386.
- Maganto, J. M.; Bartau, I. y Etxeberria, J. (2003). La participación de los hijos en el trabajo familiar. *Revista de Investigación Educativa*, 21(1), 249-269.
- Montilla, M. y Hernando, A. (2009). La orientación en secundaria en la provincia de Huelva desde la perspectiva del profesorado y miembros de equipos directivos. *Revista de Investigación Educativa*, 27(2), 353-370.
- Serrano, J. (2012). Aplicación on-line y tratamiento informático de cuestionarios. *Revista Española de Pedagogía*, 251, 61-76.

Fecha de recepción: 7 de diciembre de 2012.

Fecha de revisión: 7 de diciembre de 2012.

Fecha de aceptación: 7 de abril de 2013.