

Anexo 1. Estadística descriptiva

Descripción del proceso para el análisis estadístico de los datos

John Jairo Rivera-Trujillo
Milena Barrios-Martínez

A continuación, se presenta una descripción del proceso desarrollado con cada una de las fuentes de información que recibió tratamiento para su análisis estadístico, en el marco del proyecto *¿Para qué se lee y se escribe en la universidad colombiana?: un aporte a la consolidación de la cultura académica del país.*

Debido a que la investigación era desarrollada por diecisiete equipos de docentes, uno por universidad, se recibían diecisiete matrices de datos que debían ser depuradas y compiladas. La recepción de estas se llevaba a cabo en la Pontificia Universidad Javeriana de Bogotá, desde donde eran remitidas al equipo de estadísticos para su manipulación y análisis. La depuración de las matrices de datos consiste en la identificación de espacios vacíos o casos inconsistentes dentro de la información registrada. En este caso, se procura la recuperación de los datos faltantes, pero de no ser posible, estos registros son asignados como “casos perdidos” para que no afecten los análisis e interpretaciones posteriores.

Una vez cada una de las matrices es depurada, se procede a la compilación de las mismas en una hoja de Excel. Este procedimiento va acompañado de la asignación de un código único que permite la identificación tanto del registro en particular (tipo de fuente y número asignado en cada institución) como de la universidad de la que proviene. Una vez este procedimiento es realizado, se procede al diseño de la base de datos, que consiste en la asignación de un nombre funcional a cada

una de las variables (preguntas del instrumento) y en la creación de un directorio para cada una de ellas, es decir, la asignación de un código a cada una de las opciones de respuesta. Este procedimiento facilitó el manejo de los registros y permitió desarrollar un proceso más riguroso en la manipulación de los datos, pues minimizó la pérdida de información y la posibilidad de error en la salida de los datos.

Luego, se exportó la base de datos a los programas SPSS y/o WINS-TEPS, dependiendo del caso. Una vez allí se realizó una exploración de cada una de las variables, proceso para el que se hizo una tabla con datos descriptivos básicos (frecuencias, porcentajes, mínimo, máximo), para observar si ha sido depurada por completo o si faltan ajustes. Además, se realizó un conteo de la población para saber cuántas encuestas, programas de curso o documento de políticas había para cada una de las universidades y para cada una de las agrupaciones de la UNESCO. De igual manera, se puede visualizar el número de respuestas para cada una de las preguntas.

Una vez se ha verificado la información que contiene la base de datos, se procede a la generación de resultados, que para esta investigación consistió en la recuperación de la información por frecuencias y porcentajes. Es decir, por pregunta se realiza un conteo de respuestas en cada una de las opciones y su correspondiente porcentaje (por ejemplo, cuadro 1). Estos resultados fueron generados para cada uno de los niveles de representatividad de la muestra: Total de universidades; Universidad; Áreas UNESCO; Áreas UNESCO por Universidad (para el caso de los cursos se realizó una agrupación que diera cuenta de los cursos disciplinares en contraste con aquellos que eran destinados exclusivamente al abordaje de la lectura y la escritura).

Como puede verse, la generación de resultados responde a las necesidades particulares de la investigación, pues permite encontrar regularidades asociadas a los campos de formación disciplinar en relación con las prácticas para la enseñanza de la lectura y la escritura que se proponen en las universidades.

Cuadro 1

2. ¿Para cuáles espacios académicos se escribe?			
		Número	Porcentaje (%)
Concurso	No	3.461	93,09
	Sí	257	6,91
B. Grupo de estudio	No	2.481	66,71
	Sí	1.238	33,29
C. Asignatura	No	511	13,74
	Sí	3.208	86,26
D. Taller de escritura	No	3.114	83,73
	Sí	605	16,27
E. Semillero de investigación	No	3.164	85,08
	Sí	555	14,92
F. Evento académico	No	2.765	74,35
	Sí	954	25,65
G. Curso extracurricular	No	3.065	82,41
	Sí	654	17,59
H. Otra	No	3445	92,66
	Sí	273	7,34

Adicionalmente, se procedió a la recuperación de la información correspondiente a las preguntas abiertas, que incluía, entre otras cosas, la información adicionada a cada pregunta por los investigadores para complementar las opciones de respuesta ya existentes y las explicaciones o profundizaciones de otras respuestas (cuáles, qué, por qué). Con este fin, se generó un reporte de las variables que tenían esta característica y se recuperó cada una de las respuestas; y en función de su importancia dentro del estudio, se reagruparon como opciones de respuesta para complementar la información ya existente.

Una vez estos datos fueron analizados e interpretados, los equipos de investigadores tuvieron la posibilidad de solicitar la realización de “cruces especiales” entre las preguntas, requerimiento que se atendió con la realización de tablas de contingencia. Este procedimiento se acerca a la explicación de los fenómenos más que a su descripción, pues permite la generación de hipótesis y la comprobación de relaciones entre los distintos aspectos indagados en las fuentes.

Finalmente, la información registrada en estas tablas fue graficada para facilitar su visualización, ya fuera en barras, tortas o con líneas de continuidad, todo lo anterior con base en frecuencias y porcentajes en cada uno de los niveles de la muestra. Esta elección se realizó en función de las características de las variables, teniendo en cuenta que fueran de fácil comprensión para el lector.