

MODELO PARA LA CREACIÓN DE UN MAPA INTERACTIVO DE VULNERABILIDAD ESCOLAR UTILIZANDO TÉCNICAS DE SEGMENTACIÓN QUE PERMITAN IDENTIFICAR REGLAS DE FOCALIZACIÓN



Santo Domingo,
República Dominicana
Marzo 2016

MODELO PARA LA CREACIÓN DE UN MAPA INTERACTIVO DE VULNERABILIDAD ESCOLAR UTILIZANDO TÉCNICAS DE SEGMENTACIÓN QUE PERMITAN IDENTIFICAR REGLAS DE FOCALIZACIÓN

Investigador Principal:

Osmel Brito-Bigott

Co-investigador:

Daniel Morales Romero

Santo Domingo,
República Dominicana.

2016



Modelo para la creación de un mapa interactivo de vulnerabilidad escolar utilizando técnicas de segmentación que permitan identificar reglas de focalización

Dirección Técnica de Evaluación e Investigación

Dinorah de Lima, M. Ed.

Dr. Luis Camilo Matos de León

Juan Miguel Pérez, M. Ed.

Nery Antonio Taveras, M. Ed.

Dr. Julián Álvarez Acosta.

Ing. Dilcia Armesto Núñez

Massiel Cohén Camacho, M. Ed.

Coordinadora de proyecto IDEICE-PNUD

Rita Licelot Cruz M. A.

Corrección de estilo

Ramón Fari Rosario

Diseño y Diagramación

Lic. Yeimy Rosa Olivier Salcedo

Lic. Natasha Mercedes Arias

Centro de Documentación

Ing. Dilcia Armesto Núñez

Derechos Reservados

Ministerio de Educación de la República Dominicana

Febrero 2016

Se permite reproducir parcialmente este documento siempre que se cite la fuente

ISBN: 978-9945-499-18-6

Santo Domingo, D.N.
República Dominicana



AUTORIDADES

Lic. Danilo Medina Sánchez
Presidente de la República

Dra. Margarita Cedeño de Fernández
Vicepresidenta de la República

Lic. Carlos Amarante Baret
Ministro de Educación

Lic. Luis Enrique Matos De La Rosa
Viceministro de Educación, Encargado de Servicios Técnicos y Pedagógicos

Lic. Antonio Peña Mirabal, Viceministro de Educación
Encargado de Asuntos Administrativos y Financieros

Lic. Saturnino de los Santos
Viceministro de Educación, Encargado de Certificación Docente

Ing. Víctor Ricardo Sánchez, Viceministro de Educación
Encargado de Planificación y Desarrollo Educativo

Dr. Adarberto Martínez
Viceministro de Educación, Encargado de Supervisión y Evaluación de la Calidad Educativa

Lic. Luis de León
Viceministro de Educación, Encargado de Descentralización

Dr. Julio Leonardo Valeirón
Director Ejecutivo del IDEICE

Resumen

En el trabajo llevado a cabo en torno al Modelo para la creación de un mapa interactivo... se integraron las bases de datos del Sistema de Gestión de Centros y Pruebas Nacionales para elaborar un índice de vulnerabilidad educativa que toma en cuenta los indicadores escolares de abandono y repitencia, así como las calificaciones de las Pruebas Nacionales por materia. Se propuso un algoritmo para la elaboración del índice y se analizaron las características del sistema educativo que están relacionadas con la vulnerabilidad escolar. Dichas características se refieren al tamaño del distrito, al tamaño de la escuela, la dificultad de acceso y el sector. Si bien estas variables solo caracterizan la vulnerabilidad, lo relevante es que con la misma pudo construirse una economía de la vulnerabilidad que permitió elaborar mapas de localización de vulnerabilidad por regionales y por distritos. Estos tienen el potencial de constituir información de inteligencia del sector educativo que contribuya con la toma de decisiones y con la focalización de intervenciones que logren mayor equidad en el sistema educativo.

Indice

Introducción	1
Marco teórico	2
Metodología	3
Datos	4
Resultados y análisis	12
Conclusiones	16
Bibliografía	17
Anexos	18

Introducción

En el sistema educativo dominicano existen indicadores de vulnerabilidad escolar, pero son unidimensionales y no existe un score de vulnerabilidad, ni una forma de visualización de grandes volúmenes de datos, que faciliten información oportuna a las autoridades en torno a la localización de la vulnerabilidad.

El sistema educativo dominicano genera millones de datos que tienen la potencialidad de construir indicadores que informen, de manera oportuna, sobre diversas materias. De hecho, el Ministerio de Educación de República Dominicana (MINERD), mediante el Sistema de Gestión de Centros, de Pruebas Nacionales y del Sistema de Acompañamiento y Supervisión, genera múltiples indicadores que, por medio de *drill down*, permiten visualizar información a nivel Regional, Distrital y Escolar.

Dado que la producción de datos agregados permiten la obtención de una panorámica situacional sobre los indicadores de desempeño del sistema, no existen “reglas de negocio”, ni de visualización de datos que permitan establecer focalizaciones en atención a la vulnerabilidad. Tampoco las bases de datos de Pruebas Nacionales y el Sistema de Gestión de Centros están integradas en una base de datos y *datawarehouse* único, lo que dificulta la generación de reportes de manera oportuna.

La elaboración de un *datawarehouse* que desarrolle la potencialidad de dar respuestas a la demanda de información sobre vulnerabilidad puede proveer a las autoridades del sistema educativo información nunca antes explotada de manera dinámica al realizar análisis integrado y con técnicas usadas para caracterizar segmentos y establecer determinantes propios de cada segmento que logre identificarse.

Para lograr los objetivos de esta investigación se integró la base de datos de Pruebas Nacionales y del Sistema de Gestión de Centros, lo que permitió calcular índices para caracterizar los niveles de vulnerabilidad de cada escuela y establecer diferencias de vulnerabilidad escolar por distritos y regionales educativas.

Este trabajo está asociado a la Línea 7 del Plan Decenal (Promover la equidad educativa con apoyo a los estudiantes provenientes de los sectores sociales más vulnerables), Línea 4 de Plan Decenal (Establecer claros estándares de calidad y un sistema de evaluación que permita monitorear el desempeño); al objetivo 1 del Segundo Eje Estratégico de la Ley Estrategia Nacional de Desarrollo (END) 2012 (“Una sociedad con igualdad de derechos y oportunidades, en la que toda la población tiene garantizada educación, salud, vivienda digna y servicios básicos de calidad y que promueve la reducción progresiva de la pobreza y la desigualdad social y territorial”); el Punto 7 del Pacto Educativo (Sobre la modernización de la gestión del sistema educativo para elevar su eficacia, transparencia y rendición de cuentas en el logro de los objetivos); y también lo hace la Cuarta Dimensión del IDEC (Gestión y participación).

El modelo que se desarrolló logró la elaboración de un mapa de vulnerabilidad escolar utilizando características de segmentos que permitan identificar reglas de focalización, que sirvan a su vez como un insumo para el desarrollo de políticas públicas de seguimiento a la vulnerabilidad escolar, e informar a las autoridades de manera oportuna sobre la localización de la vulnerabilidad para ser más efectivos en la asignación de recursos de manera que se logre democratizarlos de forma más equilibrada.

En la sección 2 se presenta el marco teórico, en la sección 3 la metodología, en la 4 los datos utilizados, en la 5 los resultados y en la última se concluye.

Marco teórico

Minería de datos y segmentación

Según Berry y Linoff (2004), la minería de datos es un proceso de exploración y análisis de grandes conjuntos de datos para descubrir patrones y reglas con un significado de negocio.

La minería de datos puede ser directa e indirecta. La minería de datos directa trata de categorizar y explicar alguna variable de respuesta; por ejemplo, la vulnerabilidad escolar. La minería de datos indirecta trata de buscar patrones o similitudes entre grupos sin el uso de una variable de respuesta particular.

Con la minería de datos pueden construirse modelos, ya sean con algoritmos o con un conjunto de reglas que conectan un conjunto de insumos con una variable dependiente particular.

Collica (2007), argumenta que la segmentación es la clasificación de los datos de manera conveniente. De hecho, las variables que permiten categorizar las unidades de observación (personas, escuelas, regionales, distritos, países, etc.), sirven para establecer segmentos dentro de la población.

Todas estas definiciones son pertinentes para abordar la vulnerabilidad escolar.

Vulnerabilidad escolar

La Dirección General de Cultura y Educación de la Provincia de Buenos Aires (2010), define la vulnerabilidad educativa como “problemáticas relacionadas con las trayectorias escolares y el vínculo de escolarización de niños, niñas y jóvenes.” Más aun, sostiene que “la perspectiva de la vulnerabilidad educativa implica atender a posibles interrupciones y desfasajes [sic] en las trayectorias escolares de los alumnos desde las propias instituciones escolares.”

Hay que identificar la vulnerabilidad escolar y establecer dónde se localiza, pues resulta de suma importancia para la focalización de políticas públicas tendientes a la equidad. Por ejemplo, el Fondo Nacional de Becas (2010) de Costa Rica concibe los índices de vulnerabilidad educativa como una herramienta novedosa para mejorar el proceso de asignación de becas. Más aún, este índice lo ven como una información estratégica que los puede guiar en la disminución del abandono escolar al poder identificar los factores de riesgo relacionados con la vulnerabilidad.

El Ministerio de Educación de Chile, usa una Metodología de Construcción de Grupos Socioeconómicos en SIMCE 8vo. Básico (2009) para agrupar y clasificar a las escuelas haciendo uso de conglomerados de información a nivel escolar.

Catalá, Con y Lentini (2008) elaboran un Índice de Vulnerabilidad Educativa (IVE) que considera la sobreedad, la repitencia y el abandono escolar. El IVE es un índice simple, que suma cada uno de estos tres indicadores en una escala del 0 al 1. El IVE es un índice agregado por escuela, que a su vez puede ser agregado para caracterizar diversos segmentos de la vulnerabilidad educativa.

Si bien han existido esfuerzos para clasificar a los centros educativos mediante su desempeño en las Pruebas Nacionales como el Informe CLASE (2012) del exministro del MINERD, Melanio Paredes, así como los índices unidimensionales por escuela, distritos y regionales reportados por el propio sistema educativo. A pesar de este escenario, República Dominicana carece de un índice de vulnerabilidad escolar que aproveche todas las fuentes de datos con las que dispone el MINERD, además del desconocimiento de la aplicación de modelos de minería de datos que traten de conseguir relaciones tendientes a la vulnerabilidad educativa.

Metodología

El diseño de la investigación parte de la hipótesis de que las informaciones contenidas en las bases de datos de Pruebas Nacionales y el Sistema de Gestión de Centros son susceptibles de ser integradas en un *datawarehouse* único y aplicarles modelos de segmentación y de minería de datos que generen un score de vulnerabilidad.

Para comprobar dicha hipótesis primero se integran las bases de datos de Pruebas Nacionales y la del Sistema de Gestión de Centros, según el modo de disponerla para realizar un análisis intensivo mediante modelos de segmentación y de minería de datos.

El método analítico que debe emplearse hace una serie de definiciones operacionales para tratar los datos y convertirlos en información. De esta manera, sea X_{ij} el indicador i de la escuela j , donde $i = \{Pruebas Nacionales, Indicadores del Sistema de Gestión de Centros\}$ donde a la vez $Indicadores del SGCE = \{abandono, repitencia\}$.

La metodología se propone usar un algoritmo que contempla los siguientes pasos:

- Primero, construir los índices en base a los indicadores de la cantidad de estudiantes por condición del SGCE. Estos índices se construyeron con la siguiente regla: sea $Indice_SGC_{jhk}$ el indicador elaborado, donde h son los subcomponentes del índice $h: \{Repitencia, Abandono\}$, de la escuela j , para el año escolar k , donde $k: \{2012-2013, 2013-2014\}$. Los $Indice_{jhk}$ son el resultado de dividir la suma del componente h para media y básica entre la matrícula total de estudiantes de la escuela

$$\frac{\sum_{hij}}{\text{Matrícula total}}$$

- Segundo, se ejecuta un *ranking* para cada indicador i , donde el $Cuartil_{ij}$ es un score del cuartil al que pertenece la escuela j , de manera que $Cuartil_{ij} = \{0, 1, 2, 3\}$. Así '0' es el score de la escuela que se encuentra en el último cuartil, '1' el score de la escuela que se encuentra en el tercer cuartil, '2' la del segundo cuartil, y '3' la del primer cuartil.

1 En el caso de los rankings de los indicadores del SGCE se invirtió el sentido de los valores para que todos los indicadores tuvieran la misma dirección. Por ejemplo: se cambió 0 por 3, 1 por 2 y así sucesivamente.

- Tercero, se construyen índices a partir de los *rankings* por cuartil. Sea $Indice_Desempeño_{j,hk}$ el resultado de la sumatoria de los *rankings* por asignatura de las Pruebas Nacionales, el cual tendrá un valor mínimo de 0 y máximo de 12. Sea $Indice_Condición_{j,hk}$ el resultado de la sumatoria de los *rankings* de los $Indices_SGC_{j,hk}$, los cuales tendrán un valor mínimo de 0 y un máximo de 6.
- Cuarto, se elabora un tercer índice $Indice_Vulnerabilidad_{j,hk}$ el cual es el resultado de sumar el $Indice_Desempeño_{j,hk}$ y el $Indice_Condición_{j,hk}$. Este índice tendrá un rango de 0 a 18 puntos, donde 18 es la mejor condición.
- Finalmente, si se elabora un *ranking* del índice de vulnerabilidad se obtiene que las escuelas se clasifican por $Cuartil_{ij}=\{0,1,2,3\}$ en *Muy vulnerables*, *Vulnerables*, *Poco vulnerables*, *Sin vulnerabilidad*, respectivamente, para cada valor del *ranking*.

El diseño metodológico de esta investigación busca crear un score de vulnerabilidad que tenga propiedades de agregación a nivel Distrital y Regional. Asimismo, el desarrollo de esta metodología busca lograr una visualización múltiple de los datos que normalmente son difíciles de mostrar en una sola visualización.

Por otro lado, se realiza el estudio de la vulnerabilidad escolar de la siguiente manera:

- Nivel nacional, regional y distrital: Caracterizado por público-privada, rural-urbana, número de estudiantes.
- Comparaciones de Score Agregados del vulnerabilidad.
- Elaboración de un mapa nacional, regional, distrital de Vulnerabilidad.

En definitiva, esta investigación espera lograr construir un modelo que pueda generar un score de vulnerabilidad escolar que sirva a las autoridades para localizar las zonas y escuelas más débiles (vulnerables), y las que deberían recibir más atención en miras de lograr una mayor equidad.

Por último, esta investigación tiene la potencialidad de proporcionar un modelo para la creación de un mapa de vulnerabilidad escolar, que sirva a su vez como un insumo de política clara de seguimiento de la misma, así como también para informar a las autoridades, de manera oportuna, sobre su localización, permitiendo además una mayor efectividad en la asignación de recursos que garantizarán democratizarlos equilibradamente.

Datos

Para la elaboración del estudio de vulnerabilidad de las escuelas de República Dominicana se dispone de bases de datos suministradas por el Sistema de Gestión de Centros Educativos (SGCE) y Pruebas Nacionales (PN). Estas bases de datos tienen características que las hacen viables para ser integradas, de modo que pueda obtenerse de ellas conclusiones que respondan a la necesidad del sistema educativo en torno a sus debilidades y oportunidades, con miras a la búsqueda del mejoramiento y optimización del sistema. En esta sección, se presentan las descripciones de las variables de las bases de datos disponibles, así como los formatos de estas variables, las modificaciones realizadas, los errores encontrados, el número de observaciones disponibles, las estadísticas correspondientes, otras especificaciones necesarias, y cómo se construye el índice desde cada uno de sus subcomponentes. En el anexo se describe cómo se integraron estas dos bases de datos.

Para obtener el índice de vulnerabilidad se deben obtener previamente dos componentes. El primero es el índice de índice de condición, que se obtiene luego de aplicar el procedimiento metodológico descrito en el apartado anterior a la base de datos del Sistema de Gestión de Centros. El segundo es el índice de desempeño, que se obtiene de la base de datos de Pruebas Nacionales. Por esta razón, la obtención de cada uno de estos componentes se describe en subsecciones separadas.

Sistema de Gestión de Centros

La base de datos del Sistema de Gestión de Centro Educativos contiene la información de matrícula por grado y la cantidad de estudiantes para cada tipo de condición (promovidos, repitientes, abandono), así como la modalidad de cada centro educativo, y otras características Regionales, Distritales y lo concerniente al sector que pertenece.

El objetivo principal de esta subsección es describir cada uno de los subcomponentes del índice de condición descrito en la metodología. Este es un insumo del índice de vulnerabilidad escolar tal como se ha definido.

De esta manera, para cada centro educativo cuenta con la información que provee cada una de las variables que contiene la base de datos. En el anexo de este trabajo se describe cómo se construyeron los índices que se presentan a continuación.

En la tabla 1 se presentan las reglas de construcción de los índices relacionados a las condiciones de los estudiantes y se muestran los índices construidos a partir de los indicadores de la cantidad de estudiantes por condición.

Tabla 1
Regla de construcción de los índices relacionados a las condiciones de los estudiantes

VARIABLE	NOMBRE	REGLA DE CONSTRUCCIÓN
Índice de promoción	Índice_promoción	Es el resultado de dividir el número de promovidos de las modalidades de básica y media entre el número total de estudiantes de las escuelas.
Índice de reprobados	Índice_reprobados	Es el resultado de dividir el número de reprobados de las modalidades de básica y media entre el número total de estudiantes de las escuelas.
Índice de abandono	Índice_abandono	Es el resultado de dividir el número abandonos de las modalidades de básica y media entre el número total de estudiantes de las escuelas.

Las estadísticas descriptivas de estos índices calculados para las escuelas de básica para cada año escolar se presentan en la tabla 2. Esta tabla muestra que el 82% de los estudiantes fueron promovidos para los años escolares 2012-2013 y 2013-2014, respectivamente. Mientras que un 7% y 6% fueron reprobados respectivamente. La tasa de abandono escolar anual se ubicó en el 2%.

Tabla 2
Estadística descriptiva de los índices de las condiciones de los estudiantes de básica

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX.	OBS.	DATOS FALTANTES
índice_promocionB_12_13	0.82	0.08	0.31	1	3621	858
índice_reprobadosB_12_13	0.07	0.05	0	0.46	3621	858

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX.	OBS.	DATOS FALTANTES
índice_abandonoB_12_13	0.02	0.04	0	0.64	3621	858
índice_promocionB_13_14	0.82	0.09	0	1	3624	855
índice_reprobadosB_13_14	0.06	0.05	0	0.41	3624	855
Índice_abandonoB_13_14	0.02	0.05	0	1	3624	855

En la tabla 3 se presentan las estadísticas de los índices de las condiciones de los estudiantes en las escuelas de media. El 84% de los estudiantes fueron promovidos en los dos años escolares y el 4% de los estudiantes abandonaron la escuela en 2012-2013, mientras que en 2013-2014, el 5% de los estudiantes abandonaron. El índice de reprobados se ubicó en el 7 y 6%, respectivamente.

Tabla 3
Estadísticas de los índices de las condiciones de los estudiantes en las escuelas de media

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX.	OBS.	DATOS FALTANTES
índice_promocionM_12_13	0.84	0.10	0	1	1818	658
índice_reprobadosM_12_13	0.07	0.07	0	0.43	1818	658
índice_abandonoM_12_13	0.04	0.07	0	1	1818	658
índice_promocionM_13_14	0.84	0.12	0	1	1814	662
índice_reprobadosM_13_14	0.06	0.07	0	0.75	1814	662
índice_abandonoM_13_14	0.05	0.09	0	1	1814	662

Con estos insumos se realizan los *rankings* cuartil de los índices de promoción, reprobados y repitencia. En la tabla 4 se presenta la operacionalización de estas variables, por medio de la regla de construcción y sus modificaciones. Con cada *ranking* cuartil se opera el algoritmo descrito en la metodología para obtener el índice de condición con la sumatoria del *ranking* cuartil de reprobados y de abandono. Este índice de condición se vuelve a aplicar el *ranking* para obtener el *ranking* cuartil del índice de condición, el cual es un insumo del índice de vulnerabilidad.

Tabla 4
Regla de construcción índice de condición

VARIABLE	NOMBRE	REGLA DE CONSTRUCCIÓN	MODIFICADA	OTRAS MODIFICACIONES
Ranking Índice de promoción	Rank_Índice_promoción	Se elaboraron rankings por cuartil a partir de las variables relacionadas con las condiciones promedio de los estudiantes por escuela.	Estas variables toman valores entre 0 y 3 para cada cuartil donde 0 es el más bajo.	Para lograr que todas tuvieran el mismo sentido se invirtieron los valores en las variables ranking índice de reprobados y abandono.
Ranking Índice de reprobados	Rank_Índice_reprobados			
Ranking Índice de abandono	Rank_Índice_abandono			
Índice condición	Índice_condición	Esta variable se construyó sumando los rangos por condición de abandono y reprobados.	Contiene valores ente 0 y 6. Donde 6 es la mejor condición.	N/A
Ranking Índice condición	Rank_Índice_condición	Es un ranking cuartil del Índice de condición.	Estas variables toman valores entre 0 y 3 para cada cuartil donde 0 es el más bajo.	N/A

Pruebas Nacionales

Por otra parte, la base datos de las Pruebas Nacionales proporciona el otro subcomponente del índice de vulnerabilidad escolar, el cual es el índice de desempeño, tal como se describe en la metodología. Por lo tanto, el objetivo principal de esta subsección es describir cada uno de los subcomponentes del índice de desempeño que es un insumo del índice de vulnerabilidad escolar.

La tabla 5 muestra las descripciones de las variables relacionadas con las calificaciones obtenidas por los estudiantes en Español, Matemáticas, Naturales y Sociales.

Tabla 5
Descripciones de las variables relacionadas con las calificaciones

VARIABLE	FORMATO	DESCRIPCIÓN	OBSERVACIONES
Español	Numérico	Indica el promedio de las calificaciones de los estudiantes de cada escuela en el examen de español en las Pruebas Nacionales de los años escolares 2012-2013 y 2013-2014.	N/A
Matemáticas	Numérico	Indica el promedio de las calificaciones de los estudiantes de cada escuela en el examen de matemáticas en las Pruebas Nacionales de los años escolares 2012-2013 y 2013-2014.	N/A
Naturales	Numérico	Indica el promedio de las calificaciones de los estudiantes de cada escuela en el examen de Naturales en las Pruebas Nacionales de los años escolares 2012-2013 y 2013-2014.	N/A
Sociales	Numérico	Indica el promedio de las calificaciones de los estudiantes de cada escuela en el examen de Sociales en las Pruebas Nacionales de los años escolares 2012-2013 y 2013-2014.	N/A

A partir de la Base de Datos de la primera convocatoria de las Pruebas Nacionales se construyeron otras variables. Para operacionalizar y poder ejecutar el algoritmo del índice de desempeño, se utilizaron las siguientes reglas de construcción de las variables tal como se muestra en la tabla 6.

Con la sumatoria del *ranking* cuartil por materia se obtiene el índice de desempeño, que a su vez se le aplica un *ranking* cuartil para obtener el ranking del índice de desempeño, que es un insumo del índice de vulnerabilidad.

Tabla 6
Variables modificadas a partir de los datos de las Pruebas Nacionales

VARIABLE	NOMBRE	REGLA DE CONSTRUCCIÓN 1	MODIFICADA	OTRAS MODIFICACIONES
Ranking Español	Rank_Español	Se dispuso de cuatro variables llamadas Español, Matemáticas, Naturales y Sociales con valores de las calificaciones en las Pruebas Nacionales. Con las cuales se elaboró un rango por cuartil.	Estas variables toman valores entre 0 y 3 para cada cuartil donde 0 es el más bajo.	Se les asignó calificaciones (D, C, B, A) a los valores del rango entre 0 y 3, respectivamente.
Ranking Matemáticas	Rank_Matemáticas			
Ranking Sociales	Rank_Naturales			
Ranking Naturales	Rank_Sociales			
Índice de desempeño	Índice_desempeño	Esta variable se construyó a partir de los rangos por materia. Es la sumatoria de los valores asignados a cada cuartil por rangos por materia.	Contiene valores entre 0 y 12, donde 12 se le asigna a las escuelas que estén en los cuartiles más alto de todos los rangos por materias.	N/A

VARIABLE	NOMBRE	REGLA DE CONSTRUCCIÓN 1	MODIFICADA	OTRAS MODIFICACIONES
Ranking del Índice de desempeño	Rank_Índice_desempeño	Se elaboró un rango por cuartil a partir del Índice de vulnerabilidad.	Estas variables toman valores entre 0 y 3 para cada cuartil, donde 0 es el más bajo.	N/A

Un ejemplo del producto resultante de la construcción de variables para cada escuela es el que se presenta en la tabla 7. Por ejemplo, para la escuela Esmirna, cuyo código de planta física 16000114, le corresponden los siguientes valores:

Tabla 7
Ejemplo de Construcción de Variables

VARIABLE	REGLA DE CONSTRUCCIÓN	RANGO	CALIFICACIÓN
Rango Español	Cuartil (Español, Esmirna)	3	A
Rango Matemáticas	Cuartil (Matemáticas, Esmirna)	2	B
Rango Sociales	Cuartil (Sociales, Esmirna)	3	A
Rango Naturales	Cuartil (Naturales, Esmirna)	0	D
Índice de desempeño	Σ (Rangos por asignaturas, Esmirna)	8	N/A
Rango del índice de desempeño	Cuartil (índice de desempeño, Esmirna)	2	N/A

Índice de Vulnerabilidad

En la tabla 8 se presenta la regla de construcción del índice de vulnerabilidad, tal como se describió en la sección metodológica, que a su vez usa como insumo el índice de desempeño y el índice de condición. El índice de vulnerabilidad es un score escolar que puede tomar valores de 0 a 18. De esta manera, si una escuela está en el *ranking* cuartil más bajo en todas las categorías (Matemáticas, Español, Sociales, Naturales, Abandono y Repitencia), obtiene el valor de cero; y si está en el *ranking* cuartil más alto en todas las categorías, obtiene el valor de 18. A su vez, el ranking cuartil del índice de vulnerabilidad permite poner etiquetas categóricas de vulnerabilidad por rango cuartil.

Tabla 8
Descripción del índice de vulnerabilidad

VARIABLE	NOMBRE	REGLA DE CONSTRUCCIÓN 1	MODIFICADA	OTRAS MODIFICACIONES
Índice de Vulnerabilidad	Índice_Vulnerabilidad	Es la sumatoria del índice de desempeño y el ranking del índice de condición.	Contiene valores entre 0 y 18, donde 18 se le asigna a las escuelas que estén en el cuartiles más alto de todos los rangos por materias y condición.	N/A
Ranking del Índice de Vulnerabilidad	Rank_Índice_Vulnerabilidad	Se elaboró un rango por cuartil a partir del Índice de Vulnerabilidad	Esta variable toma valores entre 0 y 3 para cada cuartil, donde 0 es el peor Vulnerabilidad.	N/A

La tabla 9 presenta las estadísticas descriptivas de vulnerabilidad para las escuelas de media. Primero se presenta el índice de desempeño, el cual tiene un promedio de 5.99 que es aproximadamente igual a 6. Por construcción, el algoritmo hace que la media y la mediana tengan casi la misma medida para el índice de desempeño y de condición. Ahora, para el índice de vulnerabilidad hay variaciones debido a que se pierden observaciones al descartar escuelas que no tengan simultáneamente información de desempeño y condición. En cuanto al índice de condición la media es de 3.001 y el índice de vulnerabilidad de escuelas de educación media es de 9.55.

Tabla 9
Estadísticas de los índices de media

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
Índice Media Desempeño 2013 2014	5.995	3.841	0	12	2413	63
Índice Media Condición 2013 2014	3.001	1.939	0	6	1814	662
Índice Media vulnerabilidad 2013 2014	9.559	5.016	0	18	1801	675

La tabla 10 contiene las misas estadísticas de vulnerabilidad pero para educación básica. Como en la tabla anterior el análisis es el mismo por construcción del algoritmo.

Tabla 10
Estadística de los índices de básica

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
Índice Básica desempeño 2013-2014	5.996	3.849	0	12	4394	85
Índice Básica condición 2013 2014	3.000	1.806	0	6	3624	855
Índice Básica vulnerabilidad 2013-2014	9.286	4.648	0	18	3583	896

Economía de la Vulnerabilidad

El índice de vulnerabilidad puede usarse tanto para caracterizar escuelas, así como Regionales o Distritos educativos.

Las tablas 11 y 12 presentan la cantidad de escuelas para el índice de vulnerabilidad, tanto para las escuelas de media como para las de básica. Para básica, el 45.16% de las escuelas se encuentran por debajo del 50% del índice. El 3.6% de las escuelas de básica tiene los 18 puntos máximos que puede tomar el índice, es decir, 129 escuelas presentan un score de vulnerabilidad que señala que tienen fortaleza. En cambio, y en el extremo, 37 escuelas de básica están en la categoría más vulnerable, con un índice de cero. En la modalidad de media en 44.75% de las escuelas está por debajo de la mitad del índice y el 6.22% tiene el total de los 18 puntos que implica que 122 presentan altos niveles de fortaleza educativa.

Tabla 11
Frecuencia índice de vulnerabilidad de básica

ÍNDICE VULNERABILIDAD BÁSICA 2013-2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
0	37	1.03	37	1.03
1	86	2.4	123	3.43
2	121	3.38	244	6.81
3	194	5.41	438	12.22
4	204	5.69	642	17.92
5	223	6.22	865	24.14
6	255	7.12	1120	31.26
7	246	6.87	1366	38.12
8	252	7.03	1618	45.16
9	281	7.84	1899	53
10	250	6.98	2149	59.98
11	248	6.92	2397	66.9
12	210	5.86	2607	72.76
13	193	5.39	2800	78.15
14	197	5.5	2997	83.64
15	162	4.52	3159	88.17
16	150	4.19	3309	92.35
17	145	4.05	3454	96.4
18	129	3.6	3583	100
Datos Faltantes = 896				

Tabla 12
Frecuencia índice de vulnerabilidad de media

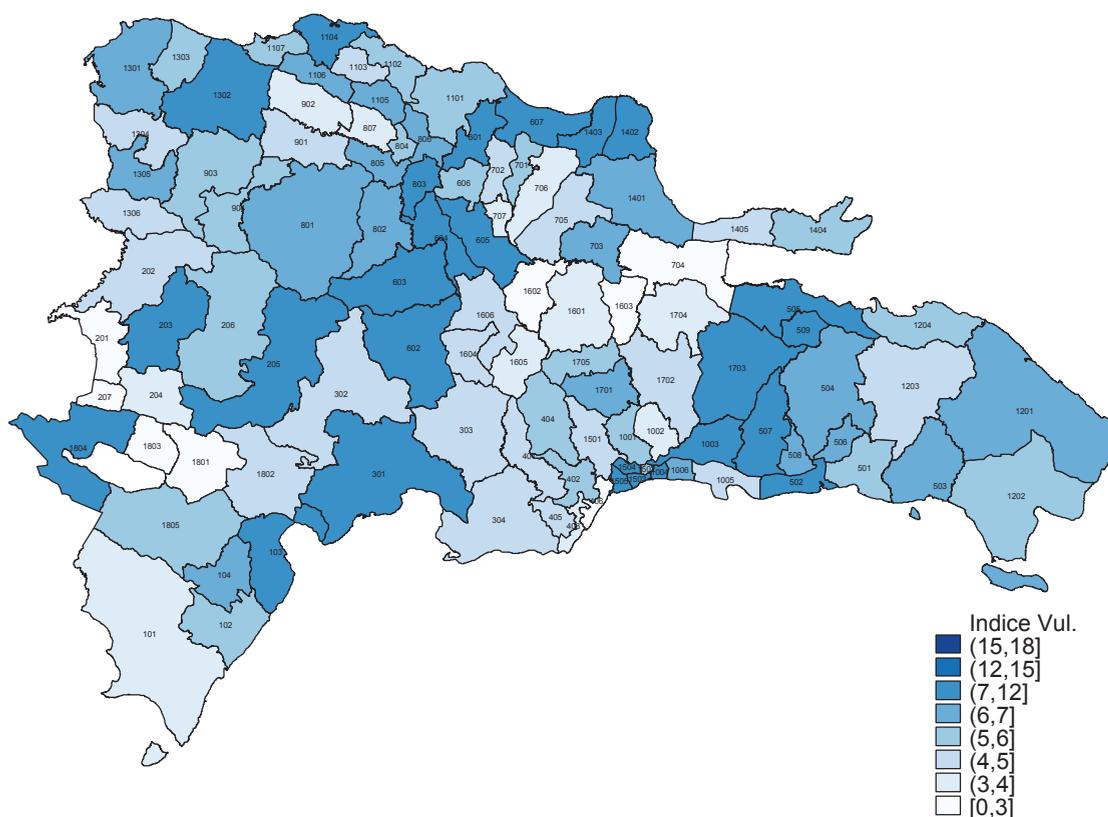
ÍNDICE MEDIA VULNERABILIDAD 2013-2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
0	28	1.55	28	1.55
1	48	2.67	76	4.22
2	70	3.89	146	8.11
3	86	4.78	232	12.88
4	112	6.22	344	19.1
5	123	6.83	467	25.93
6	113	6.27	580	32.2
7	98	5.44	678	37.65
8	128	7.11	806	44.75
9	106	5.89	912	50.64
10	114	6.33	1026	56.97
11	97	5.39	1123	62.35
12	111	6.16	1234	68.52
13	94	5.22	1328	73.74

ÍNDICE MEDIA VULNERABILIDAD 2013-2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
14	95	5.27	1423	79.01
15	86	4.78	1509	83.79
16	93	5.16	1602	88.95
17	87	4.83	1689	93.78
18	112	6.22	1801	100
Datos Faltantes = 675				

Asimismo, los índices de vulnerabilidad de básica y media pueden colapsarse para visualizar diferencias de vulnerabilidad por Regional y Distritos. ²

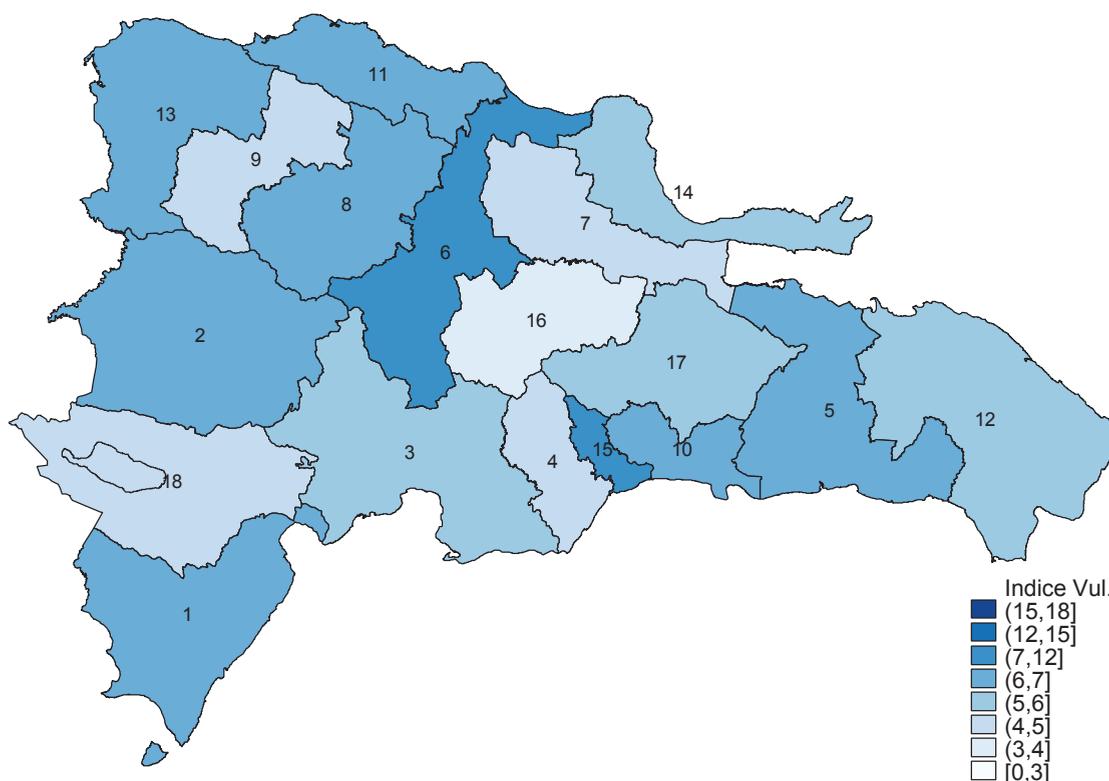
En la ilustración 1 se presenta el mapa de vulnerabilidad escolar por distrito educativo, mientras que en la ilustración 2 se presenta por Regional. De manera visual, pueden describirse diferencias en los niveles de vulnerabilidad escolar por distrito educativo y por regional, lo que puede contribuir con la focalización de la política educativa en atención a las escuelas que más lo necesiten.

Ilustración 1. Mapa de Vulnerabilidad Escolar



² En los anexos se presentan estadísticas de vulnerabilidad por Regional y Distrito.

Ilustración 2. Mapa de Vulnerabilidad Escolar por Regional



Resultados y análisis

En la tabla 13 y 14 se presentan los resultados de las estimaciones del índice de vulnerabilidad para las escuelas de Media y Básica usando Mínimos Cuadrados Ordinarios. En ambas tablas se presentan los resultados asociados a las variables explicativas: sector privado, dificultad de acceso, matrícula total de estudiantes, número de escuelas públicas por distrito.

Además se incluyen unas variables *dummies* por regional, las cuales al excluir una de ellas, permite obtener un tipo de diferencia de media respecto a la regional excluida. En este caso, se excluyó la regional 15 que era una de las que tenía menor vulnerabilidad. En este orden de ideas, si los encargados de la política educativa quisieran enfocarse en las escuelas que estén muy desventajadas con respecto a la capital, los coeficientes estimados de las variables *dummies* por regional, pueden constituir una guía de focalización.

La regresión 1 de ambas tablas busca recoger el impacto del sector privado en la vulnerabilidad de las escuelas. El coeficiente estimado tanto para Media como para Básica tiene signo positivo y es estadísticamente significativo al 1%. En el caso de Media se evidencia que las escuelas que pertenecen al sector privado tienen 4.4 puntos por encima de las escuelas públicas y semioficiales en el índice de vulnerabilidad y que en el caso de las escuelas básicas tiene 4.19 puntos por encima. Esto implica que las escuelas privadas, en más del 99% de las veces son menos vulnerables que las escuelas públicas y semioficiales.

En el caso de los resultados de las estimaciones relacionadas con el número de escuelas públicas por distrito, tanto para Básica como Media, se obtienen coeficientes con signos negativos, aunque no son estadísticamente significativos. Esto significa que el tamaño del distrito no guarda relación con el score de vulnerabilidad escolar, por lo que en atención a la misma, las consideraciones de tamaño de los distritos no deberían ser foco de la política educativa.

En el caso de la dificultad de acceso se encuentra consistentemente que esta está relacionada negativamente con el índice de vulnerabilidad de las escuelas de Media y de Básica. En la regresión 3 de las tablas 13 y 14 se encuentra que este coeficiente es estadísticamente significativo al 1%. Las escuelas de Media con mayor dificultad de acceso tienen en promedio 2.37 puntos menos que las escuelas urbanas. En el caso de Básica presentado en la tabla 26, las escuelas tienen en promedio 2.64 puntos menos que las escuelas con menor dificultad de acceso. Esto implica que las escuelas que están más alejadas de centros urbanos son más vulnerables.

Para contabilizar en término de mediciones qué significan estas regresiones en atención a la desigualdad y vulnerabilidad, las escuelas públicas, que además se encuentran más alejadas de los centros urbanos obtienen en promedio un score de vulnerabilidad de 7 puntos menos que las escuelas privadas que están en un centro urbano.

Tabla 13
Determinantes de la vulnerabilidad para las escuelas de media

VARIABLE DEPENDIENTE: ÍNDICE DE VULNERABILIDAD DE MEDIA												
	1		2		3		4		5		6	
p-valor F	(0.000)	***	(0.000)	***	(0.000)	***	(0.000)	***	(0.000)	***	(0.000)	***
r-cuadrado	0.3149		0.1636		0.1753		0.1651		0.2784		0.3299	
N. Observaciones	1775		1801		1775		1769		1769		1769	
N. de Clúster (Distrito)	104		104		104		104		104		104	
VARIABLES												
Privado	4.3861	***							5.0975	***	4.392269	***
Escuelas Públicas por distrito			-0.0193						-0.0120		-0.017229	*
Dificultad de acceso					-0.7934	***			-0.1947		-0.140156	
Matrícula total de estudiantes							0.00119	***	0.0020	***	0.001635	***
Regional 1	-2.5523	***	-5.2987	***	-5.1386	***	-5.0179	***			-2.666133	***
Regional 2	-2.8220	***	-5.0852	***	-4.9117	***	-5.1220	***			-2.473233	***
Regional 3	-2.5474	***	-3.8672	***	-4.3165	***	-4.3886	***			-1.970488	***
Regional 4	-3.1032	***	-4.6618	***	-4.4244	***	-4.7402	***			-3.250256	***
Regional 5	-0.9438		-1.6320	*	-1.5737	*	-1.5991				-1.091943	**
Regional 6	-0.8426		-1.8723	**	-1.8979	**	-2.2658	**			-0.433663	
Regional 7	-4.4750	***	-6.3672	***	-6.0361	***	-6.2978	***			-4.199624	***
Regional 8	-2.1678	**	-2.9166	***	-2.6916	**	-2.8777	**			-2.012294	**
Regional 9	-2.5725	***	-4.7841	***	-4.5731	***	-4.6433	***			-2.541393	***
Regional 10	-1.2278		-1.5200		-1.4511		-1.5267				-1.193136	
Regional 11	-1.6751	***	-3.2727	***	-2.2698	***	-2.8694	***			-1.537642	***
Regional 12	-1.0650		-2.2934		-2.4702	*	-2.5171	*			-0.544183	
Regional 13	-3.4381	***	-6.6434	***	-5.9825	***	-6.0410	***			-3.486537	***

VARIABLES												
Regional 14	-3.3936	***	-4.9189	***	-4.6403	***	-5.1364	***			-2.846239	***
Regional 15	Omitida		Omitida		Omitida		Omitida				Omitida	
Regional 16	-3.4172	***	-5.5297	***	-5.0057	***	-5.2406	***			-3.2643	***
Regional 17	-1.3461		-3.8271	***	-3.2232	**	-3.6279	***			-0.9718	
Regional 18	-2.7110	***	-6.2088	***	-5.6328	***	-5.4890	***			-3.0383	***

P-valores entre paréntesis. Se presentan los coeficientes estimados asociados a cada variable. Se señalan con asteriscos aquellos coeficientes que sean estadísticamente significativos: *** (1%), ** (5%), * (10%), respectivamente. Los errores estándares con los que se realizan las pruebas de significación individual de cada coeficiente se calculan por medio de clúster a nivel distrital.

La matrícula total de estudiantes que es un indicativo del tamaño promedio de la escuela no muestra una consistencia en los resultados. Para media parece estar relacionada positivamente con el índice de vulnerabilidad. Mientras que para básica está relacionada negativamente con los resultados de la escuela.

Asimismo, cuando se controla por todas las variables simultáneamente se encuentra que para Media el sector y la dificultad de acceso son estadísticamente significativas al 1%, y se mantienen los signos esperados para Básica y para Media. En adición a lo dicho, en el caso de Media, el número de escuelas públicas por distrito está relacionado negativamente al 10% con el índice de vulnerabilidad.

Tabla 14
Determinantes del índice de vulnerabilidad para las escuelas de básica

VARIABLE DEPENDIENTE: ÍNDICE DE VULNERABILIDAD BÁSICA												
	1		2		3		4		5		6	
p-valor F	(0.000)	***	(0.000)	***	(0.000)	***	(0.000)	***	(0.000)	***	(0.000)	***
r-cuadrado	0.2282		0.0899		0.116		0.095		0.1994		0.2334	
N. Observaciones	3577		3583		3577		3567		3567		3567	
N. de Clúster (Distrito)	104		104		104		104		104		104	
VARIABLES												
Privado	4.1984	***							4.1992	***	3.876789	***
Escuelas Públicas por distrito			-0.0176						-0.0065		-0.013388	
Dificultad de acceso					-0.8808	***			-0.2589	**	-0.286602	***
Matrícula total de estudiantes							-0.00132	***	-0.0001		-0.00025	
Regional 1	-0.1021		-2.4903		-1.9896		-2.20328				-0.476726	
Regional 2	-0.9453		-3.0133	**	-2.3329		-3.26458	**			-0.884915	
Regional 3	-1.9489	*	-3.1811	***	-3.2485	***	-3.73293	***			-1.706062	*
Regional 4	-2.6598	***	-4.3313	***	-3.3831	***	-4.22139	***			-2.665802	***
Regional 5	-1.0591		-2.07	**	-1.4143		-1.89578	*			-1.199987	*
Regional 6	0.1935		-1.2559		-0.607		-1.73994				0.469726	
Regional 7	-1.4894		-3.42	***	-2.4205	**	-3.68703	***			-1.483512	*
Regional 8	-0.7798		-1.7756		-1.2477		-1.81875				-0.777784	
Regional 9	-1.5216		-3.6928	**	-2.8774	**	-3.81218	***			-1.657728	*
Regional 10	-1.1717		-1.3188		-1.1471		-1.28977				-1.17289	

VARIABLES											
Regional 11	-0.4058		-2.3195	**	-0.8885		-2.27238	**			-0.492812
Regional 12	-1.3314		-2.7667	**	-2.1979	*	-2.93849	**			-0.971617
Regional 13	-0.7804		-3.5628	***	-2.3279	*	-3.41509	***			-1.170279
Regional 14	-1.323		-3.1511	**	-2.1099		-3.52449	***			-1.15196
Regional 15	Omitida		Omitida		Omitida		Omitida				Omitida
Regional 16	-2.4282	***	-4.2274	***	-3.4517	***	-4.39483	***			-2.490775
Regional 17	-1.0591		-3.6044	***	-2.3842	**	-3.68946	***			-1.098809
Regional 18	-3.0434	**	-6.1154	***	-5.2176	***	-5.66036	***			-3.639901

P-valores entre paréntesis. Se presentan los coeficientes estimados asociados a cada variable. Se señalan con asteriscos aquellos coeficientes que sean estadísticamente significativos: *** (1%), ** (5%), * (10%), respectivamente. Los errores estándares con los que se realizan las pruebas de significación individual de cada coeficiente se calculan por medio de clúster a nivel distrital.

También puede observarse la diferencia de vulnerabilidad por regional en comparación con la regional 15.

En la tabla 14 se muestra cómo cuando se controla por todas las variables sugeridas, la Regional 4 obtiene 2.66 puntos menos que la regional 15 en el índice de vulnerabilidad. El coeficiente estimado a la *dummy* de la regional 4 resultó estadísticamente significativo al 1%, lo que implica que el 99% de las veces las escuelas de la Regional 4 obtienen 2.66 puntos menos en el índice de vulnerabilidad que las escuelas de la Regional 15.

Del mismo modo, las regionales 16 y 18 son más vulnerables que la Regional 15. Las regionales 3, 5 y 7 obtienen 1.7, 1.2, 1.5, 1.7 menos puntos que la Regional 15 con un nivel de significancia del 10%, respectivamente.

Cuando se hace el mismo análisis para Media se obtiene que la mayoría de las regionales obtiene menos puntos que la Regional 15 y este resultado es estadísticamente significativo en casi todos los casos.

Como se ha argumentado a lo largo de toda la investigación, la metodología desarrollada y las métricas presentadas pueden servir para hacer consideraciones de política educativa en torno a la focalización en atención a la vulnerabilidad y desigualdad.

Se logró identificar características asociadas a la desigualdad, como son la dificultad de acceso y el sector al que pertenece el centro educativo. Asimismo, el índice de vulnerabilidad logró agregarse tanto para mostrar la vulnerabilidad de forma visual, así como las diferencias relativas a una regional o distrito particular.

Conclusiones

Con esta investigación se logró calcular un score de vulnerabilidad escolar que puede ser explicado por medio de diversas variables del sistema educativo como son el número de estudiantes, el número de escuelas por distrito y la dificultad de acceso a la escuela.

El *score* de vulnerabilidad logró agregarse a nivel regional y distrital. La relevancia de este índice y la manera como se presenta y se agrega, puede constituir una información que les permita a las autoridades educativas focalizar la política de seguimiento en atención a disminuir las brechas que existen entre las escuelas ya sea, por sector, por región, con políticas asociadas a disminuir la desigualdad o a reducir las brechas de vulnerabilidad.

La creación del índice de vulnerabilidad permite que los datos puedan ser desplegados en mapas, lo que facilita una comprensión visual de la economía de la vulnerabilidad educativa en la República Dominicana.

Por último, esta investigación hizo un uso intensivo y extensivo de los datos contenidos en las bases de datos más grandes del sistema educativo, lo que constituye una innovación que trata de trascender los meros informes univariados o unidimensionales que normalmente se encuentran en los informes del MINERD. Con la metodología y resultados de esta investigación, el MINERD y sus gestores pueden contar con información de inteligencia del sector educativo.

Bibliografía

Barry, M. J. & Gordon, S. (2004). *Data Mining Techniques*. (2Th. Ed.). New Jersey: Wiley.

Catalá, S., Con, M. & Lentini, L. (2008). Índice de Vulnerabilidad Educativa (IVE). *Informes temáticos de la dirección de investigación y estadística del Ministerio de Educación del GCBA*. Recuperado de <http://www.buenosaires.gov.ar/areas/educacion/dirinv/pdf/ive.pdf>

Collica, R. S. (2007). *CRM Segmentation and Clustering Using SAS Enterprise Miner*. SAS Institute.

Dirección General de Cultura y Educación de la Provincia de Buenos Aires. (2010). *Definiciones de vulnerabilidad educativa*. Recuperado de <http://servicios2.abc.gov.ar/lainstitucion/organismos/direccionprovincialplaneamiento/destacadoserie/definiciones-de-vulnerabilidad.pdf>

Fondo Nacional de Becas. (2010). *Índice de vulnerabilidad educativa*. Recuperado de <http://www.fonabe.go.cr/Informacion/EstadisticasGeneral/Documents/Índice%20Vulnerabilidad%20Educativa.pdf>

Ministerio de Educación de Chile. (2009). *Metodología de construcción de grupos socio-económicos en SIMCE 8vo Básico*. Recuperado de http://ww2.educarchile.cl/UserFiles/P0001/File/Construccion_de_grupos_comparables_de_establecimientos_SIMCE_2.pdf

Ministerio de Educación de la República Dominicana, Dirección de Análisis y Estudios Prospectivos (2014). *Base de datos del Sistema de Gestión de Centros*. Santo Domingo.

Ministerio de Educación de la República Dominicana, Dirección de Evaluación de la Calidad Educativa (2104). *Base de datos de Pruebas Nacionales*. Santo Domingo.

Secretaría de Estado de Educación de República Dominicana. (2008). *Plan Decenal de Educación 2008-2018: Un instrumento de trabajo en procura de la excelencia educativa*. Santo Domingo: SEE.

República Dominicana. (2012). *Ley Estrategia Nacional de Desarrollo*. Santo Domingo: el autor.

República Dominicana. (2014). *Pacto Educativo de la República Dominicana*. Santo Domingo: el autor.

Anexos

Sistema de Gestión de Centros

En la tabla A1 se encuentran las descripciones de las variables del Sistema de Gestión de Centros, tales como zona, sector, dirección, municipio, regional, etc.

Tabla A1
Datos del Sistema de Gestión de Centros

VARIABLE	FORMATO	DESCRIPCIÓN	COMENTARIOS
Centros	Númérico-Carácter	Está formada por un código de cinco dígitos asignado a cada escuela por el SGCE seguido del nombre correspondiente al centro y de la especificación del nombre de dicho centro.	N/A
Plantas Físicas	Númérico-Carácter	Contiene un código de ocho dígitos asignado a cada escuela además del nombre de la planta física del centro.	El nombre de la planta no necesariamente es igual al nombre de la escuela especificado en centros. Además, dos o más escuelas pueden compartir el mismo código de planta física.
Zona	Carácter	Indica la zona del país donde está ubicada la escuela. Estos lugares pueden ser: rural, rural aislada, rural turística, urbana, urbana marginal, urbana turística.	N/A
Dirección	Carácter	Indica la dirección de la escuela por número y nombre de calle.	N/A
Provincia	Carácter	Indica una de las provincias de la República Dominicana a la que pertenece la escuela.	N/A
Municipio	Carácter	Indica uno de los municipios de la República Dominicana al que pertenece la escuela.	N/A
Sector	Carácter	Indica el sector al que pertenece la escuela, este puede ser: público, privado o semioficial.	N/A
Distrito Educativo	Carácter	Es el distrito de educación al que pertenece la escuela. Esta variable está conformada por un código de 4 dígitos donde los dos primeros indican el número de la regional y los dos siguientes el número del distrito de cada regional. Además tiene el nombre por el que se identifica cada distrito.	N/A
Dirección Regional	Númérico- Carácter	Está compuesta por el número de la regional a la que pertenece la escuela seguido del nombre de esta regional.	N/A

En la tabla A2 se presentan las descripciones de las variables del Sistema de Gestión de Centros relacionadas con las modalidades y las cantidades de estudiantes como son la matrícula total, Inicial, básica y media.

Tabla A2
Datos del Sistema de Gestión de Centros con sus modificaciones

VARIABLE	FORMATO	DESCRIPCIÓN PRIMERA SUBDIVISIÓN	COMENTARIOS
Matrícula total	Númérico	Indica la cantidad total de estudiantes que tiene la escuela para todas sus modalidades y condiciones para los años escolares 2012-2013 2013-2014.	N/A

VARIABLE	FORMATO	DESCRIPCIÓN PRIMERA SUBDIVISIÓN	COMENTARIOS
Inicial	Numérico	Esta variable se subdivide en otras tres variables que son: Maternal, Kinder y Pre-primario. Cada una contiene el número de estudiantes inscritos en estos cursos para los años escolares 2012-2013, 2013-2014.	La sumatoria de las variables maternal, kínder y Pre-primario debe ser igual a la matrícula en inicial.
Básico	Numérico	Esta variable se subdivide en otras ocho variables que son los grados de primero a octavo. Cada una contiene el número de estudiantes inscritos en estos cursos para los años escolares 2012-2013, 2013-2014.	La sumatoria de las variables debe ser igual a la matrícula de básica.
Media	Numérico	Esta variable se subdivide en otras cuatro variables que son los grados de primero a cuarto de básica. Cada una contiene el número de estudiantes inscritos en estos cursos para los años escolares 2012-2013, 2013-2014.	La sumatoria de las variables: Primero, Segundo, Tercero y Cuarto debe ser igual a la matrícula de media.

La tabla A3 describe las variables del Sistema de Gestión de Centros relacionadas con la condición de los estudiantes, como el número de promovidos, reprobados y desertores.

Tabla A1
Datos del Sistema de Gestión de Centros relacionados con la condición de los estudiantes

VARIABLE	FORMATO	DESCRIPCIÓN PRIMERA SUBDIVISIÓN	COMENTARIOS
Promovido	Numérico	Indica el número de estudiantes promovidos para cada modalidad (inicial, básica, media).	N/A
Reprobado	Numérico	Indica el número de estudiantes reprobados en las modalidades (básica, media).	Esta variable no está disponible para inicial.
Abandono	Numérico	Indica el número de estudiantes que abandonan la escuela en cada modalidad (inicial, básica, media).	N/A

Los detalles de las variables que se construyeron a partir de la base de datos del Sistema de Gestión de Centro se presentan en la tabla A2. Estas variables son el detalle del número de estudiantes por modalidad. Por ejemplo: matrícula inicial, matrícula de Básica, matrícula Media, matrícula sin inicial.

La tabla A4 presenta la descripción de las variables construidas relacionadas con la ubicación y el sector al que pertenecen las escuelas. Estas son la dificultad de acceso y el indicativo de sector privado.

Tabla A2
Variables construidas del Sistema de Gestión de Centros

VARIABLE	NOMBRE	REGLA DE CONSTRUCCIÓN
Dificultad de acceso	Dificultad_acceso	Es el resultado de la asignación de valores a las zonas en las que se pueden encontrar las escuelas. Urbana es igual a 1, Urbana Marginal y Urbana turística 2, Rural es igual a 3, Rural aislada y Rural Turística igual a 4.
Privado	Privado	Es una variable dummy que toma valor 1 si la escuela es privada y cero (0) si pública o semi-oficial.

La tabla A5 muestra que la proporción de escuelas privadas de media es de 42.1% y que en promedio de acuerdo a los valores asignados hay un mayor número de escuelas fáciles de acceder.

Tabla A3
Estadística sobre la dificultad de acceso y el sector de las escuelas de Media

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
Dificultad/acceso	1.496	0.837	1	4	1796	680
Privado	0.421	0.494	0	1	1796	680

La tabla A6 señala la proporción de escuelas privadas de Básica es de 31.2% y el promedio de las escuelas difíciles de acceder el cual es mayor que el de las escuelas de Media.

Tabla A4
Estadística sobre la dificultad de acceso y el sector de las escuelas de básica

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
Dificultad acceso	1.949	1.021	1	4	3625	854
Privado	0.312	0.463	0	1	3625	854

En la tabla A7 se presenta la operacionalización de la matrícula total de cada escuela por cada uno de los niveles.

Tabla A7
Detalle de las variables modificadas de la base de datos del Sistema de Gestión de Centros

VARIABLE	FORMATO	SUBDIVISIÓN	DESCRIPCIÓN PRIMERA SUBDIVISIÓN	COMENTARIOS
Matrícula total	Numérico	Matrícula total 2012-2013 Matrícula total 2013-2014	Indica la cantidad total de estudiantes que tiene la escuela.	N/A
Inicial	Numérico	Maternal Kinder Pre-primario	Esta variable se subdivide en otras tres variables que son: maternal, kinder y pre-primario. Cada una contiene el número de estudiantes inscritos en estos cursos para los años escolares 2012-2013 y 2013-2014.	La sumatoria de las variables maternal, kinder y pre-primario debe ser igual a la matrícula en inicial.
Básico	Numérico	Primero	Esta variable se subdivide en otras ocho variables que son los grados de primero a octavo. Cada una contiene el número de estudiantes inscritos en estos cursos para los años escolares 2012-2013, 2013-2014.	La sumatoria de las variables debe ser igual a la matrícula de básica.
		Segundo		
		Tercero		
		Cuarto		
		Quinto		
		Sexto		
		Séptimo		
		Octavo		
Medio	Numérico	Primero	Esta variable se subdivide en otras cuatro variables que son los grados de primero a cuarto de Media. Cada una contiene el número de estudiantes inscritos para los años escolares 2012-2013, 2013-2014.	La sumatoria de las variables primero, segundo, tercero y cuarto debe ser igual a la matrícula de media.
		Segundo		
		Tercero		
		Cuarto		

En la tabla A8 se presenta la estadística descriptiva de la matrícula total por escuelas de educación Básica. La media de estudiantes en Básica a nivel nacional para el año escolar 2012-2013 era de 408.53 estudiantes por escuela, mientras que en el año escolar 2013-2014 fue de 394.5. Esta estadística presenta la matrícula promedio de la escuela Básica dominicana, con una desviación estándar amplia, lo que indica que hay un gran número de escuelas por debajo de la matrícula promedio, así como por encima. Un factor que llama la atención es la disminución de la matrícula promedio del año 2012-2013 al 2013-2014. Esto refleja en cierta forma que la política de construcción de escuelas está logrando desconcentrar a las mismas.

Tabla A8
Estadísticas descriptivas de la matrícula total por escuelas básicas

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX.	OBS.	DATOS FALTANTES
Total_2012_2013	408.53	371.23	1	2654	3615	842
Total_2013_2014	394.50	358.78	1	2945	3616	841

En la tabla A9 se presentan las estadísticas descriptivas de la matrícula total por Escuelas de Media. La media nacional de estudiantes en bachillerato para los años escolares 2012-2013 y 2013-2014 fue de 405.45 y 391.52, respectivamente. Al igual que en educación Básica, el promedio de matrícula por escuela disminuyó.

Tabla A9
Estadísticas descriptivas de la matrícula total por escuelas de media

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX.	OBS.	DATOS FALTANTES
Total_2012_2013	405.45	292.95	1	2214	1789	687
Total_2013_2014	391.52	286.86	1	2160	1781	695

Dificultad de Acceso y Proporción de Escuelas Privadas por Regional

Las siguientes tablas A10 y A11 muestran las estadísticas descriptivas de la dificultad de acceso y el sector privado para cada regional de las escuelas de Media y de Básica. Para Media se consigue que en la Regional tres hay un 23.3% de escuelas privadas y que la mayoría de las escuelas están en zonas no difíciles de acceder. En la Regional 11 hay una mayor proporción de escuelas difíciles de acceder y un 33% de escuelas privadas.

Tabla A10
Dificultad de acceso y sector de las escuelas de Media por Regional

REGIONAL	N.OBS	VARIABLE	PROMEDIO PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
1	48	Dificultad_acceso	1.176	0.576	1	3	34	14
		Privado	0.088	0.288	0	1	34	14
2	74	Dificultad_acceso	1.645	0.977	1	4	62	12
		Privado	0.129	0.338	0	1	62	12
3	85	Dificultad_acceso	1.384	0.860	1	4	73	12

REGIONAL	N.OBS	VARIABLE	PROMEDIO PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
		Privado	0.233	0.426	0	1	73	12
4	130	Dificultad_acceso	1.523	0.878	1	4	86	44
		Privado	0.314	0.467	0	1	86	44
5	171	Dificultad_acceso	1.256	0.585	1	4	121	50
		Privado	0.529	0.501	0	1	121	50
6	152	Dificultad_acceso	1.682	0.977	1	4	107	45
		Privado	0.346	0.478	0	1	107	45
7	103	Dificultad_acceso	1.695	0.990	1	4	82	21
		Privado	0.220	0.416	0	1	82	21
8	260	Dificultad_acceso	1.554	0.923	1	4	193	67
		Privado	0.492	0.501	0	1	193	67
9	53	Dificultad_acceso	1.410	0.850	1	4	39	14
		Privado	0.179	0.389	0	1	39	14
10	447	Dificultad_acceso	1.366	0.595	1	4	298	149
		Privado	0.587	0.493	0	1	298	149
11	93	Dificultad_acceso	2.230	1.080	1	4	74	19
		Privado	0.351	0.481	0	1	74	19
12	63	Dificultad_acceso	1.564	0.912	1	4	39	24
		Privado	0.308	0.468	0	1	39	24
13	48	Dificultad_acceso	1.568	0.959	1	4	37	11
		Privado	0.027	0.164	0	1	37	11
14	75	Dificultad_acceso	2.077	1.118	1	4	52	23
		Privado	0.231	0.425	0	1	52	23
15	470	Dificultad_acceso	1.238	0.477	1	3	349	121
		Privado	0.665	0.473	0	1	349	121
16	103	Dificultad_acceso	1.685	0.970	1	4	73	30
		Privado	0.233	0.426	0	1	73	30
17	64	Dificultad_acceso	2.043	1.074	1	4	46	18
		Privado	0.087	0.285	0	1	46	18
18	37	Dificultad_acceso	1.226	0.669	1	4	31	6
		Privado	0.032	0.180	0	1	31	6

Tabla A11
Dificultad de acceso y sector de las escuelas de Básica por Regional

REGIONAL	N.OBS	VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX.	OBS	DATOS FALTANTES
1	92	Dificultad acceso	1.548	0.898	1	4	73	19
		Privado	0.096	0.296	0	1	73	19
2	161	Dificultad acceso	2.180	1.086	1	4	133	28
		Privado	0.090	0.288	0	1	133	28
3	228	Dificultad acceso	1.755	1.057	1	4	188	40
		Privado	0.197	0.399	0	1	188	40
4	272	Dificultad acceso	2.237	1.070	1	4	215	57
		Privado	0.233	0.423	0	1	215	57
5	307	Dificultad acceso	1.823	0.978	1	4	248	59
		Privado	0.407	0.492	0	1	248	59
6	331	Dificultad acceso	2.346	1.034	1	4	286	45
		Privado	0.196	0.398	0	1	286	45
7	275	Dificultad acceso	2.388	0.982	1	4	219	56
		Privado	0.155	0.363	0	1	219	56
8	403	Dificultad acceso	1.878	1.007	1	4	352	51
		Privado	0.378	0.486	0	1	352	51
9	111	Dificultad acceso	2.109	1.104	1	4	92	19
		Privado	0.109	0.313	0	1	92	19
10	616	Dificultad acceso	1.512	0.738	1	4	490	126
		Privado	0.563	0.496	0	1	490	126
11	202	Dificultad acceso	2.583	1.035	1	4	163	39
		Privado	0.215	0.412	0	1	163	39
12	132	Dificultad acceso	2.287	1.094	1	4	94	38
		Privado	0.202	0.404	0	1	94	38
13	110	Dificultad acceso	2.202	1.033	1	4	94	16
		Privado	0.043	0.203	0	1	94	16
14	154	Dificultad acceso	2.610	1.029	1	4	123	31
		Privado	0.138	0.347	0	1	123	31
15	637	Dificultad acceso	1.373	0.572	1	4	515	122
		Privado	0.590	0.492	0	1	515	122
16	218	Dificultad acceso	2.110	1.077	1	4	163	55
		Privado	0.190	0.394	0	1	163	55
17	135	Dificultad acceso	2.518	1.033	1	4	114	21
		Privado	0.035	0.185	0	1	114	21
18	95	Dificultad acceso	1.683	0.930	1	4	63	32
		Privado	0.016	0.126	0	1	63	32

Índice de Condición: Tablas de Frecuencia

Las tablas A12 y A13 presentan las cantidades de escuelas por índice de condición para las escuelas de Media y Básica. En ellas se encuentra que de las escuelas de Básica, el 40% se encuentra en una condición menor al 50% del valor del índice mientras que para Media hay un 42%.

Tabla A12
Frecuencia índice de condición media

ÍNDICE MEDIA CONDICIÓN 2013 2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
0	229	12.62	229	12.62
1	273	15.05	502	27.67
2	261	14.39	763	42.06
3	281	15.49	1044	57.55
4	275	15.16	1319	72.71
5	265	14.61	1584	87.32
6	230	12.68	1814	100
Datos Faltantes = 662				

Tabla A13
Frecuencia índice de condición Básica

ÍNDICE CONDICIÓN BÁSICA 2013 2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
0	365	10.07	365	10.07
1	501	13.82	866	23.9
2	611	16.86	1477	40.76
3	674	18.6	2151	59.35
4	609	16.8	2760	76.16
5	493	13.6	3253	89.76
6	371	10.24	3624	100
Datos Faltantes = 855				

Pruebas Nacionales

En la tabla A14 se describen las variables que contiene la base de datos de Pruebas Nacionales como son modalidad, códigos, cantidad de alumnos y el distrito al que pertenecen las escuelas. Cada registro contiene información sobre los centros educativos participantes de la primera convocatoria de las Pruebas Nacionales para las modalidades de Básica y Media.

Tabla A14
VARIABLES de la base de Pruebas Nacionales

VARIABLE	FORMATO	DESCRIPCIÓN	COMENTARIOS
Modalidad	Carácter	Indica la modalidad en la que funciona cada escuela esta puede ser básica, modalidad general o técnico profesional.	N/A
Distrito	Numérico	Es el distrito de educación al que pertenece la escuela. Esta variable está conformada por un código de 4 dígitos donde los dos primeros indican el número de la Regional y los dos siguientes el número del distrito de cada regional.	N/A
Código de Pruebas Nacionales	Numérico	Contiene un código de ocho dígitos asignado a cada escuela. Este código es el código de planta física de las escuelas.	Varias escuelas pueden compartir el mismo código de prueba.
Código del SGC	Numérico	Es el código de centro de las escuelas.	Este es compartido por las diferentes modalidades de una misma escuela.
Alumnos	Numérico	Indica la cantidad de alumnos de cada escuela que tomaron las Pruebas Nacionales en la primera convocatoria.	Esta cantidad de estuantes corresponde a cada modalidad.
Nombre del centro	Carácter	Indica el nombre de cada escuela	N/A

Índice de Desempeño: Tablas de Frecuencia

Las tablas A15 y A16 presentan las cantidades de escuelas de Básica y Media para el índice de desempeño. En esta se muestra que para Media hay 238 escuelas con un desempeño de cero, mientras que para Básica existen 359 escuelas con un desempeño de cero (0).

Tabla A15
Frecuencia índice de desempeño para media

ÍNDICE MEDIA DESEMPEÑO 2013 2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
0	238	9.86	238	9.86
1	152	6.3	390	16.16
2	165	6.84	555	23
3	186	7.71	741	30.71
4	194	8.04	935	38.75
5	174	7.21	1109	45.96
6	203	8.41	1312	54.37
7	195	8.08	1507	62.45
8	188	7.79	1695	70.24
9	176	7.29	1871	77.54
10	126	5.22	1997	82.76
11	143	5.93	2140	88.69
12	273	11.31	2413	100
Datos Faltantes = 63				

Tabla A16
Frecuencia índice de desempeño para básica.

ÍNDICE BÁSICA DESEMPEÑO 2013 2014	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
0	359	8.17	359	8.17
1	334	7.6	693	15.77
2	358	8.15	1051	23.92
3	346	7.87	1397	31.79
4	333	7.58	1730	39.37
5	347	7.9	2077	47.27
6	345	7.85	2422	55.12
7	288	6.55	2710	61.68
8	339	7.72	3049	69.39
9	311	7.08	3360	76.47
10	277	6.3	3637	82.77
11	259	5.89	3896	88.67
12	498	11.33	4394	100
Datos Faltantes = 85				

Integración de Bases de Datos

Se describe cómo se integraron las bases de datos del Sistema de Gestión de Centros y de las Pruebas Nacionales de los años escolares 2012-2013 y 2013-2014, y se detalla el proceso de integración por el cual se combinaron las variables construidas o modificadas con las variables originales para obtener la base de datos con la cual se realizó la investigación sobre la vulnerabilidad en República Dominicana. Este proceso fue posible gracias a los códigos de centros educativos presentes en ambas bases de datos.

No obstante, cabe destacar que en el proceso de integración se encontró con problemas de duplicidad de datos por el hecho de que a pesar de que el código de centro es exclusivo, de cada centro pueden encontrarse varios registros con el mismo código, los cuales corresponden al mismo centro pero para diferentes modalidades como antes fue explicado.

Para minimizar este problema de duplicidad de datos se tomó la decisión de separar las bases de datos por modalidad, una para Básica y otra para Media para los años escolares 2012-2013 y 2013-2014. La primera cuenta con un total de 4,479 observaciones mientras que la segunda tiene la información para 2,476 escuelas. A continuación se presentan los pasos llevados a cabo para lograr estas bases de datos finales.

1. Se le asignó una clasificación de modalidad a las escuelas del Sistema de Gestión de Centros Educativos basada en si tenían o no estudiantes inscritos en las modalidades de Básica y Media.
2. Se separó la base de datos del SGCE de acuerdo con las modalidades de Básica y Media asignadas. Con esta segmentación se eliminó el problema de duplicidad del SGCE. La base de datos correspondiente a Media tiene 1,870 escuelas mientras que la de Básica tiene 6,908 escuelas.

- Del mismo modo, se separaron los datos de Pruebas Nacionales y se agruparon las informaciones de los dos años escolares disponibles. Así, se obtuvieron dos bases de datos: una para Básica con 4,479 escuelas y otra para Media con 2,476. En este caso, como los datos de Pruebas Nacionales venían con la especificación de si era Básica de adolescentes o Básica de adultos, de Media general o de Media técnico-profesional pueden existir más de dos registros con el mismo código de centro educativo. Para Básica existen 17 escuelas en esta condición mientras que para Media hay 91 escuelas.
- Finalmente, se integraron las bases de datos correspondientes a Media y Básica del SGCE y de las Pruebas Nacionales utilizando como conector el código de centro educativo que posee cada escuela en ambas bases de datos. Obtuvimos como resultado una base de datos para Básica de 4,476 y otra para Media con 2,476.

Índice de Vulnerabilidad, Condición y Desempeño por Regional

A continuación se presentan los resultados de las estadísticas de las escuelas de Media de los índices construidos para cada Regional. Esta tabla A17 y A18 permite apreciar las diferencias entre regionales. Respecto al índice de condición puede verse que las regionales que se encuentran en peor estado, son las regionales 13, 6 y 4. Las regionales que tienen un peor desempeño son la 7 y 16. Asimismo, las más vulnerables son 18, 7 y 13.

Tabla A17
Estadísticas de los índices por Regional

REGIONAL	N.OBS	VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
1	48	Índice Media Desempeño 2013 2014	5.022	3.276	0	11	46	2
		Índice Media Condición 2013 2014	2.206	1.553	0	6	34	14
		Índice Media vulnerabilidad 2013- 2014	7.235	4.000	0	15	34	14
2	74	Índice Media Desempeño 2013 2014	4.685	3.628	0	12	73	1
		Índice Media Condición 2013 2014	2.344	1.825	0	6	61	13
		Índice Media vulnerabilidad 2013-2014	7.082	4.443	0	18	61	13
3	85	Índice Media Desempeño 2013 2014	4.482	3.548	0	12	85	0
		Índice Media Condición 2013 2014	2.889	1.820	0	6	72	13
		Índice Media vulnerabilidad 2013-2014	7.889	4.362	1	18	72	13
4	130	Índice Media Desempeño 2013 2014	4.698	3.656	0	12	126	4
		Índice Media Condición 2013 2014	2.092	1.675	0	6	87	43
		Índice Media vulnerabilidad 2013-2014	7.713	4.567	0	18	87	43
5	171	Índice Media Desempeño, 2013 2014	7.006	3.515	0	12	167	4
		Índice Media Condición 2013 2014	3.098	1.727	0	6	122	49
		Índice Media vulnerabilidad 2013-2014	10.800	4.399	0	18	120	51
6	152	Índice Media Desempeño 2013 2014	6.456	3.728	0	12	149	3
		Índice Media Condición 2013 2014	2.676	2.162	0	6	111	41
		Índice Media vulnerabilidad 2013-2014	10.036	4.843	0	18	111	41
7	103	Índice Media Desempeño 2013 2014	3.147	2.861	0	12	102	1
		Índice Media Condición 2013 2014	2.543	1.803	0	6	81	22
		Índice Media vulnerabilidad 2013-2014	5.877	4.072	0	18	81	22

REGIONAL	N.OBS	VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
8	260	Índice Media Desempeño 2013 2014	6.012	3.742	0	12	253	7
		Índice Media Condición 2013 2014	2.722	1.935	0	6	198	62
		Índice Media vulnerabilidad 2013-2014	9.347	4.722	0	18	196	64
6	53	Índice Media Desempeño 2013 2014	4.769	3.579	0	12	52	1
		Índice Media Condición 2013 2014	2.231	2.096	0	6	39	14
		Índice Media vulnerabilidad 2013-2014	7.615	4.140	0	16	39	14
10	447	Índice Media Desempeño 2013 2014	6.435	3.658	0	12	430	17
		Índice Media Condición 2013 2014	3.749	1.742	0	6	303	144
		Índice Media vulnerabilidad 2013-2014	10.789	4.845	0	18	298	149
11	93	Índice Media Desempeño 2013 2014	5.891	3.532	0	12	92	1
		Índice Media Condición 2013 2014	2.808	2.032	0	6	73	20
		Índice Media vulnerabilidad 2013-2014	9.288	4.727	1	18	73	20
12	63	Índice Media Desempeño 2013 2014	6.885	3.489	0	12	61	2
		Índice Media Condición 2013 2014	2.410	1.802	0	6	39	24
		Índice Media vulnerabilidad 2013-2014	9.667	4.765	1	18	39	24
13	48	Índice Media Desempeño 2013 2014	4.042	3.115	0	12	48	0
		Índice Media Condición 2013 2014	1.541	1.260	0	4	37	11
		Índice Media vulnerabilidad 2013-2014	6.081	3.883	0	15	37	11
14	75	Índice Media Desempeño 2013 2014	4.222	3.573	0	12	72	3
		Índice Media Condición 2013 2014	2.585	1.946	0	6	53	22
		Índice Media vulnerabilidad 2013-2014	7.170	4.693	0	18	53	22
15	470	Índice Media Desempeño 2013 2014	7.877	3.762	0	12	456	14
		Índice Media Condición 2013 2014	3.676	1.948	0	6	355	115
		Índice Media vulnerabilidad 2013-2014	12.279	4.863	1	18	351	119
16	103	Índice Media Desempeño 2013 2014	3.627	3.557	0	12	102	1
		Índice Media Condición 2013 2014	2.838	1.688	0	6	74	29
		Índice Media vulnerabilidad 2013-2014	6.878	4.716	0	18	74	29
17	64	Índice Media Desempeño 2013 2014	5.333	3.690	0	12	63	1
		Índice Media Condición 2013 2014	2.935	1.679	0	6	46	18
		Índice Media vulnerabilidad 2013 2014	8.587	4.246	0	18	46	18
18	37	Índice Media Desempeño 2013 2014	4.806	2.847	0	12	36	1
		Índice Media Condición 2013 2014	1.690	1.650	0	6	29	8
		Índice Media vulnerabilidad 2013 2014	6.690	2.880	0	13	29	8

La tabla A18 muestra los resultados de los índices construidos para las escuelas de Básica. Según esta las escuelas básicas más vulnerables son 18,16, 4 y la 3.

Tabla A18
Estadística descriptiva de los índices para las escuelas de Básica.

REGIONAL	N OBS	VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
1	92	Índice Básica desempeño 2013 2014	6.767	3.598	0	12	90	2
		Índice Básica Condición 2013 2014	2.194	1.813	0	6	72	20
		Índice Básica Vulnerabilidad 2013 2014	9.333	4.494	0	18	72	20
2	161	Índice Básica desempeño 2013 2014	6.145	4.306	0	12	159	2
		Índice Básica Condición 2013 2014	1.932	1.821	0	6	132	29
		Índice Básica Vulnerabilidad 2013 2014	8.432	5.143	0	18	132	29
3	228	Índice Básica desempeño 2013 2014	5.630	3.547	0	12	227	1
		Índice Básica Condición 2013 2014	2.383	1.753	0	6	188	40
		Índice Básica Vulnerabilidad 2013 2014	7.909	4.110	0	18	187	41
4	272	Índice Básica desempeño 2013 2014	4.574	3.434	0	12	265	7
		Índice Básica Condición 2013 2014	2.531	1.763	0	6	213	59
		Índice Básica Vulnerabilidad 2013 2014	7.333	4.151	0	18	213	59
5	307	Índice Básica desempeño 2013 2014	6.954	3.610	0	12	302	5
		Índice Básica Condición 2013 2014	2.693	1.692	0	6	251	56
		Índice Básica Vulnerabilidad 2013 2014	9.665	4.251	0	18	248	59
6	331	Índice Básica desempeño 2013 2014	7.046	3.734	0	12	324	7
		Índice Básica Condición 2013 2014	2.835	1.775	0	6	285	46
		Índice Básica Vulnerabilidad 2013 2014	10.007	4.550	0	18	283	48
7	275	Índice Básica desempeño 2013 2014	4.450	3.729	0	12	269	6
		Índice Básica Condición 2013 2014	3.227	1.729	0	6	220	55
		Índice Básica Vulnerabilidad 2013 2014	8.162	4.342	0	18	216	59

REGIONAL	N OBS	VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
8	403	Índice Básica desempeño 2013 2014	6.662	3.684	0	12	390	13
		Índice Básica Condición 2013 2014	2.870	1.845	0	6	353	50
		Índice Básica Vulnerabilidad 2013 2014	9.781	4.586	0	18	342	61
9	111	Índice Básica desempeño 2013 2014	4.991	3.502	0	12	110	1
		Índice Básica Condición 2013 2014	2.793	1.825	0	6	92	19
		Índice Básica Vulnerabilidad 2013 2014	7.967	4.181	0	17	91	20
10	616	Índice Básica desempeño 2013 2014	6.005	3.743	0	12	598	18
		Índice Básica Condición 2013 2014	3.847	1.507	0	6	491	125
		Índice Básica Vulnerabilidad 2013 2014	10.215	4.434	1	18	484	132
11	202	Índice Básica desempeño 2013 2014	6.040	3.888	0	12	200	2
		Índice Básica Condición 2013 2014	3.085	1.749	0	6	164	38
		Índice Básica Vulnerabilidad 2013 2014	9.494	4.601	0	18	162	40
12	132	Índice Básica desempeño 2013 2014	6.333	3.434	0	12	126	6
		Índice Básica Condición 2013 2014	1.742	1.654	0	6	93	39
		Índice Básica Vulnerabilidad 2013 2014	8.489	4.120	1	18	90	42
13	110	Índice Básica desempeño 2013 2014	5.718	3.836	0	12	110	0
		Índice Básica Condición 2013 2014	2.564	1.682	0	6	94	16
		Índice Básica Vulnerabilidad 2013 2014	8.426	4.413	0	17	94	16
14	154	Índice Básica desempeño 2013 2014	5.052	3.743	0	12	153	1
		Índice Básica Condición 2013 2014	3.114	1.733	0	6	123	31
		Índice Básica Vulnerabilidad 2013 2014	8.285	4.305	0	18	123	31
15	637	Índice Básica desempeño 2013 2014	7.248	4.001	0	12	628	9
		Índice Básica Condición 2013 2014	3.827	1.639	0	6	514	123
		Índice Básica Vulnerabilidad 2013 2014	11.485	4.974	0	18	509	128

REGIONAL	N OBS	VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
16	218	Índice Básica desempeño 2013 2014	3.884	3.308	0	12	216	2
		Índice Básica Condición 2013 2014	3.006	1.657	0	6	163	55
		Índice Básica Vulnerabilidad 2013 2014	7.377	4.003	0	18	162	56
17	135	Índice Básica desempeño 2013 2014	5.604	3.468	0	12	134	1
		Índice Básica Condición 2013 2014	2.395	1.555	0	6	114	21
		Índice Básica Vulnerabilidad 2013 2014	8.080	3.720	0	17	113	22
18	95	Índice Básica desempeño 2013 2014	4.333	3.493	0	12	93	2
		Índice Básica Condición 2013 2014	1.565	1.585	0	6	62	33
		Índice Básica Vulnerabilidad 2013 2014	5.984	3.783	0	13	62	33

Otras Variables de control estadístico

Se describen algunas variables adicionales como son el número de escuelas públicas por distrito y unas variables categóricas *dummies* indicadoras de regional. Estas variables se presentan en las tablas A19, A20 y A21.

Tabla A19
Descripción número de escuelas

VARIABLE	FORMATO	ORIGINAL	MODIFICADA
Número Escuelas Públicas	Numérico	Indica la cantidad de escuelas públicas por distrito.	N/A

Tabla A20
Estadística descriptiva de las escuelas públicas por regional

VARIABLE	PROMEDIO/ PROPORCIÓN	DESVIACIÓN ESTÁNDAR	MÍN.	MAX	OBS	DATOS FALTANTES
Escuelas públicas	56.02	23.89	15	119	105	18

Para determinar el tamaño de las regionales se contabilizan por medio de tablas de frecuencia el número de escuelas por Regional Educativa.³ De esta manera, en las tablas 4 y 5 se presentan la cantidad de escuelas de Básica y de Media por Regional. Por ejemplo, la Regional 2 nuestra tiene 161 escuelas básicas y 73 escuelas de Media, mientras que la Regional 8 tiene 403 escuelas en Básica y 253 escuelas de Media.

³ Lo mismo se hace por distrito educativo, pero por razones de espacio, y para que no resulte repetitivo no se incluye de forma explícita en este informe, pero sí en los análisis estadísticos.

Tabla 21
Análisis de frecuencia de escuelas básicas por regional

REGIONAL	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
1	92	2.06	92	2.06
2	161	3.61	253	5.68
3	228	5.12	481	10.79
4	268	6.01	749	16.81
5	303	6.8	1052	23.6
6	329	7.38	1381	30.98
7	273	6.13	1654	37.11
8	403	9.04	2057	46.15
9	111	2.49	2168	48.64
10	613	13.75	2781	62.4
11	202	4.53	2983	66.93
12	131	2.94	3114	69.87
13	110	2.47	3224	72.34
14	154	3.46	3378	75.79
15	634	14.22	4012	90.02
16	218	4.89	4230	94.91
17	133	2.98	4363	97.89
18	94	2.11	4457	100

Estos análisis de frecuencia resultan pertinentes para la economía de la vulnerabilidad. También tiene el potencial de cruzarse con otras variables como matrícula. Esta identificación también permite agregar índices que permitan elaborar estadísticas descriptivas de vulnerabilidad educativa por Regional.

Tabla A22
Análisis de frecuencia de escuelas de media por Regional

REGIONAL	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
1	48	1.97	48	1.97
2	73	3	121	4.97
3	85	3.49	206	8.46
4	127	5.21	333	13.67
5	171	7.02	504	20.69
6	148	6.08	652	26.77
7	101	4.15	753	30.91
8	253	10.39	1006	41.3
9	53	2.18	1059	43.47
10	441	18.1	1500	61.58
11	92	3.78	1592	65.35
12	63	2.59	1655	67.94

REGIONAL	FRECUENCIA	PORCENTAJE	FRECUENCIA ACUMULADA	PORCENTAJE ACUMULADO
13	47	1.93	1702	69.87
14	74	3.04	1776	72.91
15	460	18.88	2236	91.79
16	100	4.11	2336	95.89
17	64	2.63	2400	98.52
18	36	1.48	2436	100

En la tabla A 23 se presentan las variables categóricas *dummies* por Regional.

Tabla A23. Descripción *dummies* Regional

VARIABLE	FUENTE	ORIGINAL	SUBDIVISIÓN	MODIFICADA
Regional	Numérico	Indica el número de la Regional a la que pertenece cada escuela	Dummy_regional1	Cada una de las <i>dummies</i> toma valor de 1 para la Regional que indica y de 0 para todas las demás regionales.
			Dummy_regional2	
			Dummy_regional3	
			Dummy_regional4	
			Dummy_regional5	
			Dummy_regional6	
			Dummy_regional7	
			Dummy_regional8	
			Dummy_regional9	
			Dummy_regional10	
			Dummy_regional11	
			Dummy_regional12	
			Dummy_regional13	
			Dummy_regional14	
			Dummy_regional15	
			Dummy_regional16	
			Dummy_regional17	
			Dummy_regional18	



Instituto Dominicano de Evaluación e
Investigación de la Calidad Educativa

<http://www.ideice.gob.do/>

978-9945-499-18-6

