

UNIVERSIDAD POLITÉCNICA ESTATAL DEL CARCHI

POSGRADO



MAESTRÍA EN ESTADÍSTICA APLICADA

“Revisión de métodos y propuesta de metodología para la obtención de un índice de clasificación para la priorización de instituciones educativas de sostenimiento público utilizando métodos multivariados”

Trabajo de titulación previa la obtención del
Título de Magister en Estadística Aplicada

Autor: Jhony Leandro Zabala Celi

Tutor: Andrés Alejandro Galvis Correa

Tulcán, 2025

CERTIFICADO DEL TUTOR

Certifico que el maestrante Zabala Celi Jhony Leandro con el número de cédula 1710120435 ha elaborado el trabajo de “Revisión de métodos y propuesta de metodología para la obtención de un índice de clasificación para la priorización de instituciones educativas de sostenimiento público utilizando métodos multivariados”. Este trabajo se sujeta a las normas y metodología dispuestas en la Codificación del Reglamento de Régimen Académico y de Estudiantes de la Universidad Politécnica Estatal del Carchi con RESOLUCIÓN No. 171-CSUP-2023, por lo tanto, autorizo su presentación para la sustentación respectiva.



Firmado electrónicamente por:
ANDRES
ALEJANDRO
GALVIS CORREA

f.....

Andrés Alejandro Galvis Correa

TUTOR

Tulcán, abril de 2025

AUTORÍA DE TRABAJO

El presente trabajo de titulación constituye un requisito previo para la obtención del título de Magister en Estadística Aplicada.

Yo, Zabala Celi Jhony Leandro con cédula de identidad número 1710120435 declaro: que la investigación es absolutamente original, auténtica, personal y los resultados y conclusiones a los que he llegado son de mi absoluta responsabilidad.

f.....

Jhony Leandro Zabala Celi

AUTOR

Tulcán, abril de 2025

ACTA DE CESIÓN DE DERECHOS DEL TRABAJO DE TITULACIÓN

Yo, Zabala Celi Jhony Leandro declaro ser autor/a de los criterios emitidos en el trabajo de titulación: “Revisión de métodos y propuesta de metodología para la obtención de un índice de clasificación para la priorización de instituciones educativas de sostenimiento público utilizando métodos multivariados” y eximo expresamente a la Universidad Politécnica Estatal del Carchi y a sus representantes legales de posibles reclamos o acciones legales.

f.....

Jhony Leandro Zabala Celi

AUTOR

Tulcán, abril de 2025

AGRADECIMIENTO

"A quienes me enseñaron que cada caída es solo una oportunidad para levantarme con más fuerza, que la vida se recorre mejor con el alma liviana y que la solidaridad no debería ser una elección, sino un principio inherente al ser humano.

Gracias eternas, papá y mamá."

DEDICATORIA

A quienes cada día enfrentan sus propias batallas, aquellos que, con valentía y determinación, luchan por superar sus limitaciones, demostrando que la verdadera fuerza nace en la perseverancia.

A quienes, pese al esfuerzo y sacrificio, sienten que su lucha es en vano, recordándoles que cada paso dado, por pequeño que parezca, es un avance hacia sus sueños y un testimonio de su resiliencia.

Y, sobre todo, a quienes han estado a mi lado en cada momento, en las alegrías y en las dificultades, brindándome su apoyo incondicional, su compañía sincera y su aliento cuando más lo he necesitado.

Este logro es tanto mío como suyo.

ÍNDICE

RESUMEN.....	xvi
ABSTRACT	xvii
CAPÍTULO I.....	1
PROBLEMA	1
1.1 Planteamiento del problema	1
1.2 Preguntas de investigación o hipótesis.....	2
1.3 Objetivos de investigación	3
1.3.1 Objetivo General	3
1.3.2 Objetivos Específicos	3
1.4. Justificación	3
CAPÍTULO II.....	6
FUNDAMENTACIÓN TEÓRICA	6
2.1 Antecedentes de investigación.....	6
2.2 Marco teórico.....	9
2.2.1 Inversión en Educación y equidad en la distribución de recursos.	9
2.2.2 Factores asociados a las instituciones educativas fiscales que inciden en su clasificación.	11
2.2.3 Índices e Indicadores	16
2.2.4 Análisis Multivariado	18
2.2.5 Uso de técnicas multivariadas en la construcción de índices	19
2.3 Marco legal	22
CAPÍTULO III	25
METODOLOGÍA.....	25
3.1 Descripción del área de estudio/grupo de estudio	25
3.2 Enfoque y tipo de investigación	26

3.2.1 Enfoque.....	26
3.2.2 Tipo de Investigación	27
3.3 Definición y operacionalización de variables.....	27
3.4 Procedimientos	28
3.4.1 Análisis de Correspondencias Múltiples.	30
3.4.2 Construcción del Índice de Clasificación de Instituciones Educativas a partir del Análisis de Correspondencias Múltiples.	33
3.4.3 Análisis Componentes Principales No Lineales	33
3.4.4 Método "Principal Components Analysis by Nonlinear Least Squares" PRINCALS	36
3.4.5 Construcción del Índice de Clasificación de Instituciones Educativas a partir del Análisis de Componentes Principales No Lineales.....	39
CAPÍTULO IV	40
RESULTADOS Y DISCUSIÓN	40
4.1 Resultados del Análisis de Correspondencias Múltiples.....	40
4.2 Resultados Índice de Clasificación de Instituciones Educativas aplicando el Análisis de Correspondencias Múltiples.	45
4.2.1 Identificación de las Instituciones educativas que cuentan con condiciones menos favorables.....	46
4.3 Procedimiento Aplicación del Análisis de Componentes Principales No Lineales - Categórico.....	48
4.4 Distribución de puntuaciones del Índice de Clasificación de Instituciones Educativas aplicando el Análisis de Componentes Principales No Lineales.....	56
4.4.1 Identificación de las Instituciones educativas que cuentan con condiciones menos favorables ACPNL.	57
4.5. Comparación de resultados de la aplicación de las dos técnicas de clasificación.	58
4.6. Similitudes entre el Análisis de Correspondencias Múltiples y El Análisis de Componentes Principales Categórico.	61

4.7. Diferencias entre Análisis de Correspondencias Múltiple y Análisis de Componentes Principales No Lineales.....	62
4.8. Ventajas y Desventajas del ACM.....	63
4.9. Ventajas y desventajas del ACPNL.....	63
4.10 Determinación la metodología que permitirá clasificar de manera adecuada a las instituciones educativas.....	64
CONCLUSIONES.....	66
REFERENCIAS.....	68
ANEXOS.....	72

ÍNDICE DE TABLAS

Tabla 1	Distribución de instituciones educativas por provincias.....	25
Tabla 2	Variables y dimensiones	28
Tabla 3	Seudo Matriz	31
Tabla 4	Varianza explicada 5 dimensiones.....	40
Tabla 5	Dimensiones Análisis Correspondencias Múltiple.....	44
Tabla 6	Dimensiones por varianzas.....	44
Tabla 7	Totales por fila.....	44
Tabla 8	Distribución de instituciones por quintiles, según índice de ACM.....	47
Tabla 9	Distribución instituciones menos favorecidas, según índice ACM.....	47
Tabla 10	Cuantificadores ponderados por índice de escalamiento	54
Tabla 11	Distribución de instituciones educativas por quintiles, según puntuación del índice de clasificación	57
Tabla 12	Distribución de instituciones educativas por quintiles ACPNL, según puntuación del índice de clasificación por ubicación urbana - rural	58
Tabla 13	Matriz de Contingencia quintiles obtenidos a partir del Análisis de Componentes Principales vs quintiles obtenidos a partir de Análisis de Componentes Principales No lineales.....	60
Tabla 14	Transformación variable agua.....	78
Tabla 15	Transformación variable lavamanos.....	80
Tabla 16	Transformación variable Inodoro	80
Tabla 17	Resumen Variables categóricas transformadas	83
Tabla 18	Eigenvalores y porcentaje varianza explicada.....	93
Tabla 19	Resumen puntuaciones ACPNL	102

ÍNDICE DE FIGURAS

Figura 1 Pirámide de datos	16
Figura 2 Área de estudio.....	26
Figura 3 Scree plot Distribución de la Inercia entre las dimensiones	40
Figura 4 Curva distribución de puntuaciones ACM.	46
Figura 5 Mapa de puntos de ubicación instituciones educativas pertenecientes al quintil de instituciones menos favorecidas.	48
Figura 6 Varianza explicada por componente	49
Figura 7 Matriz de Correlaciones variables ACP No Lineal.....	50
Figura 8 Rotación Varimax	50
Figura 9 Loading Plot.....	52
Figura 10 Biplot de categorías e individuos	53
Figura 11 Curva de distribución de puntuaciones ACP No Lineal.....	56
Figura 12 Mapa de ubicación instituciones educativas menos favorecidas según ACP Categórico.....	58
Figura 13 Comparación de la distribución del Índice ACM vs ACPNL	59
Figura 14 Transformación variable Autoidentificación de estudiantes	74
Figura 15 Transformación variable Autoidentificación de estudiantes	75
Figura 16 Transformación variable Discapacidad de estudiantes	76
Figura 17 Transformación variables total estudiantes, total docentes en ratio estudiantes por docente	77
Figura 18 Transformación variable docentes con título cuarto nivel.....	77
Figura 19 Transformación variable Agua.....	79
Figura 20 Transformación variable Lavamanos	80
Figura 21 Transformación variable Inodoro.....	81
Figura 22 Transformación variable Abandono.....	81
Figura 23 Transformación variable Acceso a la Institución Educativa	82

Figura 24 Trasmformación variable Internet.....	83
Figura 25 Ejemplo categorías en el espacio.	92
Figura 26 Puntuación transformada variable autoidentificación estudiantes	94
Figura 27 Puntuación transformada variable auto identificación docentes.....	95
Figura 28 Puntuación transformada variable discapacidad estudiantes	96
Figura 29 Puntuación transformada variable ratio estudiantes por docente.....	96
Figura 30 Puntuación transformada variable docentes con título de cuarto nivel	97
Figura 31 Puntuación transformada variable docentes con nombramiento.....	98
Figura 32 Puntuación transformada variable abandono	98
Figura 33 Puntuación transformada variable Área de ubicación de la Institución Educativa	99
Figura 34 Puntuación transformada variable Acceso a la Institución Educativa	99
Figura 35 Puntuación transformada variable Agua en la Institución Educativa.....	100
Figura 36 Puntuación transformada variable lavamanos en la institución educativa	101
Figura 37 Puntuación transformada variable inodoros en la institución educativa	101
Figura 38 Puntuación transformada variable servicio de internet en la Institución Educativa.....	102

ÍNDICE DE ANEXOS

Anexo A. Certificado del abstract por parte de idiomas	72
Anexo B. Tratamiento de variables previo al ACM y ACPNL.....	74
Anexo C. Análisis de Correspondencia simple paso a paso	85
Anexo D. Porcentaje de varianza explicada 13 dimensiones	93
Anexo E. Valores transformados tras la cuantificación óptima en ACP No Lineal.....	94

RESUMEN

La presente investigación tiene como objetivo diseñar un índice que, utilizando técnicas de clasificación multivariada, facilita la toma de decisiones en la priorización de instituciones educativas fiscales para ser beneficiarias de recursos educativos a través de proyectos de inversión. Se identificaron factores determinantes y se compararon dos metodologías: el Análisis de Correspondencias Múltiples -ACM- y el Análisis de Componentes Principales No Lineal -ACP No Lineal-. Los resultados evidencian que, si bien el ACM genera un índice con menor variabilidad y mayor estabilidad, el ACP No Lineal captura de forma más precisa la complejidad y variabilidad de los datos, lo que lo hace idóneo para reflejar diferencias significativas entre las instituciones. Variables como la etnia de estudiantes y docentes, la relación estudiante-docente, y el acceso a servicios básicos como agua e internet se destacan como elementos fundamentales para la diferenciación y priorización. En este contexto, el ACP No Lineal se erige como la mejor opción para la construcción del índice, ya que permite capturar una mayor variabilidad y manejar de forma eficiente distintos tipos de variables categóricas - ordinales y nominales-. Su capacidad para reflejar de manera precisa la complejidad subyacente en los datos, junto con la habilidad para gestionar no linealidades y transformar los datos en puntuaciones numéricas, posibilita la creación de un índice más robusto y adaptativo. De esta manera, la metodología desarrollada optimiza la categorización de las instituciones en función de múltiples dimensiones, contribuyendo de manera significativa a la toma de decisiones en la asignación de recursos educativos a través de proyectos de inversión.

Palabras clave: Clasificación multivariada, ACM, ACP No Lineal, Instituciones Educativas.

ABSTRACT

The present research aims to design an index that, using multivariate classification techniques, facilitates decision-making in the prioritization of public educational institutions to be beneficiaries of educational resources through investment projects. Determinant factors were identified, and two methodologies were compared: Multiple Correspondence Analysis (MCA) and Nonlinear Principal Component Analysis (NLPCA). The results show that, while MCA generates an index with lower variability and greater stability, Non-Linear PCA captures the complexity and variability of the data more accurately, making it ideal for reflecting significant differences between institutions. Variables such as the ethnicity of students and teachers, the student-teacher relationship, and access to basic services like water and internet stand out as fundamental elements for differentiation and prioritization. In this context, Non-Linear PCA stands out as the best option for constructing the index, as it allows capturing greater variability and efficiently handling different types of categorical variables -ordinal and nominal-. Its ability to accurately reflect the underlying complexity in the data, along with the capacity to manage non-linearities and transform the data into numerical scores, enables the creation of a more robust and adaptive index. In this way, the developed methodology optimizes the categorization of institutions based on multiple dimensions, significantly contributing to decision-making in the allocation of educational resources through investment projects.

Keywords: Multivariate Classification, PCA, Nonlinear PCA, Educational Institutions

CAPÍTULO I

PROBLEMA

1.1 Planteamiento del problema

Para alcanzar sus objetivos educativos nacionales e internacionales muchos países deben invertir más y mejor en sus sistemas educativos. El Banco Mundial y la UNESCO a través del Observatorio de la financiación de la educación (EFW), realiza un informe cada año, y en su última publicación realizada para el año 2024, se señala

“El gasto total en educación de los gobiernos, los hogares y los donantes ha aumentado de manera constante en la última década. Sin embargo, esto no se ha traducido en grandes incrementos de las asignaciones por niño, especialmente en los países más pobres con poblaciones en expansión. De hecho, a nivel mundial, el gasto total en educación por niño no ha aumentado”(Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO] y Banco Mundial, 2024, p. 1).

A nivel de Latinoamérica las cifras publicadas por el Banco Mundial acerca de la inversión en educación no son alentadoras, en el año 2019 el gasto fue del 4.60% del PIB, para el año 2021 se redujo al 4.16% del PIB, ha esta problemática se debe sumar lo mencionado por el Observatorio de la Financiación en la Educación que menciona “gastar más no es suficiente si no se presta atención a la eficiencia y la equidad”, es decir que los recursos asignados se utilicen de forma óptima, maximizando los resultados en aprendizaje, infraestructura y acceso a oportunidades educativas.

El Estado ecuatoriano tiene la responsabilidad de garantizar la educación como un derecho fundamental para todas las personas a lo largo de sus vidas. Según las normas constitucionales del país, el Ministerio de Educación es el encargado de gestionar de manera eficiente y eficaz los recursos invertidos en el proceso educativo en Ecuador. De esta forma, se busca cumplir con los objetivos propuestos a nivel nacional, regional y mundial en cuanto al acceso a una educación de calidad para todas las personas, sin importar su sexo, edad, condición social o étnica.

De acuerdo con información de la UNESCO, el gasto en educación en Ecuador pasó de 4.568 millones de dólares en 2019 a 3.918 millones de dólares en 2021 (UNESCO Institute for Statistics, 2024). En consonancia, el Banco Mundial reporta que el gasto en educación como porcentaje del PIB disminuyó de 4,25% en 2019 a 3,65% en 2021 y al 3,89% en el año 2023 (Banco Mundial, 2024). Esto indica que, aunque las necesidades de asignación de recursos son mayores, los fondos entregados por el Estado han disminuido. Esta realidad contrasta

con lo establecido en la Disposición Transitoria Decimoctava de la Constitución de la República del Ecuador, que señala:

"El Estado asignará de forma progresiva recursos públicos del Presupuesto General del Estado para la educación inicial básica y el bachillerato, con incrementos anuales de al menos el cero punto cinco por ciento del Producto Interno Bruto hasta alcanzar un mínimo del seis por ciento del Producto Interno Bruto" (Asamblea Nacional de la República del Ecuador, 2008, p. 271).

En este contexto de recursos económicos insuficientes, organismos internacionales como BID, UNICEF, Banco Mundial, la Agencia Francesa para el Desarrollo AFD, la Asociación Alemana para Jóvenes y Adultos, entre otras, tienen convenios con el Ministerio de Educación, con el objeto de implementar proyectos de inversión en beneficio de la Comunidad Educativa, estos proyectos suelen ser focalizados y es allí donde es responsabilidad del Ministerio de Educación realizar la priorización de las instituciones beneficiarias de las intervenciones.

Por ejemplo, en el manual operativo del programa de Apoyo a la calidad e inclusión de la educación en Ecuador de la AFD, se menciona "Para la selección de las unidades educativas conforme a lo acordado en el Convenio de Préstamo, el MINEDUC, desarrolla una metodología de implementación y valoración de los criterios de elegibilidad y priorización para selección de unidades educativas" (Agencia Francesa para el Desarrollo [AFD], 2018, p. 15).

En este contexto el Ministerio de Educación ha empleado la determinación de parroquias o circuitos priorizadas para seleccionar instituciones educativas que sean beneficiarias de recursos adicionales, siguiendo siempre las disposiciones del marco normativo vigente, sin embargo, los proyectos de inversión muchas veces necesitan que la priorización sea a un nivel aún más detallado, es decir a nivel de las instituciones educativas.

En tal sentido se considera necesario la creación de un índice que, utilizando técnicas de clasificación multivariada, facilite la toma de decisiones en la priorización de instituciones educativas, para ser beneficiarias de recursos educativos a través de proyectos de inversión.

1.2 Preguntas de investigación o hipótesis

- ¿Puede la creación de un índice que, utilizando técnicas de clasificación multivariada, facilitar la toma de decisiones en la priorización de instituciones educativas, para ser beneficiarias de recursos educativos a través de proyectos de inversión?

- ¿Cuáles son los factores determinantes asociados a las instituciones educativas fiscales que contribuyen al diseño de un índice, utilizando técnicas de clasificación multivariada, para facilitar su priorización como beneficiarias de recursos educativos a través de proyectos de inversión?
- ¿Evaluar la posibilidad de comparar dos técnicas de clasificación multivariada para optimizar la categorización de instituciones educativas fiscales, contribuye al diseño de un índice que facilite la priorización de estas como beneficiarias de recursos educativos a través de proyectos de inversión?

1.3 Objetivos de investigación

1.3.1 Objetivo General

Diseñar un índice que, utilizando técnicas de clasificación multivariada, facilite la toma de decisiones en la priorización de instituciones educativas fiscales para ser beneficiarias de recursos educativos a través de proyectos de inversión.

1.3.2 Objetivos Específicos

- Analizar los factores asociados a las instituciones educativas fiscales que contribuyen al diseño de un índice, utilizando técnicas de clasificación multivariada, para facilitar su priorización como beneficiarias de recursos educativos a través de proyectos de inversión.
- Evaluar y comparar dos técnicas de clasificación multivariada para optimizar la categorización de instituciones educativas fiscales, contribuyendo al diseño de un índice que facilite su priorización como beneficiarias de recursos educativos a través de proyectos de inversión.

1.4. Justificación

La educación es fundamental para el desarrollo de la sociedad, ya que actúa como catalizador del progreso social y económico, y es esencial que la inversión estatal en este sector se realice de manera eficiente y equitativa. A pesar de que diversos estudios, como “El capital humano y los retornos a la educación en Ecuador” (Montenegro, 2015), indican que la educación básica presenta una tasa de retorno del 8 al 9%, se evidencia que la inversión pública no siempre alcanza a los colectivos más vulnerables. Por ejemplo, Blanco (2006) señala que, en contextos con datos desagregados, los grupos en situación de desventaja como niños de zonas rurales aisladas, de extrema pobreza, indígenas o con discapacidad son precisamente aquellos que requieren la mayor atención.

El Ministerio de Educación ha establecido múltiples convenios de cooperación con organismos internacionales, lo que ha permitido la implementación de proyectos de capacitación docente y reformas educativas, tales como la capacitación de 4501 docentes en 2021 y la adaptación del Currículo GlobALE para aproximadamente 5 mil educadores. Proyectos como el “Apoyo a la Reforma Educativa en los Circuitos Focalizados (PARECF)” financian intervenciones en unidades educativas priorizadas, lo que resalta la necesidad de definir claramente cuáles son las zonas y las instituciones que requieren mayor apoyo.

Ante este escenario, surge la interrogante fundamental: ¿cómo optimizar la asignación de recursos educativos para garantizar que los proyectos de inversión lleguen a las instituciones con mayor vulnerabilidad? La respuesta a esta pregunta es crucial, dado que la focalización precisa no solo mejora la eficiencia del gasto público, sino que también promueve una mayor equidad en el acceso a la educación. En este sentido, la metodología propuesta, basada en el análisis de múltiples indicadores a través de técnicas multivariadas, ofrece una herramienta objetiva para identificar de forma sistemática a las unidades educativas que presentan mayores necesidades.

Específicamente, se plantea la utilización del Análisis de Correspondencias Múltiples (ACM) y el Análisis de Componentes Principales No Lineal (ACP No Lineal) para la construcción de un índice de priorización. Aunque el ACM ofrece menor variabilidad y mayor estabilidad, el ACP No Lineal se destaca por capturar de manera más precisa la complejidad y diversidad de los datos, permitiendo el manejo de distintos tipos de variables categóricas (ordinales y nominales) y la transformación de estos en puntuaciones numéricas. Esto da como resultado un índice más robusto y adaptativo, capaz de responder a la heterogeneidad del contexto educativo ecuatoriano.

En conclusión, la aplicación de métodos estadísticos avanzados en la creación de índices educativos se justifica por la necesidad de mejorar la focalización de los recursos públicos, garantizando que las intervenciones se dirijan a las instituciones y comunidades educativas que más lo requieren. Esta investigación, por lo tanto, no solo contribuye al desarrollo de una herramienta innovadora para la toma de decisiones, sino que también responde a la imperante demanda de una asignación de recursos más equitativa y efectiva en el ámbito educativo.

Este estudio se alinea en la consecución de los Objetivos de Desarrollo Sostenible, especialmente en lo que respecta al ODS 4 Educación de Calidad y al ODS 10 Reducción de las Desigualdades. Al desarrollar un índice de priorización basado en técnicas multivariadas, la investigación proporciona una herramienta innovadora y objetiva para

identificar y focalizar las intervenciones en aquellas instituciones educativas que requieren mayor apoyo.

Finalmente, la presente investigación se encuentra alineada a la línea de investigación de Aplicación de la estadística en la solución de problemas del entorno.

CAPÍTULO II

FUNDAMENTACIÓN TEÓRICA

2.1 Antecedentes de investigación

La educación es un derecho fundamental de todos los individuos. Como señala González Teresa, "no se puede ejercer ninguno de los derechos civiles, políticos, sociales, económicos o culturales sin un mínimo de educación" (González y Luna Corvera, 2019, p. 2).

Para garantizar este derecho Schettino Piña en el documento El Derecho a La Educación menciona,

Debe haber instituciones y programas de enseñanza en cantidad suficiente en el ámbito del Estado Parte. Las condiciones para que funcionen dependen de numerosos factores, entre otros, el contexto de desarrollo en el que actúan; por ejemplo, las instituciones y los programas probablemente necesiten edificios u otra protección contra los elementos, instalaciones sanitarias para ambos sexos, agua potable, docentes calificados con salarios competitivos (Schettino Piña, 2017, p. 8).

Complementariamente se debe exigir una permanente rendición de cuentas, "no es cuestión no solo de exigir mayores cantidades de recursos, sino también de profundizar en mecanismos de financiamiento progresivos que tengan como contrapartes, rendiciones de cuentas a la sociedad en base a objetivos y metas evidenciables" (Operti, Renato, 2022, p.5). En este contexto, el Estado emerge como el principal actor responsable de garantizar el derecho a la educación, dotando de los recursos necesarios y siendo eficiente en su utilización. Al respecto, Ramoni Perazzi afirma:

El crecimiento económico y el desarrollo social estable del país requieren, entre otros factores, la presencia de un sector público eficiente y organizado que pueda contribuir significativamente al logro del desarrollo local y nacional, y con la capacidad de superar los problemas más críticos que pueda confrontar la economía. Para ello es importante desarrollar un proceso decisorio sólido y una eficiente administración de los recursos públicos, lo que supone la utilización de instrumentos apropiados que, además, faciliten la concreción de las estrategias de desarrollo nacional, regional y local (Ramoni P. et al., 2013, p. 112).

La eficiencia y el impacto de la inversión pública se reflejan especialmente en el ámbito educativo, donde la asignación de recursos se realiza a través de diversas estrategias y normativas que buscan atender las necesidades de las instituciones beneficiarias. En Ecuador, se entregan recursos educativos como alimentación escolar, textos y uniformes.

Para estos casos, existen normativas que establecen qué tipo de instituciones son las beneficiarias.

Sin embargo, el Estado no es el único actor en la actualidad, existen organismos internacionales comprometidos con el financiamiento de proyectos que mejoren los servicios educativos.

“Las corporaciones más interesadas en la problemática educativa son: Banco Mundial (BM); Organización de las Naciones Unidas para la Educación la Ciencia y la Cultura (UNESCO); la Organización de Cooperación y Desarrollo Económico (OCDE) y, en el ámbito latinoamericano, el Banco Interamericano de Desarrollo (BID) y la Comisión Económica para América Latina (CEPAL)”

Además, se otorgan otros recursos mediante proyectos de inversión, financiados generalmente por organismos internacionales, donde los beneficiarios se determinan en función de los objetivos de los proyectos y considerando factores clave que inciden en la realidad que se desea transformar.

Un ejemplo de ello es el Proyecto de Apoyo a la Reforma Educativa en los Circuitos Focalizados (PARECF), financiado a través de un convenio legal entre el Estado ecuatoriano y el Banco Mundial, destinado a intervenir en unidades educativas ubicadas en circuitos priorizados. Sin embargo, el manual operativo de dicho proyecto señala que:

Los límites y las escuelas de estos circuitos (que constituyen una agrupación administrativa de escuelas) han cambiado año tras año, socavando la posibilidad de producir indicadores por circuito que sean estables a lo largo del tiempo. Por tal razón, se decidió utilizar la parroquia (la próxima desagregación territorial más grande), como la unidad territorial para orientar las intervenciones del Proyecto, ya que esta unidad se ha mantenido estable en el tiempo (Ministerio de Educación, 2022, p. 17).

No obstante, la utilización de la parroquia como unidad territorial presenta desafíos. Existen varias parroquias que cuentan con más de una unidad educativa, y no todas son objeto de intervención. Además, esta estrategia resulta excluyente para otras instituciones que pueden presentar igual o mayor necesidad de apoyo, pero que no son consideradas por encontrarse en una parroquia que no ha sido priorizada.

En este contexto, disponer de un índice que permita clasificar e identificar de manera precisa a las instituciones educativas resulta pertinente y constituiría un instrumento fundamental para el análisis y selección de estas.

El uso de métodos estadísticos en el análisis de problemáticas del sector educativo es esencial para obtener una comprensión profunda y objetiva de la realidad. A través de técnicas como el análisis de regresión, se pueden identificar los factores que influyen en el rendimiento académico de los estudiantes, determinando la relación entre variables como la infraestructura escolar, el nivel socioeconómico de las familias y la capacitación docente.

Mediante el análisis de componentes principales (ACP), es posible reducir la dimensionalidad de grandes conjuntos de datos, identificando patrones y variables clave que afectan el rendimiento académico o la deserción escolar.

Por otro lado, el análisis de correspondencias múltiples (ACM) facilita la agrupación de instituciones educativas o estudiantes en función de características similares, permitiendo la identificación de perfiles que requieren intervenciones específicas.

Estas herramientas permiten tomar decisiones basadas en evidencia, optimizando la asignación de recursos y el diseño de políticas que respondan a las necesidades reales del sistema educativo.

En este sentido, estudios como el de Carlos Gómez, investigador de la Universidad Nacional de Colombia, acerca de la clasificación de colegios según las pruebas Saber 11 del ICFES: un análisis usando modelos marginales, propone:

Utilizar el método de estimación a través de las ecuaciones de estimación generalizadas (EEG), propuestas por Liang y Zeger (1986) y que son una extensión multivariada de los enfoques de cuasi verosimilitud, para clasificar a los colegios según las pruebas SER, dado que incluyen la posibilidad de asumir estructuras de dependencia entre las diferentes observaciones de una unidad muestral (Gómez S., 2016, p. 80).

En Ecuador, no se han aplicado ampliamente este tipo de metodologías para determinar la focalización en el ámbito educativo. Sin embargo, las técnicas multivariadas sí han sido utilizadas en el sector social. Uno de los casos más conocidos es el actual Registro Social, definido como el "catastro en el que consta información social, económica y demográfica individualizada y a nivel de familias; que permite a las instituciones identificar a sus poblaciones objetivo, para enfocar mejor los esfuerzos hacia los grupos en condiciones de pobreza"(Naranjo Bonilla, 2008, p. 5).

Este catastro tiene sus inicios en el Bono Solidario implementado en 1998, posteriormente denominado Programa de Bono de Desarrollo Humano, que consistía en la entrega de una transferencia monetaria mensual de 30 dólares. Este subsidio familiar se otorgaba a madres, ancianos y personas con discapacidad, quienes eran identificados mediante una encuesta

socioeconómica aplicada a cada hogar. Con este insumo, se establecía un índice de bienestar que servía como base para la selección de los beneficiarios.

En este sentido, el uso de métodos estadísticos avanzados para la clasificación y focalización de recursos en el ámbito social ha cobrado relevancia en los últimos años. Un ejemplo de ello es el estudio de Morales, Flores y Salazar Méndez (2022), quienes aplicaron el Análisis de Componentes Principales No Lineal (ACP No Lineal) para la construcción de un índice de estratificación socioeconómica en Ecuador. Este enfoque metodológico permitió transformar variables categóricas en una escala continua mediante un proceso de cuantificación óptima, maximizando la varianza explicada y mejorando la precisión en la identificación de grupos socioeconómicos. Según los autores, este estudio representa una aportación novedosa, ya que hasta el momento no se han encontrado investigaciones previas en Ecuador que utilicen esta metodología.

El uso de técnicas estadísticas avanzadas en la investigación social ha demostrado ser una herramienta clave para la toma de decisiones basada en datos. Vélez-Tamay, Carriel y Castillo-Ortega (2021) aplicaron el Análisis de Componentes Principales (ACP) para la construcción de un Índice de Desarrollo Local en Ecuador, logrando reducir la dimensionalidad de un conjunto de variables económicas, sociales y ambientales. Esta metodología permitió identificar patrones de desarrollo en distintos cantones del país, demostrando su utilidad en la síntesis de información compleja y en la focalización de recursos. La aplicación de esta técnica en otros ámbitos, como el educativo, permitiría generar índices de clasificación de instituciones o identificar factores determinantes en el acceso y calidad de la educación. De este modo, el ACP se consolida como una herramienta estadística robusta con el potencial de optimizar la planificación y evaluación de políticas educativas.

La implementación de este tipo de técnicas en la clasificación de unidades educativas podría contribuir significativamente a una mejor asignación de recursos y a la focalización de intervenciones en el sector educativo.

2.2 Marco teórico

2.2.1 Inversión en Educación y equidad en la distribución de recursos.

La educación es un derecho fundamental que permite a las personas de una sociedad salir del círculo vicioso de la pobreza, proporciona herramientas básicas para comprender y participar de manera activa en la sociedad. Sin embargo, este derecho fundamental se ve

muchas veces limitado por decisiones gubernamentales que afectan especialmente a quienes asisten a instituciones de educación públicas.

Según información publicada por el Banco Mundial, en el Ecuador el gasto público en educación como porcentaje del PIB ha pasado de 4,62% en el año 2018 al 3,89% en el año 2023, evidenciando un decrecimiento sostenido en la asignación de recursos, esta situación se agrava cuando no existe una eficiente distribución de los recursos asignados, “la localización de los recursos económicos determina no sólo la infraestructura de las instituciones y la calidad de las condiciones laborales de los profesores, además, condicionan un número de beneficios adicionales para los estudiantes” (Nuevas Ideas Ecuador, 2019). Estos beneficios forman parte de una educación integral.

El Observatorio de Financiación de la Educación, que es una iniciativa del Banco Mundial juntamente con el Instituto de Estadística de la UNESCO menciona “En todos los países, especialmente en los de ingreso relativamente más bajo, la cuantía, la eficiencia y la equidad del gasto en educación son a menudo insuficientes para alcanzar los objetivos de aprendizaje” (Banco Mundial y UNESCO, 2023, p. 4).

En esta misma línea, es crucial no solo asignar los recursos necesarios, sino también garantizar que lleguen de manera efectiva a quienes más los necesitan. “Lo que se gasta no siempre se gasta bien. Lograr mejores resultados del aprendizaje requiere un uso eficiente de los recursos adecuados” *op.cit.*

Respaldando esta perspectiva, Amada Lalama, docente investigadora de la Universidad de Guayaquil, sostiene que “no necesariamente entregar más recursos a educación, significa mejor nivel educativo” (Lalama et al., 2022, p. 63). Este planteamiento se fundamenta en un análisis comparativo de países de Latinoamérica, donde se observó que naciones como Chile y Uruguay, con una calidad educativa destacada en la región, destinaron menores recursos que Costa Rica y Brasil, cuyos resultados en el informe PISA 2018 fueron más modestos.

En el documento del VI Congreso Internacional de Desarrollo Económico, Social y Empresarial de Latinoamérica, en el acápite denominado Inversión en Educación y Calidad Educativa, un análisis de los resultados en el marco de los ODS. Caso Ecuador, se menciona que

El Estado debe velar por este derecho y proveer un sistema educativo integral y de calidad que incluya procesos, infraestructura, tecnología, materiales, equipos, malla curricular y lo más importante docentes con altas competencias para formar a los estudiantes en los diferentes niveles educativos, de tal manera que se transformen en agentes activos que aporten a la sociedad. (Lalama et al., 2022, p. 54)

Garantizar un acceso equitativo a una educación de calidad no solo empodera a las personas al dotarlas de habilidades y conocimientos esenciales, sino que también impulsa el crecimiento económico y la reducción de desigualdades.

La inversión en infraestructura educativa, capacitación docente y recursos pedagógicos modernos no debe verse como un gasto, sino como una inversión estratégica que asegura un futuro más próspero, innovador y justo para las generaciones presentes y futuras.

Organismos multilaterales como el Banco Mundial (BM), el Banco Interamericano de Desarrollo (BID), la Corporación Andina de Fomento (CAF) y la Agencia Francesa de Desarrollo (AFD) han desempeñado un papel clave en el financiamiento de proyectos de inversión en educación en Ecuador.

A través de líneas de crédito y recursos no reembolsables, estas instituciones han apoyado iniciativas orientadas a mejorar la infraestructura educativa, fortalecer la calidad del aprendizaje y reducir las brechas de acceso a la educación.

Estos financiamientos permiten la implementación de programas de modernización del sistema educativo, construcción y rehabilitación de escuelas, capacitación docente y el desarrollo de metodologías innovadoras para la gestión educativa.

Isabel Gil, docente de la Universidad de Cuenca en un estudio realizado sobre los créditos otorgados por el Banco Mundial para educación menciona que “el monto entregado en calidad de préstamo para proyectos educativos al Ecuador entre los años 1967 y 2017 fue de 568,5 millones de dólares” (Gil, 2023, pp. 17–31). La magnitud de estos financiamientos refleja el compromiso de los organismos internacionales con el desarrollo educativo en Ecuador, pero también conlleva una gran responsabilidad en su gestión. Es fundamental que estos recursos sean invertidos de manera eficiente y estratégica, asegurando que lleguen a las instituciones y comunidades que más los necesitan, optimizando su impacto en la calidad educativa y en la reducción de las brechas de acceso al aprendizaje

2.2.2 Factores asociados a las instituciones educativas fiscales que inciden en su clasificación.

La construcción de un índice para clasificar las instituciones educativas fiscales requiere un análisis profundo de los factores determinantes que influyen en su desempeño y categorización. Estas instituciones educativas desempeñan un papel central en la formación de la población, especialmente en contextos de vulnerabilidad, donde representan una de las principales herramientas para reducir brechas socioeconómicas.

Su clasificación no solo refleja indicadores de rendimiento académico, sino también elementos clave como la infraestructura, la disponibilidad de recursos, la gestión

institucional y las características del entorno geográfico y socioeconómico en el que operan. Este marco teórico tiene como propósito identificar y articular las variables que inciden en esta clasificación, proporcionando un sustento conceptual para el diseño de un índice que clasifique a las instituciones educativas fiscales, permitiendo identificar las que se desempeñan en condiciones menos favorables.

- **Ubicación Geográfica**

Uno de los aspectos recurrentes en el análisis es el impacto de la ubicación geográfica, ya sea urbana o rural, sobre el desempeño y las condiciones de las instituciones educativas. Al respecto, la UNESCO señala

“Las escuelas se definen y distinguen sobre la base de numerosos atributos. Sin embargo, existen diferencias fundamentales entre las escuelas de pequeños poblados o zonas rurales y aquellas ubicadas en las ciudades, y entre escuelas públicas y privadas. En términos generales, se podría esperar que las escuelas en ciudades y privadas aventajaran a las escuelas en pequeños poblados y públicas en términos de condiciones de funcionamiento” (UNESCO, 2008a, p. 5).

Complementariamente, el informe *Inversión en educación y calidad educativa: un análisis de los resultados ODS. Caso Ecuador* confirma que las desventajas en las zonas rurales profundizan las desigualdades y perpetúan un círculo vicioso de carencias. “En estas áreas, se concentran la mayor cantidad de personas en pobreza y extrema pobreza, con acceso, en el mejor de los casos, a educación básica” (Lalama et al., 2022, p. 54).

- **Formas de acceso a la institución educativa.**

Por las condiciones geográficas y de infraestructura del país existen instituciones educativas a las cuales se accede solamente de forma aérea o fluvial, lo que implica medios de transporte diferentes, especialmente en las zonas más recónditas de país, al respecto UNESCO afirma que “dentro del contexto de las escuelas en pequeños poblados o zonas rurales, se debe tener en cuenta la disponibilidad de recursos comunitarios y la existencia de medios confiables de transporte hacia la escuela” (UNESCO, 2008, p. 6) . La conectividad del transporte, o la carencia de esta, desempeña un papel crucial en el desarrollo de cualquier país, ya que facilita el acceso a servicios esenciales, como la educación. En este contexto, el Foro Económico Mundial, en su *Informe sobre Competitividad Global 2019*, ubicó a Ecuador en el puesto 34 de 141 países en términos de calidad vial, destacándose por encima de varios países de América Latina, como Colombia, que ocupó el puesto 104.

Este avance representa un logro significativo para Ecuador, pues la mejora en la infraestructura vial no solo optimiza el transporte terrestre, sino que también amplía las

oportunidades para acceder a servicios básicos en zonas previamente marginadas, contribuyendo así al desarrollo social y económico del país.

- **Acceso a servicios básicos agua, lavamanos e inodoros en las instituciones educativas**

El acceso a servicios básicos como agua potable, lavamanos e inodoros en las escuelas es esencial para garantizar un entorno seguro, saludable y propicio para el aprendizaje. Estos servicios no solo promueven la higiene y la prevención de enfermedades, sino que impactan directamente en la asistencia, el rendimiento académico y el bienestar de los estudiantes.

La falta de infraestructura adecuada puede generar barreras significativas para la educación, especialmente en comunidades vulnerables, donde la ausencia de condiciones mínimas de salubridad afecta de manera significativa a niñas y niños.

En estudio realizado por el BID, Agua y saneamiento en escuelas menciona “La evidencia indica que el acceso universal a agua potable, saneamiento e higiene, con la consecuente prevención de enfermedades, podría generar un incremento de hasta 1900 millones de días de asistencia a la escuela a nivel mundial” (Piderit et al., 2024, p. 4).

En lo relacionado a los lavamanos, este mismo documento menciona la importancia de contar con al menos un lavamanos por cada 40 estudiantes y que los lavamanos “deberán tener en cuenta factores que garanticen el uso a todos los estudiantes, con incidencia en la accesibilidad, la inclusión y la seguridad; factores ergonómicos que permitan y faciliten el uso; y factores que repercutan en la calidad y eficiencia” *op.cit.*

En lo referente a los inodoros UNICEF da una referencia general de al menos 1 sanitario y 1 urinario para 50 niños y 1 baño para 25 niñas. Pero estas proporciones también dependerán de la edad de los estudiantes.

- **Acceso a Internet**

El internet y las tecnologías de la información y la comunicación (TIC) se han convertido en pilares fundamentales para transformar la educación. Su capacidad para ofrecer acceso ilimitado a información, herramientas interactivas y plataformas de aprendizaje personalizado permite enriquecer la enseñanza y preparar a los estudiantes para los retos de un mundo digital. En este contexto, es fundamental reconocer “el acceso a tecnología e Internet como un derecho de todos los estudiantes, asumiendo los Estados el deber de asegurar el acceso a quienes no pueden hacerlo por sí mismos” (UNESCO, 2014, p. 53).

Estas tecnologías fomentan la inclusión, conectan comunidades educativas y desarrollan habilidades tecnológicas esenciales, promoviendo una educación más dinámica, colaborativa e innovadora.

Sin embargo, el acceso a las TIC aún es limitado.

La brecha digital en sus diferentes dimensiones es un desafío aún no superado en América Latina y el Caribe. Esta brecha se hace visible entre poblaciones urbanas y rurales, a nivel socioeconómico, en la dimensión de género, y se relaciona estrechamente con desigualdades de capital sociocultural en general (Laura y Tessore, 2021, p. 15).

- **Identificación étnica de estudiantes y docentes.**

En la Declaración Mundial sobre Educación para Todos de 1990, adoptada en Jomtien Tailandia, se pidió a los países que se comprometieran activamente a “modificar las desigualdades en materia de educación”, enfatizando que los grupos indígenas, las minorías étnicas, raciales y lingüísticas, no deben sufrir ninguna discriminación en el acceso a las posibilidades de aprendizaje complementariamente la UNESCO brinda directrices sobre la educación intercultural bilingüe y establece principios dentro de los cuales tenemos:

La eliminación de todas las formas de discriminación en el sistema educativo posibilitando el acceso en pie de igualdad a todas las formas de educación para todos los grupos culturales de la población

Una formación docente inicial adecuada y una formación profesional permanente que brinde a los profesores:

una profunda comprensión del paradigma intercultural en la educación y su importancia para la transformación de la práctica cotidiana en las aulas, las escuelas y las comunidades; una conciencia crítica del papel que la educación debe desempeñar en la lucha contra el racismo y la discriminación (UNESCO, 2006).

- **Porcentaje de Docentes con título de cuarto nivel**

La formación y especialización de los docentes a través de estudios formales como maestrías es fundamental para elevar la calidad educativa en escuelas y colegios. Este tipo de preparación no solo les permite adquirir conocimientos avanzados y actualizados en su área de enseñanza, sino también desarrollar habilidades pedagógicas más efectivas, lo que repercute directamente en el aprendizaje de sus alumnos.

Los maestros mejor preparados pueden implementar metodologías innovadoras, fomentar un pensamiento crítico en sus estudiantes y adaptarse a las necesidades cambiantes de la educación moderna. “La formación de los docentes es un aspecto fundamental en la perpetuación de las desigualdades, pues son justamente los estudiantes indígenas y más pobres quienes carecen de docentes capacitados para el ejercicio de su profesión” (Larrea Maldonado, 2024, p. 24).

En escuelas con mayores carencias de recursos, servicios e instalaciones, la tarea del docente dentro del aula es primordial para transmitir conocimientos básicos y para proveer a los niños condiciones adecuadas para aprender, esto es brindarles estímulo, confianza, respeto, contenerlos psicológicamente, enseñarles a escuchar y ser escuchados, retenerlos en la escuela (Falus y Goldberg, 2011, p. 46).

En el Informe mundial sobre el personal docente, publicado por la UNESCO, se menciona Los docentes desempeñan un papel esencial en el ejercicio del derecho a la educación y son fundamentales para alcanzar todos los ODS, especialmente el ODS 4. Para lograr una educación de calidad para todos, todos los niños y niñas deben tener acceso a un docente cualificado (UNESCO, 2023, p. 3).

- Número de estudiantes por docente

El ratio de estudiantes por docente se presenta como un factor clave en la percepción de autoeficacia del profesorado, particularmente en los contextos de educación secundaria. “Según investigaciones previas, un mayor número de estudiantes por docente está asociado con una disminución en los niveles de autoeficacia, lo que podría impactar negativamente en la calidad de la enseñanza y la gestión del aula” (Tang et al., 2020, citado en González-Mayorga y Rodríguez Esteban, 2023, p.72).

Además, estudios internacionales sugieren que los efectos de esta variable no son uniformes en todos los contextos; por ejemplo, Reeves et al 2017 citado en (González-Mayorga y Rodríguez Esteban, 2023) encontraron que, “mientras en Japón un tamaño elevado de clase tenía un efecto negativo en el desempeño docente, en Estados Unidos podía resultar en efectos positivos. Esto subraya la necesidad de considerar el ratio de estudiantes por docente como una variable contextual y culturalmente sensible que influye en el rendimiento y la eficacia de los educadores” (p.72).

- **Estudiantes con discapacidad**

Contar con un indicador que registre el número de estudiantes con discapacidad en una institución educativa es fundamental para construir un índice de clasificación escolar. Este tipo de indicador no solo permite visibilizar la diversidad en el entorno escolar, sino que también impulsa a las escuelas a desarrollar políticas y prácticas que garanticen la participación activa y equitativa de todos los estudiantes.

La inclusión es un proceso que ayuda a superar las barreras que limitan la presencia, la participación y los logros de los estudiantes y la equidad consiste en garantizar la imparcialidad, donde se considera que la educación de todos los estudiantes tiene la misma importancia (UNESCO, 2021, p. 3).

- **Abandono escolar**

Una variable que está siendo permanentemente analizada por su importancia es el abandono escolar y es uno de los principales desafíos que enfrenta Latinoamérica y Ecuador en particular, según I CEPAL “el abandono escolar está estrechamente vinculado con el trabajo infantil y la informalidad laboral y que en 14 países de América Latina un 41,4% de la fuerza de trabajo de 15 años y más no completó la enseñanza secundaria” (CEPAL, 2024, p. 2).

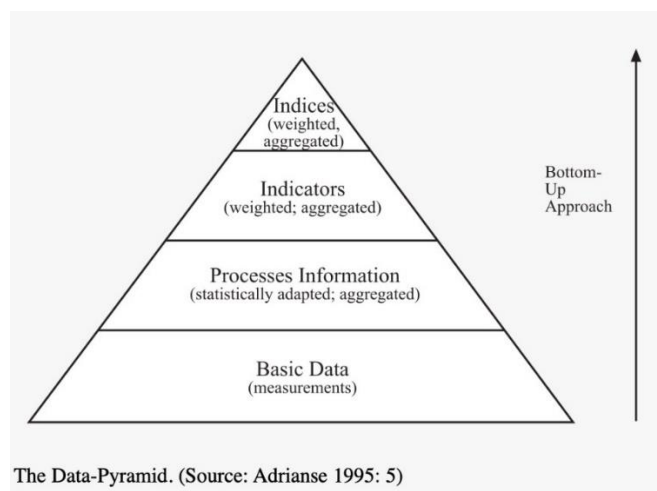
2.2.3 *Índices e Indicadores*

"La reflexión sobre los índices nació alrededor de tres siglos atrás en Inglaterra, con la necesidad de medir con un solo número las variaciones de un conjunto de precios entre dos fechas determinadas" (Moscolini Nora, 2011, p. 25). Los índices posibilitan simplificar la información y representar de manera más concisa y comprensible las relaciones entre las variables. "Los índices para medir la privación son herramientas ampliamente utilizadas en el análisis de las desigualdades" (Duque et al., 2021, p. 2).

Según el enfoque de la pirámide de datos de Adrianse (1995), se centra en la organización de la información desde los datos básicos hasta los índices (Figura 1). Los índices se encuentran en la cúspide de la pirámide y representan el nivel más alto de síntesis y agregación de información. Estos índices son ponderados y agregados, lo que implica que son una combinación de diferentes indicadores que han sido evaluados en función de su importancia y relevancia (Linser y Vienna, 1999).

Figura 1

Pirámide de datos



El análisis multivariado es una técnica estadística que permite el estudio simultáneo de múltiples variables a través de la aplicación de modelos matemáticos y estadísticos, “se

refiere a todos los métodos estadísticos que analizan simultáneamente medidas múltiples de cada objeto sometido a investigación. Cualquier análisis simultáneo de más de dos variables puede ser considerado aproximadamente como un análisis multivariado.”

Existe una multitud de técnicas de análisis multivariante, abarcando desde los métodos puramente descriptivos hasta los que utilizan técnicas de inferencia estadística para obtener conclusiones de los datos y poder elaborar modelos que no son evidentes a simple vista observando los datos obtenidos. También nos permitirán desarrollar modelos de predicción de varias variables y establecer relaciones entre las mismas” (Molina, 2019, p. 2).

Las de técnicas multivariadas se clasifican en dos:

a) Técnicas de Análisis de dependencia

Uriel menciona que “las técnicas de análisis de dependencia buscaran la existencia o ausencia de relaciones entre dos grupos de variables, basándose en un experimento controlado o en una base teórica previa el investigador clasifica dos grupos de variables: dependientes o independientes” (Joaquín y Uriel, 2017)

Las principales técnicas de dependencia son:

1. Ecuaciones estructurales
2. Regresión Lineal Múltiple
3. Análisis discriminante
4. Regresión VDL
5. MANOVA
6. Correlación Canónica

b) Técnicas de Análisis de Interdependencia

“Existen situaciones de investigación en las que es imposible separar las variables en dependientes e independientes, y lo que interesa determinar es cómo y por qué las variables están correlacionadas entre ellas (Joaquín y Uriel, 2017).

Las principales técnicas de Interdependencia son:

1. Análisis de Componentes Principales
2. Análisis Factorial
3. Análisis de Correspondencias
4. Análisis de Conglomerados
5. Escalamiento Multidimensional

Las tres primeras técnicas se utilizan para identificar las relaciones entre las variables, el Análisis de Conglomerados es útil para observar el relacionamiento entre casos y el escalamiento multidimensional se centra en la relación entre objetos.

Para el caso de la construcción del índice de clasificación de las instituciones educativas las técnicas de interdependencia son las adecuadas ya que se busca mirar las relaciones entre las variables de las instituciones educativas y a partir de estas construir el índice. Al contar con variables de tipo categórico son específicamente el Análisis de Correspondencia y el de Componentes Principales No lineales las técnicas idóneas a ser aplicadas, descartando el Análisis Factorial que está diseñado para el tratamiento de variables cuantitativas continuas. Por lo tanto, la utilización del ACM al igual que el Análisis de Componentes Principales No Lineales son herramientas efectivas y plenamente aplicables en la clasificación de instituciones educativas permitiendo la construcción de un índice que capture las múltiples facetas de estas entidades, ofreciendo una herramienta robusta para comprender y comparar de manera integral las características de cada institución.

2.2.4 Análisis Multivariado

El análisis multivariado es una técnica estadística que permite el estudio simultáneo de múltiples variables a través de la aplicación de modelos matemáticos y estadísticos, “se refiere a todos los métodos estadísticos que analizan simultáneamente medidas múltiples de cada objeto sometido a investigación. Cualquier análisis simultáneo de más de dos variables puede ser considerado aproximadamente como un análisis multivariado.”

Existe una multitud de técnicas de análisis multivariante, abarcando desde los métodos puramente descriptivos hasta los que utilizan técnicas de inferencia estadística para obtener conclusiones de los datos y poder elaborar modelos que no son evidentes a simple vista observando los datos obtenidos. También nos permitirán desarrollar modelos de predicción de varias variables y establecer relaciones entre las mismas” (Molina, 2019, p. 2).

Las de técnicas multivariadas se clasifican en dos:

c) Técnicas de Análisis de dependencia

Uriel menciona que “las técnicas de análisis de dependencia busquen la existencia o ausencia de relaciones entre dos grupos de variables, basándose en un experimento controlado o en una base teórica previa el investigador clasifica dos grupos de variables: dependientes o independientes” (Joaquín y Uriel, 2017)

Las principales técnicas de dependencia son:

7. Ecuaciones estructurales

8. Regresión Lineal Múltiple
9. Análisis discriminante
10. Regresión VDL
11. MANOVA
12. Correlación Canónica

d) Técnicas de Análisis de Interdependencia

“Existen situaciones de investigación en las que es imposible separar las variables en dependientes e independientes, y lo que interesa determinar es cómo y por qué las variables están correlacionadas entre ellas (Joaquín y Uriel, 2017).

Las principales técnicas de Interdependencia son:

6. Análisis de Componentes Principales
7. Análisis Factorial
8. Análisis de Correspondencias
9. Análisis de Conglomerados
10. Escalamiento Multidimensional

Las tres primeras técnicas se utilizan para identificar las relaciones entre las variables, el Análisis de Conglomerados es útil para observar el relacionamiento entre casos y el escalamiento multidimensional se centra en la relación entre objetos.

Para el caso de la construcción del índice de clasificación de las instituciones educativas las técnicas de interdependencia son las adecuadas ya que se busca mirar las relaciones entre las variables de las instituciones educativas y a partir de estas construir el índice. Al contar con variables de tipo categórico son específicamente el Análisis de Correspondencia y el de Componentes Principales No lineales las técnicas idóneas a ser aplicadas, descartando el Análisis Factorial que está diseñado para el tratamiento de variables cuantitativas continuas. Por lo tanto, la utilización del ACM al igual que el Análisis de Componentes Principales No Lineales son herramientas efectivas y plenamente aplicables en la clasificación de instituciones educativas permitiendo la construcción de un índice que capture las múltiples facetas de estas entidades, ofreciendo una herramienta robusta para comprender y comparar de manera integral las características de cada institución.

2.2.5 Uso de técnicas multivariadas en la construcción de índices

Savitri Abeyasekera, en el documento Métodos Multivariantes para la construcción de un índice, menciona:

Los datos determinan la forma de un índice mediante un procedimiento multivariante y que la interpretación de un índice es de un valor único que captura la información de varias variables en una medida compuesta, que suele adoptar la forma:

$$a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_pX_p$$

donde las a_1 son ponderaciones que deben determinarse a partir de los datos y las X_1 son un subconjunto apropiado de p variables medidas (Abeyasekera, n.d., p. 14).

Las técnicas multivariadas son frecuentemente utilizadas para la construcción de índices, En el artículo denominado "Socioeconomic Adversity and Higher Education: Is the Quota Law a Potential Bridge to Social Mobility in Brazil?", escrito por Fernanda Maria de Almeida, Kristinn Hermannsson y Antônio Sérgio de Araújo Fernandes, analiza el impacto de la Ley de Cuotas en Brasil en la movilidad social de estudiantes con alta adversidad socioeconómica, los autores mencionan que "El Índice de Adversidad del estudio se construyó utilizando la técnica de Análisis Factorial Confirmatorio. Basado en supuestos teóricos, esta es una técnica multivariada que permite cuantificar variables latentes no observables determinadas por un conjunto más amplio de variables observables" (Nunes et al., 2015, p. 8).

En el artículo "*A Comparison of Social Vulnerability Indices Specific to Flooding in Ecuador: Principal Component Analysis (PCA) and Expert Knowledge*", escrito por Agathe Bucherie, Carolynne Hultquist, Susana Adamo, Colleen Neely, Fernanda Ayala, Juan Bazo y Andrew Kruczkiewicz, se explora la construcción de índices de vulnerabilidad social frente a inundaciones en Ecuador, comparando dos métodos: el análisis de componentes principales (PCA) y el juicio experto.

Los autores evalúan cómo estas metodologías generan diferencias significativas en la evaluación de la vulnerabilidad, particularmente en áreas con señales de vulnerabilidad menos definidas. El PCA se utiliza para transformar variables correlacionadas en dimensiones separadas, siendo la primera componente la que explica la mayor varianza, mientras que el método de juicio experto prioriza variables basadas en el conocimiento práctico de los desastres.

Según los autores utilizarán esta técnica ya que "El PCA se usa comúnmente para reducir las dimensiones de los datos al transformar" las variables más correlacionadas en dimensiones separadas, siendo la primera componente la que explica la mayor varianza" (Bucherie et al., 2022, p. 7).

Ambos métodos aportan perspectivas complementarias y subrayan la importancia de incluir datos históricos de impacto para validar los índices en la toma de decisiones.

El proyecto MEDEA, financiado por el Fondo Europeo de Desarrollo Regional, construyó un índice de privación a partir de datos censales en grandes ciudades españolas, con el objetivo de identificar secciones censales con situaciones socio económicas más desfavorables, para lo cual utilizó el Análisis de Componentes Principales, concluyendo que “El índice de privación puede ser un instrumento útil para la planificación sanitaria al detectar áreas pequeñas de grandes ciudades con una situación socioeconómica desfavorable, puede contribuir al estudio de las desigualdades sociales en salud en España”(Duque et al., 2021, p. 179).

El Análisis de Correspondencias Múltiples (ACM) se presenta como una alternativa igualmente poderosa al Análisis de Componentes Principales Lineales y No Lineales para la construcción de un índice de clasificación de instituciones educativas.

El ACM, es una técnica de reducción de dimensionalidad que permite sintetizar información compleja y multivariada, su principal fortaleza radica en su capacidad para trabajar eficazmente con variables categóricas, lo que es particularmente útil cuando se manejan datos cualitativos, como pueden ser aspectos relacionados con el ambiente institucional, la calidad docente o las percepciones de la comunidad educativa.

“El ACM está diseñado para analizar tablas disyuntivas completas, que son tablas de contingencia de variables cualitativas. Se debe considerar que, las modalidades de cada variable son mutuamente excluyentes, y cada individuo pertenece a una y solo una de ellas” (Abril Donoso et al., 2022, p. 47).

Estudios como el de Abdeljaouad y Verme (2013) en la medición de la pobreza multidimensional en Marruecos y el trabajo de Lamfre Laura y Perren Joaquín (2023) sobre la calidad de vida en Argentina, demuestran cómo el ACM puede identificar patrones de correlación entre variables categóricas y resumir la información en componentes principales que reflejen fielmente las dimensiones analizadas.

De igual forma, en el documento realizado por Adrina Chazarreta, publicado en la revista Saberes, se propone construir un índice a partir del Análisis de Correspondencias Múltiples (ACM) que permita clasificar a las empresas según el grado de separación de la propiedad y la dirección de estas y comparar el alcance en empresas con diferentes características.

“La aplicación de esta técnica permitió cuantificar y complejizar la medición de la separación de la propiedad y el control. En lugar de pensarla como un atributo dicotómico la construcción de este índice permite definirla como una variable cuantitativa, pudiendo de esta forma identificar diferentes gradaciones en su interior” (Chazarreta, 2022, p. 195).

Sin embargo, existen diferencias entre el ACM y el ACPNL, según Will Atkinson “La diferencia es que el ACPNL maneja todos y cada uno de los tipos de datos simultáneamente (datos continuos y categóricos) y, lo que es más importante, se puede especificar de antemano si las variables categóricas siguen una estructura nominal ordinal” (Atkinson, 2024, p. 833).

2.3 Marco legal

La Constitución ecuatoriana consagra la educación como un derecho fundamental, estableciendo al Estado como el principal garante de la igualdad de oportunidades y del acceso a una educación de calidad. En este marco se reconoce la educación como un derecho universal y gratuito, sustentado en los principios de inclusión y equidad, lo que obliga al Estado a implementar políticas y asignar los recursos necesarios para garantizar el acceso de todos los ciudadanos a una educación de calidad.

El artículo 3 de la Constitución de la República dispone que son deberes primordiales del Estado garantizar, sin discriminación alguna, el efectivo goce de los derechos consagrados en la Constitución y en los instrumentos internacionales, en particular el derecho a la educación, la salud, la alimentación, la seguridad social y el acceso al agua.

De igual forma el artículo 26 señala que la educación es un derecho de las personas a lo largo de su vida y un deber ineludible e inexcusable del Estado. Constituye un área prioritaria de la política pública y de la inversión estatal, garantía de la igualdad e inclusión social y condición indispensable para el buen vivir. Las personas, las familias y la sociedad tienen el derecho y la responsabilidad de participar en el proceso educativo

Por su parte, el artículo 39 establece que el Estado garantizará los derechos de las y los jóvenes, promoviendo su efectivo ejercicio mediante políticas, programas, instituciones y recursos que aseguren y mantengan de forma permanente su participación e inclusión en todos los ámbitos, especialmente en los espacios de poder público. Asimismo, se reconoce a la juventud como actor estratégico en el desarrollo del país, comprometiéndose a garantizarles el acceso a la educación, la salud, la vivienda, la recreación, el deporte, el tiempo libre, la libertad de expresión y de asociación. El Estado fomentará, además, su incorporación al mundo laboral en condiciones justas y dignas, haciendo especial énfasis en la capacitación, el acceso al primer empleo y la promoción de sus habilidades de emprendimiento.

Finalmente, en la disposición transitoria decimoctava, se establece que “El Estado asignará de forma progresiva recursos públicos del Presupuesto General del Estado para la educación

inicial, básica y el bachillerato, con incrementos anuales de al menos el 0,5% del Producto Interno Bruto hasta alcanzar un mínimo del 6% del Producto Interno Bruto.”

La Ley Orgánica de Educación Intercultural LOEI, regula el sistema educativo en Ecuador y define el marco normativo para la organización, administración y supervisión de las instituciones educativas.

En lo relacionado a las instituciones públicas el artículo 54 determina “Los establecimientos educativos públicos son fiscales, municipales o comunitarias. La educación impartida por estas instituciones es laica y gratuita, sin costo para sus beneficiarios. La comunidad tiene derecho a la utilización responsable de las instalaciones y servicios de los establecimientos educativos públicos para actividades culturales, artísticas, deportivas, de recreación y esparcimiento, que promuevan el desarrollo comunitario, de conformidad con la Ley y demás normativa aplicable. La regulación del acceso, organización y funcionamiento de estas instituciones corresponderá a la Autoridad Educativa Nacional. Estas instituciones no servirán a intereses, ideologías o activismo de índole político.

La LOEI establece criterios que orientan la gestión, evaluación y mejora de la calidad educativa, resaltando la importancia de atender las necesidades particulares de cada institución para garantizar la igualdad de oportunidades.

El artículo 2 menciona los fundamentos que deben regir el sistema, entre ellos la calidad, la inclusión, la interculturalidad y la equidad. Se señala que la educación debe ser de calidad y garantizar igualdad de oportunidades para todos los sectores de la sociedad.

El literal c. relacionado a la Igualdad de oportunidades y de trato, hace énfasis en que “se garantizan entornos de aprendizaje accesibles y asequibles material y económicamente a todas las niñas, niños y adolescentes, respetando sus diversas necesidades, capacidades y características, eliminando todas las formas de discriminación. Se establecerán medidas de acción afirmativa para efectivizar el ejercicio del derecho a la educación

De igual forma el artículo 5 dispone que el Estado tiene la obligación ineludible e inexcusable de garantizar el derecho a la educación de todos los habitantes del territorio ecuatoriano y de los ecuatorianos en el exterior y el acceso universal a lo largo de su vida, para lo cual generará las condiciones que garanticen la igualdad de oportunidades para acceder, permanecer, movilizarse y culminar los servicios educativos.

El Estado ejerce la rectoría sobre el Sistema Educativo a través de la Autoridad Educativa Nacional de conformidad con la Constitución de la República y la Ley, y garantizará una educación pública de calidad, gratuita y laica. La educación en todas sus formas y en todos

los niveles debe tener cuatro características interrelacionadas: disponibilidad, aceptabilidad, asequibilidad y accesibilidad.

En esta misma línea, el literal c, de artículo 6 establece como obligación del Estado “Asegurar que los establecimientos educativos sean espacios democráticos de ejercicio de derechos y convivencia pacífica”

La normativa promueve procedimientos transparentes y participativos en la toma de decisiones, lo que respalda la utilización de métodos analíticos y de clasificación para fundamentar la priorización de las instituciones que serán beneficiadas por proyectos de inversión.

El artículo 20 de la ley, establece que “Dentro del presupuesto del sector educación se priorizará obligatoriamente a los segmentos sociales que están en situación de vulnerabilidad o riesgo, que requieren atención prioritaria, para compensar las desigualdades derivadas de factores socio-económicos, culturales, emergencias o desastres, geográficos o de cualquier otra índole.”

El artículo 22 literal p. dispone “Fortalecer el funcionamiento de los establecimientos educativos públicos, priorizando el sector rural, zonas de frontera, de riesgo, y aquellas afectadas por desastres naturales y antrópicos”.

En este mismo artículo el literal x. establece “Garantizar la transferencia de recursos de manera oportuna, regular y suficiente a los niveles y modalidades del Sistema Nacional de Educación”

Por otra parte, el Estatuto Orgánico de Gestión Organizacional del Ministerio de Educación expedido mediante Acuerdo Ministerial No 020 – 12, establece como misión de la Coordinación General de Planificación “Asesorar, impulsar, coordinar y articular con todos los niveles de gestión los procesos integrales de planificación estratégica, operativa y de organización de la oferta educativa sobre la base de un análisis prospectivo ajustado a los objetivos nacionales determinados en la Constitución, Ley Orgánica de Educación Intercultural, Plan Nacional del Buen Vivir y el Plan Nacional de Educación.”.

Referido instrumento, establece entre otras, las siguientes atribuciones y responsabilidades a cargo de la Coordinación General de Planificación: “Ajustar y poner en consideración de la Autoridad Educativa Nacional las propuestas de política pública para el Plan Nacional de Educación y demás planes nacionales de desarrollo, políticas para la reorganización de la oferta educativa nacional y optimización de los recursos; y, Proponer los lineamientos estratégicos de corto, mediano y largo plazo que orientarán la ejecución de la política educativa, mismos que deberán ser aprobados por la Autoridad Educativa Nacional.”

CAPÍTULO III

METODOLOGÍA

3.1 Descripción del área de estudio/grupo de estudio

Según la División Política Administrativa DPA del año 2024, publicado por el Instituto Ecuatoriano de Estadística y Censos INEC, el Ecuador se encuentra dividido en 24 provincias, 222 cantones y 1.251 parroquias urbanas y rurales, en todas ellas existe por lo menos una unidad educativa. Según el Censo Nacional de Población y Vivienda realizado en el año 2022, en el Ecuador contaba en ese año con 16.938.986 habitantes, de ellos 3.954.792 se encontraban en la edad de 5 a 17 años, que es la edad teórica para asistir a una institución educativa.

En tal sentido, la investigación se realizó considerando de inicio a 9.518 instituciones educativas que tenían información en todas las variables que son parte de este estudio de 12.189 instituciones educativas de sostenimiento público, de tipo ordinario, a nivel nacional, estas son instituciones que ofertan los servicios educativos a personas que no requieren tratamiento especializado y se encuentran en la edad teórica para su nivel de estudios, es decir, no se pueden matricular personas adultas conforme se encuentra registradas en la base. Estas instituciones se distribuyen a nivel de provincia como se muestra en la Tabla 1.

Tabla 1

Distribución de instituciones educativas por provincias

Provincia	No. IES
Azuay	576
Bolívar	337
Cañar	259
Carchi	116
Chimborazo	321
Cotopaxi	379
El Oro	450
Esmeraldas	330
Galápagos	14
Guayas	1.291
Imbabura	200
Loja	821
Los Ríos	569
Manabí	1.255
Morona Santiago	276
Napo	147
Orellana	214
Pastaza	102

Pichincha	727
Santa Elena	163
Santo Domingo De Los Tsáchilas	286
Sucumbíos	271
Tungurahua	174
Zamora Chinchipe	240
Nacional	9.518

De estas instituciones 4.948 se encuentran en el área rural, representando el 51,99% de las instituciones a nivel nacional.

En la figura 2 se muestra al Ecuador Continental y las Islas Galápagos, siendo esta el área de estudio, ya que las instituciones educativas se ubican en todo el país.

Figura 2

Área de estudio



3.2 Enfoque y tipo de investigación

3.2.1 Enfoque

“El enfoque cuantitativo es una forma de investigación en la cual se utilizan métodos numéricos y estadísticos para medir y analizar los datos. La idea central del enfoque cuantitativo es que se pueden medir las variables y establecer relaciones causales entre ellas” (Hadi et al., 2023, p. 52).

Técnicas como el Análisis de Correspondencias Múltiples y el ACP no lineal son herramientas estadísticas que responden de manera sensible a la variabilidad y estructura de

los datos. Al trabajar con información poblacional, estas metodologías permiten obtener resultados que reflejan de forma precisa las relaciones y patrones entre las variables, especialmente en aquellos casos en que dichas relaciones son no lineales o presentan una complejidad considerable. Esto conlleva a la generación de un índice de clasificación más fiable y robusto, capaz de incorporar la diversidad de la realidad educativa. En consecuencia, conforme a lo expuesto por Hadi et al. (2023), la presente investigación se enmarca en un enfoque cuantitativo.

3.2.2 Tipo de Investigación

“Los estudios correlacionales evalúan el grado de asociación entre dos o más variables, en los primeros se mide cada una de éstas, y después se cuantifican, analizan y establecen las vinculaciones” (Hernández et al.,2014, p93).

La presente investigación se fundamenta en un enfoque correlacional, ya como mencionan Hernández, Fernández y Baptista, se centra en cuantificar y analizar la relación existente entre las diversas variables que afectan a las instituciones educativas fiscales. Mediante el empleo de técnicas estadísticas multivariadas, como el Análisis de Correspondencias Múltiples y el Análisis de Componentes Principales, se busca medir y evaluar la fuerza y dirección de las asociaciones entre los factores relevantes. Este análisis permite identificar patrones y correlaciones significativas, lo cual es esencial para el diseño de un índice de clasificación que refleje de manera precisa las necesidades y fortalezas de cada institución en el contexto de la asignación de recursos educativos.

Paralelamente, el estudio adopta un enfoque explicativo al indagar en los mecanismos y procesos que subyacen a las relaciones observadas entre las variables. La comparación de las dos técnicas de clasificación multivariada no solo contribuye a la identificación de patrones, sino que también permite profundizar en la comprensión de por qué y cómo se establecen estas relaciones, ofreciendo una explicación integral de los factores que influyen en la priorización de las instituciones para recibir proyectos de inversión.

3.3 Definición y operacionalización de variables

Para la construcción del índice de clasificación de instituciones educativas, se realizó la operacionalización de variables categóricas con el propósito de evaluar la idoneidad del Análisis de Correspondencias Múltiples (ACM) y el Análisis de Componentes Principales No Lineal (ACP no lineal) en la reducción de dimensionalidad y estructuración del índice. Se seleccionaron variables cualitativas representativas de distintas dimensiones de vulnerabilidad, las cuales fueron transformadas en una matriz de indicadores disyuntivos completos para su tratamiento estadístico. A partir de ello, se aplicaron ambas técnicas con

el objetivo de comparar su capacidad para capturar la estructura latente de los datos y su estabilidad en la generación de los ejes factoriales. El ACM permitió analizar la asociación entre categorías mediante la representación de perfiles en un espacio común, mientras que el ACP no lineal evaluó la contribución de cada variable desde un enfoque basado en varianza óptima para datos categóricos.

A continuación, se muestra en la tabla 2 las variables, dimensiones e indicadores que son parte del estudio.

Tabla 2

Variables y dimensiones

Variable	Dimensión	Indicador
Estudiantes	Social	Porcentajes de estudiantes indígenas, negros y mulatos
	Salud	Porcentaje de estudiantes con discapacidad
	Social	Porcentaje de abandono escolar por institución
Docentes	Social	Porcentajes de docentes indígenas, negros y mulatos
	Entorno educativo	Ratio de estudiantes por docente
	Académica	Porcentaje de docentes con título de cuarto nivel
	Académica	Porcentaje de docentes con nombramiento
Instituciones	Infraestructura	Porcentaje estado de agua
	Infraestructura	Acceso a la Institución educativa
	Infraestructura	Porcentaje de estado de lavamanos
	Infraestructura	Área Urbana o Rural de la Institución
	Infraestructura	Porcentaje de instituciones con inodoros en buen estado
	Infraestructura	Porcentaje de instituciones con conexión a internet

3.4 Procedimientos

El objetivo de esta investigación fue diseñar un índice que, utilizando técnicas de clasificación multivariada, facilite la toma de decisiones en la priorización de instituciones educativas fiscales para ser beneficiarias de recursos educativos a través de proyectos de inversión. En esta sección se describen los procedimientos que se siguieron incluyendo la preparación de los datos, determinación de factores asociados a las instituciones educativas que contribuyen en la construcción del índice, técnicas del análisis multivariado para tratamiento de datos categóricos, Análisis de Correspondencias Múltiples y Análisis de Componentes Principales No lineales.

FASE 1. Determinación de los factores asociados a las instituciones educativas fiscales que contribuyen al diseño de un índice, utilizando técnicas de clasificación multivariada, para facilitar su priorización como beneficiarias de recursos educativos a través de proyectos de inversión.

Esta fase se centra en la determinación de los factores asociados a las instituciones educativas fiscales, identificando aquellos indicadores esenciales para el diseño de un índice que facilite la priorización en la asignación de recursos educativos. En esta etapa se recopilan variables cuantitativas como cualitativas, que abarcan aspectos como infraestructura, socio demográficos, académicos, tanto para estudiantes como docentes, permitiendo una evaluación integral de las necesidades y potencialidades de cada institución.

Para llevar a cabo esta identificación, se emplean técnicas de clasificación multivariada, destacándose el análisis de correspondencias múltiples y el análisis de componentes principales no lineales. El análisis de correspondencias múltiples facilita la exploración de las relaciones entre variables categóricas, permitiendo visualizar agrupamientos y patrones subyacentes, mientras que el análisis de componentes principales no lineales extrae dimensiones latentes en datos complejos, capturando relaciones no lineales que podrían pasar desapercibidas con métodos tradicionales. Estas técnicas, en conjunto, ofrecen una herramienta poderosa para desentrañar la estructura multidimensional de los datos educativos.

El resultado de esta fase es fundamental para la construcción de un índice de priorización robusto y representativo, ya que permite identificar y validar los factores críticos que influirán en la toma de decisiones para la asignación de recursos. La información derivada de estos métodos proporciona una base sólida para comprender la complejidad de las interacciones entre los diferentes indicadores, facilitando una clasificación objetiva y precisa de las instituciones. De esta manera, se establece un marco metodológico riguroso que contribuye a una gestión más equitativa y eficiente de la inversión en educación.

FASE 2. Evaluar y comparar dos técnicas de clasificación multivariada para optimizar la categorización de instituciones educativas fiscales, contribuyendo al diseño de un índice que facilite su priorización como beneficiarias de recursos educativos a través de proyectos de inversión.

En la Fase 2 se procede a evaluar y comparar dos técnicas de clasificación multivariada el Análisis de Correspondencias Múltiples y el Análisis de Componentes Principales No Lineal con el objetivo de optimizar la categorización de las instituciones educativas fiscales y contribuir al diseño de un índice que facilite su priorización para la asignación de recursos

educativos a través de proyectos de inversión. Esta etapa se centra en analizar detalladamente la capacidad de cada técnica para capturar patrones complejos y relaciones no lineales entre los indicadores, identificando sus fortalezas y limitaciones en el contexto del estudio. A continuación, se presenta un análisis descriptivo que explora el funcionamiento de ambas metodologías.

3.4.1 Análisis de Correspondencias Múltiples.

En la aplicación del Análisis de Correspondencias Múltiples se siguieron los siguientes pasos:

- Preparación de los datos.

Los datos provienen de los registros administrativos del Archivo Maestro de Instituciones Educativas AMIE, que recoge información detallada de las instituciones educativas, la fuente de información cuenta con variables cualitativas y cuantitativas. Entre las variables cuantitativas, por ejemplo, se encuentra el número de estudiantes que se autoidentifican como indígenas, afroecuatorianos y montubios. Durante el período académico 2022-2023 se registró una matrícula total de 3.247.260 estudiantes, de los cuales 192.708 se consideran indígenas, 54.788 se autoidentifican como afroecuatorianos y 59.709 como montubios. Estos tres grupos de autoidentificación se consolidaron en un único grupo étnico, representando a nivel nacional el 9,46% del total de estudiantes.

La transformación de esta variable consistió en identificar aquellas instituciones educativas cuyo porcentaje de estudiantes pertenecientes a este grupo étnico es menor o igual al promedio nacional del 9,46% y asignarles el valor de 1; a las instituciones que superan este porcentaje se les asigna el valor de 2. De esta manera, se crea una nueva variable denominada “etnia”, que permite clasificar de forma estandarizada a las instituciones según la proporción de estudiantes que se identifican como indígenas, afroecuatorianos o montubios, facilitando así el análisis multivariado y la posterior toma de decisiones en la priorización de recursos. En el caso de las variables cualitativas, como la ubicación de la institución educativa (área urbana o rural), se asignaron categorías numéricas basadas en la teoría, que establece que la condición en desventaja debe recibir un valor mayor. Así, se asignó el valor 1 a las instituciones ubicadas en áreas urbanas y el valor 2 a aquellas situadas en áreas rurales.

El procedimiento detallado de todas las variables se encuentra en el anexo B.

- Construcción de la tabla de contingencia:

El siguiente paso en este tipo de casos es transformar la matriz de datos en una pseudo tabla de contingencias con n filas por n columnas, donde cada una de las filas representa una de

las instituciones educativas y cada columna representa una variable que contiene un tipo de característica de la institución. Como se observa en la tabla 3.

Tabla 3

Seudo Matriz

Ins. Educat.	Var 1	Var 2	Var 3	Var ...	Var ...,	Var 11	Var 12
1	Cat1	Cat3	Cat2	Cat2	Cat3	Cat1	Cat2
2	Cat3	Cat3	Cat3	Cat2	Cat1	Cat1	Cat3
:	:	:	:	:	:	:	:
9517	Cat2	Cat2	Cat1	Cat2	Cat1	Cat2	Cat3
9518	Cat1	Cat1	Cat3	Cat1	Cat3	Cat3	Cat3

- Cálculo de Perfiles y Matrices de distancia

Se calculan los perfiles fila y columna, y utilizar la estadística chi-cuadrado para medir las distancias entre las categorías. Esto permite cuantificar la inercia y evaluar la variabilidad explicada por cada dimensión, el procedimiento es el siguiente:

Paso 1. Construcción de la tabla de frecuencias relativas

Sea $N = [n_{ij}]$ la tabla de contingencia donde n_{ij} representa el número de ocurrencias en la celda correspondiente a la fila i y la columna j , y $n = \sum_{i,j} n_{ij}$ el total de observaciones. Se define la matriz de frecuencias relativas:

$$p_{ij} = \frac{n_{ij}}{n} \quad (1)$$

Paso 2. Cálculo de las sumas marginales:

Las sumas marginales (o perfiles) para las filas y columnas se obtienen como:

$$r_i = \sum_j p_{ij} \quad \text{y} \quad c_j = \sum_i p_{ij} \quad (2)$$

Paso 3. Determinación de los perfiles

El perfil de la fila i es el vector:

$$f_i = \left(\frac{p_{ij}}{r_i} \right)_{j=1}^J \quad (3)$$

De forma análoga se procede para las columnas.

Paso 4. Cálculo de la distancia chi-cuadrado:

La distancia entre dos filas i e i' se define mediante la distancia chi-cuadrado:

$$d^2(i, i') = \sum_{j=1}^J \frac{1}{c_j} \left(\frac{p_{ij}}{r_i} - \frac{p_{i'j}}{r_{i'}} \right)^2 \quad (4)$$

Esta fórmula pondera las diferencias en los perfiles de cada fila según los pesos de las columnas, capturando así la contribución relativa de cada categoría.

- Aplicación de la Descomposición en Valores Singulares (SVD)

Aplicar la descomposición en valores singulares (SVD) a la matriz derivada (normalmente la matriz de residuales estandarizados) obtenida en el paso anterior. Esta descomposición permite extraer los valores singulares y los vectores singulares, que funcionan de manera similar a los valores y vectores propios, y que sirven para identificar las dimensiones significativas del análisis.

El algoritmo computacional para obtener las coordenadas de los perfiles de fila y columna con respecto a los ejes principales, utilizando el SVD, se describe a continuación:

Paso 1. Calcular la matriz de residuos normalizados

$$S = D_r^{-1/2} (P - rc^T) D_c^{-1/2} \quad (5)$$

Donde D_r y D_c son matrices diagonales de las frecuencias marginales de las filas y columnas, P es la matriz de frecuencias relativas, y r y c son los vectores de las frecuencias marginales.

Paso 2. Aplicar la descomposición SVD

$$S = U D_\alpha V^T \quad (6)$$

Aquí, U y V son matrices ortogonales, y D_α es la matriz diagonal de los valores singulares.

Paso 3. Calcular las coordenadas principales de las filas y columnas

1. Coordenadas principales de filas

$$F = D_r^{-1/2} U D_\alpha \quad (7)$$

2. Coordenadas principales de las columnas

$$G = D_c^{-1/2} V D_\alpha \quad (8)$$

Paso 4. Inercia y Varianza Total

$$Inercia = \Phi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(p_{ij} - r_i c_j)^2}{r_i c_j} \quad (9)$$

La inercia mide la varianza total de la matriz de datos y refleja la dispersión de los perfiles de filas y columnas respecto a un valor esperado. Este valor es comparable a un estadístico

chi-cuadrado y es útil para medir la proporción de variabilidad explicada por los primeros ejes principales.

En el anexo C se ejemplifica el Análisis de Correspondencia.

Estos pasos matemáticos permiten transformar la tabla de contingencia en un espacio de menor dimensión, facilitando la visualización e interpretación de las asociaciones entre categorías a través de gráficos como biplots.

- Representación gráfica y construcción del espacio factorial:

Proyectar las categorías y observaciones en el espacio multidimensional definido por las dimensiones seleccionadas, generando biplots u otros gráficos que faciliten la interpretación visual de las relaciones y agrupamientos.

- Interpretación y validación:

Se analiza la contribución de cada variable a las dimensiones obtenidas, interpretando el significado de las relaciones observadas.

3.4.2 Construcción del Índice de Clasificación de Instituciones Educativas a partir del Análisis de Correspondencias Múltiples.

Para la construcción del índice de clasificación de instituciones educativas se utilizó la siguiente fórmula:

$$Indice_{amie} = \frac{Total\ fila_i - \text{mínimo}(Total\ fila)}{\text{máximo}(Total\ fila) - \text{mínimo}(Total\ fila)} \times 100$$

Donde:

- *Total fila_i*, representa el resultado o puntaje obtenido para la institución *i* a partir del análisis realizado. Se trata de la suma o combinación de las puntuaciones (o coordenadas) derivadas del ACM que reflejan las características o desempeño de dicha institución.
- *mínimo(Total fila)*, es el valor más bajo obtenido entre los puntajes (Total fila) de todas las instituciones evaluadas. Este valor sirve de referencia para identificar el punto de partida en la escala de medición.
- *máximo(Total fila)*, es el valor más alto registrado en los puntajes (Total fila) de todas las instituciones. Junto con el mínimo, define el rango total de variación de los puntajes.

3.4.3 Análisis Componentes Principales No Lineales

“El Análisis de Componentes Principales Categóricos - No Lineal, es una técnica explicativa de reducción de dimensiones que proporciona datos numéricos y visuales para conjuntos de

datos entre los cuales existe una relación lineal o no lineal y que incluyen variables continuas, categóricas y discretas” (Demir y Keskin, 2022).

“El análisis de componentes principales (ACP) es un método descriptivo de uso común para manejar datos cuantitativos y puede ampliarse para tratar con datos de niveles de medición mixtos” (Mori et al., 2017), “Para el ACP extendido con tal mezcla de análisis cuantitativos y datos cualitativos, requerimos la cuantificación de datos cualitativos para poder para obtener datos de escala óptimos, ACP con escalamiento óptimo se denomina no lineal ACP” (Mori et al., 2017).

No lineal ACP que incluye escalas óptimas alterna entre la estimación de los parámetros de ACP y cuantificar datos cualitativos. El algoritmo de mínimos cuadrados alternos (ALS) se utiliza como algoritmo para ACP no lineal y puede encontrar soluciones de mínimos cuadrados mediante minimizar dos tipos de funciones de pérdida: una aproximación de rango bajo y una homogeneidad análisis con restricciones, PRINCIPALES de Young et al, (Componentes principales de datos multivariados de nivel de medición mixto: un método de mínimos cuadrados alternos con características de escala óptimas 43:279–281, 1978) y PRINCALS de Gifi (Nonlinear Analisis multivariable, Wiley, Chichester, 1990).

- Cuantificación de datos cualitativos.

A continuación, se detalla de manera descriptiva el proceso de asignación de valores numéricos a las categorías de variables cualitativas, tomado del documento “Nonlinear Principal Component Analysis and its Applications” de Mori, Yuichi y Kuroda, Masahiro y Makino, Naomichi, (2017).

El escalado óptimo es una técnica de cuantificación que asigna de manera óptima valores numéricos a escalas cualitativas y se ejemplifica de la siguiente manera:

Sea y_j de Y un vector cualitativo con K_j categorías, Para cuantificar y_j , el vector se codifica utilizando una matriz de indicadores $n \times K_j$

$$G_j = (g_{jik}) = \begin{pmatrix} g_{j11} & \cdots & g_{j1k_j} \\ \vdots & \vdots & \vdots \\ g_{jn1} & \cdots & g_{jnk_j} \end{pmatrix} = (g_{j1} \quad \cdots \quad g_{jk_j})$$

Donde:

$$g_{jik} = \begin{cases} 1, & \text{si el objeto } i \text{ pertenece a la categoría } k \\ 0, & \text{Si el objeto } i \text{ pertenece alguna otra categoría distinta de } k \end{cases}$$

Esta representación permite transformar una variable categórica en una forma adecuada para su análisis mediante métodos de reducción de dimensionalidad.

- Asignación de valores numéricos mediante escalado óptimo

En el ACP No Lineal, el objetivo es encontrar una cuantificación óptima x_{jk} para cada categoría k de la variable Y_j . Esta cuantificación se representa mediante el vector:

$$X_j = (x_{j1}, x_{j2}, \dots, x_{jk})^T \quad (10)$$

La transformación de los datos categóricos se expresa como:

$$Z_j = G_j X_j \quad (11)$$

Donde Z_j es la nueva representación cuantificada de la variable Y_j en el espacio transformado.

- Optimización de la cuantificación

Los valores X_j se determinan de manera óptima mediante un procedimiento iterativo que busca maximizar la varianza explicada en el análisis. Esto se logra resolviendo el siguiente problema de optimización:

$$\max_{X_j} \sum_{i=1}^n (Z_{ji} - \bar{Z}_j)^2 \quad (12)$$

sujeto a la restricción de que la varianza total sea igual a 1:

$$\frac{1}{n} \sum_{i=1}^n Z_{ji}^2 = 1 \quad (13)$$

Este proceso se repite iterativamente para cada variable cualitativa hasta que las cuantificaciones converjan a valores estables.

- Aplicación del ACP sobre los datos cuantificados

Una vez obtenidas las cuantificaciones Z_{ji} , se aplica el Análisis de Componentes Principales (ACP) sobre la matriz de datos transformados Z , utilizando descomposición en valores singulares (SVD):

$$Z = UDV^T \quad (14)$$

donde:

- U contiene los vectores singulares izquierdos,
- D es la matriz diagonal de valores singulares,
- V contiene los vectores singulares derechos.

Los primeros componentes principales extraídos de esta descomposición representan las dimensiones óptimas en las que los datos cualitativos han sido transformados, preservando la estructura de las relaciones originales. En el Anexo E muestran las puntuaciones óptimas transformadas.

3.4.4 Método "Principal Components Analysis by Nonlinear Least Squares" **PRINCALS**

El método PRINCALS (Principal Components Analysis by Nonlinear Least Squares) es una variante del Análisis de Componentes Principales No Lineal (ACP No Lineal) que permite transformar variables categóricas en puntuaciones cuantitativas optimizadas. A diferencia del Análisis de Correspondencias Múltiples (ACM), que representa categorías en un espacio factorial, PRINCALS asigna valores numéricos óptimos a las categorías y luego aplica ACP en estos datos transformados.

Esta transformación se realiza minimizando una función de pérdida, que mide la diferencia entre los datos originales y sus aproximaciones en el espacio de componentes principales.

$$\min \sum_{i=1}^n \sum_{j=1}^m w_{ij} (x_{ij} - \hat{x}_{ij})^2 \quad (15)$$

donde:

- n es el número de observaciones,
- m es el número de variables,
- x_{ij} son los datos observados,
- \hat{x}_{ij} son los datos modelados,
- w_{ij} son pesos asignados a las diferencias, lo que permite flexibilidad en el trato de diferentes escalas y tipos de datos.

En PRINCALS, cada categoría de las variables se asigna a una puntuación numérica, Estas puntuaciones se ajustan iterativamente para encontrar la mejor representación lineal de los datos en un espacio dimensional reducido, manteniendo al mismo tiempo la estructura ordinal o nominal de los datos originales.

El ajuste del modelo en PRINCALS implica encontrar vectores de puntuaciones para los individuos y cargas para las variables que minimicen la función de pérdida. Este proceso se realiza iterativamente mediante un algoritmo de optimización, comúnmente un algoritmo de alternancia de mínimos cuadrados.

Los componentes resultantes de PRINCALS son interpretados como las direcciones de máxima varianza en el espacio transformado, similar al ACP tradicional. Sin embargo, en PRINCALS, estas direcciones también reflejan la mejor representación lineal de las escalas ordinales o nominales subyacentes de las variables.

- Proceso iterativo de PRINCALS

En PRINCALS, el objetivo es encontrar transformaciones óptimas para variables categóricas nominales y ordinales, de modo que estas puedan integrarse en un análisis de componentes principales y maximicen la varianza explicada. El método de mínimos cuadrados alternantes (ALS) realiza este ajuste de manera iterativa.

- Transformación inicial de las variables categóricas

Dada una variable cualitativa X_i con C_i categorías, se construye una matriz indicadora Z_i de dimensión $n \times C_i$, donde cada fila indica la categoría a la que pertenece la observación.

Ejemplo para una variable con 3 categorías (A, B, C) y 4 observaciones:

$$Z_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Donde

- Las observaciones 1 y 3 pertenecen a la categoría A.
- La observación 2 pertenece a la categoría B.
- La observación 4 pertenece a la categoría C.
- Inicialización de puntuaciones:

Al inicio, se asignan valores provisionales (arbitrarios) a cada una de las categorías de la variable X_i . Esto se puede hacer asignando valores numéricos iniciales $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{iC_i})$ a cada categoría. Esta asignación inicial puede ser aleatoria o seguir alguna regla básica.

$$\mathbf{y}_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ y_{i3} \end{pmatrix}$$

Para una variable con 3 categorías (A, B, C), los valores iniciales pueden ser $y_{i1} = 0.1$, $y_{i2} = 0.3$, $y_{i3} = 0.7$.

- Ajuste de las transformaciones (Optimización)

El algoritmo ALS ajusta los valores numéricos asignados a las categorías para maximizar la varianza explicada. Esto se hace a través de una optimización iterativa:

Paso 1: Proyección en las componentes principales

Para obtener los valores ajustados de las variables categóricas, PRINCALS maximiza la varianza total de las transformaciones de las categorías. Esto implica proyectar las variables categóricas (ya transformadas en numéricas) en los componentes principales, al igual que en el Análisis de Componentes Principales (ACP).

La varianza total explicada por el modelo está relacionada con los autovalores $\lambda_1, \lambda_2, \dots, \lambda_k$ obtenidos a partir de la descomposición en valores propios de la matriz de covarianza de los datos transformados.

$$\text{Varianza explicada} = \frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^p \lambda_i} \quad (16)$$

donde

- λ_i son los autovalores de la matriz de covarianza,
- k es el número de componentes seleccionados
- p es el número total de variables (transformadas).

Paso 2: Minimización del error cuadrático

El algoritmo utiliza métodos de mínimos cuadrados para minimizar el error de ajuste entre los valores numéricos transformados de las categorías y las proyecciones en las componentes principales. Se define un criterio de optimización que minimiza la distancia entre los valores transformados de las categorías y la estructura de los componentes principales.

El error que se minimiza es la suma de los errores cuadrados entre los valores observados Z_i las variables indicadoras de las categorías y los valores transformados \mathbf{y}_i , representado por la siguiente función de error:

$$\text{Error} = \sum_{i=1}^n \sum_{j=1}^{C_i} (Z_{ij} - f(\mathbf{y}_i))^2 \quad (17)$$

donde:

- $f(\mathbf{y}_i)$ es la función que proyecta los valores numéricos transformados de las categorías sobre las componentes principales. El objetivo es encontrar la transformación \mathbf{y}_i que minimice este error cuadrático.
- Repetición (Iteración hasta convergencia)

Actualización de las transformaciones:

Después de cada iteración, se ajustan los valores de las categorías. Este ajuste se hace actualizando \mathbf{y}_i en función del resultado de la proyección en las componentes principales y minimizando el error cuadrático.

Este proceso se repite iterativamente, actualizando las transformaciones de todas las variables categóricas de manera alternante, hasta que el cambio entre las soluciones consecutivas es pequeño o insignificante (convergencia).

Matemáticamente, el algoritmo busca minimizar el error cuadrático en cada iteración, recalculando los valores óptimos \mathbf{y}_i mediante:

$$\mathbf{y}_i^{(t+1)} = \underset{\mathbf{y}_i}{\operatorname{argmin}} \sum_{i=1}^n \left(Z_i - f(\mathbf{y}_i^{(t)}) \right)^2 \quad (18)$$

donde

- $\mathbf{y}_i^{(t)}$ son los valores transformados en la iteración t y
- $\mathbf{y}_i^{(t+1)}$ son los valores actualizados.
- Criterio de convergencia

La convergencia se alcanza cuando la diferencia entre los valores numéricos transformados de las categorías en dos iteraciones consecutivas es menor que un umbral predefinido. Esto se mide utilizando una métrica como la norma euclidiana de la diferencia entre las transformaciones anteriores y las actuales:

$$\| \mathbf{y}_i^{(t+1)} - \mathbf{y}_i^{(t)} \| < \epsilon \quad (19)$$

donde

- ϵ es un umbral pequeño que indica que los valores transformados han alcanzado un punto estable.

3.4.5 Construcción del Índice de Clasificación de Instituciones Educativas a partir del Análisis de Componentes Principales No Lineales.

Para elaborar el índice de instituciones educativas fiscales, se parte de las puntuaciones transformadas de cada categoría de las variables intervinientes. En primer lugar, a la categoría con el valor mínimo se le asigna un valor de cero, y se suma dicha cantidad a las demás categorías de la misma variable para conservar las diferencias relativas sin alterarlas. Posteriormente, se selecciona el valor máximo de cada variable (tomando los valores transformados de sus categorías) y se suman estos máximos. Con el fin de obtener una puntuación final comprendida entre 0 y 100, se calcula una constante dividiendo 100 entre la suma de dichos valores máximos. Esta constante se utiliza para reescalar cada valor de categoría (ya ajustado en función del mínimo), produciendo así las puntuaciones finales que integran el índice de priorización.

CAPÍTULO IV

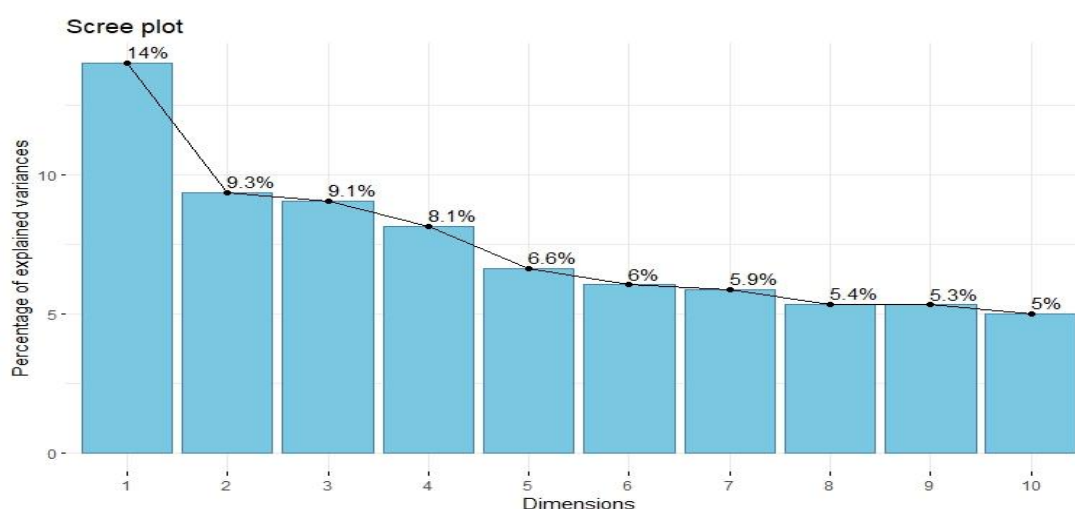
RESULTADOS Y DISCUSIÓN

4.1 Resultados del Análisis de Correspondencias Múltiples.

Al aplicar el Análisis de Correspondencias Múltiples, es necesario analizar la información detallada sobre la estructura de los datos e identificar los ejes más importantes, es decir aquellos componentes donde se presenta la inercia, que representan la mayor parte de la variabilidad presente en los datos, esto se realiza a través del gráfico Scree Plot (Figura 3).

Figura 3

Scree plot Distribución de la Inercia entre las dimensiones



Con el fin de robustecer la construcción del índice se toman las 5 primeras dimensiones que en conjunto explican el 47% de la variabilidad, conforme se muestra en la Tabla 4.

Tabla 4

Varianza explicada 5 dimensiones

Dimensión	Varianza explicada
Dimensión 1	0,1391
Dimensión 2	0,0936
Dimensión 3	0,0906
Dimensión 4	0,0813
Dimensión 5	0,0662
Total 5 Dimensiones	0,47087

A partir de la información de los cinco ejes (dimensiones) en el Análisis de Correspondencias Múltiples (ACM), puede observarse que cada uno capta un conjunto distinto de variables y categorías con alta contribución. A grandes rasgos, la contribución indica cuánto “dibuja” o

caracteriza cada categoría en el eje correspondiente, mientras que la coordenada (“Coord”) y el coseno cuadrado (“Cos2”) ayudan a interpretar la posición y la calidad de representación de la categoría. Si el \cos^2 es bajo, la categoría no está bien explicada por los ejes seleccionados.

A continuación, se presenta un resumen eje por eje, destacando las variables con mayor influencia en cada caso.

EJE 1

Tiene como principales contribuyentes a: agua_1 correspondiente al servicio de agua calificado como bueno (12.38%), luego tenemos a las instituciones educativas que cuentan con internet representados por internet_1 (9.78%), y las que no cuentan con internet representado por internet_2 (9.47%), los docentes con título de cuarto nivel doc4n_1 (9.04%), el servicio de agua calificado como malo agua_3 (6.22%), y el ratio estudiante por docente mayor al porcentaje nacional r_esdoc_2 (6.12%).

En tal sentido se puede concluir que, la disponibilidad o tipo de servicio de agua (categorías agua_1, agua_3) y conectividad a internet (internet_1, internet_2) marcan fuertemente la separación de los datos en este eje.

El nivel de formación docente (doc4n_1) y la variable relacionada con la relación estudiante-docente (r_esdoc_2) también aportan a la variabilidad.

Otras variables con contribuciones intermedias son la zona geográfica area2_1, area2_2, urbana y rural respectivamente y la presencia de discapacidades en las instituciones educativas mayores al porcentaje nacional (disc_2).

EJE 2

En este eje los principales contribuyentes son:

La etnia de los docentes autoidentificados como negros, indígenas y montubios representados por etniad_2 (24.58%), de la misma forma para el caso de los estudiantes etnia_2 (23.41%), de igual manera se tiene a la etnia de docentes auto identificados como blancos o mestizos etniad_1 (12.72%), de igual forma con los estudiantes con la autoidentificados de blancos y mestizos etnia_1 (8.66%), también se incluye el abandono en instituciones con porcentajes mayores al abandono nacional aband_2 (8.64%), y el acceso a la institución educativa de forma fluvial acceso_2 (5.17%).

En resumen, este eje está dominado por las variables etnia de docentes y estudiantes (tanto en su categoría 1 como 2), lo que indica que la auto identificación étnica o composición estudiantil tiene un peso muy importante en la diferenciación de las instituciones en esta dimensión.

Por otra parte, la variable abandono escolar aband (aband_1, aband_2) también contribuye, sugiriendo que la condición de abandono o no de la institución puede estar correlacionada con la composición étnica de la población estudiantil.

Además, acceso_2 acceso fluvial aparece con contribución notable, mostrando que el tipo de acceso a la institución también diferencia los datos en este eje.

EJE 3

Principales contribuyentes:

Inodoros en esta regular representado por inodoro_2 (27.21%), lavamanos en estado regular lavamanos_2 (25.01%), lavamanos en buen estado lavamanos_1 (14.69%), e inodoro en buen estado inodoro_1 (13.23%).

La variable inodoro (categorías 1 y 2) y la variable lavamanos (categorías 1 y 2) sobresalen ampliamente, sumando entre ellas más del 80% de la contribución total en este eje.

Esto indica que la presencia o tipo de instalaciones sanitarias (baños y lavamanos) es el factor principal que separa las instituciones a lo largo de este tercer eje.

El resto de las categorías (doc4n_1, area2_1, acceso_2, etc.) tienen contribuciones más bajas, lo que refuerza que el eje 3 se concentra en la distinción de infraestructura sanitaria.

EJE 4

Principales contribuyentes:

Inodoros en mal estado inodoro_3 (38.37%), lavamanos en mal estado lavamanos_3 (37.17%), lavamanos en estado regular lavamanos_2 (7.68%) e inodoro en estado regular inodoro_2 (4.56%).

De nuevo, las variables inodoro y lavamanos son determinantes, pero esta vez destacan específicamente las categorías “_3” (inodoro_3, lavamanos_3) con contribuciones muy altas.

Se trata de un eje que, al igual que el Eje 3, refina la diferenciación en cuanto al tipo o estado de instalaciones sanitarias.

El resto de las categorías (agua_2, agua_3, r_esdoc_2, etc.) aportan mucho menos, confirmando que este eje está muy focalizado en la variabilidad generada por las condiciones de los servicios higiénicos.

EJE 5

Principales contribuyentes:

El agua en estado regular agua_2 (20.15%), agua calificada como mala agua_3 (19.64%), ratio estudiantes por docente mayor al ratio nacional, r_esdoc_2 (12.79%), acceso a la

institución educativa de manera fluvial acceso_2 (10.3%) e inodoro en mal estado inodoro_3 (6.72%).

Aquí destacan de nuevo las categorías relacionadas con el tipo de servicio de agua, junto con r_esdoc_2 relación estudiante-docente, acceso_2 (probablemente acceso fluvial), y inodoro_3.

El patrón sugiere que en este quinto eje se combinan infraestructura de agua, acceso a la institución y ciertos aspectos de infraestructura sanitaria y recursos docentes.

En conclusión, en el análisis de correspondencias múltiple destacan factores asociados a las instituciones educativas como se detalla a continuación:

Infraestructura básica y sanitaria:

A lo largo de varios ejes (1, 3, 4 y 5), las variables “agua”, “inodoro” y “lavamanos” aparecen como factores clave, lo cual indica que la disponibilidad y tipo de servicios higiénicos tienen un fuerte poder discriminante entre las instituciones.

Composición sociocultural y localización:

El Eje 2 muestra que la autoidentificación étnica (etnia, etniad) y el tipo de acceso (acceso_2) son factores que también explican una parte sustancial de la variabilidad.

Otros factores a considerar:

Conectividad a internet, relación estudiante-docente (r_esdoc), la condición de abandono (aband) y la formación docente (doc4n) también aparecen repetidamente con contribuciones notables, aunque su relevancia varía de un eje a otro.

Obtención de los ponderadores de cada categoría.

Para la obtención de los ponderadores se realizaron los siguientes pasos:

Se consideraron las coordenadas de cada uno de los individuos en cada dimensión considerada para el análisis, estas coordenadas se obtuvieron al aplicar la función MCA.

En el Análisis de Correspondencias Múltiples, cada individuo o caso en el conjunto de datos es representado como un punto en un espacio de dimensiones múltiples. Cada dimensión del espacio resultante intenta capturar la máxima cantidad de información posible sobre las inercias - variaciones de los datos. Las coordenadas de los individuos en cada dimensión son críticas porque representan la posición del individuo respecto a las categorías de las variables incluidas en el análisis.

Dado que el primer punto de inflexión se encontraba en las dos primeras dimensiones, las cuales explican solo el 25% de la variabilidad, para la construcción del índice se utilizaron

las cinco primeras dimensiones, como se ejemplifica en la tabla 5. En este rango se encontraba el segundo punto de inflexión y se capturó el 47% de la variabilidad total.

Tabla 5

Dimensiones Análisis Correspondencias Múltiple

No.	AMIE	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
1	01H00568	0.2886	-0.1589	-0.4116	0.00393	0.09059
2	01H00630	-0.5289	0.30785	-0.1513	-0.0415	-0.24664
3	01H00635	0.6353	-0.22368	0.32517	-0.4130	-0.17632
,	,	,	,	,	,	,
,	,	,	,	,	,	,
9516	17H00680	-0.5141	0.07711	-0.1420	0.10425	0.32125
9517	17H00539	-0.6357	0.29954	-0.0675	0.09605	0.34380
9518	17H00864	-0.8418	0.29107	0.04276	0.11444	0.11768

Se Multiplicaron las coordenadas obtenidas en el paso 1 por el porcentaje de varianza explicada descrita en la Tabla 6.

Tabla 6

Dimensiones por varianzas

No	AMIE	D1*V1	D2*V2	D3*V3	D4*V4	D5*V5
1	01H00568	0,288 * 0,139	„ * „	„ * „	„ * „	0.0905 * 0,066
2	01H00630	-0,528 * 0,139	„ * „	„ * „	„ * „	-0.2466 * 0,066
3	01H00635	0,635 * 0,139	„ * „	„ * „	„ * „	-0,1746 * 0,066
,	,	,	„ * „	„ * „	„ * „	,
,	,	,	„ * „	„ * „	„ * „	,
9516	17H00680	-0.514 * 0,139	„ * „	„ * „	„ * „	0.3212 * 0,066
9517	17H00539	-0.635 * 0,139	„ * „	„ * „	„ * „	0.3438 * 0,066
9518	17H00864	-0.841 * 0,139	„ * „	„ * „	„ * „	0.1176 * 0,066

Totales por fila

Se sumó para cada AMIE, los resultados obtenidos en cada dimensión al multiplicarse por la varianza, es decir, se sumaron por filas los resultados obtenidos en el paso anterior, como se muestra en la tabla 7de totales por fila.

Tabla 7

Totales por fila

AMIE	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Totales por fila
01H00568	0,04016	-0,0129	-0,0385	0,00026	0,00821	-0,0028

01H00630	-0,0495	0,0203	-0,0137	-0,0057	-0,0200	-0,0686
01H00635	0,05755	-0,0311	0,02643	-0,03867	-0,0116	0,00250
”	”	”	”	”	”	”
”	”	”	”	”	”	”
17H00680	-0,0715	0,0062	-0,0133	0,00691	0,02910	-0,0425
17H00539	-0,0595	0,01984	-0,0061	0,01337	0,02794	-0,0045
17H00864	-0,0762	0,04050	0,00348	0,01072	0,00779	-0,0137

Re escalamiento.

Para el re escalamiento se debe considerar los valores máximo y mínimo de la columna

Totales por fila:

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
-0,14819	-0,05461	-0,01226	0,00018	0,04036	0,453671

Se seleccionó el valor mínimo de los totales por fila.

Se restó de los Totales por fila el valor mínimo de los Totales por fila

Se obtuvo el valor máximo de la columna Totales por fila y se le restó el valor mínimo de los Totales por fila.

Fórmula y ejemplo aplicación

En función de los pasos establecidos se obtiene la siguiente fórmula:

$$Indice_{amie} = \frac{Total\ fila_i - \text{mínimo}(Total\ fila)}{\text{máximo}(Total\ fila) - \text{mínimo}(Total\ fila)} \times 100$$

Aplicación de fórmula para la institución 01H00568

$$Indice_{01H00568} = \frac{-0,00284 - (-0,14819)}{0,453671 - (-0,14819)} \times 100$$

$$Indice_{01H00568} = \frac{0,14535184}{0,60186} \times 100$$

$$Indice_{01H00568} = 24,15$$

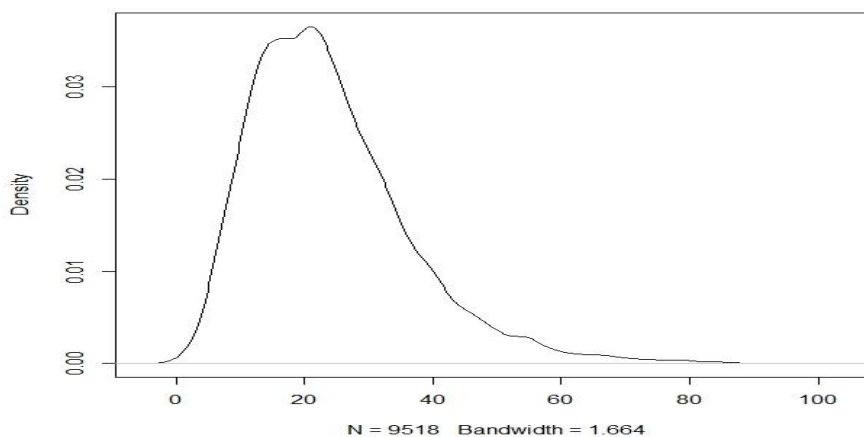
4.2 Resultados Índice de Clasificación de Instituciones Educativas aplicando el Análisis de Correspondencias Múltiples.

Una vez re escalados los valores en un rango de 0 a 100, donde los valores cercanos a 100 representan las peores condiciones y los cercanos a 0 las mejores, se obtiene la siguiente

distribución de los puntajes obtenidos por las instituciones educativas a través del ACM (Figura 4).

Figura 4

Curva distribución de puntuaciones ACM.



Como se observa en la figura 4, la distribución presenta un sesgo hacia la derecha, con un valor de asimetría de 1.1353, lo que indica una cola larga en esa dirección. Esto sugiere la presencia de algunos valores atípicos altos que afectan la forma de la distribución. La densidad disminuye gradualmente a medida que los valores aumentan, lo que revela que la mayoría de los índices de las instituciones educativas están concentrados en los valores más bajos, es decir, en instituciones con mejores condiciones.

Esta distribución facilita la identificación de necesidades específicas de intervención en ciertos grupos de instituciones educativas, lo que permite desarrollar políticas diferenciadas enfocadas en mejorar las condiciones de aquellas que se encuentran en el extremo derecho de la distribución

4.2.1 Identificación de las Instituciones educativas que cuentan con condiciones menos favorables.

El re-escalamiento de las coordenadas obtenidos por las categorías de las instituciones educativas permite obtener un listado ordenado conforme a los puntajes obtenidos, a partir de lo cual se puede clasificar de formas distintas a las instituciones, en este caso se clasifica por percentiles, más precisamente por quintiles teniendo en el quinto quintil a las de condiciones menos favorables y en el quintil uno, a las instituciones educativas con mejores condiciones, esto se puede observar en la tabla 8.

Tabla 8*Distribución de instituciones por quintiles, según índice de ACM*

Quintiles	No. Instituciones	Rango puntuaciones
Quintil 1	1,904	0,00 – 14,1557
Quintil 2	1,903	14,16491 – 19,84615
Quintil 3	1,904	19,85051 – 25,49592
Quintil 4	1,903	25,49867 – 33,60453
Quintil 5	1,904	33,61803 – 100,00

Geográficamente las instituciones que se encuentran en el quintil 5 se distribuyen conforme se muestra en la tabla 9.

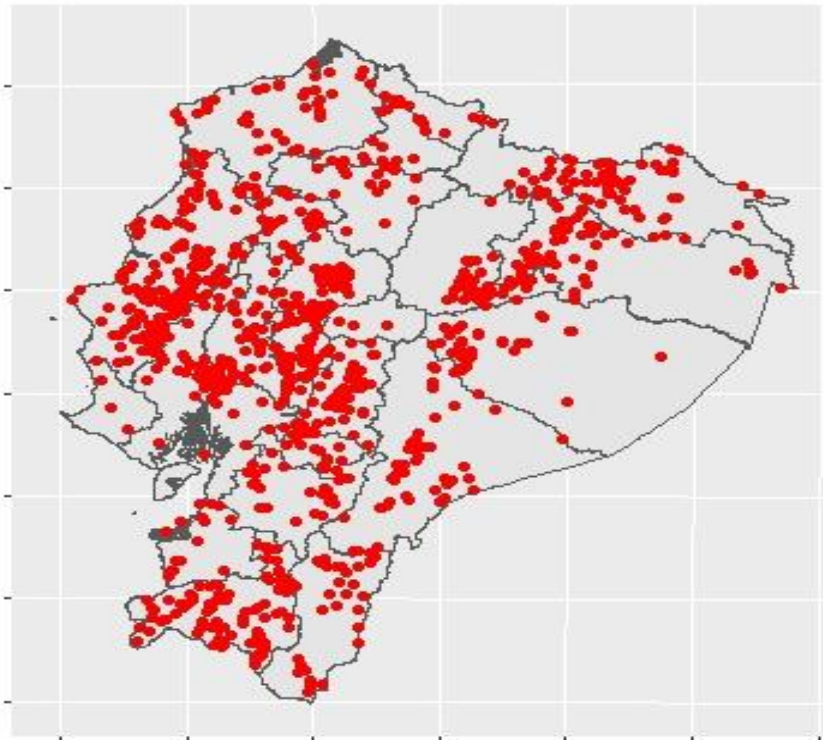
Tabla 9*Distribución instituciones menos favorecidas, según índice ACM.*

Provincia	Instituciones Q5
Manabí	260
Guayas	257
Los Ríos	180
Esmeraldas	140
Morona Santiago	117
Orellana	110
Sucumbíos	104
Loja	86
Chimborazo	79
Napo	75
Cotopaxi	74
Pichincha	60
Pastaza	57
Bolívar	55
Zamora Chinchipe	47
Azuay	43
Santo Domingo De Los Tsáchilas	43
El Oro	37
Cañar	29
Imbabura	25
Carchi	10
Santa Elena	9
Tungurahua	4
Galápagos	3

Para visualizar la distribución geográfica se elaboró un mapa de distribución de Instituciones educativas del Quintil 5 (Figura 5), conforme al cálculo realizado del Índice de Instituciones Educativas obtenido a partir de Análisis de Correspondencias Múltiples., como se puede observar, instituciones con estas características se encuentran en todo el país, con importantes concentraciones en provincias de la Costa y Amazonía.

Figura 5

Mapa de puntos de ubicación instituciones educativas pertenecientes al quintil de instituciones menos favorecidas.



4.3 Procedimiento Aplicación del Análisis de Componentes Principales No Lineales - Categórico

- Transformación de variables a formato categórico

Para homogenizar el tipo de datos, todas las variables incluidas en el análisis se transformaron a categorías según criterios establecidos (ver Anexo B). De este modo, variables inicialmente cuantitativas se reagruparon en intervalos o categorías de interés para el estudio.

- Estandarización

Una vez definidas las variables categóricas, se estandarizó la matriz de datos para uniformar las escalas. Posteriormente, se aplicó el algoritmo **PRINCALS** (implementado en la librería **GIFI** de R) a fin de realizar el Análisis de Componentes Principales No Lineales.

- Obtención de los Eigenvalores

El análisis genera componentes que resumen la variabilidad de los datos originales. Los eigenvalores o valores propios reflejan cuánta varianza explica cada componente (Figura 6).

Figura 6

Varianza explicada por componente



Para calcular el porcentaje de varianza explicada por el componente

$$\% \text{ Varianza explicada} = \left(\frac{\text{Eigenvalue}}{\sum_i^n \text{Eigenvalues}} \right) \times 100$$

Por lo tanto, la varianza explicada por el primer componente será igual a:

$$\% \text{ Varianza explicada} = \left(\frac{2,3749}{13} \right) \times 100$$

$$\% \text{ Varianza explicada} = 18,26\%$$

De igual forma para el segundo componente:

$$\% \text{ Varianza explicada} = \left(\frac{1,5758}{13} \right) \times 100$$

$$\% \text{ Varianza explicada} = 12,12\%$$

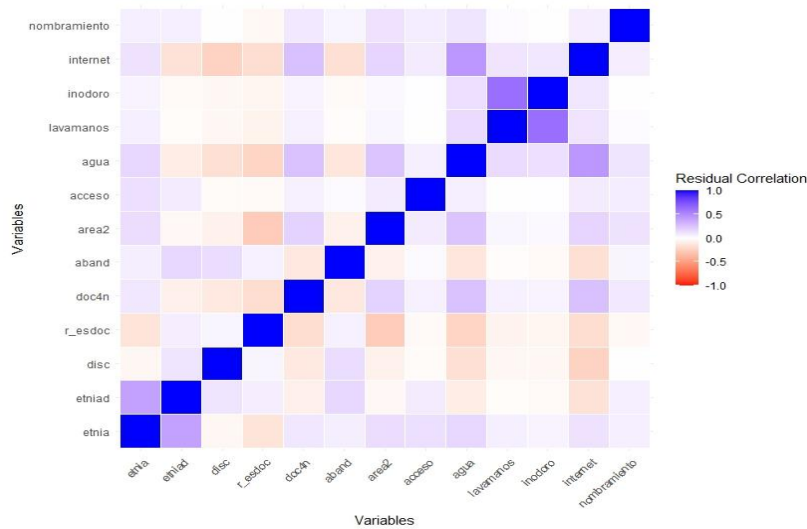
Los porcentajes explicados para los trece componentes se detallan en el Anexo D.

- Matriz de Correlación Residual

Muestra las correlaciones residuales entre las variables, se calcula a partir de la diferencia entre la matriz de correlaciones observadas y la matriz de correlaciones construidas en función de los componentes principales seleccionados (Figura 7).

Figura 7

Matriz de Correlaciones variables ACP No Lineal



La matriz de correlación residual compara las correlaciones observadas entre las variables con las correlaciones estimadas por los componentes principales seleccionados.

- Correlaciones residuales bajas (cerca de cero) indican que los componentes principales capturan bien la relación entre las variables.
- Correlaciones residuales altas sugieren que todavía hay variabilidad no explicada por los componentes retenidos (por ejemplo, agua–internet y lavamanos–inodoro mostraron correlaciones residuales significativas).
- Rotación Varimax

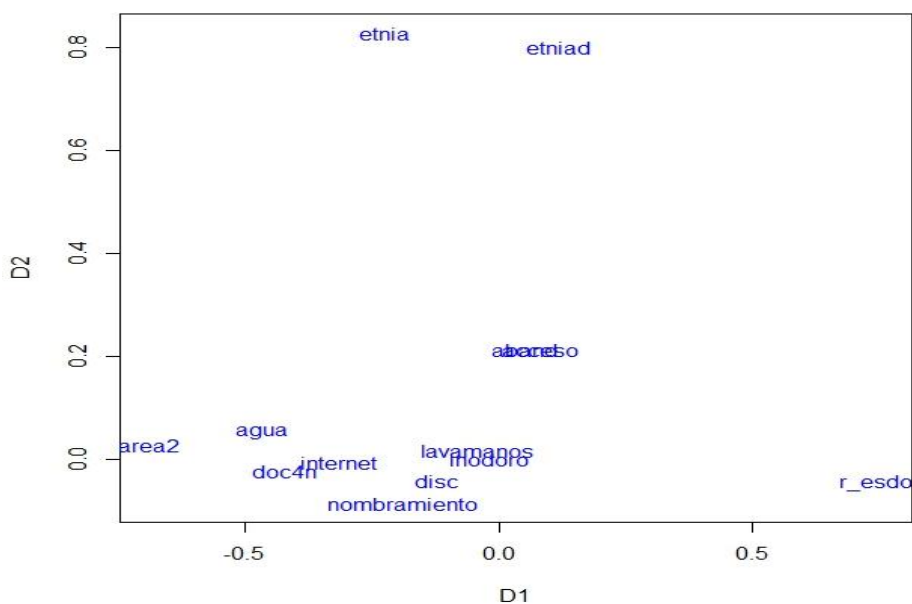
Para facilitar la interpretación, se aplicó una rotación Varimax a los componentes (Figura 8). Esta rotación ortogonal busca maximizar la varianza de las cargas factoriales al interior de cada componente, de modo que cada variable quede fuertemente asociada con un único componente y débilmente asociada con los demás. Así, la estructura se hace más clara:

Dimensión D1 (eje X): Agrupa las variables con mayor contribución a la varianza inicial.

Dimensión D2 (eje Y): Explica parte de la variabilidad que D1 no capta

Figura 8

Rotación Varimax



- Interpretación de los componentes y factores asociados

Del gráfico Rotación Varimax (Figura 8) se destacan los siguientes hallazgos:

Relación estudiantes/docente (r_esdoc)

Se ubica muy a la derecha en D1, indicando un alto peso en la explicación de la variabilidad. Por ende, la ratio estudiantes/docente resulta un factor relevante para diferenciar y priorizar instituciones.

Variables étnicas (etnia, etniad)

Se sitúan alejadas del origen en D2, reflejando una fuerte contribución de la diversidad étnica a la variabilidad capturada por el segundo componente. Estas variables sociodemográficas adquieren especial importancia para la clasificación de instituciones.

Abandono escolar (aband) y acceso (acceso)

Se superponen en el plano, sugiriendo perfiles muy similares. Su proximidad indica correlación entre la forma de llegar a la institución (terrestre, fluvial, aérea) y las tasas de abandono.

Variables de infraestructura (agua, internet, lavamanos, inodoro, etc.)

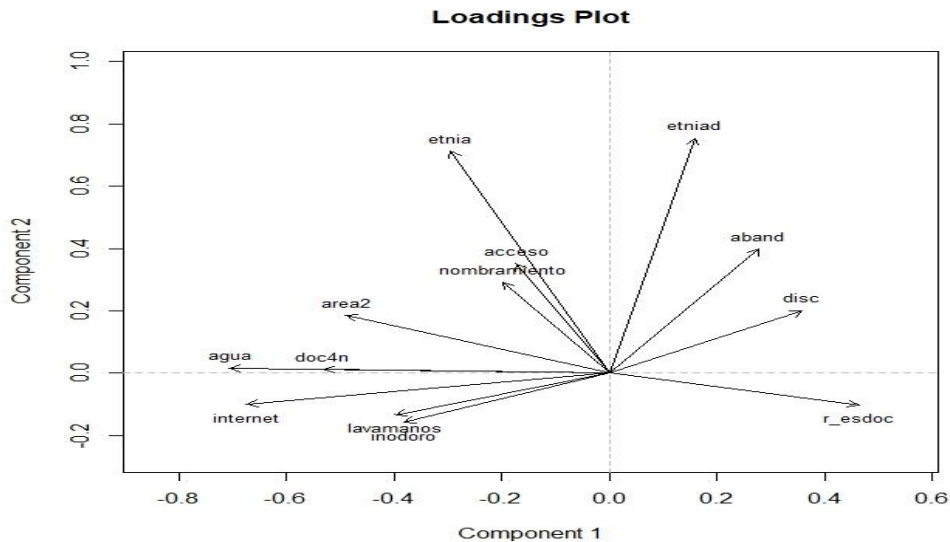
Aparecen cercanas al origen en las dos primeras dimensiones, sugiriendo que no son los principales diferenciadores en D1 y D2. No obstante, podrían cobrar relevancia en componentes posteriores o presentar correlaciones residuales entre sí (como se vio en la matriz residual).

- Puntuaciones transformadas y Loading Plot

Las puntuaciones transformadas son el resultado de cuantificar las variables categóricas y proyectarlas en las nuevas dimensiones, conforme se observa en el Loading Plot (Figura 9):

Figura 9

Loading Plot



La Figura 9 nos muestra que:

Agua, lavamanos, inodoro, internet y docentes de cuarto nivel, se agrupan en el eje del Componente 1, lo que sugiere que capturan aspectos de infraestructura y recursos.

Etnia, abandono y discapacidad, se orientan hacia el extremo positivo del Componente 1, indicando mayor influencia de factores sociodemográficos.

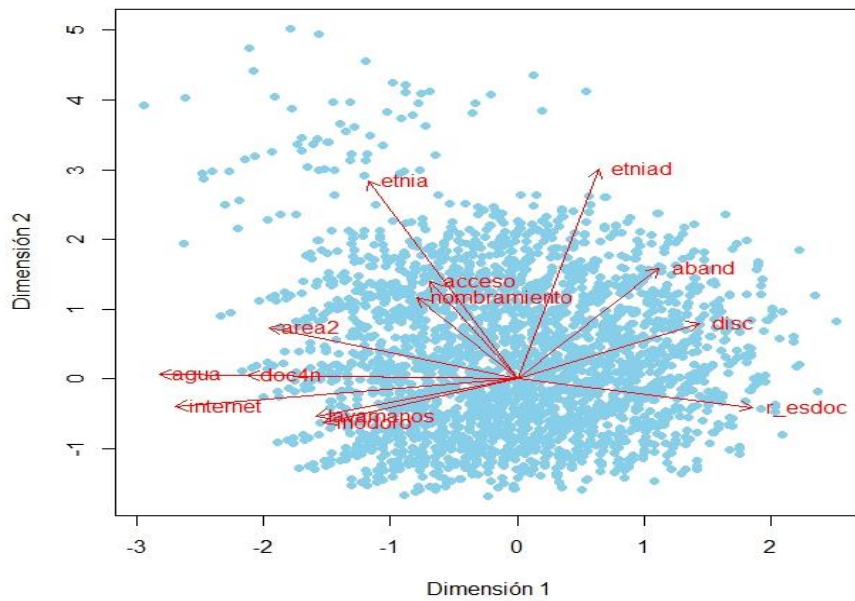
Nombramiento docente y acceso, se ubican más en el Componente 2, reflejando gestión institucional y condiciones de localización.

- Biplot de categorías e individuos.

En el Biplot (Figura 10), las variables se representan como flechas y los individuos (instituciones educativas) como puntos:

Figura 10

Biplot de categorías e individuos



Longitud de flecha: Mide la influencia de la variable en la dimensión.

Dirección de flecha: Indica cómo varían los individuos con respecto a la variable.

VARIABLES ALEJADAS DEL ORIGEN: Tienen mayor contribución a la variabilidad total (por ejemplo, etnia, etniad, aband, r_esdoc, agua, internet, disc).

Este análisis confirma que etnia, relación estudiante-docente, abandono y tipo de acceso ejercen un impacto notable en la diferenciación de las instituciones, mientras que las variables de infraestructura podrían requerir más componentes para manifestar su contribución o presentar correlaciones residuales.

- Reescalamiento de categorías para la construcción del índice

Finalmente, se elaboró un ranking a partir de las puntuaciones transformadas, siguiendo estos pasos:

1. Ajuste de mínimos: A la categoría con menor valor se le asigna 0 y se desplazan las demás categorías de la misma variable en la misma magnitud para conservar diferencias relativas.
2. Selección de máximos: Se identifica el valor máximo de cada variable y se suman dichos valores para definir el rango total.
3. Normalización a 0–100: Se divide 100 entre la suma de máximos (por ejemplo, $100 / 0,38588 \approx 259,14$), obteniendo una constante de escalamiento.
4. Cálculo del índice: Cada puntuación transformada se multiplica por dicha constante para asignar valores en la escala 0–100

Este procedimiento produce un índice final que integra los factores más relevantes (sociodemográficos, de acceso, relación estudiantes/docente, etc.) y facilita la priorización de instituciones según sus necesidades o condiciones identificadas.

Una vez obtenidos los cuantificadores ponderados del índice de escalamiento, como se muestra en la Tabla 10, estos valores se asignan a las instituciones educativas en función de la información original de cada variable transformada. En este proceso, si una institución educativa presenta un porcentaje de estudiantes que se auto perciben como indígenas, afrodescendientes o montubios inferior al promedio nacional (categorizado como 1), se le asigna una puntuación de 0. Por el contrario, si el porcentaje es igual o superior al promedio nacional, se le asigna un valor de 6.05079. Este procedimiento se aplica a todas las categorías de las variables y, al final, se suman los valores obtenidos, determinando así el puntaje final de cada institución.

Tabla 10

Cuantificadores ponderados por índice de escalamiento

No	Categoría	Valor categoría	Punt. transformador	Mínimo	Transforma	Máximo	Cuantificador ponderadas índice de escalamiento
	Estudiantes auto identificados como Afro, montubios, indígenas menor al nacional	1	-0,0060		0,00000	0,0233	0,00000
1	Estudiantes auto identificados como	2	-0,006		0,02335		6,05079

	Afro, montubios, indígenas mayor al nacional		0,0172			
	Docentes auto identificados como Afro, montubios, indígenas menor al nacional	1	-0,0073		0,0000	0,02162
	Docentes auto identificados como Afro, montubios, indígenas mayor al nacional	2	0,0142			
2				-0,0073	0,02162	5,60327
	Estudiantes con discapacidad menor al nacional	1	-0,0063		0,00000	0,02293
	Estudiantes con discapacidad mayor al nacional	2	0,0166			
3				-0,0063	0,02293	5,94291
	Ratio estudiantes por docente menor al nacional	1	-0,0069		0,00000	0,02209
	Ratio estudiantes por docente mayor al nacional	2	0,0151			
4				-0,0069	0,02209	5,72520
	Docentes cuarto nivel mayor al nacional	1	-0,0187		0,00000	0,02436
	Docentes cuarto nivel menor al nacional	2	0,0056			
5				-0,0187	0,02436	6,31210
	Docentes nombramiento mayor al Nacional	1	-0,0096		0,00000	0,02054
	Docentes nombramiento menor al Nacional	2	0,0109			
6				-0,0096	0,02054	5,32335
	Porcentaje abandono menor al nacional	1	-0,0054		0,00000	0,02466
	Porcentaje abandono mayor al nacional	2	0,0191			
7				-0,0054	0,02466	6,39156
	Área urbana	1	-0,0106		0,00000	0,02052
8	Área rural	2	0,0098			
	Acceso terrestre	1	-0,0012		0,00000	0,10675
	Acceso fluvial	2	0,0815			
9	Acceso aéreo	3	0,1054		0,08283	21,46551
	Agua buena	1	-0,0126		0,00000	0,10675
10				-0,0012	0,10675	27,66372
					0,02276	0,00000

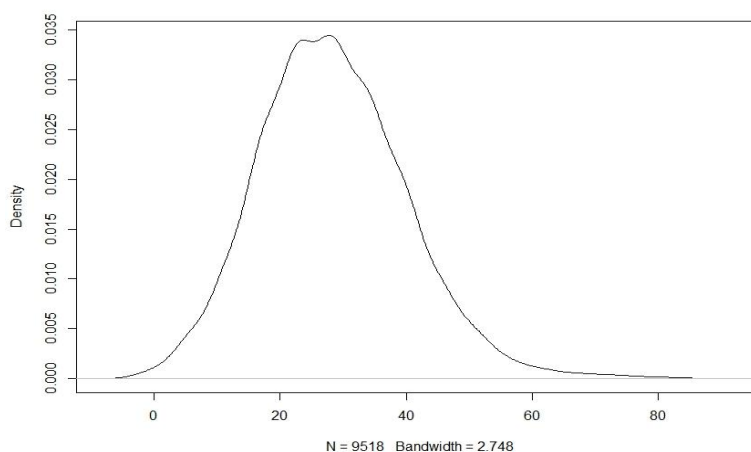
Agua Regular	2	0,0064		0,01910		4,95068
Agua Mala	3	0,0101	-0,0126	0,02276		5,89887
Lavamanos Bueno	1	-0,0073		0,00000	0,02620	0,00000
Lavamanos Regular	2	0,0129		0,02038		5,28147
11 Lavamanos Malo	3	0,0188	-0,0073	0,02620		6,79012
Inodoro Bueno	1	-0,0067		0,00000	0,02959	0,00000
Inodoro Regular	2	0,0143		0,02115		5,48205
12 Inodoro Malo	3	0,0228	-0,0067	0,02959		7,66814
Establecimiento educativo con Internet	1	-0,0104		0,00000	0,02050	0,00000
Establecimiento educativo sin Internet	2	0,0100	-0,0104	0,02050		5,31321
13 Internet	2					

4.4 Distribución de puntuaciones del Índice de Clasificación de Instituciones Educativas aplicando el Análisis de Componentes Principales No Lineales.

Una vez re escalados los valores en un rango de 0 a 100, donde los valores cercanos a 100 representan las peores condiciones y los cercanos a 0 las mejores, se obtiene la siguiente distribución de los puntajes obtenidos por las instituciones educativas a través del ACP No lineal (Figura 11).

Figura 11

Curva de distribución de puntuaciones ACP No Lineal



La distribución de los datos muestra una asimetría de 0,5155221 lo que indica que la distribución tiene una asimetría positiva, es decir que la cola derecha de la distribución es ligeramente más larga o tiene más valores extremos que la cola izquierda, por lo tanto, se cuenta con sesgo hacia las instituciones educativas con mayores niveles de carencias.

4.4.1 *Identificación de las Instituciones educativas que cuentan con condiciones menos favorables ACPNL.*

Las cuantificaciones ponderadas por el índice de clasificación sustituyen a los valores originales de cada categoría y se realiza una sumatoria de todas las cuantificaciones donde se pueden obtener valores entre 0 y 100, mientras más cercanos estén los valores a 100 significa que la institución educativa se encuentra en condiciones menos favorables, es decir que tendría mayores carencias, Por el contrario, valores cercanos a 0 indican mejores condiciones en las instituciones educativas.

Una vez obtenido el índice es factible realizar las clasificaciones conforme las necesidades lo determinen, es decir se puede obtener el 25% de instituciones con las condiciones más desfavorables dividiendo al índice en cuartiles o sí se requiere un segmento menor se podría dividir en deciles, es decir el 10% más desfavorables. En este caso se divide a las instituciones educativas en quintiles como se muestra en la Tabla 11, obteniéndose los siguientes resultados:

Tabla 11

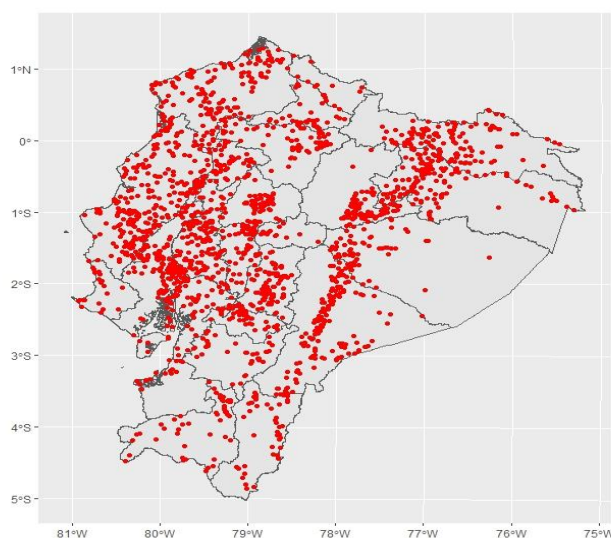
Distribución de instituciones educativas por quintiles, según puntuación del índice de clasificación

Quintil	Etiqueta	No. casos	Rango de puntuación
Q1	Instituciones educativas de mejores condiciones	1.905	0-17,61
Q2	Instituciones educativas de condiciones aceptables	1.921	17,62-23,59
Q3	Instituciones educativas de condiciones regulares	1.887	23,60-29,41
Q4	Instituciones educativas de condiciones insuficientes	1.909	29,42-38,70
Q5	Instituciones educativas deficitarias	1.896	38,71-100

Para visualizar la distribución geográfica de las instituciones educativas dentro del quintil 5, menos favorable, obtenidos a través de la aplicación del ACP No Lineal, se elaboró un mapa (Figura 12), como se puede observar, instituciones con estas características se encuentran en todo el país, con importantes concentraciones en provincias de la Costa y Amazonía.

Figura 12

Mapa de ubicación instituciones educativas menos favorecidas según ACP Categórico



En lo relacionado a la ubicación geográfica de la institución educativa, vemos que el índice obtenido a partir del ACPNL, es congruente con la teoría, de varios autores, UNESCO, Lalama, de que las instituciones con condiciones menos favorables se ubican en el área rural, el 76%, mientras que para el quintil Q1 de mejores condiciones el 26% se ubican en la ruralidad, esta información se detalla en la Tabla 12.

Tabla 12

Distribución de instituciones educativas por quintiles ACPNL, según puntuación del índice de clasificación por ubicación urbana - rural

Quintil	Instituciones ubicadas área urbana	Instituciones ubicadas área rural	Total instituciones
Q1	1414	491	1.905
Q2	1126	795	1.921
Q3	844	1043	1.887
Q4	733	1176	1.909
Q5	453	11443	1.896

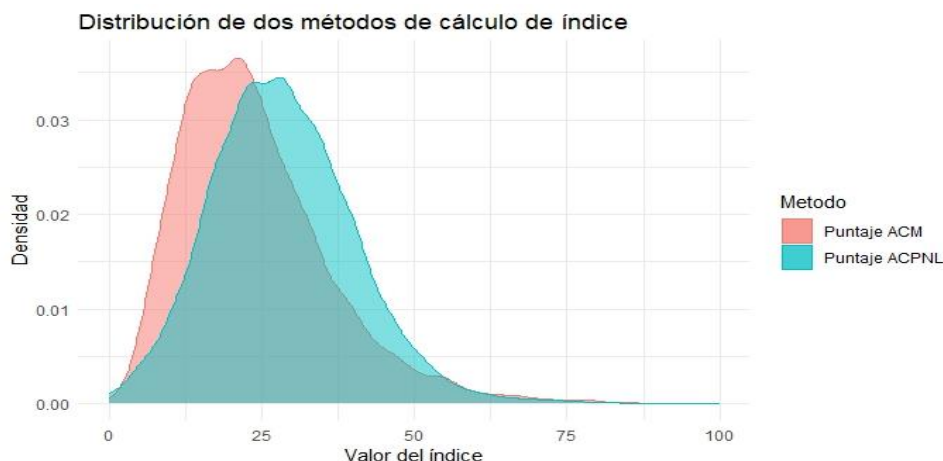
4.5. Comparación de resultados de la aplicación de las dos técnicas de clasificación.

Tanto en el Análisis de Correspondencias Múltiples (ACM) y el Análisis de Componentes Principales No Lineales (ACPNL), La densidad de los valores del índice está representada

en el eje vertical, mientras que los valores del índice se muestran en el eje horizontal (Figura 13).

Figura 13

Comparación de la distribución del Índice ACM vs ACPNL



La distribución del índice calculado mediante ACM tiene una forma asimétrica con un pico más pronunciado y está sesgada hacia la izquierda, en comparación con la distribución del índice obtenido a través del ACP no lineal. Esto indica que la mayoría de los valores del índice de instituciones educativas calculado con ACM tienden a concentrarse en un rango más estrecho de valores entre aproximadamente 10 y 35, con una menor dispersión.

La distribución del índice de instituciones educativas calculado mediante ACPNL es más amplia y tiene un pico menos pronunciado. Esto sugiere que los valores del índice tienen una mayor variabilidad, extendiéndose sobre un rango más amplio de valores que van desde valores bajos hasta aproximadamente 75.

Existe un solapamiento considerable entre las dos distribuciones en el rango de aproximadamente 10 a 35, lo que indica que en este intervalo ambos métodos producen índices comparables, sin embargo, fuera de este rango, especialmente a medida que los valores del índice aumentan, la distribución de ACPNL sigue extendiéndose, mientras que la de ACM disminuye rápidamente.

La mayor concentración y menor dispersión de los valores indican que el ACM podría estar generando un índice de clasificación de instituciones educativas más conservador, donde las observaciones tienden a agruparse en un rango limitado de valores, posiblemente debido a la forma en que el método maneja la variabilidad en los datos categóricos.

Para la curva de distribución del ACP No lineal la mayor dispersión sugiere que se está capturando más variabilidad en los datos, lo que puede ser indicativo de su capacidad para

manejar no linealidades y relaciones más complejas entre las variables categóricas, esto muestra en un índice más sensible a las diferencias existentes entre las instituciones educativas.

El gráfico sugiere que los dos métodos de cálculo de índice producen distribuciones con características diferentes, el ACM genera índices de instituciones educativas con menor variabilidad y valores más concentrados, mientras que el ACPNL produce índices de instituciones educativas con una mayor dispersión. Por lo tanto, si se busca un índice más estable y conservador, el ACM podría ser preferible; si se desea capturar una mayor variabilidad y complejidad en los datos, el ACPNL sería el método más adecuado.

- Tabla de Contingencia de comparación de resultados de los dos métodos aplicados por quintiles.

La tabla de contingencia que se muestra en la tabla 13, compara los resultados de la clasificación de observaciones en quintiles utilizando dos métodos, el Análisis de Correspondencias Múltiples (ACM) y el Análisis de Componentes Principales No Lineales (ACPNL), cada celda de la matriz muestra el número de observaciones que caen en la misma combinación de quintiles para ambos métodos

Tabla 13

Matriz de Contingencia quintiles obtenidos a partir del Análisis de Componentes Principales vs quintiles obtenidos a partir de Análisis de Componentes Principales No lineales

		Análisis Correspondencias Múltiples					
		Q1	Q2	Q3	Q4	Q5	Total
Análisis de Componentes Principales No lineales	Q1	1.086	533	236	46	4	1.905
	Q2	545	643	492	219	22	1.921
	Q3	212	478	575	479	143	1.887
	Q4	59	212	445	679	517	1.909
	Q5	2	37	156	483	1.218	1.896
Total		1.904	1.903	1.904	1.903	1.904	9.518

La diagonal principal de la matriz muestra las observaciones que fueron clasificadas en el mismo quintil por ambos métodos, estos valores son especialmente relevantes porque indican concordancia entre los dos métodos utilizados para clasificar a las instituciones educativas, además los valores más altos se ubican en esta diagonal lo que sugiere que,

aunque existen diferencias, hay una considerable concordancia entre los dos métodos, porcentualmente representa al 45%.

Por otra parte, la mayoría de las instituciones educativas en los quintiles más extremos (Q1 y Q5) coinciden entre ambos métodos, lo que puede indicar que las diferencias entre las categorías más bajas y altas son claras y evidentes en los datos, permitiendo una clasificación más consistente.

Las celdas fuera de la diagonal principal muestran el número de instituciones educativas que fueron clasificadas en diferentes quintiles por los dos métodos, esto representa la discordancia entre ACM y ACPNL en la clasificación de las observaciones, especialmente en los quintiles intermedios (Q2, Q3, Q4). Esto podría deberse a cómo cada método maneja las complejidades de los datos categóricos.

4.6. Similitudes entre el Análisis de Correspondencias Múltiples y El Análisis de Componentes Principales Categórico.

Tanto el ACM como ACPNL tienen como objetivo principal reducir la dimensionalidad de los datos categóricos, lo que permite una visualización más manejable y la identificación de patrones subyacentes. Ambos métodos buscan simplificar la estructura de los datos, representando las instituciones educativas y sus características en un espacio de menor dimensión.

Ambos métodos, ACM y ACPNL, son factibles para clasificar instituciones educativas a partir de datos categóricos, facilitando la toma de decisiones a partir de la identificación de las instituciones educativas más vulnerables.

Aunque abordan este objetivo desde enfoques distintos. Mientras que el Análisis de Correspondencias Múltiples (ACM) se centra exclusivamente en características cualitativas de las instituciones educativas, el Análisis de Componentes Principales No Lineal (ACPNL) extiende su aplicación a variables y características cualitativas con estructura tanto ordinal como nominal. Estas características permiten desarrollar y comparar ambas técnicas multivariadas, alineándose con el objetivo de construir un índice que clasifique efectivamente a las instituciones educativas de sostenimiento público. Además, tanto el ACM como el ACPNL ofrecen la ventaja de visualizar los resultados en un espacio reducido, lo que facilita la interpretación de las relaciones entre variables y observaciones. Las representaciones gráficas, propias de ambos métodos, son cruciales para comprender los patrones en los datos, permitiendo que los tomadores de decisiones en el ámbito educativo interpreten los resultados de manera intuitiva y fundamenten la priorización de recursos con base en clasificaciones claras y comprensibles.

4.7. Diferencias entre Análisis de Correspondencias Múltiple y Análisis de Componentes Principales No Lineales.

La principal diferencia es que el ACPNL toma en cuenta si las variables categóricas son ordinales o nominales, y esto es una gran ventaja, del ACPNL sobre el ACM, y esto incide en la obtención del índice, tal como lo menciona Atkinson, “el ACM se adapta mejor a múltiples variables nominales (variables categóricas sin un orden de clasificación obvio) y cuando se analizan otros tipos de variables (ordinales) puede ser preferible utilizar métodos alternativos que, no obstante, produzcan espacios multidimensionales susceptibles de datos geométricos”(Atkinson, 2024)

Otra diferencia importante entre ACM y ACPNL es cómo manejan y transforman los datos categóricos, El ACM no transforma las categorías de las variables; en cambio, se basa en la descomposición de la inercia de una tabla de contingencia para obtener las coordenadas de las categorías y las observaciones en un espacio reducido, el ACPNL, por otro lado, asigna puntuaciones numéricas a las categorías de las variables categóricas, transformándolas en variables cuantitativas antes de realizar el análisis de componentes principales. Esta transformación permite que ACPNL maneje tanto la no linealidad como la naturaleza categórica de los datos.

El método del ACP No Lineal a través de PRINCALS es más flexible que el ACM en términos del tipo de variables que puede manejar, Mientras que el ACM se aplica únicamente a variables categóricas, PRINCALS puede manejar tanto variables categóricas como ordinales y nominales, esta flexibilidad permite a PRINCALS ser aplicado en una gama más amplia de contextos donde las variables tienen diferentes niveles de medición. Esto en concordancia con lo mencionado por Demir y Keskin, “El Análisis de Componentes Principales Categóricos - No Lineal, es una técnica explicativa de reducción de dimensiones que proporciona datos numéricos y visuales para conjuntos de datos entre los cuales existe una relación lineal o no lineal y que incluyen variables continuas, categóricas y discretas” (Demir y Keskin, 2022).

La interpretación de los resultados también difiere entre los dos métodos, en el ACM, la proximidad entre puntos en el espacio reducido refleja la similitud en los perfiles de las categorías o las observaciones, en el ACPNL, la interpretación se basa en las puntuaciones asignadas a las categorías y cómo estas puntuaciones contribuyen a los componentes principales. La interpretación en el ACPNL puede ser más compleja debido a la necesidad de considerar las transformaciones aplicadas a las categorías.

El ACM asume que todas las variables categóricas tienen la misma importancia y no distingue entre diferentes tipos de variables categóricas, esto puede ser una limitación en casos donde algunas variables categóricas tienen una estructura ordinal o cuando se desea asignar diferentes pesos a las variables. El ACPNL, al permitir transformaciones de las variables, no tiene esta limitación y puede asignar diferentes pesos o manejar variables de diferentes tipos, lo que lo hace más flexible y potente.

4.8. Ventajas y Desventajas del ACM

- Ventajas Análisis de Correspondencias Múltiples

“El ACM está diseñado para analizar tablas disyuntivas completas, que son tablas de contingencia de variables cualitativas”, (Abril Donoso et al., 2022, p. 47) esto hace que sea relativamente fácil y sencillo de implementar e interpretar, lo que lo hace accesible para analistas sin una formación estadística avanzada. Su simplicidad permite su aplicación en la clasificación de instituciones educativas, con el fin de facilitar la toma de decisiones en la priorización de recursos educativos, sin requerir altos conocimientos técnicos.

La representación gráfica del ACM facilita la identificación de relaciones y patrones entre categorías y observaciones correspondientes a las instituciones educativas.

Es un método bien establecido y ampliamente aceptado en áreas como las ciencias sociales, lo que facilita la comparación de resultados con estudios anteriores.

- Desventajas Análisis de Correspondencias Múltiples

No diferencia entre tipos de variables categóricas (nominales vs ordinales), esto es una limitante importante para la construcción del índice de clasificación de instituciones educativas, ya que se requiere que el mismo refleje estas diferencias en importancia o estructura, lo que puede resultar en un índice menos preciso.

Comparado con métodos como ACPNL, el ACM es menos flexible en cuanto a la manipulación y transformación de los datos, lo que puede limitar su capacidad para reflejar con precisión las complejidades de las instituciones educativas.

4.9. Ventajas y desventajas del ACPNL

- Ventajas Análisis de Componentes Principales Categórico

El poder diferenciar y trabajar con datos categóricos ordinales es una gran fortaleza del ACP No Lineal, como menciona Atkinson “puede manejar todos y cada uno de los tipos de datos simultáneamente (datos continuos y categóricos) y, lo que es más importante, se puede especificar de antemano si las variables categóricas siguen una estructura nominal ordinal o simple/múltiple” (Atkinson, 2024, p. 833).

El ACP No Lineal utilizando PRINCALS permite asignar puntuaciones numéricas optimizadas a las categorías de las variables, lo que facilita la creación de un índice de condiciones de las instituciones educativas, que refleje de manera precisa la importancia relativa de cada variable, incluyendo la posibilidad de diferenciar entre variables nominales y ordinales.

Al maximizar la varianza explicada por los componentes principales transformados, el ACPNL permite crear un índice de clasificación de instituciones educativas que capture la mayor cantidad de información posible de los datos originales. Esto es especialmente útil en el contexto de la priorización de recursos educativos, ya que asegura que el índice construido represente de manera completa y detallada las características de cada institución, lo que facilita un análisis en función de sus necesidades específicas. Su capacidad para manejar tanto datos categóricos como ordinales y no lineales proporciona un enfoque flexible que puede adaptarse a la complejidad y diversidad de los datos educativos.

- Desventajas Análisis de Componentes Principales Categórico

Si no se maneja adecuadamente, el ACPNL podría generar un índice demasiado específico a las particularidades de ciertas instituciones educativas, lo que podría limitar su generalización a otros contextos o conjuntos de datos y, en consecuencia, afectar su utilidad en la toma de decisiones de recursos educativos a nivel fiscal. Además, este método requiere una mayor capacidad computacional que el ACM, lo cual podría ser un desafío en instituciones con limitaciones tecnológicas, especialmente al trabajar con grandes volúmenes de datos educativos.

4.10 Determinación la metodología que permitirá clasificar de manera adecuada a las instituciones educativas.

Para determinar la metodología es necesario conocer el tipo de datos que forman parte del proceso, en este caso se cuenta con datos categóricos y tanto el Análisis de Correspondencias Múltiples como el Análisis de Componentes Principales No lineales están diseñados para manejar datos este tipo de datos, sin embargo, el ACPNL tiene la ventaja de poder manejar tanto datos ordinales como nominales, y puede asignar puntuaciones numéricas a las categorías, esto le permite captar mejor la complejidad y las relaciones no lineales que pueden existir entre las variables, algo que el ACM no hace tan explícitamente.

Además, el ACPNL es más flexible en la reducción de la dimensionalidad y en la explicación de la varianza en los datos, lo que puede ser crucial cuando se busca capturar la mayor cantidad de información posible en el índice de clasificación de las instituciones educativas.

En lo relacionado a flexibilidad, el Análisis de Correspondencias Múltiples es menos flexible en términos de transformación de variables y manejo de no linealidades, además, trata a todas las variables categóricas con igual importancia, lo cual puede ser una limitante si se tiene una mezcla de variables con diferentes niveles de medición. En esto el ACPNL tiene una ventaja ya que es capaz de manejar diferentes tipos de variables ordinales o nominales y de ajustar el modelo a la estructura específica de los datos, esta flexibilidad permite la construcción de un índice que refleje mejor la importancia relativa de cada variable.

Si el objetivo es construir un índice que sea más estable y represente de manera más uniforme las carencias en la población, se recomienda el ACM, su menor dispersión indica un índice menos susceptible a variaciones pequeñas en los datos, lo que puede ser útil en aplicaciones donde la estabilidad del índice es crítica.

Si se busca un índice que capture mejor las diferencias sutiles entre las observaciones y que sea más sensible a la variabilidad en los datos, el ACPNL sería la mejor opción, ya que el método permite una representación más rica y diferenciada de las carencias, al capturar no solo las diferencias categóricas sino también la estructura subyacente más compleja.

Finalmente, factores como la etnia de estudiantes y docentes, y el número de estudiantes por docente son los factores sociales que más inciden en la clasificación de las instituciones educativas, esto probablemente está relacionado a que la población indígena y afroecuatoriana presentan tasas más altas de pobreza, en publicación realizada por la revista gestión el 26 de marzo del año 2023 se menciona “Según el Banco Mundial, la población indígena presenta tasas de pobreza que son en promedio dos veces más altas que para el resto de latinoamericanos. Esto se corrobora en el caso de Ecuador, puesto que las comunidades con mayor prevalencia de pobreza pertenecen a la población afroecuatoriana e indígena”.

<https://revistagestion.ec/analisis-sociedad/una-radiografia-de-la-pobreza-en-ecuador/>

Por otra parte, la presencia de servicios básicos como agua potable e internet en las instituciones educativas es crucial en su clasificación. Estas variables podrían reflejar el nivel de infraestructura de las escuelas, un aspecto esencial para su evaluación.

Aspectos relacionados a la vulnerabilidad de los estudiantes también es un factor a considerar. El abandono escolar representa un desafío crítico, al señalar posibles deficiencias en el entorno académico, social o económico que impiden la permanencia de los estudiantes, impactando negativamente en el desarrollo humano y en la calidad percibida de la institución. Por otro lado, la capacidad de las instituciones para atender a estudiantes con discapacidad (representada por "disc") es un reflejo directo de su nivel de accesibilidad, adaptabilidad y compromiso con la diversidad.

CONCLUSIONES

El desarrollo y la selección de técnicas de clasificación multivariada, como el Análisis de Componentes Principales No Lineal (ACP No Lineal) y el Análisis de Correspondencias Múltiples (ACM), efectivamente facilitan la toma de decisiones en la priorización de recursos educativos a nivel fiscal. Estas metodologías permiten identificar patrones complejos en los datos y categorizar instituciones educativas en función de múltiples dimensiones, como acceso a infraestructura o inclusión social. Al reducir la complejidad de los datos en dimensiones más interpretables, estas técnicas brindan una herramienta poderosa para visualizar disparidades y focalizar recursos en instituciones con mayores necesidades. Además, al priorizar de manera objetiva, basada en evidencia estadística, se promueve una distribución más equitativa y eficiente de los recursos, alineada con metas de inclusión y calidad educativa.

Esto demuestra que la aplicación de métodos multivariados no solo mejora la comprensión de las diferencias entre instituciones, sino que también optimiza las estrategias de inversión y planificación fiscal.

El ACM tiende a generar un índice de clasificación de las instituciones educativas con menor variabilidad y valores más concentrados, mientras que el ACPNL produce valores del índice con una mayor dispersión. Por lo tanto, si se busca un índice más estable y conservador, el ACM podría ser la opción más adecuada; si se desea capturar una mayor variabilidad y complejidad en los datos, el ACPNL es el método que más se ajusta a estas exigencias.

Las variables etnia estudiantes y etnia docentes destacan en su contribución con la Dimensión 2 del ACP no lineal siendo muy importantes en la priorización y diferenciación de las instituciones educativas, probablemente reflejando características sociodemográficas importantes. En esta misma línea en la Dimensión D1, la variable ratio estudiantes por docente (r_esdo), agua e internet son las de mayor contribución. Esto sugiere que la relación estudiante/docente y servicios básicos como agua e internet, son factores determinantes asociados a las instituciones educativas de sostenimiento público para diferenciar y priorizar. El Análisis de Componentes Principales No Lineales es la mejor opción para la construcción del índice ya que capturar una mayor variabilidad, maneja diferentes tipos de variables categóricas (ordinales y nominales), y refleja de manera más precisa la complejidad subyacente en los datos, su capacidad para manejar no linealidades y transformar los datos en puntuaciones numéricas permite la creación de un índice más robusto y adaptativo.

RECOMENDACIONES

El trabajo realizado ha cumplido satisfactoriamente con los objetivos planteados y ha desarrollado un índice de clasificación efectivo para las instituciones educativas de sostenimiento público. No obstante, para asegurar la implementación efectiva y la validez de la metodología propuesta, se sugiere considerar las siguientes recomendaciones:

Garantizar la completitud de la información para todas las instituciones Educativas.

Es fundamental que la información recopilada sea completa y homogénea para todas las instituciones educativas incluidas en el análisis. La calidad y la fiabilidad del índice de clasificación dependen directamente de la integridad de los datos. Por lo tanto, se recomienda establecer mecanismos para asegurar la recolección de datos de todas las instituciones educativas, con especial atención a aquellas que se encuentran en zonas remotas o que podrían tener limitaciones en el acceso a la información.

Verificación in situ de la metodología desarrollada.

Para validar la aplicabilidad y precisión del índice, es crucial que la metodología sea comprobada en el lugar de implementación. Se sugiere realizar visitas a una muestra representativa de las instituciones educativas clasificadas para evaluar si los resultados obtenidos reflejan de manera precisa la realidad de cada institución. Esto permitirá ajustar la metodología y los criterios de clasificación en función de las observaciones realizadas en el campo.

Actualización periódica del Índice de Clasificación.

Las condiciones y factores que afectan a las instituciones educativas pueden cambiar con el tiempo. Por lo tanto, se recomienda que el índice de clasificación sea revisado y actualizado periódicamente para asegurar que refleje las condiciones actuales de las instituciones.

Incorporación de nuevos factores asociados a la clasificación.

A medida que se implemente el índice, es posible que se identifiquen factores adicionales que influyan en el desempeño y las necesidades de las instituciones educativas. Se recomienda mantener un enfoque flexible y estar dispuesto a incorporar nuevas variables que puedan mejorar la precisión y efectividad del índice.

REFERENCIAS

- Abeyasekera, S. (n.d.). Multivariate methods for index construction. *Household Surveys in Developing and Transition Countries: Design, Implementation and Analysis*, 1–21. <http://unstats.un.org/unsd/hhsurveys/>
- Abril Donoso, M. E., Chariguamán Maurisaca, N. E., y Aguilar Reyes, J. E. (2022). Análisis de Correspondencias Múltiples para el Estudio de los Homicidios Intencionales en el Ecuador. *Revista Politécnica*, 50(3), 43–52. <https://doi.org/10.33333/rp.vol50n3.04>
- AFD. (2018). Manual Operativo programa AFD. 1–129.
- Atkinson, W. (2024). Charting fields and spaces quantitatively: from multiple correspondence analysis to categorical principal components analysis. *Quality and Quantity*, 58(1), 829–848. <https://doi.org/10.1007/s11135-023-01669-w>
- Banco Mundial. (2024). Government expenditure on education, total (% of GDP) [Base de datos]. <https://datos.bancomundial.org/indicador/SE.XPD.TOTL.GD.ZS>
- Blanco, G. (2006). La equidad y la inclusión social: Uno de los desafíos de la escuela. *La Equidad y La Inclusión Social: Uno de Los Desafíos de La Educación y La Escuela Hoy*, 4(3), 1–15. <https://www.redalyc.org/pdf/551/55140302.pdf>
- Bucherie, A., Hultquist, C., Adamo, S., Neely, C., Ayala, F., Bazo, J., y Kruczkiewicz, A. (2022). A comparison of social vulnerability indices specific to flooding in Ecuador: principal component analysis (PCA) and expert knowledge. *International Journal of Disaster Risk Reduction*, 73(October 2021), 102897. <https://doi.org/10.1016/j.ijdrr.2022.102897>
- CEPAL. (2024). Prevención y reducción del abandono escolar en América Latina y el Caribe. 2022, 1–22. <https://www.cepal.org/es/publicaciones/68814-prevencion-reduccion-abandono-escolar-america-latina-caribe>
- Chazarreta, A. (2022). Aproximación Empírica a La Separación De La Propiedad Y El Control Del Capital. *Construcción De Un Índice De Las Estructuras Organizativas De Las Empresas Industriales*. 2012 1. 14(2), 195–213.
- Demir, C., y Keskin, S. (2022). Introduction of Nonlinear Principal Component Analysis with an Application in Health Science Data. *Eastern Journal of Medicine*, 27(3), 394–402. <https://doi.org/10.5505/ejm.2022.09068>

- Duque, I., Domínguez-berjón, M. F., Cebrecos, A., Prieto-salceda, M. D., y Española, S. (2021). Índice de privación en España por sección censal. 35(2), 113–122.
- UNESCO, y BANK, W. (2024). EDUCATION FINANCE WATCH 2024 Introduction.
- Falus, L., y Goldberg, M. (2011). Perfil de los docentes en América Latina. Siteal, 59. <https://unesdoc.unesco.org/ark:/48223/pf0000370958>
- Gil, I. (2023). El Grupo Banco Mundial en Ecuador y el mito de la educación como mejor estrategia para salir de la pobreza The World Bank Group in Ecuador and the Myth of Education as the Best Strategy to Escape Poverty. 42(esp 2), 17–31.
- Gómez S., C. A. (2016). Schools classification according ICFES Saber 11 tests: An analysis using marginal models (MM). Sociedad y Economía, 30, 69–89.
- González-Luna Corvera, T. (2019). La educación es un derecho, no es un privilegio. Diálogos Sobre Educación, 19, 1–15. <https://doi.org/10.32870/dse.v0i19.611>
- González-Mayorga, H., y Rodríguez Esteban, A. (2023). Autoeficacia en la gestión del aula en el profesorado de primaria y secundaria: variables predictoras y perfiles docentes. Aula Abierta, 52(1), 71–80. <https://doi.org/10.17811/rifie.52.1.2023.71-80>
- Joaquín, A., y Uriel, E. (2017). Análisis Multivariante aplicado con R. In 2017 (Vol. 5, Issue 1). <https://revistas.ufrj.br/index.php/rce/article/download/1659/1508%0Ahttp://hipatiapress.com/hpjournal/index.php/qre/article/view/1348%5Cnhttp://www.tandfonline.com/doi/abs/10.1080/09500799708666915%5Cnhttps://mckinseysociety.com/downloads/reports/Educ>
a
- Lamfre Laura, Perren Joaquín, B. S. (2023). Análisis de Correspondencias Múltiple Condicionada. Una Aplicación al Estudio de la calidad de vida según clase social en la Argentina de Inicios del milenio. SaberES, 15, 157–175.
- Larrea Maldonado, A. M. (2024). Fracturas que nos marcan. Las desigualdades en la educación en Ecuador. In Fracturas que nos marcan. Las desigualdades en la educación en Ecuador. <https://doi.org/10.46546/2024-55atrio>
- Laura, A., y Tessore, M. (2021). Brechas digitales y derecho a la educación durante la pandemia por COVID-19. Propuesta Educativa, 2(56), 11–27. <http://www.scielo.org.ar/pdf/pe/n56/1995-7785-PE-56-00011.pdf>

- Linser, S., y Vienna, L. S. (1999). the assessment of sustainable development . THEORETICAL BACKGROUND OF INDICATORS AND THEORETICAL APPROACHES FOR THE ASSESSMENT OF SUSTAINABILITY. September.
- Ministerio de Educación. (2022). Manual Operativo “PARECF.”
- Molina, M. (2019). Revista electrónica de AnestesiaR. Revista Electronica de AnestesiaR, 11(6), 1–6.
- Montenegro, E. T. (2015). El capital humano y los retornos a la educación en Ecuador.
- Morales, K., Flores, M., & Salazar Méndez, Y. (2021). Análisis de componentes principales no lineales para la construcción de un índice de estratificación socioeconómica para el Ecuador. *Desarrollo y Sociedad*, (88), 43-82. <https://doi.org/10.13043/DYS.88.2>
- Mori, Y., Kuroda, M., y Makino, N. (2017). Nonlinear Principal Component Analysis and Its Applications. https://doi.org/10.1007/978-981-10-0159-8_8
- Moscolini Nora. (2011). LAS NUBES DE DATOS.
- Naranjo Bonilla, M. (2008). Ecuador: reseña de los principales programas sociales y lecciones aprendidas, 2000-2006.
- Nunes, A. O., da Silva, T. E. V., Mota, J. C. M., y de Almeida, A. L. F. (2015). Validation of the academic management evaluation instrument based on principal component analysis for engineering and technological courses | Validación de un instrumento de evaluación de la gestión académica basado en el análisis de componentes principal. *Ingeniería e Investigación*, 35(2), 96–102. <https://doi.org/10.15446/ing.investig.v35n2.47369>
- Piderit, B., González, F., y Bos, M. S. (2024). Agua y saneamiento en escuelas Planificación y diseño de instalaciones sanitarias.
- Ramoni P., J., Orlandoni M., G., Saavedra, S., y Rivas W., A. (2013). Evaluación de impactos de proyectos de inversión pública: importancia de la participación de las comunidades. *Revista Venezolana de Sociología y Antropología*, 23, 111–123. <http://www.redalyc.org/articulo.oa?id=70538668007>
- Schettino Piña, A. (2017). El Derecho a La Educación. *Revista de La Facultad de Derecho de México*, 54(241), 243. <https://doi.org/10.22201/fder.24488933e.2004.241.61597>
- UNESCO. (2006). Directrices de la UNESCO sobre educación intercultural.

- UNESCO. (2014). Enfoques Estratégicos sobre las TICS En Educación. In Enfoque Estratégico Sobre Tics En Educación En América Latina Y El Caribe. <https://unesdoc.unesco.org/ark:/48223/pf0000223251>
- UNESCO. (2021). Llegando a todos los estudiantes. 0–11. <https://doi.org/10.2760/46933>
- UNESCO. (2023). Informe mundial sobre el personal docente Afrontar la escasez de docentes PUNTOS CLAVE. 1–36. https://unesdoc.unesco.org/ark:/48223/pf0000387400_spa
- UNESCO Institute for Statistics. (2024). Government expenditure on education [Base de datos]. <https://databrowser.uis.unesco.org/view#indicatorPaths=UIS-SDG4Monitoring%3A0%3AROFST.MOD.1%2CUIS-SDG4Monitoring%3A0%3AROFST.MOD.2%2CUIS-EducationOPRI%3A0%3AX.US.FSGOV&geoMode=countries&geoUnits=&browsePath=EDUCATION%2FUIS-EducationOPRI%2Fgov-exp&timeMode=range&view=table&chartMode=multiple&tableIndicatorId=X.US.FSGOV&chartIndicatorId=ROFST.MOD.1&chartHighlightSeries=&chartHighlightEnabled=true>
- Vélez-Tamay, A., Carriel, V., & Castillo-Ortega, Y. (2021). Índice de desarrollo local para Ecuador con datos del 2010. *Desarrollo y Sociedad*, (88), 83-127. <https://doi.org/10.13043/DYS.88.3>

ANEXOS

Anexo A. Certificado del abstract por parte de idiomas



UNIVERSIDAD POLITÉCNICA ESTATAL DEL CARCHI FOREIGN AND
NATIVE LANGUAGES CENTER

ABSTRACT- EVALUATION SHEET				
<p>NAME: Jhony Leandro Zabala Celi, DATE: Martes, 15 de abril de 2025 Topic: "Revisión de métodos y propuesta de metodología para la obtención de un índice de clasificación para la priorización de instituciones educativas de sostenimiento público utilizando métodos multivariados"</p>				
MARKS AWARDED		QUANTITATIVE AND QUALITATIVE		
VOCABULARY AND WORD USE	Use new learnt vocabulary and precise words related to the topic	Use a little new vocabulary and some appropriate words related to the topic	Use basic vocabulary and simplistic words related to the topic	Limited vocabulary and inadequate words related to the topic
	EXCELLENT: 2 <input type="checkbox"/>	GOOD: 1,5 <input checked="" type="checkbox"/>	AVERAGE: 1 <input type="checkbox"/>	LIMITED: 0,5 <input type="checkbox"/>
WRITING COHESION	Clear and logical progression of ideas and supporting paragraphs.	Adequate progression of ideas and supporting paragraphs.	Some progression of ideas and supporting paragraphs.	Inadequate ideas and supporting paragraphs.
	EXCELLENT: 2 <input checked="" type="checkbox"/>	GOOD: 1,5 <input type="checkbox"/>	AVERAGE: 1 <input type="checkbox"/>	LIMITED: 0,5 <input type="checkbox"/>
ARGUMENT	The message has been communicated very well and identify the type of text	The message has been communicated appropriately and identify the type of text	Some of the message has been communicated and the type of text is little confusing	The message hasn't been communicated and the type of text is inadequate
	EXCELLENT: 2 <input checked="" type="checkbox"/>	GOOD: 1,5 <input type="checkbox"/>	AVERAGE: 1 <input type="checkbox"/>	LIMITED: 0,5 <input type="checkbox"/>
CREATIVITY	Outstanding flow of ideas and events	Good flow of ideas and events	Average flow of ideas and events	Poor flow of ideas and events
	EXCELLENT: 2 <input type="checkbox"/>	GOOD: 1,5 <input type="checkbox"/>	AVERAGE: 1 <input type="checkbox"/>	LIMITED: 0,5 <input type="checkbox"/>
SCIENTIFIC SUSTAINABILITY	Reasonable, specific and supportable opinion or thesis statement	Minor errors when supporting the thesis statement	Some errors when supporting the thesis statement	Lots of errors when supporting the thesis statement
	EXCELLENT: 2 <input checked="" type="checkbox"/>	GOOD: 1,5 <input type="checkbox"/>	AVERAGE: 1 <input type="checkbox"/>	LIMITED: 0,5 <input type="checkbox"/>
<p>TOTAL/AVERAGE</p> <p>9 - 10: EXCELLENT 7 - 8,9: GOOD 5 - 6,9: AVERAGE 0 - 4,9: LIMITED</p>		<p>TOTAL 9</p>		



**UNIVERSIDAD POLITÉCNICA ESTATAL DEL
CARCHI- FOREIGN AND NATIVE LANGUAGES
CENTER**

**Informe sobre el Abstract de Artículo Científico o
Investigación.**

Autor: Jhony Leandro Zabala Celi

Fecha de recepción del abstract: Viernes, 11 de abril de 2025

Fecha de entrega del informe: Martes, 15 de abril de 2025

El presente informe validará la traducción del idioma español al inglés si alcanza un porcentaje de: 9 – 10 Excelente.

Si la traducción no está dentro de los parámetros de 9 – 10, el autor deberá realizar las observaciones presentadas en el ABSTRACT, para su posterior presentación y aprobación.

Observaciones:

Después de realizar la revisión del presente abstract, éste presenta una apropiada traducción sobre el tema planteado en el idioma Inglés. Según la rúbrica de evaluación de la traducción en Inglés, ésta alcanza un valor de 9; por lo cual se valida dicho trabajo.

Atentamente



MA. Martha Viveros
Docente responsable del
CIDEN

Anexo B. Tratamiento de variables previo al ACM y ACPNL **Tratamiento de Variables**

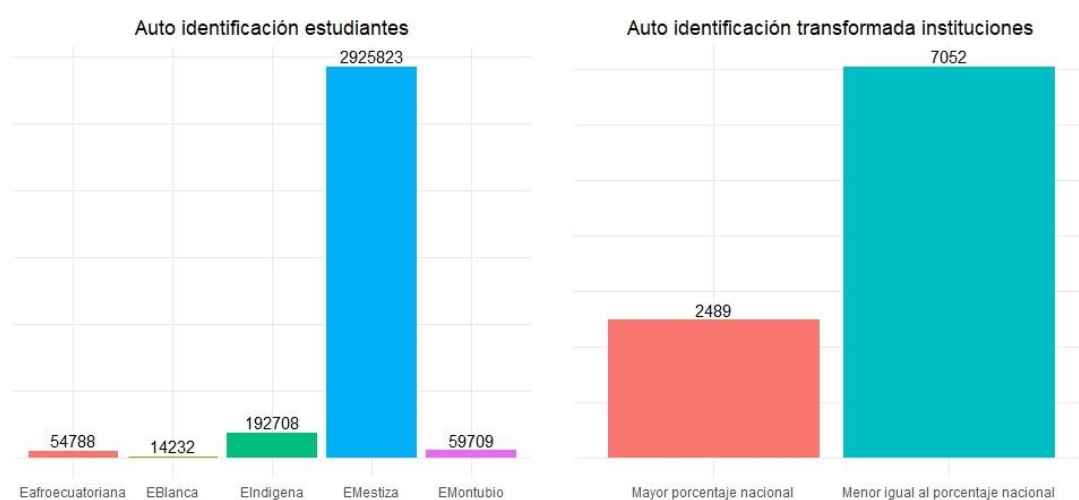
A continuación, se describen las variables que serán utilizadas en el análisis de componentes principales categóricos y el proceso realizado para su transformación previa al algoritmo PRINCALS

1. Variable Auto identificación étnica estudiantes.

En el período académico 2022-2023 Inicio, se registró una matrícula de 3.247.260 estudiantes, de los cuales 192,708 se consideran indígenas, 54.788 se auto identifican como afroecuatorianos y 59.709 como montubios, a estos tres tipos de auto identificación se les conforma en un solo grupo y representan a nivel nacional el 9,46% del total de estudiantes. La transformación consiste en identificar a todas aquellas instituciones educativas que cuentan con un valor porcentual de estudiantes identificados como indígenas, afroecuatorianos o montubios menores o iguales al porcentaje nacional que es 9,46% y se les asigna el valor de 1 y las instituciones que superan este valor se le asigna un valor de 2, entonces se obtiene una nueva variable denominada etnia, conformada por las instituciones educativas que cuentan con estudiantes identificados como indígenas, afroecuatorianos o montubios menores o iguales al porcentaje nacional y en complemento aquellas que cuentan con porcentajes mayores,

Figura 14

Transformación variable Auto identificación de estudiantes



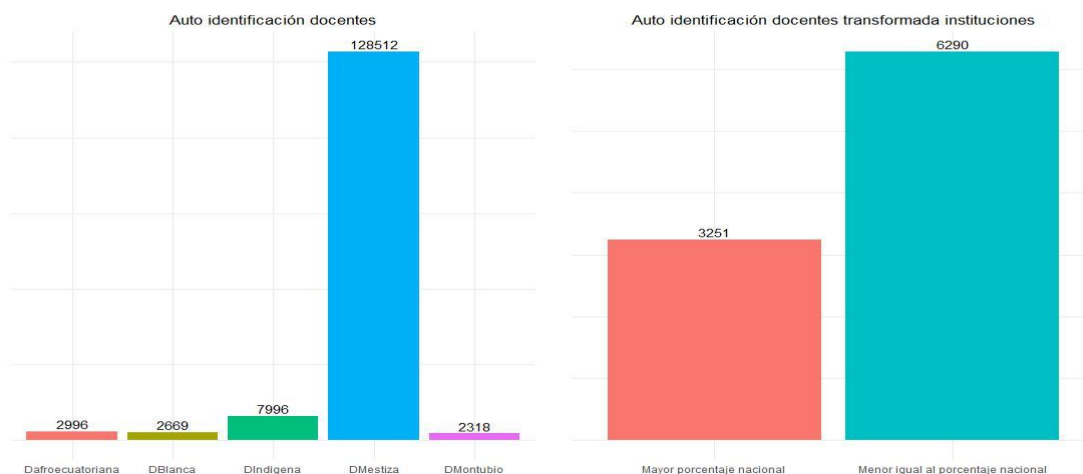
2. Variable Auto identificación étnica docentes.

En el periodo académico 2022-2023 Inicio, se registró 144.491 docentes, de los cuales 7.996 se consideran indígenas, 2.996 se auto identifican como afroecuatorianos y 2.318 como montubios, a estos tres tipos de autoidentificación se les conforma en un solo grupo y representan a nivel nacional el 9,2% del total de docentes.

La transformación consiste en identificar a todas aquellas instituciones educativas que cuentan con un valor porcentual de docentes identificados como indígenas, afroecuatorianos o montubios menores o iguales al porcentaje nacional que es 9,46% y se les asigna el valor de 1 y las instituciones que superan este valor se le asigna un valor de 2, entonces se obtiene una nueva variable denominada etniad, conformada por las instituciones educativas que cuentan con docentes identificados como indígenas, afroecuatorianos o montubios menores o iguales al porcentaje nacional y en complemento aquellas que cuentan con porcentajes mayores.

Figura 15

Transformación variable Autoidentificación de estudiantes



3. Variable discapacidad

Para el periodo académico de análisis se tiene que 23.962 estudiantes reportaron tener algún tipo de discapacidad, esto representa el 0,74%.

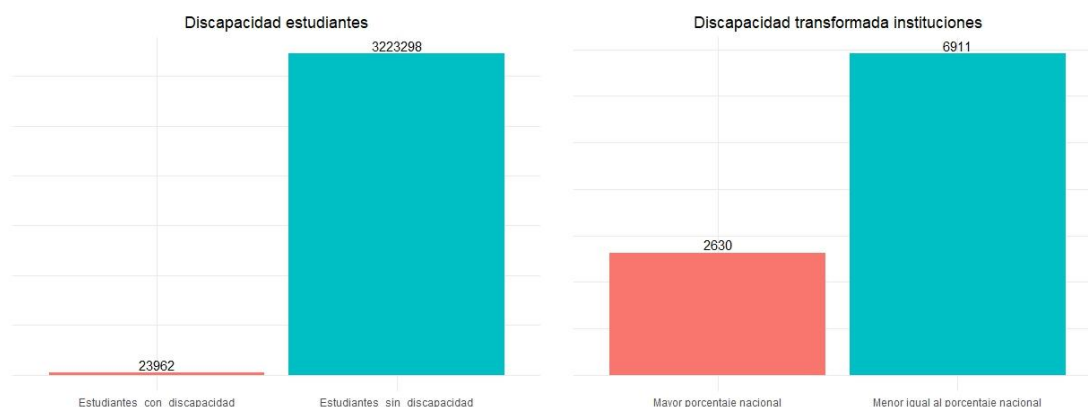
La transformación de esta variable consiste en identificar las instituciones educativas que registran valores porcentuales menores o iguales al porcentaje nacional y asignarles el valor

de 1 y a las que superan este porcentaje asignarles un valor de 2, Obteniendo una nueva variable con dos categorías.

Es necesario acotar que esta información corresponde a lo informado por el representante del estudiante y no necesariamente significa que cuente con un carné de discapacidad.

Figura 16

Transformación variable Discapacidad de estudiantes



4. Variables: Número de estudiantes, Número de docentes

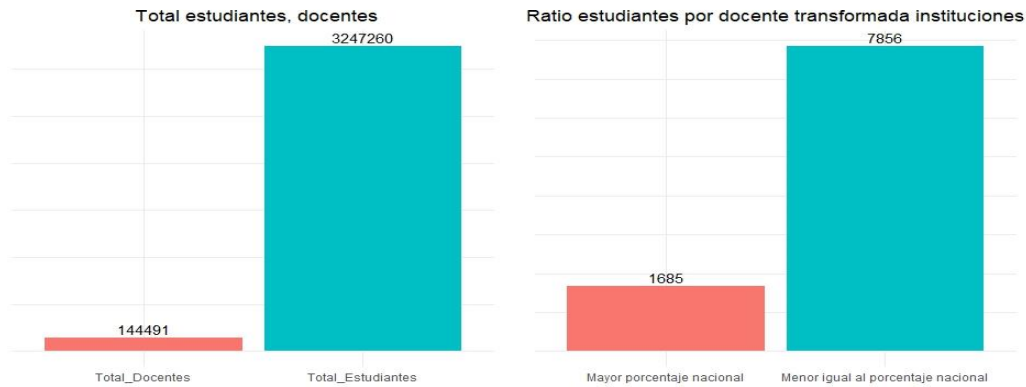
Las variables número de estudiantes y número de docentes se utilizan para calcular el ratio de estudiantes por docente, y resulta de dividir, el total de estudiantes entre el total de docentes.

El número total de estudiantes es 3.247.260 y el número total de docentes es 144.491 por lo que se obtiene un ratio de estudiantes por docente igual a 22,47 a nivel nacional.

De la misma manera, se clasifican las instituciones educativas en aquellas que cuentan con ratio de estudiantes por docente menores o iguales a 22,47, asignándoles el valor de 1 y las que superan ese valor se les categoriza como 2.

Figura 17

Trasformación variables total estudiantes, total docentes en ratio estudiantes por docente



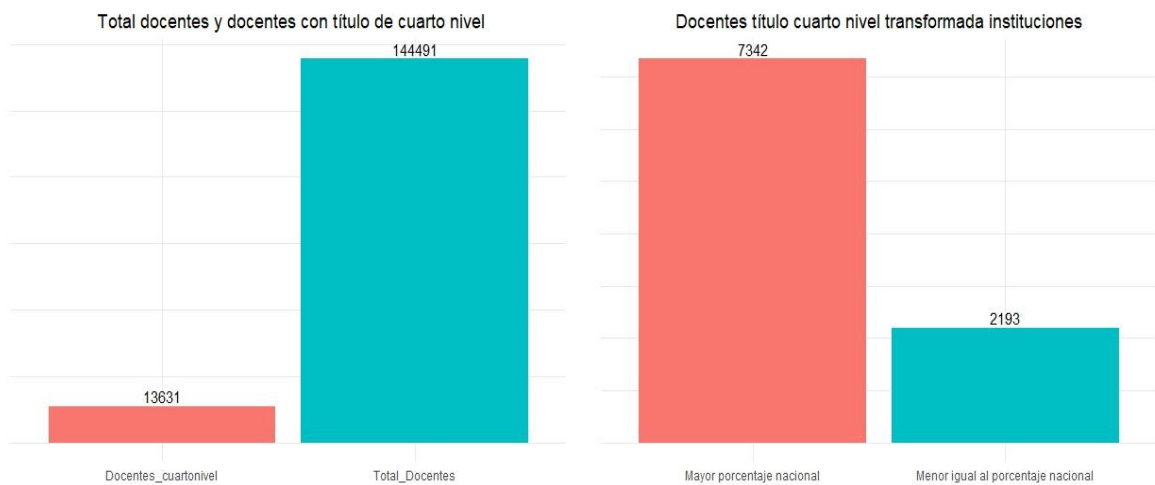
5. Variable docentes con título de cuarto nivel

En esta variable se almacena el número de docentes que cuenta con título de cuarto nivel por cada institución educativa, siendo 13.361 docentes, esto implica que el 9,43% poseen este nivel de educación.

Como se ha realizado en casos anteriores, se procede a crear una nueva variable que muestre a las instituciones educativas que cuentan con porcentajes mayores al porcentaje nacional de profesores con título de cuarto nivel y se les asigna un valor de uno, aquellas instituciones que cuentan con valores menores se les asigna la categoría 2.

Figura 18

Trasformación variable docentes con título cuarto nivel



6. Variable nombramiento docentes

En esta variable se almacena el número de docentes que la relación laboral con la institución educativa es a través de nombramiento definitivo, se calcula el porcentaje de docentes que cuentan con esta condición en cada institución educativa, se les compara con el porcentaje nacional, si el valor de la institución es mayor o igual al porcentaje nacional se asigna un valor de 1, si el porcentaje de docentes con nombramiento en la institución educativa es menor al nacional se asigna el valor de 2.

El porcentaje de docentes con nombramiento en las instituciones educativas públicas a nivel nacional es igual a 70,23%.

7. Variable Agua

Para la construcción de la variable agua, se considera la variable acceso, con las siguientes categorías: entubada, pozo, red pública, río, tanquero, vertiente, así como la variable estado con las categorías: Bueno, regular, malo y desconoce

En la transformación de la variable se ha realizado lo siguiente:

Tabla 14

Transformación variable agua

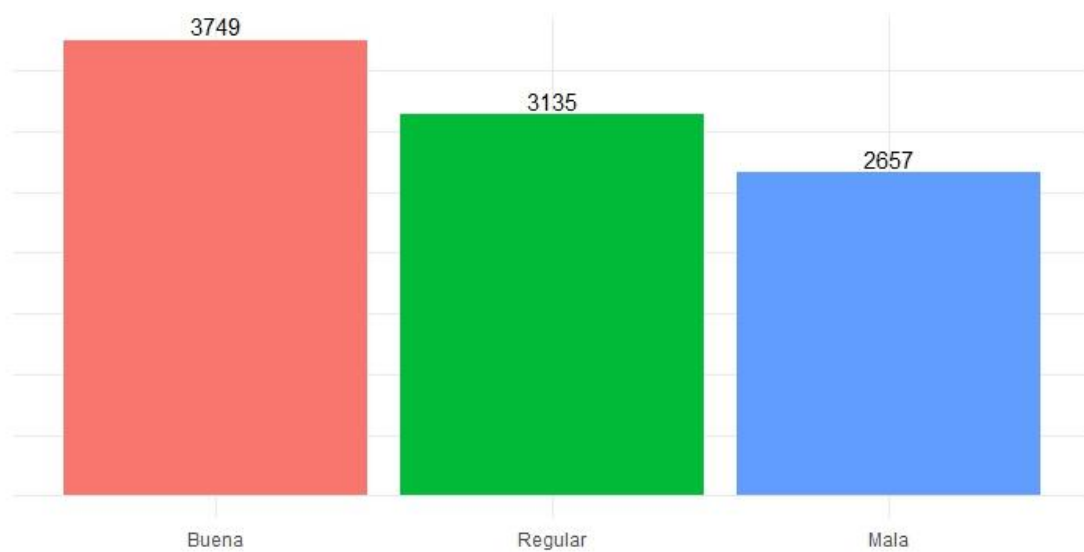
Variable Acceso	Variable Estado	Categoría	variable transformada
Red pública	Bueno	1	
	Regular	1	
	Malo	2	
	Desconoce	2	
Entubada	Bueno	2	
	Regular	2	
	Malo	3	
	Desconoce	3	
Pozo	Bueno	3	
	Regular	3	
	Malo	3	
	Desconoce	3	
Río	Bueno	3	

Tanquero	Regular	3
	Malo	3
	Desconoce	3
	Bueno	3
	Regular	3
	Malo	3
Vertiente	Desconoce	3
	Bueno	3
	Regular	3
	Malo	3
	Desconoce	3

De esta manera se obtiene la siguiente distribución por institución educativa:

Figura 19

Transformación variable Agua



8. Variable lavamanos

Esta variable cuenta directamente con el estado del lavamanos, esto es: Bueno, regular, malo y desconoce

Tabla 15*Transformación variable lavamanos*

Variable	Variable Estado	Categoría	variable transformada
Lavamanos	Bueno	1	
	Regular	2	
	Malo	3	
	Desconoce	3	

Con la variable transformada se obtiene la siguiente distribución:

Figura 20*Transformación variable Lavamanos*

9. Variable inodoro

De esta variable también se tiene directamente el estado: Bueno, Regular, Malo, Desconoce

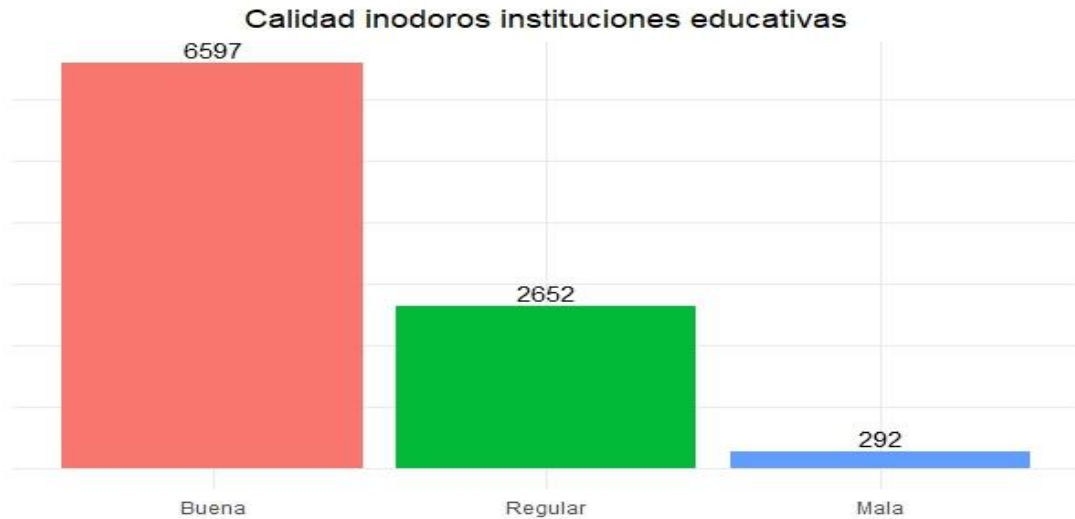
Tabla 16*Transformación variable Inodoro*

Variable	Variable Estado	Categoría	variable transformada
Inodoro	Bueno	1	
	Regular	2	
	Malo	3	
	Desconoce	3	

La variable obtiene la siguiente distribución:

Figura 21

Trasformación variable Inodoro

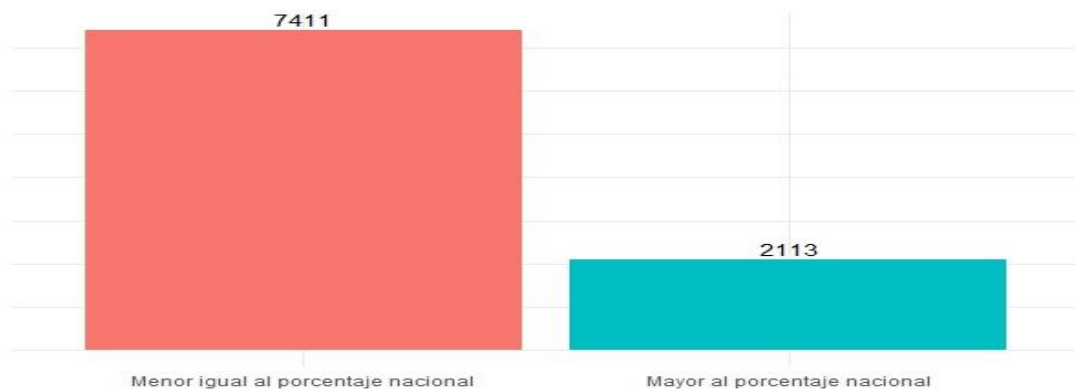


10. Variable Abandono

La variable abandono contiene el porcentaje de abandono que ha ocurrido en el periodo 2022-2023. El Ministerio de educación pública en su portal web entre otros el porcentaje de abandono, que para el período de análisis se ubicó en el 1,88%, este valor es tomado como referencia para comparar a las instituciones que cuentan con un valor menor o igual al 1,88% se les asigna con un valor de 1, mientras que aquellas instituciones que cuentan con valores superiores al porcentaje nacional se les categoriza como 2.

Figura 22

Trasformación variable Abandono



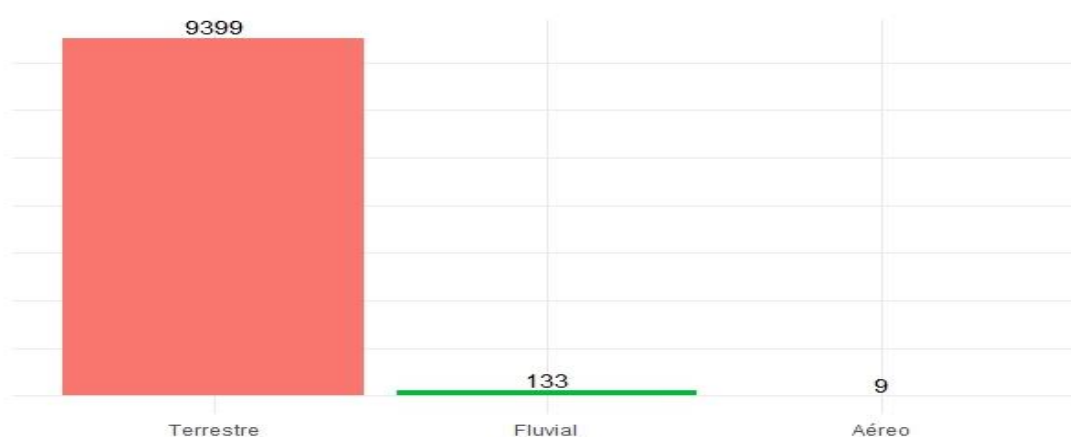
11. Variable Acceso

Esta variable indica el tipo de acceso que se tiene para llegar a la institución educativa, siendo este de tres tipos: Terrestre, cuando se accede por medio de carretera, camino, sendero, fluvial, cuando para acceder a la institución educativo en algún tramo es obligatorio navegar por río o laguna, brazo de mar, aéreo, cuando la única forma de ingresar a la institución educativa es de manera aérea, especialmente en zonas muy remotas de la Amazonía.

Esta es una variable categórica cuya distribución se muestra a continuación:

Figura 23

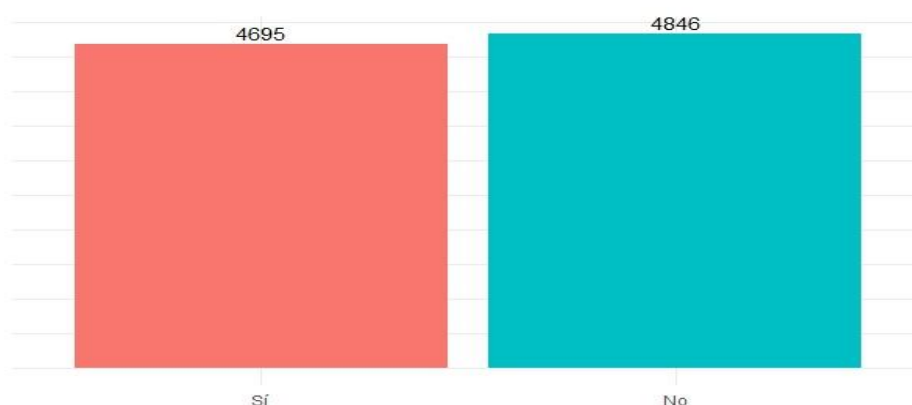
Trasformación variable Acceso a la Institución Educativa



12. Variable Internet

Esta es una variable categórica que cuenta con las categorías Sí cuando la institución educativa cuenta con internet y No cuando todavía no tiene este servicio,

La distribución de esta variable se muestra a continuación:

Figura 24*Transformación variable Internet*

Cuadro Resumen variables categóricas ACP Categórico

De esta manera se obtiene una matriz transformada con todas las variables categóricas, como se muestra en la siguiente tabla:

Tabla 17*Resumen Variables categóricas transformadas*

Variable	Tipo de variable	Categorías de respuesta	código
Codigo_Institución	Texto	NA	NA
Etnia	cualitativa	Porcentaje menor al Nacional	1
		Porcentaje mayor al Nacional	2
Discapacidad	Cualitativa	Porcentaje menor al Nacional	1
		Porcentaje mayor al Nacional	2
Ratio estudiantes por docente	Cualitativa	Ratio de estudiantes por docente menor al nacional	1
		Ratio de estudiantes por docente mayor al nacional	2
Docentes con título de cuarto nivel	Cualitativa	Porcentaje mayor al Nacional	1
		Porcentaje menor al Nacional	2
Abandono	Cualitativa	Porcentaje menor al Nacional	1
		Porcentaje mayor al Nacional	2
Área	Cualitativa	Urbana	1

		Rural	2
Acceso a la institución educativa	Cualitativa	Terrestre	1
		Fluvial	2
		Aéreo	3
Agua	Cualitativa	Buena	1
		Regular	2
		Mala	3
Lavamanos	Cualitativa	Buena	1
		Regular	2
		Mala	3
Inodoro	Cualitativa	Buena	1
		Regular	2
		Mala	3
Internet	Cualitativa	Sí	1
		No	2

Anexo C. Análisis de Correspondencia simple paso a paso

Para entender cómo se desarrollan los 7 pasos propuestos se realiza un ejercicio considerando de la base de instituciones educativas de sostenimiento fiscal una tabla de contingencia producto del cruce entre la calificación del agua y la región del país donde se encuentra ubicada la institución educativa.

Tabla 14

Calificación servicio agua

Región	Calificación servicio agua			Total
	Bueno	Regular	Malo	
Sierra	1.655	1.685	585	3.925
Costa	1.789	915	1.658	4.362
Amazonía	305	535	414	1.254
Total	3.749	3.135	2.657	9.541

En este caso lo primero que se debe realizar antes de iniciar con el paso 1, es verificar que no existe independencia entre las variables, por lo que se utiliza el test chi cuadrado, obteniéndose el siguiente

Ho: La calificación a los servicios de agua en las instituciones educativas es independiente de la región geográfica donde se encuentra ubicada

Ha: La calificación a los servicios de agua en las instituciones educativas depende de la región geográfica donde se encuentra ubicada

Pearson's Chi-squared test

$$X\text{-squared} = 838.23, df = 4, p\text{-value} = 2.2e-16$$

Como se puede observar el p-value es menor a 0.05 e incluso menor a 0.01 por lo que se puede afirmar que no existe independencia entre las variables

Calcular la matriz de residuos normalizados: $S = D_r^{-1/2}(P - r c^T) D_c^{-1/2}$

Para iniciar con este primer paso se debe obtener la matriz P, que es la misma matriz original con los valores relativizados

P = Matriz P

Región	Bueno	Regular	Malo
Sierra	0.1734619	0.17660623	0.06131433
Costa	0.187506	0.09590190	0.17377633
Amazonia	0.031967	0.05607379	0.04339168

Luego se tiene la expresión rc^T que es la matriz de valores esperados de la matriz P

Obtención de matriz de valores esperados (rc^T)

La fórmula general para la obtención de los valores esperados es la siguiente:

$$E_{ij} = \frac{Fila_i * Columna_i}{Total}$$

Donde:

E_{ij} : Es el valor esperado en la celda correspondiente.

$Fila_i$: Es el valor total de la fila i

$Columna_i$: Corresponde al valor total de la columna i

$Total$: Suma de todos los valores de la tabla

Región	Calificación servicio agua		
	Bueno	Regular	Malo
Sierra	$(3.925*3.749)/9.541$	$(3.925*3.135)/9.541$	$(3.925*2.657)/9.541$
Costa	$(4.362*3.749)/9.541$	$(4.362*3.135)/9.541$	$(4.362*2.657)/9.541$
Amazonía	$(1.254*3.749)/9.541$	$(1.254*3.135)/9.541$	$(1.254*2.657)/9.541$

Desarrollando las operaciones de la matriz anterior se obtiene.

Región	Calificación servicio agua		
	Bueno	Regular	Malo
Sierra	1.542,27	1.289,68	1.093,04
Costa	1.713,99	1.433,27	1.214,74
Amazonía	492,74	412,04	349,22

Luego se obtiene la matriz de frecuencias relativas

Región	Calificación servicio agua		
	Bueno	Regular	Malo
Sierra	0.1616468	0.13517283	0.1145627
Costa	0.17964425	0.15022265	0.1273179
Amazonía	0.05164463	0.04318643	0.0366017

Entonces claramente se puede observar que lo que hace la expresión $\mathbf{P} - \mathbf{rc}^T$, es restar lo observado menos lo esperado en términos relativos, obteniéndose el siguiente resultado

Matriz $\mathbf{P} - \mathbf{rc}^T$

	Bueno	Regular	Malo
Sierra	0,011815028	0,04143339	-0,053248421
Costa	0,007862305	-0,05432075	0,046458446
Amazonía	-0,01967733	0,01288736	0,006789974

La matriz $\mathbf{P} - \mathbf{rc}^T$ se encuentra previamente multiplicado por $\mathbf{D}_r^{-1/2}$ que es la matriz diagonal de filas elevada a la un medio negativo y posteriormente se multiplica por la matriz diagonal de las columnas elevada a la un medio negativo $\mathbf{D}_c^{-1/2}$

Matriz $\mathbf{D}_r^{-1/2}$

	[,1]	[,2]	[,3]
	1.559111	0	0
	0	1.478952	0
	0	0	2.758342

Matriz $\mathbf{D}_c^{-1/2}$

	[,1]	[,2]	[,3]
	1.595288	0	0
	0	1.478952	0
	0	0	2.758342

Con esto se cuenta con todos los elementos necesarios para calcular la matriz S que lo que hace es tener un valor estandarizado con media cero y varianza 1,

$$\text{Matriz } \mathbf{S} = \mathbf{D}_r^{-1/2}(\mathbf{P} - \mathbf{r}\mathbf{c}^T)\mathbf{D}_c^{-1/2}$$

	[,1]	[,2]	[,3]
[1,]	0.000477465	0.0014001673	-1.525071e-03
[2,]	0.0003531045	-0.002040051	1.478748e-03
[3,]	-0.000254057	0.0001391397	6.213116e-05

Calcular la SVD: $\mathbf{S} = \mathbf{U}\mathbf{D}_\alpha\mathbf{V}^T$ donde $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$

La “Descomposición en Valores Singulares” SVD es una generalización de la descomposición en autovalores de una matriz, y se puede aplicar a cualquier matriz rectangular, Para calcular la SVD, es necesario descomponer la matriz S en tres sub matrices la U, la D y la V, donde la matriz $\mathbf{U}^T\mathbf{U}$ y la matriz $\mathbf{V}^T\mathbf{V}$ son iguales a la matriz identidad, dónde la matriz U está ligada a las filas de la matriz de residuos, mientras que V está ligada a la de columnas y la matriz D es la inercia

Descomposición SVD

\$d			
[1]	2.711657e-01	1.196858e-01	4.487363e-17

\$u

	[,1]	[,2]	[,3]
[1,]	0.71042728	0.2896733	0.6413910
[2,]	-0.70183913	0.2241362	0.6761544
[3,]	0.05210491	-0.9305119	0.3625366

\$v

	[,1]	[,2]	[,3]
[1,]	0.01234084	0.7790455	-0.6268459
[2,]	0.66991116	-0.4718444	-0.5732206
[3,]	-0.74233870	-0.4128570	-0.5277143

Coordenadas principales de las filas: $\mathbf{F} = \mathbf{D}_r^{-1/2}\mathbf{U}\mathbf{D}_\alpha$

La matriz F relacionada a las coordenadas principales de las filas es el resultado de multiplicar la **Matriz** $D_r^{-1/2}$ obtenida en el punto 1, por la matriz singular U del punto 2 y por la matriz D_α que contiene a los valores singulares de la matriz S, los cuales están almacenados en el vector “d” dentro del resultado de la función svd.

Matriz $D_r^{-1/2}$

	[,1]	[,2]	[,3]
[1,]	1.559111	0	0
[2,]	0	1.478952	0
[3,]	0	0	2.758342

Matriz U

	[,1]	[,2]	[,3]
[1,]	0.71042728	0.2896733	0.6413910
[2,]	-0.70183913	0.2241362	0.6761544
[3,]	0.05210491	-0.9305119	0.3625366

Matriz D_α

	[,1]	[,2]	[,3]
[1,]	0.2711657	0,0000000	0,000000e+00
[2,]	0,0000000	0.1196858	0,000000e+00
[3,]	0,0000000	0,0000000	4.487363e-17

Siendo el producto de estas tres matrices la matriz $F = D_r^{-1/2}UD_\alpha$

Matriz F, coordenadas principales fila

	[,1]	[,2]	[,3]
[1,]	0.3003527	0.05405404	4.487363e-17
[2,]	-0.2814663	0.03967425	4.487363e-17
[3,]	0.0389728	-0.30719394	4.487363e-17

Coordenadas principales de las columnas: $G = D_c^{-1/2}VD_\alpha$

Se realiza el mismo procedimiento que en el punto 3 para las coordenadas fila, pero esta vez para las columnas.

Matriz $D_c^{-1/2}$

	1.595288	0	0
	0	1.744529	0
	0	0	1.894965

Matriz V

	[,1]	[,2]	[,3]
[1,]	0.01234084	0.7790455	- 0.6268459
[2,]	0.66991116	-0.4718444	- 0.5732206
[3,]	-0.74233870	-0.4128570	- 0.5277143

Matriz D_α

	[,1]	[,2]	[,3]
[1,]	0,2711657	0,0000000	0,000000e+00
[2,]	0,0000000	0,1196858	0,000000e+00
[3,]	0,0000000	0,0000000	4.487363e-17

Matriz de coordenadas principales de columnas: $G = D_c^{-1/2}VD_\alpha$

	[,1]	[,2]	[,3]
[1,]	0.005338494	-0.14874577	-4.4873e-17
[2,]	0.316905766	-0.09851889	-4.4873e-17
[3,]	-0.381450353	-0.09363612	-4.4873e-17

Coordenadas estándar de las filas: $X = D_r^{-1/2}U$

Para obtener es las coordenadas estándar de filas debemos multiplicar la matriz diagonal de filas por la matriz singular U

Matriz $D_r^{-1/2}$

	[,1]	[,2]	[,3]
	1.559111	0	0
	0	1.478952	0

0 0 2.758342

Matriz U

	[,1]	[,2]	[,3]
[1,]	0.71042728	0.2896733	0.6413910
[2,]	-0.70183913	-0.2241362	0,6761544
[3,]	0.05210491	-0.930511	0.362536

Matriz de Coordenadas estándar de las filas: $X = D_r^{-1/2}U$

Resulta de multiplicar la matriz diagonal de las filas por la matriz de valores singulares U

	[,1]	[,2]	[,3]
[1,]	1.1076352	0.4516329	1
[2,]	-1,0379865	-0,3314868	1
[3,]	0,1437232	2,5666703	1

Coordenadas estándar de las columnas: $Y = D_c^{-1/2}V$

Resulta de multiplicar la matriz diagonal de las columnas por la matriz de valores singulares

V

	[,1]	[,2]	[,3]
[1,]	0,0196872	-1,2428024	-1
[2,]	1,1686794	-0,8231462	-1
[3,]	-1.406705	-0.782349	-1

- $7, \text{Inercia} = \varnothing^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(p_{ij} - r_i c_j)^2}{r_i c_j}$

La inercia corresponde a los cuadrados de los valores singulares de la matriz, Cuanto mayor sea el valor de la inercia, mayor será la importancia de ese componente en la representación de la matriz original. Esto nos permite identificar los componentes más relevantes y descartar aquellos que tengan un impacto menor en el análisis o procesamiento de los datos

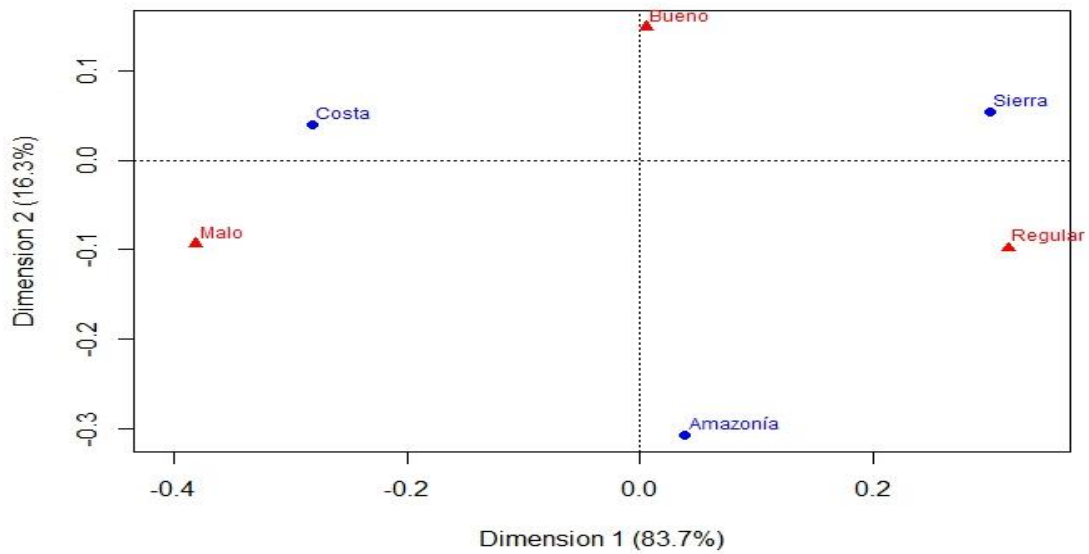
0,6937819	1.00	1.00
-----------	------	------

Una vez obtenidas las coordenadas a través del proceso matemático, es posible generar un gráfico que asocia las categorías según su proximidad. En este gráfico, se observa que las instituciones educativas ubicadas en la región Sierra tienden a asociarse con la percepción

de que el agua es regular, es decir, ni buena ni mala. Por otro lado, las instituciones de la Costa muestran una mayor cercanía a la categoría que indica malas condiciones del agua que reciben.

Figura 25

Ejemplo categorías en el espacio.



Anexo D. Porcentaje de varianza explicada 13 dimensiones

Tabla 18

Eigenvalores y porcentaje varianza explicada

Dimensión	Eigenvalues	Porcentaje varianza explicada
1	2,3749294	18,27%
2	1,5758822	12,12%
3	1,5367744	11,82%
4	1,0437560	8,03%
5	0,9911357	7,62%
6	0,9115270	7,01%
7	0,8487129	6,53%
8	0,8281772	6,37%
9	0,7678842	5,91%
10	0,7164468	5,51%
11	0,5365737	4,13%
12	0,4972633	3,83%
13	0,3709372	2,85%

Anexo E. Valores transformados tras la cuantificación óptima en ACP No Lineal

Transformación variable: Etnia de estudiantes

La variable categórica etnia de los estudiantes cuenta con las siguientes categorías:

Estudiantes autoidentificados como afroecuatorianos, montubios e indígenas.

Se tiene las categorías:

Instituciones educativas con porcentaje de estudiantes autoidentificados como afroecuatorianos, montubios e indígenas menores o igual al porcentaje nacional, categoría 1

Instituciones educativas con porcentaje de estudiantes autoidentificados como afroecuatorianos, montubios e indígenas mayor al porcentaje nacional, categoría 2

Categoría	Transformado
1	-0,006086208
2	0,017262653

Al realizar la transformación los valores transformados por categoría son

Figura 26

Puntuación transformada variable autoidentificación estudiantes



Variable Autoidentificación étnica docentes.

Se tiene las categorías:

Instituciones educativas con porcentaje de docentes autoidentificados como afroecuatorianos, montubios e indígenas menor o igual al porcentaje nacional, categoría 1

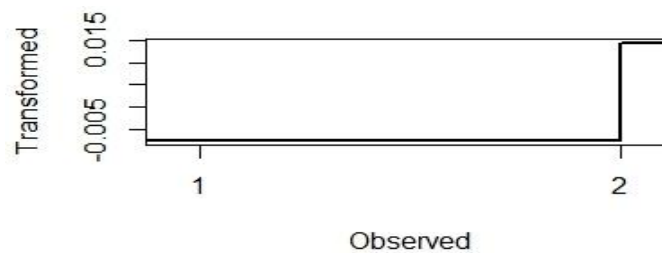
Instituciones educativas con porcentaje de docentes autoidentificados como afroecuatorianos, montubios e indígenas mayores al porcentaje nacional, categoría 2.

Al realizar la transformación los valores transformados por categoría son

Categoría	Transformado
1	-0,00737391933884899
2	0,0142480659560262

Figura 27

Puntuación transformada variable auto identificación docentes



Variable discapacidad estudiantes.

Se tiene las categorías:

Instituciones educativas con porcentaje de estudiantes que auto reportan tener una discapacidad menor al porcentaje nacional, categoría 1

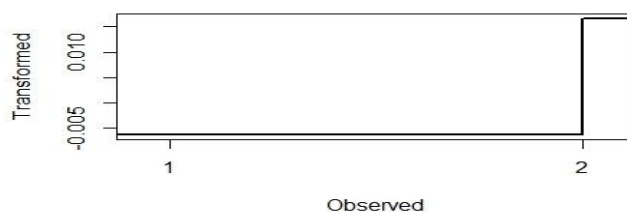
Instituciones educativas con porcentaje de estudiantes que auto reportan tener una discapacidad mayor al porcentaje nacional, categoría 2

Categoría	Transformado
1	-0,00632705888303259
2	0,0166055178301069

Al realizar la transformación los valores transformados por categoría son

Figura 28

Puntuación transformada variable discapacidad estudiantes



Variable ratio estudiantes por docente.

Se tiene las categorías:

Instituciones educativas con ratio estudiantes por docente menor al porcentaje nacional, categoría 1

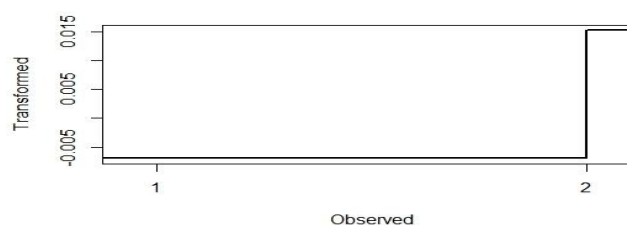
Categoría	Transformado
1	-0,00692856009860965
2	0,0151639139444613

Instituciones educativas con ratio estudiantes por docente mayor al porcentaje nacional. categoría 2

Al realizar la transformación los valores transformados por categoría son

Figura 29

Puntuación transformada variable ratio estudiantes por docente



Variable docentes con título de cuarto nivel

Se tiene las categorías:

Instituciones educativas con porcentaje de estudiantes con título de cuarto nivel mayor al porcentaje nacional, categoría 1

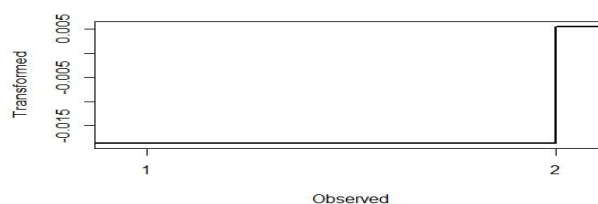
Instituciones educativas con porcentajes de docente con título de cuarto nivel menor al porcentaje nacional, categoría 2

Al realizar la transformación los valores transformados por categoría son

Categoría	Transformado
1	-0,00692856009860965
2	0,0151639139444613

Figura 30

Puntuación transformada variable docentes con título de cuarto nivel



Variable docentes con nombramiento.

Se tiene las categorías:

Instituciones educativas con porcentaje de docentes con nombramiento en la institución educativa mayor al porcentaje nacional, categoría 1

Instituciones educativas con porcentajes de docente con nombramiento menor al porcentaje nacional, categoría 2

Categoría	Transformado
1	-0,00961697906056586
2	0,0109248536814597

Al realizar la transformación los valores transformados por categoría son

Figura 31

Puntuación transformada variable docentes con nombramiento



Variable abandono estudiantes.

Se tiene las categorías:

Instituciones educativas con porcentaje de estudiantes que abandonan los estudios en la institución educativa mayor al porcentaje nacional, categoría 1

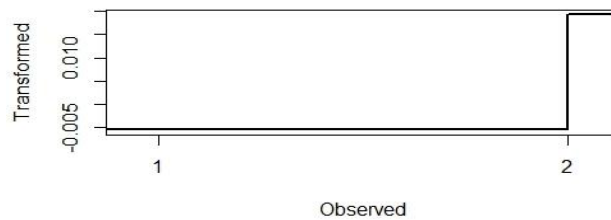
Instituciones educativas con porcentaje de estudiantes que abandonan los estudios en la institución educativa menor al porcentaje nacional, categoría 2

Categoría	Transformado
1	-0,005475381
2	0,019188452

Al realizar la transformación los valores transformados por categoría son

Figura 32

Puntuación transformada variable abandono



Variable área de ubicación de la institución educativa.

Se tiene las categorías:

Instituciones educativas ubicadas en el área rural, categoría 1

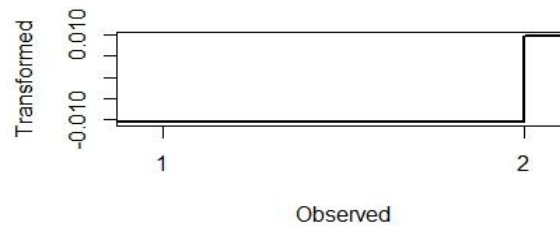
Instituciones educativas ubicadas en el área urbana, categoría 2

Al realizar la transformación los valores transformados por categoría son

Categoría	Transformado
1	-0,005475381
2	0,019188452

Figura 33

Puntuación transformada variable Área de ubicación de la Institución Educativa



Variable Forma de acceso a la institución educativa.

Se tiene las categorías:

Acceso terrestre a la Institución educativa, categoría 1

Acceso fluvial a la institución educativa, categoría 2

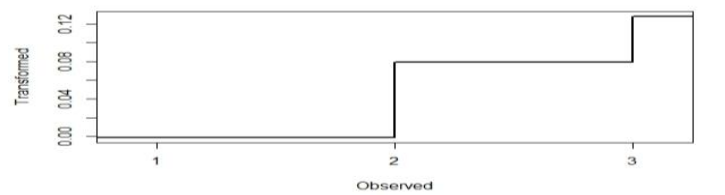
Acceso aéreo a la institución educativa

Al realizar la transformación los valores transformados por categoría son

Categoría	Transformado
1	-0,00125860410726271
2	0,0815728159891971
3	0,105490552151071

Figura 34

Puntuación transformada variable Acceso a la Institución Educativa



Variable agua en la institución educativa.

Se tiene las categorías:

Agua identificada como buena, categoría 1

Agua identificada como regular, categoría 2

Agua identificada como mala, categoría 3

Al realizar la transformación los valores transformados por categoría son

Categoría	Transformado
1	-0,012614601905897
2	0,0064891557692505
3	0,0101480444762643

Figura 35

Puntuación transformada variable Agua en la Institución Educativa



Variable lavamanos en la institución educativa.

Se tiene las categorías:

Lavamanos identificada como bueno, categoría 1

Lavamanos identificada como regular, categoría 2

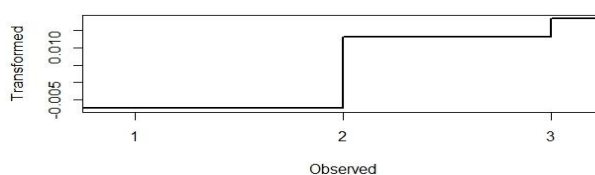
Lavamanos identificada como malo, categoría 3

Al realizar la transformación, los valores transformados por categoría son

Categoría	Transformado
1	-0,00738349280098396
2	0,0129967350327288
3	0,018818312177421

Figura 36

Puntuación transformada variable lavamanos en la institución educativa



Variable Inodoro en la institución educativa.

Se tiene las categorías:

Inodoro identificada como bueno, categoría 1

Inodoro identificada como regular, categoría 2

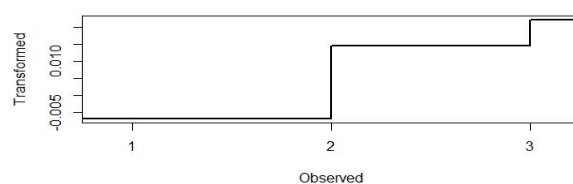
Inodoro identificada como malo, categoría 3

Al realizar la transformación, los valores transformados por categoría son

Categoría	Transformado
1	-0,00678020460228041
2	0,0143740124605365
3	0,0228097377196305

Figura 37

Puntuación transformada variable inodoros en la institución educativa



Variable servicio de Internet en la institución educativa.

Se tiene las categorías:

Instituciones educativas con servicio de Internet, categoría 1

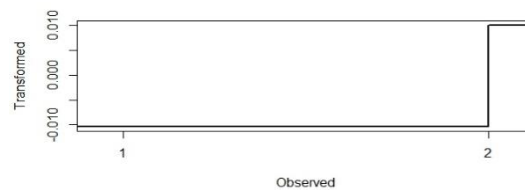
Instituciones educativas sin servicio de Internet, categoría 2

Al realizar la transformación, los valores transformados por categoría son

Categoría	Transformado
1	-0,01041291

Figura 38

Puntuación transformada variable servicio de internet en la Institución Educativa

**Tabla 19**

Resumen puntuaciones ACPNL

No.	Variable	Categoría	Valor categoría	Puntuación transformada
1	Etnia	Estudiantes auto identificados como Afro, montubios, indígenas menor al nacional 1	1	-0,00608
		Estudiantes auto identificados como Afro, montubios, indígenas mayor al nacional 2	2	0,01726
2	Etniad	Docentes auto identificados como Afro, montubios, indígenas menor al nacional 1	1	-0,00732
		Docentes auto identificados como Afro, montubios, indígenas mayor al nacional 2	2	0,01424
3	disc	Estudiantes con discapacidad menor al nacional 1	1	-0,00632
		Estudiantes con discapacidad mayor al nacional 2	2	0,01660

No.	Variable	Categoría	Valor categoría	Puntuación transformada
4	r_esdoc	Ratio estudiantes por docente menor al nacional 1	1	-0,00692
		Ratio estudiantes por docente mayor al nacional	2	0,01516
5	doc4n	Docentes cuarto nivel mayor al nacional 1	1	-0,01871
		Docentes cuarto nivel menor al nacional 2	2	0,00560
6	nombramiento	Docentes nombramiento mayor al Nacional	1	-0,00961
		Docentes nombramiento menor al Nacional	2	0,01092
7	aband	Porcentaje abandono menor al nacional	1	-0,00547
		Porcentaje abandono mayor al Nacional	2	0,01918
9	acceso	Acceso terrestre	1	-0,00125
		Acceso fluvial	2	0,08111
		Acceso aéreo	3	0,11058
10	agua	Agua buena	1	-0,0125
		Agua Regular	2	0,00607
		Agua Mala	3	0,01055
11	Lavamanos	Lavamanos Bueno	1	-0,00739
		Lavamanos Regular	2	0,01310
		Lavamanos Malo	3	0,01840
12	Inodoro	Inodoro Bueno	1	-0,00678
		Inodoro Regular	2	0,01446
		Inodoro Malo	3	0,02223
13		Establecimiento educativo con Internet	1	-0,01041

No.	Variable	Categoría	Valor categoría	Puntuación transformada
	Internet	Establecimiento educativo sin Internet	2	0,01008