

Universidad Nacional de Colombia

FACULTAD DE CIENCIAS

ESTADÍSTICA DESCRIPTIVA MULTIVARIADA

Caractetización del trabajo infantil en Colombia para 2022

Juan David Duitama Correa
Alejandro Urrego Lopez
Nicolás Alejandro Díaz Ubaque

Noviembre 2023

1. Introducción

El trabajo infantil en Colombia es una preocupación significativa que afecta no solo a la infancia del país, sino también su desarrollo social y económico. Este fenómeno se manifiesta en diversas formas, desde labores peligrosas y físicamente exigentes hasta empleos informales que privan a los niños de su educación y desarrollo integral. La problemática se enmarca dentro de un contexto complejo de desigualdad socioeconómica, falta de acceso a oportunidades educativas y vulnerabilidades estructurales.

A pesar de los avances legislativos y los esfuerzos gubernamentales por mitigar el trabajo infantil, las cifras revelan una persistencia preocupante de esta práctica. Según estimaciones recientes, alrededor del 3,4 % de los niños colombianos estaban involucrados en actividades laborales para el primer trimestre de 2022, lo que refleja una realidad que demanda una comprensión profunda y estrategias efectivas para su abordaje.

En línea con los Objetivos de Desarrollo Sostenible de la ONU, establecidos en la Asamblea Internacional de 2015, este proyecto se propone analizar y caracterizar la compleja situación del trabajo infantil en Colombia. Para este fin, se utilizarán datos provenientes del módulo de trabajo infantil de la Gran Encuesta Integrada de Hogares del cuarto trimestre de 2022, permitiendo así una exploración detallada de las condiciones sociodemográficas y culturales que rodean a esta problemática.

La elección del análisis de correspondencias múltiples como herramienta metodológica se justifica en su capacidad para revelar relaciones subyacentes entre variables multidimensionales, lo que posibilitará una comprensión más profunda de los factores que contribuyen al fenómeno del trabajo infantil en Colombia. Asimismo, este enfoque permitirá identificar posibles patrones y asociaciones que podrían guiar el diseño de políticas más efectivas y estrategias de intervención adecuadas para salvaguardar los derechos y el bienestar de los niños en el país.

Con este marco en mente, este proyecto busca no solo arrojar luz sobre una problemática de importancia social crucial, sino también proponer recomendaciones concretas que contribuyan a la creación de un entorno más seguro y propicio para el desarrollo integral de la infancia colombiana, y brindar soportes teóricos basados en el análisis de datos que permitan a expertos en ámbitos sociales plantear posibles soluciones a la problemática del trabajo infantil.

2. Marco Teórico

Estudios anteriores

Este proyecto se enmarca dentro de un contexto de investigación que ha abordado el tema del trabajo infantil en Colombia en estudios previos, tales como el boletín técnico del Departamento Administrativo Nacional de Estadística (DANE) que presentó las cifras más recientes sobre trabajo infantil en octubre de 2022.

El análisis presentado en este boletín se centró en un enfoque univariado, examinando de manera detallada aspectos clave como la segmentación por género, las razones que impulsan a los niños a ejercer actividades laborales, las ramas laborales específicas involucradas y los grupos de edad afectados.

Además, este estudio proporcionó descripciones gráficas y tabulares, y comparó las cifras actuales con datos de años anteriores, aportando así una perspectiva histórica crucial para comprender la evolución de esta problemática social. Es importante destacar que el boletín técnico del DANE también consideró una categoría ampliada de trabajo infantil, que abarca tareas domésticas que superan las 15 horas semanales, en consonancia con las pautas establecidas por organismos internacionales.

Otro estudio tenido en cuenta para este proyecto, es el de 'Caracterización del trabajo infantil del año 2018 en Colombia', realizado por Yira Del Carmen Guerra Vergara, en donde la autora indaga sobre algunos conceptos y leyes que consideran y regulan el trabajo infantil, con el fin de lograr delimitar las variables y criterios que engloban al trabajo infantil en el país, realizando de igual forma análisis de cifras de trabajo infantil por grupos de edad, sexo, rama ocupacional y razones por las que trabaja, y variables como ingreso y tasa de asistencia escolar.

Análisis de correspondencias múltiples

El análisis de correspondencias múltiples (ACM) es una técnica estadística multivariada utilizada para analizar la relación entre variables categóricas en conjuntos de datos multivariados. El ACM permite visualizar y explorar la estructura de dependencia entre categorías de variables en un espacio dimensional reducido, lo que facilita la interpretación de patrones complejos y la identificación de tendencias significativas en los datos.

Al aplicar el ACM al estudio del trabajo infantil en Colombia, se pretende examinar las relaciones entre diversas características socio-demográficas y culturales para los infantes de estudio. Esta técnica permitirá identificar posibles asociaciones entre variables como la edad, el género, la ubicación geográfica, el nivel educativo, las condiciones socioeconómicas y otras variables relevantes. Al visualizar estas relaciones en un espacio reducido, el análisis de correspondencias múltiples proporcionará una comprensión más profunda de los factores que contribuyen al fenómeno del trabajo infantil en el contexto colombiano, lo que a su vez ayudará a informar el diseño de políticas y estrategias efectivas de intervención.

Es importante destacar que el ACM se basa en técnicas de análisis exploratorio de datos y se enfoca en revelar patrones ocultos y asociaciones significativas entre variables, lo que lo convierte en una herramienta valiosa para la identificación de factores determinantes y la formulación de enfoques preventivos y correctivos más precisos y adaptados a la realidad del contexto del trabajo infantil en Colombia para el año 2022.

3. Objetivos

- Caracterizar el trabajo infantil en Colombia para el año 2022, estudiando su relación con variables sociodemográficas, culturales y socioeconómicas asociadas a los infantes del país.
- Identificar patrones de asociación entre variables, que permitan establecer grupos de infantes que sean potencialmente vulnerables frente al trabajo infantil.

- Estudiar la relación entre el trabajo infantil y la disposición a condiciones de vida dignas, así como acceso a servicios básicos y estudio.
- Analizar la relación entre las características de los infantes de estudio y la prevalencia del trabajo infantil en el país.
- Determinar si el trabajo infantil se relaciona también con otros tipos de explotación en los infantes del país.

4. Metodología

Datos

Los datos objeto de análisis para este proyecto corresponden al módulo de trabajo infantil aplicado en la Gran Encuesta Integrada de Hogares en el cuarto trimestre del 2022, agregando algunas variables que no se encuentran dentro del mismo, pero sí dentro de la GEIH.

Variable	Recategorización	Categoría	Frecuencia	Porcentaje
Asiste al colegio	COLEA	Sí	1214	39.92 %
		No	1828	60.08 %
Actividad Principal	ACTPR	Estudiando	506	16.64 %
		Oficios del hogar	2030	66.77 %
		Trabajando	506	16.64 %
Vive con su papá	PA	No	2423	79.61 %
		Sí	619	20.39 %
Es campesino	CAMP	No	2634	86.66 %
		Sí	408	13.34 %
Acueducto	ACUED	0	587	19.31 %
		1	2455	80.69 %
		4+	1375	45.17 %
Tamaño Hogar	TAMHO	1-2	272	8.94 %
		3-4	1395	45.89 %
		4+	1375	45.17 %
Edad	EDADC	14-	1032	33.91 %
		14-15	564	18.54 %
		15+	1446	47.55 %
Clase de Hogar	HOGAC	Rural	610	20.04 %
		Urbano	2432	79.96 %

Tabla 1: Descripción de variables (transpuesta, una columna) con frecuencia relativa.

Variable	Recategorización	Categoría	Frecuencia	Porcentaje
Sexo	SEXO	Hombre	1397	45.92 %
		Mujer	1645	54.08 %
Parentesco	PAREN	H	2148	70.63 %
		N	339	11.15 %
		OtroP	555	18.22 %
Afiliado Seguro Social	AFILI	Sí	2629	86.46 %
		No	413	13.54 %
Sabe Leer	ESLEE	Sí	2811	92.50 %
		No	231	7.58 %
Asiste al Colegio	COLEP	Publica	1080	35.46 %
		Privada	134	4.40 %
		NA's	1828	60.14 %
Vive con su mamá	MA	No	1918	63.01 %
		Sí	1124	36.99 %
Tipo de Hogar	HOGAT	Apto	978	32.14 %
		Casa	1953	64.13 %
		Cuarto	63	2.07 %
		Otro	48	1.58 %
Cuenta con Luz	LUZ	0	97	3.16 %
		1	2945	96.84 %
Cuenta con Gas	GAS	0	1352	44.40 %
		1	1690	55.60 %
Cuenta con Alcantarillado	ALCAN	0	1084	35.64 %
		1	1958	64.36 %
Recoleccion de Basura	BASUR	0	593	19.48 %
		1	2449	80.52 %
Origen del agua	PRAWA	Naturaleza	248	8.15 %
		Otros	131	4.31 %
		Tecnificado	2663	87.54 %
Propiedad de la Vivienda	PROPI	Arriendo	1303	42.78 %
		Otro	977	32.04 %
		Propia	762	25.18 %

Tabla 2: Descripción de variables (transpuesta, una columna) con frecuencia relativa.

En el primer ACM, las variables activas se seleccionaron para evaluar el nivel de vulnerabilidad de los infantes que viven en zonas rurales. Se consideraron el tipo de vivienda, pues con análisis previos y el contexto del país, se tiene claridad sobre la incidencia del trabajo infantil en zonas rurales, el servicio de acueducto, pues se considera el servicio de necesidad más básica respecto a los que se tienen en el estudio, y la razón de usar únicamente esta como variable activa y las demás como variables ilustrativas, radica en que todas ellas brindan información muy parecida, la cohabitación con el padre, considerándose ésta más relevante que la cohabitación con la madre, teniendo en cuenta la frecuencia en problemas de abandono paternal que se presentan en el país, el tamaño del hogar como un indicador de hacinamiento, la edad, para observar como se relacionan los distintos rangos de la misma, especialmente infancia y adolescencia, con respecto a la necesidad de ejercer actividades laborales y las oportunidades de estudio, y la asistencia a la escuela, siendo

un aspecto de importancia vital para el estudio, pues corresponde a la actividad que deberían realizar todos los infantes del país, y estudiar las características que permiten que los infantes tengan la oportunidad de hacerlo puede ayudar a plantear estrategias para afianzar estas características. Por otro lado, las variables ilustrativas se utilizaron para agregar información relevante, aunque al introducirlas se observó que algunas generaban ruido o estaban altamente correlacionadas con las variables activas, como es el caso de los servicios básicos distintos al acueducto, o la cohabitación con la madre. A pesar de esto, su inclusión permitió examinar su relación con las variables activas en el plano factorial, y caracterizar de manera más completa los grupos obtenidos.

En el segundo ACM, se añadieron variables activas específicas para los niños con actividad principal trabajando, como el salario o la razón por la cual trabajan, con el objetivo de profundizar en la comprensión de esta subpoblación.

Tratamiento de datos

A partir de la variable edad, se obtiene la variable EDAD C que tiene 3 categorías: 14- (menos de 14 años), 15+ (más de 15 años) y 14-15 (entre 14 y 15 años). La variable Tamaño del Hogar, que indica cuántas personas viven en el hogar, se convierte en la variable tam_hogarC con 3 categorías: 1-2 (1 o 2 habitantes), 4+ (más de 4 habitantes) y 3-4 (3 y 4 habitantes). El estado civil se recategorizó en 2 grupos: soltero y otro estado civil. De igual manera, el parentesco se categorizó en hijo, nieto y otro parentesco.

Para el estudio se considera la base de datos del módulo de Trabajo Infantil en la GEIH de 2022. Aún así, con el objetivo de complementar la información presente en la misma, se incluye información adicional de la GEIH, realizando una unión de los conjuntos de datos por medio del identificador de cada infante. Características generales, seguridad social en salud y educación; junto a Datos del hogar y la vivienda, fueron los conjuntos de datos adicionales seleccionados.

Se recategorizan algunas variables que cuentan con demasiadas categorías, como tipo del hogar, acceso al agua de consumo y estado de propiedad de la vivienda en la que reside, pues la interpretabilidad del ACM se ve afectada entre mayor número de grupos se tengan, y se considera que varias de estas categoría no proporcionan información relevante al estudio, por lo que se recategorizan como «otros» dentro de sus respectivas variables.

El análisis exclusivo de solo los infantes que trabajan requirió recategorización en la edad por desbalance, tres grupos de edades, 15 o menos, 16 y, 17 años; al igual que en el anterior caso, las variables del nivel educativo alcanzado, el estado de propiedad de la vivienda y el lugar de trabajo se recategorizaron en menos categorías con el fin de mejorar los procedimientos de este análisis.

Balance de datos

El módulo de trabajo infantil de la GEIH cuenta con registros de infantes que ejercen alguna actividad laboral e infantes que no. El número de infantes que se encuentran trabajando o buscando trabajo es de 506 individuos, aquellos individuos que desarrollan oficios del hogar que superan 15 horas semanales, es de 2030, mientras que el número de los que se encuentran estudiando es de 38452, por lo que se evidencia un desbalance en las proporciones de los datos.

Para solucionar esto, se decide realizar un muestreo aleatorio simple sobre los registros de los infantes que se encuentran estudiando, de forma que el número de datos para este grupo es el mismo

que de los que se encuentran trabajando, 506 individuos.

5. Analisis Infantes :

En este apartado, el objetivo es caracterizar a los infantes cuya actividad principal es el trabajo, a aquellos cuya actividad principal son los oficios del hogar y a los que tienen como actividad principal el estudio.

Selección de Ejes en el ACM

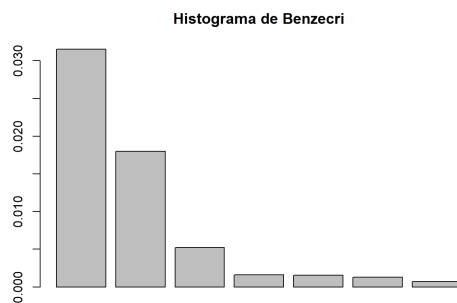
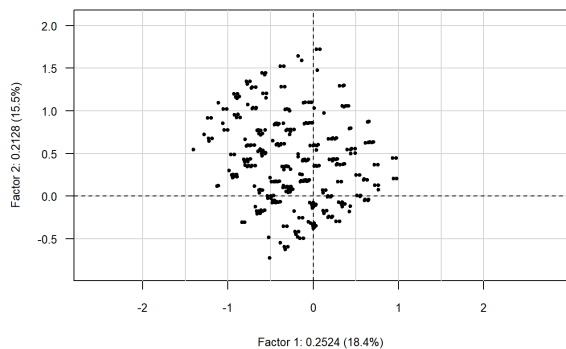
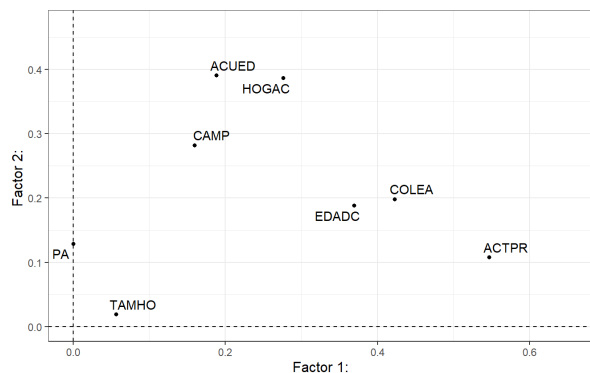


Figura 1: Criterio de Benzécri

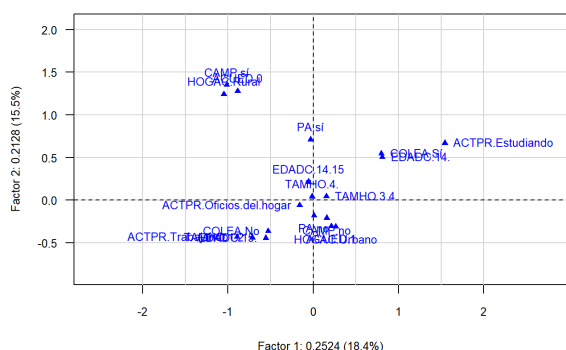
Para seleccionar el número de ejes, se utilizó el criterio de Benzécri. Claramente, sobresalen dos ejes, por lo tanto se opta por hacer el ACM con este número de ejes.



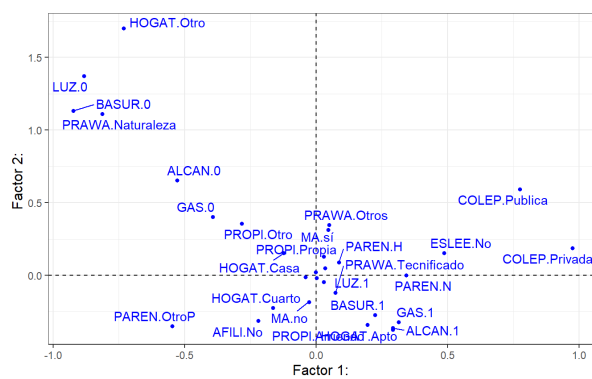
Nube de individuos



Razon de Correlacion



Variables Activas



Variables Ilustrativas

Figura 2: Primer Plano Factorial

Primer Eje Factorial

Se puede observar que en el primer eje factorial, se contraponen las categorías correspondientes a ser del campo, contra no serlo, hogares urbanos contra hogares rurales, infantes que se encuentran estudiando, contra los que se encuentran trabajando y realizando oficios del hogar, infantes de menos de 14 años contra aquellos que tienen 15 o más, y los que asisten al colegio contra los que no, teniendo en todos los casos, las primeras categorías mencionadas ubicándose en la dirección positiva del eje.

Segundo Eje Factorial

El segundo eje factorial se encarga de diferenciar a los infantes que viven con su padre, de los que no, y acentúa la segmentación que se presentó entre hogares rurales y urbanos en el primer eje. De igual manera, presenta la contraposición de los infantes provenientes del campo y los que no, e introduce una diferenciación entre infantes que tienen acceso a acueducto y los que no lo tienen.

Primer Plano Factorial

Se puede observar que en el primer cuadrante del plano factorial se encuentran las variables: menor de 14 años que asisten al colegio, vive con el papá, tiene acueducto, no vive en el campo y la actividad principal es el estudio. Por otro lado, podemos observar que en el segundo cuadrante se encuentran las variables: vive en el campo, no tiene acueducto y hogar rural. En el tercer cuadrante, las variables son: actividad principal oficios del hogar, actividad principal trabajando, tamaño de hogar 1 o 2, edad más de 14 años y no asiste al colegio. Mientras que en el cuarto cuadrante están las variables: no campesino y hogar urbano.

Variables Ilustrativas.

Índice	Valor Test	Valor Test
SEXO.Hombre	-0.096	1.025
SEXO.Mujer	0.096	-1.025

Tabla 3: Descripción de índices para Axis1 y Axis2.

Al observar los valores de test, las únicas variables que no se interpretan son 'sexo hombre', 'sexo mujer' y 'tipo de hogar cuarto'. Por lo tanto, se puede observar que la falta de servicios parece ocurrir al mismo tiempo y suele ser más común en el campo. De igual manera, al observar 'actividad principal estudiando', esta está cerca de 'colegio propio privado' y 'sabe no sabe leer'; esto debe estar asociado con la edad de los niños de este cuadrante. Luego, podemos observar que estos niños suelen contar con vivienda propia. De igual manera, podemos observar que los niños cuya actividad principal es el trabajo viven en arriendo, no viven con su mamá o con su papá, suelen vivir en apartamentos, y la relación con el dueño de la casa suele ser diferente a la de familiar.

Además, se puede observar que en el tercer cuadrante se encuentran las variables: 'no vive con la mamá', 'no está afiliado a servicio medio', 'otro parentesco', y 'vive en arriendo'.

Se puede notar que los niños que están trabajando o cuya actividad principal es "oficios del hogar" tienen menos acceso a servicios, no conviven con sus padres, no asisten al colegio y residen en la ciudad.

Como era de esperar, los niños que estudian tienen acceso a servicios, viven con sus padres o con sus abuelos, y son de la ciudad.

6. Cluster Metodo de Ward y k-means:

Número de Clases

Observando el barplot de los índices del dendrograma, se puede notar que se puede generar una partición en 6 clusters y una más fina en 7. Se decidió optar por la primera opción, ya que permite una mejor caracterización de los niños cuya actividad principal es trabajar, así como de los niños cuya actividad principal son los oficios del hogar o el estudio.

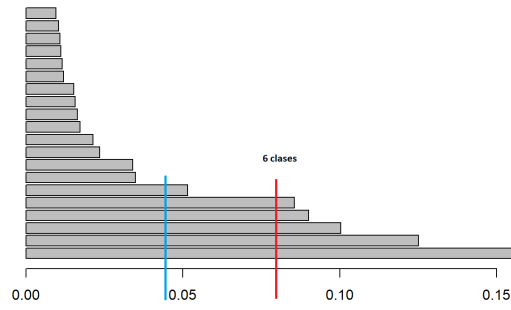


Figura 3: Barplot indices de nivel del endograma.

Numero de ejes para la clasificacion:

Al observar el histograma de los valores propios de Benzecri, se nota que solo hay 7 ejes que sobresalen. Por lo tanto, se decide tomar este número de ejes para realizar la clasificación.

Consolidacion de las clases

Bef.Size	Aft.Size	Bef.Inertia	Aft.Inertia	Bef.Weight	Aft.Weight	Bef.Dist_2	Aft.Dist_2
354	309	0.0684	0.0402	0.1164	0.1016	0.7395	0.7602
417	358	0.0840	0.0554	0.1371	0.1177	0.5857	0.6123
272	269	0.0746	0.0714	0.0894	0.0884	1.3310	1.3059
371	377	0.0897	0.1142	0.1220	0.1239	0.6613	1.0301
468	488	0.0705	0.0733	0.1538	0.1604	0.8279	0.7845
1160	1241	0.1710	0.1779	0.3813	0.4080	0.1697	0.1451
3042	3042	0.5582	0.5324	1.0000	1.0000	NA	NA

Tabla 4: Consolidacion de las clases

Al momento de consolidar las clases, 6 grupos ceden individuos, lo que sugiere que estas clases no son tan estables. Sin embargo, podemos notar que la tercera clase no cedió individuos y que la inercia ha disminuido de 0.5582 a 0.5324 .

Caracterizacion de las Clases

Al Explorar la asociacion entre las clases y las respectivas variables activas e ilustrativas mencionadas anteriormente podemos obsevar que las dos clases que caracterizan el trabajo son las clases 3 y 1,

Variable	Test.Value	p.Value	Class.Cat	Cat.Class	Global	Weight
ACTPR.Trabajando	36.247	0.000	61.1	100.0	16.6	506
EDAD.C.15+	21.421	0.000	21.2	99.0	47.5	1446
SEXO.Hombre	10.736	0.000	16.5	74.4	45.9	1397
COLEA.No	8.801	0.000	13.9	82.2	60.1	1828
MA.no	6.253	0.000	12.7	78.6	63.1	1918
ESLEE.Sí	5.206	0.000	10.9	98.7	92.4	2811
PA.no	5.142	0.000	11.5	90.0	79.7	2423
PAREN.OtroP	4.884	0.000	16.0	28.8	18.2	555
TAMHO.4+	3.996	0.000	12.5	55.7	45.2	1375
CAMP.no	3.938	0.000	10.9	93.2	86.6	2634
LUZ.1	3.096	0.002	10.4	99.4	96.8	2945
PROPI.Arriendo	3.013	0.003	12.0	50.5	42.8	1303

Tabla 5: Resultados para la Clase 1

Variable	Test.Value	p.Value	Class.Cat	Cat.Class	Global	Weight
TAMHO.1-2	Inf	0.000	98.9	100.0	8.9	272
PAREN.OtroP	10.689	0.000	21.8	45.0	18.2	555
PA.no	8.228	0.000	10.7	96.3	79.7	2423
MA.no	6.008	0.000	11.1	79.2	63.1	1918
EDAD.C.15+	5.844	0.000	12.0	64.3	47.5	1446
ACTPR.Trabajando	5.680	0.000	15.8	29.7	16.6	506
HOGAT.Cuarto	5.223	0.000	31.7	7.4	2.1	63
COLEA.No	3.476	0.001	10.2	69.5	60.1	1828

Tabla 6: Resultados para la Clase 3

Al observar ambas clases, se nota que la mayoría de los niños que trabajan son hombres mayores de 15 años que no viven con sus padres o madre. Además, no asisten al colegio, pero debido a su edad, saben leer. También parece que la mayoría no se encuentra en el campo. Esta información sugiere una tendencia en la que los niños en esta categoría se dedican al trabajo en lugar de asistir a la escuela, situación posiblemente influenciada por factores como la edad, el entorno familiar y las condiciones geográficas.

7. Análisis de infantes que trabajan

Este apartado, a comparación del anterior, explora las relaciones y describe el trabajo infantil sobre los infantes que tienen como actividad principal el estar trabajando o estar buscándolo. Se hizo uso, al igual que el análisis pasado, de un análisis de correspondencias múltiples(ACM). El resumen univariado de cada una de las variables utilizadas se encuentra presente en la tabla 7 junto a las categorías de cada variable.

Tabla 7: Resumen univariado variables activas

Variable	Abreviación	Categorías	Frecuencia
Sexo	SEXO	Hombre	376 (74.31 %)
		Mujer	130 (25.69 %)
Edad	EDAD	15 o Menos	102 (20.16 %)
		16	121 (23.91 %)
		17	283 (55.93 %)
Asiste a una institución educativa	COLEA	No	390 (77.08 %)
		Sí	116 (22.92 %)
Razón para trabajar	RAZON	Familiar	164 (32.41 %)
		Otra	342 (67.59 %)
Nivel Educativo Alcanzado	ANOGR	Bachillerato	121 (23.91 %)
		Ninguno	125 (24.70 %)
		Preescolar	154 (30.43 %)
		Primaria	81 (16.01 %)
		Otro	25 (4.94 %)
¿Es campesino?	CAMP	No	425 (83.99 %)
		Sí	81 (16.01 %)
Clase del Hogar	HOGAC	Rural	149 (29.45 %)
		Urbano	357 (70.55 %)
¿Cuenta con acueducto?	ACUED	No	119 (23.52 %)
		Sí	387 (76.48 %)
Estado de propiedad de la vivienda	PROPI	Arrendada	226 (44.66 %)
		Propia	120 (23.72 %)
		Otro	160 (31.62 %)
¿Dónde trabaja?	TRAPA	Ambulante	58 (11.46 %)
		Establecimiento	258 (50.99 %)
		Fuera	190 (37.55 %)
Sueldo	SUELD	-500Mil	373 (73.72 %)
		+500Mil	133 (26.28 %)

Tabla 8: Resumen univariado variables ilustrativas

Variable	Abreviación	Categorías	Frecuencia
Parentesco	PAREN	H	302 (59.68 %)
		N	53 (10.47 %)
		OtroP	151 (29.84 %)
Estado Civil	ESTCI	Soltero	427 (84.39 %)
		Otro	79 (15.61 %)
Afiliación seguridad social	AFILI	No	88 (17.39 %)
		No sabe	1 (0.2 %)
		Sí	417 (82.41 %)
Sabe leer y escribir	ESLEE	No	12 (2.37 %)
		Sí	494 (97.63 %)
Tipo de institución educativa	COLEP	Privado	24 (4.74 %)
		Público	92 (18.18 %)
		NA	390 (77.08 %)
Vive con su madre	VMA	Sí	120 (23.72 %)
		NA	386 (76.28 %)
Vive con su padre	VPA	Sí	65 (12.85 %)
		NA	441 (87.15 %)
Tipo de Hogar	HOGAT	Apartamento	154 (30.43 %)
		Casa	335 (66.21 %)
		Habitación	12 (2.37 %)
		Otro	5 (0.99 %)
Servicio de Luz	LUZ	No	14 (2.77 %)
		Sí	492 (97.23 %)
Servicio de Gas	GAS	No	260 (51.38 %)
		Sí	246 (48.62 %)
Servicio de Alcantarillado	ALCAN	No	211 (41.70 %)
		Sí	295 (58.30 %)
Servicio de recolección de basura	BASUR	No	129 (25.49 %)
		Sí	377 (74.51 %)
Procedencia del agua de consumo	PRAWA	Naturaleza	46 (9.09 %)
		Otros	21 (4.15 %)
		Tecnificado	439 (86.76 %)

Las variables ilustrativas fueron escogidas acorde a las frecuencias de las categorías que las componen, varias de estas se encuentran desbalanceadas, por lo que afectan el análisis multivariado, además de que la información que algunas de estas variables pueden brindar ya está suplida por algunas de las variables activas, como es el caso de los servicios públicos.

Primer eje factorial

Se requiere un solo eje para hacer la interpretación de los resultados, ya que el histograma de los valores propios lo indica (Figura 4), no obstante, para facilitar la visualización de los resultados, se utilizarán dos ejes, es decir, se hará uso de los planos factoriales. Siendo el primer eje, el de mayor importancia, este se encarga de caracterizar a las viviendas de los individuos, es decir,

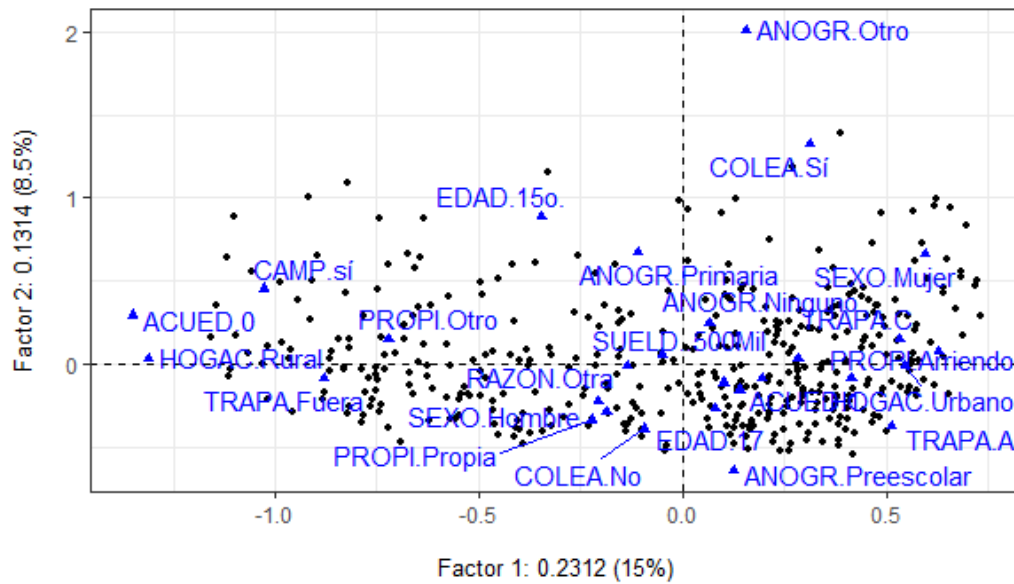
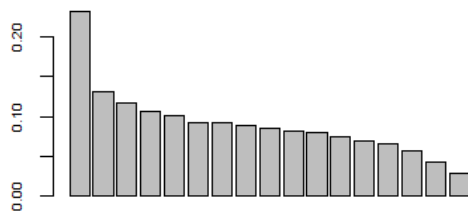
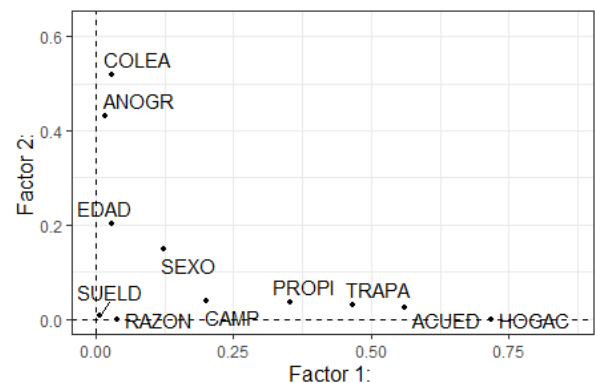


Figura 5: Primer plano factorial

las características que poseen. En el lado izquierdo del eje se encuentran las viviendas rurales, estas se encuentran muy relacionadas con el no poseer servicio de acueducto, recordemos que de los servicios públicos fue la única variable que se incluyó como activa; a su vez, ser campesino también se correlacionó con las viviendas rurales. Sucedió lo mismo con el lado derecho del eje, las viviendas ubicadas en este lado son viviendas urbanas que cuentan con servicios públicos, además de ser habitadas por infantes no campesinos.



Histograma de valores propios



Razón de correlación con el plano 1:2

Figura 4: Selección de ejes y caracterización plano 1:2

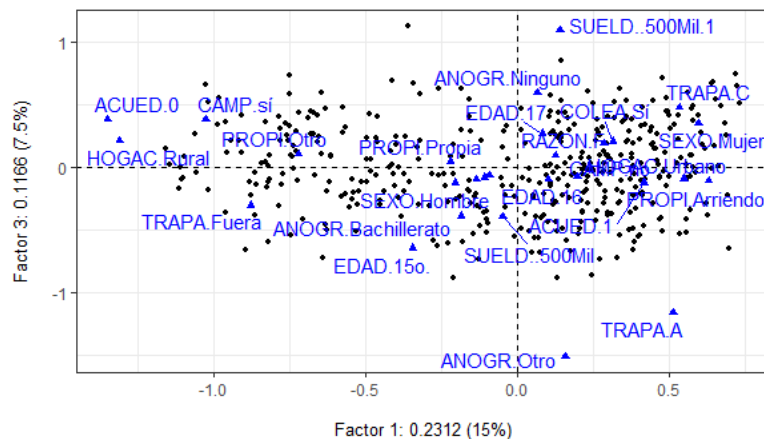
Estas viviendas urbanas están muy correlacionadas con la categoría Arriendo de la variable Estado de propiedad de la vivienda, es decir, los infantes que trabajan y viven en la ciudad; lo

hace en una vivienda arrendada. Finalmente, la variable ¿Dónde trabaja? se encuentra también caracterizada por el primer eje factorial, separando la categoría Fuera con las otras dos. Esta categoría se encuentra localizada en el sector izquierdo del primer eje, cercana a las viviendas rurales, dándonos una idea sobre la situación de los infantes que habitan esta clase de viviendas, estos deben desplazarse y trabajar en lugares como minas, campos con cultivos agrícolas, etc. Por el otro lado, las otras dos categorías, Ambulante y Establecimiento se encuentran localizadas en el sector derecho, con las viviendas urbanas, por lo que los infantes que trabajan y habitan en viviendas urbanas, en su mayoría, lo hacen en lugares como locales, otras viviendas, transporte público, semáforos, entre otros.

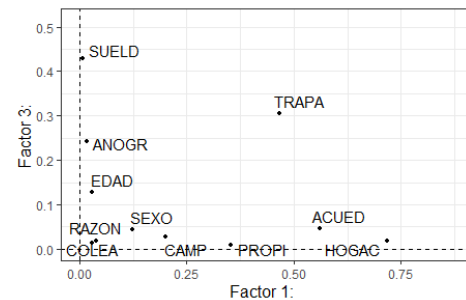
Otros ejes factoriales y variables ilustrativas

Anteriormente mencionamos que la interpretación de los resultados podrían hacerse con solo el primer eje factorial, no obstante, hay algunos resultados interesantes sobre otros ejes factoriales. El segundo eje factorial diferencia los individuos que asisten o no al colegio, como se ve en el primer plano factorial, ubicando los individuos que sí lo hacen en la parte superior del plano factorial y abajo los que no, a pesar de que la variable Nivel Educativo Alcanzado también está correlacionada con el segundo eje factorial, está confundida por la baja frecuencia de la categoría Otro.

El tercer eje factorial separa los individuos por el sueldo, aquellos individuos que ganan más de 500 mil pesos se encuentran ubicados en la parte superior del plano factorial, y los que menos de esa cantidad en la parte inferior.



Plano factorial 1:3

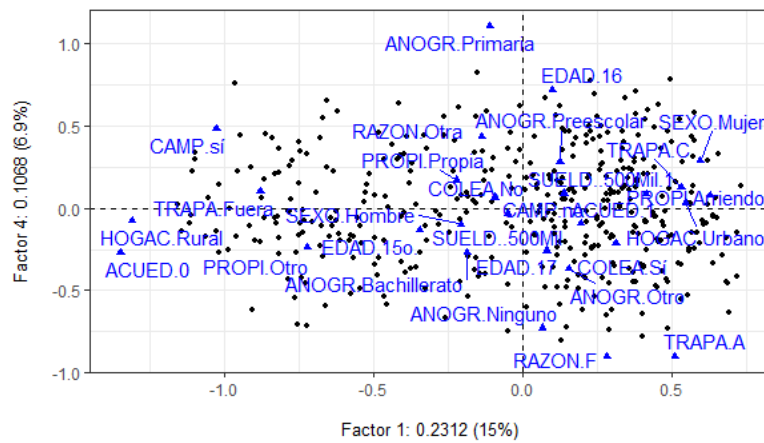


Razón de correlación con el plano 1:3

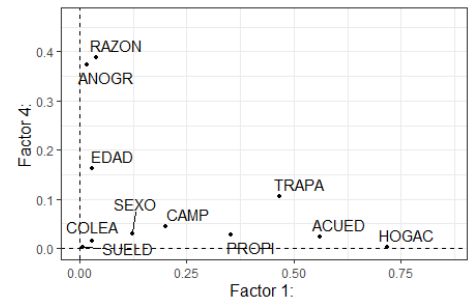
Figura 6: Plano factorial 1:3

El cuarto eje factorial separa las razones por las que trabajan los individuos, ubicando a los infantes que trabajan por razones familiares en la parte inferior del plano, y en la parte superior aquellos que lo hacen por otras razones, esta categoría incluye motivos como el trabajar por iniciativa propia. La Figura 7 indica que el cuarto eje también está muy correlacionado con la variable Nivel Educativo alcanzado, y separa los individuos con niveles educativos inferiores o iguales a primaria con los que tienen bachillerato, ninguno u otro. A pesar de que estas dos variables estén

correlacionadas con el cuarto eje, no parece haber una relación clara entre ellas.



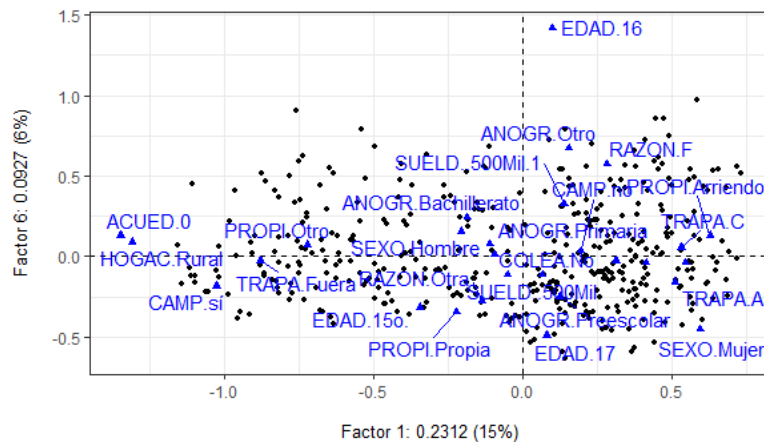
Plano factorial 1:4



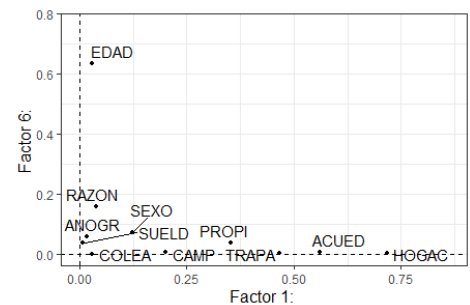
Razón de correlación con el plano 1:4

Figura 7: Plano factorial 1:4

Finalmente, el sexto eje factorial está bastante correlacionado con la edad de los individuos, este separa a los individuos con 16 años de los demás.



Plano factorial 1:6



Razón de correlación con el plano 1:6

Figura 8: Plano factorial 1:6

Las variables ilustrativas se encargaron de reforzar las características de las viviendas de los individuos, por ejemplo, las categorías relacionadas con la ausencia de servicios públicos como la luz, alcantarillado, gas, o recolección de basuras, se encuentran localizadas en el lado izquierdo del primer eje, lugar en el que se encontraban anteriormente las viviendas de clase rural. Lo mismo sucede con el lado derecho del eje, viviendas urbanas que sí gozan de estos servicios. Otra categoría

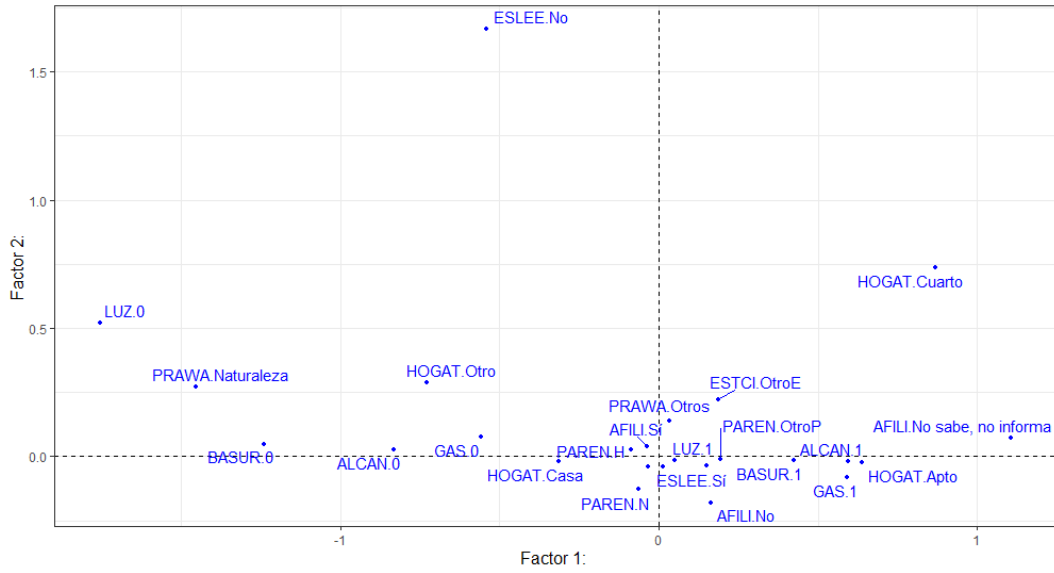


Figura 9: Variables ilustrativas sobre plano 1:2

que estuvo muy relacionada con el eje 1 fue el tipo de hogar, específicamente las categorías Apartamento y Cuarto, estos se encuentran ubicados en el lado derecho, indicando que el tipo de vivienda que habitan los infantes que trabajan en la ciudad suelen ser habitaciones o apartamentos, por el otro lado del eje se tiene que los individuos localizados en áreas rurales viven en casas u otro tipo de vivienda.

Agrupamiento

El siguiente apartado abarca la segmentación de los individuos en grupos caracterizados, se segmentarán en cuatro grupos con el fin de observar y caracterizar las poblaciones resultantes, si bien el gráfico de los índices indica que la segmentación podría hacerse en dos grupos, los objetivos de este trabajo tienen el fin de encontrar subpoblaciones vulnerables.

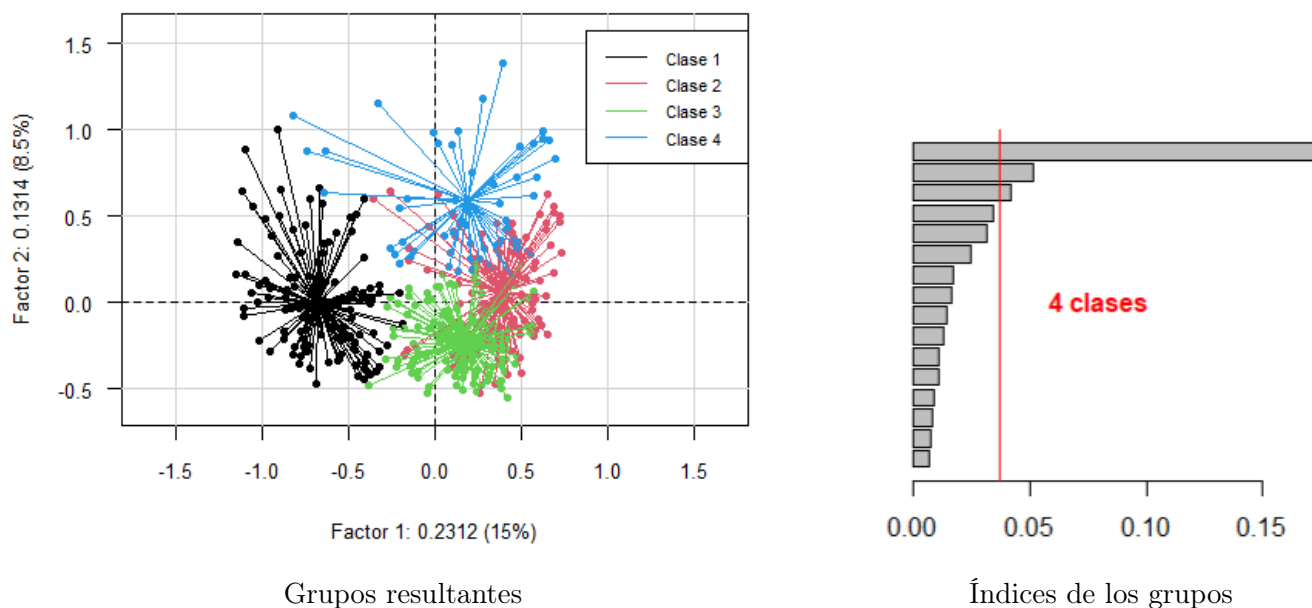


Figura 10: Resultados agrupamiento

	Tamaño antes	Tamaño después	Inercia Antes	Inercia Después
1	152.00	135.00	0.15	0.12
2	215.00	128.00	0.17	0.09
3	84.00	182.00	0.05	0.12
4	55.00	61.00	0.05	0.061
TOTAL	506.00	506.00	0.42	0.39

Luego de haber realizado el método de segmentación por medio de la combinación de k-medias y el método de Ward, las clases resultantes son las siguientes. La clase 1 se caracteriza por contener individuos que habitan hogares rurales, en casas que no cuentan con servicios públicos, el acceso al agua de estos individuos es principalmente por la naturaleza, es decir, obtienen el agua para consumo de ríos, lagunas, etc. También trabajan en su mayoría en lugares fuera de la ciudad, además de no poder asistir al colegio, en su mayoría son hombres campesinos.

Por otro lado, las 3 clases restantes se caracterizaron por tener a los individuos que habitan hogares urbanos con servicios públicos, no son campesinos en su gran mayoría, suelen vivir en apartamentos o habitaciones arrendadas, trabajan en establecimientos, o como trabajadores ambulantes en las calles. La diferencia entre estos tres grupos radica en el sueldo que reciben, aquellos individuos de la clase 2 reciben un sueldo superior a 500 mil pesos y asisten al colegio, los de la clase 3 menos de 500 mil y no asisten al colegio. La cuarta clase es la que contiene a los individuos más vulnerables, esta; además de contener a los infantes que asisten al colegio y reciben menos de 500 mil pesos como sueldo, tienen edades inferiores a los 15 años.

8. Conclusiones:

En línea con los objetivos planteados, se pueden ver disparidades significativas entre los niños que trabajan y aquellos cuya actividad principal es el estudio. Los primeros, mayoritariamente

hombres mayores de 15 años, tienden a residir en áreas urbanas sin su padre o sin su madre y, aunque no asisten al colegio, muestran suficiencia en lectura y escritura básica. Por otro lado, los niños que estudian, más comúnmente ubicados en entornos urbanos, viven con sus padres o abuelos, asisten al colegio y disfrutan de acceso a servicios esenciales. Además, se identifican subgrupos entre los niños que trabajan, donde aquellos en áreas rurales enfrentan condiciones más desafiantes en comparación con los que trabajan en entornos urbanos, algunos de los cuales reciben sueldos superiores a 500 mil pesos y asisten al colegio.

Así pues, el grupo de infantes más vulnerable, correspondería a aquellos que superan los 14 años, residen en áreas rurales, no se encuentran estudiando, y no tienen acceso a servicios básicos, teniendo en cuenta que en algunos casos también reciben sueldos inferiores a 500 mil, lo que lleva a pensar que también podrían ser víctimas de otros tipos de explotaciones producto de sus condiciones de precariedad.

Con lo anterior expuesto, se hace necesario estudiar estrategias que permitan disminuir eficientemente la tasa de trabajo infantil en el país, y aunque por cuestiones de alcance y falta de conocimiento disciplinar respecto a ámbitos sociales, este proyecto no tiene como objetivo proporcionar dichas estrategias, se recomendaría plantear iniciativas gubernamentales dirigidas a mitigar el trabajo infantil mediante la aproximación a mejorar algunas de las características asociadas al trabajo infantil, como lo son acceso a servicios públicos, oportunidades de estudio, etc. Programas que ofrezcan la oportunidad a estos infantes de estudiar, ya sea brindando horarios más flexibles, o ayudas económicas que les permitan subsistir por sí mismos, también podrían ser una opción a tener en cuenta. Por último, se hace imperioso evaluar la capacidad de programas gubernamentales como Familias en Acción y demás, que pueden llegar a ayudar a mitigar bastante la situación de muchas familias de bajos recursos del país, lo que ayuda a disminuir la necesidad de laburar de muchos infantes, por lo que corroborar que se distribuyan de manera correcta y eficiente los recursos destinados a estos programas, además de evaluar la cobertura que tienen estos sobre las familias más vulnerables, puede representar otra estrategia para mitigar el trabajo infantil en el país, pues la necesidad de abordar las condiciones desfavorables, especialmente en áreas rurales, emerge como una prioridad para promover el bienestar y la igualdad de oportunidades educativas en Colombia.

9. Bibliografía

- Carmen Y. (2019). Caracterización de trabajo infantil en 2018 en Colombia. Recuperado de <https://repositorio.unicartagena.edu.co/bitstream/handle/11227/12410/CARACTERIZACION%20DEL%20TRABAJO%20INFANTIL%20DEL%202018%20EN%20COLOMBIA.pdf?sequence=1&isAllowed=y>.
- DANE. (2022). Boletín técnico. Trabajo infantil TI. Recuperado de https://www.dane.gov.co/files/investigaciones/boletines/ech/trabajo-infantil/bol_trab_inf_oct2021-dic2021.pdf.
- Pardo C. 2020. Estadística Descriptiva Multivariada. Bogotá, Colombia. Universidad Nacional de Colombia.
- Departamento Nacional de Estadística. 2022. Módulo de Trabajo Infantil. <https://microdatos>.

[dane.gov.co/index.php/catalog/791/data-dictionary/F8?file_name=total_tri_4_2022](https://microdatos.dane.gov.co/index.php/catalog/791/data-dictionary/F8?file_name=total_tri_4_2022).

- Departamento Administrativo Nacional de Estadística. 2022. Gran Encuesta Integrada de Hogares. Departamento Administrativo Nacional de Estadística. 2022. Módulo de Trabajo Infantil. https://microdatos.dane.gov.co/index.php/catalog/791/data-dictionary/F8?file_name=total_tri_4_2022.