

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE POSGRADO**

PROYECTO DE TITULACIÓN

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

**“MAGÍSTER EN ESTADÍSTICA MENCIÓN EN GESTIÓN DE LA
CALIDAD Y PRODUCTIVIDAD”**

TEMA:

**FACTORES QUE INFLUYEN EL RENDIMIENTO ACADÉMICO DEL
ESTUDIANTE POR ETNIA: UN ENFOQUE MULTIVARIANTE**

AUTOR:

PAÚL ALBERTO JINEZ LLANGARI

Guayaquil - Ecuador

2020

RESUMEN

En Ecuador existe discriminación en la vida diaria, tales como insultos, humillación y maltratos por el color de la piel o la forma de vestir en pleno siglo XXI; los grupos étnicos que han sufrido mayor desigualdad son los pueblos y nacionalidades indígenas y afrodescendientes (El telégrafo, 2017); la educación es herramienta fundamental para reducir la desigualdad y promover el bienestar. Es la principal forma de inclusión social en las sociedades modernas, da acceso a una vida digna y ayuda a practicar los derechos sociales básicos. En la presente investigación se identificó cuáles son los factores que influyen el rendimiento académico del estudiante por etnia. Se tomará como fuente de información la evaluación denominada “Ser Bachiller” del ciclo 2017 – 2018 que realiza el Instituto Nacional de Evaluación Educativa (INEVAL) del Ecuador. Con los datos obtenidos se aplicó Análisis de Componentes Principales No Lineales que redujo de 109 a 54 variables sin perder los patrones. Para establecer que existe diferencia en el rendimiento entre los grupos étnicos a través de una Prueba de Hipótesis. Se identificó que uno de los factores más importantes es el acompañamiento que ejercen los padres hacia los estudiantes con Análisis de Correspondencia y Árbol de decisión.

ABSTRACT

In Ecuador there is discrimination in daily life, such as insults, humiliation and mistreatment because of the color of the skin or the way of dressing in the 21st century; The ethnic groups that have suffered the greatest inequality are the indigenous and Afro-descendant folk and nationalities (El telégrafo, 2017), Education is a fundamental tool to reduce inequality and promote well-being. It is the main form of social inclusion in modern societies, gives access to a decent life and helps practice basic social rights. In the present investigation it was identified what are the factors that influence the student's academic performance by ethnicity. The so-called "Bachiller" evaluation of the 2017 - 2018 cycle carried out by the National Institute of Educational Evaluation (INEVAL) of Ecuador will be taken as a source of information. With the data obtained, Analysis of Nonlinear Main Components was applied, which reduced from 109 to 54 variables without losing the patterns. To establish that there is a difference in performance between ethnic groups through a Hypothesis Test. It was identified that one of the most important factors is the support that parents exert towards students with Correspondence Analysis and Decision Tree..

DEDICATORIA

Este trabajo le dedico a Dios por ser mi fortaleza y protección en todas las etapas de mi vida, a mi mamá y mi hermano por ser un pilar fundamental para conseguir mis objetivos, a los docentes de la institución por haber compartido sus conocimientos.

AGRADECIMIENTO

Agradezco a la Escuela Superior Politécnica del Litoral abrimme las puertas y ser fundamental en mi mejora profesional, A los docentes por compartir su conocimiento, a mi madre y hermano por su ayuda constante.

DECLARACIÓN EXPRESA

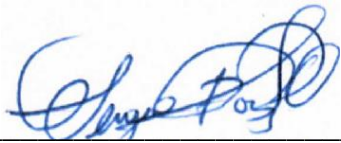
La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación, me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo, corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

A handwritten signature in black ink, enclosed within a light gray oval border. The signature is written in a cursive style and appears to read 'Paúl Alberto Jinez Llangari'.

Paúl Alberto Jinez Llangari

TRIBUNAL DE GRADUACIÓN



P.HD. Sergio Bauz Olvera
PRESIDENTE



Mgtr. Heydi Roa López
DIRECTOR



M.Sc. Francisco Moreira Villegas
VOCAL 1



Ph.D. Johny Pambabay Calero
VOCAL 2

ABREVIATURAS O SIGLAS

ACM: Análisis de Correspondencia Múltiple

ACP: Análisis de Componentes Principales

ACPNL: Análisis de Componentes Principales No Lineales

CART: Árbol de Clasificación y Regresión (Classification and Regression Trees)

INEVAL: Instituto Nacional de Evaluación Educativa

ODS: Objetivos de Desarrollo Sostenible

ONU: Organización de las Naciones Unidas

TABLA DE CONTENIDO

CAPÍTULO 1	1
1. INTRODUCCIÓN.....	1
1.1. Antecedentes	1
1.2. Descripción del problema.....	2
1.3. Objetivos	3
Objetivo General	3
Objetivo Específicos	3
1.1. Hipótesis	3
1.2. Alcance.....	3
CAPÍTULO 2	5
2. MARCO TEÓRICO.....	5
2.1. Análisis de Componentes Principales No Lineales	5
2.2. Análisis de correspondencias.....	7
2.3. Árboles de decisión.....	13
2.3.1. Elementos del árbol.....	14
2.3.2. Proceso de análisis de árboles de decisión	15
2.3.3. Construcción del árbol máximo	15
2.3.4. Calidad del nodo: Función de impureza	15
2.3.5. Poda del árbol	16
2.3.6. Selección del árbol óptimo	17
CAPÍTULO 3	19
3. METODOLOGÍA.....	19
3.1. Determinación de la existencia de diferencias en el rendimiento académico de los grupos étnicos en el ciclo 2017 - 2018.....	19
3.1.1. Control de factores	19
3.1.2. Prueba de hipótesis.....	19
Paso 1: Supuestos.....	20
Paso 2: Establecimiento de la hipótesis.....	20
Paso 3: Nivel de significancia	20
Paso 4: Estadístico de prueba	20
Paso 5: Regla de decisión	21
Paso 6: Cálculo	21
Paso 7: Conclusiones y recomendaciones	21

3.2. Reducción de la dimensionalidad del conjunto de datos de la evaluación del ciclo 2017 – 2018 a través de las variables asociadas a la etnia para facilidad de análisis.....	21
3.3. Identificación de las posibles causas o factores que influyen en los niveles de logro de los grupos étnicos en el ciclo 2017 – 2018.....	22
3.4. Construcción de un modelo estadístico que permita el pronóstico del rendimiento académico de los grupos étnicos en el ciclo 2017 – 2018.....	23
CAPÍTULO 4	24
4. RESULTADOS.....	24
4.1. Determinación de la existencia de diferencias en el rendimiento académico de los grupos étnicos en el ciclo 2017 - 2018.....	24
4.1.1. Prueba de hipótesis.....	24
Paso 1: Supuestos.....	24
Paso 2: Establecimiento de las hipótesis.....	26
Paso 3: Nivel de significancia	26
Paso 4: Estadístico de prueba	27
Paso 5: Regla de decisión	27
Paso 6: Cálculo	27
Paso 7: Conclusiones y recomendaciones	28
4.2. Reducción de la dimensionalidad del conjunto de datos de la evaluación del ciclo 2017 – 2018 a través de las variables asociadas a la etnia para un eficiente análisis	28
4.2.1. Análisis de consistencia de la información	28
4.2.2. Reducción de dimensiones	29
4.3. Identificación de las posibles causas o factores que influyen en los niveles de logro de los grupos étnicos en el ciclo 2017 – 2018.....	33
4.3.1. Factores que influyen en el rendimiento académico de la etnia blanco/mestizo.....	35
4.3.2. Factores que influyen en el rendimiento académico de la etnia indígena	35
4.3.3. Factores que influyen en el rendimiento académico de la etnia afroecuatoriano.....	36
4.3.4. Factores que influyen en el rendimiento académico de la etnia montubio	37
4.4. Construir un modelo estadístico que permita el pronóstico del rendimiento académico de los grupos étnicos en el ciclo 2017 – 2018.....	37
4.4.1. Árbol de decisión de los factores que influyen en el rendimiento académico del grupo étnico blanco/mestizo.....	41
4.4.2. Árbol de decisión de los factores que influyen en el rendimiento académico del grupo étnico montubio	45

4.4.3. Árbol de decisión de los factores que influyen en el rendimiento académico del grupo étnico indígena y afroecuatoriano	48
CAPÍTULO 5	50
5. CONCLUSIONES Y RECOMENDACIONES.....	50
5.1. Conclusiones.....	50
5.2. Recomendaciones.....	51
6. Referencias	52
7. Apéndices y anexos	1
Matriz de carga de variables rotada.....	1

LISTADO DE FIGURAS

Ilustración 1: Elementos del árbol de decisión	14
Ilustración 2: Diagrama de flujo del algoritmo CART	18
Ilustración 3: Gráfico de Correspondencia Múltiple	34
Ilustración 4: Gráfico de correspondencia Múltiple para etnia blanco/mestizo	35
Ilustración 5: Gráfico de correspondencia múltiple del grupo étnico indígena	36
Ilustración 6: Gráfico de correspondencia múltiple del grupo étnico afroecuatoriano.....	36
Ilustración 7: Gráfico de correspondencia múltiple del grupo étnico montubio	37
Ilustración 8: Diagrama de Árbol de Decisión grupo étnico blanco/mestizo.....	38
Ilustración 9: Diagrama de Árbol de Decisión de grupo étnico montubio, indígena y afroecuatoriano	43

LISTADO DE TABLAS

Tabla 1: Tabla de contingencia	8
Tabla 2. Frecuencia marginal de la variable X	8
Tabla 3: Frecuencia marginal de la variable Y	9
Tabla 4: Tabla de correspondencias de filas	9
Tabla 5: Tabla de correspondencias de columnas	10
Tabla 6: Prueba de normalidad de los datos	25
Tabla 7: Prueba de homogeneidad de varianzas	25
Tabla 8: Comparación de poblaciones.....	27
Tabla 9: Resumen del modelo de ACPNL.....	29
Tabla 10: Resumen del modelo ACPNL rotado	30
Tabla 11:Carga de componentes rotados	31
Tabla 12: Resumen del modelo de Análisis de Correspondencia Múltiple	33
Tabla 13: Árbol de decisión del grupo étnico blanco/mestizo.....	39
Tabla 14: Árbol de decisión del grupo étnico montubio	44
Tabla 15: Árbol de decisión del grupo étnico indígena y afroecuatoriano	46

CAPÍTULO 1

1. INTRODUCCIÓN

1.1. Antecedentes

La educación es una herramienta fundamental para reducir la desigualdad y promover el bienestar. Es la principal forma de inclusión social en las sociedades modernas, da acceso a una vida digna y ayuda a practicar los derechos sociales básicos, debido a esto, la Asamblea General de las Naciones Unidas en el año de 1948 aprobó la declaración universal de derechos humanos, donde el artículo 26 reconoce a la educación como un derecho humano; en relación a la igualdad, el inciso dos menciona:

“La educación tendrá por objeto el pleno desarrollo de la personalidad humana y el fortalecimiento del respeto a los derechos humanos y a las libertades fundamentales; favorecerá la comprensión, la tolerancia y la amistad entre todas las naciones y todos los grupos étnicos o religiosos; y promoverá el desarrollo de las actividades de las Naciones Unidas para el mantenimiento de la paz.” (UNESCO, 2005)

Esta declaración otorga el derecho a tener educación accesible, eliminando la discriminación en el acceso a la educación, ya sea por etnia, sexo o raza, y que las minorías elijan sus actividades curriculares en su propio idioma como política de estado.

La Organización de las Naciones Unidas en el año 2015 aprobó la Agenda 2030 para el Desarrollo Sostenible con la finalidad de brindar oportunidades para que los países y sociedades mejoren la calidad de vida de la población suscrita sin dejar a nadie atrás, en esta constan 17 Objetivos de Desarrollo Sostenible (ODS). La educación tiene su objetivo exclusivo que es ODS 4 que dice: *“Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos”*. (ONU, 2015)

La meta 4.5 indica que:

“De aquí a 2030, eliminar las disparidades de género en la educación y asegurar el acceso igualitario a todos los niveles de la enseñanza y la formación profesional para las personas vulnerables, incluidas las personas con discapacidad, los pueblos indígenas y los niños en situaciones de vulnerabilidad” (ONU, 2015)

Con el cumplimiento de la meta 4.5 de los ODS se garantiza tener una educación igualitaria de libre acceso para grupos étnicos que se encuentran en situación de vulnerabilidad.

En el Ecuador, en la década de los 60 se desarrolló una gran reforma educativa, teniendo como meta el acceso universal a la enseñanza básica que era liderada por el estado y la sociedad. Esto produjo un crecimiento importante del Sistema Educativo Público tuvo como efecto la movilidad social de los ecuatorianos, aumentando el mercado de trabajo y la distribución de la riqueza, también la implementación de un modelo social, así como difundir códigos de la modernidad, ciudadanía, generación de ingresos y productividad laboral, pero esto fue favorable para determinadas etnias, debido a que el acceso beneficiaba a determinados grupos sociales que tenían hegemonía.

La educación ha tenido un significativo progreso en la última década, pero todavía persisten los problemas de igualdad en la población ecuatoriana. Tener una educación de calidad aún depende de la situación socioeconómica, condición étnica, lugar de residencia, sexo y la edad. La población de área urbana se incorporó masivamente al sistema educativo teniendo como consecuencia una vía de ascenso social, mientras que en la población rural, especialmente en la campesina e indígena, todavía persiste la falta de recursos y oportunidades para acceder a una educación de calidad.

1.2. Descripción del problema

La educación es uno de los pilares fundamentales para el desarrollo de un país, debido a esto la ONU ha planteado Objetivos del Desarrollo Sostenible (ODS) orientados a la educación como el objetivo 4 que dice “Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos” (ONU, 2015), pero según las proyecciones realizadas por la Unesco menciona que no se alcanzará el objetivo al 2030, debido a que ha pasado un tercio del tiempo planificado. Algunos resultados obtenidos fueron que 1 de cada 6 niños de la edad comprendida entre 6 y 17 años estarán excluidos de la escuela, mientras que 6 de cada 10 jóvenes completarán la educación secundaria (Unesco, 2019).

En el Ecuador existe discriminación en la vida diaria, tales como insultos, humillación y maltratos por el color de la piel o la forma de vestir en pleno siglo XXI; los grupos étnicos que han sufrido mayor desigualdad son los pueblos y nacionalidades indígenas y afrodescendientes. En el censo de población y vivienda realizado en el 2001 se identifica que existen 830.418 personas pertenecientes a pueblos o nacionalidades, donde 14 son nacionalidades y 18 son de pueblos indígenas. En estos últimos 30 años se ha realizado una importante inversión en el sistema de educación intercultural bilingüe y además se ha mejorado el acceso al Sistema Nacional de Educación. Uno de los problemas esenciales en la educación es disminuir la tasa de analfabetismo.

La presente investigación pretende identificar los factores que intervienen en la diferencia de puntajes entre etnias y a su vez, sirva para generar lineamientos en proyectos que mitiguen la discriminación y, por lo tanto, sea una fuente de información que sirva de apoyo a la toma de decisiones.

1.3. Objetivos

Objetivo General

Determinar la influencia de los factores asociados de los estudiantes en el rendimiento académico de los grupos étnicos en el ciclo 2017 – 2018, a través de un análisis multivariante para la toma de decisiones en proyectos orientados al mejoramiento del sistema educativo.

Objetivo Específicos

- a. Determinar la existencia de diferencias en el rendimiento académico de los grupos étnicos en el ciclo 2017 - 2018.
- b. Reducir la dimensionalidad del conjunto de datos de la evaluación del ciclo 2017 – 2018 a través de las variables asociadas a la etnia para un eficiente análisis.
- c. Identificar las posibles causas o factores que influyen en los niveles de logro de los grupos étnicos en el ciclo 2017 – 2018.
- d. Construir un modelo estadístico que permita el pronóstico del rendimiento académico de los grupos étnicos en el ciclo 2017 – 2018.

1.1. Hipótesis

La hipótesis de investigación de factores que influyen el rendimiento académico del estudiante por etnia que se plantea en este estudio es:

La existencia de diferencia en el rendimiento académico de los grupos étnicos permitirá determinar los factores que inciden en la problemática.

1.2. Alcance

La presente investigación explora el rendimiento académico de los grupos étnicos a través factores asociados que tiene el estudiante a través de análisis multivariantes adecuados para esta investigación.

Se explican las causas que intervienen en el rendimiento académico con información que dispone el Instituto Nacional de Evaluación Educativa en el Ecuador (Ineval).

Esta investigación contiene cuatro etapas. En la primera etapa se busca el cumplimiento del primer objetivo de la investigación, que es el de identificar la existencia de diferencias en el rendimiento académico de los grupos étnicos involucrados en el estudio, donde se realiza la aplicación de análisis de varianza de dos vías, considerando a la variable etnia como el tratamiento a 8 niveles

("Afrodescendiente, Mulato, Negro, Montubio, Indígena, Mestizo, Blanco, Otro"), y como bloque, a la variable sexo a dos niveles (hombre y mujer).

La segunda etapa es discriminar las variables de la base de datos de factores asociados del estudiante; primero se identifica la consistencia de información de cada una de ellas, para luego realizar un Análisis de Componentes Principales no lineales que permita reducir las dimensiones para poder tener un entendimiento de la problemática sin perder las propiedades de la base de datos.

La tercera etapa es la identificación de las posibles causas que ocurren en la mencionada problemática; la técnica estadística a utilizar es Análisis de Correspondencia Múltiple (ACM) (Greenacre, 2008). Esta técnica brinda un gráfico que permite identificar la relación que tienen los atributos.

Por último, se construye un modelo estadístico que permita realizar inferencias; la técnica estadística a utilizar es Árbol de Decisión con el Método CART, donde se puede cuantificar la relación que tienen las variables de estudio en función del rendimiento académico (Pérez , Grandón, Caniupán , & Vargas , 2018).

Para dar por finalizada la investigación, se realizan las conclusiones y recomendaciones.

CAPÍTULO 2

2. MARCO TEÓRICO

En el desarrollo del presente estudio se eligieron varias técnicas estadísticas con el fin de entender los factores que influyen en el rendimiento académico de los diferentes grupos étnicos. A continuación, se presenta un breve resumen de los principales métodos necesarios para construir los resultados obtenidos en la investigación.

2.1. Análisis de Componentes Principales No Lineales

Existe una diferencia entre el Análisis de Componentes Principales (ACP) convencional y el Análisis de Componentes Principales No Lineales (ACP NL), el primero asume relaciones lineales entre las variables cuantitativas, y el segundo escala las variables categóricas a diferentes niveles, se cuantifican óptimamente en la dimensión seleccionada para poder modelar las relaciones no lineales que existen entre las variables

De la tesis “El escalonamiento óptimo con base en el Análisis de Componentes Principales No Lineales (ACP NL) para la construcción de índices de condiciones de vida y socioeconómicos” (Tapia, 2007), se obtiene lo siguiente:

El Análisis de Componentes Principales es una técnica estadística que analiza la covarianza o la correlación entre un número de variables que se asume miden un constructo y que permite reducir el conjunto original de variables en uno más pequeño, mediante la estimación de componentes no correlacionados que capturan la mayor parte de la información contenida en las variables originales.

Cuando se cuenta con variables cualitativas o categóricas, por ejemplo, el nivel de satisfacción con cierto aspecto o variables que califican (no cuantifican) como la “calidad percibida” el análisis de componentes clásico no es aplicable. El análisis de componentes principales no lineales fue diseñado para dar una respuesta a este tipo de situaciones asignando valores a las categorías de cada una de las variables y luego correlacionarlas para caracterizar o analizar la estructura de los datos. Este modelo multivariante se lo conoce también por los nombres de:

- Análisis de componentes principales categóricos por escalamiento óptimo,
- Análisis de componentes principales no métrico, o
- Análisis de componentes principales mediante mínimos cuadrados alternantes.

La valoración de las variables cualitativas con este procedimiento logra también maximizar la correlación lineal entre las variables dadas y transformar las variables cualitativas en variables cuantitativas, disponiendo así de la mejor combinación lineal posible de las variables tratadas. Esto sirve a su vez, para la construcción de modelos de regresión (en donde interesa precisamente la linealidad), o poder discriminar o caracterizar mejor los casos ambiguos, entre muchos otros análisis que se podrían efectuar.

A continuación, se realiza una síntesis de algunos aspectos básicos que sirven para la construcción de esta técnica estadística.

Una de las utilidades del Análisis de Componentes Principales No Lineales es la reducción de datos, no necesita supuestos de normalidad, se estudia p variables observables, $X_1, X_2, X_3, \dots, X_p$ a partir de estas se generan k variables no observables donde $k < p$.

$X^t = [X_1 \ X_2 \ \dots \ X_p]$, es el vector aleatorio que varía en p , con matriz de varianzas y covarianzas Σ , donde $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p \geq 0$ son valores propios y $a_1, a_2, a_3, \dots, a_p$, son vectores propios correspondientes a Σ , también se considera las combinaciones lineales

$$\begin{aligned} Y_1 &= a_1^t X = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \\ Y_2 &= a_2^t X = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \\ &\vdots \\ Y_p &= a_p^t X = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_p \end{aligned}$$

Se tiene:

$$Var(Y_i) = a_i^t \Sigma a_i \quad \text{con } i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_j) = a_i^t \Sigma a_j \quad \text{con } i, j = 1, 2, \dots, p$$

Donde Y_1, Y_2, \dots, Y_p son las combinaciones lineales de los componentes principales que no tienen correlación entre si y $Var(Y_1) \geq Var(Y_2) \geq Var(Y_3) \geq \dots \geq Var(Y_p) \geq 0$.

Cada componente principal se define como:

La primera componente principal es igual a la combinación lineal $Y_1 = a_1^t X$ que maximiza $Var(a_1^t X)$ está sujeto a $\langle a_1, a_1 \rangle = 1$.

La segunda componente principal es igual a la combinación lineal $Y_2 = a_2^t X$ que maximiza $Var(a_2^t X)$ está sujeto a $\langle a_2, a_2 \rangle = 1$ y $Cov(Y_1, Y_2) = 0$.

La tercera componente principal es igual a la combinación lineal $Y_3 = a_3^t X$ que maximiza $Var(a_3^t X)$ está sujeto a $\langle a_3, a_3 \rangle = 1$ y $Cov(Y_2, Y_3) = 0$.

En general la i -ésima componente principal es igual a la combinación lineal $Y_i = a_i^t X$ que maximiza $Var(a_i^t X)$ está sujeto a $\langle a_i, a_i \rangle = 1$ y $Cov(Y_i, Y_j) = 0$.

Generalizando el método se tiene que; la matriz de varianzas y covarianzas Σ , asociada al vector $X^t = [X_1 \ X_2 \ \dots \ X_p] \in R^p$ y sean los valores y vectores propios $(\lambda_1, a_1), (\lambda_2, a_2), (\lambda_3, a_3), \dots, (\lambda_p, a_p)$ que corresponden a la matriz Σ donde $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_p \geq 0$ entonces la i -ésima componente principal es:

$$Y_i = a_i^t X = a_{i1}X_1 + a_{i2}X_2 + \dots + a_{ip}X_p \text{ con } i = 1, 2, \dots, p$$

Cumpliendo con:

$$Var(Y_i) = a_i^t \Sigma a_i \text{ con } i = 1, 2, \dots, p$$

$$Cov(Y_i, Y_j) = a_i^t \Sigma a_j = 0 \text{ con } i \neq j$$

La proporción de varianza explicada por los componentes principales está dada por:

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + \lambda_3 + \dots + \lambda_k}, \text{ para } k = 1, 2, 3, \dots, p$$

Se escogerá el número de componentes principales en función al porcentaje de varianza que se desee explicar, con esto se cuantifica las variables categóricas y también ayuda a reducir la dimensionalidad cuando tienes un número grande de variables, por lo tanto, es difícil realizar una interpretación de las relaciones entre objetos, al reducir la dimensionalidad a un pequeño número de componentes será más fácil realizar análisis y obtener mejores conclusiones.

2.2. Análisis de correspondencias

El objetivo del análisis de correspondencias “es resumir una gran cantidad de datos en un número reducido de dimensiones, con la menor pérdida de información posible” (De la Fuente Fernández, 2011). Esta técnica se aplica solo a variables cualitativas, ya sean categóricas u ordinales.

El Análisis de Correspondencias Simples se construye a partir de tablas de contingencia (bidimensional), esto significa que las categorías de la variable se presentan en las filas y en las columnas se presentan las categorías de otra variable y se cuenta el número de ocurrencias que contengan cada una de las intersecciones, de modo que se puedan identificar las relaciones que existen entre las dos variables (De la Fuente Fernández, 2011).

Cuando se analizan más de dos variables (n-dimensional), este análisis se transforma en un análisis de correspondencia múltiple, pero se mantiene el principio del análisis de correspondencia simple.

El análisis de correspondencia permite visualizar las relaciones de dependencia e independencia de un conjunto de variables categóricas.

La tabla de contingencia se representa en forma de matriz donde X e Y son variables categóricas con valores (x_1, \dots, x_k) y (y_1, \dots, y_m) respectivamente. Cada una de las variables tienen N elementos de una población, al cruzar estas dos variables se produce una intersección entre las filas y columnas que es una celda donde se encuentra la frecuencia observada n_{ij} . De igual forma se puede obtener la frecuencia esperada. Estos dos tipos de frecuencia son importantes para el cálculo de la distribución chi-cuadrado.

Tabla 1: Tabla de contingencia

	Y_1	Y_2	Y_j	Y_m	
X_1	n_{11} (e_{11})	n_{12} (e_{12})	n_{1j} (e_{1j})	n_{1m} (e_{1m})	N_{1*}
X_2	n_{21} (e_{21})	n_{22} (e_{22})	n_{2j} (e_{2j})	n_{2m} (e_{2m})	N_{2*}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
X_i	n_{i1} (e_{i1})	n_{i2} (e_{i2})	n_{ij} (e_{ij})	n_{im} (e_{im})	N_{i*}
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
X_k	n_{k1} (e_{k1})	n_{k2} (e_{k2})	n_{kj} (e_{kj})	n_{km} (e_{km})	N_{k*}
	N_{*1}	N_{*2}		N_{*j}		N_{*m}	N_{1*}

Fuente: Análisis de Correspondencias Simples y Múltiples.
Elaborado por: De la Fuente Santiago

Se presentan las ecuaciones matemáticas que sirven para contabilizar cada celda y el total.

$$N_{i*} = \sum_{j=1}^m n_{ij},$$

$$N_{*j} = \sum_{i=1}^k n_{ij},$$

$$N_{**} = \sum_i N_{i*} = \sum_j N_{*j},$$

$$e_{ij} = \frac{N_{i*} \cdot N_{*j}}{N_{**}}$$

Las frecuencias relativas de la distribución marginal de las variables X e Y, son:

Tabla 2. Frecuencia marginal de la variable X

$X = x_i$	x_1	x_2	...	x_i	...	x_k	Total
Frecuencias relativas marginales	$\frac{N_{1*}}{N_{**}}$	$\frac{N_{2*}}{N_{**}}$...	$\frac{N_{i*}}{N_{**}}$...	$\frac{N_{k*}}{N_{**}}$	1

Fuente: Análisis de Correspondencias Simples y Múltiples.
Elaborado por: De la Fuente Santiago

Tabla 3: Frecuencia marginal de la variable Y

$Y = y_j$	y_1	y_2	...	y_j	...	y_m	Total
Frecuencias relativas marginales	$\frac{N_{*1}}{N_{**}}$	$\frac{N_{*2}}{N_{**}}$...	$\frac{N_{*j}}{N_{**}}$...	$\frac{N_{*m}}{N_{**}}$	1

Fuente: Análisis de Correspondencias Simples y Múltiples.
Elaborado por: De la Fuente Santiago

La tabla de correspondencias permite describir las distribuciones condicionadas que se pueden obtener de Y en función de X (filas), así como la distribución condicionada de la variable X en función de Y (columnas).

La distribución condicionada de Y en función de X, o filas, se obtiene a través de las siguientes expresiones.

Tabla 4: Tabla de correspondencias de filas

$f(Y/X = x_i)$	Y_1	...	Y_j	...	Y_m	Total
X_1	$\frac{n_{11}}{N_{1*}}$...	$\frac{n_{1j}}{N_{1*}}$...	$\frac{n_{1m}}{N_{1*}}$	1
X_2	$\frac{n_{21}}{N_{2*}}$...	$\frac{n_{2j}}{N_{2*}}$...	$\frac{n_{2m}}{N_{2*}}$	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
X_i	$\frac{n_{i1}}{N_{i*}}$...	$\frac{n_{ij}}{N_{i*}}$...	$\frac{n_{im}}{N_{i*}}$	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮
X_k	$\frac{n_{k1}}{N_{k*}}$...	$\frac{n_{kj}}{N_{k*}}$...	$\frac{n_{km}}{N_{k*}}$	1

Fuente: Análisis de Correspondencias Simples y Múltiples.
Elaborado por: De la Fuente Santiago

Distancias chi-cuadrado en la distribución condicionada de filas:

$$d_{ij}^f = \sum_{h=1}^m \frac{1}{N_{*h}} \left[\frac{n_{ih}}{N_{i*}} - \frac{n_{jh}}{N_{j*}} \right]^2$$

La distribución condicionada de las columnas de X en función de Y se obtiene a partir de las siguientes expresiones.

Tabla 5: Tabla de correspondencias de columnas

$f(Y/Y = y_i)$	Y_1	Y_j	Y_m
X_1	$\frac{n_{11}}{N_{*1}}$	$\frac{n_{1j}}{N_{*j}}$	$\frac{n_{1m}}{N_{*m}}$
X_2	$\frac{n_{21}}{N_{*1}}$	$\frac{n_{2j}}{N_{*j}}$	$\frac{n_{2m}}{N_{*m}}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
X_i	$\frac{n_{i1}}{N_{*1}}$	$\frac{n_{ij}}{N_{*j}}$	$\frac{n_{im}}{N_{*m}}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
X_k	$\frac{n_{k1}}{N_{*1}}$	$\frac{n_{kj}}{N_{*j}}$	$\frac{n_{km}}{N_{*m}}$
Total	1	1	1	1	1

Fuente: Análisis de Correspondencias Simples y Múltiples.
Elaborado por: De la Fuente Santiago

Distancias chi cuadrado en la distribución condicionada de columnas:

$$d_{ij}^c = \sum_{h=1}^k \frac{1}{N_{*h}} \left[\frac{n_{ih}}{N_{*i}} - \frac{n_{jh}}{N_{*j}} \right]^2$$

La distancia chi-cuadrado pondera ya sea por fila o columna, por lo que cada una tiene un peso proporcional según la importancia en el conjunto. A este peso se le conoce como masa. Considerando la masa proporcional a su frecuencia se evita privilegiar a las categorías con poca ocurrencia.

El chi-cuadrado de filas o de columnas cumple con el principio de equivalencia distribucional. Esta dice que si dos categorías tienen perfiles idénticos pueden formar una sola categoría que sea la suma de sus pesos, que modifique la distancia entre filas y columnas. Esto significa que al unir categorías los resultados se mantienen.

Una de las funciones principales de la distribución chi-cuadrado es identificar la independencia o dependencia que existe entre las variables X e Y, esto significa que identifica la relación que existe entre las dos variables a través de un contraste de hipótesis, siendo estas:

H_0 : La variable X y Y son independientes

H_1 : La variable X y Y tienen relación de dependencia

Este contraste compara las filas y columnas con los perfiles marginales según corresponda, el estadístico observado es:

$$\chi_{(k-1)(m-1)}^2 = \sum_{i=1}^k \sum_{j=1}^m \frac{(n_{ij} - e_{ij})^2}{e_{ij}}$$

Donde:

$$e_{ij} = \frac{N_{i*} \cdot N_{*j}}{N_{**}}$$

El estadístico observado se puede expresar también mediante:

$$\chi^2_{(k-1)(m-1)} = \sum_{i=1}^k \sum_{j=1}^m \frac{(n_{ij} - e_{ij})^2}{e_{ij}} = \sum_{i=1}^k \sum_{j=1}^m \frac{N_{i*} \left[\frac{n_{ij}}{N_{i*}} - \frac{N_{*j}}{N_{**}} \right]^2}{\frac{N_{*j}}{N_{**}}} = \sum_{i=1}^k \sum_{j=1}^m \frac{N_{*j} \left[\frac{n_{ij}}{N_{*j}} - \frac{N_{i*}}{N_{**}} \right]^2}{\frac{N_{i*}}{N_{**}}}$$

La regla de decisión de la chi-cuadrado:

Si $\chi^2_{observado} > \chi^2_{tabulado}$, se rechaza H_0 o que es lo mismo decir,

Si $\sum_{i=1}^k \sum_{j=1}^m \frac{(n_{ij} - e_{ij})^2}{e_{ij}} > \chi^2_{\alpha, (k-1)(m-1)}$, entonces se rechaza H_0 ; por lo tanto se acepta la hipótesis alternativa que indica que las variables X e Y tienen relación de dependencia. En caso contrario se acepta la H_1 .

Con las filas y columnas descritas en los párrafos anteriores se elabora la matriz de coordenadas (distancias) mediante el método de la distancia chi-cuadrado, que permite calibrar la magnitud entre la tabla de datos y la tabla de datos sin relación entre las variables.

Para poder desarrollar este proceso se determinan los vectores singulares de la matriz.

$$C = (r_{ij}) \text{ siendo } r_{ij} = \frac{n_{ij} - e_{ij}}{\sqrt{e_{ij}}}$$

Para construir un sistema de coordenadas asociadas a las filas y columnas de la tabla de contingencia. Además, permitirá visualizar la relación que existe entre filas y columnas. Las distancias χ^2 entre filas y columnas son las que el análisis de correspondencia reproduce en unas gráficas, donde las distancias son pitagóricas ponderadas entre filas y columnas.

Distancias chi-cuadrado de filas:

$$d_{ij}^f = \sum_{h=1}^m \frac{1}{N_{*h}} \left[\frac{n_{ih}}{N_{i*}} - \frac{n_{jh}}{N_{*j}} \right]^2$$

Distancias chi-cuadrado de columnas:

$$d_{ij}^c = \sum_{h=1}^k \frac{1}{N_{*h}} \left[\frac{n_{ih}}{N_{i*}} - \frac{n_{jh}}{N_{*j}} \right]^2$$

Las distancias se miden con respecto a la relación media de las coordenadas de la fila o columna ponderada por su masa (peso proporcional según el conjunto de datos). Este aparecerá situado en el origen de coordenadas y se le conoce como centro de gravedad. "La media de las distancias al cuadrado de cada punto de fila

al centro de gravedad se conoce como inercia de filas, o inercia de columnas cuando se trata de las columnas, e inercia total de la nube de puntos cuando se consideran todos los elementos de la tabla” (De la Fuente Fernández, 2011). Cuando la inercia es baja significa que las categorías están muy cerca del centro de gravedad, mientras que, si la inercia es alta en las categorías, esto significa que existen grandes diferencias con respecto a las medias de las filas y columnas.

Luego se procede a calcular la matriz diagonal de varianzas - covarianzas para obtener los vectores y valores propios donde se definen los ejes en los que será proyectada la nube de puntos. Cuando la inercia explicada es alta en los primeros factores entonces se podrá realizar el gráfico en dos o tres dimensiones y con esto poder visualizar las relaciones que existen entre ellas.

El análisis de correspondencias busca obtener coordenadas cartesianas por lo que se trata de encontrar dos matrices que representen a los puntos fila y puntos columnas.

$$A = \begin{bmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_k \end{bmatrix}, \text{ representa a los puntos fila } a_i = (a_{i1}, \dots, a_{ih})'$$

$$B = \begin{bmatrix} b'_1 \\ b'_2 \\ \vdots \\ b'_k \end{bmatrix}, \text{ representa a los puntos fila } b_j = (b_{j1}, \dots, b_{jh})'$$

Existen varias maneras de calcular estas matrices, una de las más usadas es la normalización simétrica o canónica (ACC), que satisface que el producto escalar $(a_i * b_j)$ sea proporcional a los residuos tipificados (r_{ij}) .

El método de normalización simétrica o canónica descompone la matriz $C = r_{ij}$, en valores singulares calculando matrices $U_{k \times h}$, D y $V_{m \times h}$, con $H = \text{mínimo} \{k - 1, m - 1\}$, tal que $C = UDV'$, siendo $U'U = V'V = I$, $D = \text{diagonal}\{\mu_1, \mu_2, \dots, \mu_H\}$ donde $\mu_i \equiv \text{valores singulares } (i = 1, \dots, H)$.

Las matrices A y B se calculan a partir de las expresiones $A = D_k^{-1/2}UD$, y $B = D_m^{-1/2}VD$ con $D_k = \text{diagonal}\{n_1, \dots, n_k\}$, $D_m = \text{diagonal}\{n_1, \dots, n_m\}$.

Del documento Análisis de Correspondencia de López-Roldán & Fachelli, 2015 se obtiene la metodología de análisis de correspondencia múltiples.

El análisis de correspondencias múltiples es la aplicación del análisis de correspondencias simple al estudio de tablas lógicas, donde se considera un

número cualquiera de variables cualitativas, pero esta tiene procedimientos de cálculo y reglas de interpretación específicas.

Notación. Consideremos la matriz X:

c individuos ($i = 1, \dots, n$)

p variables cualitativas ($j = 1, \dots, p$)

Cada variable x_{+j} tiene c categorías (diferentes según la variable) que permiten descomponer la variable en tantas modalidades o categorías

Codificación disyuntiva completa: Si un individuo i tiene en la variable j la categoría $c = c_0$, entonces tendrá el valor 1 para esta categoría, x_{ijk} , y 0 para el resto de las categorías de la variable, $x_{ijk} = 0$, si $c \neq c_0$. Se obtiene así la Matriz o Tabla Disyuntiva

Matriz o tabla disyuntiva D (matriz lógica o binaria) asociada a la matriz de datos original:

$$X = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 3 & 2 \\ 2 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} \quad D = \begin{pmatrix} 10 & 100 & 01 \\ 01 & 010 & 10 \\ 10 & 001 & 01 \\ 01 & 100 & 01 \\ 10 & 010 & 10 \end{pmatrix} \quad B = D'D = \begin{pmatrix} 3 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 & 0 & 2 & 1 \\ 1 & 1 & 2 & 0 & 0 & 0 & 2 \\ 1 & 1 & 0 & 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 2 & 0 & 2 & 0 & 2 & 0 \\ 1 & 1 & 2 & 0 & 1 & 0 & 3 \end{pmatrix}$$

La matriz o tabla de contingencia de Burt B, $B = D'D$ es el resultado de todas las posibles tablas de contingencia las p variables.

Las Propiedades que se cumple para la extensión del Análisis de Correspondencia Simple en Análisis de Correspondencia Múltiple:

- Es equivalente un ACS de la tabla de contingencia entre Y y X
- Analizar la tabla disyuntiva D (de $n - \text{filas}$ e $I + J \text{ columnas}$)
- Analizar la tabla de Burt de $I + J \text{ filas}$ e $I + J \text{ columnas}$.

A partir de la tabla de Burt, se obtienen los vectores y valores propios diagonalizando la matriz:

$$V = \frac{1}{p} D^{-1} B$$

2.3. Árboles de decisión

La metodología de árboles de decisión fue desarrollada por Breiman, Friedman, Olshen y Stone en 1984. El objetivo es encontrar que las variables independientes se construyen de manera sucesiva una decisión de la información disponible, dividiendo los datos originales en pares de grupos de las variables de en función de la variable dependiente, los árboles de decisión pueden clasificar los factores que afectan a los grupos étnicos según el rendimiento.

Los árboles de decisión son una alternativa a los análisis tradicionales de regresión y clasificación. Las ventajas del árbol CART (Classification and Regression Trees) es la robustez a outliers, la invarianza de los árboles de clasificación o de regresión a transformaciones monótonas de las variables independientes y la facilidad de interpretación (Deconinck, Zhang, Coomans, & Heyden, 2006).

Los árboles de regresión se construyen cuando la variable dependiente es cuantitativa continua, mientras que un árbol de clasificación es cuando la variable dependiente es cualitativa.

Los árboles de decisión utilizan el algoritmo llamado segmentación recursiva, donde hay tres procedimientos tales como: QUEST (Quick Unbiased Efficient Statistical Tree), CHAID (Chi-square Automatic Interaction Detector), y CART (Classification and Regression Trees).

2.3.1. Elementos del árbol

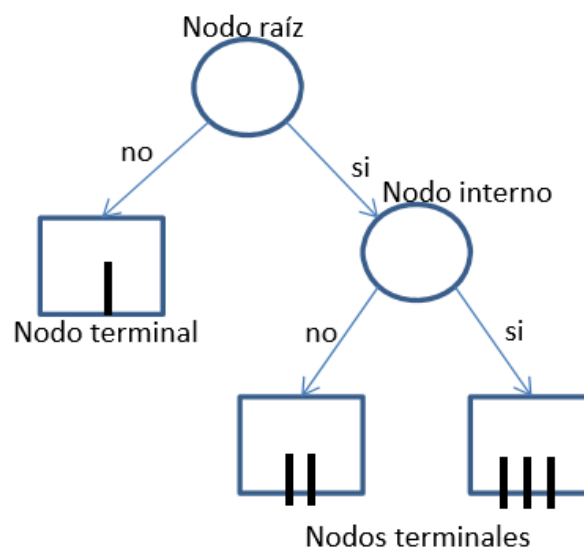


Ilustración 1: Elementos del árbol de decisión

En el primer nivel existe un único nodo llamado nodo raíz, seguido en un segundo nivel que se denomina nodo interno, y para dar por terminado a una rama de un árbol son los nodos terminales, El nodo raíz y el nodo interno son particionados por los nodos hijos o también se les conoce como ramas.

El objetivo del algoritmo de segmentación recursiva es hacer las variables resultantes en nodos terminales homogéneos como sea posible, la medida de homogeneidad es una noción de impureza (De la Fuente Fernández, 2011).

2.3.2. Proceso de análisis de árboles de decisión

El criterio de particionamiento nos permite dividir los datos donde se determinará la medida de impurezas. Para el análisis de clasificación, tanto de regresión como clasificación, se dividen en tres pasos importantes (Timofeev, 2004).

1. Construcción del árbol máximo.
2. Poda del árbol.
3. Selección del árbol óptimo mediante un procedimiento de validación cruzada (“cross-validation”).

2.3.3. Construcción del árbol máximo

De la tesis llamada “Comparación de Árboles de Regresión, Clasificación y regresión logística” de Serna (2009) se describen los pasos que sigue la metodología de árboles de decisión.

El árbol máximo es construido utilizando un procedimiento de partición binario, comenzando en la raíz del árbol. Este árbol es un modelo que describe el conjunto de entrenamiento (grupo de datos original) y generalmente es sobreajustado, es decir, contiene gran cantidad de niveles y nodos que no producen una mejor clasificación y puede ser demasiado complejo. Cada grupo es caracterizado por la distribución (respuesta categórica), o por la media (respuesta numérica) de la variable respuesta, el tamaño del grupo y los valores de las variables explicativas que lo definen. Gráficamente, el árbol se representa con el nodo raíz (los datos sin ninguna división) al iniciar, y las ramas y hojas debajo (cada hoja es el final de un grupo).

2.3.4. Calidad del nodo: Función de impureza

La función de impureza es una medida que permite determinar la calidad de un nodo, esta será denotada por $i(t)$. Existen varias medidas de impureza (criterios de particionamiento) que nos permiten analizar varios tipos de respuesta, las tres medidas más comunes presentadas por Breiman (1984), para árboles de clasificación son:

El índice de información o entropía, el cual se define como:

$$i(t) = \sum_j p(i|t) \ln p(j|t)$$

El objetivo es encontrar la partición que maximice $\Delta i(t)$ en la ecuación:

$$\Delta i(t) = - \sum_{j=1}^k p(i|t) \ln p(j|t)$$

Donde $j = 1, \dots, k$ es el número de clases de la variable respuesta categórica y $p(j|t)$ la probabilidad de clasificación correcta para la clase j en el nodo t .

El índice Gini, que tiene la forma:

$$i(t) = \sum_{i/j} p(j|t)p(i|t)$$

En donde el objetivo es encontrar la partición que maximice $\Delta i(t)$

$$\Delta i(t) = - \sum_{j=1}^k |p_j(t)|^2$$

Este índice es el más utilizado. En cada división, el índice Gini tiende a separar la categoría más grande en un grupo aparte, mientras que el índice de información tiende a formar grupos con más de una categoría en las primeras decisiones, y por último el índice "Towing".

A diferencia del índice Gini, Towing busca las dos clases que en conjunto formen más del 50% de los datos, esto define dos "súper categorías" en cada división, para las cuales la impureza es definida por el índice Gini. Aunque el índice Towing produce árboles más balanceados, este algoritmo trabaja más lento que la regla de Gini (Deconinck, Zhang, Coomans, & Heyden, 2006). Para usar el índice Towing se selecciona la partición s , que maximiza

$$\frac{p_L p_R}{4} \left[\sum_j |p(j|t_L) - p(j|t_R)| \right]^2$$

Donde t_L y t_R representan los nodos hijos izquierdo y derecho respectivamente, p_L y p_R representan la proporción de observaciones en t que pasaron a t_L y a t_R en cada caso.

2.3.5. Poda del árbol

El árbol obtenido es generalmente sobreajustado, por tanto, es podado cortando sucesivamente ramas o nodos terminales hasta encontrar el tamaño "adecuado" del árbol. Breiman, Friedman, Olshen y Stone (1984) introducen algunas ideas básicas para resolver el problema de seleccionar el mejor árbol. Computacionalmente el procedimiento descrito es complejo. Una forma es buscar una serie de árboles anidados de tamaños decrecientes (De'ath & Fabricius, 2000), cada uno de los cuales, es el mejor de todos los árboles de su tamaño. Estos árboles pequeños son comparados para determinar el óptimo. Esta comparación está basada en una función de costo complejidad, $R_\alpha(T)$. Para cada árbol T , la función costo - complejidad se define como (Deconinck, Zhang, Coomans, & Heyden, 2006):

$$R_\alpha(T) = R(T) + \alpha |\bar{T}|$$

Donde $R(T)$ es el promedio de la suma de cuadrados entre los nodos, puede ser la tasa de mala clasificación total o la suma de cuadrados de residuales total

dependiendo del tipo de árbol, $|\bar{T}|$ es la complejidad del árbol, definida como el número total de nodos del sub-árbol y α es el parámetro de complejidad. El parámetro α es un número real mayor o igual a cero. Cuando $\alpha = 0$ se tiene el árbol más grande y, a medida que α se incrementa, se reduce el tamaño del árbol. La función $R_\alpha(T)$ siempre será minimizada por el árbol más grande, por tanto, se necesitan mejores estimaciones del error. Para esto, Breiman, Friedman, Olshen, & Stone (op. cit.) proponen obtener estimadores “honestos” del error por “validación cruzada”. Computacionalmente el procedimiento es exigente pero viable, pues solo es necesario considerar un árbol de cada tamaño, es decir, los árboles de la secuencia anidada.

2.3.6. Selección del árbol óptimo

De la secuencia de árboles anidados es necesario seleccionar el árbol óptimo y para esto no es efectivo utilizar comparación o penalización de la complejidad (De'ath & Fabricius, 2000), por tanto, se requiere estimar con precisión el error de predicción y en general esta estimación se hace utilizando un procedimiento de validación cruzada. El objetivo es encontrar la proporción óptima entre la tasa de mala clasificación y la complejidad del árbol, siendo la tasa de mala clasificación el cociente entre las observaciones mal clasificadas y el número total de observaciones. El procedimiento de validación cruzada puede implementarse de dos formas:

Si se cuenta con suficientes datos se particiona la muestra, sacando la mitad o menos de los datos y se construye la secuencia de árboles utilizando los datos que permanecen, luego predecir, para cada árbol, la respuesta de los datos que se sacaron al iniciar el proceso, obtener el error de las predicciones y seleccionar el árbol con el menor error de predicción.

En general no se cuenta con suficientes datos como para utilizar el procedimiento anterior, de modo que otra forma sería:

Validación cruzada con partición en V , (v -fold cross validation, se menciona más adelante).

La idea básica de la “validación cruzada” es sacar de la muestra de aprendizaje una muestra de prueba, con los datos de la muestra de aprendizaje se calculan los estimadores y el subconjunto sacado es usado para verificar el desempeño de los estimadores obtenidos utilizándolos como “datos nuevos”. El desempeño entendido como el error de predicción, es acumulado para obtener el error medio absoluto del conjunto de prueba.

Como se mencionó anteriormente, para la metodología CART generalmente se utiliza validación cruzada con partición en V (v -fold cross validation), tamaño $V = 10$ y el procedimiento es el siguiente

- Dividir la muestra en diez grupos mutuamente excluyentes y de aproximadamente igual tamaño.

- Sacar un conjunto por vez y construir el árbol con los datos de los grupos restantes. El árbol es usado para predecir la respuesta del conjunto eliminado.
- Calcular el error estimado para cada subconjunto. Repetir los “ítems” dos y tres para cada tamaño de árbol.
- Seleccionar el árbol con la menor tasa de mala clasificación.
- Al llegar a este punto se procede a analizar el árbol obtenido.

La siguiente figura es el diagrama de flujo del algoritmo CART

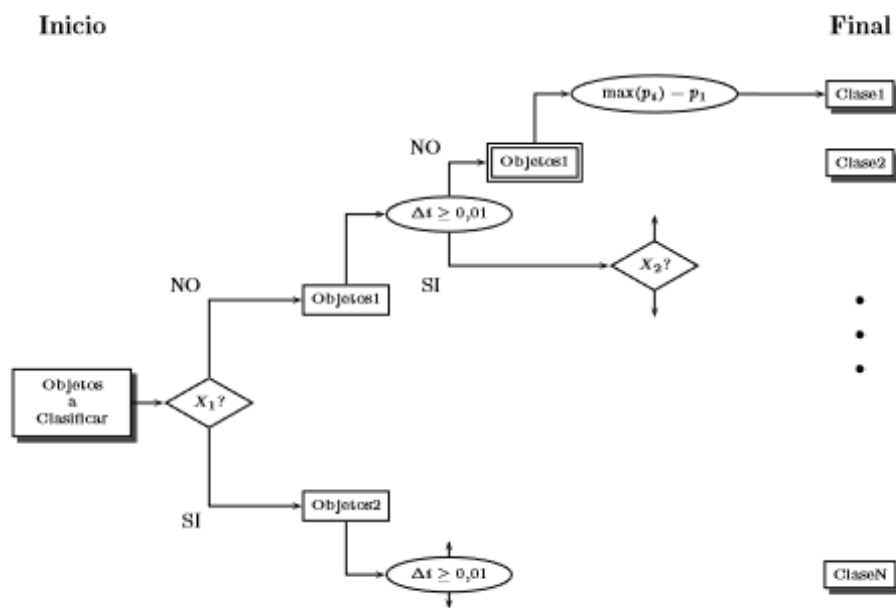


Ilustración 2: Diagrama de flujo del algoritmo CART

Fuente: Comparación de Árboles de Regresión, Clasificación y regresión logística
Elaborado por: Sandra Serna

CAPÍTULO 3

3. METODOLOGÍA

3.1. Determinación de la existencia de diferencias en el rendimiento académico de los grupos étnicos en el ciclo 2017 - 2018.

En esta sección se expresa la metodología que se utilizara para especificar si existe diferencia de rendimiento académico según los grupos étnicos.

3.1.1. Control de factores

Para el desarrollo de la investigación se toma la evaluación “Ser Bachiller” del ciclo 2017 -2018 que dispone el Instituto Nacional de Evaluación Educativa, esta evalúa el desarrollo de aptitudes y destrezas que los estudiantes deben tener al culminar la educación obligatoria; los instrumentos que dispone la institución mide cinco campos de conocimiento que son: Aptitud Abstracta, Dominio Matemático, Dominio Lingüístico, Dominio Científico y Dominio Social; Esta prueba tiene como población objetivo a los aspirantes al título de bachiller y postulantes a estudios de educación superior, la cobertura es a nivel nacional, y el instrumento consta de los siguientes elementos:

- Tipo de ítems cerrados (opción múltiple).
- 155 ítems de evaluación con 5 ítems que no aportan a la calificación.
- La duración es de 180 minutos para la población normal, para la población con algún tipo de discapacidad tiene un tiempo de duración de 240 minutos.
- Una sesión de evaluación.
- Modalidad de aplicación digital.

Además, existe un instrumento adicional que busca explicar los motivos por los cuales el estudiante ha obtenido la calificación, esta información se denomina Factores Asociados del estudiante, donde se dispone de la variable “¿Cómo te identificas según tu cultura y costumbres?”, que servirá para realizar las comparaciones de rendimiento académico.

3.1.2. Prueba de hipótesis

El método de prueba de hipótesis permitirá tomar una decisión objetiva de la existencia de diferencia en el rendimiento académico de los grupos étnicos, para en etapas posteriores del estudio, determinar los factores por los cuales se produce.

La prueba de hipótesis consta de siete pasos básicos que son: supuestos, establecimiento de hipótesis, nivel de significancia, estadístico de prueba, regla de decisión, cálculo, y por último, conclusiones y recomendaciones. A continuación, se detalla metodológicamente cada uno.

Paso 1: Supuestos

Según la teoría estadística se recomienda realizar una prueba paramétrica debido a que tiene mayor eficiencia, robustez, los errores son poco probables; para utilizar las pruebas paramétricas (Prueba Z, Prueba t–student, etc.) se debe cumplir con los supuestos de normalidad (Kolmogorov – Smirnov) y homocedasticidad (Prueba de Levene).

En el caso de no cumplir con uno de los supuestos se tiene la alternativa de realizar Pruebas no Paramétricas (U de Mann – Whitney, Prueba de Wilcoxon, Kruskal – Wallis, etc.).

Para esta investigación se plantea realizar una Prueba ANOVA si se cumplen con los supuestos de las pruebas paramétricas (normalidad, homocedasticidad). En el caso de no cumplir con uno de los supuestos se utilizará la prueba H de Kruskal - Wallis.

Paso 2: Establecimiento de la hipótesis

En este paso se redacta la hipótesis de investigación, realizando una suposición de una posible relación entre dos o más variables, esta se expresa según la experiencia que tiene el investigador.

Hipótesis de investigación

H₀: Es similar el rendimiento académico entre los grupos étnicos

H₁: Al menos un grupo étnico tiene diferencia en el rendimiento académico

Paso 3: Nivel de significancia

El nivel de significancia es la probabilidad de dar conclusiones erradas, existe el error tipo 1 que dice rechazar la hipótesis nula cuando esta es verdadera, y el error tipo 2 que dice aceptar la hipótesis nula, siendo esta falsa.

En la presente investigación se elegirá el nivel de significancia (α) del 0.05 si se desea que sea significativo.

Paso 4: Estadístico de prueba

En este paso se elige el estadístico, el cual servirá de referencia para poder tomar la decisión si existe diferencia en el rendimiento académico entre los diferentes

grupos étnicos, debido a esto se eligió la prueba no paramétrica denominada H de Kruskal – Wallis.

Estadístico no paramétrico

Supongamos el tamaño de cada grupo étnico es n_j con $j = 1, 2, \dots, k$. Entonces se asigna rangos a la muestra combinada, el estadístico es.

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1)$$

Donde:

$N = \sum_{j=1}^k n_j$, Si cada $n_j > 5$ tiene una distribución chi-cuadrado con $k-1$ grados de libertad, Si alguno de los tamaños de la muestra son ≤ 5 , se utiliza la distribución para muestras pequeñas.

Paso 5: Regla de decisión

En este paso se redacta la regla de decisión que tiene el estadístico y que permitirá concluir si existe o no diferencia significativa del rendimiento académico entre los grupos étnicos.

La regla de decisión para el estadístico H de Kruskal - Wallis es:

Se rechaza H_0 si $H < H_{\alpha/2}$ o $H > H_{1-\alpha/2}$

Paso 6: Cálculo

En el capítulo de resultados se encuentran descritos todos los cálculos del estadístico Kruskal – Wallis.

Paso 7: Conclusiones y recomendaciones

En base a los resultados obtenidos se realiza la comparación según lo descrito en el paso 5 y se procederá a redactar las conclusiones y recomendaciones de la investigación.

3.2. Reducción de la dimensionalidad del conjunto de datos de la evaluación del ciclo 2017 – 2018 a través de las variables asociadas a la etnia para facilidad de análisis.

La finalidad de la reducción de dimensiones de la base de datos es precisamente, disminuir el número de variables (componentes) sin perder patrones y comportamiento de la base de datos de la evaluación “Ser Bachiller” y por lo tanto se logre explicar los factores por los cuales influyen el rendimiento académico de cada uno de los grupos étnicos.

La técnica seleccionada para la reducción de dimensiones es Análisis de Componentes Principales No Lineales (ACPNL). Las nuevas variables o componentes son no correlacionados por lo que se permiten realizar análisis de la nueva base de datos manteniendo la esencia, pero dejando variables representativas, las componentes se ordenan en función de la cantidad de varianza explicada.

Para realizar el ACPNL se trabajó a través de la técnica de transformación denominada escalonamiento óptimo, donde a la categoría de la variable se le asigna cuantificaciones numéricas, y a partir de estas cuantificaciones se obtiene resultados.

El método de escalonamiento óptimo seleccionado para realizar la transformación se denomina “**Multiplicación**”, donde a partir de los valores de la variable se tipifican multiplicando por 10, se suma una constante donde el menor valor transformado sea uno.

Para los datos perdidos se realizó un método de imputación de categoría aleatoria donde se imputa el valor faltante de la variable con valores aleatorios en función de la distribución de las categorías de la variable

El método de normalización para el ACPNL es principal por variable. ya que el interés de optimización es la asociación entre las variables, o sea la correlación de cada una de estas.

Además, se ponderó la variable etnia en 3, debido a que esta es la variable objetivo del estudio, por lo que se busca que el comportamiento de esta variable sea determinante para las conclusiones del estudio.

3.3. Identificación de las posibles causas o factores que influyen en los niveles de logro de los grupos étnicos en el ciclo 2017 – 2018.

Mediante la técnica de Análisis de Correspondencias se busca encontrar las relaciones existentes entre las categorías de las variables cualitativas. Se construye a través de tablas de contingencia, dando por resultado un gráfico de relaciones entre categorías y estas permiten describir la concordancia entre las variables

Para el análisis de correspondencia múltiple se construyó, a través de la técnica estadística de transformación denominada escalonamiento óptimo, donde a la categoría de la variable se le asigna cuantificaciones numéricas, y a partir de estas

cuantificaciones se obtiene resultados, el método seleccionado para realizar la transformación es multiplicación donde a partir de los valores de la variable se tipifican, se multiplica por 10 y se suma una constante donde el menor valor transformado sea uno.

Los valores perdidos se excluyen para las correlaciones, esto quiere decir que los valores perdidos de las variables no contribuyen en el análisis de correspondencia. Para las correlaciones escaladas óptimamente se selecciona la moda como método de imputación.

El método de normalización para el Análisis de Correspondencia Múltiple es **principal por objeto**, ya que el interés de optimización es la asociación entre los objetos, es decir, la correlación de cada una de las categorías.

Además, se ponderó en 2 la variable “¿Cómo te identificas según tu cultura y costumbres?”, y con 2 el nivel de logro global. Con esto se aseguró el cumplimiento del objetivo de estudio de la investigación.

3.3.1. Construcción de un modelo estadístico que permita el pronóstico del rendimiento académico de los grupos étnicos en el ciclo 2017 – 2018.

Mediante la técnica de árbol de decisión se busca cuantificar los motivos por los cuales existen diferencias que existe en el rendimiento académico entre cada uno de los grupos étnicos, a través de una representación gráfico - cuantitativo de las posibles alternativas que influyen en la mencionada problemática.

Para la construcción del árbol de decisión se tomó como variable dependiente la calificación promedio entre las materias matemáticas, lenguaje, sociales y ciencias naturales mientras que las variables independientes son las variables que fueron seleccionadas a partir de la técnica de ACPNL, pero se le da importancia a la variable “¿Cómo te identificas según tu cultura y costumbres?”, se selecciona la variable de influencia para la investigación llamada el quintil del nivel socioeconómico.

El método de crecimiento seleccionado para la construcción del árbol de decisión es CART (Classification and Regression Trees), ya que brinda la facilidad de realizar una regresión.

Se realizó la poda para evitar el sobreajuste y se especificó la diferencia máxima aceptable en el podado de 1 error estándar.

La profundidad máxima del árbol de decisión es 5, los casos mínimos del nodo padre son igual a 100 y los casos mínimos del nodo hijo es de 50.

CAPÍTULO 4

4. RESULTADOS

4.1. Determinación de la existencia de diferencias en el rendimiento académico de los grupos étnicos en el ciclo 2017 - 2018.

En esta sección se describe el resultado de la existencia de diferencia de rendimiento académico de los grupos étnicos cumpliendo con todos los pasos para probar una hipótesis.

4.1.1. Prueba de hipótesis

Para poder determinar la existencia de diferencia de rendimiento académico de los grupos étnicos se realizó una prueba de hipótesis, la misma que permitió concluir objetivamente la existencia o no de la misma. Esta prueba de hipótesis se comprobó siguiendo los pasos que se detallan a continuación.

Paso 1: Supuestos

Según la literatura la primera opción de análisis es a través de pruebas paramétricas debido a su robustez y fiabilidad, para esto se comprobó con supuestos básicos que los datos provengan de una población normal y que sean homocedásticos (igualdad de varianzas).

Normalidad

Para probar si los datos provienen de una distribución normal se utilizó el estadístico Kolmogorov – Smirnov, este nos permitió identificar si los datos analizados siguen una distribución normal.

Hipótesis

H_0 : Los datos siguen una distribución Normal
 H_1 : Los datos no siguen una distribución Normal

Resultados

Tabla 6: Prueba de normalidad de los datos

Prueba de Kolmogorov-Smirnov para una muestra

		Puntaje global
N		268525
Parámetros normales ^{a,b}	Media	765,9456215
	Desv. Desviación	88,99807974
Máximas diferencias extremas	Absoluto	0,046
	Positivo	0,046
	Negativo	-0,027
Estadístico de prueba		0,046
Sig. asintótica(bilateral)		,000 ^c

a. La distribución de prueba es normal.

b. Se calcula a partir de datos.

c. Corrección de significación de Lilliefors.

A través de la prueba de Kolmogorov – Smirnov se indica que los datos no provienen de una distribución normal, la hipótesis nula se rechaza, es decir que la significancia asintótica (p) es menor que el nivel de significancia al 0,05 establecido para la investigación.

Homogeneidad

Otro de los supuestos de las pruebas paramétricas es de igualdad de varianzas, esto significa que los grupos de análisis (grupos étnicos) poseen igualdad de varianzas para esto se utiliza la prueba de Levene.

Hipótesis

Usar un término más estadístico, homocedasticidad podría ser.

H_0 : Existe homocedaticidad de los datos

H_1 : Los datos son heterocedasticos

Resultados

Tabla 7: Prueba de homogeneidad de varianzas

		Prueba de homogeneidad de varianzas ^{a,b}			
		Estadístico de Levene	gl1	gl2	Sig.
Puntaje global	Se basa en la media	238,900	4	268065	0,000
	Se basa en la mediana	231,060	4	268065	0,000
	Se basa en la mediana y con gl ajustado	231,060	4	267697,561	0,000

Prueba la hipótesis nula de que la varianza de error de la variable dependiente es igual entre grupos.

a. Variable dependiente: Puntaje global

b. Diseño : Intersección + etnibbe

Según la prueba de Levene de homogeneidad de varianzas se indica que los grupos étnicos no tienen igualdad de varianza ya sea por la media, mediana, mediana y con los grados de libertad ajustados o media recortada, es decir que para cada uno de los diferentes casos la significancia asintótica (p) es menor que el nivel de significancia al 0.05 establecido.

Para concluir la etapa de supuestos se tiene como resultado, que los datos no provienen de una distribución normal y las varianzas son desiguales, debido a esto, es imposible utilizar pruebas paramétricas, por esto se planteó realizar la comparación a través de pruebas no paramétricas, el estadístico pertinente para este estudio fue H de Kruskal – Wallis.

Paso 2: Establecimiento de las hipótesis

Como se comprobó en el paso 1, no se pueden utilizar estadísticos paramétricos por lo que se realizó la prueba con el estadístico no paramétrico denominado H de Kruskal - Wallis que es la generalización para más de dos muestras, esta prueba consiste en comparar las poblaciones a través de las medianas por lo que las hipótesis establecidas fueron:

La hipótesis nula no se redacta con negación.

Recordar que en el contraste se plantea parámetros, por tal, reemplazar M1 por $\tilde{\mu}_1$ y así sucesivamente

H_0 : *Es similar el rendimiento académico entre los grupos étnicos*

$$H_0: \tilde{\mu}_1 = \tilde{\mu}_2 = \dots = \tilde{\mu}_k$$

H_1 : *Al menos un grupo étnico tiene diferencia en el rendimiento académico*

$$H_1: \tilde{\mu}_i \neq \tilde{\mu}_j \text{ al menos para un par } (i, j)$$

Donde:

$\tilde{\mu}_k$: 1, ..., k grupos étnicos

Paso 3: Nivel de significancia

En la presente investigación se eligió el nivel de significancia (α) del 0.05 dado que se deseaba que sea significativo, pero en caso de desear mayor exactitud el nivel de significancia será al 0.01 o el . si es altamente significativo.

$$\alpha = 0.05 \text{ o } 0.01$$

Paso 4: Estadístico de prueba

Debido a que del resultado del paso 1 no se pudo determinar si existe diferencia del rendimiento a través de una prueba paramétrica entonces se eligió la prueba no paramétrica denominada H de Kruskal – Wallis, siendo su estadístico el siguiente:

$$H = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j^2}{n_j} - 3(N+1)$$

En la sección de la metodología del capítulo 3 se detalla cada uno de los parámetros de la ecuación.

Paso 5: Regla de decisión

En este paso se redactó la regla de decisión que tiene el estadístico que permitió concluir si existe o no diferencia significativa del rendimiento académico de los grupos étnicos de estudio.

La regla de decisión para el estadístico H de Kruskal - Wallis fue:

Aclarar, la hipótesis nula se rechaza o no se rechaza, en estadístico se concluye así..

Si se habla de resultados significativo cuando se ha rechazado la hipótesis nula, aclarar esta regla de decisión en términos estadístico.

Si $H_{cal} > H_{tab}$, se rechaza la hipótesis nula.

Si $H_{cal} \leq H_{tab}$, se acepta la hipótesis nula.

O también se puede concluir:

Si $p \leq 0,05$, no existe evidencia significativa para aceptar la hipótesis nula, en el caso contrario se acepta la hipótesis alternativa.

Paso 6: Cálculo

En este paso se presenta el resultado del estadístico H de Kruskal – Wallis que servirá de referencia para poder tomar la decisión si existe o no la diferencia del rendimiento académico según los grupos étnicos que se analizan en el estudio.

Tabla 8: Comparación de poblaciones

Estadísticos de prueba ^{a,b}	
	Puntaje global
H de Kruskal-Wallis	5004,472
Gl	3
Sig. Asintótica	0,000

a. Prueba de Kruskal Wallis

b. Variable de agrupación: ¿Cómo te identificas según tu cultura y costumbres?

El estadístico de H de Kruskal-Wallis nos da por resultado de 5004.472 que es calculado con 3 grados de libertad y el valor p (significancia asintótica) es un valor muy cercano a cero.

Paso 7: Conclusiones y recomendaciones

A partir de los resultados del cálculo de la prueba H de Kruskal-Wallis se tuvo que el valor p es cercano a cero, por lo tanto, es menor al nivel de significancia (0,05 y 0,01) entonces se tiene que hay evidencia suficiente para rechazar la hipótesis nula, debido a esto se concluye que al menos un grupo étnico tiene diferencia en el rendimiento académico.

4.2. Reducción de la dimensionalidad del conjunto de datos de la evaluación del ciclo 2017 – 2018 a través de las variables asociadas a la etnia para un eficiente análisis

En esta sección se busca reducir el número de variables para mayor facilidad de análisis sin perder los patrones.

4.2.1. Análisis de consistencia de la información

Para la realización de esta investigación se tomó información de la evaluación “Ser Bachiller” del ciclo 2017 – 2018 que realiza el Instituto Nacional de Evaluación Educativa donde consta de dos secciones, la primera es de niveles de logro donde se encuentra información demográfica del estudiante, códigos únicos, Notas alcanzadas (cualitativas y cuantitativas), nivel socioeconómico, información de la institución; en total son 55 variables. La segunda es de factores asociados que posiblemente afecte al estudiante en el rendimiento académico, esta consta de 295 variables.

La información de niveles de logro es consistente, posee más del 95% de la información de cada una de las variables; mientras que en la información de factores asociados existe una variabilidad en las respuestas de los estudiantes, debido a esto, se realizó un análisis de consistencia para quedarse únicamente con las variables que posean información suficiente para poder realizar análisis eficientes en etapas posteriores.

Para poder determinar patrones que tiene la base de datos y pueda brindar información estable, confiable, coherente, de calidad, se seleccionaron a las variables que disponen del 80% o más de información.

De las 295 variables de factores asociados, 186 variables tuvieron información menor al 80% que representa el 63% de las variables, mientras que 109 variables tuvieron información del 80% o más, que sería el 37%.

4.2.2. Reducción de dimensiones

Debido a que 109 variables es un número grande de variables, para poder realizar un análisis eficiente, se buscó disminuir las variables sin perder patrones, comportamiento de la información y el objetivo de la investigación, para esto se escogió la técnica Análisis de Componentes Principales No Lineales (ACPNL), la misma que realiza un mapeo lineal de variables correlacionadas y las transforma en un conjunto menor de datos con variables no correlacionadas, las nuevas componentes se van ordenando en función de la cantidad de varianza explicada. A continuación, se presentan el resultado de la implementación de la técnica estadística de Análisis de Componentes Principales No Lineales (ACPNL) ponderando la variable etnia.

Tabla 9: Resumen del modelo de ACPNL

Dimensión	Alfa de Cronbach	Varianza contabilizada para	
		Total (autovalor)	% de varianza
1	0,921	11,436	9,944
2	0,874	7,508	6,528
3	0,785	4,513	3,924
4	0,740	3,754	3,265
5	0,728	3,598	3,128
6	0,610	2,528	2,199
7	0,587	2,394	2,081
8	0,575	2,324	2,021
9	0,510	2,024	1,760
10	0,491	1,949	1,695
11	0,465	1,855	1,613
12	0,430	1,742	1,515
13	0,367	1,571	1,366
14	0,359	1,553	1,351
15	0,321	1,468	1,276
16	0,273	1,372	1,193
17	0,241	1,314	1,142
18	0,206	1,258	1,094
19	0,198	1,245	1,082
20	0,166	1,197	1,041
21	0,132	1,151	1,001
22	0,116	1,129	0,982
23	0,102	1,111	0,966
24	0,075	1,081	0,940
25	0,065	1,069	0,930

26	0,040	1,041	0,906
27	0,037	1,038	0,902
28	0,003	1,001	0,871
29	-0,007	0,993	0,863
30	-0,012	0,988	0,859
31	-0,028	0,973	0,846
32	-0,039	0,963	0,838
Total	,994 ^a	69,140	60,121

a. Se utiliza el total de alfa de Cronbach en el autovalor total.

En la tabla 9 del resumen del modelo se tiene que el autovalor (varianza explicada) es mayor a 1 en 28 componentes, mientras que en el porcentaje de varianza explicada se tiene que 21 componentes son mayores al 1%, debido a esto, y para mejorar la precisión del modelo se realizó la rotación.

Tabla 10: Resumen del modelo ACPNL rotado

Resumen del modelo rotado^a

Dimensión	Alfa de Cronbach	Varianza contabilizada para	
		Total (autovalor)	% de varianza
1	0,800	4,139	3,599
2	0,790	4,085	3,552
3	0,845	3,757	3,267
4	0,881	3,546	3,083
5	0,802	3,379	2,939
6	0,841	2,786	2,422
7	0,816	2,619	2,278
8	0,854	2,567	2,232
9	0,782	2,333	2,029
10	0,681	2,270	1,974
11	0,813	2,241	1,948
12	0,642	2,166	1,884
13	0,786	2,140	1,861
14	0,591	2,121	1,844
15	0,818	2,116	1,840
16	0,630	2,079	1,808
17	0,804	1,972	1,715
18	0,743	1,950	1,696
19	0,745	1,938	1,685
20	0,759	1,905	1,656
21	0,747	1,886	1,640
22	0,477	1,778	1,546
23	0,757	1,777	1,545
24	0,638	1,660	1,444
25	0,629	1,448	1,259
26	0,544	1,367	1,188
27	0,631	1,332	1,158

28	0,440	1,315	1,143
29	0,655	1,307	1,136
30	0,232	1,074	0,934
31	0,191	1,070	0,930
32	0,039	1,017	0,885
Total	,994 ^b	69,137	60,120

a. Método de rotación: Varimax con normalización Kaiser.

b. Se utiliza el total de alfa de Cronbach en el autovalor total.

En la tabla 10 se muestra el resumen del modelo rotado a través del método VARIMAX (varianza máxima) con normalización de Kaiser y se tiene que en 29 componentes el autovalor es mayor que 1.3, y el porcentaje de varianza explicada es mayor al 1%, además el alfa de Cronbach (fiabilidad de la escala) es mayor al 0,544 y en el componente 28 tiene un valor de 0,440.

Después de obtener el resultado de 29 componentes se identificaron las variables que representarán a cada componente (dimensión), para esto se tuvo la matriz de carga de componentes que ayudó a elegir el nuevo conjunto de datos.

Tabla 11: Carga de componentes rotados

Variables	Carga	Dimensión
¿Tienes hijos?	-0,825	Dimensión 1
¿Con quién vives? Selecciona todas las que aplican. Hijo(s)	-0,808	Dimensión 1
¿Con quién vives? Selecciona todas las que aplican. Cónyuge, conviviente o pareja	-0,749	Dimensión 1
¿Quién es el jefe de tu hogar?	0,708	Dimensión 1
¿Cómo te identificas según tu cultura y costumbres?	0,990	Dimensión 2
¿A cuál de las siguientes nacionalidades/pueblos indígenas perteneces?	-0,989	Dimensión 2
¿Los maestros llegaban tarde a clases?	0,743	Dimensión 3
¿Los maestros faltaban a clases?	0,725	Dimensión 3
¿Los maestros se iban más temprano?	0,721	Dimensión 3
¿Los maestros nos dejaban solos en las horas de clase?	0,705	Dimensión 3
¿La lectura es uno de tus pasatiempos favoritos?	-0,724	Dimensión 5
Marca la alternativa que más se parece a tu forma de pensar y actuar. Leer es entretenido	0,716	Dimensión 5
Marca la alternativa que más se parece a tu forma de pensar y actuar. Los libros son aburridos	0,709	Dimensión 5
Señala el nivel educativo más alto que han completado el jefe de tu hogar	0,784	Dimensión 6
Señala el nivel educativo más alto que han completado tu padre	0,748	Dimensión 6
Señala el nivel educativo más alto que han completado tu madre	0,738	Dimensión 6
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Desagüe o alcantarillado	0,809	Dimensión 7
¿Qué tipo de servicio higiénico tiene tu vivienda?	-0,763	Dimensión 7
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Computadora de escritorio (desktop)	0,780	Dimensión 8

¿Crees que con los conocimientos adquiridos en la escuela puedes aplicar a un trabajo y defenderte en la vida?	-0,810	Dimensión 9
Sobre los conocimientos adquiridos en tu colegio. Siento que los conocimientos adquiridos me motivan a investigar más, desarrollar nuevas ideas y llevarlas a la práctica	0,709	Dimensión 9
Sobre los conocimientos adquiridos en tu colegio. Siento que lo que aprendí cumple con mis expectativas de aprendizaje.	0,703	Dimensión 9
En el último mes de clase que tuviste, ¿con qué frecuencia ocurría lo siguiente? Falté a clases un día completo	0,760	Dimensión 10
En el último mes de clase que tuviste, ¿con qué frecuencia ocurría lo siguiente? No asistía a clases en algunas materias	0,716	Dimensión 10
En el último mes de clase que tuviste, ¿con qué frecuencia ocurría lo siguiente? Llegaba tarde al colegio	0,715	Dimensión 10
En tu hogar, ¿cuáles de las siguientes actividades se realizan? Selecciona todas las que aplican. Alguien está registrado en una red social	0,701	Dimensión 11
¿Crees que leer es una pérdida de tiempo?	0,937	Dimensión 12
Marca la alternativa que más se parece a tu forma de pensar y actuar. Leer es perder el tiempo	0,932	Dimensión 12
Si hubieras tenido que cambiarte de colegio, ¿cómo te hubieses sentido?	0,741	Dimensión 13
¿Qué tan orgulloso te sientes o sentías en tu colegio?	0,719	Dimensión 13
Sexo del estudiante	0,953	Dimensión 14
¿Eres hombre o mujer?	0,953	Dimensión 14
¿En tu aula los estudiantes son buenos amigos?	-0,976	Dimensión 16
En tu colegio, ¿con qué frecuencia pasaban estas cosas? Tengo buenos amigos en el aula de clase	0,974	Dimensión 16
¿Cuántos cuartos de baño con ducha tiene tu hogar?	0,834	Dimensión 17
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Baños	-0,816	Dimensión 17
¿Había ruido y desorden en el aula?	0,725	Dimensión 18
Señala la labor que más se parece al trabajo de tu padre	-0,833	Dimensión 20
Señala la labor que más se parece al trabajo del jefe de tu hogar	-0,816	Dimensión 20
¿Con quién vives? Selecciona todas las que aplican. Abuelos	0,786	Dimensión 22
¿Con quién vives? Selecciona todas las que aplican. Tíos	0,769	Dimensión 22
En su mayor parte, ¿de qué material son las paredes exteriores de tu casa?	0,757	Dimensión 23
¿Algunas de estas cosas te pasan o pasaban en el colegio? Me siento amenazado por algún compañero	0,776	Dimensión 24
¿Algunas de estas cosas te pasan o pasaban en el colegio? Mis compañeros se burlan de mí	0,771	Dimensión 24
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Teléfonos celulares sin conexión a internet	0,777	Dimensión 25
¿Qué carrera te gustaría seguir en la universidad?	0,836	Dimensión 26
¿Cuántas personas en tu hogar están afiliados o cubiertos por un seguro de salud privado (con o sin hospitalización), seguro internacional o seguro de vida?	0,763	Dimensión 27
Señala qué hace regularmente tu madre	-0,768	Dimensión 29
¿Con quién vives? Selecciona todas las que aplican. Madre	0,639	Dimensión 1
¿Con quién vives? Selecciona todas las que aplican. Hermanos	0,550	Dimensión 1
¿Con quién vives? Selecciona todas las que aplican. Padre	0,459	Dimensión 1
¿Cuál es tu principal ocupación?	0,336	Dimensión 1
¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	-0,631	Dimensión 26
Señala qué hace regularmente tu padre	-0,612	Dimensión 29

La tabla 11 muestra las variables seleccionadas según cada componente (dimensión), ya que estas tienen una carga mayor a 0,7 o menor a -0,7 esto quiere

decir que dichas variables son las más representativas en cada componente y no posee ninguna correlación entre ellas. En total son 48 variables, además se seleccionaron 6 variables adicionales que no cumplían con esta condición, pero debido a su importancia para los análisis posteriores y a su aporte para las conclusiones, se las eligió, por lo tanto, tenemos un resultado de 54 variables.

4.3. Identificación de las posibles causas o factores que influyen en los niveles de logro de los grupos étnicos en el ciclo 2017 – 2018.

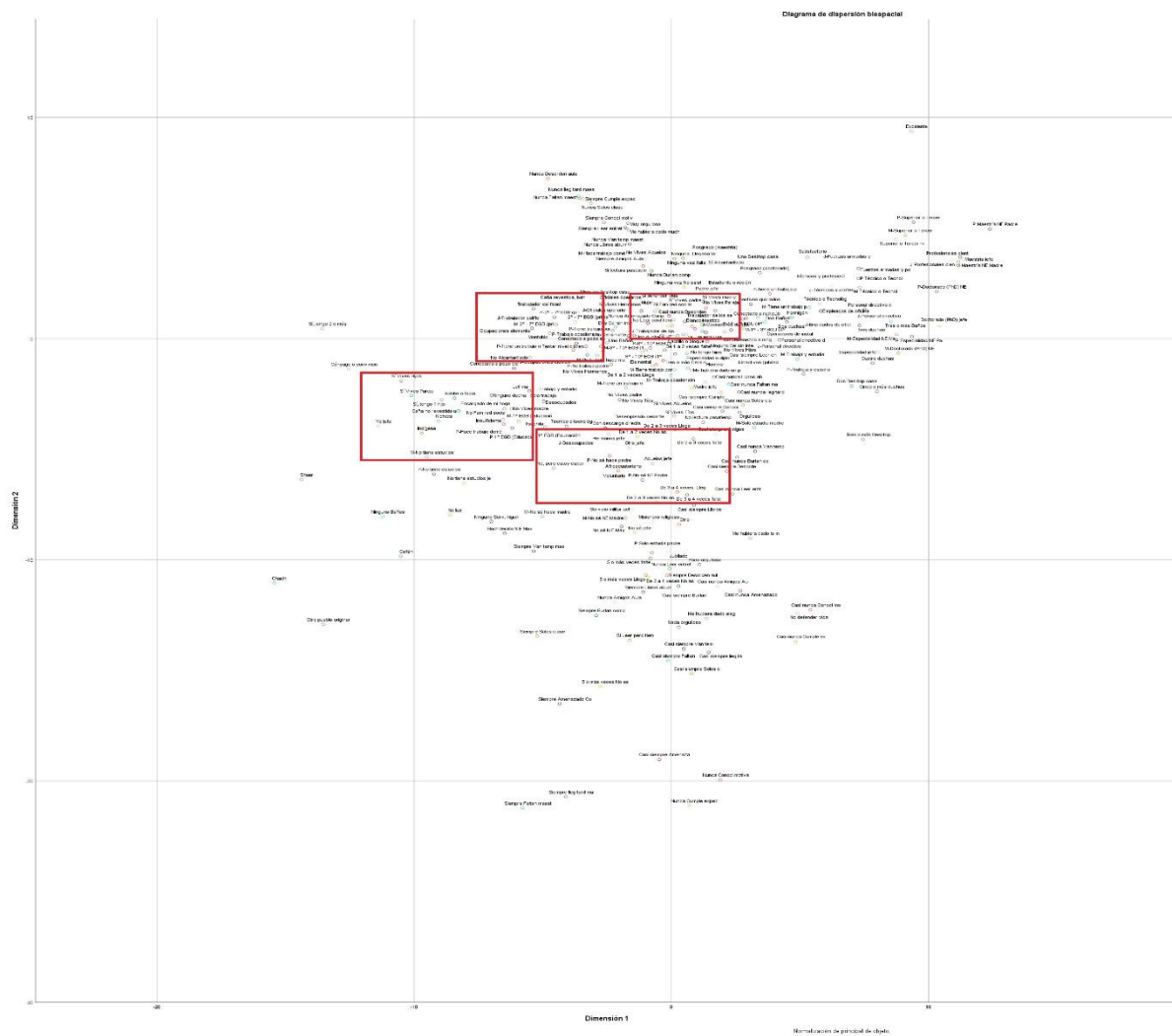
Esta investigación identificó las causas que influyen en los niveles de logro que tiene el estudiante a través de los factores asociados que se dispone de la base de datos de la evaluación “Ser Bachiller” del ciclo 2017 – 2018 del Instituto Nacional de Evaluación Educativa, a través de la técnica estadística denominada Análisis de Correspondencia Múltiple (ACM). Para esto se utilizaron las 54 variables que se obtuvieron en el Análisis de Componentes Principales No Lineales (ACPNL) detalladas en la sección anterior.

Tabla 12: Resumen del modelo de Análisis de Correspondencia Múltiple

Resumen del modelo				
Dimensión	Alfa de Cronbach	Varianza contabilizada para		
		Total (autovalor)	Inercia	
1	0,862	6,499	0,123	
2	0,793	4,498	0,085	
Total		10,997	0,207	
Media	,834 ^a	5,499	0,104	

a. La media de alfa de Cronbach se basa en la media de autovalor.

El modelo de Análisis de Correspondencia Múltiple explica el 0,207 de la proporción de la varianza en las dos dimensiones. La fiabilidad de la escala y el promedio de las correlaciones entre las variables mide el Alfa de Cronbach si se acerca a 1 significa que la escala es confiable en nuestro caso la dimensión 1 es de 0.862 y en la dimensión 2 es de 0.793, por lo tanto, la escala de medición es confiable.



- Centroides ¿A cuál de las siguientes nacionalidades/pueblos indígenas perteneces?
- Centroides ¿Algunas de estas cosas te pasan o pasaban en el colegio? Me siento amenazado por algún compañero
- Centroides ¿Algunas de estas cosas te pasan o pasaban en el colegio? Mis compañeros se burlan de mí
- Centroides ¿Cómo te identificas según tu cultura y costumbres?
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Abuelos
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Conyuge, conviviente o pareja
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Hermanos
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Hijos(as)
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Madre
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Padre
- Centroides ¿Con quién vives? Selecciona todas las que aplican. Tíos
- Centroides ¿Crees que con los conocimientos adquiridos en la escuela puedes aplicar a un trabajo y defenderte en la vida?
- Centroides ¿Crees que leer es una pérdida de tiempo?
- Centroides ¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?
- Centroides ¿Cuál es tu principal ocupación?
- Centroides ¿Cuántos cuartos de baño con ducha tiene tu hogar?
- Centroides ¿Eres hombre o mujer?
- Centroides ¿Había ruido y desorden en el aula?
- Centroides ¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican. Desague o alcantarillado
- Centroides ¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican. Luz eléctrica
- Centroides ¿La lectura es uno de tus pasatiempos favoritos?
- Centroides ¿Los maestros llegaban tarde a clases?
- Centroides ¿Los maestros nos dejaban solos en las horas de clase?
- Centroides ¿Los maestros se iban más temprano?
- Centroides ¿Que tan orgullosos te sientes o sientas en tu colegio?
- Centroides ¿Qué tipo de servicio higiénico tiene tu vivienda?
- Centroides ¿Quién es el jefe de tu hogar?
- Centroides ¿Tienes hijos?
- Centroides De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Baños
- Centroides De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Computadora de escritorio (desktop)
- Centroides De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Teléfonos celulares sin conexión a internet
- Centroides En el último mes de clase que tuviste, ¿con qué frecuencia ocurría lo siguiente? Faltaba a clases un día completo
- Centroides En el último mes de clase que tuviste, ¿con qué frecuencia ocurría lo siguiente? Llegaba tarde al colegio
- Centroides En el último mes de clase que tuviste, ¿con qué frecuencia ocurría lo siguiente? No asistía a clases en algunas materias
- Centroides En su mayor parte, ¿de qué material son las paredes exteriores de tu casa?
- Centroides En tu colegio, ¿con qué frecuencia pasaban estas cosas? Tengo buenos amigos en el aula de clase
- Centroides En tu hogar, ¿cuáles de las siguientes actividades se realizan? Selecciona todas las que aplican. Alguien está registrado en una red social
- Centroides Marca la alternativa que más se parece a tu forma de pensar y actuar. Leer es entretenido
- Centroides Marca la alternativa que más se parece a tu forma de pensar y actuar. Los libros son aburridos
- Centroides Nivel de logro puntaje global
- Centroides Señala el nivel educativo más alto que han completado el jefe de tu hogar
- Centroides Señala el nivel educativo más alto que han completado tu madre
- Centroides Señala el nivel educativo más alto que han completado tu padre
- Centroides Señala la labor que más se parece al trabajo del jefe de tu hogar
- Centroides Señala la labor que más se parece al trabajo de tu padre
- Centroides Señala qué hace regularmente tu madre
- Centroides Señala qué hace regularmente tu padre
- Centroides Si hubieras tenido que cambiarte de colegio, ¿cómo te hubieses sentido?
- Centroides Sobre los conocimientos adquiridos en tu colegio. Siento que lo que aprendí cumple con mis expectativas de aprendizaje
- Centroides Sobre los conocimientos adquiridos en tu colegio. Siento que los conocimientos adquiridos me motivan a investigar más, desarrollar nuevas ideas y llevarlas a la práctica

Ilustración 3: Gráfico de Correspondencia Múltiple

4.3.1. Factores que influyen en el rendimiento académico de la etnia blanco/mestizo

A partir del gráfico de Correspondencia Múltiple se identificaron los factores que afectan el rendimiento académico de blancos y mestizos, para facilidad y mejor entendimiento se realiza un corte del gráfico y se presenta a continuación:

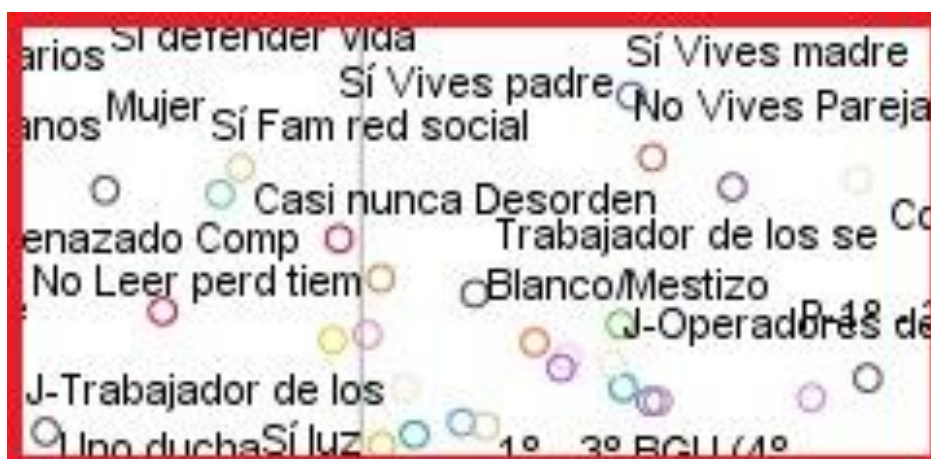


Ilustración 4: Gráfico de correspondencia Múltiple para etnia blanco/mestizo

Los factores que afectan al rendimiento académico del grupo étnico blanco/mestizo son:

La educación del jefe de familia y del padre es de 1ro a 3ro de Bachillerato General Unificado completo, el padre es Trabajador de los servicios y comerciantes u Operadores de instalaciones, máquinas y ensambladores, el estudiante vive con el padre o la madre y no vive con pareja, en el aula casi nunca hay desorden, no se siente amenazado por los compañeros, cree que leer no es una pérdida de tiempo, dispone de ducha, luz eléctrica, tiene conexión a red pública de alcantarillado.

4.3.2. Factores que influyen en el rendimiento académico de la etnia indígena

A partir gráfico de correspondencia múltiple se obtuvieron los factores que afectan al rendimiento académico del grupo étnico indígena.

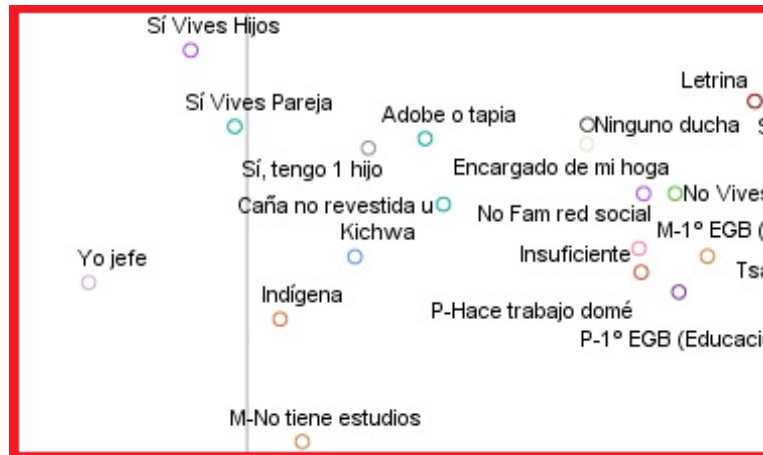


Ilustración 5: Gráfico de correspondencia múltiple del grupo étnico indígena

Los factores que influyen en el rendimiento académico al grupo étnico indígena son:

El estudiante es el jefe de hogar, vive con su pareja, no vive con su madre, tiene un hijo, es encargado de su hogar, la familia no tiene ninguna red social, la madre y el padre tienen 1º de Educación General Básica o no tiene estudios, su vivienda es de adobe o tapia u otros materiales, no tienen ducha, disponen de letrina, pertenecen a la nacionalidad Kichwa o Tsáchila, el rendimiento en la prueba ser bachiller es insuficiente.

4.3.3. Factores que influyen en el rendimiento académico de la etnia afroecuatoriano

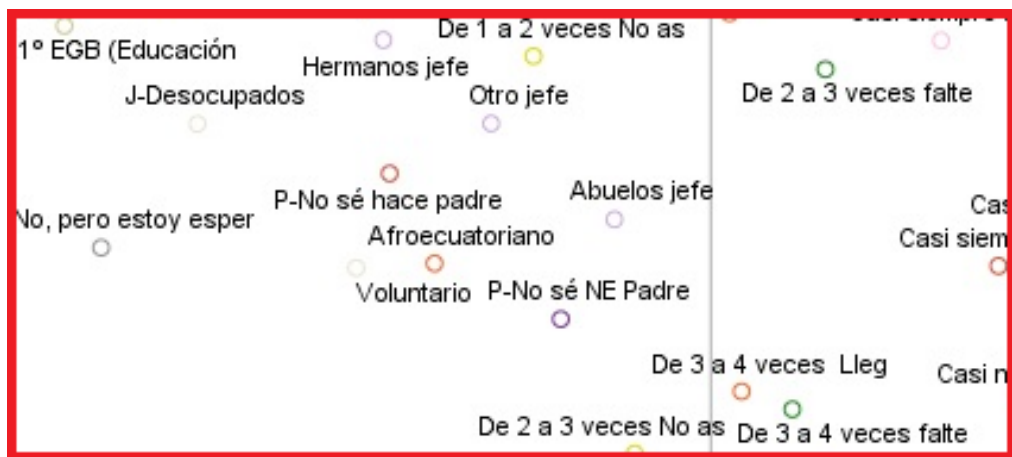


Ilustración 6: Gráfico de correspondencia múltiple del grupo étnico afroecuatoriano

A partir del gráfico de correspondencias múltiple se identificaron los factores que influyen en el rendimiento académico del grupo étnico afroecuatoriano, que son:

El estudiante tiene como jefe de hogar a los abuelos, hermanos u otra persona, está desocupado, el estudiante no tiene hijos pero está esperando uno, no sabe en que trabaja el padre, no sabe qué nivel de estudios tiene el padre, han faltado más de dos veces a clase todo el día, en el último mes no asistió a clases de 2 a 3 veces a algunas materias, en el último mes llegaba de 3 a 4 veces atrasado a clase, el

nivel de estudio del jefe de hogar es de 1 de educación general básica (educación inicial).

4.3.4. Factores que influyen en el rendimiento académico de la etnia montubio

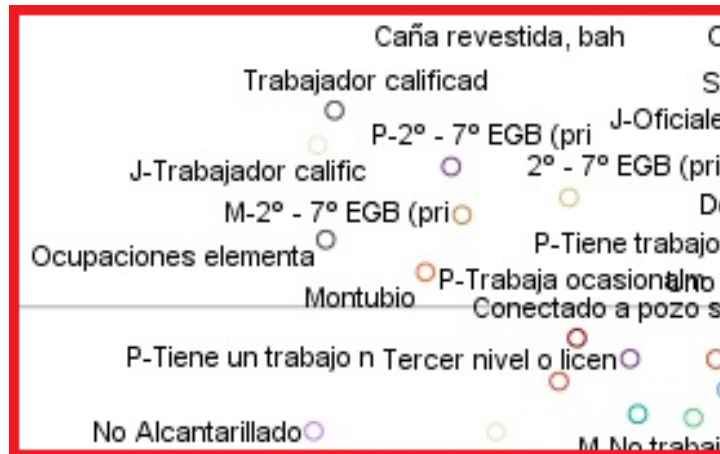


Ilustración 7: Gráfico de correspondencia múltiple del grupo étnico montubio

El gráfico de correspondencias múltiple nos ayudó a identificar los factores que influyen en el rendimiento académico del grupo étnico montubio, que son:

Los padres de los estudiantes montubios se dedican a Ocupaciones elementales (limpiadores, trabajo doméstico, peones, ayudantes de preparación de alimentos, vendedores ambulantes, trabajo no remunerado, tiene trabajo por temporadas, trabaja ocasionalmente o es trabajador calificado agropecuario y pesquero, el jefe de familia es trabajador calificado agropecuario y pesquero, , el jefe de hogar y el padre tienen estudios de 2° - 7° de Educación General Básica (primaria) o tercer nivel o licenciatura, la madre no trabaja, la casa es de Caña revestida, bahareque o madera, no dispone de alcantarillado, la vivienda está conectada a un pozo séptico.

4.4. Construir un modelo estadístico que permita el pronóstico del rendimiento académico de los grupos étnicos en el ciclo 2017 – 2018.

Para poder realizar el pronóstico de rendimiento se utilizó la técnica de Árbol de Decisión con el método CART (Classification and Regression Trees), a continuación, se presentan los resultados.

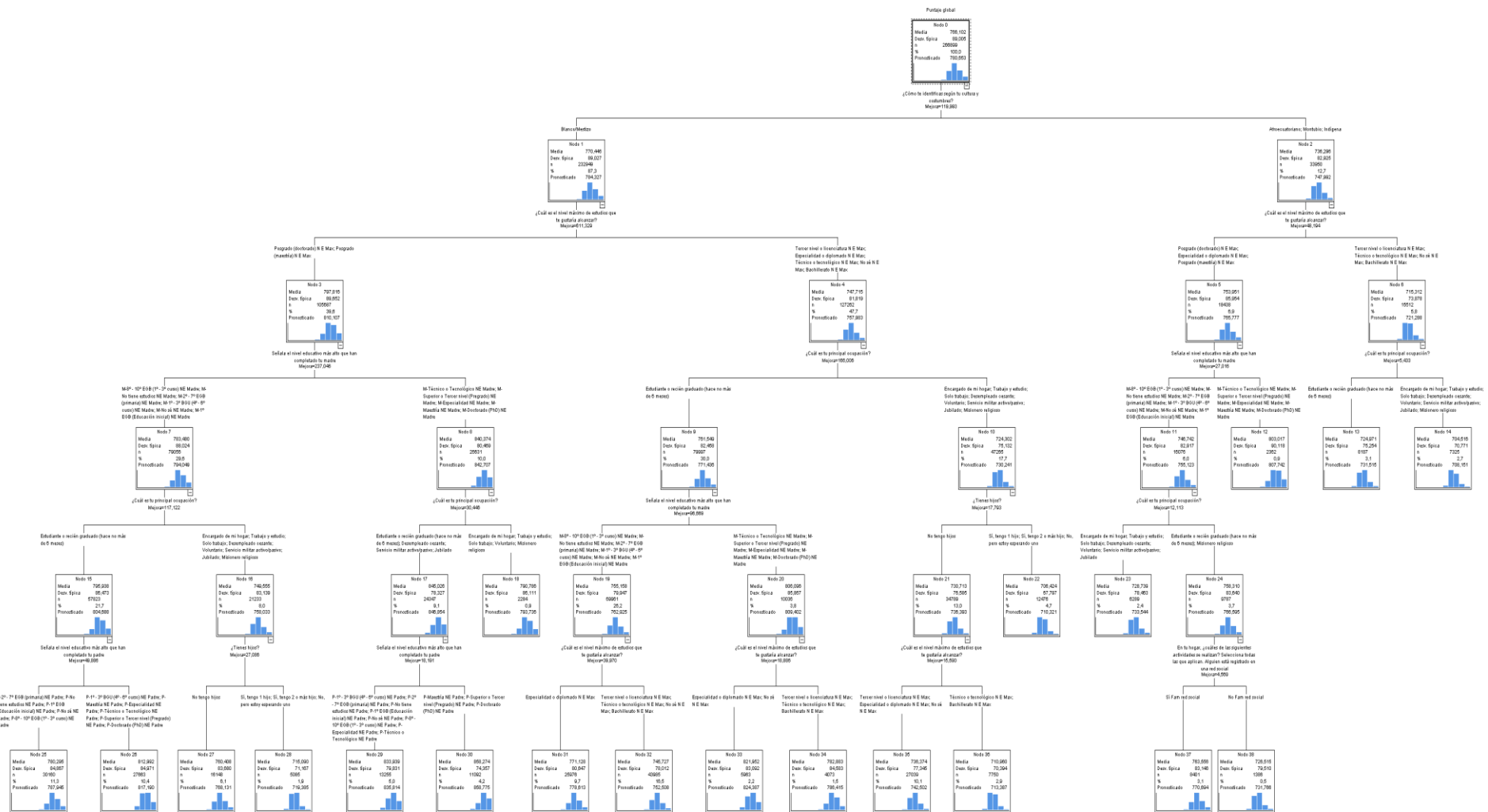


Ilustración 8: Diagrama de Árbol de Decisión grupo étnico blanco/mestizo

Tabla 13: Árbol de decisión del grupo étnico blanco/mestizo

Tabla de árbol nodo Blanco/Mestizo									
Nodo	Media	Desv. Desviación	N	Porcentaje	Media pronosticada	Nodo padre	Variable independiente primaria		
							Variable	Mejora	Valores de división
0	766,10	89,00	266.899	100%	780,65				
1	770,45	89,03	232.949	87%	784,33	0	¿Cómo te identificas según tu cultura y costumbres?	119,99	Blanco/Mestizo
3	797,82	89,65	105.687	40%	810,11	1	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	611,33	Posgrado (doctorado) N E Max; Posgrado (maestría) N E Max
4	747,72	81,82	127.262	48%	757,98	1	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	611,33	Tercer nivel o licenciatura N E Max; Especialidad o diplomado N E Max; Técnico o tecnológico N E Max; No sé N E Max; Bachillerato N E Max
7	783,48	88,02	79.056	30%	794,05	3	Señala el nivel educativo más alto que han completado tu madre	237,05	M-8º - 10º EGB (1º - 3º curso) NE Madre; M-No tiene estudios NE Madre; M-2º - 7º EGB (primaria) NE Madre; M-1º - 3º BGU (4º - 6º curso) NE Madre; M-No sé NE Madre; M-1º EGB (Educación inicial) NE Madre
8	840,37	80,47	26.631	10%	842,71	3	Señala el nivel educativo más alto que han completado tu madre	237,05	M-Técnico o Tecnológico NE Madre; M-Superior o Tercer nivel (Pregrado) NE Madre; M-Especialidad NE Madre; M-Maestría NE Madre; M-Doctorado (PhD) NE Madre
9	761,55	82,46	79.997	30%	771,44	4	¿Cuál es tu principal ocupación?	166,01	Estudiante o recién graduado (hace no más de 6 meses)
10	724,30	75,13	47.265	18%	730,24	4	¿Cuál es tu principal ocupación?	166,01	Encargado de mi hogar; Trabajo y estudio; Solo trabajo; Desempleado cesante; Voluntario; Servicio militar activo/pasivo; Jubilado; Misionero religioso
15	795,94	86,47	57.823	22%	804,69	7	¿Cuál es tu principal ocupación?	117,12	Estudiante o recién graduado (hace no más de 6 meses)
16	749,55	83,14	21.233	8%	758,03	7	¿Cuál es tu principal ocupación?	117,12	Encargado de mi hogar; Trabajo y estudio; Solo trabajo; Desempleado cesante; Voluntario; Servicio militar activo/pasivo; Jubilado; Misionero religioso
17	845,03	78,33	24.347	9%	846,85	8	¿Cuál es tu principal ocupación?	30,45	Estudiante o recién graduado (hace no más de 6 meses); Desempleado cesante; Servicio militar activo/pasivo; Jubilado
18	790,79	86,11	2.284	1%	793,74	8	¿Cuál es tu principal ocupación?	30,45	Encargado de mi hogar; Trabajo y estudio; Solo trabajo; Voluntario; Misionero religioso
19	755,16	79,95	69.961	26%	762,92	9	Señala el nivel educativo más alto que han completado tu madre	96,67	M-8º - 10º EGB (1º - 3º curso) NE Madre; M-No tiene estudios NE Madre; M-2º - 7º EGB (primaria) NE Madre; M-1º - 3º BGU (4º - 6º curso) NE Madre; M-No sé NE Madre; M-1º EGB (Educación inicial) NE Madre

20	806,10	85,87	10.036	4%	809,40	9	Señala el nivel educativo más alto que han completado tu madre	96,67	M-Técnico o Tecnológico NE Madre; M-Superior o Tercer nivel (Pregrado) NE Madre; M-Especialidad NE Madre; M-Maestría NE Madre; M-Doctorado (PhD) NE Madre
21	730,71	76,58	34.789	13%	736,39	10	¿Tienes hijos?	17,79	No tengo hijos
22	706,42	67,80	12.476	5%	710,32	10	¿Tienes hijos?	17,79	Sí, tengo 1 hijo; Sí, tengo 2 o más hijo; No, pero estoy esperando uno
25	780,30	84,87	30.160	11,3%	787,94	15	Señala el nivel educativo más alto que han completado tu padre	49,89	P-2º - 7º EGB (primaria) NE Padre; P-No tiene estudios NE Padre; P-1º EGB (Educación inicial) NE Padre; P-No sé NE Padre; P-8º - 10º EGB (1º - 3º curso) NE Padre
26	812,99	84,97	27.663	10,4%	817,19	15	Señala el nivel educativo más alto que han completado tu padre	49,89	P-1º - 3º BGU (4º - 6º curso) NE Padre; P-Maestría NE Padre; P-Especialidad NE Padre; P-Técnico o Tecnológico NE Padre; P-Superior o Tercer nivel (Pregrado) NE Padre; P-Doctorado (PhD) NE Padre
27	760,41	83,68	16.148	6%	768,13	16	¿Tienes hijos?	27,09	No tengo hijos
28	715,09	71,17	5.085	2%	719,38	16	¿Tienes hijos?	27,09	Sí, tengo 1 hijo; Sí, tengo 2 o más hijo; No, pero estoy esperando uno
29	833,94	79,83	13.255	5%	835,81	17	Señala el nivel educativo más alto que han completado tu padre	18,19	P-1º - 3º BGU (4º - 6º curso) NE Padre; P-2º - 7º EGB (primaria) NE Padre; P-No tiene estudios NE Padre; P-1º EGB (Educación inicial) NE Padre; P-No sé NE Padre; P-8º - 10º EGB (1º - 3º curso) NE Padre; P-Especialidad NE Padre; P-Técnico o Tecnológico NE Padre
30	858,27	74,36	11.092	4%	858,77	17	Señala el nivel educativo más alto que han completado tu padre	18,19	P-Maestría NE Padre; P-Superior o Tercer nivel (Pregrado) NE Padre; P-Doctorado (PhD) NE Padre
31	771,13	80,65	25.976	10%	778,61	19	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	39,97	Especialidad o diplomado N E Max
32	745,73	78,01	43.985	16%	752,51	19	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	39,97	Tercer nivel o licenciatura N E Max; Técnico o tecnológico N E Max; No sé N E Max; Bachillerato N E Max
33	821,95	83,09	5.963	2%	824,39	20	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	18,89	Especialidad o diplomado N E Max; No sé N E Max
34	782,88	84,58	4.073	2%	786,41	20	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	18,89	Tercer nivel o licenciatura N E Max; Técnico o tecnológico N E Max; Bachillerato N E Max
35	736,37	77,35	27.039	10%	742,50	21	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	15,59	Tercer nivel o licenciatura N E Max; Especialidad o diplomado N E Max; No sé N E Max
36	710,96	70,39	7.750	3%	713,39	21	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	15,59	Técnico o tecnológico N E Max; Bachillerato N E Max

4.4.1. Árbol de decisión de los factores que influyen en el rendimiento académico del grupo étnico blanco/mestizo

A partir de la Tabla 13, se tiene cuantificado los factores que influyen en el rendimiento académico, que parte del nodo del grupo étnico blanco/mestizo.

- Para los estudiantes que quieren alcanzar el nivel máximo de educación Posgrado (doctorado) y Posgrado (maestría) tienen un rendimiento promedio de 797,82 y el rendimiento pronosticado de 810,11, mientras que los estudiantes que aspiran nivel máximo de educación de Tercer nivel o licenciatura, Especialidad o diplomado, Técnico o tecnológico, No sabe, Bachillerato tienen un rendimiento promedio de 747,72 y pronosticado de 757,98
- Los estudiantes que tienen madres con nivel educativo de 8^o - 10^o EGB (1^o - 3^o curso), No tiene estudios, 2^o - 7^o EGB (primaria), 1^o - 3^o BGU (4^o - 6^o curso), No sabe, 1^o EGB (Educación inicial) el rendimiento promedio es de 783,48 y pronosticado de 794,05, mientras que los estudiantes que tienen madres con nivel educativo de Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Especialidad, Maestría, Doctorado (PhD) tiene un rendimiento promedio de 840,37 y pronosticado de 842,71.
- De los estudiantes que quieren llegar a un nivel máximo de educación de Tercer nivel o licenciatura, Especialidad o diplomado, Técnico o tecnológico, No sabe, Bachillerato, la ocupación principal del alumno es Estudiante o recién graduado (hace no más de 6 meses) su rendimiento promedio es de 761,55 y pronosticado de 771,44, mientras los de ocupación principal del alumno es Encargado de mi hogar, Trabajo y estudio, Solo trabajo, Desempleado cesante, Voluntario, Servicio militar activo/pasivo, Jubilado, Misionero religioso, tienen un rendimiento promedio de 724,30 y pronosticado de 730,24.
- De los estudiantes que tienen el nivel de preparación de la madre es de 8^o - 10^o EGB (1^o - 3^o curso), No tiene estudios, 2^o - 7^o EGB (primaria), 1^o - 3^o BGU (4^o - 6^o curso), No sabe, 1^o EGB (Educación inicial) la ocupación principal del alumno es Estudiante o recién graduado (hace no más de 6 meses) su rendimiento promedio es de 795,94 y pronosticado de 804,69, mientras los alumnos de ocupación principal Encargado de mi hogar, Trabajo y estudio, Solo trabajo, Desempleado cesante, Voluntario, Servicio militar activo/pasivo, Jubilado, Misionero religioso, tienen un rendimiento promedio de 749,55 y pronosticado de 758,03.
- De los estudiantes que el nivel de preparación de la madre es Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Especialidad, Maestría, Doctorado (PhD) la principal ocupación del alumno es Estudiante o recién graduado (hace no más de 6 meses), Desempleado cesante, Servicio militar activo/pasivo, Jubilado la ocupación principal del alumno es Estudiante o recién graduado (hace no más de 6 meses), Desempleado cesante, Servicio militar activo/pasivo, Jubilado su rendimiento promedio es 845,03 y

pronosticado 846,85, mientras de los alumnos que la ocupación principal es Encargado de mi hogar, Trabajo y estudio, Solo trabajo, Voluntario, Misionero religioso el rendimiento promedio es 790,79 y pronosticado 793,74.

- Lo estudiantes que su madre ha completado el nivel educativo 8^o - 10^o EGB (1^o - 3^o curso), No tiene estudios, 2^o - 7^o EGB (primaria), 1^o - 3^o BGU (4^o - 6^o curso), No sé, 1^o EGB (Educación inicial), el rendimiento promedio es 755,16 y pronosticado 762,92, mientras que los alumnos que su madre ha completado el nivel educativo Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Especialidad, Maestría, Doctorado (PhD) el rendimiento promedio es 806,10 y pronosticado de 809,40.
- Los estudiantes que no tienen hijos el rendimiento promedio es 730,71 y pronosticado de 736,39, mientras que los estudiantes que tienen hijos o están esperando uno el rendimiento promedio es de 706,42 y pronosticado 710,32.
- Los estudiantes que el padre tiene el nivel educativo completo 2^o - 7^o EGB (primaria), No tiene estudios, 1^o EGB (Educación inicial), No sé, 8^o - 10^o EGB (1^o - 3^o curso) el rendimiento promedio es 780,30 y el pronosticado es 787,94, mientras que los alumnos que el padre tienen nivel educativo completo 1^o - 3^o BGU (4^o - 6^o curso), Maestría, Especialidad, Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Doctorado (PhD) el rendimiento promedio es 812,99 y el pronosticado de 817,19.

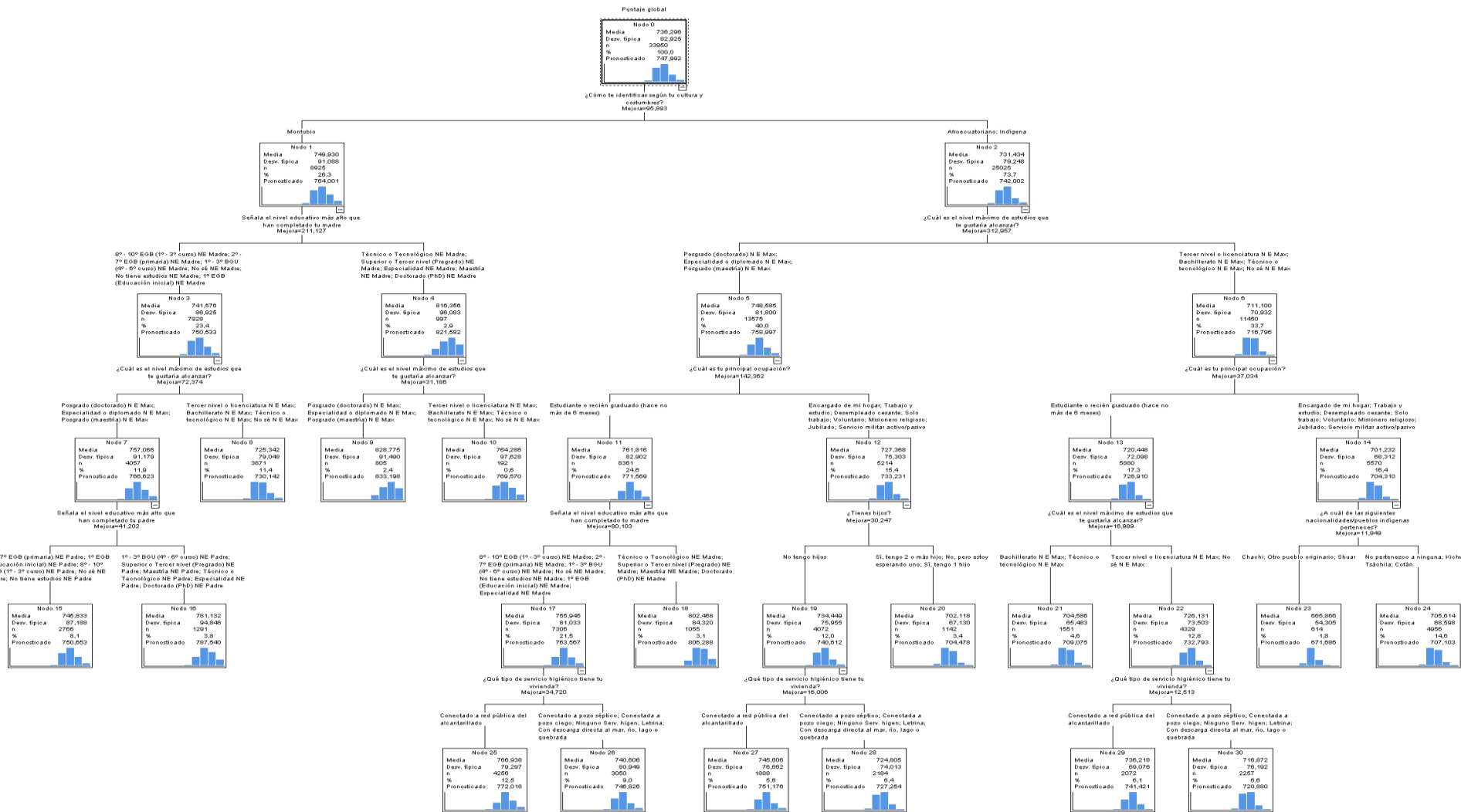


Ilustración 9: Diagrama de Árbol de Decisión de grupo étnico montubio, indígena y afroecuatoriano

Tabla 14: Árbol de decisión del grupo étnico montubio

Tabla de árbol nodo montubio									
Nodo	Media	Desv. Desviación	N	Porcentaje	Media pronosticada	Nodo padre	Variable independiente primaria		
							Variable	Mejora	Valores de división
0	736,30	82,92	33.950,00	100,0%	747,99				
1	749,93	91,09	8.925,00	26,3%	764,00	0	¿Cómo te identificas según tu cultura y costumbres?	95,89	Montubio
3	741,58	86,92	7.928,00	23,4%	750,53	1	Señala el nivel educativo más alto que han completado tu madre	211,13	8° - 10° EGB (1° - 3° curso) NE Madre; 2° - 7° EGB (primaria) NE Madre; 1° - 3° BGU (4° - 6° curso) NE Madre; No sé NE Madre; No tiene estudios NE Madre; 1° EGB (Educación inicial) NE Madre
4	816,36	96,08	997,00	2,9%	821,58	1	Señala el nivel educativo más alto que han completado tu madre	211,13	Técnico o Tecnológico NE Madre; Superior o Tercer nivel (Pregrado) NE Madre; Especialidad NE Madre; Maestría NE Madre; Doctorado (PhD) NE Madre
7	757,07	91,18	4.057,00	11,9%	766,62	3	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	72,37	Posgrado (doctorado) N E Max; Especialidad o diplomado N E Max; Posgrado (maestría) N E Max
8	725,34	79,05	3.871,00	11,4%	730,14	3	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	72,37	Tercer nivel o licenciatura N E Max; Bachillerato N E Max; Técnico o tecnológico N E Max; No sé N E Max
9	828,78	91,49	805,00	2,4%	833,20	4	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	31,19	Posgrado (doctorado) N E Max; Especialidad o diplomado N E Max; Posgrado (maestría) N E Max
10	764,29	97,63	192,00	0,6%	769,57	4	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	31,19	Tercer nivel o licenciatura N E Max; Bachillerato N E Max; Técnico o tecnológico N E Max; No sé N E Max
15	745,83	87,19	2.766,00	8,1%	750,65	7	Señala el nivel educativo más alto que han completado tu padre	41,20	2° - 7° EGB (primaria) NE Padre; 1° EGB (Educación inicial) NE Padre; 8° - 10° EGB (1° - 3° curso) NE Padre; No sé NE Padre; No tiene estudios NE Padre
16	781,13	94,85	1.291,00	3,8%	787,54	7	Señala el nivel educativo más alto que han completado tu padre	41,20	1° - 3° BGU (4° - 6° curso) NE Padre; Superior o Tercer nivel (Pregrado) NE Padre; Maestría NE Padre; Técnico o Tecnológico NE Padre; Especialidad NE Padre; Doctorado (PhD) NE Padre

4.4.2. Árbol de decisión de los factores que influyen en el rendimiento académico del grupo étnico montubio

De la Tabla 14 del árbol de decisión se tiene cuantificado los factores que influyen en el rendimiento académico del nodo grupo étnico montubio.

- De los estudiantes que el nivel educativo más alto de la madre es 8º - 10º EGB (1º - 3º curso), 2º - 7º EGB (primaria), 1º - 3º BGU (4º - 6º curso), No sé, No tiene estudios, 1º EGB (Educación inicial), tiene un rendimiento promedio de 741,58 y pronosticado de 750,53 mientras que de los alumnos que la madre tiene el nivel educativo Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Especialidad, Maestría, Doctorado (PhD), el rendimiento promedio es 816,36 y pronosticado de 821,58.
- Si la madre tiene nivel educativo 8º - 10º EGB (1º - 3º curso), 2º - 7º EGB (primaria), 1º - 3º BGU (4º - 6º curso), No sé, No tiene estudios, 1º EGB (Educación inicial), los estudiantes aspiran a un nivel de máximo de educación Posgrado (doctorado), Especialidad o diplomado, Posgrado (maestría) el rendimiento promedio es 757,07 y pronosticado de 766,62, mientras que los estudiantes que aspiran a Tercer nivel o licenciatura, Bachillerato, Técnico o tecnológico, No sé el rendimiento promedio es 725,34 y pronosticado 730,14.
- Si la madre tiene nivel educativo Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Especialidad, Maestría, Doctorado (PhD), los estudiantes aspiran a un nivel de máximo de educación Posgrado (doctorado), Especialidad o diplomado, Posgrado (maestría) el rendimiento promedio es 828,78 y pronosticado de 833,20, mientras que los estudiantes que aspiran a Tercer nivel o licenciatura, Bachillerato, Técnico o tecnológico, No sé el rendimiento promedio es 764,29 y pronosticado 769,57.
- De los estudiantes que el nivel de estudio del padre es 2º - 7º EGB (primaria), 1º EGB (Educación inicial), 8º - 10º EGB (1º - 3º curso), No sé, No tiene estudios, el rendimiento promedio 745,83 y pronosticado 750,65, mientras que los alumnos que tienen padres con nivel de estudios 1º - 3º BGU (4º - 6º curso), Superior o Tercer nivel (Pregrado), Maestría, Técnico o Tecnológico, Especialidad, Doctorado (PhD) el rendimiento promedio es 781,13 y pronosticado 787,54

Tabla 15: Árbol de decisión del grupo étnico indígena y afroecuatoriano

Tabla de árbol nodo indígena y afroecuatoriano									
Nodo	Media	Desv. Desviación	N	Porcentaje	Media pronosticada	Nodo padre	Variable independiente primaria		
							Variable	Mejora	Valores de división
0	736,30	82,92	33.950	100,0%	747,99				
2	731,43	79,25	25.025	73,7%	742,00	0	¿Cómo te identificas según tu cultura y costumbres?	95,89	Afroecuatoriano; Indígena
5	748,58	81,80	13.575	40,0%	759,00	2	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	312,96	Posgrado (doctorado) N E Max; Especialidad o diplomado N E Max; Posgrado (maestría) N E Max
6	711,10	70,93	11.450	33,7%	716,80	2	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	312,96	Tercer nivel o licenciatura N E Max; Bachillerato N E Max; Técnico o tecnológico N E Max; No sé N E Max
11	761,82	82,90	8.361	24,6%	771,57	5	¿Cuál es tu principal ocupación?	142,36	Estudiante o recién graduado (hace no más de 6 meses)
12	727,37	75,30	5.214	15,4%	733,23	5	¿Cuál es tu principal ocupación?	142,36	Encargado de mi hogar; Trabajo y estudio; Desempleado cesante; Solo trabajo; Voluntario; Misionero religioso; Jubilado; Servicio militar activo/pasivo
13	720,45	72,10	5.880	17,3%	726,91	6	¿Cuál es tu principal ocupación?	37,03	Estudiante o recién graduado (hace no más de 6 meses)
14	701,23	68,31	5.570	16,4%	704,31	6	¿Cuál es tu principal ocupación?	37,03	Encargado de mi hogar; Trabajo y estudio; Desempleado cesante; Solo trabajo; Voluntario; Misionero religioso; Jubilado; Servicio militar activo/pasivo
17	755,95	81,03	7.306	21,5%	763,57	11	Señala el nivel educativo más alto que han completado tu madre	80,10	8º - 10º EGB (1º - 3º curso) NE Madre; 2º - 7º EGB (primaria) NE Madre; 1º - 3º BGU (4º - 6º curso) NE Madre; No sé NE Madre; No tiene estudios NE Madre; 1º EGB (Educación inicial) NE Madre; Especialidad NE Madre
18	802,47	84,32	1.055	3,1%	806,29	11	Señala el nivel educativo más alto que han completado tu madre	80,10	Técnico o Tecnológico NE Madre; Superior o Tercer nivel (Pregrado) NE Madre; Maestría NE Madre; Doctorado (PhD) NE Madre
19	734,45	75,96	4.072	12,0%	740,61	12	¿Tienes hijos?	30,25	No tengo hijos

20	702,12	67,13	1.142	3,4%	704,48	12	¿Tienes hijos?	30,25	Sí, tengo 2 o más hijo; No, pero estoy esperando uno; Sí, tengo 1 hijo
21	704,59	65,48	1.551	4,6%	709,07	13	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	16,99	Bachillerato N E Max; Técnico o tecnológico N E Max
22	726,13	73,50	4.329	12,8%	732,79	13	¿Cuál es el nivel máximo de estudios que te gustaría alcanzar?	16,99	Tercer nivel o licenciatura N E Max; No sé N E Max
23	665,87	54,30	614	1,8%	671,69	14	¿A cuál de las siguientes nacionalidades/pueblos indígenas perteneces?	11,95	Chachi; Otro pueblo originario; Shuar
24	705,61	68,60	4.956	14,6%	707,10	14	¿A cuál de las siguientes nacionalidades/pueblos indígenas perteneces?	11,95	No pertenezco a ninguna; Kichwa; Tsáchila; Cofán
25	766,94	79,30	4.256	12,5%	772,02	17	¿Qué tipo de servicio higiénico tiene tu vivienda?	34,72	Conectado a red pública del alcantarillado
26	740,61	80,95	3.050	9,0%	746,83	17	¿Qué tipo de servicio higiénico tiene tu vivienda?	34,72	Conectado a pozo séptico; Conectada a pozo ciego; Ninguno Serv. higen; Letrina; Con descarga directa al mar, río, lago o quebrada
27	745,61	76,66	1.888	5,6%	751,18	19	¿Qué tipo de servicio higiénico tiene tu vivienda?	16,01	Conectado a red pública del alcantarillado
28	724,80	74,01	2.184	6,4%	727,25	19	¿Qué tipo de servicio higiénico tiene tu vivienda?	16,01	Conectado a pozo séptico; Conectada a pozo ciego; Ninguno Serv. higen; Letrina; Con descarga directa al mar, río, lago o quebrada
29	736,22	69,08	2.072	6,1%	741,42	22	¿Qué tipo de servicio higiénico tiene tu vivienda?	12,51	Conectado a red pública del alcantarillado
30	716,87	76,19	2.257	6,6%	720,88	22	¿Qué tipo de servicio higiénico tiene tu vivienda?	12,51	Conectado a pozo séptico; Conectada a pozo ciego; Ninguno Serv. higen; Letrina; Con descarga directa al mar, río, lago o quebrada

4.4.3. Árbol de decisión de los factores que influyen en el rendimiento académico del grupo étnico indígena y afroecuatoriano

Debido al comportamiento similar entre los grupos étnicos indígena y afroecuatoriano se realizó el análisis en conjunto de los factores que influyen en el rendimiento académico del nodo indígena y afroecuatoriano que presenta la Tabla 15.

- Los estudiantes que les gustaría alcanzar el nivel máximo de educación Posgrado (doctorado), Especialidad o diplomado, Posgrado (maestría), el rendimiento promedio es 748,58 y pronosticado 759,00 mientras que los alumnos desean alcanzar el nivel de educación Tercer nivel o licenciatura, Bachillerato, Técnico o tecnológico, No sabe el rendimiento promedio es 711,10 y pronosticado de 716,80.
- De los estudiantes que les gustaría alcanzar el nivel máximo de educación Posgrado (doctorado), Especialidad o diplomado, Posgrado (maestría), y su principal ocupación es Estudiante o recién graduado (hace no más de 6 meses) su rendimiento promedio es 761,82 y pronosticado 771,57, mientras los alumnos que su principal ocupación es Encargado de mi hogar; Trabajo y estudio; Desempleado cesante; Solo trabajo; Voluntario; Misionero religioso; Jubilado; Servicio militar activo/pasivo su rendimiento promedio es 727,37 y pronosticado 733,23.
- De los estudiantes que les gustaría alcanzar el nivel máximo de educación Tercer nivel o licenciatura, Bachillerato, Técnico o tecnológico, No sabe y su principal ocupación es Estudiante o recién graduado (hace no más de 6 meses) su rendimiento promedio es 720,45 y pronosticado 726,91, mientras los alumnos que su principal ocupación es Encargado de mi hogar; Trabajo y estudio; Desempleado cesante; Solo trabajo; Voluntario; Misionero religioso; Jubilado; Servicio militar activo/pasivo su rendimiento promedio es 701,23 y pronosticado 704,31.
- De los estudiantes que su madre a completado el nivel de estudios 8^o - 10^o EGB (1^o - 3^o curso), 2^o - 7^o EGB (primaria), 1^o - 3^o BGU (4^o - 6^o curso), No sé, No tiene estudios, 1^o EGB (Educación inicial), Especialidad el rendimiento promedio es 755,95 y pronosticado 763,57, mientras que los alumnos que su madre a completado en nivel de estudios Técnico o Tecnológico, Superior o Tercer nivel (Pregrado), Maestría, Doctorado (PhD), su rendimiento promedio es 802,47 y pronosticado 806,29.
- De los estudiantes que no tienen hijos su rendimiento promedio es de 734,45 y pronosticado 740,61, mientras que los estudiantes q tienen o están esperando hijos el rendimiento promedio es 702,12 y pronosticado 704,48.
- De los estudiantes que pertenecen a las nacionalidades Chachi, Otro pueblo originario, Shuar, el rendimiento promedio es 665,87 y pronosticado 671,69 mientras que los alumnos que pertenecen a las nacionalidades No pertenezco a ninguna, Kichwa, Tsáchila, Cofán, el rendimiento promedio es 705,61 y pronosticado 707,10,

- De los estudiantes que su vivienda está Conectado a red pública del alcantarillado el rendimiento promedio es de 745,61 y pronosticado 751,18, mientras que los alumnos que su vivienda tiene Conectado a pozo séptico, Conectada a pozo ciego, Ningún Servicio higiénico, Letrina, Con descarga directa al mar, río, lago o quebrada su rendimiento promedio es de 740,61 y pronosticado 746,83.

CAPÍTULO 5

5. CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

A partir de los resultados del capítulo 4 se puede concluir:

1. Existe diferencia de rendimiento académico entre los grupos étnicos, donde los mestizos/blancos tienen mejor rendimiento que el resto.
2. De 295 variables de factores asociados del estudiante se redujo a 54 variables sin perder los patrones de la base de datos.
3. El rendimiento de los estudiantes que pertenecen o se autoidentifican en el grupo étnico blanco/mestizo se ven afectados debido que el jefe de familia o el padre han completado el nivel educativo de bachillerato, la labor principal a la que se dedica el padre es brindar servicios, comerciantes o la manufactura (operadores de instalaciones, maquinaria, ensambladores), en sus viviendas tienen todos los servicios básicos (luz eléctrica, ducha, alcantarillado); Los estudiantes viven con el padre o la madre, no viven con pareja; les gusta leer, se sienten seguros en la institución a la que pertenecen.
4. Además, los estudiantes que tienen aspiraciones de tener un título de cuarto nivel (maestría, doctorado) poseen mejor rendimiento de los estudiantes que no tienen esas aspiraciones. Si el nivel educativo de la madre tiene tercer nivel o más es mejor el rendimiento académico que las que no poseen. Los estudiantes que solo se dedican a estudiar tienen mejor rendimiento de los que tienen alguna otra labor. Los estudiantes que no tienen hijos tienen mejores calificaciones de los que tienen 1 hijo o más. Los estudiantes que, el padre tiene estudios de tercer nivel o más, poseen mejor rendimiento comparado con los que no tienen.
5. De los resultados de los estudiantes que pertenecen o se autoidentifican como montubios el rendimiento se ve afectado por la ocupación del padre o jefe de familia es limpiadores, peones, ayudantes de preparación de alimentos, vendedores ambulantes, trabajo no remunerado, trabaja por temporadas, trabaja ocasionalmente o es trabajador agropecuario y pesquero estos tienen estudios primarios o de tercer nivel. La madre no trabaja. La vivienda es de caña revestida, bahareque o madera, tiene pozo séptico.
6. Además, los estudiantes que la madre tiene un nivel de estudio de tercer nivel o mayor tienen mejor rendimiento de los estudiantes que las madres no. Los estudiantes que tienen aspiración de tener un título de cuarto nivel tienen mejor rendimiento que los que no tienen esta motivación. Los estudiantes que el padre

tiene nivel de estudios de bachillerato o más, tienen mejor rendimiento de los que no tienen.

7. El estudiante indígena es jefe de hogar por lo tanto, es encargado de todo el hogar, vive con su pareja tienen un hijo, el padre y la madre tienen nivel de estudios de primaria o no disponen, la vivienda no tiene ducha y solo tienen letrina, el nivel de logro en la prueba "Ser bachiller" es insuficiente.
8. El estudiante afroecuatoriano tiene como jefe de hogar a sus abuelos, hermano u otras personas, estos están en la desocupación y su nivel de estudios es de educación inicial (1° EGB). El estudiante no tiene hijos, pero está esperando uno, falta ocasionalmente a clases, no entra algunas materias y periódicamente llega atrasado a clases.
9. Además, los factores que influyen en el rendimiento de los estudiantes que pertenecen a los grupos étnicos afroecuatorianos e indígenas son los que aspiran llegar al cuarto nivel de estudios, tienen mejor notas que los que no desean. Los estudiantes cuya principal ocupación es estudiar tienen mejor rendimiento que los que tienen otras ocupaciones (ajo, voluntariado, etc). Los estudiantes que el nivel educativo de la madre es de tercer nivel o más tienen mejores calificaciones de las madres que no tienen. Los que no tienen hijos, sus calificaciones son mejores de los que están esperando o tienen hijos. Los estudiantes que tienen alcantarillado en su vivienda tienen mejor rendimiento de los que no tienen.

5.2. Recomendaciones

1. Se recomienda realizar una investigación de los factores que afectan el rendimiento de los grupos étnicos por campo de estudio (Matemáticas, Lenguaje, Ciencias sociales, Ciencias Naturales).
2. Se recomienda realizar un estudio más profundo sobre los factores asociados que afectan al grupo étnico afroecuatoriano e indígena debido a que estos tienen una problemática mayor en su rendimiento académico comparando con las otras.

6. Referencias

- Arroyo, J. (31 de 01 de 2016). Métodos de reducción de dimensionalidad: Análisis comparativo de los métodos APC, ACPP y ACPK. *30, 1*, 115-122. Heredia, Costa Rica: UNICIENCIA. doi:<http://dx.doi.org/10.15359/ru.30-1.7>
- Autiero , G., & Nese, A. (21 de Noviembre de 2017). Ethnic and Academic Identity: What Role for Children’s Scholastic Effort? Descripción, Salerno, Italia: Forum for Social Economics. doi:10.1080/07360932.2017.1394899
- Ba, Z., Han, Z., Gong, Z., Li, F., Zhang, H., & Zhang, G. (Septiembre de 2019). Ethnic differences in experiences of school bullying in China. *104*. (ScienceDirect, Ed.) China: Children and Youth Services Review. doi:<https://doi.org/10.1016/j.chilyouth.2019.104402>
- De la Fuente Fernández, S. (2011). Análisis de correspondencias simples y múltiples. Madrid, España: Universidad Autonoma de Madrid. Obtenido de <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/REDUCIR-DIMENSION/CORRESPONDENCIAS/correspondencias.pdf>
- De’ath, G., & Fabricius, K. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *81, 11*, 3178–3192. Ecology.
- Debrosse, R., Rossignac-Milon, M., Taylor, D., & Destin, M. (7 de Junio de 2018). Can Identity Conflicts Impede the Success of Ethnic Minority Students? Consequences of Discrepancies Between Ethnic and Ideal Selves. *44*, 1725 - 1738. Personality and Social Psychology Bulletin. doi:<https://doi.org/10.1177/0146167218777997>
- Deconinck, E., Zhang, M., Coomans, D., & Heyden, Y. (2006). Classification tree models for the prediction of blood-brain barrier passage of drugs. *46, 3*, 1410–1419. Journal of Quemical Information and Modeling.
- Fuente Fernández, S. (2011). Análisis de correspondencias simples y múltiples. madrid, España: Facultad de Ciencias Economicas y Empresariales. Obtenido de <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/REDUCIR-DIMENSION/CORRESPONDENCIAS/correspondencias.pdf>
- Jambu, M. (1989). L'Exploration informatique et statistique des données. Paris, Francia.
- ONU. (2015). *Objetivos de Desarrollo Sostenibles*. Obtenido de <https://www.un.org/sustainabledevelopment/es/>
- Tapia, J. E. (Agosto de 2007). El escalonamiento óptimo con base en el análisis de componentes principales no lineales para la construcción de índices de condiciones de vida y socioeconómicos. *El escalonamiento óptimo con base en el análisis de componentes principales no lineales para la construcción de índices de condiciones de vida y socioeconómicos*, 34 - 36. Quito, pichincha, Ecuador: Escuela Superior Politecnica Nacional. Obtenido de <https://bibdigital.epn.edu.ec/bitstream/15000/537/1/CD-1037.pdf>

- Timofeev, R. (20 de Diciembre de 2004). Classification and Regression Trees (CART). Berlin, Alemania: Humboldt University. Obtenido de https://pdfs.semanticscholar.org/c77b/4aff121e3c2c2c18bef72b36286bb3a6d375.pdf?_ga=2.134514353.1506869224.1561853007-786490804.1561853007
- UNESCO. (2005). La educación como derecho humano. UNESCO. Obtenido de <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=5&cad=rja&uact=8&ved=2ahUKEwjLvpjDqvTIAhXjw1kKHdt4D1gQFjAEegQIBhAI&url=http%3A%2F%2Fwww.unescoetxea.org%2Fdokumentuwww.unescoetxea.org>
- Unesco. (09 de 07 de 2019). *Data for Sustainable Development*. Obtenido de The World is Off Track to Deliver on its Education Commitments by 2030: <http://uis.unesco.org/en/blog/world-track-deliver-its-education-commitments-2030>

7. Apéndices y anexos

Matriz de carga de variables rotada

	Po nd er aci ón de var iab le	Dimensión																															
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
¿Tienes hijos?	1	-0,825	-0,004	-0,072	-0,109	-0,015	-0,071	0,047	0,016	-0,021	0,094	-0,054	0,014	-0,040	-0,083	0,060	0,003	-0,001	-0,049	-0,041	0,000	-0,009	-0,176	-0,087	-0,022	0,099	0,004	0,029	-0,041	0,019	-0,066	0,009	0,042
¿Con quién vives? Selecciona todas las que aplican. Hijo(s)	1	-0,808	-0,012	-0,060	-0,105	-0,011	-0,073	0,048	0,022	-0,021	0,088	-0,053	0,015	-0,036	-0,098	0,073	0,005	0,003	-0,039	-0,045	0,001	-0,002	-0,168	-0,096	-0,018	0,108	0,004	0,039	-0,049	0,020	-0,074	-0,028	0,048
¿Con quién vives? Selecciona todas las que aplican. Cónyuge, conviviente o pareja	1	-0,749	-0,012	-0,038	-0,054	0,002	-0,024	-0,020	-0,021	-0,034	0,077	-0,065	0,003	-0,015	-0,024	0,028	0,002	-0,025	-0,023	-0,002	-0,007	-0,012	-0,105	-0,031	-0,011	0,022	0,032	0,018	-0,039	-0,021	-0,016	-0,036	0,028
¿Quién es el jefe de tu hogar?	1	0,708	-0,046	0,040	-0,037	0,004	-0,061	0,036	0,139	0,013	-0,064	-0,008	-0,001	-0,013	0,026	0,122	-0,008	0,077	-0,007	0,018	0,045	0,007	-0,380	-0,090	-0,004	0,144	-0,009	0,114	-0,024	-0,001	-0,084	-0,133	0,039
¿Con quién vives? Selecciona todas las que aplican. Madre	1	0,639	-0,026	0,030	-0,023	-0,007	-0,020	0,054	0,116	0,022	-0,041	0,009	0,012	-0,015	-0,004	0,137	-0,001	0,052	0,018	-0,035	0,014	0,017	-0,263	-0,097	-0,009	0,145	0,005	0,057	-0,013	0,081	-0,061	0,047	0,027
¿Con quién vives? Selecciona todas las que aplican. Hermanos	1	0,550	-0,044	0,012	-0,030	-0,041	-0,105	0,060	0,127	-0,004	-0,002	0,046	0,032	-0,015	0,013	0,144	0,001	0,079	0,029	-0,159	-0,007	0,070	-0,187	-0,131	-0,029	0,160	0,016	0,073	-0,075	0,063	-0,063	0,099	0,040
Estado civil del sustentante	1	-0,513	-0,015	-0,033	-0,058	-0,011	-0,030	0,037	0,063	0,008	-0,029	-0,017	0,000	-0,010	0,019	0,034	-0,007	-0,005	-0,034	0,011	0,004	-0,025	-0,114	-0,024	-0,003	0,070	-0,006	0,054	0,005	-0,018	-0,044	0,079	0,002
¿Con quién vives? Selecciona todas las que aplican. Padre	1	0,459	-0,061	0,028	-0,057	0,016	-0,127	0,009	0,197	0,006	-0,071	-0,051	0,001	-0,030	0,039	0,154	-0,016	0,145	-0,039	0,044	0,061	-0,015	-0,341	-0,122	0,007	0,193	-0,003	0,221	-0,054	-0,081	-0,118	0,020	0,068
¿Con quién vives? Selecciona todas las que aplican. Otros	1	-0,430	0,004	-0,016	0,021	-0,034	-0,065	0,019	-0,021	0,008	0,032	0,141	-0,002	-0,021	-0,047	0,070	0,006	0,116	0,041	-0,181	-0,025	0,047	0,052	-0,055	-0,011	0,040	0,018	-0,005	0,004	0,024	0,020	-0,023	0,023
¿Cuál es tu principal ocupación?	1	0,336	0,050	0,045	0,049	0,058	0,219	0,052	0,030	0,014	-0,133	0,230	-0,064	-0,016	-0,244	-0,015	-0,023	0,018	0,027	0,024	-0,047	-0,013	0,032	0,105	-0,035	-0,051	-0,048	0,016	0,042	0,015	-0,002	-0,229	0,046
¿Cómo te identificas según tu cultura y costumbres?	3	-0,009	0,990	0,006	0,031	0,010	0,035	0,039	0,024	0,002	0,016	0,048	-0,003	0,001	-0,003	0,079	-0,002	0,018	0,007	-0,008	0,013	-0,004	-0,002	0,055	-0,005	-0,009	-0,010	0,013	-0,002	0,017	0,016	0,000	0,000
¿A cuál de las siguientes nacionalidades/pueblos indígenas perteneces?	1	0,009	-0,989	-0,006	-0,031	-0,010	-0,035	-0,039	-0,024	-0,002	-0,016	-0,048	0,003	-0,001	0,003	-0,079	0,002	-0,018	-0,007	0,008	-0,013	0,004	0,002	-0,055	0,005	0,009	0,011	-0,013	0,001	-0,017	-0,016	0,000	0,000
¿Los maestros llegaban tarde a clases?	1	0,073	0,010	0,743	0,054	0,086	0,019	0,013	0,024	0,065	0,051	0,051	0,007	0,084	-0,013	0,014	0,008	0,016	0,146	-0,017	0,007	-0,010	0,008	0,007	0,068	-0,002	0,004	0,002	0,016	0,018	0,020	-0,012	0,007
¿Los maestros faltaban a clases?	1	0,078	0,018	0,725	0,024	0,078	0,011	-0,015	0,013	0,013	0,059	0,046	-0,015	0,079	-0,036	0,019	-0,007	0,004	0,137	-0,025	-0,001	-0,039	0,013	-0,005	0,059	-0,002	0,012	0,002	-0,010	0,002	0,032	-0,010	0,002
¿Los maestros se iban más temprano?	1	0,036	-0,018	0,721	0,027	0,049	-0,006	0,003	0,018	0,063	0,045	0,011	0,032	0,039	0,022	-0,010	0,023	0,007	-0,021	0,023	0,008	0,087	0,012	0,003	0,085	0,011	-0,005	0,013	-0,010	0,011	-0,015	-0,009	0,003

¿Los maestros nos dejaban solos en las horas de clase?	1	0,025	0,005	0,705	0,044	0,075	0,004	0,029	0,038	0,083	0,044	0,042	0,043	0,061	0,016	0,000	0,018	0,007	0,208	0,004	0,017	0,051	0,010	0,021	0,023	0,005	-0,010	0,004	-0,009	0,024	-0,016	-0,010	0,013
¿Los maestros nos dejaban hacer las actividades que deseamos, mientras él hace otras cosas?	1	0,017	0,004	0,602	0,020	0,052	-0,008	-0,025	0,027	0,072	0,042	-0,031	0,061	0,011	0,027	0,040	0,022	-0,002	0,257	0,004	0,011	0,099	0,015	0,017	-0,018	0,003	-0,004	0,004	-0,107	0,004	-0,032	0,003	0,023
¿Los maestros se preocupaban de que aprovechemos el tiempo al máximo?	1	0,019	0,011	0,520	0,034	0,058	0,061	0,047	-0,018	0,229	-0,012	0,036	0,006	0,249	0,002	0,027	0,026	0,023	-0,131	0,005	0,009	0,252	-0,025	0,008	0,068	-0,019	-0,015	0,010	0,163	-0,023	0,013	0,002	0,035
En tu colegio, ¿con qué frecuencia pasaban estas cosas? Las mayoría de los maestros están interesados en que los estudiantes estemos bien	1	0,035	0,002	0,411	0,056	0,060	0,099	0,086	-0,021	0,269	-0,025	0,052	0,003	0,345	-0,016	0,040	0,120	0,023	-0,099	-0,007	0,015	0,188	-0,038	-0,009	0,099	-0,036	0,005	0,017	0,306	-0,018	0,020	0,027	0,063
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Cámara de video digital	1	0,018	0,018	0,042	0,627	-0,005	0,061	0,058	0,118	0,014	0,020	0,027	0,028	0,009	0,036	0,102	0,008	0,009	-0,010	0,080	0,033	0,025	-0,025	0,012	0,034	0,018	-0,016	0,011	-0,008	0,025	-0,029	0,032	0,000
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Computadora portátil (laptop)	1	0,038	0,013	0,037	0,617	0,008	0,193	0,050	0,244	0,035	-0,039	0,133	-0,008	0,010	-0,044	-0,090	-0,003	0,216	-0,013	0,005	0,088	0,010	-0,014	0,021	-0,008	-0,081	-0,005	0,069	0,000	0,077	0,107	0,002	0,035
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Videojuegos	1	0,012	0,029	0,044	0,593	0,043	0,115	0,063	0,074	0,026	0,035	0,036	0,024	0,013	0,183	0,117	-0,005	0,052	0,033	0,014	0,034	0,034	-0,023	0,026	0,015	-0,041	-0,021	0,027	-0,036	0,024	-0,029	0,020	0,002
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Radios o equipos de sonido	1	0,062	-0,003	0,023	0,568	0,033	-0,154	0,068	0,096	-0,008	0,010	0,054	-0,016	0,003	0,015	0,217	-0,005	0,055	0,079	0,145	0,033	0,026	0,079	-0,009	-0,033	0,103	0,072	0,018	0,070	0,038	0,012	-0,068	0,024
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Televisores	1	-0,044	-0,066	-0,026	-0,473	-0,089	0,059	-0,166	-0,131	-0,008	-0,013	-0,108	0,041	-0,002	0,001	-0,390	0,012	-0,197	-0,064	-0,115	-0,045	-0,008	-0,080	-0,091	0,050	0,006	-0,030	-0,044	-0,045	-0,077	-0,003	0,097	0,023
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Microondas	1	0,017	0,049	0,025	0,400	0,041	0,139	0,127	0,153	0,071	0,011	0,045	0,019	-0,011	0,017	0,303	0,001	0,110	-0,005	0,119	0,071	-0,004	0,038	0,107	0,014	-0,033	-0,027	0,046	0,084	0,078	-0,077	-0,024	0,031
En tu hogar, ¿cuáles de las siguientes actividades se realizan? Selecciona todas las que aplican. Compran ropa en centros comerciales	1	0,046	0,004	0,051	0,396	0,035	0,010	0,021	-0,075	-0,071	0,079	0,075	0,018	0,022	-0,032	0,244	-0,010	-0,023	0,052	0,260	0,079	0,037	0,010	0,050	-0,007	-0,142	0,041	0,082	-0,147	0,038	-0,052	0,046	0,044
De los siguientes bienes, ¿cuántos hay en	1	0,074	0,032	0,026	0,374	0,063	0,093	0,036	0,186	0,046	-0,026	0,048	-0,016	-0,023	-0,001	0,092	-0,011	0,355	-0,022	0,090	0,161	-0,018	-0,062	0,047	0,031	-0,019	-0,030	0,147	0,015	-0,030	-0,018	-0,021	0,031

¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Agua potable o entubada	1	0,000	-0,013	0,017	0,080	0,015	0,003	0,622	0,074	0,012	-0,006	0,034	-0,005	0,000	-0,001	-0,002	0,000	0,030	-0,006	0,034	0,039	0,005	0,002	0,151	-0,001	-0,003	0,001	0,005	-0,028	0,033	0,275	-0,004	0,007
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Recolección de basura	1	-0,053	0,053	0,018	0,039	-0,005	0,082	0,527	0,054	-0,016	-0,009	0,111	-0,005	0,026	-0,003	0,160	-0,004	0,003	0,020	0,016	0,019	0,001	-0,023	0,107	-0,001	-0,031	-0,031	-0,022	-0,174	0,040	0,225	0,042	0,005
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Computadora de escritorio (desktop)	1	0,063	0,020	0,033	0,128	0,013	0,066	0,086	0,780	0,004	-0,009	0,059	-0,003	0,039	0,022	0,122	0,007	0,001	0,019	0,082	0,038	0,015	0,031	0,104	0,010	-0,004	-0,020	-0,013	0,027	0,023	-0,014	-0,003	0,010
De los siguientes bienes, ¿Cuántos hay en tu hogar? Computadores	1	0,075	0,023	0,050	0,506	0,017	0,182	0,104	0,666	0,031	-0,043	0,154	-0,013	0,034	-0,019	0,016	0,002	0,151	0,003	0,063	0,083	0,014	0,009	0,091	0,000	-0,081	-0,018	0,039	0,021	0,074	0,074	-0,003	0,019
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Impresora	1	0,065	0,029	0,033	0,303	0,025	0,137	0,040	0,639	0,001	-0,033	0,064	-0,009	0,020	-0,024	0,089	0,001	0,082	0,003	0,096	0,066	-0,004	0,000	0,090	0,010	-0,044	-0,020	0,044	0,020	0,052	0,018	0,004	0,000
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Conexión a internet	1	0,071	0,062	0,041	0,142	0,030	0,123	0,227	0,511	0,060	-0,048	0,234	-0,003	0,021	0,004	0,190	0,003	0,119	0,019	0,088	0,075	-0,014	0,019	0,104	0,002	-0,400	-0,045	0,070	0,034	0,059	0,037	-0,015	0,002
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Teléfono fijo	1	0,069	0,072	0,032	0,091	0,028	0,182	0,282	0,350	0,076	-0,064	0,085	0,008	-0,019	0,008	0,182	-0,005	0,157	-0,003	0,123	0,096	-0,042	0,096	0,029	0,031	-0,195	-0,031	0,133	0,041	0,020	-0,023	-0,045	0,011
¿Crees que con los conocimientos adquiridos en la escuela puedes aplicar a un trabajo y defenderte en la vida?	1	-0,006	-0,001	-0,066	-0,035	-0,005	-0,039	-0,011	-0,018	-0,810	-0,021	0,000	-0,025	-0,038	-0,019	-0,021	-0,082	-0,002	-0,064	0,010	-0,020	0,017	-0,006	-0,010	-0,023	0,003	0,013	0,000	0,019	-0,006	0,002	-0,010	0,008
Sobre los conocimientos adquiridos en tu colegio. Siento que los conocimientos adquiridos me motivan a investigar más, desarrollar nuevas ideas y llevarlas a la práctica	1	0,034	0,003	0,164	0,051	0,154	0,051	0,023	0,013	0,709	0,028	0,026	0,011	0,155	-0,012	0,017	0,064	0,021	0,041	-0,015	0,019	0,222	0,001	0,021	0,036	-0,019	0,021	0,000	0,043	0,006	-0,008	0,000	0,007
Sobre los conocimientos adquiridos en tu colegio. Siento que lo que aprendí cumple con mis expectativas de aprendizaje.	1	0,028	0,001	0,203	0,038	0,066	0,041	0,027	0,019	0,703	0,054	0,029	0,004	0,258	-0,002	0,007	0,081	0,007	0,061	-0,036	0,015	0,158	0,004	0,016	0,063	-0,008	-0,005	0,000	0,063	0,007	-0,004	0,009	0,009
¿Recibiste una adecuada orientación vocacional en el colegio para la selección	1	0,017	0,004	0,155	-0,017	0,074	-0,004	0,016	0,024	0,400	0,015	0,005	-0,020	0,313	-0,042	-0,036	-0,023	0,025	0,027	-0,087	0,031	-0,139	-0,011	-0,020	0,014	-0,006	0,040	-0,010	0,067	-0,004	0,029	0,009	0,008

Marca la alternativa que más se parece a tu forma de pensar y actuar. Leer es perder el tiempo	1	-0.006	-0.007	0.061	0.017	0.224	-0.018	-0.021	-0.013	0.014	0.031	-0.046	0.932	0.024	0.006	0.024	0.009	-0.001	0.012	-0.007	0.015	0.044	0.013	-0.001	0.092	-0.022	0.044	-0.005	0.005	-0.006	-0.003	0.005	-0.007
Si hubieras tenido que cambiarte de colegio, ¿cómo te hubieses sentido?	1	0.013	0.003	0.104	0.028	0.052	-0.007	0.017	0.048	0.151	0.071	0.015	0.059	0.741	0.119	-0.040	0.078	0.018	0.100	0.030	0.025	0.028	0.013	0.030	-0.030	0.021	-0.031	0.004	-0.140	0.047	-0.051	-0.021	0.026
¿Qué tan orgulloso te sientes o sentías en tu colegio?	1	0.022	0.000	0.151	0.016	0.094	-0.003	-0.009	0.025	0.219	0.084	-0.018	0.042	0.719	0.046	-0.006	0.044	0.018	0.121	-0.028	0.014	0.096	0.013	0.023	-0.009	0.012	0.010	0.005	-0.116	0.027	-0.023	-0.021	0.022
¿Algunas de estas cosas te pasan o pasaban en el colegio? Me siento seguro cuando estoy en el colegio	1	0.014	-0.008	0.179	-0.008	0.061	-0.011	-0.010	0.014	0.114	0.060	-0.027	-0.019	0.464	-0.073	0.070	0.110	-0.027	0.061	-0.087	-0.022	0.281	0.003	-0.013	0.176	0.000	0.044	-0.011	0.141	-0.021	0.036	0.027	0.031
En tu colegio, ¿con qué frecuencia pasaban estas cosas? Se llevan bien entre compañeros de aula	1	0.014	0.020	0.176	0.018	0.048	0.066	0.068	-0.011	0.102	0.014	0.041	-0.063	0.378	-0.167	0.048	0.280	-0.007	0.136	-0.041	-0.009	0.188	-0.022	-0.005	0.192	-0.040	0.018	-0.011	0.255	-0.011	0.020	0.033	0.048
Sexo del estudiante	1	0.108	-0.004	0.000	0.082	0.120	0.015	-0.003	0.000	-0.007	0.001	0.006	0.004	0.026	0.953	0.011	-0.014	0.004	-0.020	0.022	-0.003	0.049	0.030	-0.007	0.031	-0.006	0.087	0.008	0.023	-0.007	0.009	0.003	0.003
¿Eres hombre o mujer?	1	0.108	-0.003	0.000	0.082	0.120	0.015	-0.003	0.000	-0.007	0.001	0.006	0.004	0.026	0.953	0.011	-0.014	0.004	-0.020	0.022	-0.003	0.049	0.030	-0.007	0.031	-0.005	0.087	0.008	0.023	-0.007	0.009	0.003	0.003
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Cocina con horno	1	0.004	0.076	0.022	0.225	0.001	0.039	0.054	0.093	0.016	0.016	0.075	0.029	0.010	0.022	0.528	-0.001	0.035	0.016	0.067	0.013	0.022	0.094	0.127	-0.027	0.052	0.001	0.016	0.038	0.074	-0.017	-0.054	0.030
De los siguientes bienes, ¿cuántos hay en funcionamiento en tu hogar? Refrigeradora	1	0.009	0.145	0.022	0.187	0.016	-0.017	-0.004	0.108	-0.002	0.004	0.085	0.002	0.012	0.001	0.508	-0.006	0.147	0.037	-0.007	0.026	0.010	0.127	0.113	-0.011	0.060	-0.006	0.007	0.006	0.032	0.163	-0.120	0.006
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Lavadora de ropa	1	-0.012	0.138	0.039	0.164	0.039	0.182	0.044	0.144	0.039	0.057	0.125	0.019	0.007	0.026	0.489	0.002	0.081	0.027	0.132	0.037	0.001	-0.021	0.137	0.016	-0.156	-0.004	0.044	-0.105	0.061	0.043	-0.012	0.005
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Televisión por cable o satelital	1	0.008	0.063	0.040	0.237	0.043	0.107	0.012	0.021	-0.019	0.057	0.046	-0.001	-0.026	-0.040	0.408	-0.010	0.081	-0.002	0.133	0.045	0.006	0.009	0.031	0.064	-0.214	0.027	0.136	-0.247	-0.016	0.068	0.021	0.025
Tipo de área de la institución educativa	1	-0.038	0.074	-0.015	-0.027	-0.019	0.202	0.231	0.018	-0.035	0.034	-0.044	0.028	-0.003	0.005	0.358	0.008	-0.051	0.083	-0.104	0.019	-0.018	-0.009	0.061	-0.056	-0.072	-0.075	-0.189	0.224	-0.046	-0.050	0.199	0.011
¿En tu aula los estudiantes son buenos amigos?	1	0.011	0.004	-0.049	0.006	-0.017	0.007	0.006	-0.004	-0.099	-0.035	0.027	-0.012	-0.108	0.005	0.009	-0.976	0.003	-0.036	0.029	-0.009	-0.009	-0.005	0.001	-0.060	0.000	-0.008	0.006	-0.001	0.001	0.009	-0.004	0.007
En tu colegio, ¿con qué frecuencia pasaban estas cosas? Tengo buenos amigos en el aula de clase	1	-0.011	-0.004	0.057	-0.007	0.020	-0.006	-0.004	0.004	0.102	0.035	-0.026	0.012	0.120	-0.007	-0.008	0.974	-0.003	0.037	-0.031	0.008	0.017	0.004	-0.002	0.065	0.000	0.008	-0.006	0.008	-0.002	-0.009	0.005	0.006
¿Cuántos cuartos de baño con ducha tiene tu hogar?	1	0.031	0.027	0.024	0.228	0.001	0.150	0.073	0.087	0.010	0.024	0.053	0.019	0.022	0.009	0.128	0.002	0.834	0.011	0.106	0.060	0.026	0.062	0.080	0.010	-0.040	-0.023	0.027	-0.019	0.044	0.005	0.013	0.008
De los siguientes bienes, ¿cuántos hay en	1	-0.048	-0.040	-0.024	-0.264	-0.016	-0.171	-0.091	-0.100	-0.021	-0.005	-0.090	0.011	-0.022	0.001	-0.131	0.006	-0.816	-0.031	-0.120	-0.066	-0.014	-0.054	-0.090	0.009	0.050	0.032	-0.038	-0.010	-0.050	-0.008	0.011	0.004

¿Cuál es la principal razón por la que desearías estudiar tu carrera?	1	-0,058	-0,043	-0,001	-0,041	0,059	-0,125	-0,096	-0,010	-0,012	0,065	-0,273	0,047	0,004	0,021	0,074	0,018	-0,083	0,018	-0,014	0,020	0,025	0,026	-0,054	0,068	0,042	0,307	0,009	-0,196	-0,056	-0,025	0,080	0,002
¿Cuántas personas en tu hogar están afiliados o cubiertos por un seguro de salud privado (con o sin hospitalización), seguro internacional o seguro de vida?	1	0,004	0,007	0,013	0,220	0,006	0,060	0,029	-0,017	-0,022	0,038	-0,034	0,019	0,016	0,001	0,012	-0,002	0,022	0,030	0,003	0,032	0,036	0,040	0,061	-0,001	-0,054	-0,002	0,763	0,001	-0,017	0,005	0,039	0,004
¿Cuántas personas en tu hogar están afiliados o cubiertos por el seguro del IESS (general, voluntario o campesino), el seguro del ISSFA o del ISSPOL?	1	0,030	0,048	0,020	0,041	0,014	0,152	-0,019	0,123	0,014	-0,044	0,129	-0,027	-0,015	0,015	0,071	-0,011	0,066	0,005	0,044	0,101	-0,023	0,057	0,036	-0,006	-0,016	-0,015	0,674	0,072	0,188	0,038	-0,049	0,024
¿Cuánto tiempo tardas o tardabas en llegar al colegio?	1	-0,010	-0,013	-0,034	-0,046	-0,010	-0,059	-0,111	0,053	0,058	0,052	0,010	0,036	-0,084	0,068	-0,033	-0,008	-0,002	0,074	-0,001	0,031	-0,030	-0,041	-0,017	-0,054	0,058	-0,065	0,057	0,681	0,026	0,005	0,000	0,039
En tu colegio, ¿con qué frecuencia pasaban estas cosas? Los estudiantes se llevan bien con la mayoría de los maestros	1	0,067	-0,011	0,326	0,049	0,076	0,096	0,084	-0,031	0,111	-0,033	0,047	-0,047	0,350	-0,082	0,049	0,134	0,010	0,038	-0,012	-0,006	0,227	-0,034	-0,015	0,093	-0,044	0,027	0,019	0,352	-0,031	0,033	0,033	0,059
¿Cuál fue tu preparación para rendir el examen Ser Bachiller?	1	-0,125	-0,015	-0,054	-0,182	-0,035	-0,139	-0,181	-0,117	-0,133	0,118	-0,088	0,041	0,004	0,069	0,025	0,020	-0,007	0,007	-0,189	-0,034	0,148	-0,006	-0,061	-0,011	0,077	0,059	-0,040	-0,239	-0,103	0,138	0,049	0,064
Señala qué hace regularmente tu madre	1	-0,025	-0,033	-0,021	-0,109	0,007	-0,213	-0,049	-0,039	-0,017	0,003	-0,051	0,004	0,002	0,018	-0,004	0,000	-0,002	0,005	-0,016	-0,051	-0,013	-0,042	-0,059	-0,012	0,054	0,014	-0,010	-0,033	-0,768	-0,082	0,011	0,022
Señala qué hace regularmente tu padre	1	-0,028	-0,020	-0,010	-0,029	-0,016	0,005	-0,068	-0,055	0,009	0,025	-0,029	0,007	-0,038	-0,005	-0,107	0,004	-0,049	-0,005	-0,068	-0,068	0,037	0,051	-0,018	0,002	0,022	0,042	-0,100	0,006	-0,612	0,086	-0,020	0,028
¿Hay algunos de estos bienes o servicios en tu hogar? Señala todos los que aplican Luz eléctrica	1	0,022	0,053	0,001	-0,029	0,009	-0,025	0,164	0,025	0,005	-0,013	0,011	-0,004	-0,027	0,014	0,095	-0,014	0,007	0,010	0,055	0,005	-0,008	-0,041	0,050	-0,023	0,029	-0,033	0,038	0,010	-0,015	0,845	-0,015	0,015
¿Con quién vives? Selecciona todas las que aplican. Solo	1	-0,011	0,004	-0,018	-0,003	0,013	-0,024	-0,003	-0,014	0,027	0,030	0,003	0,004	-0,007	0,006	-0,103	0,009	-0,001	0,001	0,020	0,003	-0,006	0,033	0,015	0,011	-0,006	0,028	0,001	-0,003	0,019	-0,010	0,919	0,005
¿Tienes alguna discapacidad?	1	0,013	0,000	-0,006	-0,020	0,009	-0,017	0,026	-0,002	-0,004	0,005	-0,006	-0,008	-0,001	0,002	0,022	0,007	0,002	-0,008	0,005	0,007	0,010	-0,025	-0,016	0,039	0,009	0,024	0,013	0,023	-0,010	-0,014	0,002	0,968

Normalización de principal de variable.

a. Método de rotación: Varimax con normalización Kaiser. La rotación no ha podido converger en 13 iteraciones. (Convergencia=,000).