



MUNICIPIO DE GUARNE

APOYO TÉCNICO A LA FORMULACIÓN DEL PLAN DE DESARROLLO DE LA
CIUDAD INTERMEDIA DE GUARNE POR TODO LO QUE NOS UNE 2024 -
2027

ANÁLISIS ESTRATÉGICO DE LAS PRUEBAS SABER 11 DEL MUNICIPIO DE
GUARNE 2011 - 2022

Ricardo Ramírez - Especialista en ciencia de datos

Mayo de 2024

Municipio de Guarne, Antioquia, Colombia. - NIT:890982055-7
Carrera 50 No.50-02, Parque Principal - Código Postal: 054050 - Teléfono: (604)551 00 25
gestiondocumental@guarne-antioquia.gov.co - www.guarne-antioquia.gov.co

@alcaldiaguarne   

1. **Resumen ejecutivo.**

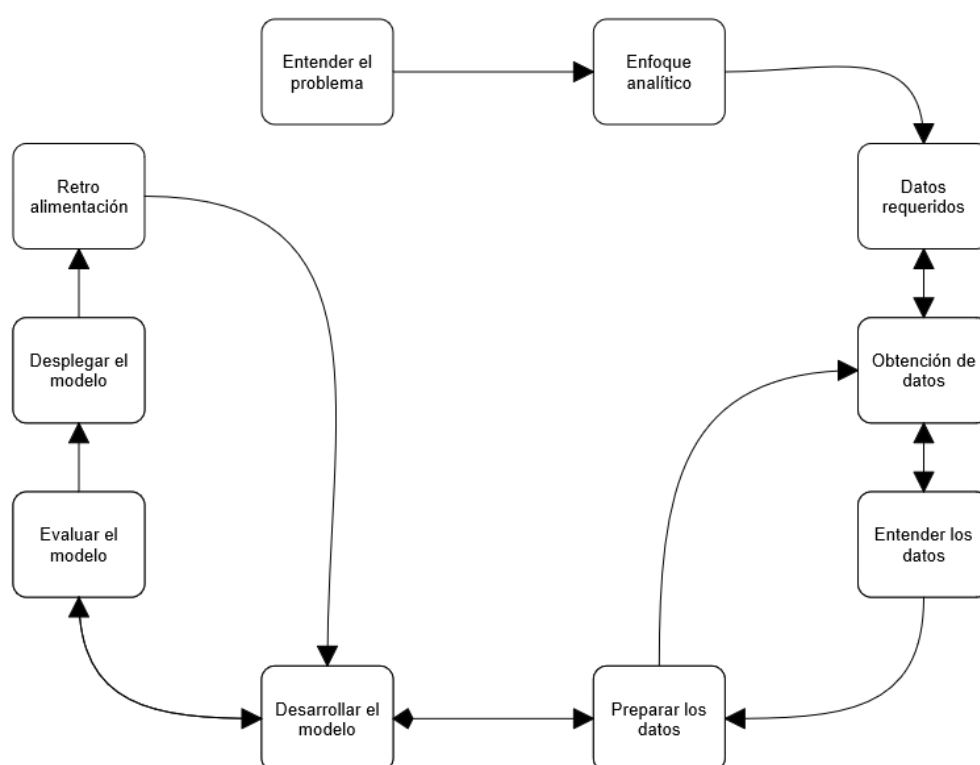
Mediante la ciencia de datos, utilizando el algoritmo de clusterización K-means, mediante una machine learning, analizando 17 variables asociadas a las pruebas SABER 11, al sistema educativo y a las condiciones socioeconómicas de los estudiantes que presentaron las pruebas, se identificaron y caracterizaron cinco segmentos o agrupaciones de estudiantes. Con base en esto, se identificaron las variables que pudieran tener mayor capacidad de influir sobre los resultados de las pruebas SABER 11.

Con base en los hallazgos, se hicieron recomendaciones estratégicas para la ejecución del Plan de Desarrollo 2024-2027 del municipio de Guarne.

2. Introducción

El enfoque analítico utilizado en el presente estudio es el de la ciencia de datos. Esto implica que las respuestas a las preguntas de la investigación se obtendrán tomando como objeto de estudio los datos. La ciencia de datos es una disciplina que tiene su propia metodología de trabajo, y aplica tecnologías como lenguajes de programación, bases de datos, e inteligencia artificial.

Ciclo de un proyecto de ciencia de datos



Los datos utilizados en este estudio cumplen con las siguientes condiciones: Precisión: Los datos son precisos y fiables, sin errores ni incoherencias. Exhaustividad: El conjunto de datos contiene toda la

información pertinente necesaria para el análisis. Coherencia: Los datos tienen un formato y una estructura coherentes en los distintos registros y variables. Pertinencia: Los datos son pertinentes para el problema o la cuestión que se aborda. Actualidad: Los datos están actualizados y reflejan la situación actual. Validez: Los datos representan con exactitud los conceptos que pretenden medir. Representatividad: El conjunto de datos contiene todo el conjunto de estudiantes que presentaron las pruebas SABER 11 en el período 2011 - 2022. Consideraciones éticas: Los datos se recopilaron, almacenaron y utilizaron de forma ética, respetando la privacidad, la confidencialidad y la normativa legal. Accesibilidad: Los datos se obtuvieron de fuentes públicas, accesibles y estaban disponibles para su análisis.

Problema. El problema abordado mediante este estudio fue:

“Cómo mejorar la calidad de la educación media en el municipio de Guarne, a partir de los resultados de las pruebas SABER 11”.

Preguntas. Las preguntas que se formularon y respondieron en esta investigación:

1. ¿Existe un patrón de segmentación en los resultados de las pruebas SABER 11 en el período de análisis?
2. Si tienen un patrón de segmentación ¿Cómo se segmentan los 5450 estudiantes de Guarne que presentaron las pruebas SABER 11 en el período de análisis?
3. ¿Cuál es la característica de cada segmento de estudiantes?

4. ¿Cuáles son las principales variables que influyen sobre los resultados de las pruebas SABER 11 en el municipio de Guarne?
5. ¿Qué objetivos y acciones debe plantearse Guarne para mejorar la calidad de la educación media, en el período 2024-2027?

3. Metodología.

Los datos utilizados en este estudio se obtuvieron de la plataforma pública [Datos Abiertos](#). Datos Abiertos es una plataforma del Ministerio de Tecnologías de la Información y las Comunicaciones, usada como repositorio de datos de las entidades públicas, y como fuente de datos del sector público de Colombia para el público en general. Para su descarga se aplicaron técnicas de ETL (extract, transform, load), que consisten en tareas de extracción de los datos desde la fuente; transformación para ponerlos aptos para su uso; y almacenamiento en un set de datos disponible para ser utilizado en los análisis, que se explicarán adelante. Este set de datos de 23 variables y 5450 entradas correspondientes a todos los estudiantes del municipio de Guarne que presentaron las pruebas SABER 11 durante los años 2011 - 2022, puede consultarse aquí: [SET DE DATOS](#) en formato csv.

Tabla 1 Variables utilizadas en el modelo analítico

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5450 entries, 0 to 5449
Data columns (total 51 columns):
#   Column                                     Non-Null Count  Dtype
---  -
0   PERIODO                                     5450 non-null   int64
1   ESTU_TIPODOCUMENTO                       5450 non-null   object
2   ESTU_CONSECUTIVO                         5450 non-null   object
3   COLE_AREA_UBICACION                     5450 non-null   object
4   COLE_BILINGUE                           4146 non-null   object
5   COLE_CALEDARIO                           5450 non-null   object
6   COLE_CARACTER                            5341 non-null   object
7   COLE_COD_DANE_ESTABLECIMIENTO           5450 non-null   int64
8   COLE_COD_DANE_SEDE                      5450 non-null   int64
9   COLE_COD_DEPTO_UBICACION                5450 non-null   int64
10  COLE_COD_MCPIO_UBICACION                 5450 non-null   int64
11  COLE_CODIGO_ICFES                        5443 non-null   float64
12  COLE_DEPTO_UBICACION                     5450 non-null   object
13  COLE_GENERO                             5450 non-null   object
14  COLE_JORNADA                             5450 non-null   object
15  COLE_MCPIO_UBICACION                     5450 non-null   object
16  COLE_NATURALEZA                          5450 non-null   object
17  COLE_NOMBRE_ESTABLECIMIENTO              5450 non-null   object
18  COLE_NOMBRE_SEDE                         5450 non-null   object
19  COLE_SEDE_PRINCIPAL                     5450 non-null   object
20  ESTU_COD_DEPTO_PRESENTACION              5450 non-null   int64
21  ESTU_COD_MCPIO_PRESENTACION              5450 non-null   int64
22  ESTU_COD_RESIDE_DEPTO                    5436 non-null   float64
23  ESTU_COD_RESIDE_MCPIO                    5436 non-null   float64
24  ESTU_DEPTO_PRESENTACION                  5450 non-null   object
25  ESTU_DEPTO_RESIDE                        5436 non-null   object
26  ESTU_ESTADOINVESTIGACION                 5450 non-null   object
27  ESTU_ESTUDIANTE                          5450 non-null   object
28  ESTU_FECHANACIMIENTO                     5450 non-null   object
29  ESTU_GENERO                              5446 non-null   object
30  ESTU_MCPIO_PRESENTACION                  5450 non-null   object
31  ESTU_MCPIO_RESIDE                        5436 non-null   object
32  ESTU_NACIONALIDAD                       5450 non-null   object
33  ESTU_PAIS_RESIDE                         5450 non-null   object
34  ESTU_PRIVADO_LIBERTAD                    5450 non-null   object
35  FAMI_CUARTOSHOGAR                        5373 non-null   object
36  FAMI_EDUCACIONMADRE                      5333 non-null   object
37  FAMI_EDUCACIONPADRE                      5325 non-null   object
38  FAMI_ESTRATOVIVIENDA                     5320 non-null   object
39  FAMI_PERSONASHOGAR                       5380 non-null   object
40  FAMI_TIENEAUTOMOVIL                      5374 non-null   object
41  FAMI_TIENECOMPUTADOR                     5379 non-null   object
42  FAMI_TIENEINTERNET                       5338 non-null   object
43  FAMI_TIENELAVADORA                       5384 non-null   object
44  DESEMP_INGLES                            5450 non-null   object
45  PUNT_INGLES                              5450 non-null   float64
46  PUNT_MATEMATICAS                         5450 non-null   float64
47  PUNT_SOCIALES_CIUDADANAS                 3522 non-null   float64
48  PUNT_C_NATURALES                         3522 non-null   float64
49  PUNT_LECTURA_CRITICA                     3522 non-null   float64
50  PUNT_GLOBAL                              3522 non-null   float64
dtypes: float64(9), int64(7), object(35)
memory usage: 2.1+ MB
(5450, 51)
```

4. Resultados.

Los datos descargados se integraron en un solo set de datos mediante técnicas de procesamiento de datos, se cargaron en un Google colab notebook de trabajo, y se convirtieron en un Pandas Dataframe, para ser analizado con Python y sus librerías. La preparación de los datos se hizo mediante los siguientes procesos:

- Descargue del set de datos al notebook y creación de un Pandas dataframe
- Revisión de datos nulos y corrección de errores.
- Revisión de tipos de datos y ajustes para análisis.

De la preparación de los datos se obtuvo un nuevo set de datos de 17 variables relevantes y 3522 unidades de observación, que corresponden a estudiantes de grado 11, que presentaron las pruebas SABER 11, entre 2011 y 2022. Se descartaron variables no relevantes o con muy pocos datos para el análisis, como: tipo de documento, colegio bilingüe, calendario del colegio, los códigos Dane del departamento, municipio y colegio, y otros que se pueden identificar en el cuadro anterior. Se descartaron los estudiantes que no tenían resultados en todas las pruebas.

Python es el lenguaje de programación de la ciencia de datos <https://www.python.org/>. Para el análisis se utilizaron diferentes librerías de Python. Las librerías son herramientas especializadas para aplicarse a tareas específicas de análisis de datos. En este caso, se utilizaron las siguientes:

- Pandas, para crear y operar estructuras de datos.
- Numpy, para crear y operar estructuras de datos.
- Seaborn, para visualización de datos.
- Matplotlib.pyplot, para visualización de datos.
- Sklearn, para aplicar algoritmos de reducción de la dimensionalidad de los datos (PCA Principal Component Analysis), y para desarrollar un modelo analítico de clusterización, mediante una Machine Learning (K-means).

Análisis de los datos. Para responder las preguntas del estudio se ejecutaron los siguientes procesos:

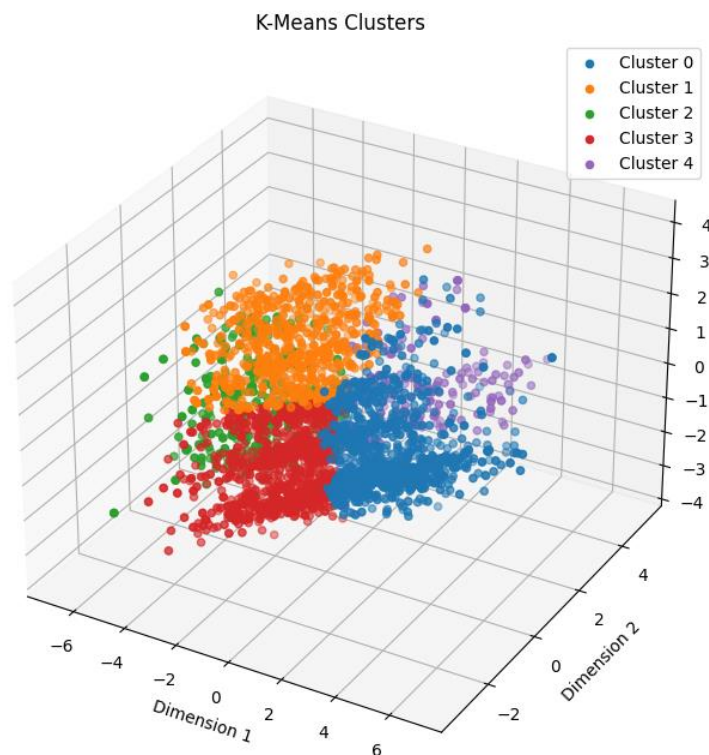
- Análisis descriptivo de todas las variables, para identificar patrones iniciales y tener una visión panorámica de los datos.
- Construcción de tablas de frecuencias, para identificar la distribución de cada variable.
- Construcción de histogramas de frecuencias de cada variable, para identificar visualmente la distribución de cada variable
- Preparación de los datos para alimentar la Machine Learning: 1. Construcción de una array de datos. 2. Estandarización de los datos, mediante Standar Scaler.
- Reducción de la dimensionalidad de los datos, de 17 dimensiones a 3 dimensiones, con la aplicación del algoritmo PCA Principal Component Analysis.
- Identificación de los clusters de estudiantes, mediante la aplicación de una Machine Learning, con el uso del algoritmo K-means. Los clusters resultantes son: 0, 1, 2, 3, 4.
- Visualización de los clusters de estudiantes.

- Caracterización de cada cluster de estudiantes, mediante el cálculo de la media de cada variable, para las variables numéricas, y de la frecuencia (%), para las variables categóricas.
- Visualización de las medias y las frecuencias de cada variable en gráficas de barras.

Todos los procesos descritos antes, se pueden revisar en este [NOTEBOOK](#)

4.1. Clusters identificados

Gráfica 1 Clusters identificados



En esta visualización, cada punto corresponde a un estudiante que presentó la prueba saber 11 en el período 2011-2022. Cada color corresponde a una agrupación homogénea de estudiantes; esto

significa que los estudiantes que pertenecen a cada segmento o grupo tienen valores similares en las 17 variables. Esta visualización ayuda a responder la pregunta ¿Existe un patrón de segmentación en los resultados de las pruebas SABER 11 en el período de análisis? Y la respuesta es: De acuerdo con las variables de análisis, los estudiantes de grado 11, del municipio de Guarne, se segmentan de acuerdo con el patrón que muestra la gráfica, y que se describen a continuación:

Se identificaron cinco clusters (segmentos o grupos) que contienen 3522 estudiantes del grado 11 del municipio de Guarne. En cada segmento sus integrantes (estudiantes) son semejantes entre sí, y diferentes a los integrantes de los otros segmentos.

Cluster 0: 1.105 estudiantes.

Cluster 1: 630 estudiantes.

Cluster 2: 532 estudiantes

Cluster 3: 983 estudiantes.

Cluster 4: 272 estudiantes.

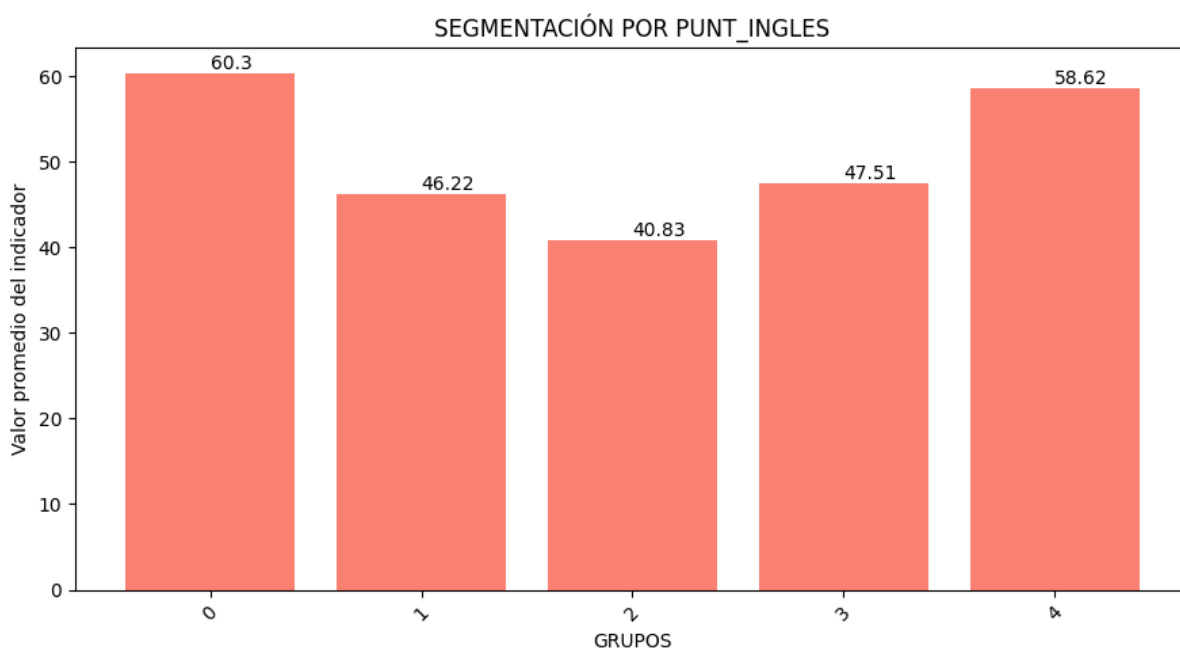
4.2. Comportamiento de las variables en cada cluster

A continuación, se identificarán los hallazgos para cada variable analizada, en cada cluster. Esta descripción, variable por variable, es

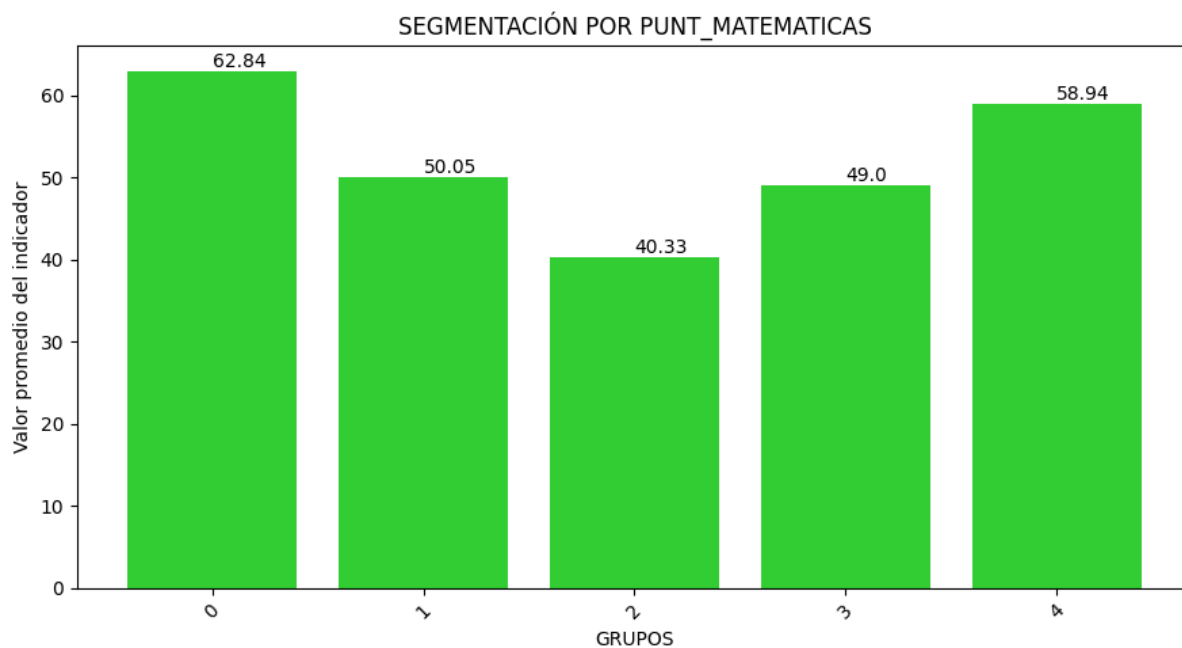
solo una referencia, que describe la realidad todavía fragmentada. El resultado complejo se muestra adelante en la caracterización de cada grupo, teniendo en cuenta la interacción de todas las variables.

4.2.1. Segmentación por resultados de las pruebas SABER 11

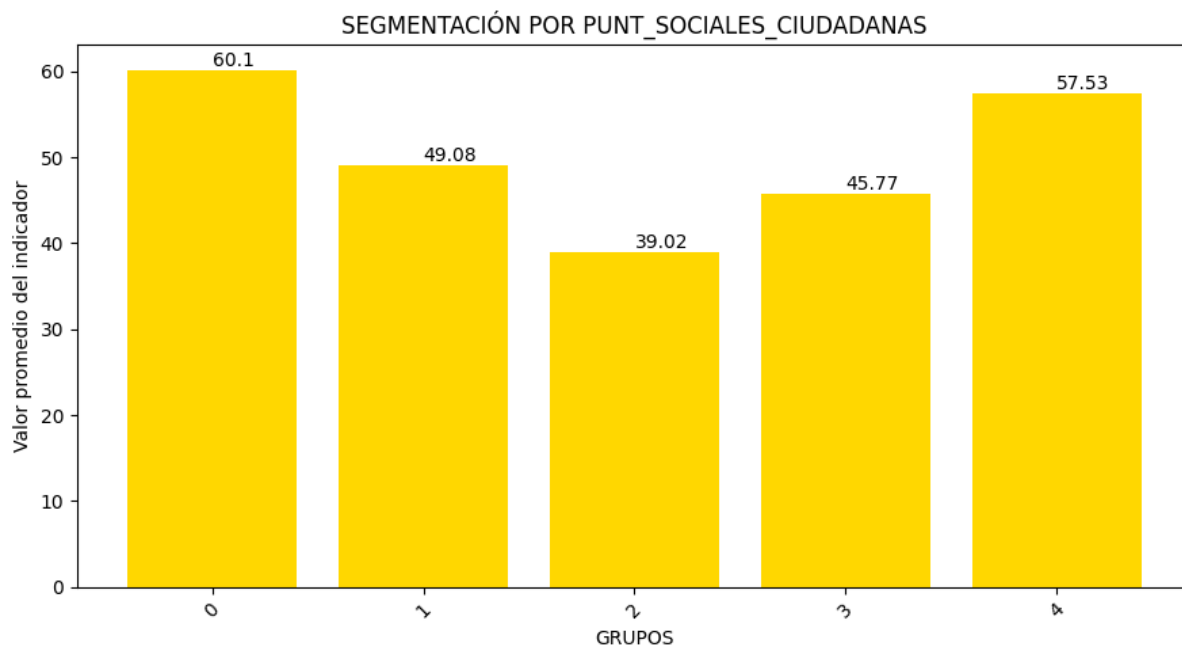
Gráfica 2 INGLÉS (ICFES, 2011-2022, puntos)



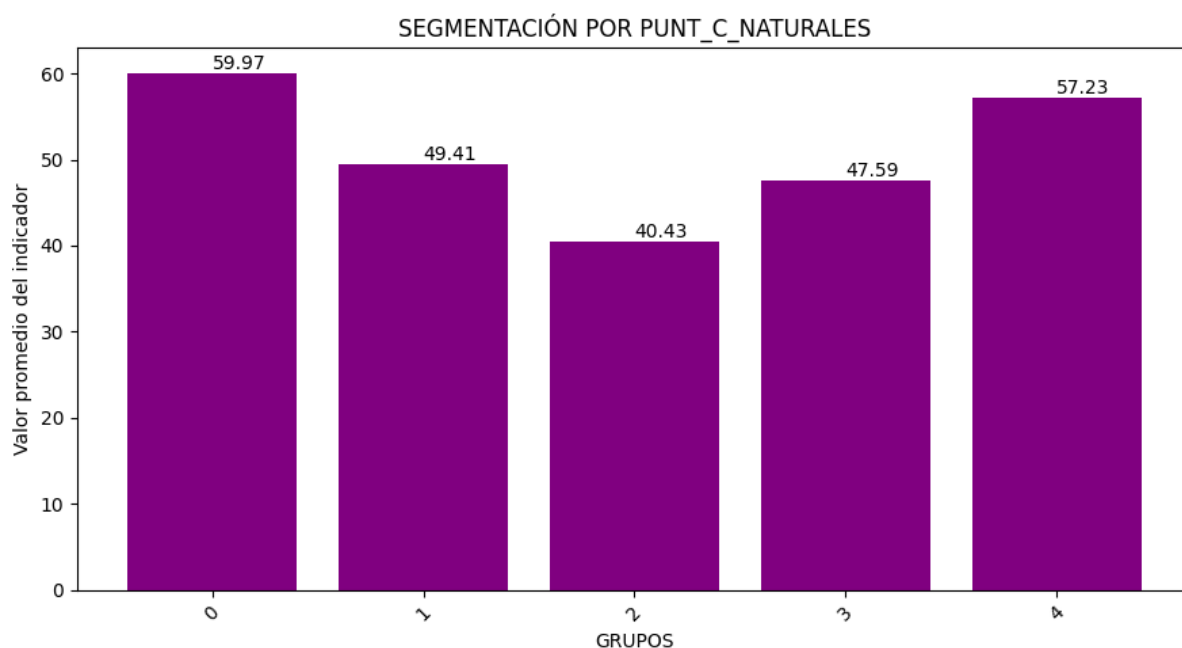
Gráfica 3 MATEMÁTICAS (ICFES, 2011-2022, puntos)



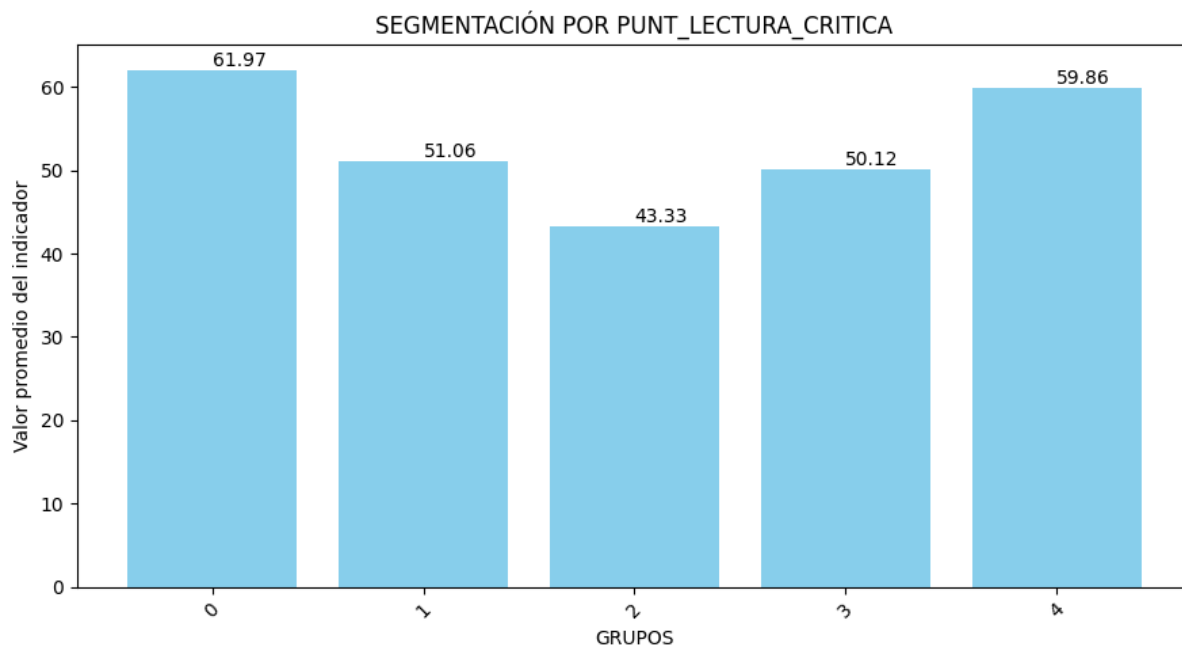
Gráfica 4 SOCIALES (ICFES, 2011-2022, puntos)



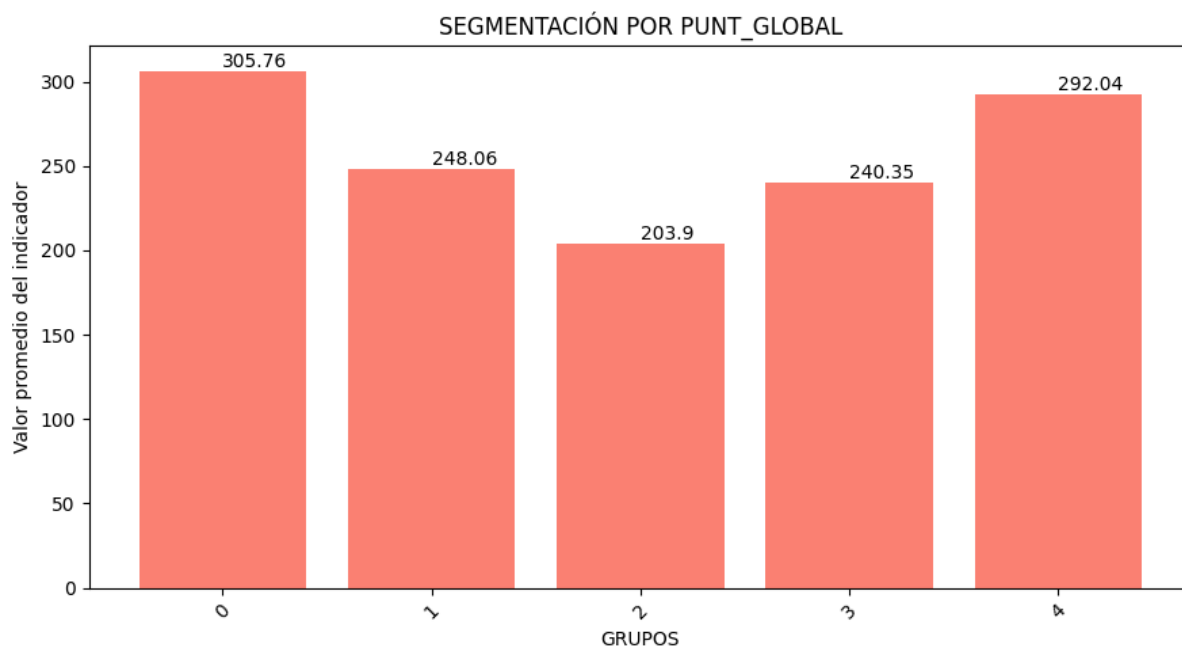
Gráfica 5 CIENCIAS NATURALES (ICFES, 2011-2022, puntos)



Gráfica 6 LECTURA CRÍTICA (ICFES, 2011-2022, puntos)



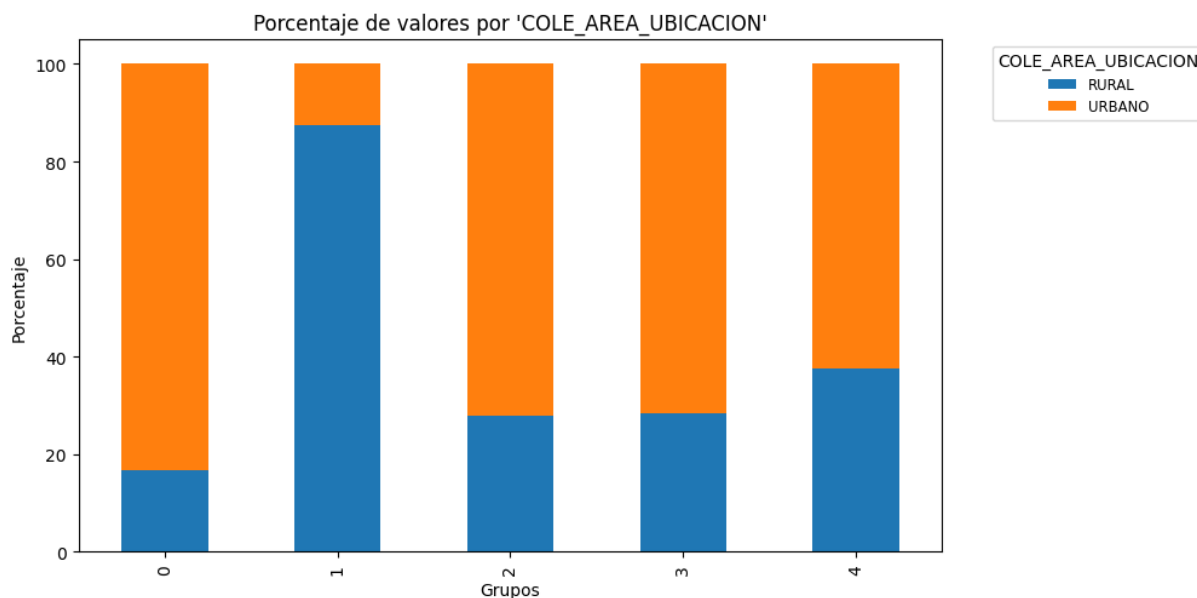
Gráfica 7 PUNTAJE GLOBAL (ICFES, 2011-2022, puntos)



Como se puede ver en las gráficas, el patrón de comportamiento de los resultados en todas las pruebas es similar: Los grupos 0 y 4 tienen los mejores resultados; los grupos 1 y 3 tienen resultados intermedios; y el grupo 2 tiene el desempeño más bajo. Teniendo en cuenta este patrón, a continuación, se examinan los resultados en las variables sociales y del sistema educativo, se caracteriza cada grupo o cluster de acuerdo con los valores resultantes en esas variables, y, finalmente, se busca identificar relaciones entre cada grupo o cluster y el patrón de resultados de las pruebas.

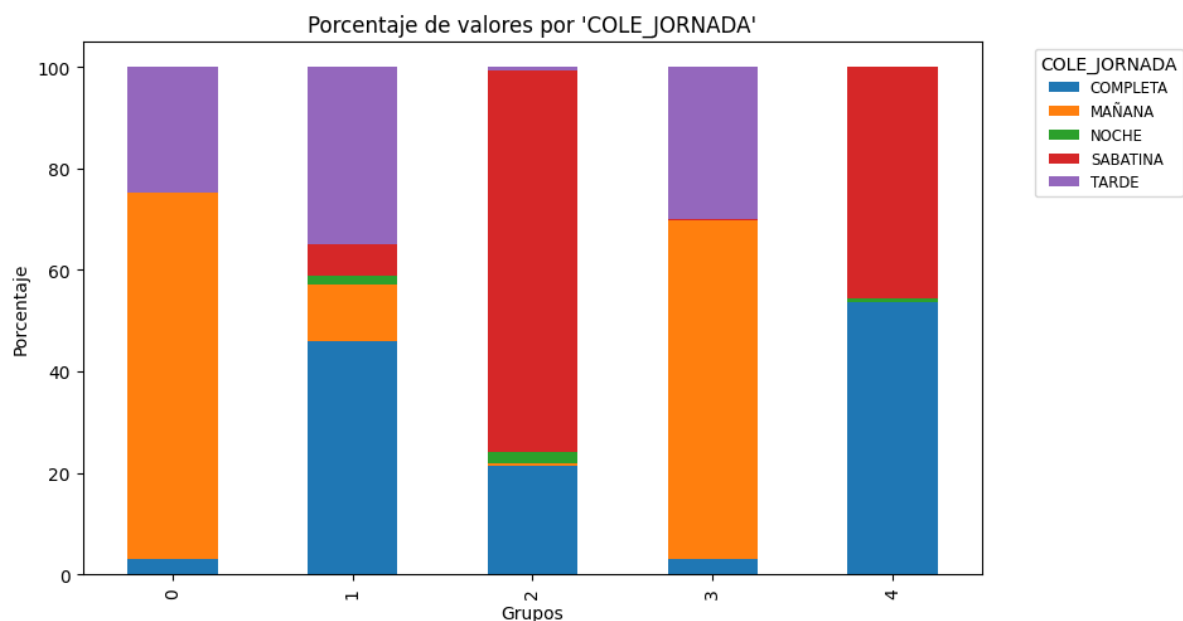
4.2.2. Variables sociales y del sistema educativo de Guarne

Gráfica 8 UBICACIÓN GEOGRÁFICA DEL COLEGIO (ICFES, 2011 - 2022, %)



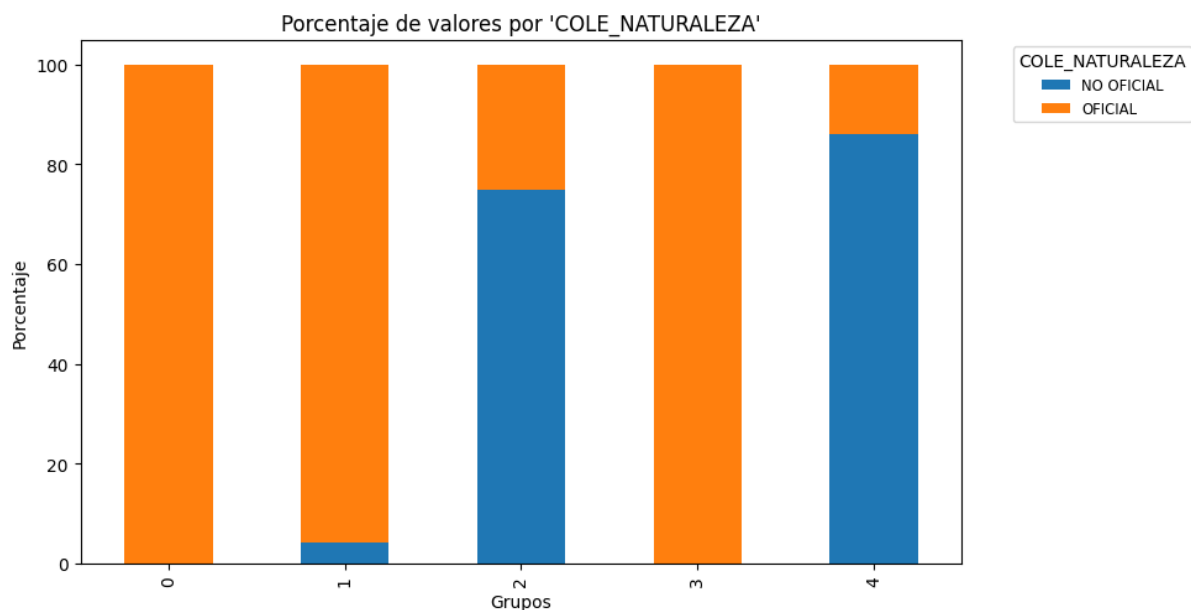
La mayoría de los estudiantes del grupo 1 pertenecen a colegios ubicados en sectores rurales del municipio de Guarne. En los demás grupos, los estudiantes pertenecen a colegios ubicados en sectores urbanos, entre el 70 y 80 %. Si se contrasta este patrón con el patrón de resultados de los estudiantes en las pruebas SABER 11 (gráficas anteriores) no parece haber una relación directa entre los resultados obtenidos por los estudiantes en las pruebas SABER 11 y la localización de los colegios.

Gráfica 9 JORNADA EDUCATIVA DE LOS COLEGIOS (ICFES, 2011 - 2022, %)



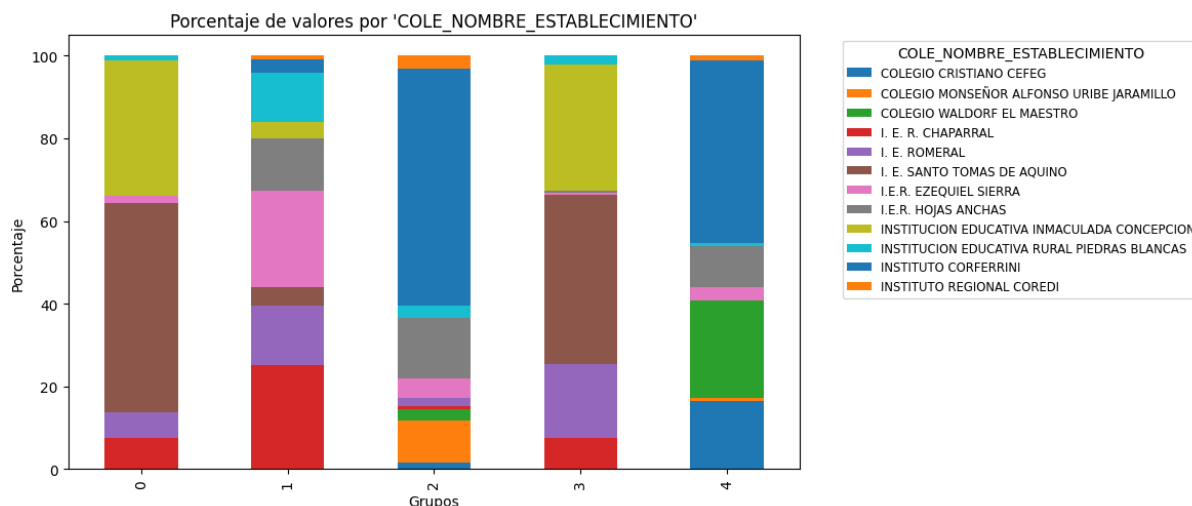
En el grupo 2 la mayoría de los estudiantes, cerca del 80 %, pertenecen a instituciones educativas de jornada sabatina, y cerca del 20 % a instituciones educativas de jornada completa. En los grupos 0 y 3 la mayoría de los estudiantes pertenecen a instituciones educativas de jornada de la mañana (75 %) y de la tarde (25 %). En los grupos 1 y 4 entre el 45 y 50 % de los estudiantes pertenecen a colegios de jornada completa. Podría haber una conexión directa, que es necesario verificar, entre la jornada sabatina de los colegios alta (grupo 2) y los bajos resultados en las pruebas saber 11.

Gráfica 10 NATURALEZA DE LOS COLEGIOS (ICFES, 2011 - 2022, %)



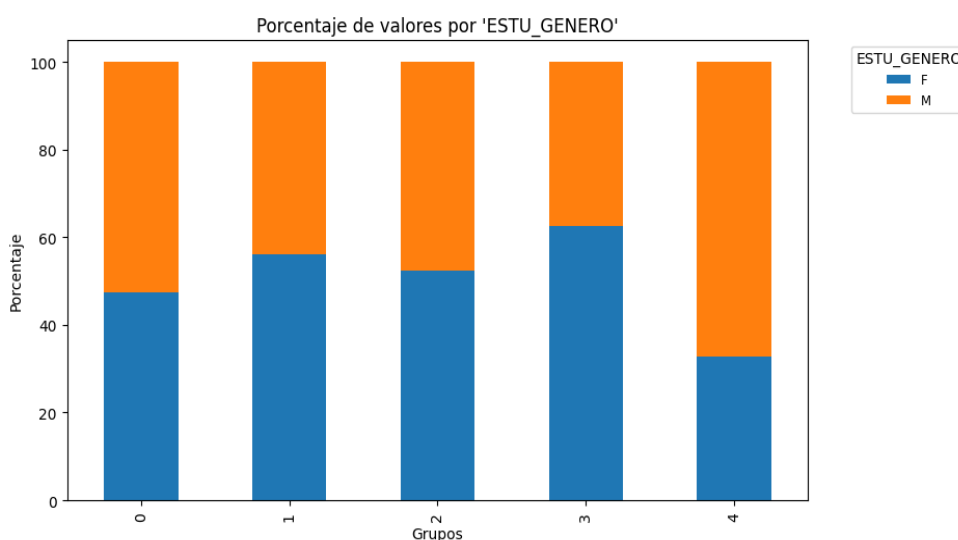
En los grupos 0, 1 y 3 la mayoría de los estudiantes pertenecen a instituciones educativas de carácter “oficial” (95 al 100 %), mientras que en los grupos 2 y 4 la mayoría de los estudiantes pertenecen a instituciones educativas de carácter “no oficial” (75 al 85 %). No parece haber una relación directa entre la naturaleza de los colegios y el puntaje obtenido por los estudiantes en las pruebas SABER 11.

Gráfica 11 ESTABLECIMIENTOS EDUCATIVOS (ICFES, 2011 - 2022, %)



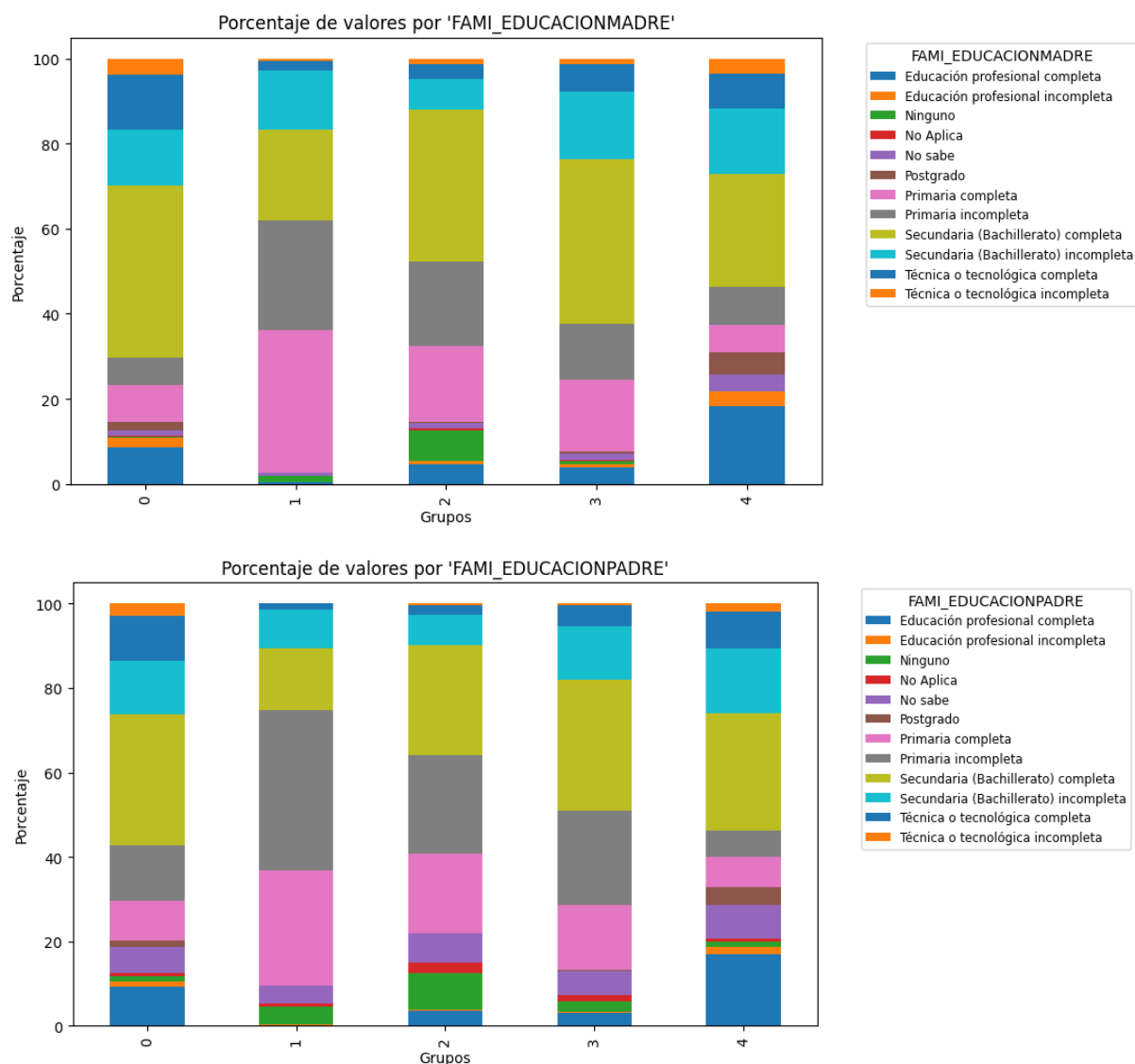
Los estudiantes que presentaron las pruebas SABER 11 están dispersos en 12 instituciones educativas del municipio de Guarne, y no se evidencia conexión directa entre la pertenencia de los estudiantes a las instituciones educativas y los puntajes obtenidos en las pruebas SABER 11. Es necesario explorar una posible relación entre el grupo 2, de los estudiantes que obtuvieron menores puntajes en las pruebas SABER 11 y su pertenencia al Instituto Corferrini.

Gráfica 12 GÉNERO DE LOS ESTUDIANTES (ICFES, 2011 - 2022, %)



En los grupos 0 y 4, la mayoría de los estudiantes son hombres (50 al 65 %) y en los grupos 1, 2 y 3 la mayoría son mujeres. No parece haber una relación directa entre el género de los estudiantes y los resultados obtenidos en las pruebas SABER 11.

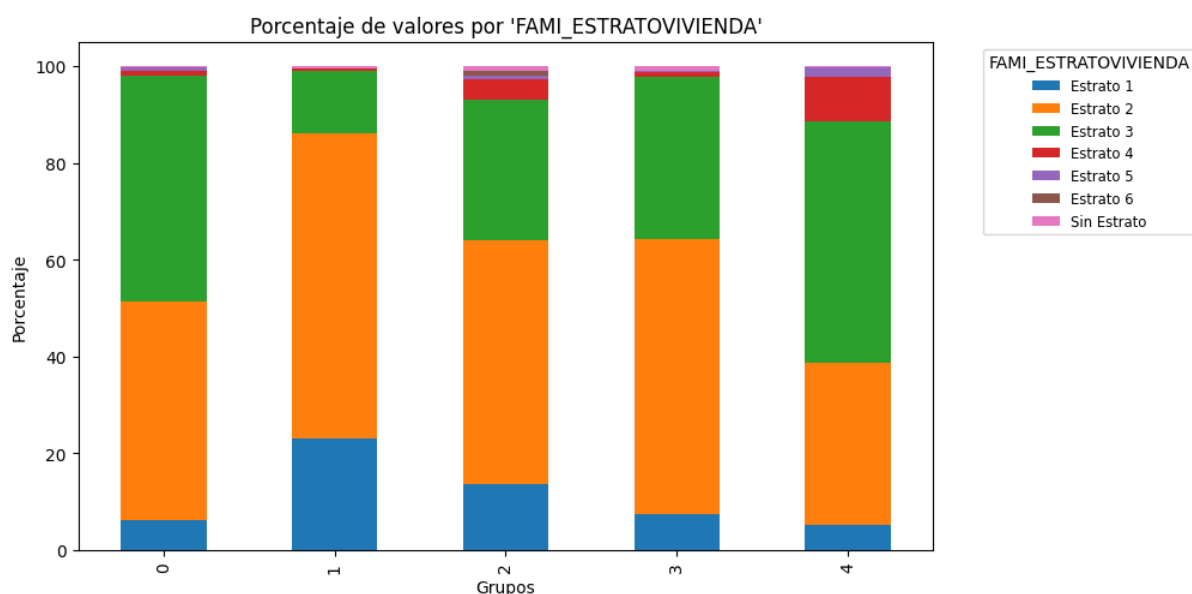
Gráfica 13 NIVEL EDUCATIVO DE LA MADRE Y EL PADRE (ICFES, 2011 - 2022, %)



Podría haber una relación directa entre los niveles educativos secundaria, educación técnica y educación profesional completa de

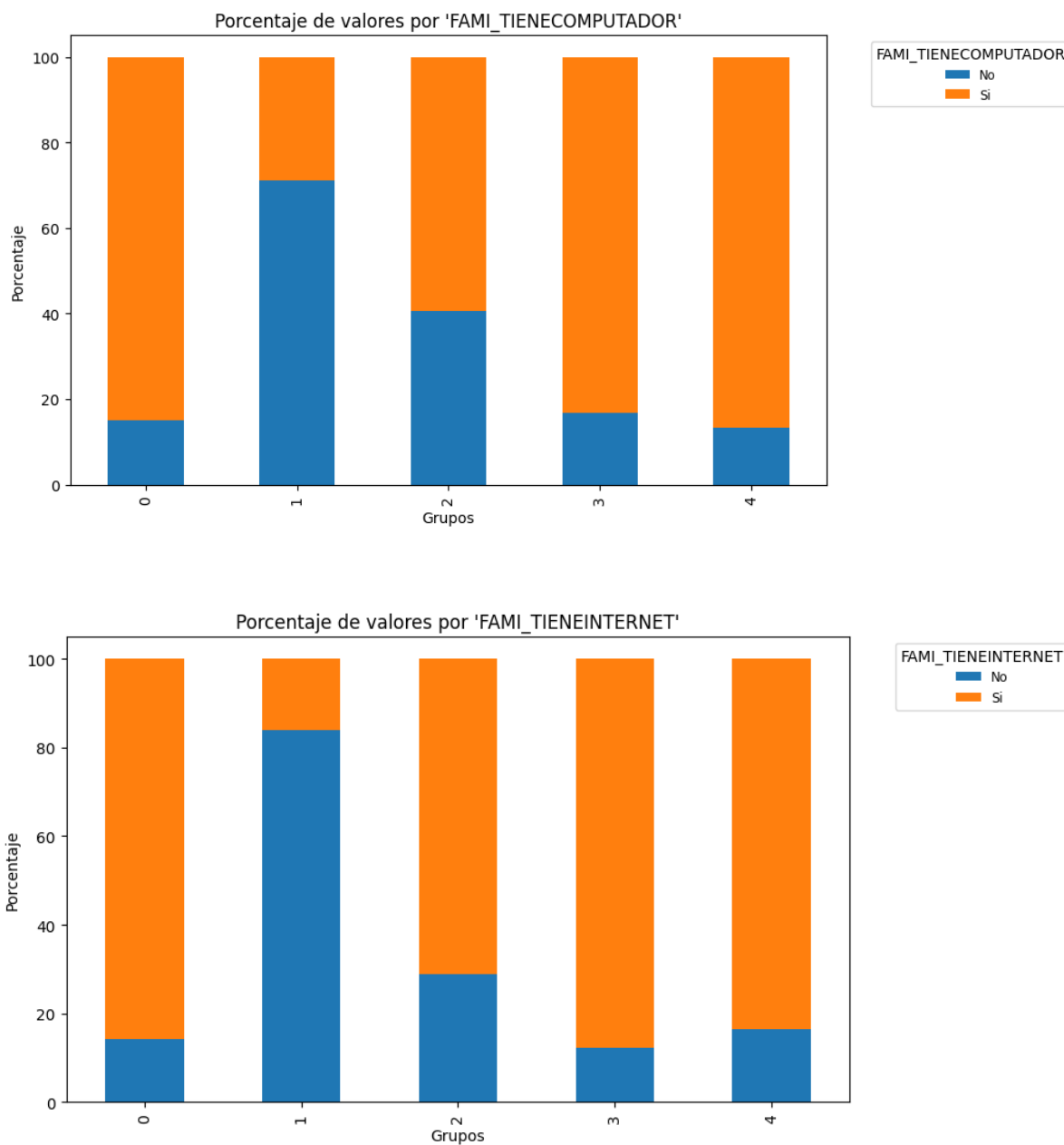
la madre y del padre, en los grupos 0 y 4, con los puntajes más elevados obtenidos por los estudiantes de esos grupos en las pruebas SABER 11.

Gráfica 14 ESTRATO DE LA VIVIENDA (ICFES, 2011 - 2022, %)



En los grupos 1, 2 y 3 la mayoría de los estudiantes viven en viviendas clasificadas en los estratos 1 y 2, mientras que los estudiantes de los grupos 0 y 4, la mayoría viven en viviendas clasificadas en los estratos 2 y 3. Podría explorarse la existencia de una relación entre la pertenencia al estrato con los puntajes obtenidos en las pruebas SABER 11, teniendo en cuenta que los grupos 0 y 4 tienen un mayor estrato y, a su vez, los mejores resultados en las pruebas.

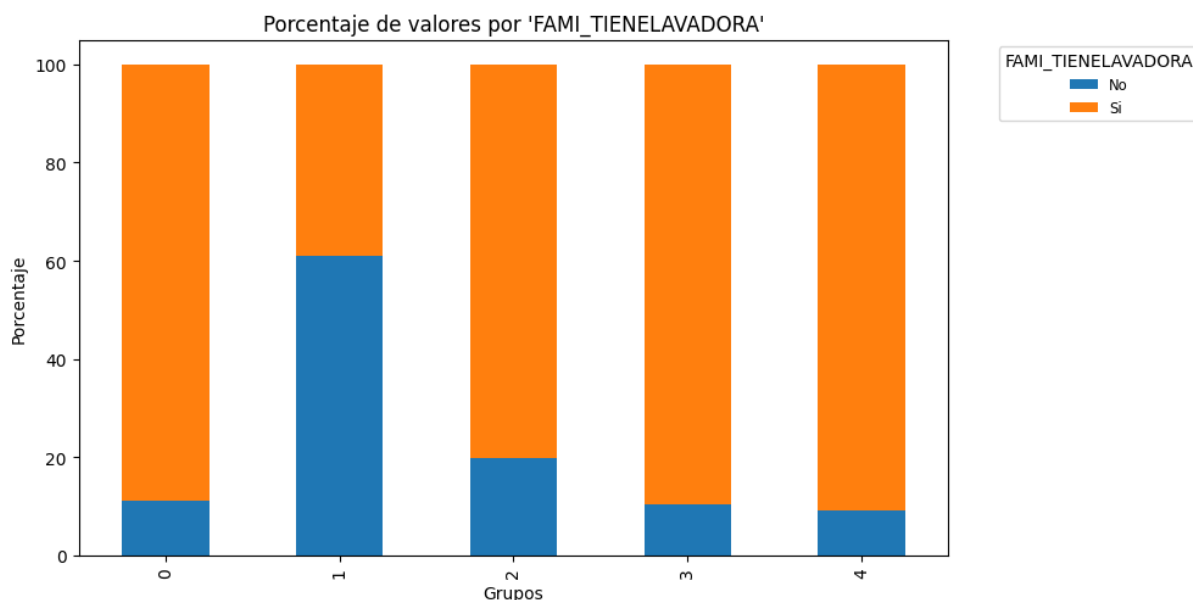
Gráfica 15 FAMILIA TIENE COMPUTADOR Y ACCESO A INTERNET EN LA CASA (ICFES, 2011 - 2022, %)



En los grupos 0, 3 y 4 más del 90% de las familias de los estudiantes tiene computador y acceso a internet en su vivienda. Podría haber una relación directa entre esta situación con los altos puntajes

obtenidos en las pruebas SABER 11 por los estudiantes de los grupos 0 y 4.

Gráfica 16 FAMILIA TIENE LAVADORA (ICFES, 2011 - 2022, %)



Este es un indicador asociado a la variable “dotación de la vivienda”, que, a su vez, está asociado a la dedicación de tiempo de la persona a cargo de las labores de cuidado de la casa y de las personas que integran la familia. En el grupo 1 es mayor el número de familias (60 %) que no tienen lavadora en la casa, seguido del grupo 2 (20 %).

4.2.3. Caracterización de los segmentos o grupos

A continuación, se caracteriza cada segmento o grupo identificado, con relación a todas las variables analizadas. La primera variable que se describe en todos los grupos es el resultado en las pruebas SABER 11 de los estudiantes que pertenecen al grupo. Esta variable se utiliza como la variable objetivo, y se busca evaluar el comportamiento de

las demás variables (predictivas) en relación con esta variable resultado u objetivo.

GRUPO 0: Los Líderes

El grupo 0 está conformado por 1105 estudiantes, y tiene las siguientes características: los estudiantes que conforman este grupo tienen **los mejores resultados** en las pruebas SABER 11, tanto en las específicas como en la global; es el grupo con mayor porcentaje de colegios en el área urbana (Más del 80 %); la mayoría de los colegios pertenecen a las jornadas mañana o tarde (98 %); todos los colegios son de naturaleza “oficial”; más del 80 % de sus estudiantes pertenecen a las instituciones educativas Santo Tomás de Aquino, Inmaculada Concepción y Romeral; la proporción de género de los estudiantes es 52 % masculino y 48 % femenino; la mayoría de madres y padres tienen bachillerato completo e incompleto y técnica, tecnológica o profesional completa o incompleta; más del 90 % de las familias viven en viviendas de los estratos 2 y 3 (98 %); la mayoría de las familias (83 %) tienen computador en su casa y acceso a internet (95 %); y la mayoría de las familias tienen lavadora en su casa (95 %).

Este es el grupo referencia, lo que significa que se debería plantear la hipótesis que si los estudiantes de Guarne tuvieran unas condiciones similares a las de este grupo, podrían tener mejores resultados en las pruebas SABER 11; y, teniendo en cuenta que las pruebas SABER 11 son un indicador asociado a la variable “calidad de la educación aprehendida por los estudiantes”, se podría extender la hipótesis a

que los estudiantes podrían tener un mejor nivel académico o de calidad de su educación si tienen condiciones sociales, económicas, y de calidad de la oferta educativa semejantes a la de este segmento.

GRUPO 2: Deben Mejorar

Este grupo 2 está conformado por 630 estudiantes, y tiene las siguientes características:

Los estudiantes que conforman este grupo tienen **los resultados más bajos** en las pruebas SABER 11, tanto en las específicas como en la global; la mayoría de colegios están en el área urbana (70 %); la mayoría de los colegios **pertenecen a las jornadas sabatina (78 %) o completa (20 %)**; la mayoría de los colegios son de naturaleza “no oficial” (78 %) ; más del 60 % de sus estudiantes pertenecen a la **institución educativa CORFERRINI, Hojas Anchas o Regional COREDI**; la proporción de género de los estudiantes es 50 % masculino y 50 % femenino; la mayoría de madres y padres tienen **bachillerato completo e incompleto y primaria completa o incompleta**; la mayoría de las familias viven en viviendas de los estratos **2 y 3**; **el 60 % de las familias tienen computador en su casa y el 70 % tiene acceso a internet** ; y la mayoría de las familias tienen lavadora en su casa (80 %).

La principal diferencia de este grupo, en la variable objetivo: puntaje en las pruebas SABER 11, con el grupo referencia, es que están en los extremos. El grupo referencia, que es el Grupo Líder, tiene el

promedio de puntaje más alto en las pruebas Saber, mientras que este grupo tiene el promedio más bajo.

Es necesario explorar si las principales diferencias de este grupo con el anterior: el grupo referencia, en las variables predictivas, está en una interacción de las siguientes variables: 1. La jornada educativa: en este grupo es sabatina, mientras que en el grupo referencia es “mañana o tarde”; 2. Las instituciones educativas: son diferentes en cada grupo; 3. El nivel educativo de los padres: en el grupo de referencia es mayor el porcentaje de padres con educación técnica, tecnológica o profesional; 4. El acceso a computadores e Internet: es más alto en el grupo referencia; 5. La tenencia de lavadora (que es un indicador de la dotación del hogar y del tiempo de los cuidadores ocupado en labores domésticas), que es mayor en el grupo de referencia, lo que podría representar un mayor tiempo de los cuidadores dedicado a la atención y cuidado de sus hijos.

GRUPO 1: Los del Campo

El grupo 1 está conformado por 532 estudiantes, y tiene las siguientes características:

Con respecto de la variable objetivo, los estudiantes que conforman este grupo tienen **resultados intermedios** en las pruebas SABER 11, tanto en las específicas como en la global.

En lo relacionado con las variables predictivas: **la mayoría de colegios están en el área rural (90 %)**; la mayoría de los colegios

pertenecen a las jornadas completa (43 %) o tarde (38 %); la mayoría de los colegios son de naturaleza “oficial” (98 %) ; **la mayoría de sus estudiantes pertenecen a la instituciones educativas rurales: Chaparral, Ezequiel Sierra, Piedras Blancas, y Hojas Anchas**; la proporción de género de los estudiantes es 42 % masculino y 58 % femenino; **la mayoría de madres y padres tienen bachillerato completo e incompleto y primaria completa o incompleta**; **la mayoría de las familias viven en viviendas de los estratos 1 y 2**; la mayoría de las familias **(70 %) no tienen computador en su casa ni acceso a internet (82 %)**; y la mayoría de las familias **no tienen lavadora en su casa (60 %)**.

Las principales diferencias de este grupo con el grupo de referencia (resaltadas en negrilla) son: Con respecto de la variable objetivo, el grupo Líder, de referencia, obtuvo los puntajes más altos en las pruebas SABER 11, mientras que este grupo obtuvo resultados intermedios. Con respecto de las demás variables, las principales diferencias entre los dos grupos son: en la localización de los colegios; en el grupo líder, la mayoría de las instituciones educativas están en el área urbana, mientras que en este grupo están en el área rural; las instituciones educativas son diferentes; en el grupo líder, la mayoría de madres y padres tienen bachillerato completo e incompleto y técnica, tecnológica o profesional completa o incompleta, mientras que en este grupo, la mayoría de madres y padres tienen bachillerato completo e incompleto y primaria completa o incompleta; en el grupo Líder, más del 90 % de las familias viven en viviendas de los estratos 2 y 3, mientras que en este grupo, la mayoría de las familias viven en viviendas de los estratos 1 y

2; en el grupo Líder, la mayoría de las familias (83 %) tienen computador en su casa y acceso a internet (95 %), mientras que en este grupo, la mayoría de las familias (70 %) no tienen computador en su casa ni acceso a internet (82 %) ; en el grupo Líder, la mayoría de las familias tienen lavadora en su casa (95 %), mientras que en este grupo la mayoría de las familias no tienen lavadora en su casa (60 %).

GRUPO 3: Los del Medio

El grupo 3 está conformado por 983 estudiantes, y tiene las siguientes características:

Los estudiantes que conforman este grupo tienen **resultados intermedios** en las pruebas SABER 11, tanto en las específicas como en la global; la mayoría de colegios están en el área urbana (70 %); la mayoría de los colegios pertenecen a las jornadas mañana (63 %) o tarde (30 %); todos los colegios son de naturaleza “oficial”; la mayoría de sus estudiantes pertenecen a la instituciones educativas Inmaculada Concepción, Santo Tomás de Aquino y Romeral; la proporción de género de los estudiantes es 40 % masculino y 60 % femenino; **la mayoría de madres y padres tienen bachillerato completo e incompleto y primaria completa o incompleta; la mayoría de las familias viven en viviendas de los estratos 2 y 3**; la mayoría de las familias (82 %) tienen computador en su casa y acceso a internet (83 %); y la mayoría de las familias tienen lavadora en su casa (90 %).

La diferencia de este grupo con el grupo Líder, en lo que respecta a la variable objetivo, es: los estudiantes del grupo Líder obtuvieron los máximos puntajes en las pruebas SABER 11, mientras que en este grupo Los del campo, obtuvieron puntajes intermedios.

Con respecto de las demás variables, las principales diferencias están en el nivel educativo de las madres y padres: en el grupo líder, la mayoría de madres y padres tienen bachillerato completo e incompleto y técnica, tecnológica o profesional completa o incompleta, mientras que en este grupo, la mayoría de madres y padres tienen bachillerato completo e incompleto y primaria completa o incompleta; por otra parte, si bien, en los dos grupos, la mayoría de las familias viven en viviendas de los estratos 2 y 3, en este grupo hay una mayor porción de familias en el estrato 2.

GRUPO 4: Los Colíderes

El grupo 4 está conformado por 272 estudiantes, y tiene las siguientes características:

Los estudiantes que conforman este grupo tienen **los segundos mejores resultados** en las pruebas SABER 11, tanto en las específicas como en la global; **el mayor porcentaje de colegios están en el área urbana (63 %)**; **la mayoría de los colegios pertenecen a las jornadas completa o sabatina (99 %)**; **la mayoría de los colegios son de naturaleza “no oficial” (90 %)**; 90 % de sus estudiantes pertenecen a las instituciones educativas **Colegio Cristiano CEFEG, Colegio Waldorf El Maestro e Instituto CORFERRINI**; la proporción

de género de los estudiantes es 70 % masculino y 30 % femenino; la mayoría de madres y padres tienen bachillerato completo e incompleto y técnica, tecnológica o profesional completa o incompleta; **más del 90 % de las familias viven en viviendas de los estratos 2, 3 y 4 (97 %)**; la mayoría de las familias (84 %) tienen computador en su casa y acceso a internet (85 %); y la mayoría de las familias tienen lavadora en su casa (90 %).

Las principales diferencias entre este grupo y el grupo Líder son: con respecto de la variable objetivo, los estudiantes del grupo Líder obtuvieron los mayores puntajes en las pruebas SABER 11, mientras que los estudiantes de este grupo obtuvieron los segundos mejores puntajes.

Con respecto de las otras variables, las principales diferencias se dan en: en el grupo Líder hay una mayor porción de instituciones educativas en el área urbana; en el grupo Líder, la mayoría de las instituciones educativas pertenecen a las jornadas mañana o tarde, mientras que en este grupo Colider, la mayoría de las instituciones educativas son de las jornadas completa o sabatina; en el grupo Líder, la mayoría de las instituciones educativas son de carácter oficial, en este grupo Colider, la mayoría son de carácter no oficial; las instituciones educativas son diferentes en los dos grupos; en el grupo Líder, la mayoría de las familias viven en viviendas de los estratos 2 y 3, mientras que en grupo Colider la mayoría de las familias viven en viviendas de los estratos 2, 3 y 4 (97 %).

5. Conclusiones y recomendaciones

Los estudiantes del municipio de Guarne que presentaron las pruebas SABER 11, en el período 2011 - 2022 no conforman un conjunto de elementos dispersos e inconexos. Existen segmentos o grupos homogéneos de estudiantes que comparten realidades semejantes. Los estudiantes en estudio se segmentan en cinco grupos homogéneos, con base en los 17 indicadores de análisis.

En los cinco grupos no es posible identificar una sola variable que tenga la potencia para determinar el resultado de la variable objetivo: los resultados en las pruebas SABER 11, que, a su vez, es un indicador de la calidad de la educación en el municipio de Guarne. Sin embargo, se tienen los siguientes hallazgos:

- Si bien la ubicación geográfica de las instituciones educativas no parece ser un factor que influya de manera directa sobre los resultados de los estudiantes en las pruebas SABER 11, sí existen situaciones asociadas con la localización geográfica que influyen sobre este resultado, como la disponibilidad de Internet, y el estrato socioeconómico de las familias.
- En los grupos de mejores resultados en las pruebas SABER 11: grupos 0 y 4, la educación de la madre y el padre es de un mayor nivel que en los demás grupos. Lo que podría indicar que la influencia de los padres, basada en su nivel académico, es importante, a su vez, en la formación académica de sus hijos, y que, en consecuencia, se podría buscar crear una

estrategia que ayude a reforzar esta influencia benéfica de los padres sobre los hijos sobre su educación.

- Otro resultado que resalta en los grupos 0, 3 y 4 es que la proporción de las familias que tienen computador y acceso a internet en la casa es la más alta de todos los grupos. Esta conexión entre mejores resultados en las pruebas SABER 11, que, a su vez es un indicador de la calidad de la educación captada por los estudiantes, y el acceso a la tecnología de computación, de interconexión y de información, aunque hoy pueda ser obvia, no es menos importante porque esto podría significar un medio poderoso, a la mano del municipio de Guarne, para incrementar la calidad de la educación.
- Una variable, que tiene relación con la dotación de la vivienda, y con el tiempo que la persona que cuida de la casa y de la familia (en la mayoría de los casos es la madre) dedica a este trabajo, es la tenencia de lavadora en la casa. El tener una mejor dotación de electrodomésticos en la casa podría significar un mayor tiempo de dedicación de los cuidadores a los hijos. Este podría ser el significado de mayores puntajes en las pruebas SABER 11 en los estudiantes de hogares que tienen una lavadora en casa, que pertenecen a los grupos 0, 3 y 4.
- No se encuentra relación entre la naturaleza de los colegios: “oficial” y “no oficial” con los resultados en las pruebas SABER 11.
- Podría haber una relación entre el estrato socioeconómico y los resultados de las pruebas SABER 11 de los estudiantes. En los grupos 0 y 4, que tienen los mayores puntajes, hay una mayor proporción de hogares de estratos 3 y 4. La pertenencia

a estratos más altos, podría tener conexión con el acceso a computadores e internet, y a equipamiento del hogar, que, a su vez, puede estar conectado con los mejores resultados en las pruebas SABER 11, y, a su vez, con una mejor calidad de la educación recibida.

CONCLUSIONES

- Los hallazgos encontrados apuntan a una conexión entre el nivel socioeconómico de las familias con los resultados obtenidos por los estudiantes en las pruebas SABER 11, que, a su vez, es un indicador de la calidad de la educación aprehendida por los estudiantes.
- El nivel socioeconómico de las familias se asocia con el nivel educativo de los padres, con el acceso a tecnologías de la computación, la comunicación y la información, y con la dotación de los hogares.
- A su vez, el nivel educativo de los padres, el acceso a tecnologías de la computación, la comunicación y la información, y la dotación de los hogares, se asocian con mejores resultados en las pruebas SABER 11.

RECOMENDACIONES

Los hallazgos encontrados y las conclusiones obtenidas en este estudio delinean un patrón importante para que el municipio de Guarne avance en decisiones hacia el mejoramiento de la calidad de la educación. Sin embargo, es necesario continuar estudiando y profundizando acerca de estas variables y conexiones.

Podría decirse que la variable estratégica a intervenir, para mejorar la calidad de la educación, es el nivel socioeconómico de las familias. Sin embargo, esa variable es el objetivo estratégico de todos los programas de desarrollo, y requiere de decisiones e inversiones de gran tamaño, y de impactos en el largo plazo.

Por otra parte, hay decisiones tácticas, que apuntan hacia el objetivo estratégico de elevar el nivel socioeconómico de las familias, que se pueden tomar en la escala de un plan de desarrollo de un municipio, en el período de cuatro años de un alcalde. De este estudio se desprenden las siguientes recomendaciones generales:

1. Diseñar y ejecutar un programa de dotación de computadores, con énfasis a estudiantes de estratos 1 y 2, de manera prioritaria.
2. Diseñar y ejecutar un programa de expansión del acceso a internet a hogares, con énfasis en zonas rurales y estratos 1, 2 y 3. El objetivo de estos dos programas es ampliar el acceso de los estudiantes a las herramientas tecnológicas de la computación y la interconexión, y, por esta vía, de la obtención de mejores resultados académicos y de preparación para el mundo laboral.
3. Diseñar y ejecutar un programa de mejoramiento del equipamiento de los hogares, de manera individual o colectiva: por ejemplo, lavadoras comunitarias, con énfasis en los estratos 1 y 2.
4. Diseñar y ejecutar un programa de atención a las/los cuidadoras/cuidadores de los hijos en la casa, que en la mayoría de los hogares resultan siendo las madres. El objetivo de estos dos programas debe ser liberar las cargas de trabajo de los/las

cuidadoras/cuidadores, para, por una parte, aliviar la intensidad y duración de su jornada de trabajo, que en muchas ocasiones acumula trabajo fuera del hogar con trabajo dentro del hogar; y, por otra parte, capacitar y permitir que los/las cuidadoras/cuidadores dediquen una parte de su jornada al apoyo a sus hijos en relación con el estudio.