

**Análisis de la deserción escolar en Colombia: factores socioeconómicos y estrategias de  
intervención para reducir el abandono educativo**

Cristian Camilo Vanegas Paredes

Asesor

Jorge Eliecer Ospino Portillo

Universidad Nacional Abierta y a Distancia UNAD  
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI  
Especialización en Ciencia de Datos y Analítica

2025

## Resumen

La deserción escolar en Colombia constituye un desafío crítico que afecta especialmente a las regiones más vulnerables, donde factores como la pobreza, la violencia intrafamiliar y la deficiente calidad educativa dificultan la permanencia de los estudiantes en el sistema escolar. Las adversas condiciones socioeconómicas, junto con la carencia de apoyo pedagógico y familiar, perpetúan ciclos de exclusión social y reducen las oportunidades de movilidad económica. (Gómez-Restrepo et al., 2016)

A pesar de los esfuerzos a través de políticas públicas y programas sociales, los índices de deserción escolar permanecen alarmantes, evidenciando la necesidad de enfoques más integrales. Este estudio busca identificar los factores determinantes del abandono escolar y proponer estrategias efectivas que promuevan una educación inclusiva y de calidad, con un enfoque particular en los contextos de mayor desigualdad. La investigación destaca la importancia de abordar esta problemática desde una perspectiva multidimensional que integre aspectos educativos, sociales y económicos.

**Palabras claves:** Deserción, Escolar, Abandono, Educativo, Factores Socioeconómicos, Estrategias de Intervención, Análisis, Cualitativo, Cuantitativo

## **Abstract**

School dropout rates in Colombia are a critical challenge that particularly affect the most vulnerable regions, where factors such as poverty, domestic violence, and poor educational quality make it difficult for students to remain in the school system. Adverse socioeconomic conditions, coupled with a lack of pedagogical and family support, perpetuate cycles of social exclusion and reduce opportunities for economic mobility. (Gómez-Restrepo et al., 2016)

Despite efforts through public policies and social programs, school dropout rates remain alarming, highlighting the need for more comprehensive approaches. This study seeks to identify the determining factors of school dropout and propose effective strategies that promote inclusive, quality education, with a particular focus on contexts of greatest inequality. The research highlights the importance of addressing this problem from a multidimensional perspective that integrates educational, social, and economic aspects.

**Keywords:** School Dropout, Educational Dropout, Socioeconomic Factors, Intervention Strategies, Analysis, Qualitative, Quantitative

## Tabla de Contenido

Introducción .....	6
Justificación .....	8
Objetivos.....	9
Objetivo General .....	9
Objetivos Específicos.....	9
Ajuste y Ampliación del Marco de Datos.....	10
Definición de los Datos a Usar.....	10
Identificación de las Variables a Extraer .....	11
Software Empleado .....	11
Comunicación y Seguimiento del Desarrollo .....	12
Limpieza y Exploración de Datos (EDA).....	15
Análisis Descriptivo y Gráficas .....	18
Tendencia General Decreciente entre 2011 y 2014 .....	18
Periodo de Estabilización con Ligeros Repuntes (2014–2019) .....	19
Caída Abrupta en 2020 (Pandemia de COVID-19) .....	19
Incremento muy Pronunciado Post-pandemia (2021–2022).....	20
Análisis Data ICFES .....	20
Modelo de Predicción de Riesgo de Deserción .....	26
Importancia de Variables .....	26
Hacia el “Riesgo de un Estudiante”- Modelo de datos .....	30
Conclusiones .....	32
Referencias Bibliográficas .....	34

## Lista de Figuras

<b>Figura 1</b> <i>Acceso a la Base de Datos ICFES</i> .....	10
<b>Figura 2</b> <i>Primera Evidencia Contacto con Asesor del Proyecto de Grado</i> .....	12
<b>Figura 3</b> <i>Segunda Evidencia Contacto con Asesor del Proyecto de Grado</i> .....	12
<b>Figura 4</b> <i>Tercera Evidencia Contacto con Asesor del Proyecto de Grado</i> .....	13
<b>Figura 5</b> <i>Cuarta Evidencia Contacto con Asesor del Proyecto de Grado</i> .....	13
<b>Figura 6</b> <i>Quinta Evidencia Contacto con Asesor del Proyecto de Grado</i> .....	14
<b>Figura 7</b> <i>Sexta Evidencia Contacto con Asesor del Proyecto de Grado</i> .....	14
<b>Figura 8</b> <i>Segmentación y Filtrado de las Datas</i> .....	15
<b>Figura 9</b> <i>Cargue de Datos</i> .....	16
<b>Figura 10</b> <i>Limpieza y Selección de Variable - MEN</i> .....	16
<b>Figura 11</b> <i>Limpieza y Agrupación ICFES</i> .....	17
<b>Figura 12</b> <i>Limpieza y Agrupación DANE</i> .....	17
<b>Figura 13</b> <i>Tasa de Deserción por Año</i> .....	18
<b>Figura 14</b> <i>Promedio de Aprobación del ICFES por Año</i> .....	20
<b>Figura 15</b> <i>Relación entre Aprobación ICFES y Deserción Escolar</i> .....	22
<b>Figura 16</b> <i>Importancia de las Variables</i> .....	26
<b>Figura 17</b> <i>Predicción Riesgo de Deserción</i> .....	30

## Introducción

La propuesta titulada “*Análisis de la Deserción Escolar en Colombia: Factores Socioeconómicos y Estrategias de Intervención para Reducir el Abandono Educativo*” plantea un abordaje integral a una de las problemáticas más críticas del país, el abandono escolar, que limita el acceso a la educación, perpetúa ciclos de pobreza y restringe las oportunidades de movilidad social, especialmente en las regiones más vulnerables. A pesar de los esfuerzos gubernamentales reflejados en programas como el de Alimentación Escolar o los modelos flexibles de formación, las cifras de deserción se mantienen en niveles preocupantes, lo que evidencia la necesidad de diseñar estrategias más efectivas y concretas.

Se reconoce que la deserción escolar es un fenómeno multidimensional, condicionado por factores socioeconómicos como la pobreza, la baja escolaridad de los padres, la violencia intrafamiliar y las limitaciones estructurales de las instituciones educativas, además de variables asociadas a la calidad y pertinencia de los procesos pedagógicos. Desde el enfoque de capacidades propuesto por Amartya Sen, se subraya la importancia de la educación como motor de equidad y desarrollo humano, señalando que su ausencia amplía las desigualdades y reduce las posibilidades de superación de la pobreza. En este sentido, la propuesta busca no solo comprender las causas del problema, sino también contribuir con soluciones prácticas y sostenibles que fortalezcan la permanencia escolar.

El proyecto se orienta a proponer estrategias de intervención que respondan a las necesidades reales de los estudiantes y comunidades más afectadas. Para ello, plantea como primera fase la identificación de los factores socioeconómicos determinantes del abandono escolar, seguido de una evaluación crítica de las políticas y programas actualmente vigentes, con el fin de reconocer tanto sus alcances como sus limitaciones. Finalmente, se propone el diseño de

estrategias basadas en evidencia que integren aspectos educativos, sociales y económicos, garantizando la pertinencia y la sostenibilidad de las intervenciones.

Metodológicamente, la investigación adopta un enfoque mixto que combina técnicas cualitativas y cuantitativas. Se prevé la revisión exhaustiva de literatura y documentos oficiales, el análisis de datos secundarios provenientes de fuentes como el Ministerio de Educación Nacional y el DANE, la realización de estudios de caso en instituciones educativas ubicadas en zonas vulnerables y la aplicación de encuestas y entrevistas a estudiantes, padres y docentes. Entre las herramientas analíticas destacan los métodos descriptivos, los modelos de regresión y el análisis de clúster, que permitirán identificar patrones y grupos de riesgo en torno al fenómeno de la deserción escolar.

El proyecto, con una duración estimada de nueve meses, se estructurará en fases que incluyen la revisión bibliográfica, el diseño del marco conceptual, la recolección y análisis de datos, la validación de los resultados mediante la consulta a expertos y la redacción y difusión de los hallazgos. Para su ejecución se han considerado recursos humanos especializados, herramientas tecnológicas, materiales de oficina.

Se espera como resultado la elaboración de un informe detallado que explique los factores asociados a la deserción, la formulación de al menos tres estrategias de intervención validadas por especialistas en educación y la realización de talleres dirigidos a docentes, padres y comunidades, orientados a sensibilizar sobre la importancia de la permanencia escolar. De esta manera, el proyecto busca generar conocimiento útil para instituciones educativas, formuladores de políticas públicas y comunidades en riesgo, contribuyendo a la construcción de un sistema educativo más inclusivo, equitativo y sostenible en Colombia.

## Justificación

La deserción escolar en Colombia es un fenómeno complejo que afecta tanto el desarrollo social como el económico del país, especialmente en las regiones más vulnerables. Este problema, intensificado por factores socioeconómicos como la pobreza, el bajo nivel educativo familiar, la falta de motivación y las limitaciones estructurales de las instituciones educativas, requiere un abordaje integral. Las investigaciones previas coinciden en la importancia de la articulación entre actores educativos, políticas públicas efectivas y la participación comunitaria.

La educación es un derecho fundamental y un motor esencial para la movilidad social y la equidad, como lo señala el enfoque de capacidades de Amartya Sen, que subraya el impacto de la educación en la calidad de vida y la superación de la pobreza. Sin embargo, la ausencia de estrategias efectivas en ciertas regiones perpetúa las desigualdades y limita las oportunidades de cientos de niños y jóvenes, como lo evidencian los altos índices de deserción en Colombia.

(Manzano López & Ramírez Zambrano, 2023)

Estudios recientes resaltan que las políticas públicas, como las establecidas en el Plan Nacional de Desarrollo 2018-2022, han contribuido a la reducción de la deserción escolar mediante programas de alimentación, transporte escolar y fortalecimiento de trayectorias educativas. Sin embargo, estos esfuerzos no han sido suficientes para abordar las disparidades regionales ni para consolidar sistemas educativos inclusivos y sostenibles. Además, la necesidad de actualizar indicadores y herramientas de análisis, como las propuestas del Observatorio de Trayectorias Educativas, es crucial para mejorar la toma de decisiones. (Ministerio de Educación Nacional, 2022)

## **Objetivos**

### **Objetivo General**

Diseñar y plantear estrategias de intervención efectivas, basadas en evidencia, para reducir el abandono educativo en Colombia, con especial atención a las zonas más vulnerables del país.

### **Objetivos Específicos**

Analizar los principales factores socioeconómicos, familiares y educativos que inciden en la deserción escolar en las regiones con mayores índices de vulnerabilidad.

Evaluar la efectividad y el alcance de las políticas públicas y programas vigentes orientados a la reducción de la deserción escolar en Colombia.

Formular propuestas de intervención sustentadas en los hallazgos del análisis de datos, enfocadas en mejorar las condiciones socioeconómicas y fortalecer el acceso y permanencia en el sistema educativo.

## Ajuste y Ampliación del Marco de Datos

### Definición de los Datos a Usar

Se definió un rango de análisis de hasta cinco años, comprendido entre 2018 y 2023, incorporando el periodo de la pandemia con el fin de comparar el comportamiento antes y después de este evento y así identificar los factores que inciden en la deserción escolar.

Como fuentes de información seleccionadas se consideran las siguientes:

- 1) MEN (Ministerio de Educación Nacional)

[https://www.datos.gov.co/browse?Dataset-Information\\_Agency=Ministerio+de+Educaci%C3%B3n+Nacional&am...](https://www.datos.gov.co/browse?Dataset-Information_Agency=Ministerio+de+Educaci%C3%B3n+Nacional&am...)

- 2) ICFES

### Figura 1

*Acceso a la Base de Datos ICFES*



*Nota.* Tomada de

[https://icfesgovco.sharepoint.com/sites/DataIcfes2.0/SitePages/Home.aspx?e=4%3ahMDo3v&w...  
eb=1&at=9](https://icfesgovco.sharepoint.com/sites/DataIcfes2.0/SitePages/Home.aspx?e=4%3ahMDo3v&w...)

3) DANE (GEIH/ECV)

<https://microdatos.dane.gov.co/index.php/catalog/central/about>

4) Datos locales (Bogotá o ciudades específicas)

<https://datosabiertos.bogota.gov.co/dataset>

### **Identificación de las Variables a Extraer**

Se generó un dataset (Excel o CSV) consolidando todas las fuentes con las siguientes columnas estándar:

- Año
- Departamento / Municipio
- Institución (si aplica)
- Nivel (básica, media, superior)
- Matrícula inicial
- Matrícula final
- Tasa de deserción (%)
- Género
- Zona (urbano/rural)
- Variables socioeconómicas (de ICFES o DANE: nivel SISBEN, ingresos,

educación padres, etc.)

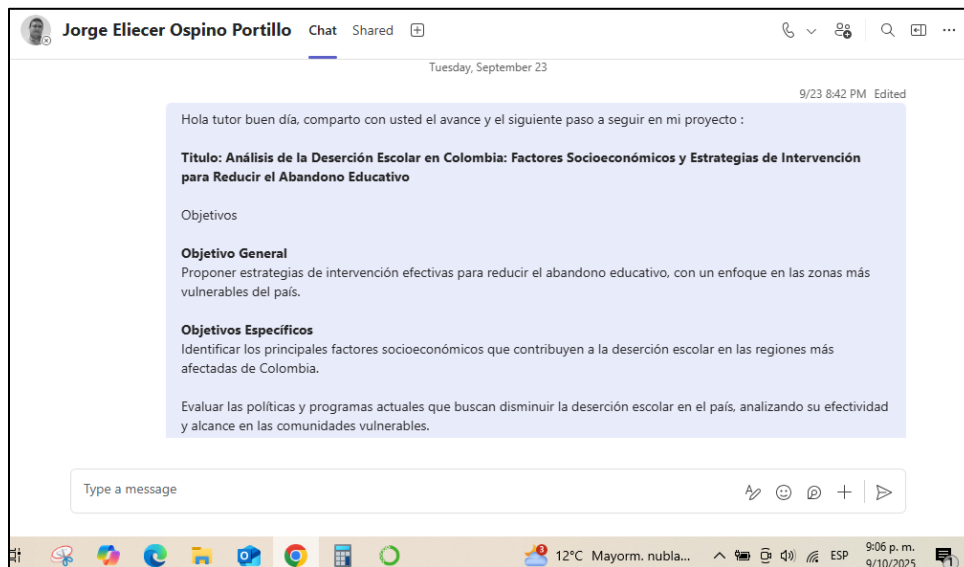
### **Software Empleado**

- Excel (Power Query): para unificación básica de la información.
- Python (pandas) : Para automatizar, realizar limpieza y unión de datasets.

## Comunicación y Seguimiento del Desarrollo

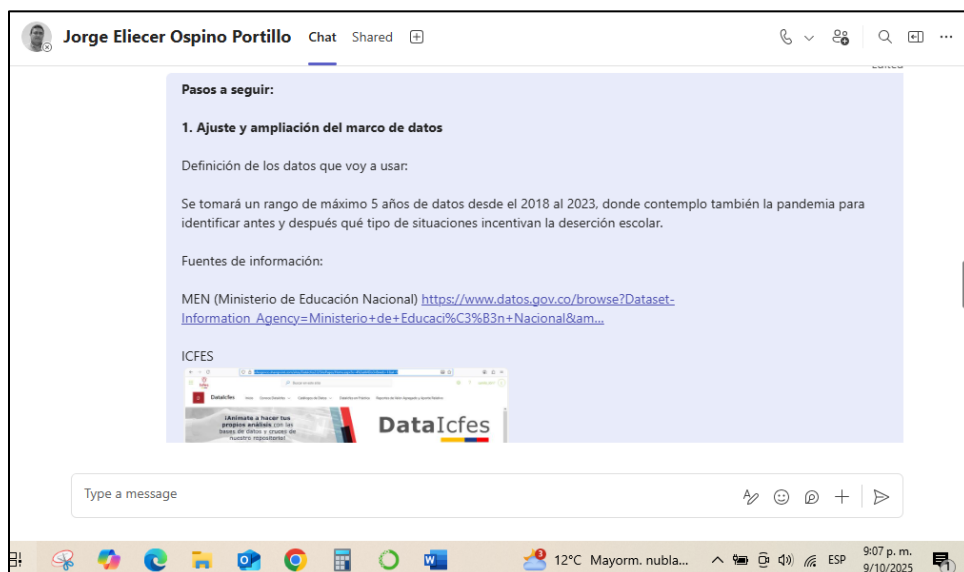
**Figura 2**

*Primera Evidencia Contacto con Asesor del Proyecto de Grado*



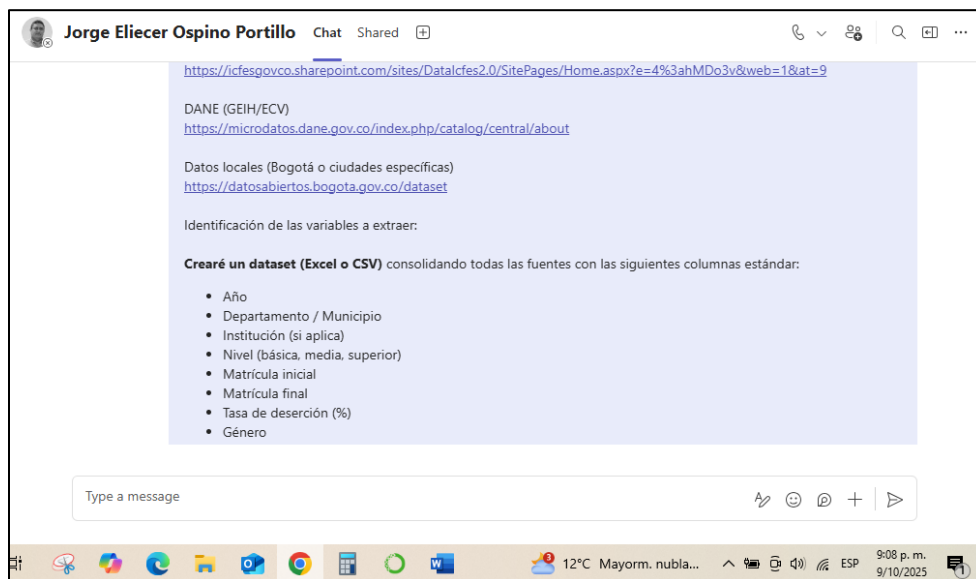
**Figura 3**

*Segunda Evidencia Contacto con Asesor del Proyecto de Grado*



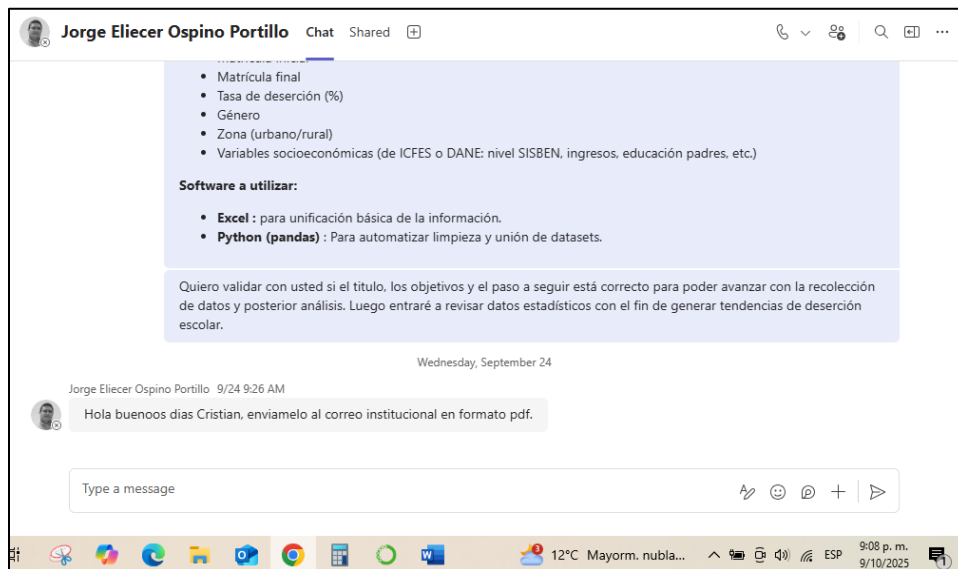
## Figura 4

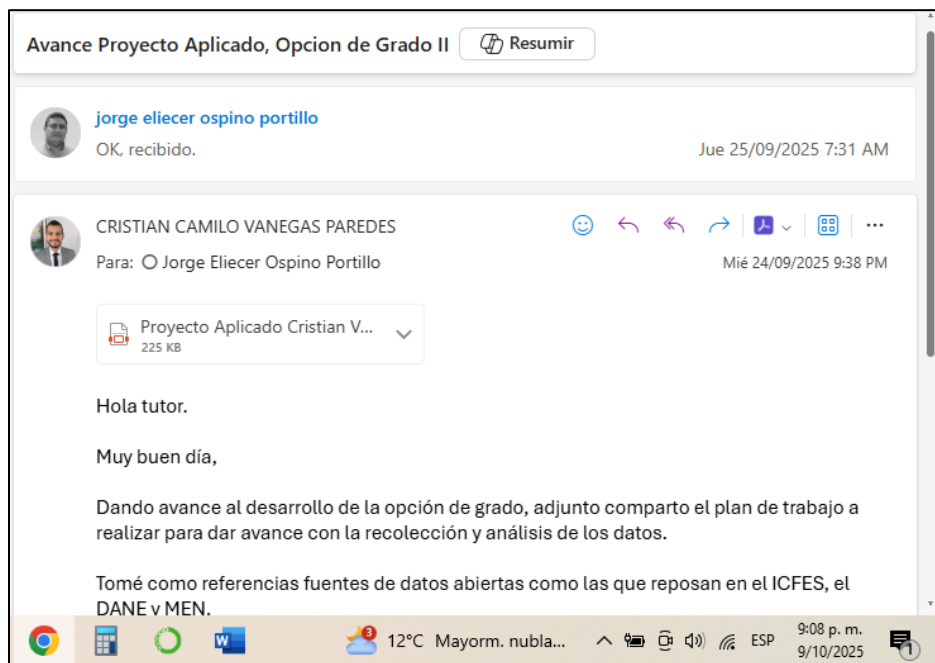
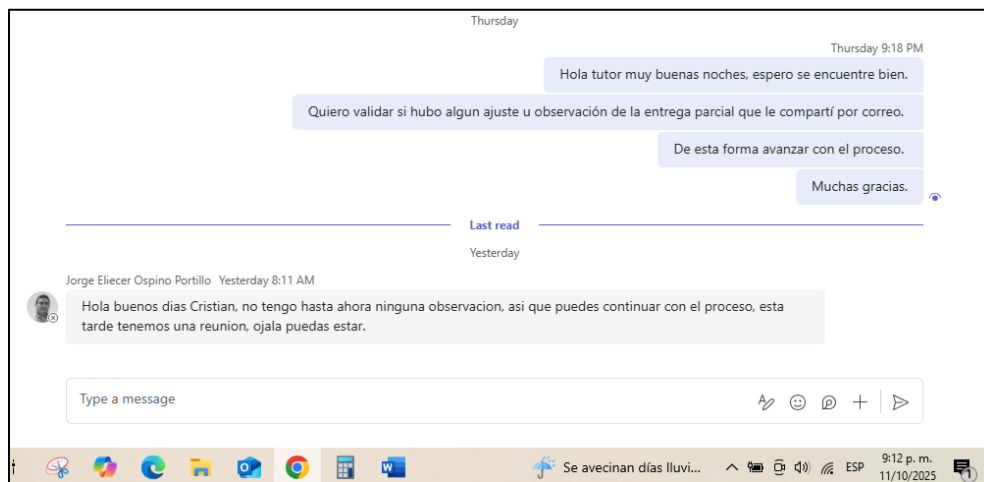
### Tercera Evidencia Contacto con Asesor del Proyecto de Grado



## Figura 5

### Cuarta Evidencia Contacto con Asesor del Proyecto de Grado



**Figura 6***Quinta Evidencia Contacto con Asesor del Proyecto de Grado***Figura 7***Sexta Evidencia Contacto con Asesor del Proyecto de Grado*

## Limpieza y Exploración de Datos (EDA)

Con los datos recopilados, se desarrolla un análisis descriptivo que incluye la construcción de gráficos, tablas y segmentaciones que permiten identificar patrones y tendencias relevantes.

Para este proceso, en primera instancia se utiliza el editor de Power Query, a través del cual se cargan, filtran y transforman las bases de datos provenientes de diferentes formatos, con el propósito de depurar y estructurar la información de manera adecuada para su posterior análisis.

### Figura 8

#### Segmentación y Filtrado de las Datas

The screenshot displays the Microsoft Power Query Editor window titled "GEIH DANE - Editor de Power Query". The main area shows a data table with the following columns: "Source-Name", "DIRECTORIO", "SECUENCIA\_P", "ORDEN", and "HOGAR". The table contains 23 rows of data, all with "Área - Características generales (Personas).csv" in the "Source-Name" column. The "DIRECTORIO" column contains unique identification numbers ranging from 4585999 to 4587126. The "SECUENCIA\_P" column contains values 1, 3, 2, 1, 1, 4, 2, 1, 2, 1, 1, 1, 2, 3, 2, 7, 2, 5, 2, 2, 2, 2. The "ORDEN" column contains values 1, 1. The "HOGAR" column contains values 1, 3, 1, 1, 1, 2, 4, 2, 2, 1, 1, 1, 2, 3, 2, 7, 2, 5, 2, 2, 2, 2.

The interface includes a ribbon with tabs for "Inicio", "Transformar", "Agregar columna", and "Vista". The "Transformar" tab is active, showing options like "Elegir columnas", "Quitar columnas", "Conservar filas", "Quitar filas", "Dividir columna", "Agrupar por", and "Reemplazar los valores". The right-hand pane shows the "Configuración de la consulta" and "PASOS APLICADOS" sections. The "PASOS APLICADOS" section lists several steps, including "Filas filtradas", which is currently selected.

Figura 9

## Cargue de Datos

```

Cargar y estandarizar cada fuente

# =====
# 1. Cargar las bases de datos
# =====
ruta_men = "MEN.xlsx"
ruta_icfes = "ICFES.xlsx"
ruta_dane = "DANE.xlsx"

men = pd.read_excel(ruta_men)
icfes = pd.read_excel(ruta_icfes)
dane = pd.read_excel(ruta_dane)

print("MEN:", men.shape)
print("ICFES:", icfes.shape)
print("DANE:", dane.shape)

MEN: (429, 37)
ICFES: (994, 88)
DANE: (479, 40)

```

Figura 10

## Limpieza y Selección de Variable - MEN

	AÑO	COD_DEPTO	DEPARTAMENTO	TASA_DESERCIÓN_TOTAL	DESERCIÓN_TRANSICIÓN	\
0	2011	5	Antioquia	397	362.0	
1	2011	8	Atlántico	276	26.0	
2	2011	11	Bogotá, D.C.	395	1005.0	
3	2011	13	Bolívar	314	185.0	
4	2011	15	Boyacá	307	24.0	
	DESERCIÓN_PRIMARIA	DESERCIÓN_SECUNDARIA	DESERCIÓN_MEDIA	\		
0	365	457	371			
1	306	242	261			
2	53	196	255			
3	293	379	313			
4	224	403	351			
	TASA_APROBACION_TOTAL	APROBACIÓN_TRANSICIÓN	APROBACIÓN_PRIMARIA	\		
0	9398	7	9456			
1	967	12	9649			
2	9605	0	9469			
3	9476	46	9548			
4	942	17	961			
	APROBACIÓN_SECUNDARIA	APROBACIÓN_MEDIA	COBERTURA_NETA	COBERTURA_BRUTA		
0	457	9333	9385	10681		
1	242	9664	9905	10788		
2	196	9745	9029	9778		
3	379	932	914	11041		
4	403	9323	8611	10415		

ICFES: desempeño académico + año + departamento

**Figura 11***Limpieza y Agrupación ICFES*

	AÑO	COD_DEPTO	n_estudiantes	punt_global_promedio	\
0	2018	5.0	16	278.250000	
1	2018	8.0	12	254.916667	
2	2018	11.0	26	262.230769	
3	2018	13.0	12	234.583333	
4	2018	15.0	7	255.285714	
			punt_matematicas_promedio	punt_lectura_promedio	\
0			58.187500	59.187500	
1			49.000000	54.250000	
2			54.538462	52.730769	
3			46.083333	50.083333	
4			49.857143	52.714286	
			percentil_global_promedio	tasa_aprobacion_icfes	
0			65.250000	0.625000	
1			52.750000	0.500000	
2			59.307692	0.615385	
3			41.333333	0.333333	
4			49.571429	0.428571	

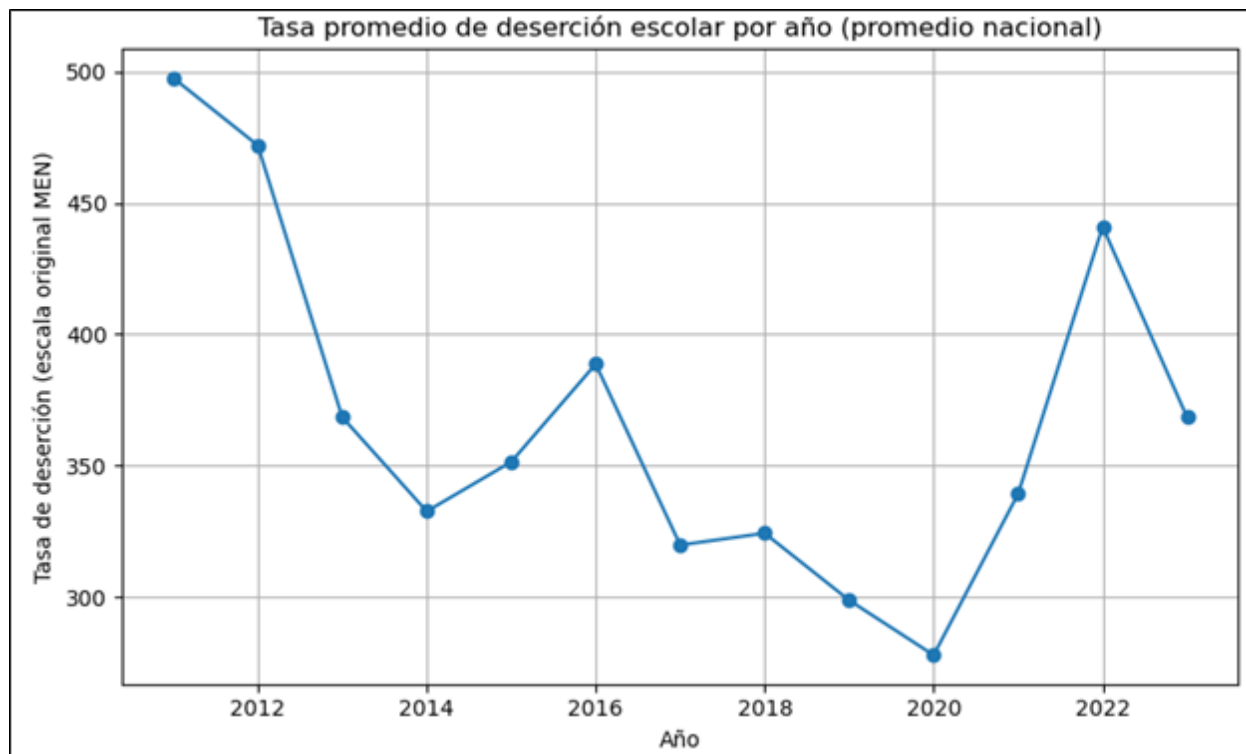
**Figura 12***Limpieza y Agrupación DANE*

	AÑO	COD_DEPTO	n_personas	tasa_pobreza	escolaridad_promedio
0	2011	5.0	22	0.181818	7.928571
1	2011	8.0	23	0.391304	7.357143
2	2011	11.0	10	0.100000	11.666667
3	2011	13.0	13	0.230769	12.166667
4	2011	15.0	7	0.000000	NaN

## Análisis Descriptivo y Gráficas

**Figura 13**

*Tasa de Deserción por Año*



La gráfica muestra la evolución anual de la deserción escolar entre aproximadamente 2011 y 2023.

### **Tendencia General Decreciente entre 2011 y 2014**

La deserción inicia en valores muy altos (cerca de 500) alrededor de 2011.

Entre 2011 y 2014 se observa una reducción sostenida, llegando a valores cercanos a 330.

Interpretación: Es posible que este periodo coincida con políticas de permanencia educativa y estrategias de cobertura desarrolladas por el Ministerio de Educación Nacional, como:

- Ampliación del Programa de Alimentación Escolar (PAE)

- Jornada única piloto
- Fortalecimiento de campañas de permanencia
- En tu análisis puedes destacar que esta tendencia sugiere un impacto positivo de

políticas sociales y educativas iniciales.

### **Periodo de Estabilización con Ligeros Repuntes (2014–2019)**

Entre 2014 y 2016 hay un pequeño repunte (de ~330 a ~390).

Entre 2017 y 2019 vuelve a descender gradualmente hasta niveles cercanos a 300.

Interpretación: Esto indica que, aunque hubo mejoras previas, hubo factores que generaron fluctuaciones, posiblemente asociados a:

Diferencias territoriales en cobertura y calidad.

Factores socioeconómicos (aumento del desempleo juvenil, migración venezolana).

Disparidades urbano–rural.

Aquí puedes escribir que la deserción, aunque controlada, sigue siendo sensible a cambios económicos y sociales.

### **Caída Abrupta en 2020 (Pandemia de COVID-19)**

En 2020 se observa un mínimo histórico (~275).

Esto parece contradictorio, ya que se sabe que el impacto de la pandemia aumentó la deserción, no la redujo. Por tanto, esta disminución no es interpretativa, sino una señal de que:

La forma en que el MEN registró la deserción en 2020 cambió por:

Cierre generalizado de instituciones.

Dificultad para registrar matrícula real.

Transición a educación remota sin control formal de permanencia.

Cambios en denominadores poblacionales o mecanismos de reporte.

Conclusión técnica: Ese “descenso” es artefactual y debe explicarse como una anomalía estadística causada por la pandemia, no como una mejora real.

### **Incremento muy Pronunciado Post-pandemia (2021–2022)**

En 2021 sube de ~275 a ~340.

En 2022 sube todavía más, llegando a ~440 (casi regresando a los niveles de 2011).

Interpretación: Este aumento sí es consistente con lo que reportaron:

- El Banco Mundial
- El BID
- El MEN
- Estudios de la UNESCO sobre pérdida de aprendizaje y rezago educativo

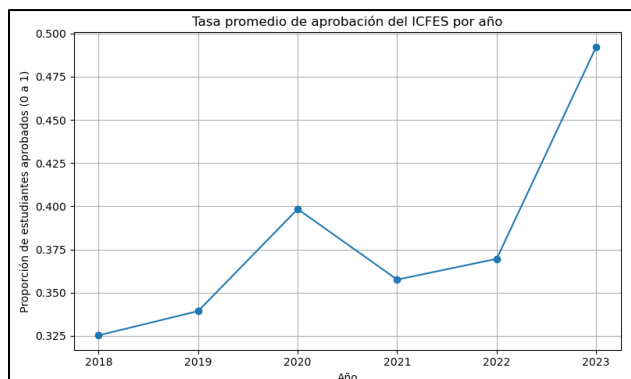
La reapertura escolar en 2021 reveló:

- Rezago acumulado en aprendizaje
- Estudiantes que no volvieron a clases
- Incrementos en trabajo infantil
- Crisis económica de los hogares durante y después del confinamiento

### **Análisis Data ICFES**

#### **Figura 14**

*Promedio de Aprobación del ICFES por Año*



La evolución de la tasa promedio de aprobación en las pruebas Saber 11 (ICFES) entre 2018 y 2023 evidencia cambios relevantes en el desempeño académico de los estudiantes a nivel nacional. En 2018 la proporción de estudiantes que superaron el umbral de aprobación se ubicó en 0.325, aumentando ligeramente en 2019 hasta 0.339. Este comportamiento relativamente estable sugiere que, antes de la pandemia, el rendimiento académico medido por el ICFES presentaba variaciones moderadas entre cohortes.

En 2020 se observa un incremento importante en la tasa de aprobación, que alcanza 0.398. Este valor debe interpretarse con cautela, ya que coincide con el periodo de pandemia por COVID-19, durante el cual se modificaron las condiciones de aplicación de las pruebas, los procesos educativos y los mecanismos de evaluación. Es posible que este aumento no refleje únicamente una mejora real en el desempeño, sino también cambios en los criterios, el número de estudiantes evaluados o las condiciones excepcionales de ese año académico.

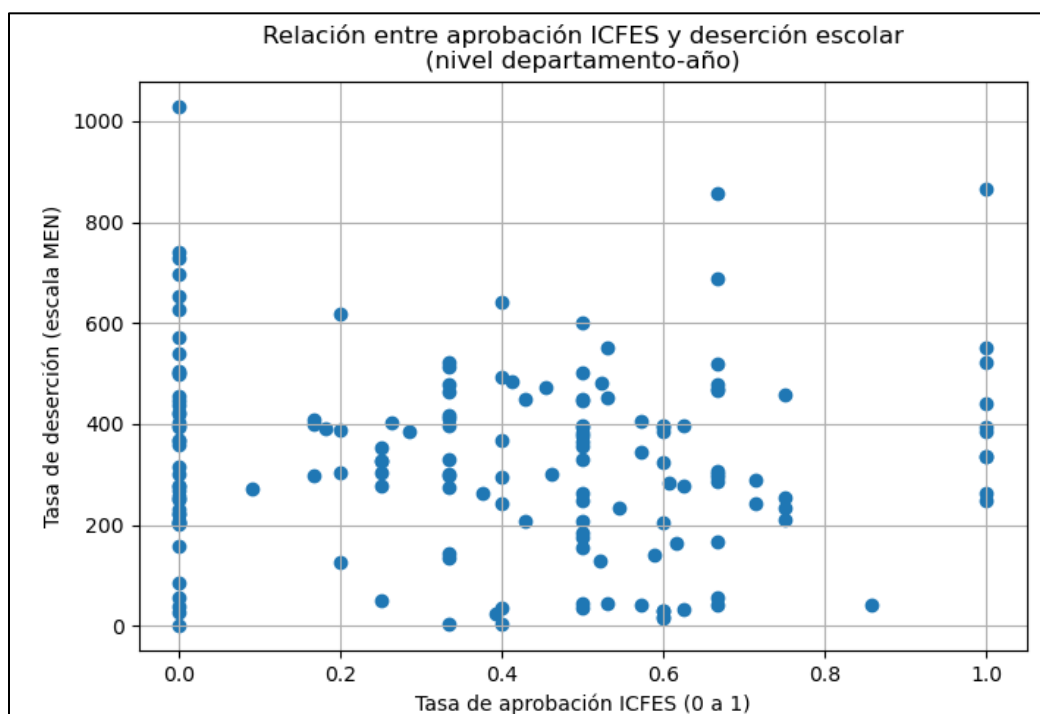
A partir de 2021 la proporción de aprobación desciende nuevamente a niveles cercanos a los de 2019 (0.358), lo que puede indicar un impacto académico negativo derivado del retorno a la presencialidad después de un largo periodo de educación remota, rezago acumulado y brechas en acceso a recursos educativos. En 2022 se mantiene una tendencia levemente ascendente, alcanzando 0.370, lo que podría sugerir un proceso gradual de recuperación en el rendimiento.

Finalmente, en 2023 se registra el valor más alto del periodo analizado (0.492), lo cual constituye un aumento significativo respecto a los años anteriores. Este comportamiento puede interpretarse como un posible efecto de estabilización del sistema educativo posterior a la pandemia, un fortalecimiento en las estrategias institucionales de acompañamiento académico, o un mejoramiento en la preparación de los estudiantes.

En conjunto, los resultados muestran que el rendimiento académico medido por la aprobación del ICFES no se ha comportado de forma lineal durante los años recientes, y que eventos sistémicos como la pandemia influyeron de manera marcada en los patrones de desempeño. Para efectos del análisis de la deserción escolar, esta variabilidad resulta relevante, ya que la literatura señala una relación directa entre bajo rendimiento académico y mayor probabilidad de abandono. Por tanto, las fluctuaciones observadas en la aprobación pueden contribuir a explicar tendencias posteriores en las tasas de deserción a nivel nacional.

### Figura 15

*Relación entre Aprobación ICFES y Deserción Escolar*



Interpretación:

La gráfica presenta la relación entre:

- Eje X: tasa de aprobación del ICFES (proporción de estudiantes aprobados en un departamento–año).
- Eje Y: tasa de deserción escolar reportada por el MEN (escala original).
- Cada punto representa un departamento en un año específico.

A partir de la distribución de los datos, se pueden extraer las siguientes conclusiones relevantes:

No existe una relación lineal simple entre aprobación del ICFES y deserción

La nube de puntos no forma una tendencia clara ascendente ni descendente, lo que indica que:

La aprobación en el ICFES por sí sola no explica completamente la deserción escolar.

Esto refuerza que el abandono escolar es un fenómeno multidimensional, influido por factores educativos, socioeconómicos y contextuales.

Se observan muchos departamentos con baja aprobación y alta deserción

En el cuadrante izquierdo (tasa de aprobación entre 0.0 y 0.3), aparece una alta concentración de puntos con:

Aprobación muy baja en ICFES

Deserción elevada (300 a 600 o más)

Interpretación: Este comportamiento coincide con lo esperado:

Departamentos donde un menor porcentaje de estudiantes alcanza un desempeño mínimo tienden a mostrar mayores tasas de abandono escolar.

Esto puede explicarse por:

Dificultades acumuladas de aprendizaje

Escuelas con menor calidad educativa

Limitaciones en seguimiento académico

Contextos socioeconómicos más vulnerables

Departamentos con buenas tasas de aprobación pueden tener deserción alta

Observas varios puntos hacia la derecha de la gráfica (aprobación  $\geq 0.6$  e incluso 1.0) que sin embargo muestran:

Deserción alta (300–500)

Incluso casos por encima de 800

Esto es muy importante para tu tesis, porque significa:

Un buen desempeño en el ICFES no garantiza automáticamente una baja deserción.

Esto refleja que:

Factores externos al rendimiento académico (económicos, sociales, territoriales) siguen siendo determinantes.

La permanencia no depende únicamente de las capacidades del estudiante, sino también de condiciones estructurales.

Amplia dispersión en casi todos los rangos de aprobación

La variabilidad en el eje Y (deserción) es grande en casi todos los niveles de aprobación.

Esto sugiere:

Los departamentos con niveles similares de desempeño académico pueden comportarse muy distinto en términos de permanencia escolar.

Por tanto, el ICFES no es un predictor suficiente, aunque sí aporta información.

La relación entre la tasa de aprobación en el ICFES y la deserción escolar muestra una asociación compleja y no lineal. Aunque se observa que varios departamentos con bajas tasas de aprobación presentan niveles altos de deserción, el gráfico también evidencia que existen

territorios con buenos niveles de desempeño académico que, aun así, mantienen deserciones elevadas. Esto indica que el rendimiento en las pruebas Saber 11, si bien es un factor asociado al abandono escolar, no es suficiente por sí solo para explicar las variaciones entre departamentos.

La amplia dispersión de puntos en todos los rangos de aprobación sugiere que la deserción escolar responde a múltiples dimensiones, entre ellas condiciones socioeconómicas, calidad educativa, acceso efectivo al sistema, e incluso dinámicas territoriales particulares. En consecuencia, la aprobación del ICFES debe interpretarse como un indicador parcial dentro de un conjunto más amplio de factores que determinan la permanencia escolar. Este resultado concuerda con los hallazgos del modelo predictivo construido, donde las variables asociadas a cobertura educativa mostraron una mayor importancia relativa que los indicadores de rendimiento académico.

### Modelo de Predicción de Riesgo de Deserción

MAE (Error absoluto medio): 161.90

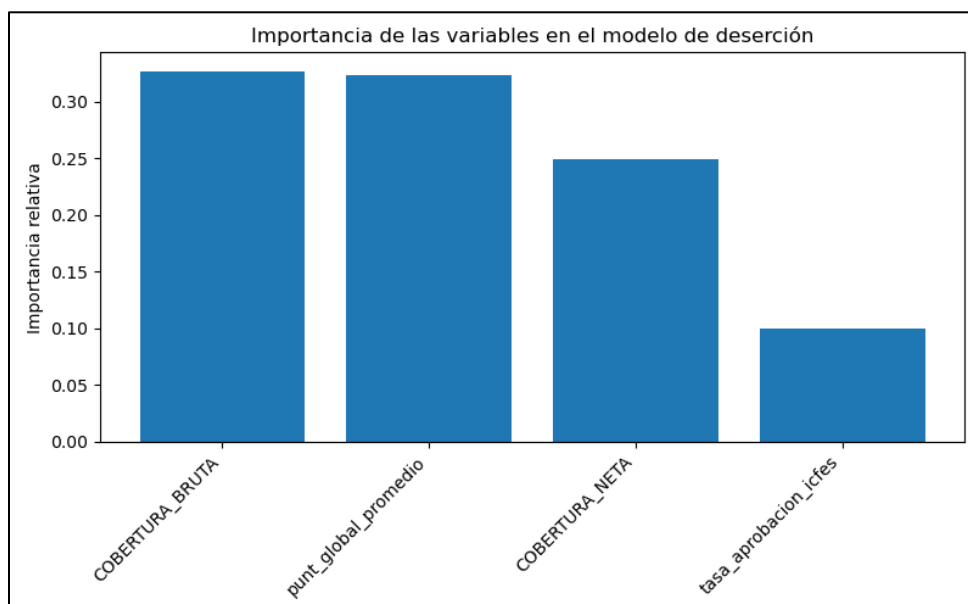
R<sup>2</sup> (coeficiente de determinación): 0.026

El modelo predictivo entrenado mostró un desempeño limitado (MAE = 161.9; R<sup>2</sup> = 0.026), lo que indica que los indicadores educativos utilizados no son suficientes para capturar la complejidad del fenómeno de la deserción escolar. Esto sugiere que la deserción depende en mayor medida de factores socioeconómicos y contextuales no incluidos aún en el modelo, tales como condiciones de pobreza, nivel educativo del hogar, características territoriales y condiciones institucionales. Por tanto, este modelo debe considerarse una aproximación preliminar que será fortalecida en fases posteriores mediante la incorporación de variables adicionales y técnicas más robustas.

### Importancia de Variables

#### Figura 16

##### *Importancia de las Variables*



El modelo Random Forest evaluó la contribución relativa de cada variable para predecir la tasa de deserción escolar a nivel departamento/año. Los resultados fueron los siguientes:

Variable Importancia COBERTURA\_BRUTA 0.327 punt\_global\_promedio 0.323  
COBERTURA\_NETA 0.249 tasa\_aprobacion\_icfes 0.100

Estos valores indican cómo cada predictor contribuye a la reducción del error en el modelo. A partir de ellos se pueden derivar las siguientes conclusiones:

La gráfica muestra el peso relativo que el modelo asigna a cada variable al momento de predecir la tasa de deserción escolar. Cuanto mayor es la barra, mayor es la contribución de esa variable a la capacidad del modelo para explicar y predecir la deserción.

COBERTURA\_BRUTA — Variable más influyente en la deserción (0.327)

La cobertura bruta mide la cantidad total de estudiantes matriculados, independientemente de la edad para el grado correspondiente.

Interpretación:

Es la variable más relevante del modelo.

Una mayor cobertura bruta se asocia con menores niveles de deserción, indicando que los territorios con más estudiantes dentro del sistema educativo tienden a mantenerlos matriculados.

Sugiere que los departamentos con buena capacidad para absorber a la población estudiantil (infraestructura, cupos disponibles, acceso) logran mayor retención escolar.

Implicación para políticas educativas:

Incrementar los niveles de cobertura, especialmente en zonas donde la matrícula bruta es baja, puede ser una estrategia clave para reducir la deserción escolar.

punt\_global\_promedio: Peso casi igual al de la cobertura bruta (0.323)

Este indicador representa el promedio de los puntajes globales obtenidos por los estudiantes en las pruebas Saber 11.

Interpretación:

El rendimiento académico explica casi la misma proporción de la deserción que la cobertura bruta.

Departamentos con mejores resultados en el ICFES tienden a presentar menores tasas de deserción, reforzando la relación entre desempeño escolar y permanencia.

El rendimiento académico actúa como un predictor directo del riesgo educativo.

Conclusión educativa:

Los estudiantes con bajo desempeño académico presentan mayor probabilidad de abandonar el sistema, lo que subraya la necesidad de estrategias de apoyo pedagógico.

COBERTURA\_NETA Importancia intermedia (0.249)

La cobertura neta mide la proporción de estudiantes matriculados en el grado correspondiente a su edad.

Interpretación:

Es un indicador importante, pero menos decisivo que la cobertura bruta.

Sugiere que estar en el grado adecuado para la edad favorece la permanencia, pero no con el mismo impacto que el volumen total de matrícula o el rendimiento académico.

La cobertura neta puede reflejar problemas en trayectorias educativas interrumpidas, rezago escolar o repetición.

Implicación:

Mejorar la regularidad en las trayectorias escolares puede contribuir a disminuir la probabilidad de abandono.

tasa\_ aprobacion\_ icfes Menor importancia relativa (0.100)

Este indicador refleja el porcentaje de estudiantes que superan un umbral mínimo en el ICFES.

Interpretación:

Aunque influye en el modelo, su peso es mucho menor comparado con las otras variables.

Esto sugiere que aprobar o no el examen Saber 11 tiene menos poder explicativo, comparado con factores estructurales como la cobertura o el desempeño promedio.

Es posible que la aprobación sea un indicador demasiado simplificado, mientras que el puntaje promedio captura más matices del desempeño académico.

La aprobación mínima del examen es menos informativa que la calidad o consistencia del desempeño promedio del estudiantado.

Los resultados muestran que la deserción escolar está determinada por una combinación de factores estructurales y académicos. La cobertura bruta y el puntaje global promedio del ICFES emergen como los predictores más influyentes, indicando que tanto la capacidad del sistema educativo para mantener a los estudiantes matriculados como el nivel de desempeño académico son elementos clave en la permanencia escolar. La cobertura neta también presenta un papel relevante, mientras que la tasa de aprobación del ICFES aporta información adicional pero con menor peso relativo. Estos hallazgos refuerzan la naturaleza multidimensional de la deserción escolar y destacan la necesidad de intervenciones integrales que combinen estrategias de acceso, calidad educativa y acompañamiento académico.

### Hacia el “Riesgo de un Estudiante”- Modelo de datos

Con este modelo a nivel agregado se logra:

Clasificar segmentos de alto riesgo

Por ejemplo, crear un indicador binario:

1 = “alto riesgo” si la tasa de deserción predicha está en el 25% superior.

0 = “riesgo bajo o medio” en otro caso.

Asignar ese riesgo a los estudiantes

Cada estudiante en el ICFES pertenece a un departamento y a unas condiciones socioeconómicas.

Asignarle el riesgo del segmento de su departamento/año.

#### Figura 17

*Predicción Riesgo de Deserción*

	AÑO	COD_DEPTO	tasa_desercion_predicha	alto_riesgo_segmento
0	2018	5	313.825	0
1	2018	8	244.460	0
2	2018	11	322.990	0
3	2018	13	236.750	0
4	2018	15	267.845	0

¿Qué explica que no aparezca ningún departamento como alto riesgo?

Las variables del modelo (rendimiento del ICFES y coberturas) no producen grandes contrastes

Esto ya se observó en las gráficas:

La dispersión de la deserción es alta,

La relación con ICFES no es lineal,

La cobertura sí influye, pero con variación moderada.

Por lo tanto, el modelo predice diferencias, pero no diferencias abruptas en estos departamentos.

Los resultados del modelo muestran que, para el año 2018, ninguno de los departamentos analizados fue clasificado dentro del segmento de alto riesgo de deserción escolar. Las tasas predichas oscilan entre 236.75 y 322.99, lo que evidencia una variabilidad moderada entre los territorios. Aunque algunos departamentos presentan valores superiores dentro del grupo (como el departamento 11), estos no superan el umbral establecido por el cuartil 75% de las predicciones a nivel nacional, por lo cual permanecen en la categoría de riesgo medio.

Este comportamiento sugiere que, en el contexto del modelo actual, las diferencias en deserción están relativamente concentradas y que los predictores educativos utilizados principalmente la cobertura educativa y el desempeño académico— no generan contrastes extremos entre estos departamentos. Asimismo, la ausencia de territorios clasificados como de alto riesgo puede deberse a que otros departamentos del país presentan valores considerablemente más altos de deserción predicha, desplazando a los aquí analizados hacia una posición intermedia dentro de la distribución general. líticas de permanencia en este departamento.

Los demás departamentos analizados presentan tasas predichas considerablemente menores, ubicándose en niveles de riesgo medio o bajo. Esto evidencia diferencias territoriales en los determinantes de la deserción, lo cual coincide con estudios previos sobre desigualdades regionales en cobertura, desempeño académico y condiciones socioeconómicas. La clasificación generada por el modelo puede servir como una herramienta temprana de focalización para orientar intervenciones educativas.

## Conclusiones

El desarrollo del proyecto “Análisis de la Deserción Escolar en Colombia: Factores Socioeconómicos y Estrategias de Intervención para Reducir el Abandono Educativo” permitirá generar una comprensión integral del fenómeno de la deserción, apoyada en evidencia empírica y herramientas analíticas que orienten la toma de decisiones educativas.

El plan de trabajo definido establece una ruta clara y progresiva que inicia con la recopilación y depuración de datos provenientes de fuentes oficiales como el Ministerio de Educación Nacional, el DANE e ICFES, garantizando la confiabilidad de la información. La posterior exploración y análisis estadístico facilitará la identificación de patrones de riesgo y variables determinantes en distintos contextos territoriales y socioeconómicos.

La implementación de modelos predictivos y de agrupamiento contribuirá a la formulación de estrategias focalizadas, permitiendo reconocer comunidades o instituciones con mayor vulnerabilidad al abandono escolar. A su vez, este proceso fortalecerá la capacidad analítica para diseñar políticas públicas basadas en datos y con enfoque territorial.

Finalmente, la comunicación constante con el asesor académico y la validación de los resultados con expertos asegurarán la pertinencia metodológica y la aplicabilidad práctica de las conclusiones. Con ello, el proyecto no solo busca describir el problema, sino proponer soluciones sostenibles que promuevan la permanencia escolar, la equidad educativa y la reducción de brechas sociales en Colombia.

El análisis de la deserción escolar en Colombia realizado en el presente proyecto aplicado permitió comprender la naturaleza multidimensional de este fenómeno y evidenció la necesidad de abordarlo desde una perspectiva que articule indicadores educativos, socioeconómicos y de desempeño académico. A partir de la integración de diversas fuentes de información oficiales —

MEN, ICFES y DANE— y del desarrollo de procesos de exploración, transformación, análisis y modelado predictivo, fue posible generar una aproximación rigurosa al comportamiento de la deserción en los departamentos del país durante el periodo analizado.

En primer lugar, el análisis descriptivo mostró tendencias claras en la evolución de la deserción escolar, destacándose una disminución progresiva entre 2011 y 2014, seguida de fluctuaciones intermedias antes de la pandemia, y un comportamiento crítico posterior a 2020. Aunque en los registros del MEN se observó una aparente reducción abrupta en la deserción durante 2020, la interpretación de este hallazgo debe hacerse con cautela, pues las condiciones excepcionales de la pandemia alteraron los mecanismos de reporte y matrícula, lo que pudo generar una subestimación temporal del indicador. Los años subsiguientes (2021 y 2022) reflejan el impacto real del rezago educativo y las desigualdades socioeconómicas intensificadas, evidenciándose un aumento considerable en la deserción que refuerza la urgencia de estrategias focalizadas de permanencia escolar.

En cuanto al desempeño académico, la evolución de la tasa de aprobación del ICFES reveló variaciones importantes entre años y la presencia de marcadores que coinciden con eventos críticos del sistema educativo. Si bien existe cierta relación entre bajo rendimiento y mayor deserción, los análisis de dispersión mostraron que esta asociación no es lineal ni suficiente para explicar el abandono escolar en su totalidad. Departamentos con niveles similares de aprobación presentan comportamientos muy distintos en términos de permanencia, lo cual confirma que el rendimiento académico es solo un componente dentro de un entramado más amplio de determinantes.

### Referencias Bibliográficas

- Aravena Domich, M., & Pérez Campillo, C. C. (2022). La (no) deserción escolar. Análisis de estudiantes de 9.º grado en Colombia. *Revista Andina de Educación*, 5(1).  
<https://doi.org/10.32719/26312816.2022.5.1.11>
- Artunduaga Parra, D. (2023). Factores de aprendizaje que influyen en la educación primaria. *Ciencia Latina Revista Científica Multidisciplinar*, 7(4).  
[https://doi.org/10.37811/cl\\_rcm.v7i4.6873](https://doi.org/10.37811/cl_rcm.v7i4.6873)
- Ayala Vásquez, Y. I. (2024, June 1). Vista de Factores de deserción y repetición escolar en las instituciones educativas públicas en Medellín | *Revista Multidisciplinaria Investigación Contemporánea*. *Revista multidisciplinaria investigación Contemporánea*.  
<https://revmic.com/index.php/IC/article/view/56/78>
- Castillo, J., Barrera, H., & Moreno, H. (2022). Causas de la deserción del sistema de educación escolar. Un estudio de caso. *Gestionar: Revista de Empresa y Gobierno*, 2(1).  
<https://doi.org/10.35622/j.rg.2022.01.004>
- Chalela-Naffah, S., Valencia-Arias, A., Ruíz-Rojas, G. A., & Cadavid-Orrego, M. (2020). Factores psicosociales y familiares que influyen en la deserción en estudiantes universitarios en el contexto de los países en desarrollo. *Revista Lasallista de Investigación*, 17(1). <https://doi.org/10.22507/rli.v17n1a9>
- DANE. (Año de consulta). *Gran Encuesta Integrada de Hogares – GEIH*.  
<https://www.dane.gov.co>
- Deyser Gutiérrez, A. A., Fredy Vélez Díaz, J., & López, J. M. (2021). Indicadores de deserción universitaria y factores asociados. *Revista EducaT: Educación Virtual, Innovación y Tecnologías*, 2(1).

Gómez Molina, S., Valencia-Arías, A., Saldarriaga Ríos, J. G., Vélez Holguín, R. M., & Soto Giraldo, J. O. (2022). Deserción escolar de niños y niñas en Colombia en tiempos de pandemia. *Telos Revista de Estudios Interdisciplinarios En Ciencias Sociales*, 24(3).  
<https://doi.org/10.36390/telos243.11>

Gómez-Restrepo, C., Padilla Muñoz, A., & Rincón, C. J. (2016). Deserción escolar de adolescentes a partir de un estudio de corte transversal: Encuesta Nacional de Salud Mental Colombia 2015. *Revista Colombiana de Psiquiatría*, 45, 105–112.  
<https://doi.org/10.1016/J.RCP.2016.09.003>

Instituto Colombiano para la Evaluación de la Educación – ICFES. (Año de consulta). *Base de datos de resultados de las pruebas Saber 11*. <https://www.icfes.gov.co>

Manzano López, D. J., & Ramírez Zambrano, J. R. (2023). Interrelación entre la deserción escolar y las condiciones socioeconómicas de las familias: el caso de la ciudad de Cúcuta (Colombia)\*. <http://www.scielo.org.co/pdf/ecoca/n10/n10a07.pdf>

MEJÍA HIGUITA, L. M. (2022). ANÁLISIS COMPARATIVO DE LA DESERCIÓN ESCOLAR EN COLOMBIA: LOS CASOS DE MEDELLÍN Y BOGOTÁ.  
<https://rehip.unr.edu.ar/server/api/core/bitstreams/10721548-54a7-4c13-a240-2f565ddf179a/content>

Ministerio de Educación Nacional. (2022). DESERCIÓN ESCOLAR EN COLOMBIA: ANÁLISIS, DETERMINANTES Y POLÍTICA DE ACOGIDA, BIENESTAR Y PERMANENCIA. Nota Técnica.

Ministerio de Educación Nacional. (Año de consulta). *Sistema de Información de la Educación Básica y Media (SIEE) – Indicadores educativos*. <https://www.mineducacion.gov.co>

RAMSÉS VARGAS LAMADRID. (2017). Deserción escolar en Colombia, por Ramsés Vargas Lamadrid. Revista Semana.

Sosa Chinome, J. A., & Rolando Vargas, G. (2023). Consecuencias de la deserción escolar de niños y jóvenes víctimas del conflicto armado en Colombia. CITAS.

<https://doi.org/10.15332/24224529.7267>

Vela Hernández, S. P., Quintero-Zabala, S. P., Peluffo-Arias, N. P., & Escobar-Díaz, F. A. (2022). Implementación de un programa de alimentación escolar en una institución educativa durante la pandemia, Bogotá, 2021. Revista de Salud Pública, 24(3).

<https://doi.org/10.15446/rsap.v24n3.97562>