

# **PROYECTO FINAL DE ANÁLISIS DE DATOS NIVEL EXPLORADOR**

TÍTULO DEL PROYECTO

**Modelo: Análisis exploratorio de factores asociados a la deserción estudiantil en la Universidad de Antioquia: una mirada desde la interseccionalidad étnico-racial, socioeconómica y académica**

ELABORADO POR

**Sebastian Cortes Sanchez**

MINISTERIO DE LAS TECNOLOGÍAS DE LA INFORMACIÓN Y LAS  
COMUNICACIONES – MINTIC  
TALENTO TECH

MEDELLÍN  
2025

## **Introducción**

En Colombia, la permanencia universitaria se ve afectada por una serie de condiciones estructurales de orden social, económico, étnico y académico. Este proyecto busca analizar, desde una perspectiva exploratoria, los factores que inciden en la deserción estudiantil en la Universidad, en este caso la Universidad de Antioquia utilizando una base de datos adaptada de instituciones públicas similares. La intención es identificar patrones, correlaciones y posibles brechas, para aportar a la toma de decisiones institucionales que fomenten una educación superior más pertinente.

## **Tabla de contenido**

1. Introducción
2. Objetivos del Proyecto
  - 2.1 Objetivo General
  - 2.2 Objetivos Específicos
3. Desarrollo del Proyecto
  - 3.1 Identificación de la Problemática
  - 3.2 Recolección de Datos
  - 3.3 Exploración y Limpieza de Datos

3.4 Análisis Descriptivo

3.5 Interpretación de Resultados

4. Conclusiones

5. Resultados Esperados

## **Objetivos del Proyecto**

### **Objetivo General**

Realizar un análisis exploratorio de datos sobre la deserción estudiantil en la Universidad de Antioquia, enfocándose en los estudiantes con condiciones de vulnerabilidad socioeconómica, académica o étnico-racial, con el fin de identificar factores asociados a la deserción y posibles rutas de intervención institucional.

### **Objetivos Específicos**

- Analizar variables socioeconómicas y académicas para identificar factores que inciden en la deserción.
- Examinar el impacto de las condiciones socioeconómicas en la deserción estudiantil.
- Evaluar la relación entre rendimiento académico y abandono de estudios.
- Analizar la deserción por autoreconocimiento étnico (afrodescendientes e indígenas).
- Identificar posibles factores adicionales (sociales, culturales, discriminatorios) que afecten a poblaciones específicas.

## **Desarrollo del Proyecto**

### **3.1 Identificación de la Problemática**

La deserción estudiantil es uno de los principales retos que enfrentan las universidades públicas en Colombia. En este análisis, se aborda la deserción en la Universidad de Antioquia, considerando las condiciones socioeconómicas, académicas y de vulnerabilidad de los estudiantes. Este enfoque permitirá aportar al diseño de estrategias institucionales para mitigar el abandono, especialmente entre estudiantes con mayor riesgo.

### **3.2 Recolección de Datos**

Recolección de Datos Se utilizaron datos de los estudiantes de la institución Pascual Bravo, adaptados para simular la caracterización de la Universidad de Antioquia. La base de datos contiene información de 99,350 estudiantes con variables académicas, demográficas, económicas, de identidad étnica y condiciones de vulnerabilidad. Se añadió una variable adicional de "departamento de nacimiento" a partir del municipio de origen.

### **3.3 Exploración y Limpieza de Datos**

- Uso de bibliotecas de Python: pandas, numpy, matplotlib y seaborn.
- Revisión de estructura y tipos de variables.
- Identificación y tratamiento de valores faltantes.
- Limpieza de datos atípicos y estandarización de nombres.
- Selección de variables relevantes para el análisis.

### **3.4 Análisis Descriptivo**

A continuación, se presentan los principales cruces de variables con la variable de deserción y adaptados para simular la caracterización de la Universidad de Antioquia. La base de datos contiene información de 99,350 estudiantes con

variables académicas, demográficas, económicas, de identidad étnica y condiciones de vulnerabilidad. Se añadió una variable adicional de "departamento de nacimiento" a partir del municipio de origen.

### 3.3 Exploración y Limpieza de Datos

- Uso de bibliotecas de Python: pandas, numpy, matplotlib y seaborn.
- Revisión de estructura y tipos de variables.
- Identificación y tratamiento de valores faltantes.
- Limpieza de datos atípicos y estandarización de nombres.
- Selección de variables relevantes para el análisis.

## 1. Forma de pago vs. Deserción

Se observa que aunque programas de apoyo como Matrícula Cero y Sapiencia tienen alta cobertura, no necesariamente garantizan la permanencia. La deserción sigue siendo significativa en estos grupos, lo cual sugiere que el apoyo económico debe complementarse con acompañamiento psicosocial y académico.

## Autoreconocimiento étnico vs. Deserción

Aunque no hay una diferencia abismal en los porcentajes, se nota una leve mayor

deserción en estudiantes que se autoreconocen como afrodescendientes. Esto pone de relieve la necesidad de acciones afirmativas e inclusión con enfoque diferencial en las políticas institucionales.

### **Carga académica vs. Deserción**

La menor carga académica (menos asignaturas inscritas) está asociada con mayor tasa de deserción. Esto puede indicar que muchos estudiantes con baja carga académica trabajan o tienen otras responsabilidades que afectan su permanencia.

### **Tipo de programa vs. Deserción**

Los programas técnicos presentan mayor deserción que los programas profesionales. Esto podría relacionarse con las expectativas de los estudiantes

frente al valor del título o con condiciones más precarias de ingreso y permanencia.

### **Modalidad y jornada vs. Deserción**

La modalidad virtual muestra mayor vulnerabilidad frente a la presencial, y la jornada nocturna presenta mayores niveles de deserción, posiblemente por la doble carga laboral-académica.

### **Ser víctima del conflicto armado vs. Deserción**

Se detecta una ligera mayor deserción entre víctimas del conflicto armado, lo cual resalta la urgencia de medidas diferenciadas para esta población en cuanto a apoyo emocional, legal y académico.



## **Correlación entre variables numéricas**

La matriz de correlaciones muestra que la deserción tiene una relación inversa significativa con la aprobación de asignaturas y créditos. Esto ratifica el rol del rendimiento académico como factor crítico de permanencia. Aunque el autoreconocimiento étnico no muestra alta correlación directa, su relevancia emerge desde una perspectiva interseccional que cruza vulnerabilidades estructurales.

### 3.5 Interpretación de Resultados

- **Rendimiento académico:** Las variables de aprobación de asignaturas y créditos tienen correlaciones negativas significativas con la deserción.
- **Carga académica:** Estudiantes con menor número de asignaturas matriculadas tienden a desertar más.
- **Forma de pago:** Matrícula Cero y Sapiencia muestran correlaciones positivas leves con la deserción, lo que podría sugerir que el apoyo económico por sí solo no es suficiente.
- **Modalidad y jornada:** Se evidencian patrones distintos según modalidad (presencial vs. virtual) y tipo de jornada, siendo la virtual más vulnerable.
- **Condiciones de vulnerabilidad:** Ser víctima del conflicto armado se relaciona levemente con mayor deserción. El autoreconocimiento étnico no presenta fuertes correlaciones cuantitativas, pero se identifican patrones cualitativos significativos.

El análisis realizado indica que los factores académicos, especialmente el bajo rendimiento y la carga académica, son los principales predictores de deserción. Esto se refleja en la correlación negativa entre deserción y aprobación de asignaturas/créditos. La dimensión étnico-racial, aunque no presenta correlación numérica fuerte, debe interpretarse desde una mirada cualitativa e interseccional: los estudiantes afrodescendientes están más expuestos a múltiples formas de exclusión que afectan su permanencia.

Además, se evidencia que los programas de apoyo económico, por sí solos, no garantizan la permanencia estudiantil. La jornada nocturna, modalidad virtual, pertenencia a población víctima y programas técnicos también se relacionan con mayor deserción.

Por tanto, se requiere una política institucional integral que combine acompañamiento académico, emocional, financiero y cultural, con enfoque étnico-racial.

## Conclusiones

El análisis evidencia que la deserción en el modelo está principalmente influenciada por el bajo rendimiento académico y una carga de estudio reducida. Esto se pudiese deber a que los estudiantes con baja carga académica tienen otra ocupación como trabajo u realizan doble pregrado. Aunque las variables socioeconómicas y de identidad étnica no presentan correlaciones fuertes por sí solas, su interacción con otras condiciones estructurales sugiere que las políticas de permanencia deben ser integrales. Es fundamental un acompañamiento que incluya orientación académica, apoyo emocional, reconocimiento de trayectorias de vida diversas y acción afirmativa.

## Resultados Esperados

- Resultados Esperados
- Identificación de los factores más asociados con la deserción en estudiantes de la UdeA.
- Sistematización de las variables clave para políticas institucionales.
- Visibilización de la necesidad de acompañamiento integral más allá del apoyo económico.
- Generación de insumos para sistemas de alerta temprana.
- Recomendaciones para investigaciones futuras que profundicen en el impacto de la identidad étnico-racial en la permanencia universitaria.

Carpeta: <https://drive.google.com/drive/folders/19XepolrDsdydUJzzKEEd70ovZAuh-X0-?usp=sharing>