

**Diseño de un Sistema de Información con VBA Excel, para el Análisis del Comportamiento de la Matrícula, Continuidad y Graduación de los Estudiantes de la Universidad Nacional Abierta y a Distancia (UNAD) Pertenecientes al CEAD Bucaramanga de la Zona Centro Oriente**

Rafael Dionisio Ortega Almeida

Asesor

Jorge Armando Amador Moncada

Universidad Nacional Abierta y a Distancia UNAD  
Escuela de Ciencias Básicas Tecnologías e Ingenierías  
Especialización en Ciencia de Datos

2024

## Resumen

Este trabajo propone el diseño de una herramienta de análisis de datos que integra Microsoft Excel para abordar la deserción estudiantil en el CEAD Bucaramanga. El objetivo es proporcionar apoyo a la toma de decisiones estratégicas mediante un análisis exhaustivo de los datos del Registro y Control Académico, identificando variables relevantes que afectan la deserción. Se sugiere desarrollar un sistema que permita procesar grandes volúmenes de información y ejecutar algoritmos de análisis predictivo, facilitando así la minería de datos y el análisis estadístico.

La propuesta incluye la realización de una prueba piloto con Power BI para evaluar la efectividad de la herramienta, contando con la participación de directivos del CEAD para ajustar el modelo según sus retroalimentaciones. La metodología de investigación es mixta e incluye revisión de literatura, recopilación y análisis de datos, desarrollo de macros en Excel y presentación de resultados. El marco teórico abarca conceptos clave como la deserción estudiantil y la visualización de datos. En resumen, la investigación busca abordar la deserción estudiantil a través del análisis de datos y el desarrollo de tecnologías que favorezcan decisiones informadas en el CEAD Bucaramanga.

**Palabras clave:** Deserción, Análisis de datos, Microsoft Excel, Power BI, Toma de decisiones, Minería de datos.

## Abstract

This paper proposes the design of a data analysis tool that integrates Microsoft Excel to address student dropout at CEAD Bucaramanga. The objective is to provide support for strategic decision-making through a thorough analysis of Academic Registry and Control data, identifying relevant variables that affect dropout. It is suggested to develop a system that allows processing large volumes of information and executing predictive analysis algorithms, thus facilitating data mining and statistical analysis.

The proposal includes conducting a pilot test with Power BI to evaluate the effectiveness of the tool, with the participation of CEAD managers to adjust the model according to their feedback. The research methodology is mixed and includes literature review, data collection and analysis, development of macros in Excel, and presentation of results. The theoretical framework covers key concepts such as student dropout and data visualization. In summary, the research seeks to address student dropout through data analysis and the development of technologies that favor informed decisions at CEAD Bucaramanga.

**Keywords:** Attrition, Data analysis, Microsoft Excel, Power BI, Decision making, Data mining

## Tabla de Contenido

Introducción .....	9
Planteamiento del problema.....	11
Justificación .....	14
Objetivos .....	17
Objetivo General.....	17
Objetivos Específicos.....	17
Marco Referencial.....	18
Marco Teórico.....	18
Deserción Estudiantil .....	18
Análisis de Datos en Educación.....	18
Integración de VBA y Aplicaciones de Excel .....	18
Visualización de Datos con Power BI .....	20
Marco Legal .....	23
Metodología .....	25
Fase 1: Diagnóstico.....	27
Etapa 1: Revisión de Literatura y Conceptualización.....	27
Etapa 2: Recopilación de Datos .....	27
Etapa 3: Análisis Exploratorio de Datos (EDA).....	27
Fase 2: Diseño.....	28
Etapa 1: Desarrollo del Sistema de Integración VBA y Microsoft Excel.....	28
Etapa 2: Análisis de Datos .....	28
Etapa 3: Desarrollo del Modelo de Visualización de Datos .....	28
Fase 3: Evaluación .....	28

Etapa 1: Prueba Piloto del Modelo de Visualización.....	28
Etapa 2: Evaluación de la Herramienta Diseñada.....	29
Fase 4: Cierre .....	29
Etapa 1: Documentación y Redacción del Trabajo de Grado .....	29
Etapa 2: Presentación y Defensa del Trabajo de Grado.....	29
Resultados .....	30
Análisis de Datos del Proceso de Registro y Control Académico (RyC) del CEAD	
Bucaramanga.....	30
Recolección y Preprocesamiento de los Datos .....	30
Análisis Estadístico Descriptivo .....	31
Identificación de Patrones y Variables Relevantes .....	32
Variables Relevantes para la Deserción Estudiantil .....	32
Aplicar Técnicas de Minería de Datos y Análisis Estadístico en VBA – Aplicaciones de Excel para Evidenciar Correlaciones Significativas Entre las Variables y la Deserción Estudiantil, Permitiendo Así una Mejor Comprensión de los Factores que Afectan el Rendimiento Académico .....	
Técnicas de Minería de Datos y Análisis Estadístico en VBA/Excel.....	35
Análisis de Correlación.....	35
Regresión Lineal y Regresión Múltiple .....	36
Árboles de Decisión (Clasificación) .....	37
Visualización de Datos (Power BI y Excel) .....	38
Interactividad .....	38
Implementación de Funciones de Power BI .....	39

Evaluar la Efectividad del Modelo de Visualización de Datos en Power BI Mediante la Retroalimentación de los Usuarios, Favoreciendo la Toma de Decisiones Estratégicas en el CEAD Bucaramanga en Temas Relacionados con la Deserción Estudiantil .....	40
Evaluación de la Efectividad del Modelo de Visualización de Datos en Power BI .....	40
Métodos para Recoger la Retroalimentación .....	40
Análisis de Retroalimentación y Ajustes .....	44
Conclusiones .....	46
Recomendaciones .....	59
Referencias Bibliográficas .....	61

**Lista de Tablas**

<b>Tabla 1</b> <i>Características de los Estudiantes</i> .....	34
--	----

## Lista de Figuras

<b>Figura 1</b> <i>Esquema de la metodología utilizada</i> .....	24
<b>Figura 2</b> <i>Análisis de datos del proceso de R y C Académico del CEAD Bucaramanga</i> .....	32
<b>Figura 3</b> <i>Código VBA Excel para Correlación de Pearson</i> .....	33
<b>Figura 4</b> <i>Código VBA Excel para Regresión Lineal Simple</i> .....	34
<b>Figura 5</b> <i>Código VBA Excel para Regresión Lineal Simple</i> .....	34
<b>Figura 6</b> <i>Código VBA Excel para Regresión Múltiple</i> .....	35
<b>Figura 7</b> <i>Variables Relevantes para la Deserción Estudiantil</i> .....	37
<b>Figura 8</b> <i>Modelo de entrevista</i> .....	39
<b>Figura 9</b> <i>Reporte inicial en Excel de Alertas tempranas</i> .....	48
<b>Figura 10</b> <i>Aplicación de Macros VBA Excel</i> .....	48
<b>Figura 11</b> <i>Reporte procesado de Alertas tempranas</i> .....	49
<b>Figura 12</b> <i>Resultados de reporte procesado de Alertas tempranas</i> .....	49
<b>Figura 13</b> <i>Resultados de reporte procesado de Alertas tempranas</i> .....	50
<b>Figura 14</b> <i>Visualización resultados en Power BI</i> .....	52
<b>Figura 15</b> <i>Visualización resultados en Power BI</i> .....	52
<b>Figura 16</b> <i>Visualización resultados en Power BI</i> .....	53
<b>Figura 17</b> <i>Visualización resultados en Power BI</i> .....	54

## Introducción

En esta propuesta se plantea el diseño de una herramienta de análisis de datos que integre Microsoft Excel para abordar el problema de la deserción estudiantil en el CEAD Bucaramanga, como una herramienta de apoyo a la toma estratégica de decisiones. Para ello se hace necesario partir de un análisis robusto de los datos obtenidos en el proceso de Registro y Control Académico con el fin de identificar variables relevantes relacionadas la problemática de deserción académica; se sugiere desarrollar un sistema de integración de las herramientas anteriormente mencionadas que favorezca el tratamiento eficiente de volúmenes altos de información y la ejecución de algoritmos de análisis predictivos, que permitan diseñar y aplicar técnicas de minería de datos y análisis estadístico que faciliten la identificación de comportamiento de las variables, así como su correlación con la deserción estudiantil.

Finalmente, para valorar la eficacia y exactitud de la herramienta creada, se propone llevar a cabo un ensayo piloto del modelo de visualización de datos empleando Power BI, con la participación de los usuarios de nivel directivo del CEAD Bucaramanga, con el objetivo de realizar modificaciones conforme a las retroalimentaciones recibidas.

En esta propuesta se pretende desarrollar una metodología de investigación de tipo mixta, en la cual se incluye revisión de literatura, recopilación de datos, análisis exploratorio de datos, desarrollo de macros en Excel, análisis de datos, desarrollo del modelo de visualización de datos, prueba piloto del modelo, evaluación de la herramienta diseñada, documentación y presentación de resultados ante jurado académico.

Dentro de los contenidos del marco teórico se mencionan conceptos importantes como la deserción estudiantil, el análisis de datos en instituciones educativas y la visualización de datos con la herramienta Power BI.

Es importante mencionar que esta investigación se fundamenta en la necesidad de abordar el problema de la deserción estudiantil mediante el análisis de datos y el desarrollo de herramientas tecnológicas que apoyen una toma de decisiones más informada y estratégica en el CEAD Bucaramanga.

## Planteamiento del problema

La deserción estudiantil es un indicador de calidad en las instituciones de educación superior, y es un fenómeno complicado que se ha examinado durante mucho tiempo en el mundo, Lamas Dos Santos Lopes, M. I., Sismeiro Da Silva Pereira, A. P., & Fortunato Vaz, P. M. (2023) plantean que la deserción escolar es un problema que no solo afectan al estudiante, sino también al sistema de educación superior y por ende a la sociedad, y que la falta de conocimiento de los factores que contribuyen a la deserción limita la capacidad de las universidades para desarrollar estrategias que permitan reducir la deserción de sus estudiantes.

En Colombia Sotomonte Castro, J. E., Rodríguez Rodríguez, C. C., Montenegro Marín, C. E., Gaona García, P. A., & Gabriel Castellanos, J. (2016), plantean que si bien es cierto este es un problema latente en la educación superior, en Colombia solo a partir del año 2003 se inician de manera formal los estudios para conocer cuáles son las situaciones que propician el abandono de los estudios.

Este flagelo incluye a la Universidad Nacional Abierta y a Distancia (UNAD) y específicamente en el CEAD Bucaramanga, La falta de comprensión de los factores que favorecen a la deserción restringe la capacidad de la institución para desarrollar estrategias que maximicen la retención estudiantil.

En el CEAD Bucaramanga de la UNAD, el abandono escolar representa un problema importante que impacta tanto a los alumnos como a la institución en general. Pese a que se han puesto en marcha algunas acciones para tratar este problema, su eficacia ha sido restringida debido a la ausencia de un estudio completo y actualizado de los factores que influyen en la decisión de los estudiantes de abandonar sus estudios. El acceso a una amplia gama de

información a través de las bases de datos de la UNAD ofrece una oportunidad inigualable para tratar este problema a través del diseño de un sistema de información que facilite el análisis y entendimiento más profundo de los factores relacionados con la deserción estudiantil en el CEAD Bucaramanga. Sin embargo, actualmente no existe un sistema integral que facilite este análisis de manera eficiente y precisa.

Rodríguez Molano, J. I., Forero Zea, L. D., & Piñeros Reina, Y. F. (2019), plantean que El Machine Learning, surge como una de las técnicas de la inteligencia artificial, en la cual, a través de algoritmos, accede a los datos y los utiliza para aprender y predecir resultados, convirtiéndose en una herramienta fundamental para disminuir los niveles de deserción académica.

Además, Ramírez, P. E., y Grandón, E. E. (2018). realizar un análisis de abandono escolar fundamentado en arboles de decisión, mientras que Urbina-Nájera, A. B. (2021) identificar las variables que afectan el rendimiento académico de estudiantes de posgrado, a través de la implementación de técnicas de Learning Analytics basadas en el modelo Chatti.

Con las experiencias enunciadas anteriormente y otras no presentadas, surge la propuesta de diseñar un sistema de información que, mediante el análisis de las bases de datos de la UNAD, permita identificar y comprender los factores que más inciden en la deserción de los estudiantes en el CEAD Bucaramanga. Este sistema proporcionará a la institución información valiosa y actualizada que servirá de base para el desarrollo e implementación de estrategias efectivas de retención estudiantil, contribuyendo así a mejorar las tasas de graduación y el éxito académico de los estudiantes en el CEAD Bucaramanga y, en última instancia, en la UNAD en su conjunto.

Por lo expuesto anteriormente, se plantea la siguiente pregunta objeto de investigación del estudio:

¿Cuáles son las variables clave que influyen en la retención estudiantil en el CEAD Bucaramanga y cómo pueden ser identificadas y analizadas mediante un sistema de gestión de datos y análisis integral para facilitar la toma oportuna de decisiones estratégicas?

Aunque el proyecto se puede aplicar a la población de la UNAD, pero en esta primera etapa solo se utilizara la población del CEAD Bucaramanga.

## Justificación

El abandono estudiantil es un asunto complicado en la educación universitaria que impacta no solo a los alumnos, sino también a las instituciones de enseñanza y a la sociedad en general, la United Nations Development Programme. (2016), afirma que la deserción en la educación superior agrava las desigualdades sociales y económicas al excluir a ciertos grupos de la población de oportunidades educativas y laborales.

Ya en el campo de la gestión administrativa de la educación superior, Du, K. (2023), plantea en su estudio que, para los administradores en el campo educativo, los datos creados en el proceso son una limitante para lograr una gestión eficaz debido al gran volumen de información, estructura compleja y múltiples características, por ello Madrid Orrego, D. M., Figueroa Royero, L., Berrio Caballero, H., Martínez Choles, A. Y., & Charris Rodríguez, J. A. (2023).esbozan que la minería de datos es una herramienta fundamental en la toma de decisiones que, articulada con inteligencia de negocios, favorecerá a una mejor planeación en el área administrativa, docente y a los procesos psicopedagógicos para evitar la deserción estudiantil y apoyar de manera continua a los estudiantes. De igual forma Vilorio, A., Guliany, J. G., Núñez, W. N., Palma, H. H., & Núñez, L. N. (2020), exponen que este tipo de trabajos ha dado lugar a la denominada Minería de Datos Educativos (EDM) y que esta nueva área de investigación permite conocer mejor a los estudiantes y los contextos en el que se desarrollan.

Casualmente este proyecto se origina en la necesidad de contar con sistema de gestión en el CEAD Bucaramanga orientado a procesar y unificar diversos reportes que se pueden descargar de Analítica de la UNAD, orientados a verificar el estado de aprobación y de ingreso a los cursos.

Para la UNAD en particular, el diseño de un sistema de información que facilite el análisis de los factores que inciden en la deserción estudiantil en el CEAD Bucaramanga es esencial para optimizar la retención estudiantil, mejorar su sostenibilidad holística, promover el éxito académico y robustecer el prestigio institucional. Esta investigación es crucial para abordar un problema que afecta no solo a los estudiantes en el logro de sus propósitos individuales, sino también a la calidad y nombre de la institución educativa, al desarrollo regional de un territorio y un país que lo necesita.

La herramienta de análisis de datos propuesta en este trabajo de grado, que integra el análisis numérico hecho en Microsoft Excel y que podrá visualizarse en Power BI, se presenta como una solución innovadora y eficiente para abordar la problemática de la deserción estudiantil. Esta herramienta permitirá identificar los factores que influyen en la decisión de los estudiantes de abandonar sus estudios, proporcionando a la institución información valiosa para la toma estratégica de decisiones. De igual forma, facilita la caracterización de los estudiantes tanto en datos demográficos como en datos socioeconómicos, pues a través de los filtros de visualización los directores de Centro, los directores de curso, los tutores y demás usuarios, podrán entender el comportamiento de los estudiantes en cada periodo académico, favoreciendo la elaboración de alertas tempranas.

Al diseñar esta herramienta, se busca no solo comprender los motivos detrás de la deserción estudiantil, sino también desarrollar estrategias específicas y personalizadas para intervenir a tiempo y brindar el apoyo necesario a los estudiantes en riesgo. Además, se pretende establecer un sistema de monitoreo continuo que permita evaluar la efectividad de las medidas implementadas y realizar ajustes según sea necesario.

En resumen, este trabajo de grado se justifica por la necesidad de proporcionar al CEAD Bucaramanga una herramienta efectiva y basada en datos para enfrentar el desafío de la deserción estudiantil, con el objetivo último de promover el éxito académico y personal de los estudiantes y fortalecer la calidad educativa de la institución.

Por último, se puede afirmar que el impacto de esta investigación específicamente beneficia no solo a los estudiantes individuales, a la institución en su conjunto, creando un ambiente más favorable para el éxito académico y el desarrollo personal y profesional de los estudiantes y su entorno familiar y social, al sector educativo pues este trabajo servirá de base para el desarrollo de futuros proyectos que permitan seguir avanzando en el propósito de minimizar el flagelo de la deserción en la educación superior no solo en Colombia sino en el mundo, aportando al desarrollo de la humanidad en cumplimiento de los objetivos de desarrollo sostenible.

## **Objetivos**

### **Objetivo General**

Implementar un sistema de gestión de datos y análisis integral, utilizando herramientas de visualización de datos y técnicas de análisis predictivos, que favorezcan la toma oportuna de decisiones estratégicas asociadas a la retención estudiantil en el CEAD Bucaramanga.

### **Objetivos Específicos**

Analizar los datos del proceso de Registro y Control Académico del CEAD Bucaramanga utilizando técnicas estadísticas para identificar patrones y variables relevantes que contribuyan a la deserción estudiantil, con el fin de fundamentar estrategias de intervención.

Aplicar técnicas de minería de datos y análisis estadístico en VBA – Aplicaciones de Excel para evidenciar correlaciones significativas entre las variables y la deserción estudiantil, permitiendo así una mejor comprensión de los factores que afectan el rendimiento académico.

Evaluar la efectividad del modelo de visualización de datos en Power BI mediante la retroalimentación de los usuarios, favoreciendo la toma de decisiones estratégicas en el CEAD Bucaramanga en temas relacionados con la deserción estudiantil.

## **Marco Referencial**

A continuación, se presentan los referentes teóricos que soportan el siguiente trabajo de grado.

### **Marco Teórico**

#### ***Deserción Estudiantil***

El abandono escolar es un fenómeno complicado que ha sido extensamente analizado en el marco de la educación universitaria. Según Tinto (1975), la deserción estudiantil puede ser entendida como el abandono prematuro de los estudios por parte de un estudiante antes de completar su programa académico. Este autor propone el Modelo de Integración y Compromiso (Model of Institutional Departure) que destaca la importancia de factores académicos, sociales e institucionales en la decisión de un estudiante de abandonar sus estudios.

#### ***Análisis de Datos en Educación***

El análisis de datos en el ámbito educativo ha ganado relevancia en la identificación de factores asociados con la deserción estudiantil. Según Díaz & Espinoza (2017), el uso de técnicas de minería de datos y análisis estadístico permite identificar patrones y correlaciones significativas entre variables académicas, sociales y personales y la probabilidad de deserción estudiantil.

#### ***Integración de VBA y Aplicaciones de Excel***

Visual Basic for Applications (VBA) es un lenguaje de programación que permite automatizar tareas y desarrollar aplicaciones personalizadas dentro de las aplicaciones de

Microsoft Office, como Excel. A través de VBA, los usuarios pueden escribir scripts que interactúan con las hojas de cálculo, facilitando la manipulación de datos y la creación de soluciones automatizadas.

Es importante tener en cuenta que Microsoft Excel es una de las herramientas más utilizadas en el ámbito empresarial y académico para la gestión y análisis de datos. Ofrece potentes funcionalidades como tablas dinámicas, gráficos y fórmulas que permiten a los usuarios realizar análisis complejos de manera intuitiva. Sin embargo, a medida que las tareas se vuelven más repetitivas y complejas, la capacidad de Excel puede verse limitada por su naturaleza manual. Es por ello que la integración de VBA en Excel permite a los usuarios automatizar procesos, como la entrada de datos, la generación de informes y la actualización de gráficos. Esto no solo ahorra tiempo, sino que también reduce el riesgo de errores humanos.

De igual manera, VBA permite a los usuarios crear aplicaciones personalizadas dentro de Excel. Esto incluye formularios de usuario que facilitan la entrada de datos, así como funciones personalizadas que pueden extender las capacidades de Excel más allá de las fórmulas estándar. Estas aplicaciones pueden ser diseñadas específicamente para abordar necesidades particulares de análisis de datos, haciendo de Excel una herramienta aún más potente. VBA también permite la interacción con otras aplicaciones de Microsoft Office y fuentes de datos externas, como bases de datos Access o SQL Server. Esto significa que los usuarios pueden integrar datos de diferentes fuentes y realizar análisis más complejos de manera eficiente.

En el contexto de la Ciencia de Datos, la integración de VBA y Excel es especialmente valiosa. Los analistas de datos pueden utilizar VBA para automatizar procesos de limpieza y transformación de datos, así como para realizar análisis preliminares antes de aplicar técnicas

más avanzadas en otras herramientas de programación. Esto facilita un flujo de trabajo más ágil y permite centrarse en el análisis y la interpretación de resultados (Walkenbach, J., 2019)

La integración de VBA con aplicaciones de Excel representa una poderosa combinación que permite a los usuarios maximizar la eficiencia en la manipulación y análisis de datos. A través de la automatización y la personalización, esta integración contribuye a la mejora de procesos y toma de decisiones en diversos ámbitos, incluidos la educación y la gestión académica.

### ***Visualización de Datos con Power BI***

Power BI es una herramienta de visualización de datos desarrollada por Microsoft que permite crear tableros interactivos y paneles de control para analizar datos y compartir información de manera efectiva. Según Rodríguez & Vargas (2019), Power BI ofrece una amplia gama de visualizaciones y funciones de análisis que facilitan la identificación de tendencias, patrones y áreas críticas relacionadas con la deserción estudiantil.

#### **Integración y Flujo de Trabajo en la combinación de Power BI, Excel y VBA**

Utilizando VBA en Excel, se pueden limpiar y transformar datos de manera automatizada, preparando conjuntos de datos listos para el análisis. Acto seguido, se realiza la organización de estos datos que pueden ser exportados a Power BI. Esta última permite crear visualizaciones interactivas, facilitando realizar un análisis en mayor profundidad.

Gracias a la automatización con VBA, se pueden configurar procesos que actualicen periódicamente los datos en Power BI, asegurando que las visualizaciones reflejen información actualizada sin necesidad de intervención manual constante. Esta integración proporciona múltiples beneficios, pues la automatización de tareas repetitivas libera tiempo para que los

analistas se concentren en el análisis y la interpretación de datos. Favorece la precisión de la información, pues con la reducción de los procesos manuales en el manejo de datos a través de la automatización y la vinculación entre herramientas mejora la calidad de la información presentada.

De igual manera, Power BI ofrece capacidades de visualización que permiten a los usuarios explorar datos de manera más interactiva y dinámica, facilitando la identificación de insights clave.

Ahora bien, cabe mencionar que esta integración es especialmente valiosa en contextos donde se requiere tomar decisiones rápidas basadas en datos. En entornos académicos o empresariales, los informes generados pueden ayudar a identificar áreas críticas, tendencias de rendimiento y oportunidades de mejora, lo que permite a los líderes tomar decisiones informadas. (Gaskins, K, s.f.)

En la revisión bibliográfica efectuada, se ha destacado que el abandono escolar es un problema de múltiples aspectos al que se enfrentan las instituciones educativas alrededor del mundo. Se ha transformado en un asunto de creciente relevancia para investigadores y expertos en educación, dado que tiene repercusiones importantes en el desempeño escolar de los alumnos y en la efectividad de los programas de enseñanza. Investigaciones recientes, como la llevada a cabo por Madrid Orrego et al. (2023), han destacado la importancia de la minería de datos para identificar los factores que contribuyen a la deserción estudiantil en universidades públicas. Estudios como este han demostrado que el análisis de grandes conjuntos de datos puede revelar patrones y correlaciones significativas que pueden ser utilizados para desarrollar estrategias efectivas de retención estudiantil.

La aplicación de técnicas de minería de datos en el ámbito educativo ha ganado atención debido a su capacidad para analizar grandes volúmenes de datos y extraer información valiosa para la toma de decisiones. Investigaciones como la realizada por Du (2023) sobre la aplicación de sistemas de big data en la educación han demostrado cómo la fusión de múltiples fuentes de información y el uso de algoritmos de aprendizaje automático pueden mejorar la comprensión de los procesos educativos y ayudar a prevenir la deserción estudiantil. Esta investigación resalta la importancia de adoptar enfoques innovadores basados en datos para abordar los desafíos relacionados con la retención estudiantil y mejorar la experiencia educativa.

Además, la literatura académica ha explorado diversas variables que influyen en la deserción estudiantil, desde factores personales y académicos hasta aspectos sociales y contextuales. Estudios como el de Barragán Moreno y González Támara (2017) han destacado la importancia de la integración social y académica en la retención estudiantil. Esta investigación subraya el análisis del fenómeno de la deserción estudiantil requiere tener en cuenta una variedad extensa de factores y resalta la complejidad del problema, así como la importancia de abordarlo desde múltiples perspectivas y utilizando enfoques interdisciplinarios.

## Marco Legal

Este trabajo de grado se enmarca en varias disposiciones legales que regulan la educación superior en Colombia:

Ley 30 de 1992 establece las normas que rigen la educación superior en Colombia. Promueve la calidad educativa y la permanencia de los estudiantes, proporcionando un marco para la creación de estrategias que aborden la deserción.

Ley 1732 de 2014 es especialmente relevante para el proyecto, ya que establece medidas y estrategias para prevenir la deserción en la educación superior, resaltando la importancia de identificar a los estudiantes en riesgo y proporcionarles apoyo.

En la Ley 1581 de 2012 se establecen las disposiciones generales para la protección de datos personales en Colombia. El proyecto debe garantizar que la recolección, almacenamiento y análisis de datos de los estudiantes cumpla con esta normativa, asegurando la privacidad y la confidencialidad de la información personal.

Dentro de las disposiciones legales que deben ser contempladas en este trabajo de grado se deben contemplar las consignadas en el Código de Ética de la Ingeniería de Sistemas, puesto que el diseño de los sistemas de información deben alinearse a los principios relevantes que se han establecido por el Copnia.

Responsabilidad y Calidad: Al desarrollar un sistema que identificará a estudiantes en riesgo de deserción, es fundamental actuar con responsabilidad y garantizar la calidad en el diseño y funcionamiento del software, así como en el análisis de datos.

Honestidad e Integridad: Los ingenieros involucrados en el proyecto deben asegurarse de que los datos utilizados y las técnicas de análisis sean verídicos y transparentes, evitando cualquier tipo de manipulación que pueda afectar los resultados.

También es importante mencionar que el trabajo de grado cumplirá los lineamientos y políticas internas de la UNAD, puesto que universidad tiene protocolos establecidos para el manejo de datos académicos y la implementación de sistemas de información, los cuales deben ser respetados en el desarrollo del proyecto. Además, la universidad promueve iniciativas que fomentan la retención estudiantil y el uso de tecnologías de información y comunicación.

## Metodología

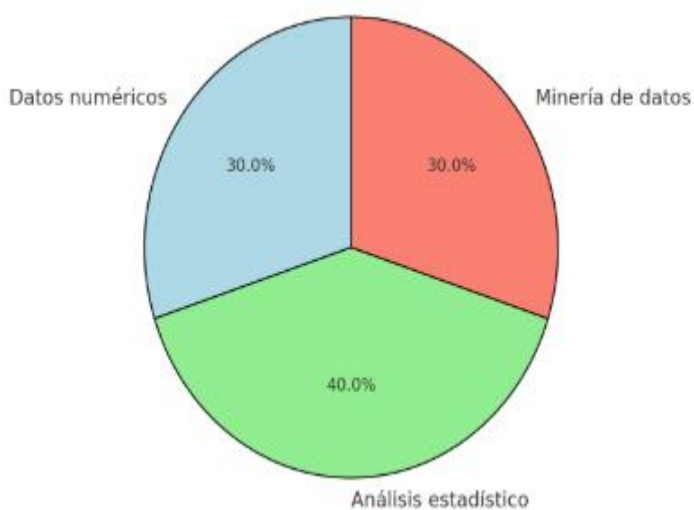
De acuerdo con la metodología de investigación propuesta por Hernández Sampieri, el tipo de investigación que se pretende abordar en esta propuesta es una investigación mixta pues se realiza una integración sistemática de los métodos cuantitativo y cualitativo.

En su mayoría, este trabajo obedece a una investigación cuantitativa; el enfoque se centra en recopilar datos numéricos del proceso de Registro y Control Académico del CEAD Bucaramanga, realizar análisis estadísticos y aplicar técnicas de minería de datos para identificar patrones, correlaciones y variables relevantes relacionadas con la deserción estudiantil.

Fase 1: Análisis de datos del proceso de Registro y Control Académico del CEAD Bucaramanga .

### Figura 1

*Distribución de enfoque en la investigación*



En la figura 1 observamos la distribución de enfoque en el trabajo de grado respecto a los datos que se procesan de los datos de Registro y Control.

## Figura 2

*Esquema de la metodología utilizada*



El esquema de la metodología utilizada en este trabajo la podemos observar en la grafico numero 2, en ella podemos observar las tres etapas de la metodología utilizada.

En cuanto al componente cualitativo que será usado en la fase de evaluación de la herramienta diseñada, se recopilarían opiniones y comentarios de los usuarios sobre la efectividad y utilidad de la herramienta mediante entrevistas o encuestas cualitativas.

En resumen, la investigación sería predominantemente cuantitativa, con elementos cualitativos adicionales para la evaluación y mejora de la herramienta diseñada.

El diseño metodológico para desarrollar la presente propuesta de trabajo de grado con los objetivos planteados anteriormente, se propone una estructura general que enmarcada en las siguientes fases:

## **Fase 1: Diagnóstico**

### ***Etapa 1: Revisión de Literatura y Conceptualización***

Revisión bibliográfica de investigaciones previas sobre deserción estudiantil, herramientas de análisis de datos, integración de, entre otros

Conceptualización de términos clave y teorías relevantes VBA y Microsoft Excel.

### ***Etapa 2: Recopilación de Datos***

Obtención de datos del proceso de Registro y Control Académico del CEAD Bucaramanga.

Aseguramiento de la calidad de los datos y preparación para el análisis.

### ***Etapa 3: Análisis Exploratorio de Datos (EDA)***

Exploración inicial de los datos para identificar patrones, tendencias y posibles problemas.

Visualización de datos utilizando herramientas como gráficos en Excel y bibliotecas VBA.

## **Fase 2: Diseño**

### ***Etapa 1: Desarrollo del Sistema de Integración VBA y Microsoft Excel***

Diseño e implementación de un sistema que permita la integración eficiente de VBA y Microsoft Excel

Desarrollo de scripts y funciones en VBA para manipular datos y ejecutar algoritmos de análisis predictivo.

### ***Etapa 2: Análisis de Datos***

Aplicación de técnicas de minería de datos y análisis estadístico utilizando VBA para identificar variables relevantes y correlaciones con la deserción estudiantil.

Implementación de algoritmos de aprendizaje automático para el análisis predictivo en caso de que la herramienta lo requiera.

### ***Etapa 3: Desarrollo del Modelo de Visualización de Datos***

Diseño y desarrollo de un modelo de visualización de datos utilizando Power BI.

Creación de tableros interactivos que presenten de manera clara y comprensible los resultados del análisis de datos.

## **Fase 3: Evaluación**

### ***Etapa 1: Prueba Piloto del Modelo de Visualización***

Ejecución de una prueba piloto del modelo de visualización con usuarios seleccionados.

Documentar la retroalimentación y sugerencias de usuarios para mejorar el modelo diseñado.

***Etapa 2: Evaluación de la Herramienta Diseñada***

Evaluación de la efectividad y precisión de la herramienta diseñada mediante el análisis de retroalimentación de usuarios.

Identificación de ajustes y mejoras necesarias en la herramienta.

**Fase 4: Cierre**

***Etapa 1: Documentación y Redacción del Trabajo de Grado***

Documentación detallada de todo el proceso metodológico, incluyendo los pasos desarrollados, herramientas utilizadas, análisis y resultados obtenidos.

Redacción del informe final del trabajo de grado.

***Etapa 2: Presentación y Defensa del Trabajo de Grado***

Sustentación escrita y oral de los hallazgos, desarrollo y conclusiones del trabajo de grado ante Jurado académico.

## Resultados

A continuación, se presentan los resultados obtenidos en el desarrollo de los objetivos formulados.

### **Análisis de Datos del Proceso de Registro y Control Académico (RyC) del CEAD**

#### **Bucaramanga**

Para dar solución a este objetivo, se realizó el análisis estadístico sobre los datos históricos de los estudiantes que han formado parte de la UNAD y que han sido registrados en el proceso de RyC. Con ayuda de herramientas estadísticas avanzadas dentro de VBA y complementadas por funciones estadísticas en Excel, se desarrolló a través de varias etapas estructuradas permitiendo extraer factores que influyen en las decisiones del Centro. A continuación, se detallan las principales fases:

#### ***Recolección y Preprocesamiento de los Datos***

En primera instancia, se procedió a reunir los datos del proceso de Registro y Control Académico de los estudiantes en la UNAD. Este conjunto de datos incluye información extraída de los aplicativos que contienen los repositorios de matrícula del periodo vigente (RyC), datos de Core y/o Analítica, los cuales proveen la Dirección de Centro, Consejería Académica. De aquí se puede extraer lo siguiente

**Datos Académicos.** calificaciones de los estudiantes, ingresos a cursos, actividades pendientes de los estudiantes en cada curso, promedio académico.

**Datos Demográficos.** edad, sexo, ubicación geográfica, convenio de matrícula (incluye gratuidad), nivel socioeconómico, caracterización del estudiante.

**Datos de Matrícula.** como el año de ingreso, el número de periodos que el estudiante ha cursado, número de matrículas por curso.

Una vez recopilados estos datos y consolidados en una hoja de cálculo, se procedió a realizar un preprocesamiento de la información; este procedimiento se realizó de la siguiente manera:

**Limpieza de Datos.** este paso se hace con el fin de asegurar que los datos están completos y libres de errores, es decir, no existen dentro de la data a procesare valores nulos o registros duplicados.

**Transformación.** normalización de las variables, en este punto se identifica si es necesario, formular o crear algunas variables que pueden ser de utilidad en el análisis que se hará posteriormente, en algunos casos, variables como la tasa de aprobaciones, promedio de calificaciones, números cursos en los que ha estado ausente, entre otros.

### ***Análisis Estadístico Descriptivo***

Con el fin de abordar un análisis estadístico descriptivo para identificar las principales características del conjunto de datos, se hace necesario utilizar herramientas de VBA y Power BI que permiten considerar métricas como:

**Frecuencias.** ¿cuántos estudiantes desertan cada período? ¿Cuántos días lleva sin ingresar al curso?

**Promedios y medias.** ¿qué tan alto o bajo es el rendimiento académico promedio de los estudiantes que desertan frente a los que permanecen?

**Distribución de variables.** ¿cómo se distribuyen las calificaciones de los estudiantes que abandonan?

**Cruces de Variables.** evaluar la correlación entre diversas variables (como la calificación y la permanencia en la UNAD).

### ***Identificación de Patrones y Variables Relevantes***

Con los datos preparados y los análisis descriptivos realizados, las visualizaciones de patrones y variables relevantes se identifican de mejor forma, es así como se empezó a aterrizar estos factores asociados a la deserción estudiantil. Aquí se recurre a técnicas de exploración más avanzada:

**Análisis de Correlación.** en este punto se realiza la revisión si existe una correlación negativa entre la cantidad de cursos aprobadas y la deserción; de igual forma si existe alguna correlación entre la edad o el nivel socioeconómico tienen alguna relación con el abandono de los procesos de formación

**Análisis de Regresión.** esta es una de las herramientas que se contempla usar para determinar qué variables predicen de manera significativa la deserción estudiantil; la regresión logística puede favorecer el modelado de la probabilidad que tiene un estudiante de abandonar su proceso según sus características académicas y demográficas.

**Análisis de Clasificación.** en este análisis se usará la técnica árboles de decisión para identificar características específicas que discriminan entre los estudiantes que desertan y los que no; la tendencia es que en los estudiantes que tienen calificaciones inferiores a cierto umbral tienen más probabilidades de abandonar.

### ***Variables Relevantes para la Deserción Estudiantil***

A partir del análisis de los datos, se identifican las siguientes variables relevantes para el análisis de la deserción estudiantil:

**Rendimiento Académico.** los estudiantes con bajo rendimiento académico, es decir, aquellos con un número elevado de cursos reprobados o con calificaciones por debajo de un promedio inferior a 3.2, tienen más probabilidades de abandonar sus procesos.

**Edad.** en algunos casos, los estudiantes de edad mayor (que tienen que equilibrar estudios con otras responsabilidades) pueden estar más expuestos a abandonar sus cursos.

**Situación Laboral.** de acuerdo con algunos datos históricos analizados en la estrategia UNAD Siempre Contigo, según manifestaron los estudiantes que trabajan mientras estudian, estos tienen una tasa de deserción más alta, ya que el trabajo puede interferir con el tiempo de estudio.

**Nivel Socioeconómico.** los estudiantes con menores recursos económicos pueden tener dificultades para continuar con sus estudios, especialmente si enfrentan problemas como el pago de matrículas o la falta de accesibilidad a elementos como conectividad (internet y PC).

**Promedio de Créditos Aprobados.** los estudiantes que no avanzan en la carga de créditos planificada de acuerdo con el plan de estudios o que repiten varios cursos tienen más probabilidad de desertar período a período.

**Distancia de Residencia.** si los estudiantes viven lejos de las áreas rurales, con baja cobertura a redes que permitan acceder al campus están expuestos a la deserción.

Una vez identificado un conjunto de variables relevantes, a continuación, se presenta un resumen de características más importantes de los estudiantes que abandonan en comparación con aquellos que permanecen.

Figura 3

*Análisis de datos del proceso de Registro y Control Académico del CEAD Bucaramanga*



Tabla 1

*Características de los estudiantes*

Estudiantes	
Abandonan	Permanecen
Bajo rendimiento académico	Alto rendimiento académico
Recursos económicos bajos	Recursos económicos medios
Zona Rural	Zona Urbana

## Aplicar Técnicas de Minería de Datos y Análisis Estadístico en VBA – Aplicaciones de Excel para Evidenciar Correlaciones Significativas Entre las Variables y la Deserción Estudiantil, Permitiendo Así una Mejor Comprensión de los Factores que Afectan el Rendimiento Académico

Para dar cumplimiento al segundo objetivo específico, se han utilizado varias técnicas que permitirán identificar y entender las relaciones entre las variables. A continuación, se detallan algunas de las técnicas que han sido usadas en VBA:

### Técnicas de Minería de Datos y Análisis Estadístico en VBA/Excel

#### *Análisis de Correlación*

**Correlación de Pearson.** para calcular la relación de las variables continuas y la relación lineal, se utilizó esta técnica para medir la fuerza y dirección de la relación entre dos variables.

La función se utilizó en VBA se utilizó el siguiente código:

#### Figura 4

#### *Código VBA Excel para Correlación de Pearson*

```
Sub CalcularCorrelacionPearson()
    Dim Rangol As Range
    Dim Rango2 As Range
    Dim Correlacion As Double

    ' Selección de rangos de datos
    On Error Resume Next
    Set Rangol = Application.InputBox("Seleccione el primer rango de datos", Type:=8)
    If Rangol Is Nothing Then Exit Sub
    Set Rango2 = Application.InputBox("Seleccione el segundo rango de datos", Type:=8)
    If Rango2 Is Nothing Then Exit Sub
    On Error GoTo 0

    ' Validación de rangos con igual número de filas
    If Rangol.Rows.Count <> Rango2.Rows.Count Then
        MsgBox "Los rangos deben tener el mismo número de filas.", vbExclamation, "Error"
        Exit Sub
    End If

    ' Cálculo de la correlación de Pearson
    On Error Resume Next
    Correlacion = WorksheetFunction.Correl(Rangol, Rango2)
    If Err.Number <> 0 Then
        MsgBox "No se pudo calcular la correlación. Verifique los datos.", vbCritical, "Error"
        Exit Sub
    End If
    On Error GoTo 0

    ' Mostrar el resultado
    MsgBox "La correlación de Pearson es: " & Format(Correlacion, "0.0000"), vbInformation, "Resultado"
End Sub
```

## *Regresión Lineal y Regresión Múltiple*

**Regresión Lineal Simple.** en este caso, para examinar una variable independiente como el rendimiento académico que afecta a una variable como lo es la deserción, en VBA, se utilizó el siguiente código:

### **Figura 5**

#### *Código VBA Excel para Regresión Lineal Simple*

```

Sub RegresionLinealSimple()
    Dim RangoX As Range
    Dim RangoY As Range
    Dim Pendiente As Double
    Dim Intercepto As Double
    Dim RSquared As Double
    Dim i As Integer
    Dim PrediccionY As Double
    Dim Residual As Double
    Dim SSresidual As Double
    Dim SStotal As Double
    Dim MediaY As Double

    ' Selección de rangos de datos
    On Error Resume Next
    Set RangoX = Application.InputBox("Seleccione el rango de datos para X", Type:=8)
    If RangoX Is Nothing Then Exit Sub
    Set RangoY = Application.InputBox("Seleccione el rango de datos para Y", Type:=8)
    If RangoY Is Nothing Then Exit Sub
    On Error GoTo 0

    ' Validación de rangos con igual número de filas
    If RangoX.Rows.Count <> RangoY.Rows.Count Then
        MsgBox "Los rangos deben tener el mismo número de filas.", vbExclamation, "Error"
        Exit Sub
    End If

    ' Cálculo de pendiente y el intercepto
    Pendiente = WorksheetFunction.Slope(RangoY, RangoX)
    Intercepto = WorksheetFunction.Intercept(RangoY, RangoX)

    ' Calcular el coeficiente de determinación R^2
    SSresidual = 0
    SStotal = 0
    MediaY = WorksheetFunction.Average(RangoY)

```

**Regresión Múltiple.** para analizar variables múltiples predictoras, como rendimiento académico, edad, situación laboral, se utilizó la regresión múltiple pues a través de esta se puede

modelar la deserción en función de varias variables a la vez. En VBA, el código anterior se puede adaptar a una regresión con múltiples variables.

## Figura 6

### *Código VBA Excel para Regresión Múltiple*

```

Sub RegresionLinealMultiple()
    Dim RangoX As Range
    Dim RangoY As Range
    Dim Resultados As Variant
    Dim Coeficientes() As Double
    Dim Intercepto As Double
    Dim RSquared As Double
    Dim i As Integer
    Dim Salida As String

    ' Selección de rangos de datos
    On Error Resume Next
    Set RangoY = Application.InputBox("Seleccione el rango de datos para Y (dependiente)", Type:=8)
    If RangoY Is Nothing Then Exit Sub
    Set RangoX = Application.InputBox("Seleccione el rango de datos para X (independientes)", Type:=8)
    If RangoX Is Nothing Then Exit Sub
    On Error GoTo 0

    ' Validación de filas en los rangos
    If RangoX.Rows.Count <> RangoY.Rows.Count Then
        MsgBox "Los rangos deben tener el mismo número de filas.", vbExclamation, "Error"
        Exit Sub
    End If

    ' Calcular los coeficientes de regresión utilizando la función LINEST
    On Error Resume Next
    Resultados = WorksheetFunction.LinEst(RangoY, RangoX, True, True)
    If Err.Number <> 0 Then
        MsgBox "No se pudo calcular la regresión. Verifique los datos.", vbCritical, "Error"
        Exit Sub
    End If
    On Error GoTo 0

    ' Extraer los resultados
    ReDim Coeficientes(1 To UBound(Resultados, 2))
    For i = 1 To UBound(Resultados, 2)

```

### *Árboles de Decisión (Clasificación)*

Esta herramienta se utilizó en VBA con componentes como Solver, para implementar el algoritmo de árbol de decisión; a través de este algoritmo se clasifican a los estudiantes de acuerdo al riesgo de deserción basándose en características como rendimiento académico, edad, situación socioeconómica, entre otros, asignando una clasificación de alto, medio o bajo riesgo de deserción, lo cual favorece la intervención temprana.

### ***Visualización de Datos (Power BI y Excel)***

Esta técnica no es propiamente una técnica de minería de datos en sí, sin embargo, da una utilidad importante en la visualización de los resultados en cada una de las etapas para comprender las correlaciones y relaciones entre las variables. Power BI permite crear dashboards interactivos y visualizaciones más avanzadas.

En Power BI, se implementarán lo siguientes tipos de gráficos para visualizar la información:

Gráficos de dispersión para ver la relación entre dos variables como rendimiento académico vs. Deserción.

Diagramas de barras y líneas para mostrar tendencias temporales y la relación entre las variables, así como mostrar tendencias temporales de deserción a lo largo de los semestres o años.

Mapas de calor para mostrar patrones de riesgo de deserción por áreas geográficas o programas académicos.

Indicadores KPI (Key Performance Indicators) para mostrar las métricas clave relacionadas con la retención estudiantil, como la tasa de deserción, la tasa de aprobación, entre otros.

### ***Interactividad***

Dentro del modelo se contempla la creación de filtros dinámicos que permitan a los usuarios segmentar la información por carrera, período académico, ubicación.

Segmentación por cohortes para ver cómo diferentes grupos de estudiantes, es decir, comparación de nuevos estudiantes vs. Antiguos, para ver si tienen tasas de deserción distintas.

Alertas tempranas que identifiquen a los estudiantes con mayor riesgo de deserción, usando colores o indicadores visuales claros.

### ***Implementación de Funciones de Power BI***

Dentro de Power BI, se contempla la inclusión de varias funciones avanzadas para enriquecer el dashboard, tales como:

**Power Query:** Para importar, limpiar y transformar los datos desde las diferentes fuentes como Excel, bases de datos, entre otras.

**DAX (Data Analysis Expressions):** Para crear medidas y columnas calculadas, como la tasa de deserción, la probabilidad de deserción y otros indicadores clave.

**Dashboards de varias páginas:** se crearán diferentes vistas que permitan a los usuarios explorar diversos aspectos de los datos, como el rendimiento académico, las características demográficas, y los patrones temporales de deserción.

### **Figura 7**

#### *Variables Relevantes para la Deserción Estudiantil*



## **Evaluar la Efectividad del Modelo de Visualización de Datos en Power BI Mediante la Retroalimentación de los Usuarios, Favoreciendo la Toma de Decisiones Estratégicas en el CEAD Bucaramanga en Temas Relacionados con la Deserción Estudiantil**

Para cumplir con el objetivo específico 3 es importante definir el proceso de evaluación que permita medir cómo las herramientas de visualización impactan en la toma de decisiones estratégicas relacionadas con la deserción estudiantil en el CEAD Bucaramanga. Este sentido se hace necesario involucrar tanto la creación del modelo de visualización en Power BI, como la recolección y análisis de retroalimentación de los usuarios para ajustar y mejorar el sistema.

Aquí te doy una guía detallada sobre cómo puedes abordar este objetivo.

### ***Evaluación de la Efectividad del Modelo de Visualización de Datos en Power BI***

Para evaluar la efectividad del modelo de visualización de datos, posterior al diseño interactivo del dashboard en Power BI de acuerdo con las necesidades de los usuarios, es decir, a la alta dirección del CEAD Bucaramanga. El modelo debe abordar la retroalimentación de los usuarios que interactúan con el dashboard para evaluar su efectividad. En este caso los usuarios que se contemplan se encuentran el director Zonal del CEAD, Líderes Zonales de Escuela, consejeros académicos y personal administrativo que esté involucrado en el seguimiento de la retención estudiantil.

### ***Métodos para Recoger la Retroalimentación***

Dentro de los métodos utilizados para realizar la recolección de información se encuentran los siguientes:

**Encuestas a Usuarios.** aplicación de una encuesta previamente diseñada evalúe la utilidad y la efectividad del dashboard, que permita identificar si el dashboard facilita la identificación de estudiantes en riesgo de deserción, si la herramienta favorece o no la

visualización de datos y su entendimiento, la utilidad de las alertas tempranas y los indicadores KPI en la toma de decisiones y si el dashboard permite realizar un seguimiento eficaz de las acciones correctivas.

## Figura 8

### *Modelo de entrevista*

#### **1. Información General**

1. **Cargo del encuestado:**
  - Docente
  - Coordinador académico
  - Director de programa
  - Otro: \_\_\_\_\_
2. **Nivel de experiencia en el uso de dashboards (Seleccione una):**
  - Ninguno
  - Básico
  - Intermedio
  - Avanzado

#### **2. Utilidad y efectividad general del dashboard**

1. En una escala de 1 a 5, donde 1 es "Nada efectivo" y 5 es "Muy efectivo", ¿cómo calificaría la utilidad del dashboard para identificar estudiantes en riesgo de deserción?  
(1 - Nada efectivo, 5 - Muy efectivo)
  - 1
  - 2
  - 3
  - 4
  - 5
2. ¿Qué tan fácil es para usted comprender la información presentada en el dashboard?  
(1 - Muy difícil, 5 - Muy fácil)
  - 1
  - 2
  - 3
  - 4
  - 5
3. ¿El dashboard incluye toda la información que necesita para tomar decisiones sobre estudiantes en riesgo?
  - Sí
  - No (Por favor, especifique qué información adicional sería útil): \_\_\_\_\_

**Entrevistas a usuarios.** Realiza entrevistas o sesiones de grupo con los usuarios clave para obtener comentarios más profundos sobre la experiencia del usuario, para identificar inconvenientes en la navegación por el dashboard, posibles mejoras que los usuarios consideren necesarias y casos específicos en los que las visualizaciones ayudaron o no ayudaron en la toma de decisiones.

**Evaluación de Uso.** Se realizará un seguimiento de la interacción con el dashboard, para identificar cuántas veces se consulta el dashboard, qué funcionalidades son las más utilizadas.

### **Indicadores Clave de la Evaluación de Uso**

Los resultados se presentarán basados en los siguientes indicadores:

Consultas al Dashboard

Total de consultas realizadas: Número total de veces que el dashboard fue accedido durante el período de evaluación.

Promedio diario de consultas: Consultas totales divididas por el número de días evaluados.

Picos de uso: Días y horas con mayor actividad.

Funcionalidades Más Utilizadas

Ranking de funcionalidades: Listado de las funcionalidades más consultadas, ordenadas por frecuencia de uso.

Distribución porcentual: Gráfica de pastel o barras que muestre el porcentaje de uso de cada funcionalidad.

c. Frecuencia de Acceso por Usuario

Usuarios únicos: Número de personas diferentes que accedieron al dashboard.

Frecuencia de acceso: Promedio de accesos por usuario.

Segmentación por rol: Análisis del uso según el rol del usuario (docente, coordinador académico, etc.).

#### d. Tendencias de Uso

Tendencia semanal: Gráfica de líneas que muestre la evolución de las consultas a lo largo del tiempo.

Horas pico: Identificación de los momentos del día con mayor actividad.

**Indicadores de Impacto.** Ahora bien, se hará un comparativo de la toma de decisiones tanto de las decisiones basadas en los datos visualizados antes del dashboard y después del dashboard, con el fin de revisar que tanto impacto ha tenido la retención estudiantil, identificando si hubo o no disminución en la tasa de deserción en grupos de estudiantes identificados como de alto riesgo y si hubo mejora en las intervenciones preventivas.

### Indicadores Clave para Evaluación Comparativa

#### a. Tasa de Deserción Estudiantil

Fórmula:

$$\text{Tasa de desercion} = \left( \frac{\text{Numero de estudiantes que abandoraon}}{\text{Totalde estudinales inscritos}} * 100 \right)$$

#### b. Identificación de Estudiantes en Riesgo

Porcentaje de estudiantes identificados como en riesgo:

*Porcentaje de riesgo de deserción*

$$= \left( \frac{\text{Numero de estudiantes identificados en riesgo}}{\text{Totalde estudinales}} * 100 \right)$$

c. Efectividad de las Intervenciones Preventivas

Métrica: Número de estudiantes en riesgo que fueron retenidos tras las intervenciones preventivas.

Indicador: Porcentaje de retención en el grupo de riesgo.

$$Retencion\ en\ riesgo = \left( \frac{Estudiantes\ retenidos\ en\ riesgo}{Estudinales\ identificados\ en\ riesgo} * 100 \right)$$

d. Tiempos de Respuesta

Antes: Promedio de tiempo en identificar estudiantes en riesgo e implementar acciones.

Después: Reducción en el tiempo promedio de identificación y respuesta gracias al dashboard.

e. Calidad de las Decisiones

Indicador subjetivo: Encuestas a los responsables para evaluar la percepción de mejora en la calidad de las decisiones basadas en los datos del dashboard.

***Análisis de Retroalimentación y Ajustes***

Con la retroalimentación de los usuarios se procederá a la identificación de áreas de mejora, incluyendo si es el caso, nuevos filtros, ajustes de diseño y la interactividad del dashboard.

En cuanto a las visualizaciones, si es el caso, se modificarán las visualizaciones para que sean más claras y útiles, con el fin de favorecer la efectividad de los gráficos y la interpretación de la información.

Si las métricas de deserción o las alertas tempranas no son lo suficientemente precisas, se revisará tanto el código como las variables involucradas para el cálculo de estas.

De igual manera, se asegura que el modelo de Power BI sea sostenible y escalable, es decir, que el modelo pueda mantenerse actualizado con el paso del tiempo y con la obtención de

nuevos datos. A futuro, se debe evaluar si el modelo puede ser adaptado fácilmente a otras áreas o Centros de la ZCORI.

## Conclusiones

El uso de herramientas de visualización de datos como Power BI ha demostrado ser eficaz en la toma de decisiones estratégicas en el CEAD Bucaramanga. Al proporcionar información clara y accesible sobre las variables que afectan la retención estudiantil, tanto los directivos como los docentes pueden hacer uso de dicha información para tomar decisiones que eviten en la mayor medida posible la deserción.

Las técnicas estadísticas, como la regresión lineal y la minería de datos, permiten la identificación de patrones significativos que influyen en la deserción estudiantil. Estas herramientas ayudan a predecir el comportamiento de los estudiantes y proporcionan una base sólida para la implementación de estrategias preventivas.

Los análisis han revelado una fuerte correlación entre ciertos factores académicos como el puntaje en actividades académicas, el ingreso a cursos, y la deserción estudiantil. Estos hallazgos son fundamentales para orientar las acciones de intervención y mejorar las tasas de retención.

La implementación de alertas tempranas basado en los datos recolectados y analizados permite que sean una herramienta clave para la detección temprana de estudiantes en riesgo de deserción, favoreciendo la intervención oportuna y la aplicación de estrategias que aumentan las probabilidades de retención.

Las estrategias de intervención basadas en los análisis de datos constituyen un impacto positivo en la retención estudiantil, pues la personalización de las intervenciones, basada en los

resultados de los modelos predictivos, ha favorece la mejora del rendimiento académico y el compromiso de los estudiantes.

Aunque los resultados obtenidos hasta el momento son prometedores, es necesario continuar con la mejora y ajuste del modelo de análisis de datos. Esto incluye la actualización continua de los datos, la inclusión de nuevas variables y la capacitación constante de los usuarios del sistema para que puedan aprovechar al máximo las herramientas de visualización y análisis.

La implementación de un sistema de gestión de datos y análisis predictivo resulta ser viable desde el punto de vista técnico y organizacional. No obstante, es esencial asegurar la sostenibilidad del proyecto a largo plazo, tanto en términos de infraestructura tecnológica como de capacitación del personal.

La creación de un informe sobre la deserción de estudiantes de la UNAD en Power BI debe enfocarse en identificar las causas, tendencias y estadísticas relacionadas con este fenómeno.

## **Pasos para construir el informe**

### **1. Recopilación de Datos**

Datos internos: Recoge información de la UNAD sobre matrículas, desertores, índices de retención, programas académicos y modalidades.

Variables clave: Edad, género, programa académico, semestre, modalidad (virtual, presencial, dual), promedio académico, estado socioeconómico.

### **2. Preparación de Datos**

Limpia los datos para eliminar duplicados y valores inconsistentes.

Normaliza y transforma las columnas necesarias para el análisis (e.g., codificación de categorías, fechas uniformes).

Relaciona tablas relevantes como estudiantes, programas académicos, y periodos académicos.

### 3. **Diseño de Indicadores Clave (KPIs)**

Tasa de deserción (porcentaje de desertores en relación con los inscritos).

Segmentación por programa, modalidad y región.

Promedio académico de desertores.

Evolución de la deserción por semestre.

### 4. **Construcción del Informe en Power BI**

Página de Inicio:

Resumen general de la deserción en gráficos interactivos.

KPIs principales como la tasa general de deserción y su evolución.

Segmentación por Programa Académico:

Comparativa de deserción por programa y facultad.

Top 5 programas con mayores tasas de deserción.

Análisis Regional:

Mapa de calor mostrando la deserción por región.

Factores socioeconómicos que afectan a cada zona.

Análisis Temporal:

Tendencias de deserción por periodo y año.

Impacto de políticas o cambios académicos en las tasas.

Factores Asociados:

Gráficos de dispersión para mostrar relación entre promedio académico y deserción.

Análisis de encuestas sobre motivos de deserción.

Recomendaciones:

Página con hallazgos clave y sugerencias de intervención.

## 5. Visualizaciones Clave

Barras y columnas apiladas para comparar modalidades y programas.

Líneas de tendencia para evolución temporal.

Mapas interactivos para análisis regional.

Gráficos de dispersión para correlaciones.

## Recomendaciones para el Informe

Interactividad: Usa filtros para que los usuarios puedan explorar por modalidad, región, género, etc.

Análisis Predictivo: Incorpora un modelo para prever tasas de deserción futuras basado en patrones actuales.

Narrativa Visual: Acompaña cada página con explicaciones para facilitar la comprensión.

Exportación: El informe pueda exportarse a PDF o PowerPoint para ser compartido.

**Figura 9**

*Reporte inicial en Excel de Alertas tempranas (se ocultan datos personales de estudiantes)*

UNIVERSIDAD NACIONAL ABIERTA Y A DISTANCIA - UNAD																
Avance de Calificaciones [SE 4.0]																
Periodo: 2024 I PERIODO 15-04 (1701)																
Zona: Zona CentroOriente																
Documento	Nombres y ap. Correo Institución	Zona	Centro	Programa	Número de en Escuela	Código del co Curso	Número de la Grupo en el c	Créditos del c	Estado de ma Actividades	Puntaje estuad	Puntaje califi	Resultado [p				
67494755	JANA RAEL B...	Zona Centro	MTaga	MAESTRÍA EN	8	ECCEDU	522010	PERSPECTIVA	1	1	1	Matriculado	5	25	25	En Desarroll
83227940	CARLOS H...	Zona Centro	MTaga	MAESTRÍA EN	11	ECCEDU	519009	SEMINARIO D	1	8	1	Matriculado	4	24	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	MUSICA	5	ECSAH	410004	LECTO-ESCR	1	2	6	Matriculado	8	74	75	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	MUSICA	5	ECSAH	410008	REPERTORIO	1	5	1	Matriculado	5	25	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	MUSICA	5	ECSAH	410017	TEORIA DE LA	1	3	1	Matriculado	5	20	25	En Desarroll
13120440	CARLOS H...	Zona Centro	DuaTe	TECNOLOGIA	8	ECSA	40003	ÉTICA Y CIUD.	1	1532	1	Matriculado	5	0	[Sin dato]	En Desarroll
13120440	CARLOS H...	Zona Centro	DuaTe	TECNOLOGIA	8	ECSA	322609	SERVICIO AL E	1	67	1	Matriculado	5	23	25	En Desarroll
13120440	CARLOS H...	Zona Centro	DuaTe	TECNOLOGIA	8	ECSA	152004	DIPLOMADO I	1	24	10	Matriculado	6	20	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Pamplona	DERECHO	8	ECP	251006	DERECHO CO	1	13	2	Matriculado	4	25	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Pamplona	DERECHO	8	ECP	252517819	DERECHO CO	1	24	1	Matriculado	5	25	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Pamplona	DERECHO	8	ECP	252517820	DERECHO PEN	1	4	1	Matriculado	5	90	125	En Desarroll
13120440	CARLOS H...	Zona Centro	Pamplona	DERECHO	8	ECP	252517821	ARGUMENTA	1	6	1	Matriculado	5	25	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Pamplona	DERECHO	8	ECP	252517822	SEGURIDAD S	1	17	2	Matriculado	4	20	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	MAESTRÍA EN	10	ECCEDU	519011	SEMINARIO D	1	4	1	Matriculado	5	25	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	MAESTRÍA EN	10	ECCEDU	522006	PERSPECTIVA	1	2	1	Matriculado	5	23	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	MAESTRÍA EN	10	ECCEDU	522008	DISEÑO DE AF	1	3	1	Matriculado	5	23	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	ESPECIALIZAC	9	ECBTI	203008067	MÁCHINE LLJ	1	13	2	Matriculado	5	71	75	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	ESPECIALIZAC	9	ECBTI	203008068	GESTIÓN Y AL	1	13	2	Matriculado	5	121	125	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	ESPECIALIZAC	9	ECBTI	203008069	PROGRAMACI	1	13	1	Matriculado	5	127	140	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	ESPECIALIZAC	9	ECBTI	203008070	PROYECTO DE	1	10	2	Matriculado	4	18	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	ESPECIALIZAC	9	ECBTI	203008077	BIG DATA INT	1	5	2	Matriculado	4	19	25	En Desarroll
13120440	CARLOS H...	Zona Centro	Bucaramanga	INGENIERIA E	6	ECAPMA	80008	PEDAGOGÍA P	1	168	2	Matriculado	4	20	25	En Desarroll

Figura 10

*Aplicación de Macros VBA Excel para el procesamiento de datos*

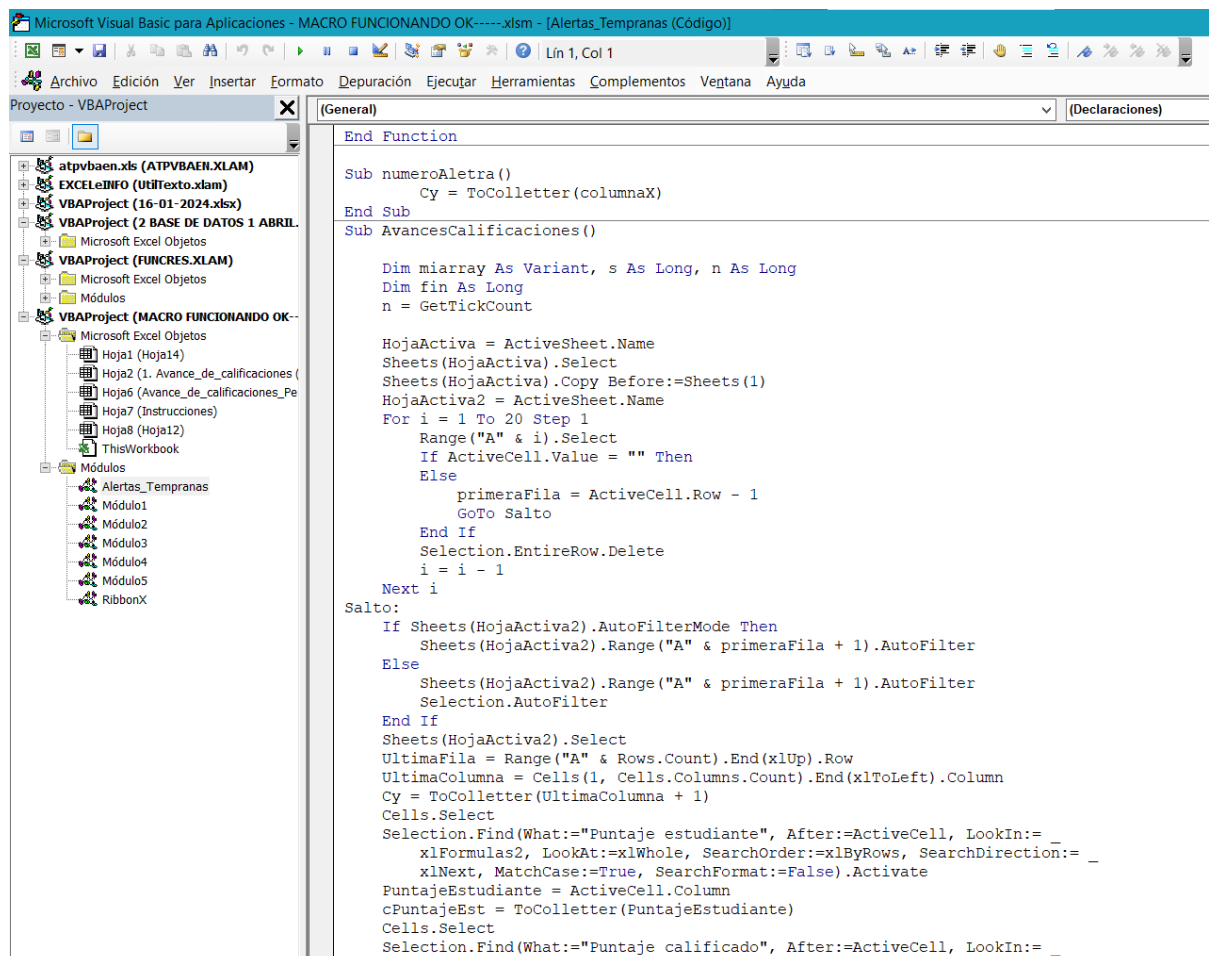


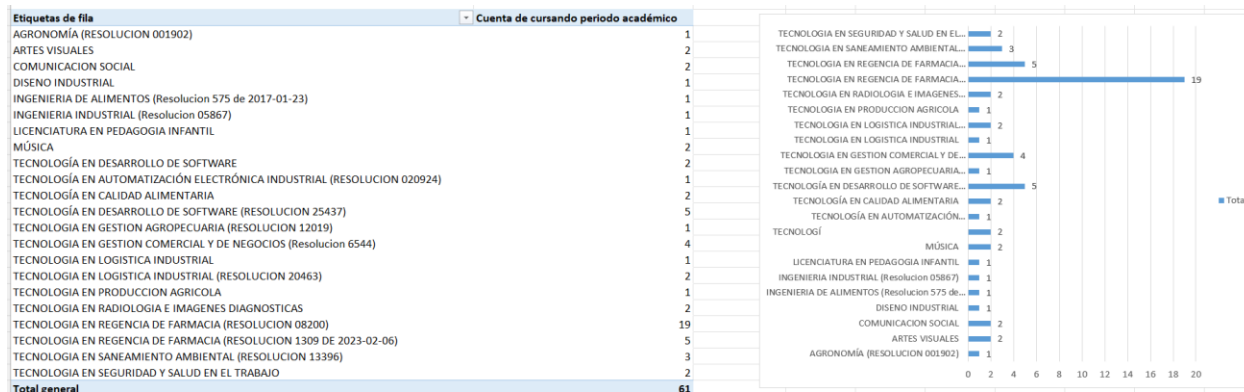
Figura 11

Reporte procesado de Alertas tempranas con colores en los cursos no aprobados

Convenio aplicado	Cantidad de Cursos con problemas	Llamar	OK	Curso
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2019)	9	9	0	PROYECTO D PROTOCOLO SOCIOLOGIA ANTROPOLO GESTION AM INVESTIGACI SERVICIO AL SISTEMAS Y SLOGISTICA COMERCIAL
UNAD: GENERACION E VIGENCIA 2022 I (NUEVOS)	9	9	0	GUITARRA PROTOCOLO ACONDICION COMUNIDAD PSICOPATOLOGIA SOCIAL ACCION PSIC ETNOPSICOL INTELIGENCIA Y CREATIV
UNAD: GENERACION E VIGENCIA 2022 I (NUEVOS)	9	9	0	ETICA Y CIUC PROTOCOLO DISCAPACIDAD INGLÉS I INTRODUCCI INGLÉS O BIOLOGIA CE HERRAMIENTAS FUNDAMENTOS DE MATE
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	8	8	0	ANTROPOLO PSICOLOGIA METODOLOG DISEÑO DE P POLITICA AG ESTADISTICA MANEJO DE CULTIVOS DE CLIMA CALIDO
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2019)	8	8	0	ETICA Y CIUC PEDAGOGIA CULTURA PO INGLÉS COP ESCENARIOS ENGLISH IV LEGISLACION LENGUA MATERNA II
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	8	8	0	PROYECTO D METODOS DE ECONOMIA SALUDOSARI FINANZAS GEOPOLITICA PROSPECTIV LOGISTICA COMERCIAL
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 I)	8	8	0	FUNDAMENT PENSAMIENT TEORIA DEL FUNDAMENT HISTORIA Y T SISTEMAS PO HISTORIA Y T ACONDICIONAMIENTO FISICO Y BIENE
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 I)	8	8	0	ETICA Y CIUC COMPETENC FUNDAMENT DIDACTICA INTRODUCCI LENGUA MAI ENGLISH IV INTRODUCCION TO LINGUISTICS
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 I)	8	8	0	PROTOCOLO INTRODUCCI PRIMEROS AI FUNDAMENT BIOLOGIA CE HERRAMIENTAS ADMINISTRACATEDRA UNADISTA (E_LEARNING)
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 I)	8	8	0	ETICA Y CIUC INGLÉS I INTRODUCCI METODOLOG FARMACOLOG TOXICOLOGI LEGISLACION ATENCION FARMACEUTICA
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 II)	8	8	0	METODOS P FORMULACION PROCESOS DE EMPRENDIM PRODUCTIVAMODULOS Y LOGISTICA Y INGENIERIA AMBIENTAL
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 II)	8	8	0	BIOLOGIA COMPETENC CULTURA PO INGLÉS O BIOLOGIA QUIMICA GEN INTRODUCCI PASTOS Y FERRAJES
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 II)	8	8	0	PEDAGOGIA ADMINISTRAC EDUCACION ESCENARIOS ENGLISH VII ENGLISH PHC MATERIALS E LEGISLACION EDUCATIVA
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 II)	8	8	0	ETICA Y CIUC COMPETENC TEATRO FUNDAMENT INVESTIGACI INTRODUCCI ENGLISH III INTRODUCCION TO LINGUISTICS
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2021 I)	8	8	0	ENGLISH COP TEACHING ET ENGLISH V ENGLISH PHC MATERIALS DIDACTICS O ETICA DE LA LENGUA MATERNA II
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2021 I)	8	8	0	ETICA Y CIUC GUITARRA GESTION DE EXPEDIENTE PROTECCION INGLÉS A1 DISEÑO DE P DERECHO ADMINISTRATIVO
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2021 I)	8	8	0	EDUCACION ESCENARIOS MATEMATICO ENGLISH III FOREIGN LAR MATERIALS E DIDACTICS O LEGISLACION EDUCATIVA
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	8	8	0	PROYECTO D EDUCACIONES FISICA ELECT METODOS DE ESTADISTICA ESTADISTICA Y PRESTACION INGLÉS B1
UNAD: GENERACION E VIGENCIA 2022 II (NUEVOS)	8	8	0	SERVICIO AL EXCEL AVAN PROYECTO ETNOGRAFIA FORMA Y MC HISTORIA DE INGLÉS A1 APLICACION DE LAS IMAGENES DEL A
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 I)	9	8	1	PROGRAMAC CALCULO DI MATEMATICA ECONOMIA S ADMINISTRAC GEOGRAFIA I PROSPECTIV SOCIOLOGIA TECNICAS DE INVESTIGAC
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2020 II)	9	8	1	GUITARRA TEATRO ACONDICION APRENDIZAJ OBSERVACION ANTROPOLO PSICOLOGIA ACCION PSIC INGLÉS B1
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	METODOLOG TECNICAS DE SEMINARIO I DESARROLLO DINAMICA DE PROGRAMAC AUTOMATIAS Y LENGUAJES FORMALES
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	METODOS DE ADMINISTRAC FORMULACION PROYECTO D EMPRENDIM GESTION TEC PRACTICA PROFESIONAL
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	PROTOCOLO INGLÉS III TECNICAS DE LEGISLACION MACROECON COMERCIO Y DISEÑO DE PROYECTOS
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	PROTOCOLO ACONDICION HISTORIA DE EPISTEMOLOG PROCESOS O PERSONALID CIBERCULTURA
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	EPSTEMOLOG GESTION DE DESARROLLO SALUD OCUP UNIM TELEMATICA DINAMICA DE SISTEMAS
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	GUITARRA FUNDAMENT FUNDAMENT LA IMAGEN Y VIDEO INGLÉS A2 APERCIACION DE LAS IMAGENES DEL ARTE
IO PLAZO 0% (CERO POR CIENTO)	7	7	0	ETICA Y CIUC AMBIEL PRINCIPIOS Y VALORACION EVALUACION FUENTES DE PENSAMIENTO LOGICO DIVERGENTE (E_LEARNING)
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	TEATRO TEORIAS DEL INVESTIGACI ENGLISH VI MATERIALS E DIDACTICS O ETICA DE LA PROFESION DOCENTE
UNAD: MEN - UNAD MATRICULA CERO 2022 - I	7	7	0	INCLUSION S PROTOCOLO TEATRO BIOLOGIA ACONDICION PRACTICA PR INGLÉS B1
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2019)	7	7	0	COMERCIO Y CONTABILID PROCESO AD MERCADO I LIDERAZGO HERRAMIENTAS INGLÉS A1
UNAD: GENERACION E VIGENCIA 2021 II (CONTINUIDAD AÑO 2019)	7	7	0	PSICOPATOLOGI ANTROPOLO ACCION PSIC PSICOLOGIA SEMINARIO E PRACTICA PR INGLÉS B1
UNAD: GENERACION E VIGENCIA 2022 II (CONTINUIDAD AÑO 2019)	7	7	0	ETICA Y CIUC INVESTIGACIONES DE P A ARTE Y ETICA PEDAGOGICA PEDAGOGIA INGLÉS A1

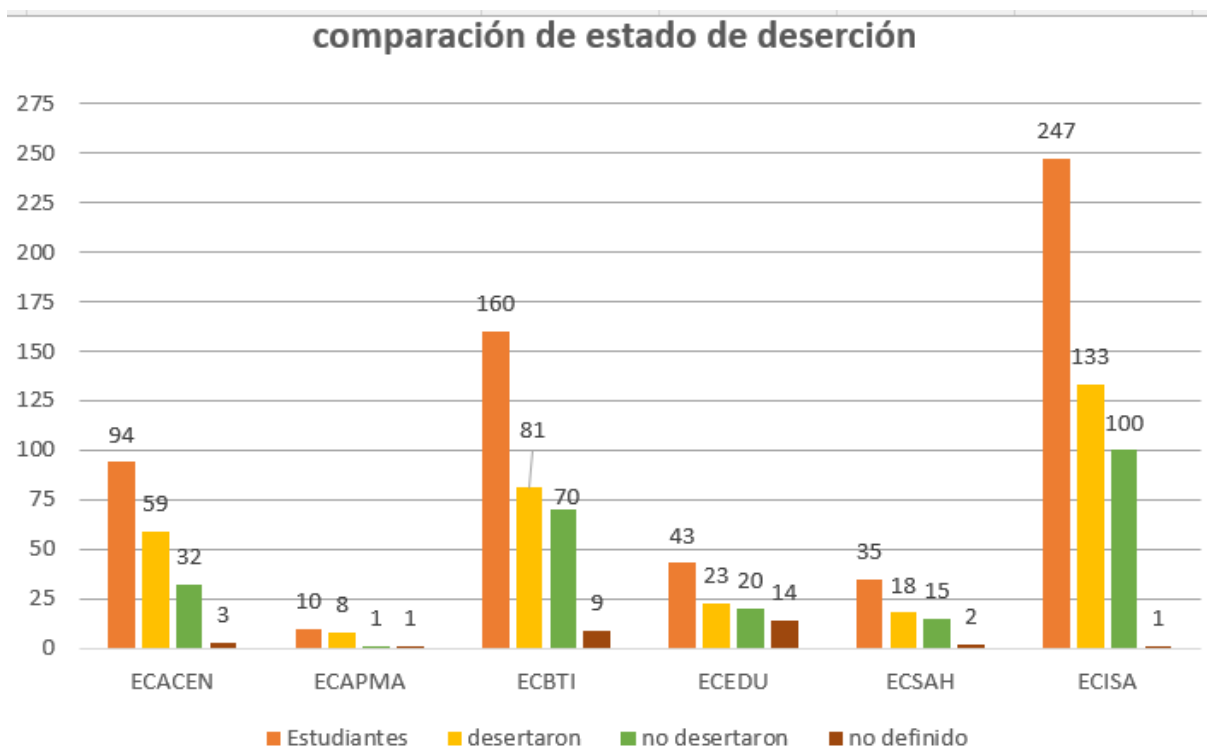
Figura 12

Resultados de reporte procesado de Alertas tempranas



**Figura 13**

*Resultados de reporte procesado de Alertas tempranas*



En la figura 13 se muestran los resultados de deserción por Escuela, se destaca la Escuela ECISA en la mayor cantidad de estudiantes.

### **Visualizaciones Avanzadas en Power BI**

#### 1. Gráfico de Heatmap (Mapa de Calor)

Propósito: Mostrar tasas de deserción por programa y semestre.

Pasos:

#### 2. Gráfico de Líneas con Segmentación

Propósito: Mostrar la evolución de la deserción con un desglose por modalidad.

### 3. Mapas Interactivos

Propósito: Analizar la distribución regional de la deserción.

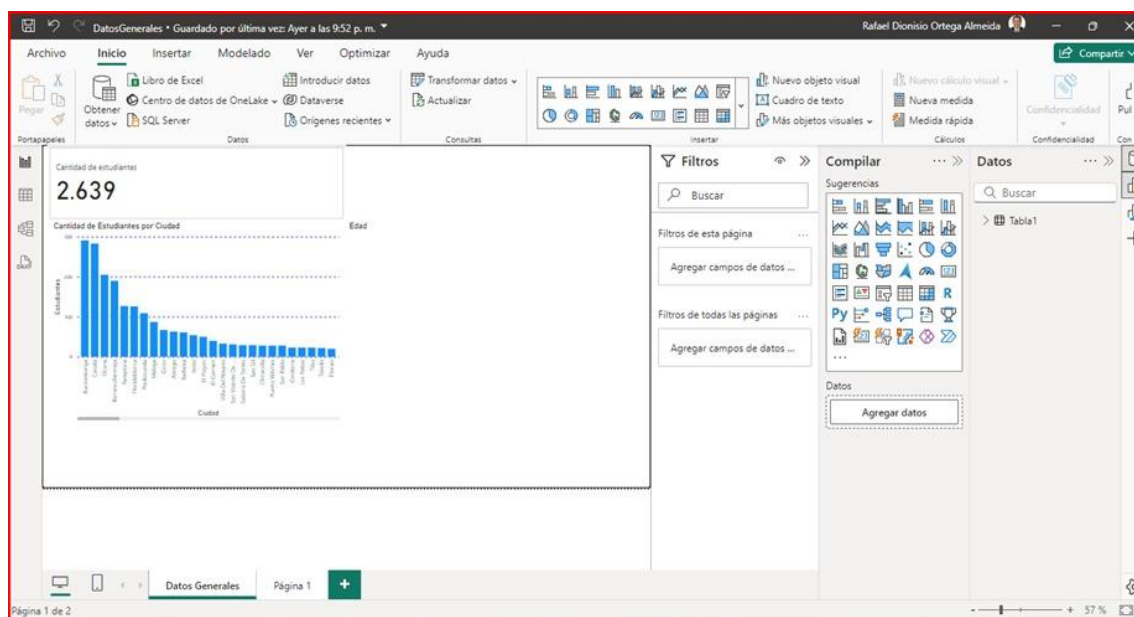
### 4. Gráfico de Dispersión Avanzado

Propósito: Relacionar promedio académico con tasa de deserción.

En la figura 14 se observa visualizaciones en Power BI de la cantidad total de estudiantes que se tomaron para la caracterización del reporte que entrega Consejería ZCORI, en la gráfica observamos las ciudades en las que se encuentra matriculado el estudiante; en la grafica 15 vemos el resultado de clasificación de las edades de los estudiantes matriculados en la ZCORI; en la grafica 16 observamos la identificación de genero de dichos estudiantes y en la grafica 17 estos estudiantes por estrato social.

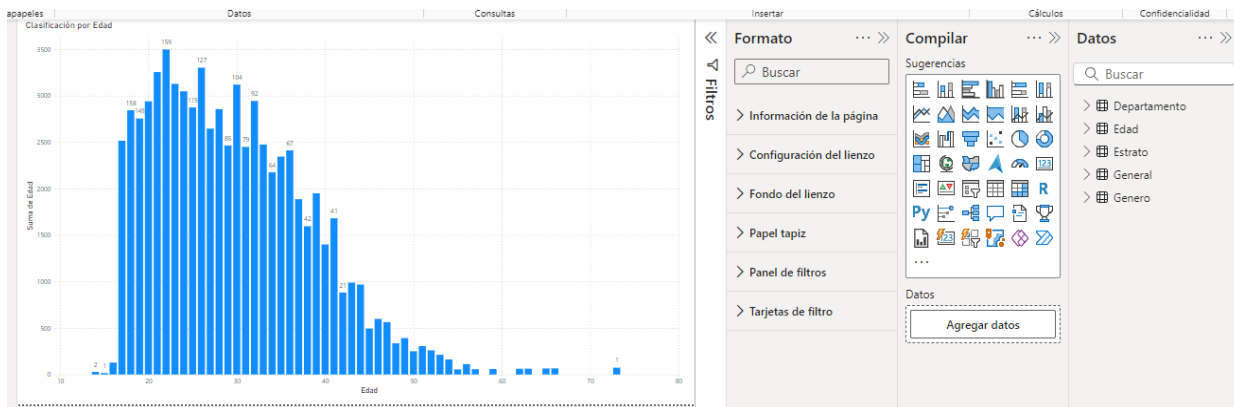
**Figura 14**

*Visualización resultados en Power BI*



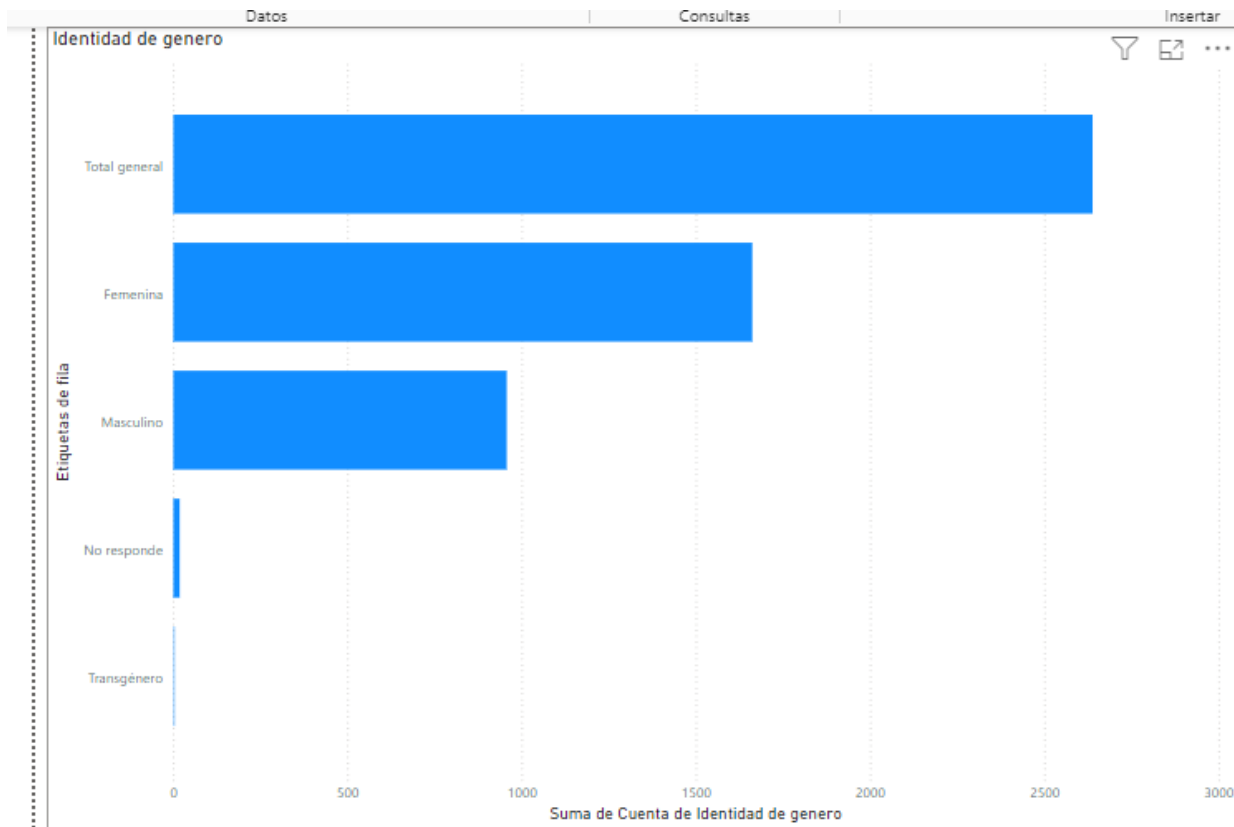
**Figura 15**

*Visualización resultados en Power BI*



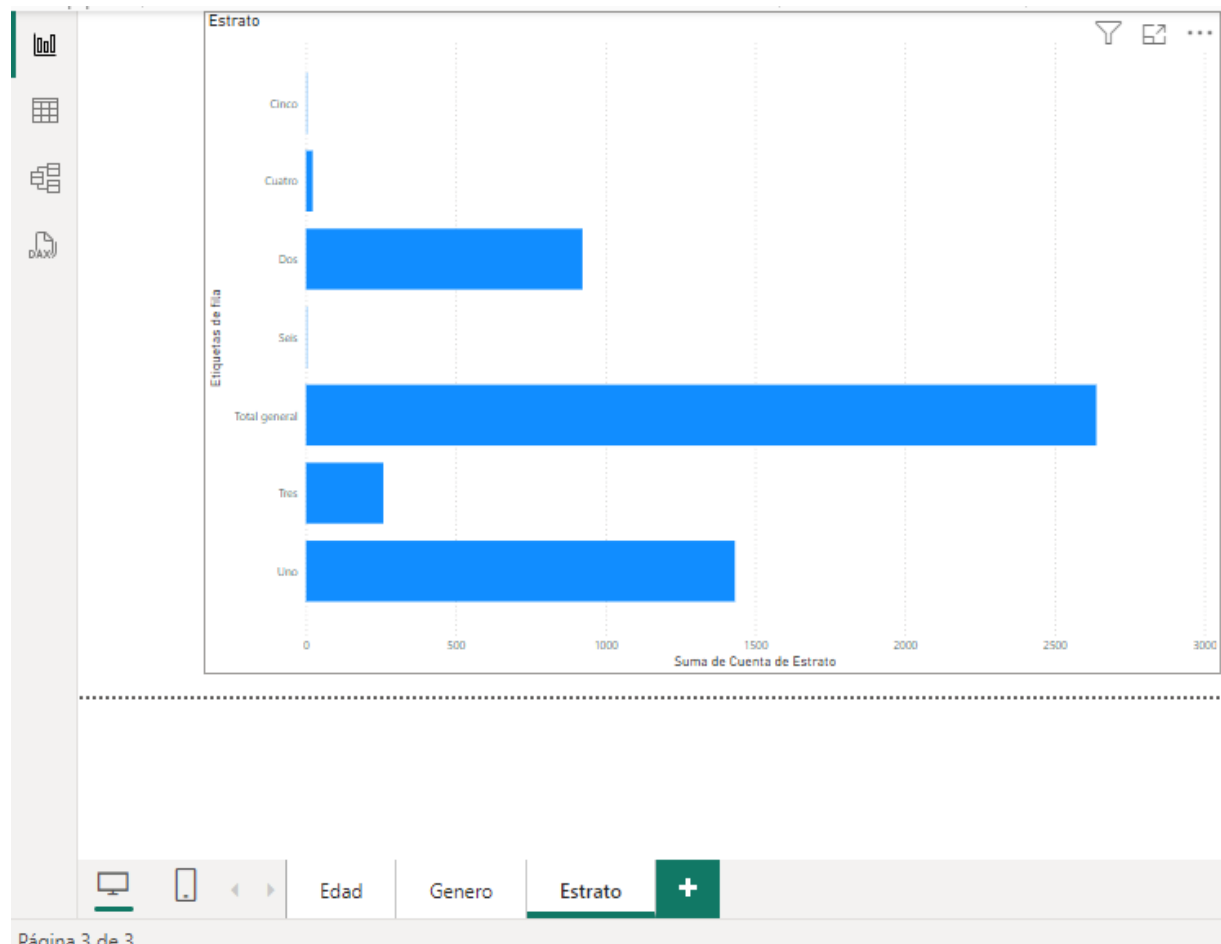
**Figura 16**

*Visualización resultados en Power BI*



**Figura 17**

*Visualización resultados en Power BI*



## Análisis de resultados

En la clasificación de estudiantes en riesgo de deserción en la UNAD (Universidad Nacional Abierta y a Distancia), se destacan algunos resultados clave que suelen derivarse de investigaciones en esta área:

### Identificación de patrones de deserción:

La combinación de datos académicos (como calificaciones y tasas de aprobación), socioeconómicos (situación laboral y nivel educativo familiar) y demográficos permite identificar patrones significativos asociados con el riesgo de deserción.

### Segmentación del riesgo:

Se pueden clasificar a los estudiantes en diferentes niveles de riesgo (bajo, medio, alto) mediante algoritmos de machine learning, como árboles de decisión, regresión logística o modelos de clasificación más avanzados como Random Forest y SVM.

### Factores más influyentes:

En muchas investigaciones, los factores que tienen mayor peso son el compromiso académico, la carga laboral del estudiante, el acceso a recursos tecnológicos y la percepción de soporte por parte de la institución.

### Predicción con alta precisión:

Dependiendo del modelo utilizado y la calidad de los datos, se pueden lograr niveles de precisión en la predicción del riesgo de deserción que superan el 85%. Esto facilita la implementación de estrategias focalizadas.

Intervenciones más efectivas:

Los resultados también permiten diseñar estrategias de apoyo como tutorías, becas, o cambios en la carga académica, que han mostrado ser efectivas para estudiantes identificados como de alto riesgo.

Algunos resultados relevantes de las intervenciones realizadas sobre los estudiantes del UNAD en el CEAD Bucaramanga y discriminados por Escuelas, nos muestra que la Escuela con más alto porcentaje de estudiantes que son recurrentes en la no entrada a los cursos y que sus calificaciones muestran bajas calificaciones es la ECSAH con 56% del total de Estudiantes y en consecuencia el programa con más alto porcentaje es el programa de Tecnología en Regencia de Farmacia, el cual aparecen un total de 64 estudiantes reportados con no ingreso a los cursos y/o que van reprobando al menos un curso de los matriculados.

Otro hallazgo es que el rango de edad de estudiantes con problemas en sus cursos cada vez en menor el rango de edad, esto es debido a que en cada periodo el número de estudiantes nuevos matriculados está en el rango de edad de 25 a 30 años.

### Figura 17

*Escuela a la que pertenecen los estudiantes que aparecen en los reportes de los que van reprobando cursos*

Escuela	Total estudiantes intervenidos en los dos reportes sin repetir	Estudiantes aparecen en reporte 1 y 2		Programa con mas estudiantes reportados	
		Cantidad estudiantes	Porcentaje	Programa (Estudiantes)	Porcentaje
ECACEN	326	79	23%	Administración de Empresas (154)	44%
ECAPMA	128	34	21%	Agronomía (61)	38%
ECBTI	468	180	32%	Ingeniería Industrial (252)	45%
ECEDU	332	180	43%	Licenciatura En Lenguas Extranjeras Con Enfoque En Inglés (213)	50%
ECISA	67	52	56%	Tecnología En Regencia De Farmacia (64)	69%
ECJP	54	34	48%	Ciencia Política (38)	54%
ECSAH	320	144	37%	Psicología (239)	61%
	1695	703			

## Recomendaciones

Se recomienda al CEAD Bucaramanga continuar con la mejora de los modelos predictivos que asisten la identificación de patrones de deserción estudiantil. Los modelos de regresión lineales y otras técnicas de minería de datos que aplicados periódicamente permite predecir y detectar a los estudiantes en riesgo de deserción, permitiendo la intervención temprana.

Es importante que el CEAD Bucaramanga continúe fortaleciendo su sistema de alertas tempranas mediante el uso de las herramientas de análisis desarrolladas, esto puede incluir la creación de umbrales para las alertas, basados en los datos obtenidos de las variables académicas, socioeconómicas y de participación estudiantil.

Para asegurar una retención estudiantil sostenible, se debe realizar un monitoreo continuo de las variables que afectan la deserción, como el rendimiento académico, la participación en actividades extracurriculares y las condiciones socioeconómicas. Además, se recomienda que el sistema de gestión de datos se mantenga actualizado con la información más reciente para que las decisiones sean siempre pertinentes.

Es recomendable que se proporcione capacitación periódica a los directivos y al personal administrativo en el uso de las herramientas de visualización de datos como Power BI y Excel VBA, con el fin de garantizar que los usuarios del sistema puedan interpretar correctamente los datos y tomar decisiones informadas basadas en ellos.

Se sugiere desarrollar intervenciones personalizadas para los estudiantes identificados como potenciales desertores. Estas intervenciones pueden incluir tutorías académicas, asesoría psicopedagógica y un seguimiento más cercano para mejorar su rendimiento y compromiso.

Es recomendable ampliar las variables de análisis, incorporando otros factores que puedan influir en la deserción, como la salud mental, el acceso a recursos tecnológicos, y la satisfacción con la metodología educativa de la UNAD. La inclusión de más variables podría mejorar la precisión de los modelos predictivos.

## Referencias Bibliográficas

Departamento Administrativo Nacional de Estadística (DANE). (2024, febrero 16). Estadísticas de educación. Enlace Web: <https://www.dane.gov.co/index.php/estadisticas-por-tema/educacion>.

Du, K. (2023). Application of English education big data system based on multi-source information fusion and machine learning. *Soft Computing*.  
<https://doi.org/10.1007/s00500-023-08429-w>

Enlace Web: <https://aurea2.unad.edu.co/analitica/>

Gaskins, K. (s.f.) "Using VBA in Excel to Automate Reports."

Lamas Dos Santos Lopes, M. I., Sismeiro Da Silva Pereira, A. P., & Fortunato Vaz, P. M. (2023). Abandono escolar no ensino superior e fatores concorrentes. (Portuguese). *EduSer: Revista de Educação*, 15(1), 50–64. <https://doi-org.bibliotecavirtual.unad.edu.co/10.34620/eduser.v15i1.236>

Madrid Orrego, D. M., Figueroa Royero, L., Berrio Caballero, H., Martínez Choles, A. Y., & Charris Rodríguez, J. A. (2023). Minería de datos para la identificación de la Deserción Estudiantil en Universidades Públicas: Data mining for the identification of Student Dropout in Public Universities. *LATAM Revista Latinoamericana De Ciencias Sociales Y Humanidades*, 4(2), 1464–1476. <https://doi.org/10.56712/latam.v4i2.697>.

Observatorio Laboral para la educación (OLE), (2024, febrero 16). Sistemas de información. Enlace Web: <https://ole.mineduacion.gov.co/portal/Sistemas-de-informacion/>

- Ramírez, P. E., & Grandón, E. E. (2018). Predicción de la Deserción Académica en una Universidad Pública Chilena a través de la Clasificación basada en Árboles de Decisión con Parámetros Optimizados. *Formación Universitaria*, 11(3), 3–10.
- Rodríguez Molano, J. I., Forero Zea, L. D., & Piñeros Reina, Y. F. (2019). Proposal of Architecture And Application of Machine Learning (MI) as A Strategy for the Reduction of University Desertion Levels Due to Academic Factors. *Ingeniería Solidaria*, 15(29), 1–23. <https://doi-org.bibliotecavirtual.unad.edu.co/10.16925/2357-6014.2019.03.06>
- Sistema Nacional de Información de la Educación Superior (SNIES). (2024, febrero 16). Bases consolidadas. Enlace Web:  
<https://snies.mineduacion.gov.co/portal/ESTADISTICAS/Bases-consolidadas/>
- Sotomonte Castro, J. E., Rodríguez Rodríguez, C. C., Montenegro Marín, C. E., Gaona García, P. A., & Gabriel Castellanos, J. (2016). Hacia la construcción de un modelo predictivo de deserción académica basado en técnicas de minería de datos. *Revista Científica*, 3(26), 35–48. <https://doi-org.bibliotecavirtual.unad.edu.co/10.14483/23448350.11089>
- United Nations Development Programme. (2016)
- Universidad Nacional Abierta y a Distancia (UNAD). (2024, FEBRERO16). Analítica.
- Urbina-Nájera, A. B., Camino-Hampshire, J. C., & Cruz Barbosa, R. (2020). Deserción escolar universitaria: Patrones para prevenirla aplicando minería de datos educativa. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 26(1).  
<https://doi.org/10.7203/relieve.26.1.16061>

Viloria, A., Guliany, J. G., Núñez, W. N., Palma, H. H., & Núñez, L. N. (2020). Retraction: Data mining applied in school dropout prediction (journal of physics: Conference series 1432 012092). *Journal of physics. Conference series*, 1432(1), 012107.  
<https://doi.org/10.1088/1742-6596/1432/1/012107>.

Walkenbach, J. (2019). *Excel 2019 Power Programming with VBA*. Wiley.