

**Herramienta de monitoreo y análisis: *Dashboard* interactivo del desempeño académico de
los estudiantes de Ciencia Política en la UNAD**

Blademir Quiguanas Chila

David Leonardo Arias Bocanegra

Asesor

Isaac Esteban Camargo Freile

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

2024

Resumen

El presente proyecto aplicado tiene como objetivo desarrollar un *dashboard* como medio de visualización y análisis de datos que permita acceder a *KPIs* específicos relacionados con el desempeño académico de los estudiantes de Ciencia Política de la Universidad Nacional Abierta y a Distancia- UNAD, para la identificación de información significativa que promuevan acciones de intervención efectivas para el mejoramiento continuo del sistema educativo. El conjunto de datos a utilizar se obtiene de los registros de la Escuela de Ciencias Jurídicas y Políticas de la UNAD, alojados en SII 5.0 – Analítica, desde el año 2019 al 2024. El enfoque de la investigación es cuantitativo, con alcance descriptivo y se implementa la metodología Diseño *Scrum*. Finalmente, se desarrolla el *dashboard* en *Power BI*.

Palabras claves: *Dashboard, business intelligence, desempeño académico, ciencia política.*

Abstract

The objective of this applied project is to develop a dashboard as a means of visualization and data analysis that allows access to specific KPIs related to the academic performance of political science students at the National Open and Distance University - UNAD, for the identification of information significant that promote effective intervention actions for the continuous improvement of the educational system. The data set to be used is obtained from the records of the UNAD School of Legal and Political Sciences, hosted in SII 5.0 - Analitica from 2019 to 2024. The research approach is quantitative, with descriptive scope and the Scrum Design methodology is implemented. Finally, the dashboard is developed in Power BI.

Keywords: Dashboard, business intelligence, academic performance, political science.

Tabla de Contenido

Introducción	8
Planteamiento del Problema	10
Justificación	12
Objetivos	14
Objetivo General	14
Objetivos Específicos	14
Marco Conceptual y Teórico	15
Marco Conceptual	15
Dashboard	15
Desempeño Académico	16
Marco Teórico	17
Antecedentes	17
Learning Analytics	19
Visualización de Datos	21
Metodología	23
Tipo de Investigación	23
Diseño de Investigación	24
Calidad de la Información del SII 5.0 del Aplicativo Analítica	25
Procesamiento de las Inconsistencias en los Archivos	25
Manejo de los Periodos Académicos	26
Evaluación de la Calidad de los Datos	26
Desafíos Específicos de los Periodos Académicos y su Manejo	27

Métricas Relacionadas con el Desempeño Académico	28
Tasa de Deserción y Continuidad por Periodo y Zona Geográfica.....	28
Análisis de Desempeño Académico por Corte Histórico.....	28
Impacto del Estrato Socioeconómico en la Deserción.....	29
Proporción de Género y Distribución por Edad.....	29
Desempeño Laboral vs Continuidad Académica	30
Herramienta de Monitoreo y Análisis: Dashboard Interactivo	31
Principales Resultados Derivados del Dashboard.....	34
Conclusiones.....	37
Recomendaciones	38
Referencias Bibliográficas	39
Apéndice	45

Lista de Figuras

Figura 1 <i>Dashboard Interactivo del Programa de Ciencia Política de la UNAD</i>	31
Figura 2 <i>Cantidad de Estudiantes por Zona y Proporción Desempeño Laboral</i>	32
Figura 3 <i>Cantidad de Estudiantes por Periodo</i>	32
Figura 4 <i>Edad de los Estudiantes</i>	33
Figura 5 <i>Proporción por Área Geográfica</i>	33
Figura 6 <i>Cantidad de Estudiantes por Estrato</i>	34
Figura 7 <i>Desempeño Académico de los Estudiantes del Programa de Ciencia Política</i>	34

Lista de Apéndices

Apéndice A <i>Código en Python para la Consolidación de Datos</i>	45
--------------------------------------------------------------------------------	----

Introducción

En el contexto educativo actual, el manejo y análisis de grandes volúmenes de datos se ha convertido en un factor clave para la toma de decisiones informadas. Las instituciones de educación superior, al enfrentarse a retos como la deserción estudiantil y la medición del desempeño académico, requieren herramientas que permitan un análisis detallado y eficiente de sus datos. El uso de tecnologías de visualización de datos como los *dashboards* interactivos, ha demostrado ser una solución eficaz para monitorear, analizar y mejorar la gestión académica. Este proyecto aplicado se enfoca en el desarrollo de un *dashboard* que facilite la visualización y análisis del desempeño académico de los estudiantes del programa de Ciencia Política de la Universidad Nacional Abierta y a Distancia (UNAD).

El propósito principal de este proyecto es ofrecer a la institución una herramienta que permita no solo centralizar la información, sino también analizar patrones y tendencias en el rendimiento estudiantil. Utilizando datos históricos desde 2019 hasta 2024, se pretende identificar indicadores clave como la tasa de deserción, la continuidad académica, el desempeño por estrato socioeconómico y la distribución geográfica de los estudiantes. Esta información será esencial para formular políticas y estrategias de intervención que promuevan la mejora continua del sistema educativo.

El diseño metodológico de este proyecto se basa en un enfoque cuantitativo con un alcance descriptivo, empleando la metodología Scrum para la implementación del *dashboard*. La herramienta se desarrollará en Power BI, una plataforma robusta para la visualización de datos, que permitirá la interacción con los distintos conjuntos de datos y la generación de informes detallados. La automatización en el manejo de los datos, mediante el uso de Python y su

biblioteca Pandas, será crucial para consolidar la información de manera eficiente, garantizando la calidad y consistencia de los datos utilizados en el análisis.

La importancia de este proyecto radica en su capacidad para transformar los datos académicos en información valiosa que apoye la toma de decisiones en la UNAD. La identificación de áreas críticas como la deserción o el bajo desempeño permitirá a los gestores educativos actuar de manera oportuna y eficaz, mejorando tanto el proceso de enseñanza-aprendizaje como las políticas de retención estudiantil. Además, el *dashboard* proporcionará una visión clara y estructurada de los indicadores clave de rendimiento, lo que facilitará la planificación y seguimiento de las acciones correctivas implementadas por la institución.

Por último, el desarrollo de este *dashboard* interactivo contribuirá significativamente al fortalecimiento del sistema educativo de la UNAD, ofreciendo una herramienta tecnológica avanzada para el monitoreo y análisis del desempeño académico. A través de esta plataforma, se espera que la universidad pueda optimizar sus procesos internos, mejorar la calidad educativa y garantizar una mayor retención de sus estudiantes, asegurando un impacto positivo en la formación de futuros profesionales en el campo de la Ciencia Política.

Planteamiento del Problema

La UNAD, como la primera megauniversidad pública del país y acreditada en alta calidad, con más de 37 años de trayectoria y más de 100 mil estudiantes activos, posee condiciones operativas y de calidad suficientes para continuar su crecimiento de manera significativa (UNAD, 2018). Este aumento va de la mano con los datos que se almacenan en el Sistema Integrado de Información - SII 5.0 por cada periodo académico, donde se oferta los diferentes programas de pregrado y posgrado de las escuelas de la institución a nivel nacional, obteniendo un alto volumen de información disponible y sin una organización definida para su interpretación y el monitoreo de diferentes factores claves con el fin de mantener una evaluación constante de la calidad educativa.

Sellán (2021) indica que la gran cantidad de datos generados en los sistemas de información de las instituciones de educación superior, junto con la falta de su adecuado procesamiento y la ausencia de indicadores académicos, impide a los actores claves en el sistema educativo conocer de manera oportuna el rendimiento del estudiante.

En este contexto, la adopción de sistemas de *Business Intelligence* se vincula estrechamente con la transformación digital. Este proceso requiere una gestión integral que promueva cambios en la mentalidad, la cultura organizacional, los estilos de liderazgo y la colaboración en equipo. La cultura organizacional juega un papel crucial en el éxito de la implementación de estos sistemas, ya que no basta con recopilar información; es esencial utilizarla de manera efectiva en la toma de decisiones, convirtiéndola en un recurso valioso para el proceso de mejora continua (Mora et al., 2021).

Para hacer frente a esta problemática, se han desarrollado modelos de madurez de *Business Intelligence and Analytics (BIA)* con el propósito de atender las necesidades

particulares de IES y obtener una ventaja competitiva en un mundo cada vez más orientado por los datos (Cardoso y Su, 2022; Stewart y Dewan, 2022). Además, se han llevado a cabo otras investigaciones que se centran en diseñar e implementar un prototipo de *dashboard* utilizando datos de egresados o en la generación de informes que revelan el comportamiento de los estudiantes en la plataforma Moodle. Esto se realizó con el propósito de monitorear su progreso y facilitar mejoras en los contenidos pedagógicos (Sellán, 2021; Castillo et al., 2018).

Según Cardoso y Su (2022) los sistemas de BIA son vitales en las organizaciones, al brindar datos procesables que facilitan la toma de decisiones informadas por parte de los usuarios empresariales. Sin embargo, muchas IES carecen de modelos accesibles y prácticos para guiar el desarrollo progresivo de soluciones BIA y maximizar su utilidad. Esta falta se acentúa en un entorno educativo cada vez más complejo y dinámico, marcado por la globalización y el rápido avance de las tecnologías de la información.

Teniendo en cuenta lo anterior, se plantea la siguiente pregunta de investigación: ¿Cómo puede el desarrollo y la implementación de un *dashboard* para la visualización y análisis de datos relacionados con el desempeño académico de los estudiantes de Ciencia Política en la UNAD, facilitar la identificación de información significativa específica y promover acciones de intervención efectivas para el mejoramiento continuo del sistema educativo?

Justificación

Las instituciones de educación superior, en sus procesos administrativos y académicos, acumulan un volumen significativo de datos que, en muchas ocasiones, no están debidamente organizados. Por lo general, estos datos se encuentran dispersos en diversos sistemas, lo que dificulta su análisis y su utilización para la toma de decisiones o el monitoreo constante de variables clave, con el objetivo de mejorar continuamente la calidad educativa (Destiandi y Hermawan, 2018).

En la literatura se evidencia un sinnúmero de investigaciones enfocadas en la implementación del *Machine Learning* en el contexto educativo y las cuales han mostrado una tendencia creciente en su uso para el análisis de todos los niveles de educación existente (Forero y Negre, 2024). Especialmente en la predicción del rendimiento académico a través de técnicas como árbol de decisión y K vecinos más cercanos y redes neuronales (Gutiérrez et al., 2022; Rico y Gaytan, 2022; Contreras et al., 2020) y en el pronóstico de la deserción universitaria (Valero et al., 2022). Sin embargo, poco se ha explorado este campo en la línea de *Business Intelligence (BI)*, exactamente el desarrollo de un *dashboard* como insumo para la organización e interpretación de la información en Instituciones de Educación Superior (IES) acreditadas en alta calidad.

Por tanto, el desarrollo de un *dashboard* para el programa de Ciencia Política de la Escuela de Ciencias Jurídicas y Políticas -ECJP de la UNAD, incentiva la eficiencia y transparencia de la información, la rendición de cuentas dentro de la universidad, la planificación, minimizar las dificultades de acceso y las asimetrías informativas existentes en la Administración Pública (Azevedo, 2021). También aporta a la literatura con *Key performance*

indicators (KPIs) identificados a partir de los datos suministrados por la institución como *proxys* al desempeño académico.

Como resultado del proyecto aplicado, la ECJP contará con una herramienta de monitoreo y análisis, con información centralizada y organizada para gestionar estos datos de manera eficiente. En otras palabras, el desarrollo del *dashboard* permitirá visualizar y analizar los datos de manera clara y comprensible, lo que facilitará la labor de los responsables de la toma de decisiones al identificar con mayor precisión las unidades o procesos de intervención y mejora.

Objetivos

Objetivo General

Desarrollar un *dashboard* como medio de visualización y análisis de datos que permita acceder a *KPIs* específicos relacionados con el desempeño académico de los estudiantes de Ciencia Política de la UNAD, para la identificación de información significativa concreta que promuevan acciones de intervención efectivas para el mejoramiento continuo del sistema educativo

Objetivos Específicos

Investigar la calidad de la información de las bases de datos internas de la UNAD relacionadas con el desempeño académico de los estudiantes de Ciencia Política desde el año 2019.

Establecer qué métricas pueden aportar a un mayor entendimiento del desempeño académico de los estudiantes de Ciencia Política de la UNAD.

Desarrollar un diseño visual intuitivo para el *dashboard*, que facilite la comprensión de la información relacionada con el desempeño académico.

Marco Conceptual y Teórico

Marco Conceptual

El presente marco proporciona la estructura conceptual sobre la cual se fundamenta el desarrollo del proyecto “*Herramienta de Monitoreo y Análisis: Dashboard Interactivo del desempeño académico de los estudiantes de Ciencia Política en la UNAD*”. Este marco se compone de dos elementos principales: *Dashboard* y Desempeño académico.

Dashboard

Dashboard se traduce comúnmente como "tablero de mando" o "panel de control". Según Sarikaya et al. (2018) lo definen como un tablero de mando es una representación visual de datos utilizada para monitorear y comprender un proceso específico, mediante infografías o visualizaciones narradas. Su evolución ha pasado de simples presentaciones con una sola vista a interfaces interactivas con múltiples vistas, incorporando enfoques en comunicación, aprendizaje, motivación y apoyo en la toma de decisiones.

Para Few (2004) un *dashboard* es una herramienta visual que representa datos importantes y métricas relevantes gráficamente para apoyar la toma de decisiones. Proporciona la información de forma clara y resumida en una única pantalla, permitiendo a los usuarios supervisar el rendimiento, reconocer tendencias y responder de manera ágil. En otras palabras, esta herramienta facilita la visualización de problemas y la toma de decisiones para corregir posibles errores. Su objetivo principal es convertir los datos en información útil para guiar nuestra estrategia hacia el logro de los objetivos establecidos y garantiza el seguimiento y control de los KPIs (índices claves del desempeño académico) (Bocanegra, 2019).

Por tanto, los *dashboards*, o tableros de mando, son herramientas vitales en el ámbito del análisis y la gestión de datos. Su evolución hacia interfaces interactivas ha permitido incorporar

enfoques más complejos que van más allá de simplemente mostrar datos, ahora también pueden facilitar la comunicación, el aprendizaje, la motivación y apoyar la toma de decisiones.

Desempeño Académico

Según Delgado et al. (2007) la evaluación del desempeño puede basarse en la calificación del profesor o en el promedio obtenido por el alumno, siendo este último un resumen del rendimiento escolar reconocido. En cambio, Aldana et al. (2010) lo define como el conjunto de experiencias y éxitos educativos de los estudiantes, resultado de su interacción con la institución de educación superior y las contribuciones que esta realiza a su desarrollo profesional y formación integral como individuos capaces de pensamiento crítico y acción reflexiva.

En general, la complejidad del rendimiento académico comienza desde su definición, a veces se refiere como aptitud escolar, desempeño académico o rendimiento escolar, aunque estas diferencias conceptuales suelen ser meramente semánticas, ya que comúnmente se utilizan de manera intercambiable en los textos para referirse a la vida estudiantil y la labor docente (Navarro, 2003).

Por otro lado, hay acuerdo en la literatura en que los factores que influyen en el rendimiento académico pueden dividirse en dos categorías principales: los determinantes personales y los determinantes sociales. Los primeros incluyen aspectos como la inteligencia, las habilidades, la asistencia a clases, el género y las calificaciones de acceso a la universidad. Los segundos abarcan el entorno familiar, el contexto socioeconómico, variables demográficas como sexo, edad y estado civil, así como el nivel educativo de los padres. Se considera que comprender estos factores asociados al rendimiento puede ayudar a las autoridades universitarias a desarrollar políticas que mejoren la calidad académica de los estudiantes de pregrado (Barahona, 2014).

Según Medina et al. (2021), los factores individuales, sociales e institucionales, así como todas sus dimensiones, parecen influir en el rendimiento académico, lo que sugiere una conexión entre estos elementos. Además, el rendimiento académico de los estudiantes en la educación secundaria tiene capacidad predictiva sobre su rendimiento en la educación universitaria (Soria et al., 2014).

Lo anterior, indica que la evaluación del desempeño académico puede basarse en la calificación del profesor o en el promedio obtenido por el alumno, reflejando el rendimiento escolar reconocido y el impacto de la institución en su desarrollo integral. La complejidad conceptual del rendimiento académico abarca diversas terminologías, pero la literatura coincide en que los factores personales y sociales influyen en él. Además, la conexión entre el rendimiento académico en la educación secundaria y universitaria destaca la importancia de comprender y abordar estos factores para mejorar la calidad educativa.

Marco Teórico

Antecedentes

En los tiempos actuales, las organizaciones enfrentan la necesidad de adoptar métodos más eficientes para el tratamiento de datos a lo largo de la cadena de valor. Pascal et al. (2017) destacan la importancia de optimizar los sistemas de gestión para controlar el conocimiento interno y externo y planificar el futuro de manera efectiva. En este contexto, la Facultad de Ingeniería de la Universidad Nacional de Lomas de Zamora implementó un *Data Warehouse* como parte de un proyecto de Tecnologías de la Información y Comunicación, mejorando así la integración del conocimiento de todos los *stakeholders* y facilitando la toma de decisiones estratégicas.

Por su parte, Castillo et al. (2018) abordan la carencia de información eficiente en las instituciones de educación superior y la falta de indicadores que evalúen a los egresados. Su investigación se centra en el desarrollo de un Sistema de Información Web para Egresados en la Universidad Tecnológica de Panamá, que permite la generación de *dashboards* útiles para la toma de decisiones. Esto refleja cómo la Inteligencia de Negocios (BI) puede transformar la retroalimentación y mejorar los procesos de enseñanza-aprendizaje.

En una línea similar, Mora et al. (2021) exploran la implementación de un sistema de inteligencia de negocios en la Escuela de Ciencias de la Administración de la Universidad Estatal a Distancia - UNED de Costa Rica. Utilizando una metodología mixta y herramientas en línea, identificaron las necesidades específicas de los actores involucrados. Los resultados demostraron la pertinencia y efectividad de estas tecnologías, subrayando la importancia de una adecuada preparación técnica y cultural en la organización para su implementación.

Sellán (2021) aborda la generación masiva de datos en instituciones de educación superior y la falta de tratamiento adecuado de estos datos. Su estudio utiliza *Power BI* para crear un prototipo que visualiza la información académica, facilitando la toma de decisiones de docentes y coordinadores. La integración de estos *dashboards* en plataformas como Moodle demuestra cómo la BI puede consolidar y resumir información, mejorando la calidad educativa.

Apraxine y Stylianou (2017) investigan cómo las prácticas de inteligencia de negocios pueden aplicarse a la industria educativa, mejorando el proceso de toma de decisiones en varios departamentos de una institución. Su enfoque de estudio de caso, utilizando entrevistas cualitativas y análisis de contenido, resalta la percepción positiva de los gerentes hacia la implementación de BI en la Universidad de Nicosia y los beneficios derivados de ello.

Azevedo et al. (2021) discuten la gobernanza en las instituciones públicas y privadas como un modelo contemporáneo para la evaluación y monitoreo, además de promover la transparencia. Proponen la construcción de un prototipo de un tablero de control de BI para una institución pública brasileña enfocada en la educación. Este enfoque no solo apoya la toma de decisiones de los gestores, sino que también mejora la eficiencia y la transparencia pública, minimizando las dificultades en la gestión de sistemas de BI.

Destiandi y Hermawan (2018) presentan el periodo de vida de la Inteligencia de Negocios como un método estructurado para desarrollar aplicaciones de soporte a la toma de decisiones, como el Tablero Académico. En Indonesia, la implementación de estos sistemas permite generar informes académicos rápidos y precisos, facilitando la toma de decisiones estratégicas. La metodología Kimball y herramientas como Microsoft Power BI son clave para manejar grandes volúmenes de datos y elevar la información a un nivel gerencial.

La investigación en el ámbito de la inteligencia de negocios en la educación superior demuestra su impacto positivo en la toma de decisiones y la mejora de los procesos académicos. Desde la integración de datos hasta la creación de *dashboards* y sistemas de información, estas tecnologías ofrecen una solución robusta para los desafíos contemporáneos en la gestión educativa

Learning Analytics

Durante la última década, ha surgido el término *Learning Analytics (LA)*, conocido en español como analítica del aprendizaje o análisis del aprendizaje. Este término fue presentado por primera vez en la conferencia internacional LAK (*Learning Analytics and Knowledge*) en 2011 y ha sido adoptado por SOLAR (*Society for Learning Analytics*) (Haro, 2018).

Según Vélez (2023) *Learning Analytics* se refiere al proceso de recolección, análisis y presentación de información relacionada con los estudiantes y sus interacciones en entornos de aprendizaje, con el objetivo de profundizar en la comprensión del aprendizaje, mejorar los ambientes educativos y potenciar los resultados del proceso de enseñanza-aprendizaje. Se trata de un área de investigación en desarrollo que emplea técnicas estadísticas, minería de datos y representación visual para examinar información educativa con el propósito de comprender y perfeccionar el proceso de enseñanza-aprendizaje (Siemens, 2013).

Clow (2013) señala que el Learning Analytics busca alcanzar varios objetivos, incluyendo entender el comportamiento de los estudiantes mediante el análisis de sus patrones de acceso a recursos, participación en actividades y desempeño en evaluaciones. También persigue la identificación temprana de estudiantes en situación de riesgo, la personalización del aprendizaje según las necesidades y estilos individuales, y la mejora de la efectividad de cursos y programas mediante la evaluación y ajustes de materiales y actividades.

Aunque hay diferentes formas de clasificar los tipos de Learning Analytics, Clow (2013) propone cuatro categorías principales:

- a) **Descriptivo:** Este tipo de análisis se centra en describir lo que ha sucedido en el pasado. Se utiliza para comprender el comportamiento de los estudiantes y sus interacciones con el entorno de aprendizaje.
- b) **Predictivo:** El análisis predictivo se enfoca en predecir futuros comportamientos o resultados académicos de los estudiantes.
- c) **Prescriptivo:** Este tipo de análisis se centra en proporcionar recomendaciones o intervenciones específicas para mejorar el aprendizaje de los estudiantes.

d) **Diagnóstico:** El análisis diagnóstico busca comprender las causas subyacentes de los resultados observados. Se utiliza para identificar áreas de mejora en los cursos o programas de estudio.

En esencia, el análisis del aprendizaje se centra en la recopilación de conjuntos de datos educativos provenientes de sistemas digitales de aprendizaje, sistemas de gestión y plataformas de redes sociales. El propósito fundamental de esta práctica es analizar dichos datos con el fin de mejorar la efectividad del proceso de aprendizaje (Haro, 2018).

Visualización de Datos

La visualización de datos implica representar gráficamente la información y los datos, utilizando elementos visuales como gráficos, mapas y cuadros. Estas herramientas permiten una comprensión más accesible de tendencias, valores atípicos y patrones en los datos (Tableau, 2023).

El *LA* emplea métodos de representación gráfica de datos para mostrar de manera clara y precisa los resultados del análisis. Entre los tipos de visualización de datos utilizados en *Learning Analytics* se encuentran:

- a) **Gráficos de líneas:** son particularmente efectivos para mostrar tendencias y cambios continuos en los datos a lo largo de distintos intervalos de tiempo (Few, 2006).
- b) **Diagrama de barras:** Se crean barras en cantidad igual al número de categorías que posee la variable, de manera que la altura de cada una refleje la frecuencia o el porcentaje de casos en cada categoría (Pértega y Pita, 2001).
- c) **Histogramas:** son especialmente eficaces para explorar la forma y la estructura de la distribución de los datos y para comparar la frecuencia de diferentes valores dentro de un conjunto de datos. (Friendly, 2000).

d) Mapas de calor: son útiles para proporcionar una representación visual intuitiva de la densidad o intensidad de una variable en un espacio bidimensional, lo que permite identificar fácilmente áreas de mayor concentración o actividad (Wilkinson, et al., 2005).

Tanto la visualización de datos como el *LA* son esenciales para comprender y comunicar de manera efectiva información compleja. La variedad de técnicas de representación gráfica, que incluyen gráficos de líneas, diagramas de barras, histogramas y mapas de calor, ofrecen diversas opciones para analizar y presentar datos. Estas herramientas no solo ayudan a identificar tendencias y patrones, sino que también revelan información crucial que puede guiar la toma de decisiones informadas. En conjunto, la combinación de visualización de datos y métodos de *Learning Analytics* proporciona un enfoque completo para comprender y aprovechar el potencial de los datos en la era digital.

Metodología

Tipo de Investigación

El enfoque de la investigación es cuantitativo y con alcance descriptivo, debido a que su objetivo es identificar y especificar las propiedades y características de personas, grupos, procesos u otros fenómenos analizados. Este tipo de estudio se centra en medir o recolectar información sobre conceptos o variables, sin necesariamente explorar relaciones entre ellos. En este tipo de investigaciones, es crucial que el investigador pueda definir o al menos visualizar qué aspectos serán medidos y sobre qué o quiénes se recopilarán los datos, ya sea individuos, grupos, comunidades o eventos. En resumen, el enfoque descriptivo busca proporcionar una comprensión detallada y precisa de los fenómenos estudiados, sin enfocarse en establecer relaciones causales o de correlación (Hernández et al., 2014).

La investigación descriptiva, según Dulock (1993), se caracteriza por su enfoque cuantitativo, que se centra en recopilar datos cuantificables para su análisis estadístico, especialmente en estudios de mercado. Además, este tipo de investigación no manipula variables, utilizando métodos de observación para obtener información sin intervenir en el comportamiento de las variables. Por lo general, se trata de estudios transversales que examinan diferentes aspectos de un grupo en un momento específico. Los datos recopilados proporcionan una base sólida para investigaciones futuras, ayudando a orientar la selección de métodos de investigación para estudios posteriores.

Finalmente, se implementa la metodología Diseño *Scrum*, teniendo en cuenta su uso en la gestión de proyectos, especialmente en el desarrollo de software. Scrum se centra en la entrega de incrementos de producto de trabajo funcionales y de valor agregado de manera rápida y eficiente. El cual se alinea al objetivo principal del proyecto aplicado, al igual que la estructura

de esta, que es: inicio del proyecto, planificación, ejecución, revisión, ajuste, cierre y repetición del Sprint (Mariño y Alfonzo, 2014).

Diseño de Investigación

Para realizar la investigación se utiliza el conjunto de datos de la Escuela de Ciencias Jurídicas y Políticas - ECJP, que se almacena en el Sistema Integrado de Información - SII 5.0, en el aplicativo analítica de la Universidad Nacional Abierta y a Distancia – UNAD. Se acota la información para el programa de Ciencia Política, desde el año 2019 hasta el 2024¹, con el fin de organizar la *data* de los últimos cinco años. Lo que permite obtener un horizonte completo del desempeño académico desde su ingreso hasta su graduación. Es importante destacar que el programa de Ciencia Política inició sus actividades académicas con la recepción de sus primeros estudiantes en la cohorte 2019-2.

La herramienta que se utiliza para la normalización de los datos y el desarrollo del *Dashboard* es *Power BI*. Entendiendo que, es una plataforma integral para el análisis y la visualización de datos. Permite a los usuarios conectar, transformar y representar datos de manera efectiva. Con herramientas intuitivas, facilita la creación de informes interactivos, paneles personalizados y visualizaciones dinámicas. Además, ofrece capacidades avanzadas de análisis predictivo y empresarial. Es una herramienta esencial para la toma de decisiones informadas en organizaciones de todos los tamaños y sectores (Microsoft, 2021).

¹ Inicialmente se solicitaron los datos a la Escuela de Ciencias Básicas, Tecnología, e Ingeniería – ECBTI, pero no se obtuvo respuesta. Por tanto, se realizó el proyecto aplicado a la ECJP.

Calidad de la Información del SII 5.0 del Aplicativo Analítica

El uso de Python, en particular de la biblioteca Pandas, fue crucial para consolidar la información proveniente de los 114 archivos descargados del SII 5.0 del aplicativo Analítica, suministrados por la Escuela de Ciencias Jurídicas y Políticas (ECJP) de la UNAD.

El primer conjunto de datos contiene variables relacionadas con el ámbito académico, como permanencia, continuidad, graduados, ausentes y desempleo académico, entre otras. El segundo conjunto de datos corresponde a la caracterización socioeconómica de los estudiantes de nuevo ingreso al programa de Ciencia Política, recopilados en distintos periodos académicos.

Mediante el código desarrollado, se logró automatizar el proceso de unificación de los datos, gestionando las diferencias entre columnas y garantizando la correcta integración de la información.

Procesamiento de las Inconsistencias en los Archivos

Uno de los principales desafíos que se identificaron al trabajar con estos archivos fue la variabilidad en las columnas. A pesar de que los archivos tenían una estructura similar, algunos de ellos presentaban columnas adicionales o ciertos campos que estaban ausentes en otros documentos. Este tipo de inconsistencias pueden afectar la calidad y completitud de la información. Para resolver este problema, se utilizó la función `pd.concat()` de pandas, que permite combinar múltiples archivos y alinear las columnas, incluso cuando algunas columnas están ausentes en ciertos archivos.

Al usar el argumento `sort=False` en la función `concat()`, se evitó que pandas reordenara las columnas de forma alfabética, asegurando que las columnas comunes se alinearan correctamente y que aquellas que solo estaban presentes en ciertos archivos se mantuvieran sin alterar la estructura general del dataset consolidado.

Manejo de los Periodos Académicos

Un desafío significativo fue la falta de una columna específica que indicara el periodo académico de los estudiantes. Esto era importante ya que los datos recopilados cubrían varios periodos académicos, y el análisis del rendimiento y continuidad de los estudiantes necesitaba considerar las diferencias entre estos periodos. Para resolver este problema, se decidió agregar manualmente una columna que indicara el periodo correspondiente, basándose en los nombres de los archivos originales.

Se añadió una columna llamada "Periodo" en la cual se ingresó manualmente el periodo académico correspondiente a cada archivo. Esto fue un paso laborioso pero necesario, dado que los datos no contenían esta información de manera explícita.

Esta columna permitió segmentar los datos por periodos y facilitó el análisis longitudinal del rendimiento académico y la continuidad o deserción de los estudiantes a lo largo del tiempo.

Evaluación de la Calidad de los Datos

El análisis también incluyó la evaluación de la calidad de los datos en términos de duplicación y consistencia. Dado que los archivos reflejaban el estado del estudiante en diferentes momentos (por ejemplo, un estudiante podía aparecer en varios archivos con diferentes estados, como "Permanencia", "Ausencia" o "Graduado"), fue necesario gestionar la posible duplicación de información.

Para evitar la duplicación no deseada, se mantuvo el historial completo de los estados del estudiante, permitiendo una visión más rica y detallada del proceso académico de cada individuo. Sin embargo, para ciertos análisis (como la tasa de deserción por periodo), se filtraron los datos de manera que solo se considerara el estado más reciente o relevante para el periodo académico en cuestión.

Desafíos Específicos de los Periodos Académicos y su Manejo

El corte de información en diferentes periodos académicos fue un desafío central, ya que algunos estudiantes pueden cambiar de estado (por ejemplo, de "Permanente" a "Desertor") a lo largo de los periodos. Este cambio en el estado del estudiante es crucial para entender su evolución, pero puede complicar el análisis si no se manejan correctamente los cortes entre periodos.

Se segmentaron los datos en función de la nueva columna "Periodo" y se utilizó esta segmentación para analizar el comportamiento académico de los estudiantes en cada periodo de manera independiente.

Se calculó la tasa de deserción y continuidad por periodo académico, lo que permitió obtener una visión más clara del rendimiento en cada etapa y evaluar si existen patrones que se repiten en diferentes periodos.

Finalmente, la unificación de los datos mediante pandas y la cuidadosa incorporación de la columna "Periodo" permitieron garantizar una calidad aceptable de los datos para el análisis. A pesar de las inconsistencias iniciales en las columnas y la falta de ciertos campos clave, se lograron consolidar de manera efectiva, permitiendo llevar a cabo un análisis detallado del desempeño académico y los factores que influyen en la deserción y continuidad de los estudiantes de Ciencia Política en la UNAD (Ver apéndice A – Código en Python para la consolidación de datos).

Métricas Relacionadas con el Desempeño Académico

Con el conjunto de datos obtenido tras completar cada uno de los pasos descritos en la sección: calidad de la Información del SII 5.0 del aplicativo Analítica, se han identificado varias métricas clave que pueden aportar una visión integral del desempeño académico de los estudiantes. Estas métricas están basadas en datos demográficos, académicos y sociales, y ayudan a entender mejor los patrones de rendimiento y deserción. A continuación, se detallan las métricas más relevantes:

Tasa de Deserción y Continuidad por Periodo y Zona Geográfica

El análisis de la tasa de deserción por zona geográfica y periodo académico permite identificar patrones de deserción a lo largo del tiempo y entre diferentes regiones del país.

Estas métricas permiten identificar zonas con mayor deserción, lo que puede estar relacionado con factores como acceso a recursos educativos o condiciones socioeconómicas. Un hallazgo clave es que ciertas zonas, como la Zona Centro Bogotá Cundinamarca, presentan una mayor concentración de estudiantes, lo que podría influir en las tasas de deserción.

Referencia: En varios estudios, como el artículo *Designing and Implementing a Dashboard with Key Performance Indicators for a Higher Education Institution* de Azevedo et al. (2021), se destaca la importancia de evaluar la tasa de deserción en diferentes contextos geográficos para tomar decisiones informadas.

Análisis de Desempeño Académico por Corte Histórico

Esta métrica permite analizar la evolución del desempeño académico en cada corte académico, observando la proporción de estudiantes en categorías como "desempeño insuficiente", "aceptable" y "superior".

Lo anterior ayuda a evaluar si las estrategias implementadas por la universidad están teniendo un impacto positivo en el rendimiento de los estudiantes. También permite detectar si los estudiantes que desertan suelen tener un desempeño académico bajo antes de abandonar el programa.

Este enfoque está alineado con los estudios sobre la predicción del rendimiento académico como indicador clave para evaluar el éxito o fracaso, tal como se menciona en el trabajo Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería de Contreras et al. (2020).

Impacto del Estrato Socioeconómico en la Deserción

La información sobre el estrato socioeconómico es crucial para identificar cómo las condiciones económicas afectan la continuidad de los estudiantes en el programa.

Permite a la universidad diseñar políticas de apoyo financiero y académico más efectivas para los estudiantes de estratos bajos, que pueden tener mayores dificultades para continuar con sus estudios. Por ejemplo, los datos muestran una mayor proporción de estudiantes de estratos 1 y 2, lo que resalta la importancia de políticas de apoyo enfocadas en estos grupos.

El artículo Diseño e implementación de un *dashboard* de soporte académico basado en datos de entornos virtuales de aprendizaje de Haro (2017-2018) menciona cómo el análisis socioeconómico de los estudiantes permite identificar áreas donde se requieren intervenciones específicas.

Proporción de Género y Distribución por Edad

La proporción de género y la distribución de la edad al ingreso pueden influir en la tasa de deserción y en el desempeño académico.

Estas métricas permiten detectar si existen diferencias significativas entre grupos de estudiantes en términos de deserción o rendimiento. El análisis preliminar muestra una ligera sobrerrepresentación de estudiantes masculinos, lo cual puede tener implicaciones sobre el diseño de programas de apoyo específicos por género.

El estudio de Factores determinantes del rendimiento académico de los estudiantes de la Universidad de Atacama de Barahona (2014) destacan la importancia del análisis por género y edad para entender cómo estas variables afectan el desempeño de los estudiantes.

Desempeño Laboral vs Continuidad Académica

El *dataset* incluye información sobre la situación laboral de los estudiantes (empleado, independiente, desempleado), lo cual puede influir en su capacidad para continuar con los estudios.

Los estudiantes que trabajan, especialmente aquellos con empleos a tiempo completo, pueden tener una mayor probabilidad de desertar debido a la falta de tiempo para dedicar a sus estudios. Identificar esta relación permite a la universidad adaptar sus políticas de apoyo académico para ayudar a los estudiantes que trabajan.

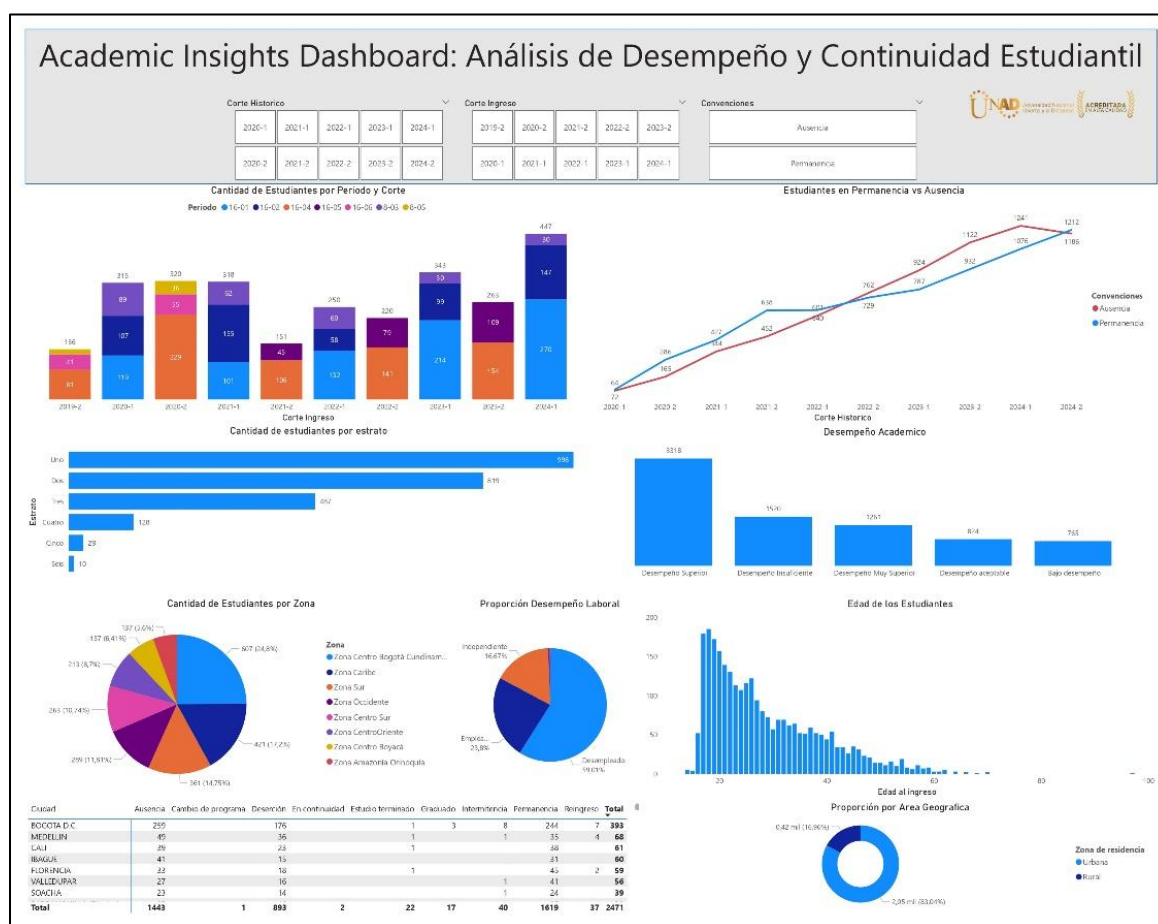
En la investigación *Designing and Implementing a Dashboard with Key Performance Indicators for a Higher Education Institution* de Azevedo et al. (2021) también señala la importancia de los datos laborales de los estudiantes para comprender las dinámicas de deserción, particularmente en aquellos que necesitan equilibrar trabajo y estudios.

Herramienta de Monitoreo y Análisis: *Dashboard* Interactivo

Luego de la normalización de los datos y elección de las métricas que se relacionan con el desempeño académico, se desarrolla un diseño visual intuitivo para el *dashboard* que facilita la comprensión de la información. En la Figura 1, muestra una serie de visualizaciones que resumen el análisis del desempeño académico y la continuidad estudiantil. Entre los gráficos incluidos se encuentran:

Figura 1

Dashboard Interactivo del Programa de Ciencia Política de la UNAD

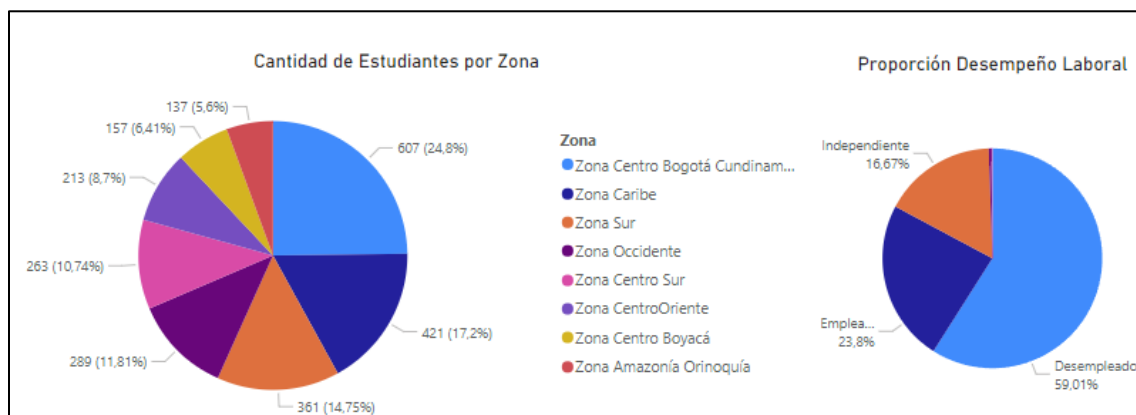


En la Figura 2, se observa un diagrama de torta que detalla el número de estudiantes distribuidos por diversas zonas, destacando Bogotá como la zona con mayor número de

estudiantes, seguido de otras regiones como la Zona Caribe, Zona Sur, y otras. También, está la proporción de desempeño laboral de los estudiantes, siendo la categoría desempleado la más grande con el 59.01% de la población estudiantil, lo que indica que más de la mitad de estos no tienen un empleo actualmente.

Figura 2

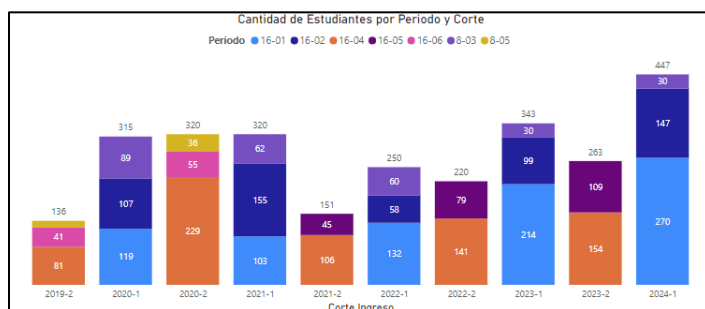
Cantidad de Estudiantes por Zona y Proporción Desempeño Laboral



La Figura 3 muestra la evolución del número de estudiantes en diferentes cortes de ingreso y periodos académicos, permitiendo observar tendencias en la matrícula a lo largo del tiempo.

Figura 3

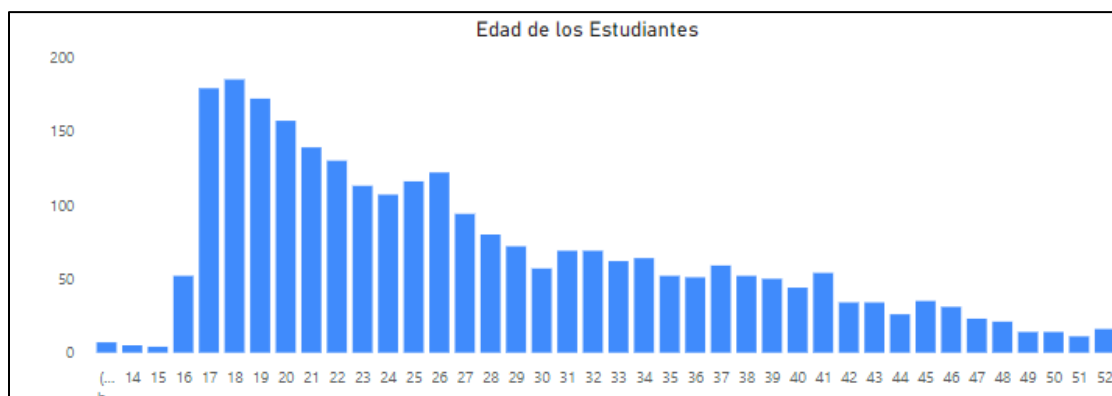
Cantidad de Estudiantes por Periodo



En la Figura 4 se identifica un diagrama de barras que muestra la distribución de la edad de los estudiantes al momento de ingresar, con intervalos que abarcan diversas edades.

Figura 4

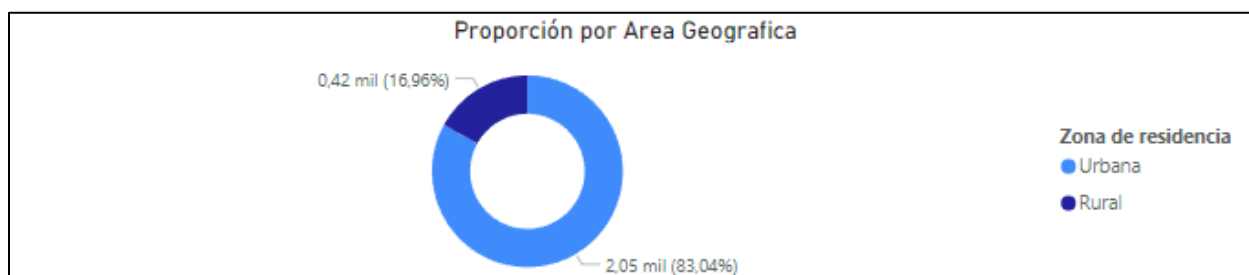
Edad de los Estudiantes



En la Figura 5, se puede observar un gráfico de anillo que ilustra la proporción de estudiantes por área geográfica, destacando una mayoría significativa de estudiantes urbanos frente a rurales.

Figura 5

Proporción por Área Geográfica



El diagrama de barras desglosa la cantidad de estudiantes según su estrato, con la mayor parte concentrada en los estratos 1, 2 y 3 (ver figura 6).

Figura 6

Cantidad de Estudiantes por Estrato

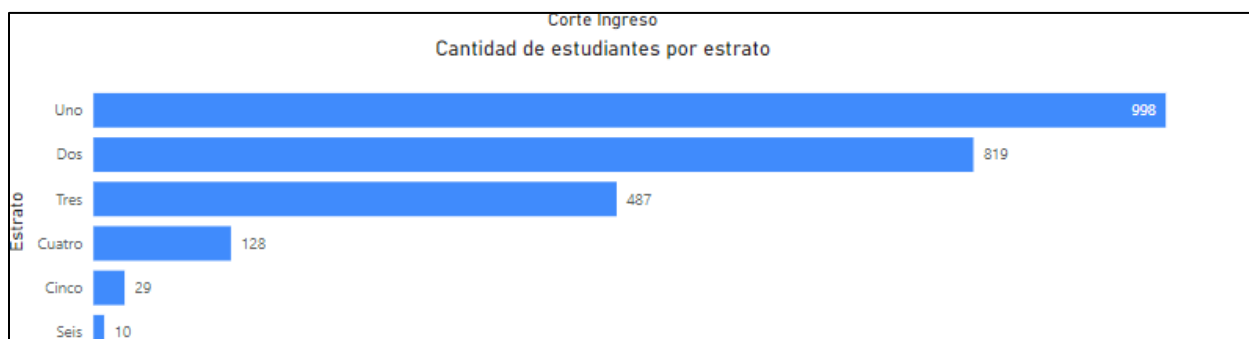
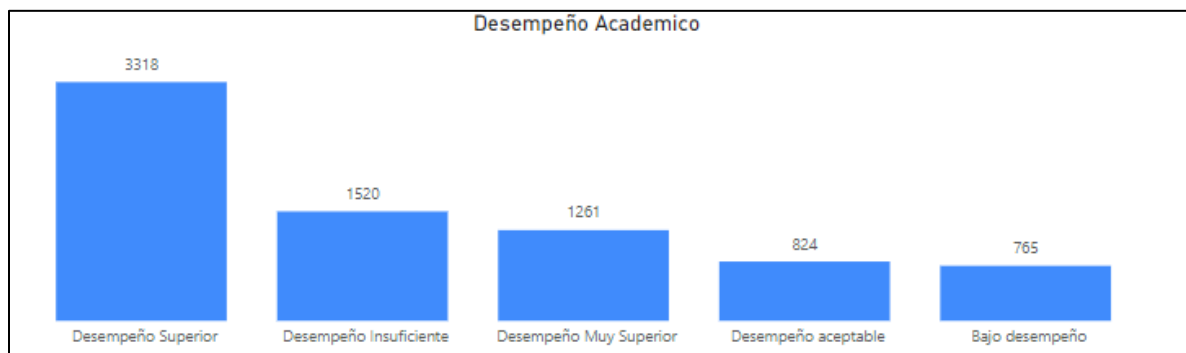


Diagrama de barras que desglosa la cantidad de estudiantes según su desempeño académico, donde se identifica que la mayoría de los estudiantes tienen un desempeño positivo (Superior y Muy Superior), pero hay una proporción significativa en las categorías de Desempeño Insuficiente y Bajo Desempeño que necesitan atención.

Figura 7

Desempeño Académico de los Estudiantes del Programa de Ciencia Política



Para una visualización detallada del *dashboard*, ingrese al siguiente enlace:

[Archivo de *dashboard* Final](#)

Principales Resultados Derivados del *Dashboard*

La Zona Centro Bogotá Cundinamarca concentra el mayor porcentaje de estudiantes, con un 24,8%, seguida de la Zona Caribe (17,2%) y la Zona Sur (14,75%). Esto indica que la

mayoría de los estudiantes provienen de zonas urbanizadas con mayor acceso a la educación. En contraste, la Zona Amazonía Orinoquía presenta la menor representación, con un 5,6%, lo que podría deberse a factores geográficos o socioeconómicos que limitan el acceso a la educación en esa región.

En cuanto al nivel socioeconómico, la mayoría de los estudiantes pertenecen a los estratos 1 y 2, con 998 y 819 estudiantes respectivamente. Esto refleja que una gran parte de la población estudiantil enfrenta condiciones socioeconómicas adversas, lo que podría influir en la tasa de deserción si no se implementan políticas de apoyo adecuadas.

El análisis del empleo revela que el 59,01% de los estudiantes está desempleado, mientras que un 23,8% tiene empleo y el 16,67% trabaja de forma independiente. Este alto índice de desempleo puede impactar negativamente en la continuidad académica, resaltando la importancia de políticas de apoyo financiero o laboral que ayuden a mejorar la retención estudiantil.

Los datos de desempeño académico muestran una distribución notable entre "desempeño aceptable" y "bajo desempeño". Aunque existe un grupo de estudiantes con desempeño superior y muy superior, una proporción significativa presenta un desempeño insuficiente, lo que podría estar relacionado con la tasa de deserción en ciertos periodos. Es importante analizar si existe una correlación entre el bajo desempeño académico y el abandono estudiantil para abordar el problema de manera efectiva.

Finalmente, en la Figura 1, específicamente en la gráfica de Permanencia vs. Ausencia por corte histórico, se observa una fluctuación en los últimos periodos. Los cortes más recientes muestran un aumento en la cantidad de estudiantes en situación de ausencia (abandono o intermitencia). Por ejemplo, en el corte 2024-1, se nota una disminución significativa en la

permanencia en comparación con periodos anteriores, lo que sugiere la necesidad de intervenciones rápidas y efectivas para reducir la deserción

Conclusiones

La librería Pandas de Python, empleada en la fase inicial del proyecto aplicado titulado “Herramienta de monitoreo y análisis: *Dashboard* interactivo del desempeño académico de los estudiantes de Ciencia Política en la UNAD”, facilitó la consolidación y normalización de los datos provenientes de 114 archivos del SII 5.0 del aplicativo Analítica. Esto permitió abordar la falta de consistencia entre columnas y la ausencia de una categorización clara para los periodos académicos. Como resultado, se automatizó el proceso mediante un código desarrollado específicamente, logrando integrar información sobre permanencia, graduados, ausentes y deserción académica, así como la caracterización socioeconómica de los estudiantes.

A partir de esta integración, se definieron métricas clave que proporcionaron una visión integral del desempeño académico de los estudiantes. Entre estas métricas se destacan: la tasa de deserción por zona geográfica y periodo académico, el impacto del estrato socioeconómico en la continuidad académica, la relación entre la situación laboral y el rendimiento académico, y la distribución de estudiantes según género y edad al ingreso. Estos indicadores resultan fundamentales para evaluar tanto el desempeño académico como la eficiencia de la herramienta desarrollada.

Finalmente, cabe resaltar que el diseño y construcción del *dashboard*, caracterizado por sus visualizaciones intuitivas, permitió resumir información relevante sobre el rendimiento y las tendencias de los estudiantes analizados. Esta funcionalidad facilita la toma de decisiones articulada y basada en datos, contribuyendo al logro eficiente de objetivos y planes de acción.

Recomendaciones

El desarrollo del proyecto aplicado tuvo como valor agregado la integración de la información del SII 5.0 con el aplicativo Analítica y la construcción de un panel de control interactivo sobre el desempeño académico de los estudiantes de Ciencia Política en la UNAD. No obstante, durante su ejecución se identificaron ciertos vacíos y falencias que podrían abordarse como una extensión de este trabajo. A continuación, se presentan las principales recomendaciones derivadas de estos hallazgos:

1. Estandarización de formatos: Implementar un formato uniforme para los archivos descargados del SII 5.0, garantizando que todos los campos requeridos estén presentes y estandarizados. Se recomienda incluir una columna obligatoria para el “Periodo Académico” que facilite la organización y el análisis de los datos.
2. Monitoreo continuo de métricas clave: Realizar un seguimiento constante de indicadores clave, como las tasas de deserción y el desempeño académico, mediante un panel de control actualizado periódicamente. Esto permitirá identificar tendencias emergentes y responder de manera oportuna.
3. Análisis avanzados de factores contextuales: Profundizar en el estudio del impacto de variables como la situación laboral y las diferencias de género en el rendimiento académico y la deserción. Este enfoque permitirá diseñar intervenciones más efectivas y específicas.
4. Fomento de la inclusión regional: Diseñar estrategias que incrementen la representación de estudiantes en regiones con menor acceso a la educación superior, como la Zona Amazonía-Orinoquía, asegurando así una mayor equidad en la cobertura educativa.

Referencias Bibliográficas

- Aldana, K., de Roberti, R. P., y Miranda, A. R. (2010). *Visión del desempeño académico estudiantil en la Universidad Centroccidental Lisandro Alvarado*. *Compendium*, 13(24), 5-21. <https://www.redalyc.org/articulo.oa?id=88017685002>
- Apraxine, D., y Stylianou, E. (2017). *Business intelligence in a higher educational institution: The case of University of Nicosia*. IEEE Global Engineering Education Conference (EDUCON) (pp. 1735-1746).
- Azevedo, A., Azevedo, J. M., & Hayakawa, M. E. (2021). *Designing and Implementing a Dashboard with Key Performance Indicators for a Higher Education Institution*. *CSEDU* (1) (pp. 165-172).
- Barahona, P. (2014). *Fatores determinantes do rendimento acadêmico dos estudantes da Universidade de Atacama*. *Estudios pedagógicos (Valdivia)*, 40(1), 25-39. <https://dx.doi.org/10.4067/S0718-07052014000100002>
- Bocanegra, S. (2019). *Uso del Dashboard digital para el monitoreo de indicadores de las Unidades de Investigación de la Universidad Nacional de San Martín-Tarapoto*. [Tesis de grado, Universidad Nacional de San Martín - Tarapoto]. Repositorio institucional. <http://hdl.handle.net/11458/3253>
- Cardoso, E., & Su, X. (2022). *Designing a Business Intelligence and Analytics Maturity Model for Higher Education: A Design Science Approach*. *Applied Sciences*, 12(9), 4625. <https://doi.org/10.3390/app12094625>
- Castillo, J., González, A., y Muñoz, L. (2018). *Inteligencia de Negocios como apoyo a Sistemas de Información de Egresados de Instituciones de Educación Superior*. *Memorias De Congresos UTP*, 1(1), 81-88

- Contreras, L. E., Fuentes, H. J., y Rodríguez, J. I. (2020). *Predicción del rendimiento académico como indicador de éxito/fracaso de los estudiantes de ingeniería, mediante aprendizaje automático*. *Formación universitaria*, 13(5), 233-246. <http://dx.doi.org/10.4067/S0718-50062020000500233>
- Clow, D. (2013). *An overview of learning analytics*. *Teaching in Higher Education*, 18(6), 683–695. <https://doi.org/10.1080/13562517.2013.827653>
- Delgado, P., Raúl, J., y Palos, P. A. (2007). *Desempeño académico y conductas de riesgo en adolescentes*. *Revista de educación y desarrollo*, 7(1), 5-16. https://www.cucs.udg.mx/revistas/edu_desarrollo/anteriores/7/007_Palacios.pdf
- Destiandi, N., y Hermawan, A. (2018). *Business Intelligent Method For Academic Dashboard*. *bit-Tech: Binary Digital-Technology*, 1(2), 55-64.
- Dulock, H. L. (1993). *Research Design: Descriptive Research*. *Journal of Pediatric Oncology Nursing*, 10(4), 154–157. <https://doi.org/10.1177/104345429301000406>
- Few, S. (2004). *Dashboard Confusion*. Perceptual Edge. https://www.perceptualedge.com/articles/ie/dashboard_confusion.pdf
- Few, S. (2006). *Information dashboard design: The effective use of data visualization to create a dashboard that tells a story*. Sebastopol, CA: O'Reilly Media. <http://public.magendanz.com/Temp/Information%20Dashboard%20Design.pdf>
- Forero-Corba, W., y Bennasar, F. N. (2024). *Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación: una revisión sistemática*. *RIED-Revista Iberoamericana De Educación a Distancia*, 27(1), 1-34. <https://doi.org/10.5944/ried.27.1.37491>

- Friendly, M. (2000). *Visualizing Categorical Data: Data, Stories, and Pictures*.
https://www.researchgate.net/publication/2431674_Visualizing_Categorical_Data_Data_Stories_and_Pictures
- Gutiérrez Villaverde, H. E., Linares Barbero, M., Agüero Correa, Á. A., y Pérez Núñez, J. R. (2022). *Predicción de rendimiento académico de alumnos usando machine learning*.
<https://repositorio.ulima.edu.pe/handle/20.500.12724/16703>
- Haro, V. (2018). *Diseño e implementación de un dashboard de soporte académico basado en datos de entornos virtuales de aprendizaje*. . [Tesis de Maestría, Universitat Politècnica de València]. Repositorio institucional. <http://hdl.handle.net/10251/111761>
- Hernández, R., Fernández, C. y Baptista, P. (2014). *Metodología de la investigación*. McGraw-Hill Interamericana. <https://www.esup.edu.pe/wp-content/uploads/2020/12/2.%20Hernandez,%20Fernandez%20y%20Baptista-Methodolog%C3%ADa%20Investigacion%20Cientifica%206ta%20ed.pdf>
- Mariño, S. I., & Alfonzo, P. L. (2014). *Implementación de SCRUM en el diseño del proyecto del Trabajo Final de Aplicación*. *Scientia Et Technica*, 19(4), 413-418.
<https://www.redalyc.org/pdf/849/84933912009.pdf>
- Medina, J., Pinzón, K., y Salazar, Y. (2021). *Determinantes del Rendimiento Académico de los Estudiantes de una Universidad Pública Ecuatoriana*. *Revista Politécnica*, 47(2), 53-62.
<https://doi.org/10.33333/rp.vol47n2.05>
- Microsoft. (2021). *¿Qué es Power BI? Microsoft Power BI*. <https://powerbi.microsoft.com/es-es/what-is-powerbi/>
- Mora-Vicarioli, F. R., Arce-Solano, J. L. ., Padilla-Romero , K. ., y Muñiz-Umaña , G. . (2021). *Implementación de un sistema de inteligencia de negocios*. *Escuela de Ciencias de la*

- Administración UNED. Revista Electrónica Calidad En La Educación Superior*, 12(1), 76–103
- Navarro, R. E. (2003). *El rendimiento académico: concepto, investigación y desarrollo*. REICE. Revista Iberoamericana sobre calidad, eficacia y cambio en educación, 1(2).
<https://www.redalyc.org/articulo.oa?id=55110208>
- Pértega Díaz, S., y Pita Fernández, S. (2001). *Representación gráfica en el análisis de datos*. Cad Aten Primaria, 8, 112-117.
https://wikieducator.org/images/6/66/Representaci%C3%B3n_gr%C3%A1fica_en_el_An%C3%A1lisis_de_Datos.pdf
- Rico-Páez, A., & Gaytan-Ramírez, N. D. (2022). *Modelos predictivos del rendimiento académico a partir de características de estudiantes de ingeniería*. IE Revista de Investigación Educativa de la REDIECH, 13, 1-18.
https://doi.org/10.33010/ie_rie_rediech.v13i0.1426
- Sarikaya, A., Correll, M., Bartram, L., Tory, M., y Fisher, D. (2018). *What do we talk about when we talk about dashboards?* IEEE transactions on visualization and computer graphics, 25(1), 682-692. 10.1109/TVCG.2018.2864903
- Sellán, J. (2021). *Implementación de Business Intelligence sobre la plataforma de aprendizaje Moodle para determinar el comportamiento de los estudiantes de una institución de educación superior*. [Tesis de grado]. Escuela Superior Politécnica del Litoral.
- Siemens, G. (2013). *Learning analytics: The power of data for improving learning*. Emerald Group Publishing Limited.

- Soria, K., y Zúñiga, S. (2014). *Aspectos determinantes del éxito académico de estudiantes universitarios*. *Formación universitaria*, 7(5), 41-50. <https://dx.doi.org/10.4067/S0718-50062014000500006>
- Stewart, C. L., & Dewan, M. A. A. (2022). *A Systemic Mapping Study of Business Intelligence Maturity Models for Higher Education Institutions*. *Computers*, 11(1), 153. <https://doi.org/10.3390/computers11110153>
- Tableau. (2023). *¿Qué es la visualización de datos?* <https://www.tableau.com/products/desktop>
- Vélez, E. J. (2023). *Diseño e implementación de un prototipo de Dashboard para seguimiento académico de los estudiantes de la Facultad de Ingeniería de la Universidad Católica de Santiago de Guayaquil*. [Tesis de grado, Universidad Católica de Santiago de Guayaquil]. Repositorio institucional. <http://repositorio.ucsg.edu.ec/handle/3317/20851>
- Universidad Nacional Abierta y a Distancia. (2018, 17 de agosto). *La UNAD se constituye en la primera megauniversidad pública de Colombia*. <https://noticias.unad.edu.co/index.php/unad-noticias/todas/2362-la-unad-se-constituye-en-la-primera-megauniversidad-publica-de-colombia>
- Universidad Nacional Abierta y Distancia. (2013). *Reglamento Estudiantil. Capítulo 8. De las opciones de grado*. En Acuerdo 0029 del 13 de diciembre de 2013. https://sgeneral.unad.edu.co/images/documentos/consejoAcademico/acuerdos/2013/COAC_ACUE_029_20131213.pdf
- Valero, J. E., Navarro, Á. F., Larios, A. C., & Julca, J. D. (2022). *Deserción universitaria: Evaluación de diferentes algoritmos de Machine Learning para su predicción*. *Revista de ciencias sociales*, 28(3), 362-375. <https://dialnet.unirioja.es/servlet/articulo?codigo=8526463>

Wilkinson, L., Anand, A., y Odhner, R. (2005). *Visualizing data with Tableau*. Sebastopol, CA:
O'Reilly Media.

Apéndices

Apéndice A

Código en Python para la Consolidación de Datos

Para la unificación de los 114 archivos Excel utilizados en este estudio, se utilizó Python como herramienta de automatización, lo que permitió combinar los datos de manera eficiente y estructurada. El siguiente código implementado en Google Colab facilitó el proceso de integración de los datos:

```
import pandas as pd
import glob
from google.colab import drive

# Monta Google Drive
drive.mount('/content/drive')

# Define la carpeta donde están los archivos, asegurando que la ruta termine con una barra
folder_path = '/content/drive/MyDrive/otros datos/' # Usa glob para encontrar todos los archivos
Excel en la carpeta
files = glob.glob(folder_path + "*.xlsx")

# Verificar si se han encontrado archivos
print("Archivos encontrados:", files)

if len(files) == 0:
    raise ValueError("No se encontraron archivos Excel en la carpeta especificada.")

# Crea una lista vacía para almacenar los DataFrames
df_list = []

# Itera sobre los archivos y los lee
for file in files:
    # Lee el archivo en un DataFrame
    df = pd.read_excel(file)

    # Agrega el nombre del archivo como una nueva columna
    df['Nombre del archivo'] = file.split("/")[-1] # Solo toma el nombre del archivo, no la ruta completa
    # Añade el DataFrame a la lista
    df_list.append(df)

# Combina todos los DataFrames en uno solo, alineando las columnas
combined_df = pd.concat(df_list, ignore_index=True, sort=False)

# Guarda el DataFrame combinado en un nuevo archivo
```

```
Excelcombined_df.to_excel('/content/drive/MyDrive/Datos/archivo_combinado.xlsx',  
index=False)  
print("Archivos combinados exitosamente.")
```