

**Análisis predictivo y de perfilamiento de patrocinadores en la ONG World Vision  
Colombia durante el segundo semestre del año 2024**

William Ricardo Berbeo Buitrago

Asesor

Andrea Ximena Robles Uriza

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

2024

## Resumen

World Vision es una organización de ayuda humanitaria que ha trabajado en Colombia por más de 40 años a través de apoyo internacional y nacional de patrocinadores que ayudan a financiar proyectos de sostenibilidad y ayuda a comunidades vulnerables del país. El proyecto Análisis predictivo y de perfilamiento de patrocinadores en la ONG World Vision Colombia durante el segundo semestre del año 2024 aborda la falta de herramientas analíticas avanzadas para comprender y predecir el comportamiento de los patrocinadores. El problema se centra en identificar edades clave para el patrocinio, determinar la duración del compromiso, predecir cancelaciones y anticipar mayores compromisos. Usando técnicas de análisis de datos y modelos predictivos, el proyecto permitirá mejorar la gestión de patrocinadores, optimizar recursos y aumentar el impacto de los programas. La metodología incluye la recopilación y limpieza de datos históricos y actuales, la segmentación de datos y la implementación de modelos como regresión logística y árboles de decisión. Se desarrollarán estrategias de marketing personalizadas y planes de retención basados en datos, mejorando la toma de decisiones estratégicas.

Enmarcado en la línea de investigación de desarrollo tecnológico de la ECBTI, este proyecto contribuirá al avance tecnológico de World Vision Colombia, permitiendo decisiones basadas en datos que optimicen la captación y retención de patrocinadores y maximicen el impacto de sus programas.

***Palabras clave:*** Patrocinadores, ONG, Retención, Cancelaciones, Marketing.

## **Abstract**

World Vision is a humanitarian aid organization that has been operating in Colombia for over 40 years, supported by international and national sponsors who fund sustainability projects and provide assistance to vulnerable communities across the country. The project Predictive Analysis and Profiling of Sponsors in the NGO World Vision Colombia during the Second Semester of 2024 addresses the lack of advanced analytical tools to understand and predict sponsor behavior. The core issues include identifying key ages for sponsorship, determining commitment duration, predicting cancellations, and anticipating higher engagement levels. By leveraging data analysis techniques and predictive models such as logistic regression and decision trees, the project aims to enhance sponsor management, optimize resources, and amplify program impact. The methodology involves collecting and cleaning historical and current data, segmenting datasets, and implementing data-driven strategies for personalized marketing and retention plans.

Framed within the technological development research line of ECBTI, this project will contribute to the technological advancement of World Vision Colombia, enabling data-driven decision-making to optimize sponsor acquisition and retention while maximizing the impact of its programs.

***Keywords:*** Sponsors, NGO, Retention, Cancellations, Marketing.

## Tabla de Contenido

Introducción .....	8
Planteamiento del Problema .....	9
Justificación .....	11
Objetivos .....	12
Objetivo General .....	12
Objetivos Específicos .....	12
Marco Conceptual .....	13
Análisis Predictivo .....	13
Perfilamiento de Usuarios .....	13
Big Data y Análisis de Datos .....	13
Marco Teórico .....	14
Metodología .....	15
Recopilación de Datos .....	15
Pre Procesamiento de Datos .....	15
Modelos Predictivos .....	15
Análisis de Resultados .....	16
Implementación de Estrategias .....	16
Organización, Depuración y Evaluación de los Datos Recopilados (Recopilación de Datos) .....	17
Caracterización y Clasificación (Procesamiento de Datos) .....	18
Tipo Numérico .....	18
Tipo Categórico .....	18
Tipo Fecha .....	19

Características Descartables de los Datos Obtenidos.....	19
Representaciones Gráficas y Métricas de los Datos .....	21
Método de Predicción para la Conversión de Patrocinadores de Unipatrocinadores a Multipatrocinadores .....	29
Segmentación de Patrocinadores por Árboles de Decisión por Variables Numéricas.....	36
Segmentación de Patrocinadores por Árboles de Decisión por Variables Categóricas .....	40
Modelo Predictivo de Conversión de Patrocinadores Unipatro a Multipatro con Regresión Logística .....	41
Análisis de Resultados .....	47
Conclusiones .....	55
Recomendaciones .....	57
Bibliografía .....	58

## Lista de Tablas

<b>Tabla 1</b> <i>Recopilación de Datos SIMMA</i> .....	17
<b>Tabla 2</b> <i>Datos Obtenidos después del Proceso de Imputación y Normalización</i> .....	20
<b>Tabla 3</b> <i>Resumen de Clusters Numéricos</i> .....	34
<b>Tabla 4</b> <i>Resumen de Clusters Categóricos</i> .....	38
<b>Tabla 5</b> <i>Resultado Promedio Validación Cruzada-Kfold M. Unipatro a Multipatro</i> .....	41
<b>Tabla 6</b> <i>Resultado Rendimiento en la Clasificación de Patrocinadores ( Unipatro vs Multipatro)</i> .....	42
<b>Tabla 7</b> <i>Matriz de Confusión Unipatro vs Multipatro</i> .....	43
<b>Tabla 8</b> <i>Resultado Promedio Validación Cruzada-Kfold</i> .....	44
<b>Tabla 9</b> <i>Patrocinadores con Probabilidad de Cancelación</i> .....	45
<b>Tabla 10</b> <i>Matriz de Confusión Patrocinadores con Probabilidad de Cancelación</i> .....	45

## Tabla de Figuras

<b>Figura 1</b> <i>Relación de Patrocinadores con Estado Civil y Categoría Unimulti</i> .....	21
<b>Figura 2</b> <i>Mapa de Calor de Motivación vs Clasificación de Patrocinadores</i> .....	22
<b>Figura 3</b> <i>Mapa de Calor de Compromisos Asignados por Regiones</i> .....	23
<b>Figura 4</b> <i>Porcentaje de Patrocinadores por Edad de los Niños</i> .....	24
<b>Figura 5</b> <i>Cantidad de Ocurrencias de Patrocinadores Clasificados por su Edad</i> .....	25
<b>Figura 6</b> <i>Métodos de Pago por el Tipo de Patrocinador</i> .....	26
<b>Figura 7</b> <i>Métodos de Pago por el Tipo de Patrocinador</i> .....	27
<b>Figura 8</b> <i>Matriz de Correlación de Variables Numéricas</i> .....	29
<b>Figura 9</b> <i>Matriz de Correlación de Variables Categóricas</i> .....	31
<b>Figura 10</b> <i>Clusters por Variables Numéricas</i> .....	33
<b>Figura 11</b> <i>Árbol de Decisión para Variables Numéricas</i> .....	36
<b>Figura 12</b> <i>Clusters por Variables Categóricas</i> .....	38
<b>Figura 13</b> <i>Árbol de Decisión para Variables Categóricas</i> .....	40

## Introducción

La gestión de patrocinadores es un pilar fundamental para el funcionamiento y la sostenibilidad de los programas sociales en la ONG World Vision Colombia. Sin embargo, la falta de herramientas analíticas avanzadas ha limitado la capacidad de la organización para entender y predecir el comportamiento de los patrocinadores, afectando la retención y la captación de nuevos apoyos. Este desafío se vuelve particularmente relevante en un contexto donde el apoyo constante de los patrocinadores es esencial para mantener el impacto positivo en las comunidades vulnerables de Colombia.

Ante esta situación, surge la necesidad de un enfoque que aproveche el potencial de la ciencia de datos para transformar el proceso de gestión de patrocinadores. Este proyecto desarrolla modelos predictivos y de perfilamiento que permiten identificar patrones de comportamiento de los patrocinadores, anticipar posibles cancelaciones y descubrir oportunidades para fortalecer el compromiso de aquellos con mayor potencial de contribución. Mediante el análisis de datos históricos y actuales, se busca proporcionar a la organización una herramienta que mejore la toma de decisiones y optimice la distribución de recursos.

La propuesta se enmarca en la línea de proyecto aplicado, enfocado en la integración de técnicas de análisis de datos para crear soluciones prácticas y eficientes. Esto permitirá a World Vision Colombia no solo mejorar la retención de patrocinadores, sino también desarrollar estrategias personalizadas, incrementando así la efectividad de sus campañas de marketing y de recaudación de fondos. En última instancia, la implementación de este proyecto tiene el potencial de fortalecer la capacidad de la organización para generar un impacto sostenible y duradero en las comunidades que atiende, asegurando un uso óptimo de los recursos y una mayor efectividad en sus programas sociales.

## **Planteamiento del Problema**

La gestión de patrocinadores en World Vision Colombia enfrenta un reto fundamental la falta de herramientas analíticas avanzadas para comprender de manera integral el comportamiento de los patrocinadores. Este vacío analítico ha generado dificultades para predecir tendencias críticas como la duración del compromiso de los patrocinadores, la probabilidad de cancelación de sus aportes y la identificación de oportunidades para aumentar su nivel de compromiso. Como consecuencia, la organización enfrenta riesgos significativos en términos de pérdida de apoyo financiero y disminución de la capacidad para sostener programas de alto impacto social a largo plazo.

La dificultad para identificar patrones de comportamiento de los patrocinadores no solo limita la capacidad de retención, sino que también impide anticiparse a posibles cancelaciones. Sin un análisis adecuado, la organización no puede tomar medidas proactivas que permitan retener a los patrocinadores y asegurar la estabilidad de los recursos que financian los programas sociales. Esto se traduce en una subutilización de los datos disponibles y en una gestión reactiva en lugar de una estrategia preventiva, lo que reduce la efectividad de las campañas de captación y fidelización de patrocinadores.

Además, World Vision Colombia carece de un modelo de análisis que permita identificar las características demográficas y de comportamiento de aquellos patrocinadores con mayor probabilidad de ampliar su compromiso o involucrarse en nuevas iniciativas. Sin estas herramientas, la organización pierde la oportunidad de optimizar sus esfuerzos de marketing y de dirigir mejor sus estrategias hacia los segmentos más receptivos.

El problema central que aborda este proyecto se centra en la necesidad de dotar a la organización de un enfoque predictivo que le permita transformar sus datos históricos y actuales

en información valiosa para la toma de decisiones. La implementación de modelos predictivos y de perfilamiento no solo mejorará la capacidad de la organización para anticiparse a las cancelaciones, sino que también permitirá a la organización diseñar estrategias personalizadas para los patrocinadores, maximizando así la retención y el impacto de las donaciones. Este proyecto busca, por tanto, transformar la gestión de patrocinadores de World Vision Colombia en un proceso más eficiente, estratégico y orientado a resultados, asegurando una mayor sostenibilidad de sus iniciativas en beneficio de las comunidades atendidas.

## **Justificación**

La gestión de patrocinadores es un aspecto crucial para la sostenibilidad de los programas de World Vision Colombia. Sin embargo, la falta de herramientas analíticas avanzadas limita la capacidad de la organización para entender y predecir el comportamiento de los patrocinadores. Esta carencia impacta directamente en la retención de patrocinadores y en la captación de nuevos apoyos, lo que a su vez compromete la efectividad de los programas sociales que dependen de estas contribuciones. Sin una comprensión clara de las dinámicas de patrocinio, la organización enfrenta el riesgo de perder valiosas oportunidades para optimizar sus recursos y asegurar el financiamiento a largo plazo.

La justificación de este proyecto radica en la necesidad de implementar un enfoque predictivo y de perfilamiento que permita anticipar cancelaciones, identificar edades clave para el patrocinio y detectar oportunidades de mayor compromiso por parte de los patrocinadores. Al contar con modelos de análisis que ofrezcan una visión más precisa del comportamiento de los patrocinadores, World Vision Colombia puede tomar decisiones estratégicas más acertadas y diseñar campañas de retención más efectivas. Esto no solo incrementa la eficiencia en la gestión de recursos, sino que también potencia el impacto social de la organización, al garantizar que los programas lleguen a las comunidades que más lo necesitan.

Finalmente, la implementación de esta iniciativa contribuirá a la modernización de la gestión de datos dentro de la organización, fortaleciendo la capacidad de World Vision para adaptarse a un entorno cada vez más orientado al uso de tecnologías de la información y análisis de datos. Este avance no solo beneficiará a la organización en el corto plazo, sino que también consolidará una base sólida para el desarrollo de estrategias futuras basadas en datos, asegurando la sostenibilidad y el crecimiento de sus iniciativas sociales en Colombia.

## **Objetivos**

### **Objetivo General**

Crear un modelo predictivo para mejorar la gestión de patrocinadores de World Vision Colombia, mediante el análisis de datos históricos y actuales, que permita optimizar la retención de patrocinadores y captar nuevos compromisos durante el segundo semestre del año 2024.

### **Objetivos Específicos**

Crear un modelo predictivo para la identificación temprana de cancelaciones de patrocinio, con el fin de permitir la implementación de acciones preventivas que reduzcan la deserción.

Identificar oportunidades de crecimiento en el compromiso de los patrocinadores, analizando los patrones de comportamiento que indiquen la disposición de los patrocinadores actuales a aumentar sus contribuciones o involucrarse en nuevos proyectos.

Implementar herramientas de visualización de datos que faciliten la interpretación de los resultados obtenidos a través de los modelos.

## **Marco Conceptual**

### **Análisis Predictivo**

El análisis predictivo hace referencia al uso de datos históricos y técnicas analíticas avanzadas, como modelos estadísticos, algoritmos de aprendizaje automático y minería de datos, para predecir comportamientos o eventos futuros. Según Hernández et al. (2018), el análisis predictivo identifica patrones y tendencias que facilitan la toma de decisiones estratégicas.

### **Perfilamiento de Usuarios**

Este término describe el proceso de segmentación de una población en función de características específicas, tales como demografía, comportamiento o historial de interacciones. Según Brenda Díaz-Landa et al. (2021), el perfilamiento permite diseñar estrategias personalizadas, maximizando el impacto de las campañas.

### **Big Data y Análisis de Datos**

La capacidad de manejar grandes volúmenes de datos (Big Data) es esencial para proyectos analíticos. GAIA Group (2017) destaca cómo estas tecnologías permiten procesar datos complejos para identificar patrones relevantes y optimizar decisiones. Retención y Fidelización de Patrocinadores, estrategias orientadas a reducir la deserción de patrocinadores mediante el uso de análisis predictivos y personalización. Según Urrutia Aguilar et al. (2021), estas metodologías son cruciales para mantener la sostenibilidad y compromiso en las organizaciones.

## Marco Teórico

- Modelos Predictivos para la Retención

Oswaldo Sifuentes Bitocchi (2018) desarrolló modelos para anticipar deserciones en estudiantes, técnica que se adapta para predecir la cancelación de patrocinios. Los algoritmos de clasificación, como la regresión logística, permiten identificar señales tempranas de abandono.

- Segmentación y Perfilamiento de Usuarios:

Rodríguez Sabiote et al. (2017) demostraron que la segmentación basada en características específicas optimiza estrategias dirigidas, lo cual es esencial para identificar grupos de patrocinadores con mayor potencial de compromiso.

- Optimización de Recursos mediante Big Data:

La integración de herramientas de Big Data, como lo expone GAIA Group (2017), mejora la eficiencia al identificar patrones en grandes volúmenes de datos. Esto permite diseñar campañas más efectivas y asignar recursos estratégicamente.

- Impacto del Análisis Predictivo en ONGs:

Brenda Díaz-Landa et al. (2021) muestran cómo el análisis correlacional predictivo puede adaptarse para evaluar y mejorar el impacto de estrategias en ONGs, aumentando el compromiso de los patrocinadores.

- Justificación y Relevancia Tecnológica

La implementación de estas técnicas no solo optimiza los recursos de World Vision Colombia, sino que también representa un avance tecnológico en su gestión, alineándose con tendencias globales en el uso de datos para decisiones estratégicas (Hernández et al., 2018).

## **Metodología**

### **Recopilación de Datos**

Fuente de Datos: Se utilizarán bases de datos internas en SQL Server de World Vision Colombia, organizadas en tablas por patrocinadores. Además, se integrarán datos históricos de interacciones, encuestas a patrocinadores y reportes de actividades previas.

VARIABLES Consideradas: Edad, género, duración del patrocinio, frecuencia de donaciones, historial de cancelaciones, participación en eventos, métodos de pago, motivaciones, designaciones, región, ocupación, profesión y tipo de patrocinador.

Objetivo: Garantizar que la base de datos sea representativa y adecuada para los análisis posteriores.

### **Pre Procesamiento de Datos**

- Limpieza de Datos: Eliminación de valores atípicos, duplicados y registros incompletos.

- Transformación de Datos:

- Normalización de variables continuas para evitar sesgos.

- Codificación de variables categóricas utilizando técnicas como One-Hot

Encoding.

- Resultado Esperado: Un conjunto de datos limpio, estructurado y listo para el análisis predictivo.

### **Modelos Predictivos**

- Selección de Modelos:

- Regresión Logística: para predecir cancelaciones y conversión de unipatrocinadores a multipatrocinadores.

- Árboles de Decisión: Para identificar patrones de comportamiento.
- Algoritmo de Clustering: para segmentar a los patrocinadores en grupos

homogéneos.

- Entrenamiento de Modelos: utilización de validación cruzada k-fold para evaluar la robustez y generalización de los modelos.

- Evaluación de Modelos : Uso de métricas como precisión, recall y el área bajo la curva ROC(AUC-ROC) para determinar la efectividad.

- Resultado Esperado: Modelos predictivos validados y listos para generar insights.

### **Análisis de Resultados**

Interpretación de Resultados: Análisis de las variables más influyentes en la predicción de cancelaciones y en la identificación de potenciales incrementos en el compromiso.

### **Implementación de Estrategias**

Desarrollo de Estrategias: Formulación de campañas de marketing personalizadas y planes de retención basados en los insights obtenidos.

Monitoreo y Ajustes: Seguimiento continuo de las predicciones y ajustes periódicos a los modelos y estrategias basados en nuevos datos.

### **Organización, Depuración y Evaluación de los Datos Recopilados (Recopilación de Datos)**

Se realiza análisis exploratorio de los datos proporcionados por World Vision Colombia con el objetivo de revisar y corregir inconsistencias, como la presencia de datos nulos, vacíos o incongruentes. Esta revisión busca depurar la fuente de datos y normalizarla para mejorar la precisión de los resultados en las predicciones relacionadas con la conversión de patrocinadores de unipatro a multipatro y la identificación de aquellos con riesgo de cancelar su patrocinio.

El propósito principal es aumentar la cantidad de patrocinios y reducir las cancelaciones futuras. La base de datos incluye información desde 1999 hasta la fecha actual. Este proceso ha implicado la centralización de los registros de patrocinadores en un único reporte, consolidando datos de pagos, niños patrocinados, información personal de los patrocinadores, motivaciones y designaciones en una consulta única para facilitar su análisis. El acceso a esta información se realiza directamente desde el sistema de bases de datos SQL-Server, trabajando sobre la base de datos original.

#### **Tabla 1**

##### *Recopilación de Datos SIMMA*

Base de datos	SIMMA
Filas	27390
Columnas	49

*Nota.* Recopilación de Datos obtenidos y trabajados desde SQL Server 2019.

## Caracterización y Clasificación (Procesamiento de Datos)

Una vez cargado el dataframe, se procedió a realizar una caracterización detallada de las variables, clasificándolas según su tipo (categóricas, numéricas, fecha, binarias, etc) y proporcionando una descripción exhaustiva de cada una. Este proceso es fundamental para comprender la naturaleza de los datos y garantizar que las etapas posteriores de limpieza, transformación y modelado se realicen de manera efectiva, minimizando inconsistencias y maximizando la calidad de los análisis.

Selección de variables.

### Tipo Numérico

De tipo numérico se seleccionan 6 variables.

- Amount : Monto en dinero asociado al patrocinio o pago.
- Num: cantidad de niños asignados.
- tiempoPatrocinio : Tiempo (en años) que lleva el patrocinador con el patrocinio.
- Edad: Edad del patrocinador.
- Pagos: Número de pagos realizados por el patrocinador.
- EdadNiño : Edad del niño

### Tipo Categórico

De tipo categórico se seleccionan 20 variables.

- Id: Identificador único del registro
- Mpago: Método de pago.
- City : Ciudad de residencia.
- State: Estado o región.
- Tmotivacion: Tipo de motivación del patrocinador.

- Dmotivacion: Descripción de la motivación.
- Región: Región de residencia o ubicación asociada
- GeneroPatro: Género del patrocinador
- EstadoCivil: Estado civil del patrocinador
- Ocupación: Ocupación del patrocinador.
- Profesión: Profesión del patrocinador.
- Pledgetype: Tipo de patrocinio.
- Designation: Designación general del patrocinio.
- DesignacionMotivacion2: Detalle adicional de la designación del patrocinio.
- Mora: Estado de cartera del patrocinador.
- Unimulti: indicador de si el patrocinado único o multiple.
- Frequency: Frecuencia de pago.
- GenderCode: Género del niño patrocinado.
- Tipopatro: Tipo de patrocinio.
- Generación: Generación a la que pertenece el patrocinador.

### **Tipo Fecha**

De tipo fecha se seleccionan 2 variables.

- AddDate: Fecha que se añadió el registro.
- FNacimientoNin@: Fecha de nacimiento del niño patrocinado.

### **Características Descartables de los Datos Obtenidos**

Tras un análisis exhaustivo de todas las variables, se procedió a identificar y descartar aquellas entidades que no aportan valor significativo a los modelos predictivos. Adicionalmente, se normalizaron los datos para garantizar coherencia y comparabilidad, y se implementaron

técnicas como imputación o eliminación de datos para campos nulos con el objetivo de preservar la calidad y fiabilidad del conjunto de datos final.

Información cargada en el dataframe después de la caracterización, normalización y depuración de datos.

**Tabla 2**

*Datos Obtenidos después del Proceso de Imputación y Normalización*

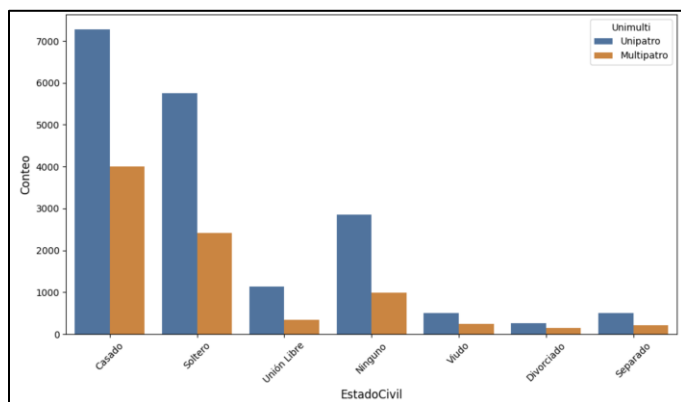
Base de datos	SIMMA
Filas	26618
Columnas	28

*Nota.* Datos obtenidos después del proceso de imputación y normalización, tomados de SIMMA.

## Representaciones Gráficas y Métricas de los Datos

**Figura 1**

*Relación de Patrocinadores con Estado Civil y Categoría Unimulti*



*Nota.* Gráfica del comportamiento por estado civil por patrocinadores.

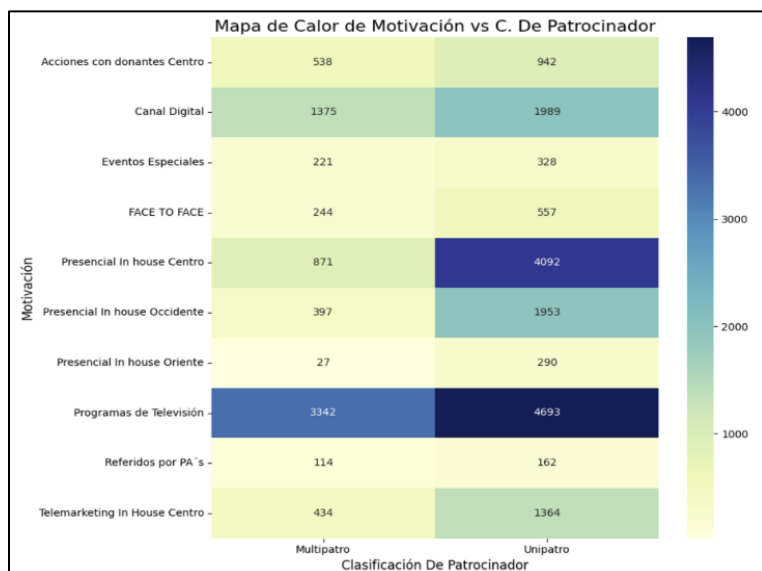
La figura 1 permite identificar patrones clave en la relación entre el estado civil y la conversión de unipatrocinadores a multipatrocinadores. Las categorías "Casado" y "Soltero" destacan como los segmentos con mayor número de patrocinadores, tanto en "Unipatro" como en "Multipatro". Sin embargo, el volumen considerable de unipatrocinadores en estas categorías indica un potencial significativo para fomentar conversiones hacia multipatrocinadores.

En particular, los casados, al representar un compromiso establecido, podrían ser un objetivo estratégico para campañas que promuevan el apoyo múltiple, apelando a valores familiares y estabilidad económica. La categoría "Unión Libre" presenta una menor proporción de multipatrocinadores, lo que sugiere que este grupo podría requerir estrategias diferenciadas, quizá basadas en motivaciones más emocionales o personalizadas. Los patrocinadores sin estado civil declarado ("Ninguno") también ofrecen una oportunidad para aumentar la conversión, al representar un segmento con potencial de compromiso adicional si se identifican motivaciones clave. La transición de unipatrocinadores a multipatrocinadores en función del estado civil puede

ser optimizada mediante campañas específicas que resalten el impacto ampliado de su contribución.

## Figura 2

### Mapa de Calor de Motivación vs Clasificación de Patrocinadores

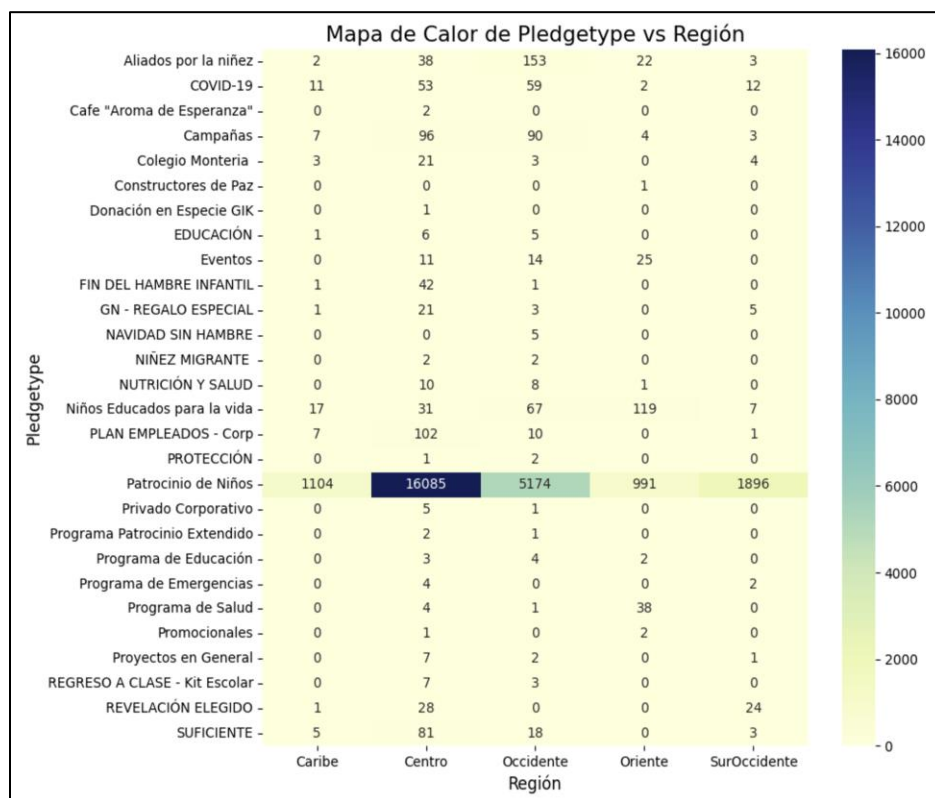


*Nota.* Mapa de calor de patrocinadores organizados por los canales de motivación.

La figura 2 muestra cómo las diferentes motivaciones se distribuyen entre los patrocinadores clasificados como unipatrocinadores y multipatrocinadores. Los programas de televisión son la motivación más fuerte para ambos grupos, destacándose especialmente entre los unipatrocinadores. Las actividades presenciales, como "In house Centro" y "Occidente", también tienen una alta incidencia, con un impacto mayor en unipatrocinadores. En menor medida, motivaciones como "Canal Digital" y "Telemarketing" presentan una distribución más balanceada entre ambos tipos de patrocinadores. Las categorías menos frecuentes, como "Eventos Especiales" o "Referidos por PA's", muestran menor representación, lo que puede indicar áreas de oportunidad para atraer más multipatrocinadores.

**Figura 3**

*Mapa de Calor de Compromisos Asignados por Regiones.*

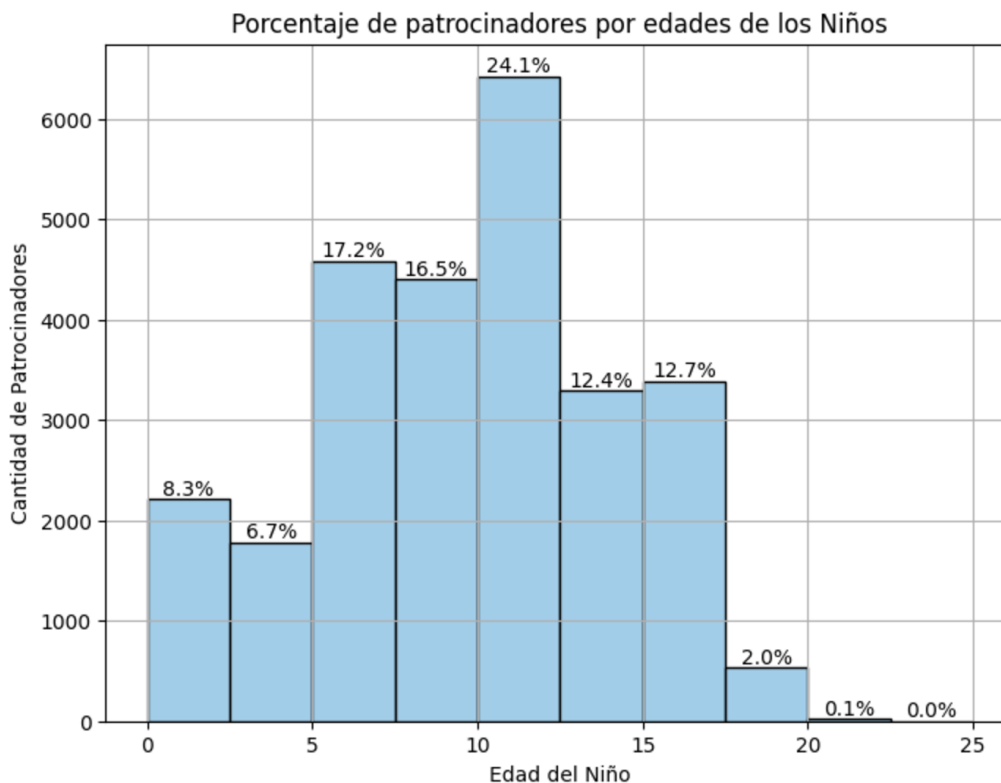


*Nota.* Mapa de calor de patrocinadores organizados por región.

La figura 3 refleja cómo las diferentes categorías de Pledgetype se distribuyen en las distintas regiones. "Patrocinio de Niños" domina todas las regiones, con su mayor concentración en la región Centro, seguido de Occidente. Esto sugiere que esta categoría representa el principal foco de patrocinio en el modelo operativo actual. Otras categorías, como "Niños Educados para la Vida" y "Nutrición y Salud", muestran una menor representación, destacando potenciales nichos a fortalecer. Las regiones como Caribe y SurOccidente tienen menor actividad general en comparación con el Centro, indicando posibles oportunidades para expandir estrategias en estas áreas menos atendidas.

**Figura 4**

*Porcentaje de Patrocinadores por Edad de los Niños.*



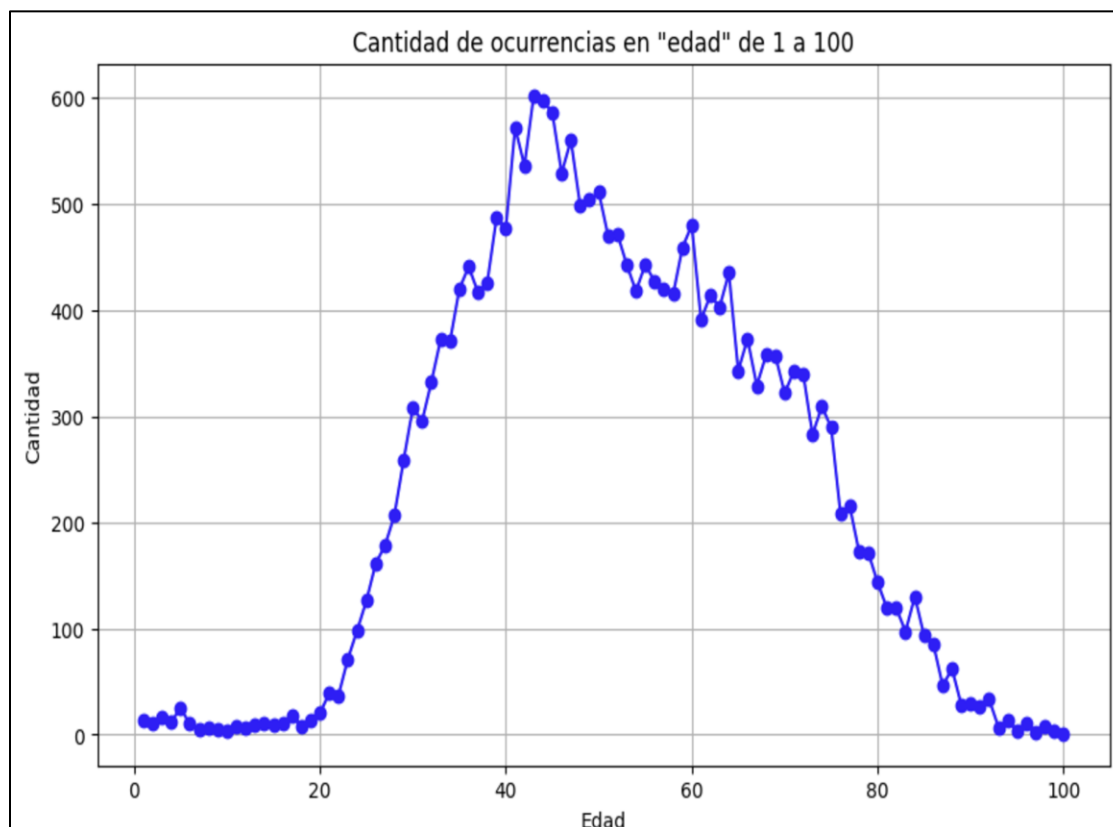
*Nota.* Gráfica de patrocinadores organizado por edad de niños patrocinados.

La figura 4 muestra una caída significativa en el porcentaje de patrocinadores a partir de los 15 años, indicando que este rango de edad es crítico para implementar estrategias de retención. Aunque los niños de 10 años tienen el mayor porcentaje de patrocinadores (24.1%), este interés disminuye de manera consistente conforme los niños se acercan a los 15 años y más allá. Este patrón puede estar relacionado con el fin del ciclo de apadrinamiento o una percepción de menor impacto por parte del donador cuando el niño se acerca a la edad adulta. Enfocar campañas de retención específicamente en el grupo de 15 años en adelante resulta esencial para mitigar la deserción, destacando historias de impacto o el potencial futuro de estos jóvenes en sus

comunidades. Este enfoque podría contrarrestar la tendencia de abandono observada, maximizando la continuidad del compromiso de los patrocinadores en esta etapa crucial.

### Figura 5

*Cantidad de Ocurrencias de Patrocinadores Clasificados por su Edad.*



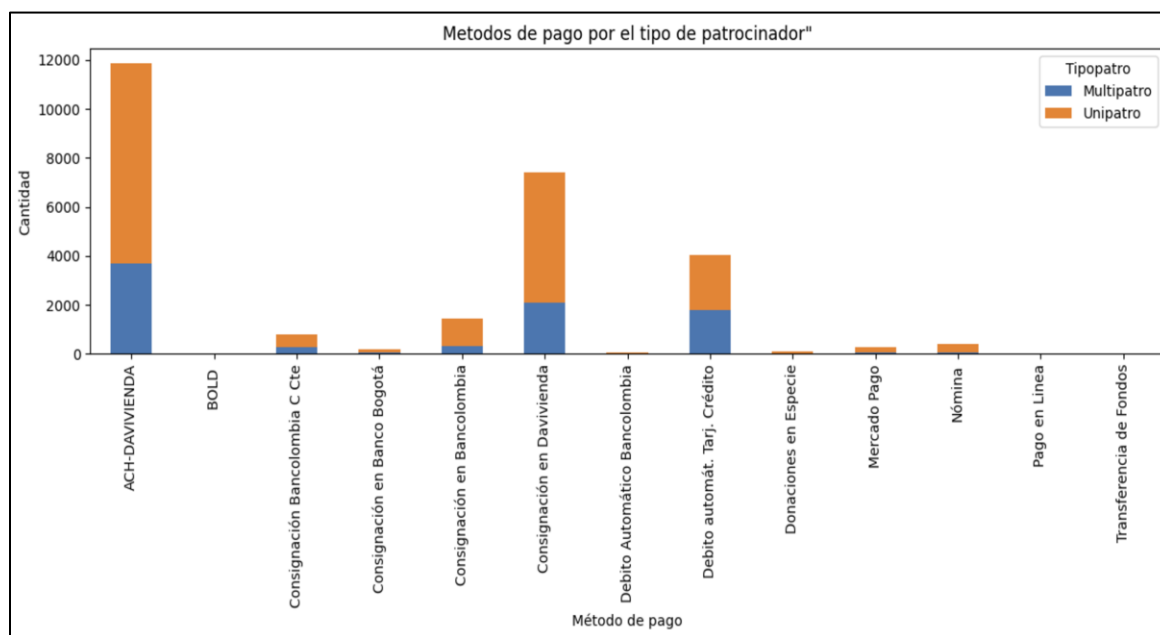
*Nota.* Gráfica de patrocinadores clasificados por edades.

La figura 5 muestra una curva representativa de la distribución de edad de los patrocinadores, con un aumento pronunciado entre los 20 y 40 años, alcanzando el pico máximo cerca de los 35-40 años. Posteriormente, se observa un declive gradual a partir de los 50 años. Este comportamiento sugiere que los adultos jóvenes y de mediana edad (20-40 años) son el segmento predominante y más activo en el patrocinio, mientras que los mayores de 50 presentan una menor incidencia. Para estrategias de conversión de unipatro a multipatro, el enfoque debe

centrarse en el grupo de 30-40 años, ya que este segmento combina estabilidad económica y una fuerte presencia en la base de datos. Las estrategias de retención deben reforzarse para los patrocinadores mayores de 50 años, considerando que podrían estar en riesgo de desvinculación debido a factores asociados con ingresos fijos o cambios en prioridades.

## Figura 6

*Métodos de Pago por el Tipo de Patrocinador.*



*Nota.* Clasificación de métodos de pago predominantes.

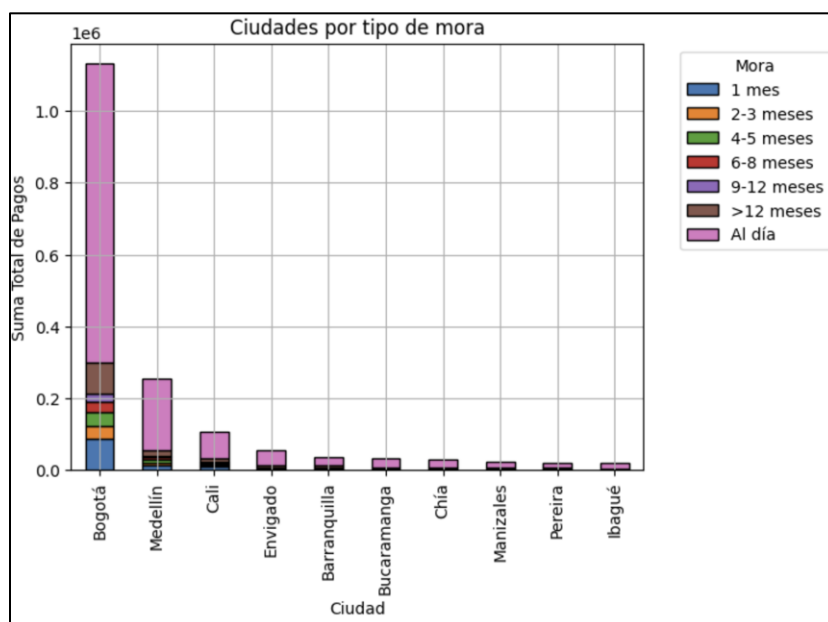
La figura 6 aporta información clave para estrategias de retención y conversión de patrocinadores al analizar los métodos de pago más utilizados por unipatrocinadores y multipatrocinadores. El método ACH-Davivienda domina ampliamente, lo que sugiere que la mayoría de los patrocinadores prefieren mecanismos automatizados y confiables. Esto también indica que campañas dirigidas a patrocinadores de este segmento pueden ser altamente efectivas al promover la migración de unipatro a multipatro, reforzando la conveniencia del método.

Los métodos alternativos, como débito automático y consignaciones, tienen menor participación pero muestran oportunidades para diversificar estrategias. La baja representación de métodos digitales como Mercado Pago o Pago en Línea destaca áreas de mejora para captar patrocinadores más jóvenes o tecnológicamente conectados. Este análisis refuerza la importancia de facilitar transacciones en los métodos más populares mientras se promueven alternativas digitales para captar nuevos segmentos de patrocinadores.

Ciudades por tipo de mora.

### Figura 7

*Métodos de Pago por el Tipo de Patrocinador.*



*Nota.* Clasificación de efectividad de pagos de patrocinadores por ciudades.

La figura 7 ofrece información valiosa para estrategias de retención y recuperación de pagos en función de la ubicación geográfica. Bogotá destaca como la ciudad con la mayor suma de pagos, pero también refleja una presencia significativa en categorías de mora, como "9-12 meses" y ">12 meses". Esto sugiere que, aunque es el principal centro de patrocinadores activos,

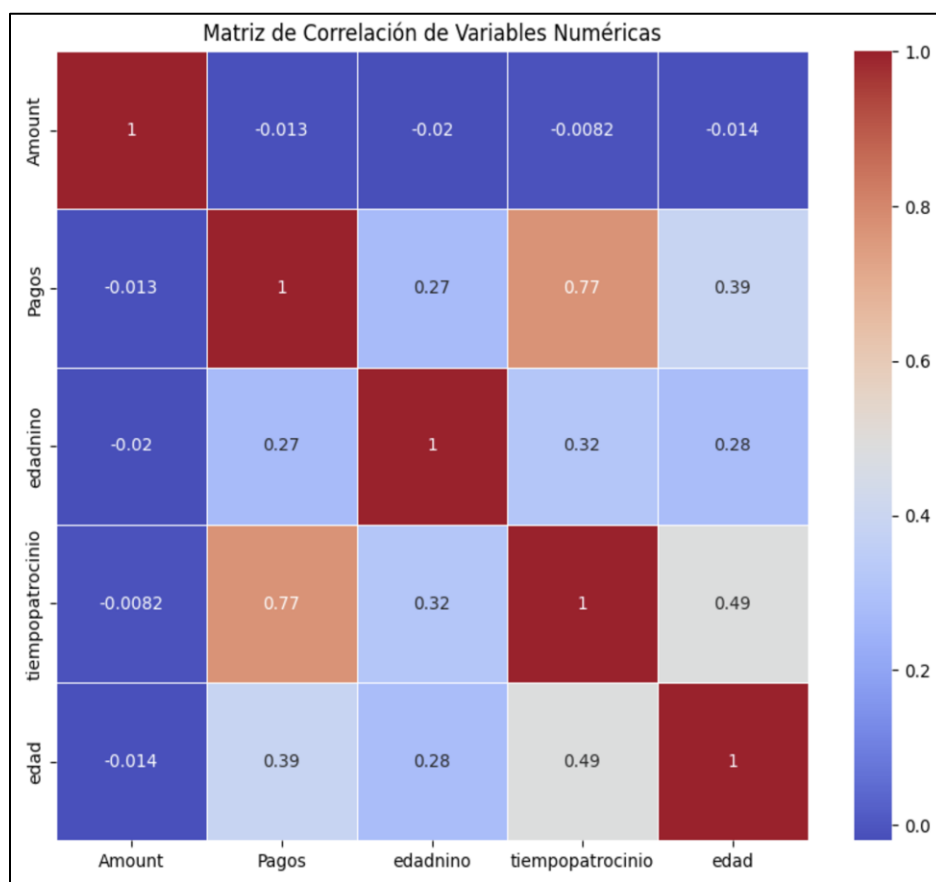
requiere estrategias focalizadas para reducir el atraso en los pagos. Medellín y Cali, aunque con menores volúmenes totales, presentan patrones similares, lo que indica oportunidades para abordar la mora antes de que aumente. Las ciudades con menor participación, como Pereira y Ibagué, muestran potencial para desarrollar estrategias de captación y fidelización. Estas ciudades podrían beneficiarse de campañas dirigidas que refuercen la importancia de la puntualidad en los pagos, especialmente en patrocinadores con tendencia a atrasos recurrentes.

## Método de Predicción para la Conversión de Patrocinadores de Unipatrocinadores a Multipatrocinadores

Se analizan las variables numéricas y categóricas para evaluar su correlación, lo que facilita la construcción de modelos predictivos orientados a optimizar la retención de patrocinadores. El objetivo es identificar patrones que reduzcan la intención de cancelación y, al mismo tiempo, fomentar la captación de nuevos compromisos.

### Figura 8

*Matriz de Correlación de Variables Numéricas*



*Nota.* Clasificación de variables numéricas organizadas en la matriz de correlación.

La figura 8 representa la matriz de correlación numérica la cual revela relaciones clave entre las variables para los objetivos de retención y captación de patrocinadores.

#### Relación entre pagos y tiempo de patrocinio (0.769)

Existe una fuerte correlación positiva, indicando que a mayor duración del patrocinio, mayor es el número de pagos realizados. Esto sugiere que los patrocinadores con mayor antigüedad tienen un compromiso más estable, lo que puede ser aprovechado para estrategias de retención.

#### Tiempo de Patrocinio y Edad del Patrocinador(0.49)

Una correlación moderada refleja que los patrocinadores mayores tienden a mantener un compromiso más prolongado. Las campañas de fidelización deberían reforzarse en este segmento para evitar fugas.

#### Pagos y Edad del Niño(0.27)

Aunque la relación es débil, indica que los pagos tienden a ser más frecuentes cuando los niños son menores. Esto valida que los esfuerzos de retención deben intensificarse conforme los niños alcanzan los 15 años o más, como lo muestran otros análisis.

#### Amount y Otras Variables(todas < 0.02)

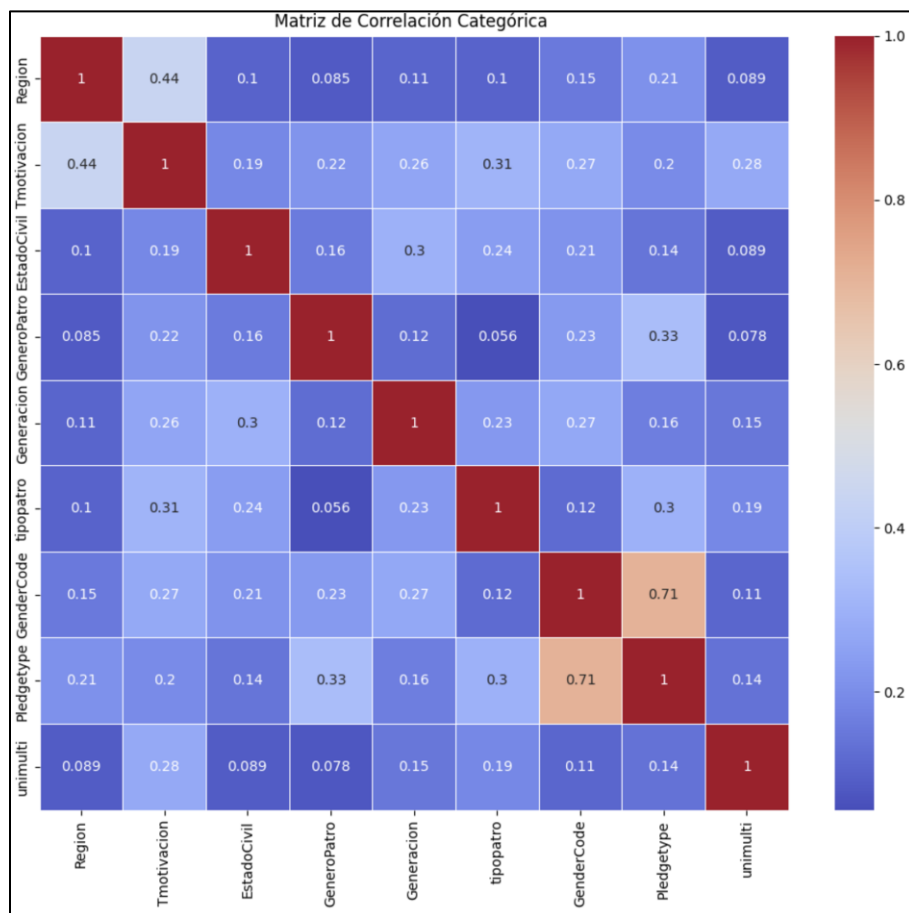
La falta de correlación significativa con el monto sugiere que este no está directamente influenciado por las variables analizadas, requiriendo un enfoque separado para entender patrones relacionados con las contribuciones.

La correlación entre edad del patrocinador y otras variables (pagos y tiempo de patrocinio) indica que los segmentos con patrocinadores más experimentados son clave para estrategias de retención, mientras que la relación débil con la edad del niño refuerza la necesidad

de intervenir antes de que el patrocinio decline. Este análisis confirma que las campañas personalizadas deben enfocarse en las variables que influyen en la estabilidad del compromiso.

### Figura 9

*Matriz de Correlación de Variables Categóricas.*



*Nota.* Clasificación de variables categóricas organizadas en la matriz de correlación.

La figura 9 es la matriz de correlación categórica utilizando Cramér's V proporciona información clave para identificar relaciones entre las variables categóricas, lo que ayuda a priorizar los factores más relevantes para la retención y conversión de patrocinadores. Existe una fuerte relación entre el tipo de patrocinio y el género del niño patrocinado. Esto sugiere que

ciertos tipos de patrocinio pueden ser más populares dependiendo del género del niño, lo cual puede guiar estrategias personalizadas según el perfil del patrocinador.

#### Tipopatro y Pledgetype (0.30)

La relación moderada indica que el tipo de patrocinio puede estar relacionado con características específicas del compromiso, como la frecuencia o modalidad, lo cual es relevante para diseñar estrategias de conversión.

#### Tmotivacion y Pledgetype (0.20)

Una correlación moderada sugiere que las motivaciones específicas influyen en el tipo de patrocinio elegido. Esto es útil para segmentar campañas de retención o promoción basadas en las motivaciones clave.

#### Tmotivacion y Unimulti (0.27)

Indica que las motivaciones también afectan la clasificación entre unipatrocinadores y multipatrocinadores. Identificar qué motivaciones están más asociadas con multipatrocinadores puede ser clave para estrategias de conversión.

#### EstadoCivil y Generacion (0.29)

Muestra que la generación del patrocinador puede estar relacionada con su estado civil, lo que refuerza la idea de segmentar campañas considerando estas características demográficas.

#### Region y Tmotivacion (0.44)

Una correlación relativamente alta refleja que la ubicación geográfica del patrocinador influye en sus motivaciones. Esto puede ayudar a diseñar estrategias regionales enfocadas en las motivaciones predominantes.

#### GeneroPatro y Pledgetype (0.33)

Sugiere que el género del patrocinador tiene un impacto en el tipo de patrocinio, lo cual puede ser útil para personalizar estrategias en función de su perfil.

Unimulti y otras variables (bajas correlaciones)

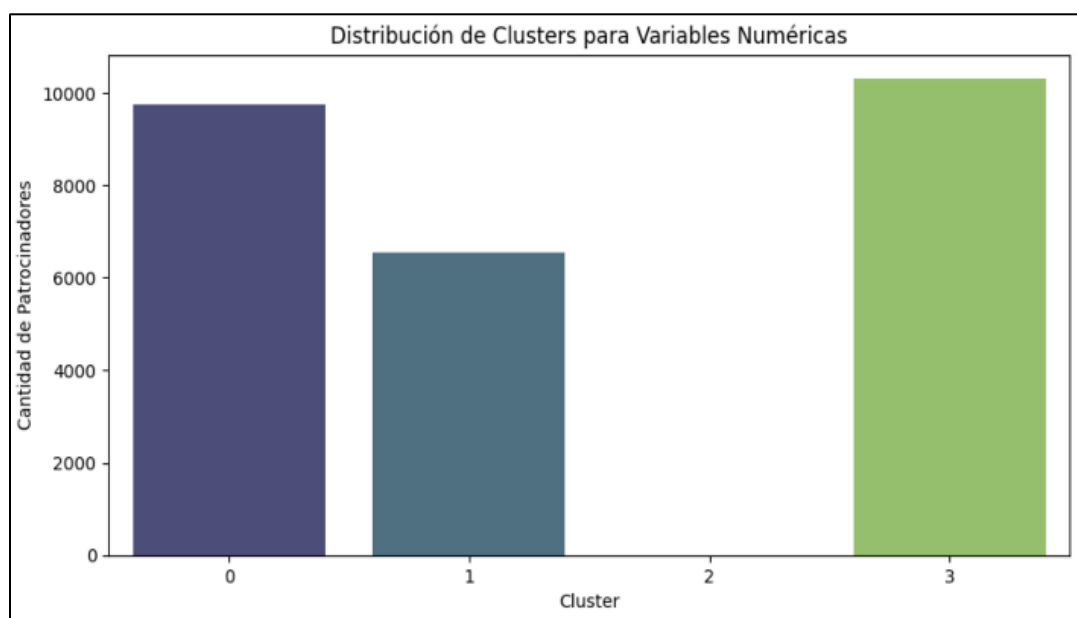
Las correlaciones más bajas ( $\leq 0.20$ ) indican que la clasificación como unipatrocinador o multipatrocinador no está fuertemente influenciada por variables como EstadoCivil, Region o Generacion. Esto sugiere que la conversión puede depender más de factores no categóricos, como montos o duración del patrocinio.

Segmentación de Patrocinadores por Clustering

Se organizan los clustering según el tipo de variables con el propósito de identificar patrones medibles entre las variables objetivo. Esto con el fin de obtener resultados más precisos al agrupar a los patrocinadores y sus modalidades de patrocinio en grupos homogéneos y comprensibles.

## Figura 10

### *Clusters por Variables Numéricas*



*Nota.* Distribución de clusters y método del codo para la selección del número de clusters.

**Tabla 3***Resumen de Clusters Numéricos*

Cluster	Valor	Pagos	Edad_Niño	Tiempo_Patro	Edad_Patro
0	4.759074e+04	156.055140	11.199447	14.802911	59.725940
1	4.845545e+04	18.387668	5.143620	2.422924	16.863858
2	7.333333e+06	5.666667	0.000000	9.333333	0.000000
3	4.844112e+04	50.063555	10.529594	5.172327	48.364157

*Nota.* Resumen de Clusters Numéricos tomado de SIMMA

En la tabla se representa los datos obtenidos por cada cluster para las variables numéricas.

Tras analizar los resultados de la segmentación de patrocinadores basada en variables numéricas, representados en la figura 10, y al revisar el resumen de los clusters obtenidos de tabla 3, se puede interpretar lo siguiente:

Cluster 0:

- Patrocinadores Mayores y Consolidados
- Edad promedio alta de 59.7 años.
- Realizan un número significativo de pagos (156 en promedio).
- Tiempo de patrocinio prolongado (14.8 años).
- Los niños patrocinados tienen una edad promedio de 11.2 años.
- Perfil: Este grupo representa patrocinadores con alto compromiso, posiblemente

jubilados o con ingresos estables, que han sostenido relaciones largas con la ONG. Son clave para estrategias de retención y fidelización, destacando el impacto de su apoyo a lo largo de los años.

#### Cluster 1:

- Nuevos Patrocinadores Jóvenes
- Edad promedio baja de 16.8 años, lo que puede incluir a patrocinadores muy jóvenes, posiblemente familiares de niños apadrinados o personas que recién comienzan su apoyo.
- Tiempo de patrocinio corto (2.4 años en promedio), número bajo de pagos (18.3 en promedio), los niños patrocinados son relativamente jóvenes (5.1 años en promedio),
- Perfil: Este segmento se compone de patrocinadores que están en la etapa inicial de su relación con la ONG. Es crucial implementar estrategias de engagement para aumentar su participación y prolongar su compromiso, incentivando su conversión a largo plazo.

#### Cluster 2:

- Empresas y Organizaciones como Patrocinadores
- Edad de los patrocinadores es "0", lo que refleja que este cluster está compuesto por empresas o entidades corporativas, no individuos.
- Los niños patrocinados tienen una edad promedio de 0 años, lo cual es lógico al incluir recién nacidos.
- Montos muy elevados en los pagos promedio (7.3 millones).
- Tiempo de patrocinio relativamente prolongado (9.3 años en promedio).
- Perfil: Este grupo representa a empresas o grandes organizaciones que realizan aportes significativos y suelen apadrinar niños recién nacidos. Su impacto financiero es alto, lo que los convierte en aliados estratégicos para la sostenibilidad de los programas. Las estrategias deben enfocarse en fortalecer la relación con estas entidades, demostrando los resultados y el impacto positivo generado por sus contribuciones.

### Cluster 3:

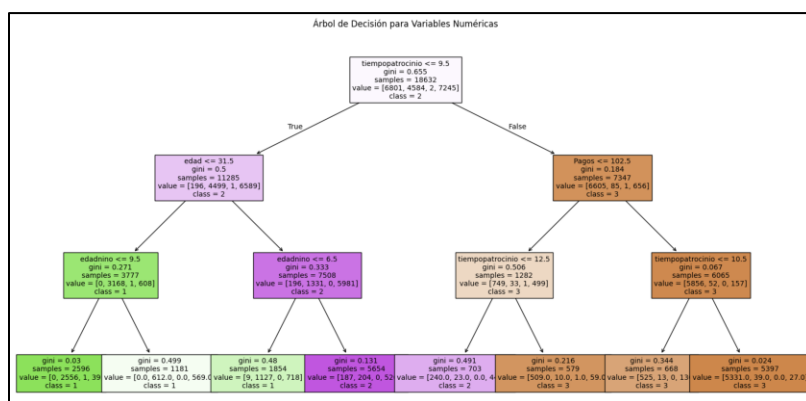
- Patrocinadores Adultos con Compromiso Moderado
- Edad promedio de 48.3 años.
- Tiempo de patrocinio intermedio (5.1 años).
- Realizan 50 pagos en promedio.
- Los niños patrocinados tienen una edad promedio de 10.5 años.
- Perfil: Este grupo incluye patrocinadores con una trayectoria significativa, aunque no tan prolongada como en el Cluster 0. Representan un segmento maduro, con ingresos estables y potencial para aumentar su nivel de compromiso. Son ideales para estrategias de conversión a multipatrocinadores, mostrando cómo su apoyo adicional puede amplificar el impacto en las comunidades atendidas.

no tan prolongada como en el Cluster 0. Representan un segmento maduro, con ingresos estables y potencial para aumentar su nivel de compromiso. Son ideales para estrategias de conversión a multipatrocinadores, mostrando cómo su apoyo adicional puede amplificar el impacto en las comunidades atendidas.

## Segmentación de Patrocinadores por Árboles de Decisión por Variables Numéricas

**Figura 11**

### Árbol de Decisión para Variables Numéricas



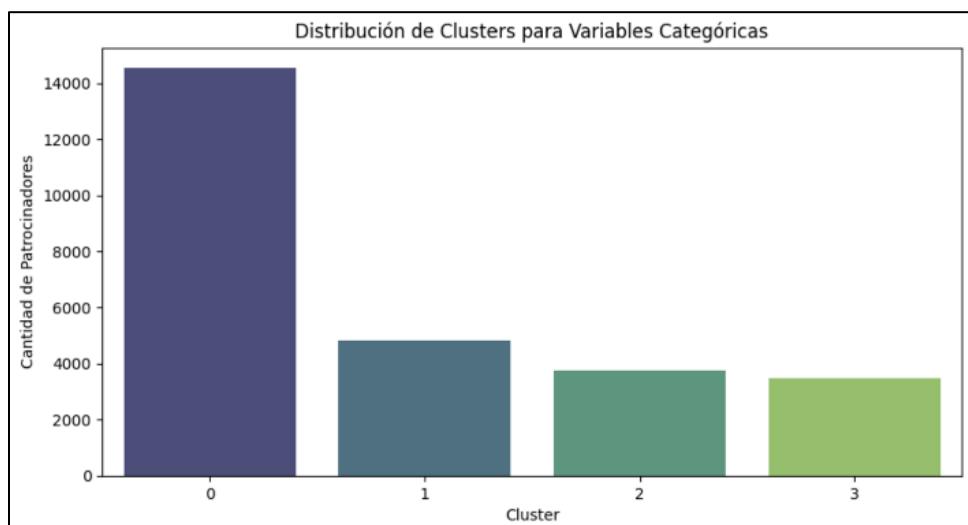
*Nota.* Estructura de árbol de decisión para variables numéricas.

En la figura 11 el árbol de decisión obtenido identifica patrones clave en los datos de los patrocinadores, alineándose con los resultados del clustering. La variable más relevante es tiempo de patrocinio, usada como el primer criterio para dividir los datos. Patrocinadores con un tiempo de patrocinio menor o igual a 9.5 años se agrupan principalmente en los clusters 1 y 2, mientras que aquellos con valores superiores tienden hacia el cluster 3, donde los pagos son también un factor relevante.

En el subárbol izquierdo, la variable edad separa los patrocinadores en función de su perfil etario. Los patrocinadores con una edad menor o igual a 31.5 años tienden a pertenecer al cluster 2, mientras que aquellos con niños menores de 9.5 años se asocian principalmente al cluster 1.

En el subárbol derecho, los pagos (Pagos) y nuevamente el tiempo de patrocinio determinan la asignación al cluster 3. Los patrocinadores con mayores contribuciones financieras y tiempos de patrocinio más largos forman este grupo.

Este árbol confirma que los patrones clave identificados en los clusters del resumen (edad, pagos, tiempo de patrocinio y edad de los niños patrocinados) se reflejan en el comportamiento de los datos. La precisión del árbol (87.5%) respalda su utilidad para comprender y segmentar las características de los patrocinadores.

**Figura 12***Clusters por Variables Categóricas*

*Nota.* Distribución de clusters basada en variables categóricas método del codo para la selección óptima del número de clusters.

**Tabla 4***Resumen de Clusters Categóricos*

Cluster	Region	T.Motivacion	Estado_Civil	Genero Patro	Generacion
0	Centro	Program. Tv	Casado	Female	
1	Centro	Canal digital	Soltero	Male	Generation X
2	Centro	P. in house	Ninguno	Male	
3	Occidente	P. in house	Soltero	Female	
,	Tipo_Patro	Genero Niño	Pledge_Type	Uni_Multi	
0	Individuo	F	Patro. Niños	Unipatro	
1	Individuo	M	Patro. Niños	Multipatro	
2	Individuo	M	Patro. Niños	Unipatro	
3	Individuo	M	Patro. Niños	Unipatro	

*Nota.* Distribución de variables organizada por tipo de cluster categórico.

Tras analizar los resultados de la segmentación de patrocinadores basada en variables categóricas, representados en la figura 12, y al revisar el resumen de los clusters obtenidos de tabla 4, se puede interpretar lo siguiente:

Cluster 0: La mayoría de los patrocinadores se concentran en esta categoría. Se asocian con la región Centro, tienen como motivación principal los Programas de Televisión, predominan los Casados, el género predominante es Female, y pertenecen a la Generación X. Su tipo de patrocinio es Unipatro.

Este cluster representa una base sólida de patrocinadores con estabilidad personal y emocional, lo que sugiere potencial para campañas de conversión de Unipatro a Multipatro. Las estrategias deben resaltar el impacto familiar y comunitario de ampliar su compromiso.

Cluster 1 Predominan en la región Centro, motivados principalmente por el Canal Digital, en su mayoría Solteros, con género Male y pertenecen a la generación Millennials. Son tanto Unipatro como Multipatro.

Este cluster muestra una mayor afinidad por canales digitales, indicando un perfil más joven y tecnológicamente conectado. Las campañas deben centrarse en experiencias personalizadas, utilizando medios digitales como redes sociales o marketing por correo electrónico para incrementar la retención y promover su migración a Multipatro.

Cluster 2 Residen mayoritariamente en la región Centro, motivados por el contacto presencial (Presencial In house Centro). La mayoría no declara un estado civil (Ninguno), con género Male, y tienen como categoría generacional 0. Son principalmente Unipatro.

Este grupo presenta características ambiguas en cuanto a estado civil y generación, lo que sugiere que podrían ser patrocinadores menos comprometidos o nuevos. Las campañas deberían

enfocarse en fortalecer su vínculo emocional con la causa y en destacar el impacto directo de su apoyo en la comunidad.

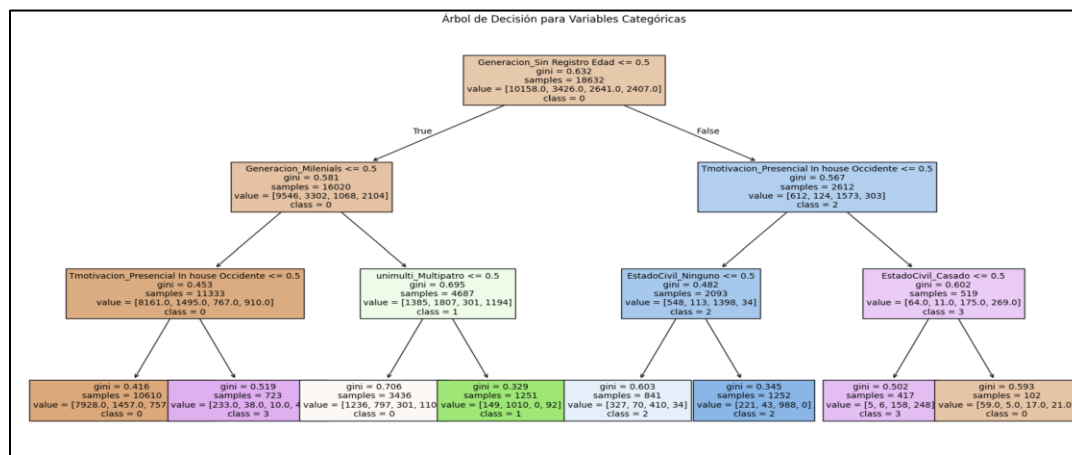
Cluster 3 Se concentran en la región Occidente, motivados por el enfoque presencial en su región (Presencial In house Occidente). Predominan los Solteros, con género Female, generación Millennials, y tienen mayoritariamente Unipatro.

Este grupo podría beneficiarse de estrategias locales que refuercen su sentido de comunidad e impacto directo en su región. Enfocar campañas en historias locales de éxito y en cómo su contribución puede transformarse en apoyo múltiple sería clave.

### Segmentación de Patrocinadores por Árboles de Decisión por Variables Categóricas

**Figura 13**

*Árbol de Decisión para Variables Categóricas*



*Nota.* Estructura de árbol de decisión para variables categóricas.

La figura 13 representa el árbol de decisión obtenido refleja los patrones identificados durante la segmentación exploratoria basada en clustering. Las variables seleccionadas, como "Generación", "Tmotivación", "unimulti" y "EstadoCivil", muestran su relevancia en la segmentación de patrocinadores en diferentes clases.

La "Generación" destaca como la variable más influyente, dividiendo inicialmente a los patrocinadores en segmentos clave como "Sin Registro Edad" y "Milenials". Este patrón confirma su importancia, previamente evidenciada en el clustering, donde generaciones específicas se asociaron con ciertas características demográficas y de patrocinio.

En ramas más profundas, "Tmotivación\_Presencial In house Occidente" y "unimulti\_Multipatro" segmentan aún más los datos según las modalidades de motivación y el tipo de patrocinio (unipatrocinio o multipatrocinio). Finalmente, "EstadoCivil" refina las clases, revelando que variables sociales también juegan un papel en las características del patrocinio.

Este árbol complementa los resultados del clustering, proporcionando una herramienta visual para entender las decisiones de segmentación y sus bases en los patrones de los datos.

### **Modelo Predictivo de Conversión de Patrocinadores Unipatro a Multipatro con Regresión Logística**

El propósito del modelo es identificar a los patrocinadores que tienen el potencial de convertirse de unipatro a multipatro. Esto es esencial para optimizar las estrategias de conversión, priorizar recursos y maximizar el impacto de los programas de patrocinio.

#### **Tabla 5**

*Resultado Promedio Validación Cruzada-Kfold M. Unipatro a Multipatro*

Validación Cruzada K-Fold (5)	
Accuracy promedio	0.8221
AUC-ROC promedio	0.8720
Precisión promedio	0.7450
Recall promedio	0.6520

*Nota.* Resultado de validación promedio por validación cruzada-Kfold M.

Accuracy promedio: 82.21% Representa la proporción de predicciones correctas del modelo en todos los pliegues. Este alto valor indica que el modelo es preciso al clasificar patrocinadores como unipatro o multipatro en un 82.21% de los casos.

AUC-ROC promedio: 87.20% El AUC-ROC refleja la capacidad del modelo para distinguir entre unipatrocinadores y multipatrocinadores en términos de probabilidad. Un valor del 87.20% demuestra que el modelo tiene un rendimiento robusto para separar correctamente ambas clases.

Precisión promedio: 74.50% La precisión muestra que el modelo identifica correctamente a multipatrocinadores en el 74.50% de los casos, minimizando falsos positivos. Esto es crucial para campañas específicas, ya que garantiza que las estrategias dirigidas a convertir unipatrocinadores a multipatrocinadores sean efectivas.

Recall promedio: 65.20% El recall indica que el modelo logra identificar al 65.20% de los multipatrocinadores reales. Aunque es menor que la precisión, sugiere que se pueden perder algunos multipatrocinadores potenciales en el proceso de clasificación.

#### Resultados Modelo Predictivo

**Tabla 6**

*Resultado Rendimiento en la Clasificación de Patrocinadores ( Unipatro vs Multipatro)*

Resultados clasificación de patrocinadores	
Accuracy	0.8181
AUC-ROC	0.8660
Precisión	0.7331
Recall promedio	06240

*Nota.* Resultados de clasificación de patrocinadores de unipatro a multipatro.

**Tabla 7***Matriz de Confusión Unipatro vs Multipatro*

Matriz de confusión	
3276	336
644	1038

*Nota.* Resultados de clasificación de patrocinadores de unipatro a multipatro.

El modelo presenta una precisión promedio del 73.31%, lo que indica que 3 de cada 4 casos clasificados como multipatro son correctos. Basándonos en la matriz de confusión promedio, se identificaron 1038 verdaderos positivos (TP), que son multipatrocinadores correctamente predichos, y 336 falsos positivos (FP), clasificados como multipatro pero que realmente son unipatro. Esto significa que, de los 1374 registros clasificados como multipatro, el 75.55% tienen alta probabilidad de conversión real.

Por otro lado, existen 644 falsos negativos (FN), patrocinadores que realmente son multipatro pero que fueron clasificados como unipatro. Este grupo, aunque no identificado directamente por el modelo, representa una oportunidad latente de conversión. Al combinarlos con los 1038 verdaderos positivos, el modelo nos proporciona un marco estratégico para priorizar esfuerzos en estos 1682 patrocinadores como posibles casos para acciones de conversión.

#### Modelo Predictivo Identificar Patrocinadores con Alta Probabilidad de Cancelar o Inactivar su Patrocinio

El propósito del modelo es identificar a los patrocinadores que tienen el potencial de cancelar o inactivar su patrocinio. Lo cual es esencial para reforzar las estrategias de fidelización de patrocinadores

**Tabla 8***Resultado Promedio Validación Cruzada-Kfold*

Validación Cruzada K-Fold (5)	
Accuracy promedio	0.8840
AUC-ROC promedio	0.9056
Precisión promedio	0.7711
Recall promedio	0.8261

*Nota.* Resultado de validación promedio por validación cruzada-Kfold M para cancelar su patrocinio.

El modelo predictivo diseñado para identificar patrocinadores con alta probabilidad de cancelar su patrocinio presenta resultados sólidos tras la validación cruzada con 5 pliegues. Con una exactitud promedio de 88.40% y un AUC-ROC promedio de 90.56%, demuestra una capacidad robusta para diferenciar entre patrocinadores en riesgo y aquellos que mantendrán su compromiso. La precisión promedio de 77.11% refleja que la mayoría de las predicciones positivas realizadas por el modelo son correctas, lo que es crucial para enfocar recursos en intervenciones efectivas. Por otro lado, el recall promedio de 82.61% muestra que el modelo identifica una proporción significativa de patrocinadores en riesgo real, reduciendo pérdidas potenciales. Los resultados obtenidos en los 5 pliegues confirman la consistencia del modelo, proporcionando confianza en su generalización a nuevos datos. Esto permite a World Vision Colombia implementar estrategias personalizadas de retención, priorizando los patrocinadores más vulnerables. Además, el balance entre precisión y recall optimiza el uso de recursos, garantizando que las campañas se dirijan a quienes más lo necesitan. Este modelo es una herramienta clave para mejorar la sostenibilidad financiera y el impacto social de la

organización, cumpliendo los objetivos del proyecto al aprovechar técnicas de análisis predictivo para decisiones estratégicas más informadas.

**Tabla 9**

*Patrocinadores con Probabilidad de Cancelación.*

Resultados clasificación de patrocinadores	
Accuracy	0.8901
AUC-ROC	0.9133
Precisión	0.7423
Recall promedio	06518

*Nota.* Clasificación de patrocinadores con probabilidad de cancelación.

**Tabla 10**

*Matriz de Confusión Patrocinadores con Probabilidad de Cancelación*

Matriz de confusión	
3450	451
123	1300

*Nota.* Resultados de patrocinadores con probabilidad de cancelación.

El modelo presenta una precisión del 89%, lo que indica que cerca de 9 de cada 10 casos han sido correctamente clasificados. Según la matriz de confusión, se identificaron 1300 verdaderos positivos (TP), es decir, patrocinadores correctamente predichos como en riesgo de cancelar, lo que permite focalizar las estrategias de retención de manera eficiente. Además, se identificaron 3450 verdaderos negativos (TN), que corresponden a patrocinadores correctamente clasificados como sin intención de cancelar, asegurando que los esfuerzos no se desperdicien en intervenciones innecesarias. Por otro lado, el modelo generó 451 falsos positivos (FP),

patrocinadores clasificados como en riesgo, pero que realmente no tienen intención de cancelar. Aunque este grupo no requiere intervención inmediata, representa un área de mejora para minimizar el uso de recursos en casos que no lo necesitan. Asimismo, hubo 123 falsos negativos (FN), patrocinadores que están en riesgo pero no fueron identificados por el modelo. Este grupo representa una oportunidad latente para ajustes futuros en el modelo, con el objetivo de reducir posibles pérdidas financieras.

En total, al combinar los 1300 verdaderos positivos con los 123 falsos negativos, se identifican 1423 patrocinadores en riesgo potencial, proporcionando un marco claro para priorizar estrategias de retención e intervención directa.

## Análisis de Resultados

Evaluación resultado del modelo predictivo de conversión de patrocinadores unipatro a multipatro, para interpretar los resultados y analizar las variables más influyentes en la predicción del modelo de conversión de patrocinadores de unipatro a multipatro, se deben considerar las métricas de desempeño, el impacto de las variables utilizadas, y cómo estas influyen en las decisiones predictivas.

Métricas de desempeño, la exactitud del 82.21% indica que el modelo clasifica correctamente la mayoría de los casos, con un balance adecuado entre unipatro y multipatro.

El AUC-ROC de 87.20% sugiere que el modelo distingue de manera robusta entre ambas clases, lo que confirma su capacidad de identificar patrones significativos.

La precisión promedio del 74.50% refleja que, cuando el modelo clasifica un patrocinador como multipatro, existe una alta probabilidad de que esta predicción sea correcta.

El recall promedio del 65.20% implica que el modelo identifica una proporción significativa de multipatrocinadores reales, aunque hay margen para mejorar en la reducción de falsos negativos.

Análisis de Variables Más Influyentes, Tmotivacion: Se observa que las motivaciones específicas de los patrocinadores juegan un papel importante en la probabilidad de conversión. Las categorías de motivación asociadas con canales digitales o programas televisivos tienen una incidencia directa en el compromiso de los patrocinadores. Esto sugiere que la personalización de campañas basada en la motivación podría mejorar la efectividad en conversiones.

tiempopatrocio: Este es uno de los predictores clave en el modelo. Los patrocinadores con mayor duración de compromiso muestran una mayor tendencia a convertirse en multipatro, lo que valida su importancia para segmentar y priorizar estrategias en este grupo.

**Pagos:** El número de pagos realizados tiene una correlación fuerte con la estabilidad del patrocinio. Patrocinadores con un historial más largo de pagos representan un segmento con alto potencial de conversión, ya que han demostrado compromiso financiero sostenido.

**EstadoCivil:** El estado civil es un factor relevante, con categorías como "Casado" mostrando una mayor predisposición a la conversión. Esto refuerza la idea de diseñar estrategias específicas para este segmento, destacando valores familiares y de continuidad en el patrocinio.

**GeneroPatro y Generacion:** La segmentación basada en género y generación es crucial. Por ejemplo, los patrocinadores de la Generación X y Millennials, con características demográficas distintas, podrían requerir enfoques diferentes para motivar la conversión.

**Región:** Las regiones específicas como el Centro y Occidente tienen particularidades en términos de conversión. Esto subraya la importancia de estrategias geolocalizadas que respondan a las dinámicas regionales de patrocinio.

**edadnino:** La edad del niño patrocinado influye en el comportamiento del patrocinador, con un mayor compromiso observado en etapas tempranas del patrocinio. Esto sugiere que los esfuerzos deben enfocarse en retener a patrocinadores cuando los niños alcanzan edades críticas, como 15 años o más para asegurar la conversión.

**Pledgetype y tipopatro:** Los tipos específicos de patrocinio están relacionados con la probabilidad de conversión. Patrocinios con enfoques más amplios o diversificados parecen tener mayor impacto en la retención y el aumento del compromiso.

**Patrones y relación entre variables,** los resultados del modelo confirman que la interacción entre variables como tiempo patrocinio, pagos y estado civil es fundamental para comprender el comportamiento de los patrocinadores. La correlación positiva entre estas variables respalda la idea de que un compromiso más prolongado y una relación más estable

conducen a un mayor potencial de conversión. Las motivaciones (Tmotivacion) y las modalidades de patrocinio (tipopatro) también son factores clave. Esto indica que es crucial entender las razones subyacentes del patrocinio para diseñar campañas dirigidas y optimizar la conversión.

Evaluación del resultado del modelo predictivo con probabilidad de cancelación e inactivación, para interpretar los resultados y analizar las variables más influyentes en la predicción del modelo de cancelación de patrocinadores, se deben considerar las métricas de desempeño, el impacto de las variables utilizadas y cómo estas influyen en las decisiones estratégicas para evitar que los patrocinadores cancelen su apoyo.

Métricas de desempeño , exactitud promedio del 88.40%: Esto indica que el modelo clasifica correctamente la gran mayoría de los casos, proporcionando una alta confiabilidad para identificar patrocinadores en riesgo de cancelación.

AUC-ROC de 90.56%: El modelo muestra una sólida capacidad para distinguir entre patrocinadores que cancelarán y aquellos que mantendrán su compromiso. Este nivel de desempeño sugiere que el modelo captura patrones críticos para predecir la inactivación.

Precisión promedio del 77.11%: Una proporción significativa de los casos clasificados como "en riesgo" efectivamente corresponde a patrocinadores con altas probabilidades de cancelar, lo que ayuda a focalizar recursos en estrategias de retención efectivas.

Recall promedio del 82.61%: El modelo identifica una amplia mayoría de los patrocinadores realmente en riesgo, asegurando que se puedan implementar acciones preventivas en los casos más relevantes.

Análisis de variables más influyentes, mora: Este es el indicador principal del riesgo de cancelación. Los patrocinadores con atrasos significativos (más de seis meses) muestran una alta probabilidad de inactivación.

La mora prolongada debe considerarse una señal crítica para implementar estrategias inmediatas, como recordatorios de pago personalizados o acuerdos flexibles.

tiempopatrocio: La duración del compromiso es un factor clave. Los patrocinadores con tiempos más cortos son más vulnerables a cancelar.

Las estrategias deben enfocarse en reforzar la conexión emocional con la causa durante los primeros años del patrocinio, utilizando narrativas de impacto y seguimiento frecuente.

Pagos: El historial de pagos refleja estabilidad financiera y compromiso. Los patrocinadores con pagos regulares y consistentes tienen menor riesgo de cancelar.

Es crucial identificar patrones de disminución en la frecuencia de pagos para intervenir antes de que la relación se deteriore.

Mpago: Los métodos de pago manuales, como consignaciones y depósitos, se asocian con un mayor riesgo de cancelación debido a la falta de automatización.

Promover el uso de métodos de pago automáticos, como débitos programados, puede reducir significativamente este riesgo.

Tmotivacion: Las motivaciones iniciales del patrocinador influyen en su probabilidad de cancelar. Aquellos que ingresan motivados por campañas emocionales suelen ser más leales, mientras que los captados por canales impersonales presentan mayores riesgos.

Personalizar las campañas según las motivaciones específicas puede fortalecer el compromiso.

**Edad:** Los patrocinadores tienden a cancelar cuando los niños alcanzan edades críticas, como 15 años o más, posiblemente debido a la percepción de menor impacto.

Es importante mantener la comunicación con historias de éxito o mostrar cómo el patrocinio sigue siendo relevante en esas etapas de transición.

**Región:** Las ubicaciones geográficas también influyen. Regiones con menor estabilidad económica presentan un mayor riesgo de cancelaciones.

Diseñar estrategias regionales adaptadas, que incluyan beneficios o incentivos específicos, puede ayudar a mantener el compromiso.

**Frecuencia:** La frecuencia de los pagos está directamente relacionada con el riesgo. Pagos menos frecuentes suelen ser un indicador de desinterés o dificultad financiera.

Incentivar pagos más regulares a través de planes de fidelización podría reducir la probabilidad de cancelación.

**Generación:** Las generaciones más jóvenes, como los Millennials, tienen un riesgo mayor de cancelar debido a cambios en prioridades financieras o desconexión con la causa.

Este segmento puede beneficiarse de estrategias innovadoras, como contenido digital interactivo o experiencias inmersivas.

**Patrones y Relación entre Variables:** El modelo confirma que la combinación de mora, tiempo de patrocinio y frecuencia de pagos es crucial para identificar patrocinadores en riesgo. La correlación entre estas variables destaca la importancia de detectar señales tempranas, como retrasos en pagos o compromisos recientes, para evitar la cancelación. Adicionalmente, factores como las motivaciones iniciales y los métodos de pago subrayan la necesidad de personalizar las intervenciones. Implementación de estrategias, estrategias de Marketing para Conversión de Patrocinadores (Unipatro a Multipatro), campañas Personalizadas Basadas en Motivaciones

Identificar las principales motivaciones de los patrocinadores unipatro utilizando los datos de Tmotivacion, priorizando aquellas relacionadas con programas de televisión y canales digitales. Diseñar campañas enfocadas en resaltar cómo un mayor apoyo multipatrocinado puede amplificar el impacto de estas motivaciones específicas.

Usar testimonios de casos de éxito vinculados a esas motivaciones para conectar emocionalmente y aumentar el compromiso.

## 2. Estrategias Regionalizadas

Aprovechar las diferencias regionales detectadas, como el mayor número de patrocinadores en la región Centro, para diseñar campañas focalizadas. Enfatizar el impacto local que puede lograrse con un compromiso más amplio.

Implementar eventos presenciales o virtuales para los patrocinadores actuales en estas regiones, mostrando cómo el apoyo multipatro beneficia directamente a sus comunidades.

## 3. Segmentación por Edad y Estado Civil

Desarrollar estrategias dirigidas a patrocinadores entre 30-40 años y casados, segmentos con alta predisposición a la conversión. Crear materiales que destaquen el impacto intergeneracional del apoyo adicional, apelando a valores familiares y estabilidad.

Introducir programas de "multipatrocinio familiar", ofreciendo opciones para que toda la familia participe en el patrocinio.

## Estrategias para Retención de Patrocinadores en Riesgo de Cancelar

1. Intervenciones Preventivas Basadas en Mora, identificar patrocinadores con mora entre 6 y 12 meses y diseñar estrategias de reactivación mediante mensajes personalizados, destacando historias de impacto generadas con sus contribuciones, ofrecer planes de pago flexibles y

recordatorios automatizados para facilitar la continuidad del apoyo sin que el patrocinador perciba presión financiera.

2. Refuerzo del Compromiso en Edades Críticas, enfocar esfuerzos en patrocinadores cuyos niños apadrinados estén próximos a los 15 años, destacando cómo su apoyo puede ser crucial para la transición del niño hacia una etapa de mayor autonomía y éxito educativo, presentar informes de impacto detallados y personalizados que resalten los logros del niño patrocinado como resultado directo de su contribución.

3. Optimización de Métodos de Pago, promover la migración de patrocinadores con métodos manuales hacia sistemas automatizados como ACH-Davivienda, destacando la conveniencia y consistencia de este método, implementar incentivos para quienes adopten métodos digitales, como acceso a contenido exclusivo sobre los proyectos apoyados.

Monitoreo y Ajustes:

Seguimiento Continuo de Predicciones

Implementar dashboards dinámicos que muestren en tiempo real el estado de conversión y retención, priorizando las métricas clave como mora, frecuencia de pagos y duración del patrocinio. Utilizar estos dashboards para evaluar la efectividad de las estrategias en tiempo real, ajustando las campañas según las tasas de conversión y cancelación observadas.

Retroalimentación Basada en Resultados

Realizar encuestas a los patrocinadores convertidos o retenidos para identificar factores que impulsaron su decisión, integrando estos insights en futuras estrategias.

Analizar continuamente la precisión de los modelos predictivos y recalibrarlos si se detectan patrones emergentes o inconsistencias.

Iteración Estratégica

Implementar ciclos trimestrales de evaluación donde los modelos predictivos sean actualizados con nuevos datos, garantizando que las estrategias sigan alineadas con el comportamiento reciente de los patrocinadores.

Generar reportes periódicos que comparen el impacto de las campañas implementadas con las predicciones iniciales, enfocándose en cerrar cualquier brecha identificada.

## Conclusiones

Identificación de patrones clave para la conversión de patrocinadores (Objetivo Específico 1)

El análisis predictivo permitió identificar que los patrocinadores con una duración de patrocinio mayor a 9.5 años y un historial de pagos consistente presentan una alta probabilidad de convertirse en multipatrocinadores. Este hallazgo se sustenta en el modelo de regresión logística, donde las variables tiempo patrocinio y Pagos mostraron las correlaciones más fuertes con la probabilidad de conversión (Tabla 5). Además, se evidenció que las motivaciones específicas, como programas de televisión y canales digitales, influyen significativamente en este proceso, siendo más predominantes en patrocinadores casados y pertenecientes a la Generación X (Figura 12)

Estos resultados aportan a la solución del problema al proporcionar bases para la implementación de estrategias segmentadas, tales como campañas personalizadas dirigidas a patrocinadores casados y mensajes enfocados en motivaciones específicas. Esto permitirá aumentar la tasa de conversión de unipatrocinadores a multipatrocinadores, optimizando los recursos y maximizando el impacto organizacional.

Predicción efectiva de patrocinadores en riesgo de cancelación (Objetivo Específico 2)

El modelo predictivo desarrollado identificó con una precisión promedio del 89% a los patrocinadores con alta probabilidad de cancelar su apoyo, mostrando la relevancia de las variables como la mora acumulada, el tiempo de patrocinio y los métodos de pago manuales, como consignaciones (Tabla 8). Además, se descubrió que los niños con edades superiores a 15 años representan un factor crítico en el riesgo de cancelación, posiblemente debido a la

percepción de menor impacto en esta etapa (Figura 4). También se destacó que regiones con menor estabilidad económica, como el Caribe, presentan mayores índices de cancelación.

Con base en estos hallazgos, se pueden implementar estrategias específicas, como la reducción de mora mediante recordatorios personalizados y el fortalecimiento del compromiso emocional de los patrocinadores con niños mayores de 15 años. Esto asegurará una base financiera más estable y contribuirá al sostenimiento de los programas sociales.

Visualización y segmentación como herramientas clave para la toma de decisiones  
(Objetivo Específico 3)

Las herramientas de visualización, como los mapas de calor de motivaciones y las matrices de correlación, permitieron identificar relaciones importantes, como la fuerte influencia de las motivaciones en las decisiones de los patrocinadores (Figura 2). Asimismo, el clustering reveló cuatro grupos principales de patrocinadores con características demográficas y comportamentales distintas, lo cual facilita la segmentación estratégica (Tabla 4). Por ejemplo, se destacó que las campañas regionalizadas tienen un impacto mayor en áreas como el Centro y Occidente, donde las motivaciones principales están vinculadas a programas presenciales y de televisión (Figura 3).

Estos hallazgos permiten a la organización dirigir recursos de manera más eficiente, maximizando las conversiones y reduciendo las cancelaciones. Además, la segmentación clara basada en datos fortalece el impacto social de las estrategias organizacionales y asegura un mejor uso de los recursos disponibles.

## **Recomendaciones**

### **Transparencia en el Uso de los Datos**

Establecer un sistema de documentación accesible para los patrocinadores y otros stakeholders que explique cómo se utilizan los datos, los objetivos del análisis predictivo y las medidas implementadas para salvaguardar su privacidad.

### **Infraestructura Tecnológica Escalable**

Garantizar que los servidores y bases de datos utilizados para la implementación de los modelos predictivos tengan capacidad suficiente para manejar la carga actual y futura de datos. Esto incluye la integración de herramientas de análisis en tiempo real y una arquitectura basada en la nube para mayor flexibilidad y escalabilidad.

### **Capacitación del Personal**

Proveer formación técnica al equipo encargado de gestionar y operar los modelos, garantizando que comprendan tanto las limitaciones como las fortalezas de las herramientas implementadas. Esto asegura decisiones más informadas y efectivas.

## Bibliografía

- Aguilar, U. (2021). *Análisis predictivo y multidimensional de la salud mental en estudiantes de medicina*. Obtenido de Investigación en educación médica:  
<https://www.redalyc.org/articulo.oa?id=349770251002>
- Brenda Díaz-Landa. (2021). *Análisis correlacional-predictivo de la influencia de la asistencia a clase en el rendimiento académico universitario*. Obtenido de Universidad de Granada:  
<https://www.redalyc.org/journal/993/99368569008/>
- ECBTI. (2016). *Líneas de Investigación ECBTI Universidad Nacional Abierta y a Distancia UNAD*. Obtenido de UNAD: <https://academia.unad.edu.co/investigacion-ecbti/cadenas-de-formacion>
- GROUP, G. (2017). *Big Data: una exploración de investigaciones, tecnologías y casos de aplicación*. Obtenido de Web of Science:  
<https://revistas.itm.edu.co/index.php/tecnologicas/article/view/685/671>
- Hernández, R. &. (2018). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta*. Obtenido de <https://www-ebooks7-24-com.bibliotecavirtual.unad.edu.co/?il=6443>
- Oswaldo Sifuentes Bitocchi. (2018). *Modelos predictivos de la deserción estudiantil en una universidad privada peruana*. Obtenido de Universidad Privada Peruana:  
<https://www.redalyc.org/journal/816/81658967008/>
- Sabiote, R. (2017). *Un análisis exploratorio y predictivo sobre las oporientaciones de meta y sobre el contenido de las meas de los*. Obtenido de Universidad de Granada:  
<https://redalyc.org/articulo.oa?id=97912446002>