

REVISTA PRISMA SOCIAL N° 46

LA TECNOLOGÍA DIGITAL Y
EL DESARROLLO SOCIOEDUCATIVO
EN LOS CONTEXTOS MUNICIPALES

3ER TRIMESTRE, JULIO 2024 | SECCIÓN TEMÁTICA | PP. 00-00

RECIBIDO: 8/5/2024 – ACEPTADO: 17/7/2024

HERRAMIENTAS DIGITALES
Y NIVEL ACADÉMICO EN
BOGOTÁ Y MUNICIPIOS
CIRCUNDANTES

DIGITAL TOOLS AND ACADEMIC
LEVELS IN BOGOTA
AND SURROUNDING MUNICIPALITIES

CESAR AUGUSTO CAÑÓN OSPINA / CESARA.CANONO@UTADEO.EDU.CO

UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO, BOGOTÁ, COLOMBIA

SANDRA PATRICIA BARRAGÁN MORENO / SANDRA.BARRAGAN@UTADEO.EDU.CO

UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO, BOGOTÁ, COLOMBIA

LEANDRO GONZÁLEZ TÁMARA / LEANDRO.GONZALEZ@UTADEO.EDU.CO

UNIVERSIDAD DE BOGOTÁ JORGE TADEO LOZANO, BOGOTÁ, COLOMBIA



prisma
social
revista
de ciencias
sociales

RESUMEN

Las herramientas digitales ostentan un rol crucial en la educación porque mejoran el acceso a la información, motivan al conocimiento, personalizan el aprendizaje, promueven la colaboración y proporcionan retroalimentación oportuna. Así, el objetivo de este artículo fue segmentar y caracterizar la población de Bogotá y sus municipios circundantes respecto al nivel académico y otras variables demográficas, considerando el uso de herramientas digitales y la conectividad. Como metodología se usó CRISP-DM implementando árboles de decisión para el modelado de los microdatos anonimizados de la Encuesta Multipropósito de 2021 cuyo operador de campo es el Departamento Administrativo Nacional de Estadística. La muestra consistió en 278.023 personas, representando un universo de 9'248.000 personas. El análisis descriptivo mostró la asociación de variables demográficas, uso de las tecnologías de información y comunicación, del empleo y los ingresos sobre el nivel educativo. El árbol de decisión mostró con un porcentaje de correcta clasificación del 82,8% que, las personas con un alto nivel educativo suelen ser mayores de 28 años, gastan significativamente en tecnología, tienen acceso a internet y lo usan para educarse.

PALABRAS CLAVE

Modelo estadístico; nivel de enseñanza; encuesta por muestreo; tecnología de la información; equipamiento de telecomunicaciones; gobierno municipal.

ABSTRACT

Digital tools play a crucial role in education as they improve access to information, motivate knowledge acquisition, personalize learning, promote collaboration, and provide appropriate feedback. Thus, the objective of this article was to segment and characterize the population of Bogotá and its surrounding municipalities regarding academic levels and other demographic variables, taking the use of digital tools and connectivity into consideration. CRISP-DM was selected as the methodology for implementing decision trees for modeling anonymized micro-data from the 2021 Multipurpose Survey, for which the National Administrative Department of Statistics was the field operator. The sample was composed of 278,023 people, representing a universe of 9,248,000 people in total. The descriptive analysis showed the association of demographic variables, the use of information technology and communication, employment, and revenue over academic levels. Subsequently, with the identified variables, the decision tree evidenced, with a correct classification percentage of 82,8%, that those people who have a high academic level tend to be older than 28 years old, spend a significant amount of money on technology devices, and software, have at least one place to access the internet and use it for educational purposes.

KEYWORDS

Statistical model; education level; sample survey; information technology; telecommunication equipment; municipal government.

1. INTRODUCCIÓN

En la última década, el panorama digital ha experimentado transformaciones significativas, generando un profundo impacto en la forma en que las sociedades y los individuos interactúan, aprenden, comunican y acceden a la información (Andrade *et al.*, 2023; Cicik Wijayanti *et al.*, 2024; Levano *et al.*, 2019; Miguel, 2020). El uso de las herramientas digitales como computadores, teléfonos inteligentes e internet se ha vuelto omnipresente en la vida diaria facilitando tareas como obtener información, trabajar, estudiar y realizar trámites (Hernández & Rayón, 2021). Más aún a nivel educativo,

La era digital se está imponiendo cada vez en las aulas de clases, con una amplia gama de herramientas didácticas, permitiendo a los educandos incrementar sus conocimientos por medio de vías rápidas y accesibles a la información de manera multidimensional (Mero-Ponce, 2021, p. 714).

Por lo anterior, se consideró que la disponibilidad de estas, al interior de los hogares, puede no solamente crear un espacio con mejor calidad de vida para las personas que hacen parte de estos, sino que también impulsa las posibilidades de acceso a la información, a la educación y al empleo (Guevara *et al.*, 2024; Miguel, 2020), lo que a largo plazo puede traer consigo tanto mejoras en la calidad de vida como en el aprovechamiento de oportunidades. Es en este marco que se consideró el propósito de la presente investigación.

1.1. PROPÓSITO DE LA INVESTIGACIÓN

La evolución del uso de las herramientas digitales, la conectividad y el acceso a internet han generado interés en la investigación social. Comprender esto permite a los responsables de políticas analizar el impacto en hogares sin acceso, afectando el nivel educativo y las oportunidades de formación e información (Pillajo *et al.*, 2023) y la obtención o ejecución de un empleo (Simões y Marta, 2024), en el contexto de las ventajas actuales que ofrecen las mencionadas herramientas.

A medida que las tecnologías digitales se integran en la vida cotidiana, la disparidad en el acceso a herramientas digitales se convierte en un objeto de estudio social (Shwedeh, 2024). La literatura ha identificado factores como la infraestructura tecnológica, la alfabetización digital y las diferencias socioeconómicas (Cicik Wijayanti *et al.*, 2024; Ventrella y Cotnam-Kappel, 2024). No obstante, las investigaciones que analizan estas disparidades segmentadas según el nivel académico son insuficientes, a pesar de que la educación juega un papel crucial para desarrollar habilidades digitales y oportunidades de acceso (Ventrella y Cotnam-Kappel, 2024).

La teoría de la brecha digital sugiere que, las diferencias en el acceso a la tecnología están intrínsecamente ligadas a las disparidades económicas y educativas (Nkansah y Oldac, 2024). La teoría del capital social digital, por otro lado, sugiere que el acceso a las herramientas digitales no es solo un tema de recursos, sino también de habilidades y conexiones en línea (Ventrella & Cotnam-Kappel, 2024; De Zúñiga *et al.*, 2022).

Estas cuestiones pueden ser abordadas desde la analítica de datos, utilizando información de repositorios de información gubernamentales, que proporcionan un gran volumen de datos

confiables para diferentes análisis y para comprender aspectos de la vida de los/as habitantes de un país como demografía, salud, educación o trabajo

Bogotá, epicentro de la diversidad cultural y económica en Colombia, está experimentando una rápida transformación digital. En este contexto, la brecha digital es un desafío fundamental que puede profundizar las disparidades existentes. Por lo cual, el propósito del análisis de la mencionada brecha digital enmarcado en teorías que abordan tanto las dimensiones tecnológicas como las sociales se observa como una oportunidad natural para la identificación y la aproximación a las desigualdades estructurales que limitan el acceso equitativo a las Tecnologías de Información Comunicación (TIC), así como para la promoción de estrategias que potencien el capital social y favorezcan el desarrollo comunitario sostenible.

1.2. ANTECEDENTES

Dado que los datos que se utilizaron para el desarrollo de este proyecto provienen de una fuente estadística, de carácter gubernamental, como lo es la Encuesta Multipropósito en su aplicación de 2021 (EM2021) (Departamento Administrativo Nacional de Estadística, 2022), es importante resaltar que esta recolección de datos masiva ha tenido un alto impacto a la hora de realizar un análisis profundo abarcando el mayor número de variables y escenarios que no son evidentes, esto proporciona conclusiones, acciones y finalmente un impacto en aspectos como el desarrollo de nuevo conocimiento, nuevas tecnologías o nuevas estrategias orientadas a la mejora de la calidad de vida de las personas, desde diferentes puntos de vista. Como ejemplo se tienen las siguientes investigaciones realizadas utilizando como fuentes de información encuestas de diferentes tipos y que, gracias al uso de estos datos, su procesamiento y su interpretación estadística, lograron generar conclusiones que potencialmente pueden poner la mirada sobre temas relevantes y que no han sido lo suficientemente explorados.

El potencial de exploración de datos provenientes de las encuestas multipropósito puede ser observado en Autores (2021), donde se pudo observar que al utilizar datos de la Encuesta Multipropósito y la Encuesta Bienal de Culturas del año 2017 para la ciudad de Bogotá, se encontró que existe una relación entre el nivel de educación de las personas y su comportamiento respecto a actitudes discriminatorias, estableciendo que la forma más común de discriminación se da hacia personas de orientaciones sexuales diversas, por encima de formas de discriminación por religión, raza, etnia, sexo e incluso discriminación por apariencia física. En esta investigación también se encontró que variables como el nivel educativo, el estrato socioeconómico de la vivienda, la edad y el sexo del observador pueden ser influyentes a la hora de identificar comportamientos discriminatorios hacia personas de los sectores LGBTIQ+.

De igual forma, en una investigación realizada por Autores (2022), basada en datos recopilados con la Encuesta Multipropósito en aplicaciones anteriores a 2021 y con enfoque a identificar los motivos por los que personas de los sectores LGBTIQ+ podrían entrar en procesos de deserción académica. Autores (2022) observaron que variables como la edad, el sexo y el tipo de tenencia de vivienda influyeron en la decisión de entrar a estudiar o no, sobre esta población. Se identificaron las variables como los costos educativos, la necesidad de trabajar, y la consideración de culminación de los estudios, como causas de la deserción académica en dicha población.

De manera amplia, también existen estudios sobre los beneficios del uso de dispositivos digitales. En un estudio realizado por Gatto y Tak (2008) se halló que, cuando un grupo de adultos/as mayores se expuso al trabajo con un computador se revelaron barreras y beneficios proporcionando una alternativa para introducir estos dispositivos a esta parte de la población y demostrando que esta tarea es realizable, ya que puede ser ejecutada por cuidadores/as o trabajadores/as de la salud que interactúen con estas personas. Entre los beneficios encontrados se obtuvo el aprendizaje sobre herramientas como buscadores en línea, compras en línea, envío de correos y actividades de entretenimiento.

Dada la importancia de las herramientas digitales en procesos académicos como lo resaltan Mucundanyi y Woodley (2021) y Greener y Wakefield (2015), se deben considerar las creencias, los sesgos o la falta de habilidades de los/as profesores/as, ya que pueden oponerse a la implementación de las herramientas digitales en algunas regiones. Mohammed *et al.* (2019) y Esteve-Mon *et al.* (2023) fomentaron desde sus estudios diferentes modelos de implementación de estas herramientas en las aulas de clase, promoviendo un acompañamiento institucional a los/as profesores/as en la implementación de estas herramientas y teniendo en cuenta estímulos en el uso de las herramientas para los/as estudiantes en actividades colaborativas.

En complemento a lo anterior, es importante resaltar que la experticia en el uso de herramientas digitales por parte de los/as profesores/as facilita el aprovechamiento de ellas y genera confianza en el/la estudiante para realizar tareas de búsqueda autónoma de información y de autoaprendizaje (García-Martínez *et al.*, 2020).

Como lo mencionaron Hrynevych *et al.* (2021), en el marco de la cuarta revolución industrial, la implementación de las herramientas digitales en las aulas de clase resulta vital para los procesos de aprendizaje de los/as estudiantes en distintos niveles, ya que, al promover el aprovechamiento de estas, las actividades de aprendizaje pueden enriquecer los conocimientos de las nuevas generaciones con perspectivas y herramientas tecnológicas diferentes como la inteligencia artificial, los software de simulación y los laboratorios virtuales, lo cual mejora no solamente la calidad en la educación que reciben sino también sus oportunidades a futuro como profesionales o al ingreso a una etapa laboral.

1.3. MARCO TEÓRICO

1.3.1. Fuente de datos

La fuente de datos seleccionada para esta investigación proviene de la EM2021 aplicada en Bogotá D.C. y en 21 municipios del departamento de Cundinamarca en la Región Andina de Colombia. La aplicación en Bogotá D.C. se hizo considerando la subdivisión en 80 Unidades de Planeación Zonal (UPZ), 15 agrupaciones de UPZ, y 20 localidades (19 urbanas y una rural). El formulario de la encuesta constó de 13 capítulos, en los que se recopilaban datos relacionados con la identificación de la vivienda, el entorno de la vivienda, las condiciones habitacionales del hogar, los servicios públicos en el hogar, composición demográfica, salud, atención a niños menores de 5 años, educación, fuerza de trabajo, uso de las TIC, uso de redes sociales, percepción de condiciones de vida, gastos en alimentos y gastos del hogar.

En total, estos 13 capítulos administraron 1.606 preguntas, de las que para efectos del objetivo de este artículo, se seleccionaron 143 que gravitaron en torno a atributos o variables sobre los niveles educativos y el uso de herramientas digitales (Departamento Administrativo Nacional de Estadística, 2022). Es decir que, en la base de datos anonimizada consolidada para la EM2021 fueron empleados esta investigación: Composición del hogar y demografía, Salud, Educación, Uso de tecnologías de la información, TIC y Fuerza de trabajo. Con estos capítulos se seleccionaron variables relevantes para la variable de salida NIVEL_EDUCATIVO, de tipo cadena de texto. Considerando las bondades de esta base de datos se formuló el objetivo de investigación.

1.4. OBJETIVO DE INVESTIGACIÓN

El presente artículo tuvo como objetivo segmentar y caracterizar la población de Bogotá y sus municipios circundantes (alrededor) respecto al nivel académico y otras variables demográficas, considerando el uso de herramientas digitales y la conectividad.

1.5. JUSTIFICACIÓN

Esta investigación se basó en datos recabados a través de la EM2021, proporcionando una panorámica actualizada y detallada. En Colombia, el Departamento Administrativo Nacional de Estadística (DANE) es la entidad gubernamental líder en actividades de recolección y análisis de datos demográficos con fines estadísticos. Esta entidad lidera y ejecuta múltiples procesos de recolección de datos mediante encuestas por muestreo y otras modalidades, generando diversas fuentes de información como la Encuesta de Calidad de Vida, la Encuesta Nacional de Hogares, la Encuesta de Seguridad y Convivencia Ciudadana y los censos poblacionales.

La EM2021, como parte de este abanico de encuestas del DANE, es una fuente significativa de datos demográficos y de otros ámbitos, abordando temas como educación, empleo, salud o tecnologías digitales. Se consideró la EM2021 como un insumo valioso para el análisis y el modelado de datos, enfocándose en acceso a herramientas digitales, TIC, y distintas variables demográficas relacionadas con la calidad de vida de las personas y con el nivel educativo de estas. El capítulo I de la EM2021 se dedica al uso de TIC con 83 preguntas específicas.

Esta investigación puede aportar en el entendimiento de la brecha digital en Bogotá, al explorar específicamente las disparidades según el nivel académico. La originalidad radica en la aplicación de un enfoque segmentado y en la modelación estadística para la identificación de patrones únicos y desafíos específicos que no son evidentes en estudios más amplios. Con base en los resultados de la EM2021, este estudio propició no solo cuantificar la brecha, sino también la comprensión de las complejas interacciones entre el nivel académico, el acceso a herramientas digitales y la conectividad en la capital colombiana.

2. DISEÑO Y MÉTODO

Tanto las actividades de implementación de metodologías para la definición del universo a estudiar como la muestra representativa de este universo y las metodologías para la captura de datos están debidamente documentadas en la Ficha Técnica para la EM2021, redactada y editada por el Equipo temático y Técnico de la Encuesta de Multipropósito del Dirección de

Metodología y Producción Estadística (DIMPE) (Dirección de Metodología y Producción Estadística - DIMPE, 2022).

En los apartados de Ámbito de estudio y Muestra, incluidos más adelante en este artículo se resumió la información concerniente en la ficha técnica ya mencionada.

2.2. OBJETO FORMAL

Como se indicó antes, esta investigación tuvo como objetivo segmentar y caracterizar la población de Bogotá y sus municipios circundantes respecto al nivel académico y otras variables demográficas, considerando el uso de herramientas digitales y la conectividad.

2.2. TIPO DE INVESTIGACIÓN

La presente investigación es de tipo cuantitativo cimentada en la información disponible a partir la recolección de datos en la EM2021, la cual es de tipo probabilístico, multietápico, estratificado y de conglomerados (Dirección de Metodología y Producción Estadística - DIMPE, 2022).

2.3. HIPÓTESIS

El acceso a las herramientas digitales crea oportunidades de mejora en la calidad de vida de las personas que tienen acceso a estas, sea en el hogar, en la academia o en su lugar de trabajo. Esta hipótesis sugirió que, el acceso a estas herramientas puede traducirse a largo plazo en mejores oportunidades en los ámbitos académico y laboral, en vista de que promueve y facilita la búsqueda de información de manera autónoma repercutiendo en la calidad de vida de las personas.

2.4. ÁMBITO DE ESTUDIO

2.4.1. Universo

De acuerdo con la ficha metodológica de la EM2021, el universo de trabajo estuvo conformado por la población civil no institucional residente en la cabecera entendida como la parte central de un municipio, centro poblado y rural disperso de Bogotá, Soacha, Chía, Cota, Funza, La Calera, Mosquera y Sopó. Asimismo, exclusivamente la cabecera de los municipios de Facatativá, Zipaquirá, Fusagasugá, Madrid, Cajicá, Sibaté, El Rosal, Tocancipá, Tabio, Tenjo, Bojacá, Gachancipá, Subachoque y Zipaquirá. La población total que conforma este universo es el resultado de aplicar un factor de expansión asignado a cada registro de la muestra encuestada inicialmente, obteniendo un universo compuesto por un total de 8'519.127 personas distribuidas como en la Tabla 1.

Tabla 1. Distribución poblacional por sexo de la muestra encuestada y el universo obtenido a partir del factor de expansión.

Sexo	Muestra	Expansión	Porcentaje
Hombre	118.776	4'046.966	46,11%
Mujer	138.818	4'472.161	53,89%
Intersexual	0	0	0%
Total	257.594	8'519.127	100%

2.4.2. Unidades de trabajo

Las unidades de observación y de análisis fueron las unidades de viviendas, hogares y personas. Más aún, las unidades de muestreo se representaron bajo la Unidad Primaria de Muestreo (UPM) representando un conglomerado de 10 viviendas en promedio, también llamadas segmentos. Las muestras utilizaron criterios probabilísticos, estratificados y de conglomerados.

2.5. MUESTRA

La muestra dada en segmentos conformadas por 10 viviendas en promedio estuvo compuesta así:

- Bogotá cabecera: 8.011 segmentos
- Bogotá resto: 496 segmentos
- 21 municipios de Cundinamarca: 1.401 segmentos
- Resto de 7 municipios de Cundinamarca: 416 segmentos

Este muestreo contó con método probabilístico construido a partir de información estadística de densidad de viviendas en manzanas y secciones rurales del Marco Geoestadístico Nacional (MGN) y fue ejecutado utilizando la técnica de barrido en donde se encuestaron a todas las personas dentro de todas las viviendas del segmento, incluso si alguna de estas personas no se encontraba prevista para su participación el día de la captura de información (Departamento Administrativo Nacional de Estadística-DANE, 2018).

2.6. VARIABLES DE ANÁLISIS

A partir de las 83 variables correspondientes al capítulo sobre TIC (capítulo I), se seleccionaron las variables de análisis: Nivel Educativo Alto, Edad, Sexo, Municipio, Frecuencia de Uso del Internet, Uso del Internet para la Educación, Acceso a Internet en el Hogar, Razón de no Estudio, Estudia, Ingreso del mes pasado por trabajo, Cantidad de lugares para acceso a internet, Valor de compra de computadores, dispositivos y software anual, y Satisfacción en la Educación.

2.7. TÉCNICAS DE INVESTIGACIÓN UTILIZADAS

La técnica de investigación utilizada para el objetivo de esta investigación, se basa en la aplicación de la metodología *Cross-Industry Standard Process for Data Mining* (CRISP-DM) para la ciencia de datos que, consta de una serie de pasos que comprenden desde el entendimiento del contexto de los datos, la limpieza, estandarización y depuración de estos para garantizar que el trabajo con estos genere un producto confiable y finalmente la implementación de un modelo de inteligencia artificial o de regresión a partir del que se puedan generar conclusiones o predicciones a raíz de los datos con los que se trabajó (Wirth y Hipp, 2000).

Para esta investigación, la aplicación de la metodología CRISP-DM comprendió los pasos desde entender cómo se recogieron los datos en la EM2021 y con qué objetivo, para posteriormente recopilar en un solo conjunto de datos todas las variables relevantes para el objetivo de este trabajo.

Seguido de esto, un proceso de limpieza y estandarización de datos fue implementado para identificar datos atípicos, faltantes y erróneos para garantizar que los datos con los que se trabajó el modelo analítico posterior fueran de confianza.

Previo a la ejecución de un modelo analítico, se realizó una caracterización de los datos disponibles, esto para establecer el contexto actual de los habitantes de Bogotá y los municipios de Cundinamarca en relación con su nivel académico, asociado al uso de herramientas digitales y las variables demográficas seleccionadas.

Enseguida, se implementó una técnica de análisis y modelamiento de datos con la que se realizaron las clasificaciones o segmentaciones sobre los datos disponibles, teniendo en cuenta las características de los datos provenientes tanto de las variables de entrada como de la variable de salida, y de esta manera, generando reglas de clasificación que permitan segmentar estos datos en grupos que presenten rezago o atraso en el nivel académico. La técnica que se utilizó para ello fue el modelamiento mediante árboles de decisión. Los árboles de decisión son modelos estadísticos que permiten realizar segmentaciones de los datos a partir de las características de estos y encuentran la manera óptima de subdividir estos datos en agrupaciones más pequeñas con características similares (Kingsford y Salzberg, 2008).

3. TRABAJO DE CAMPO Y ANÁLISIS DE DATOS

Como se encuentra documentado en la Ficha Técnica para la EM2021 (Dirección de Metodología y Producción Estadística - DIMPE, 2022), a cada integrante del grupo de recolección se le asignaron los segmentos o viviendas a encuestar, teniendo en cuenta su tamaño y la distancia entre manzanas. Mediante la técnica de barrido, cada integrante del grupo de recolección realizó una visita a las viviendas del segmento para aplicar el cuestionario a todo integrante de los hogares en una vivienda. Este proceso se llevó a cabo entre abril y noviembre de 2021.

Los análisis de datos i para esta investigación fueron los siguientes:

1. Análisis descriptivo: Utilizando visualizaciones para establecer relaciones entre variables.
2. Modelación estadística:

La modelación estadística se hizo mediante árboles de decisión relacionando atributos con el nivel educativo de la población. Esto permitió categorizar datos y etiquetar segmentos vulnerables o rezagados en cuanto a nivel educativo, generando la estructura del árbol de decisión, produciendo conclusiones y hallazgos explicados al final de este artículo.

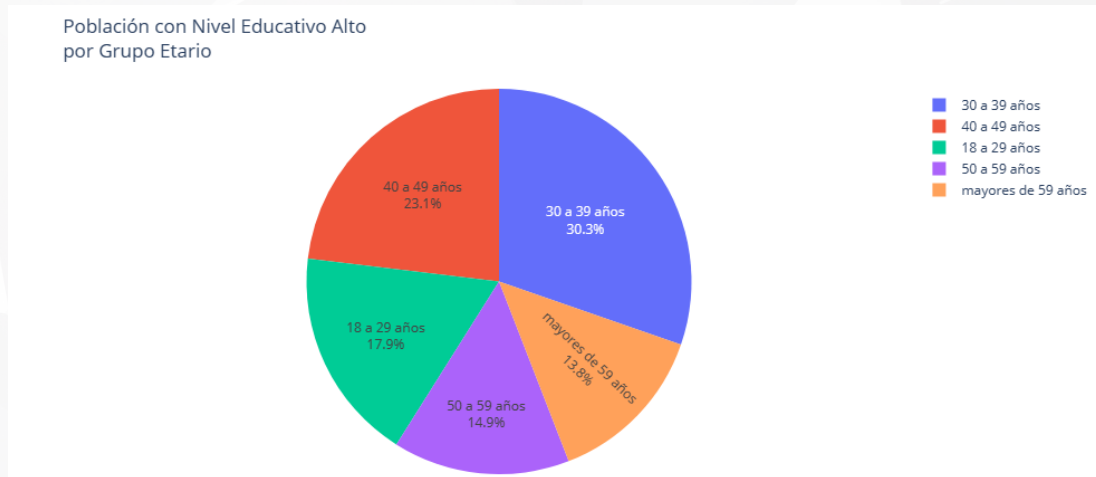
4. RESULTADOS

Para segmentar y caracterizar la población de Bogotá y sus municipios circundantes respecto al nivel académico y otras variables demográficas, considerando el uso de herramientas digitales y la conectividad, se incorporaron una serie de visualizaciones para comparar el comportamiento de las variables, con el fin de establecer qué relaciones existen entre ellas.

Inicialmente, se abordó la caracterización de las personas desde un punto de vista etario, con el fin de establecer si existe alguna relación entre la edad y los niveles académicos, así como entre la edad y el uso de TIC. Bajo la definición de un nivel académico bajo que comprendió a todos los niveles inferiores al pregrado finalizado y un nivel académico alto que agrupó al pregrado

finalizado y subsiguientes, la Figura 1 muestra que, al realizar una distribución por grupos etarios dentro de los niveles académicos, el grupo de edad que presentó una mayor población con un nivel académico alto fue el grupo de los 30 a 39 años, en contraste, las personas mayores de 59 tuvieron una menor proporción con nivel académico alto.

Figura 1. Distribución de niveles educativos por grupo etario.

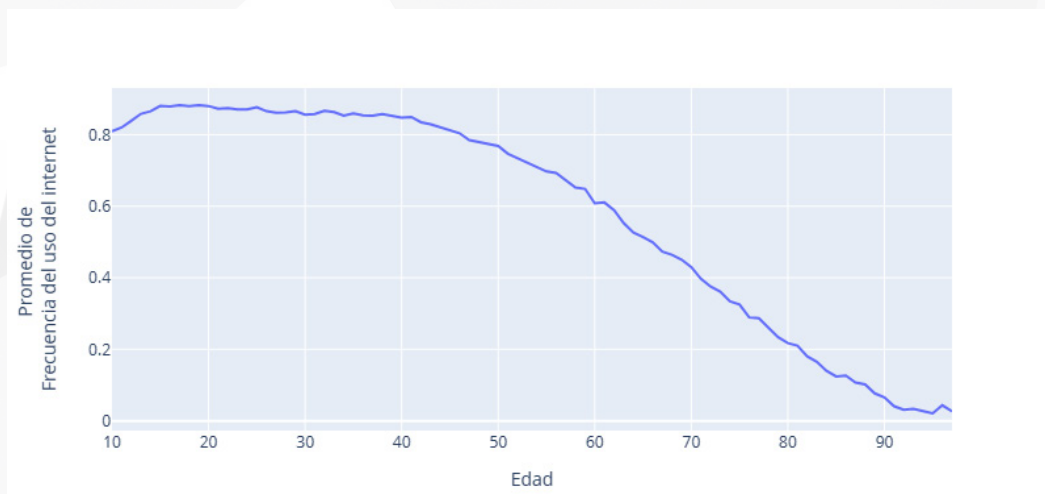


Fuente: Elaboración propia con base en DANE (2021).

De manera similar, para las visualizaciones que comparaban la variable de edad con la variable de Frecuencia del Uso del Internet, se requirió una codificación de esta última de manera que pudiera ser interpretada por el árbol de decisión. En un principio se encontraba etiquetada de 1 a 5, siendo 1 un uso diario del internet y 5 un uso nulo de este, hecho que se consideró incongruente al tener al valor mayor para representar el mínimo de la frecuencia. Lo anterior podía dar cabida a errores en la interpretación de esta variable, por lo cual se decidió etiquetarla usando valores de 0 a 1, siendo 0 un uso nulo, 1/365 uso una vez al año, 12/365 uso una vez al mes, 60/365 uso una vez a la semana y 1 un uso diario.

La Figura 2 presenta el comparativo de la variable de edad contra la variable de frecuencia del uso del internet. Esta frecuencia, en promedio, es más alta y estable en un rango de edad desde los 15 años hasta los 41 años, punto en el mostró un declive conforme avanza la edad.

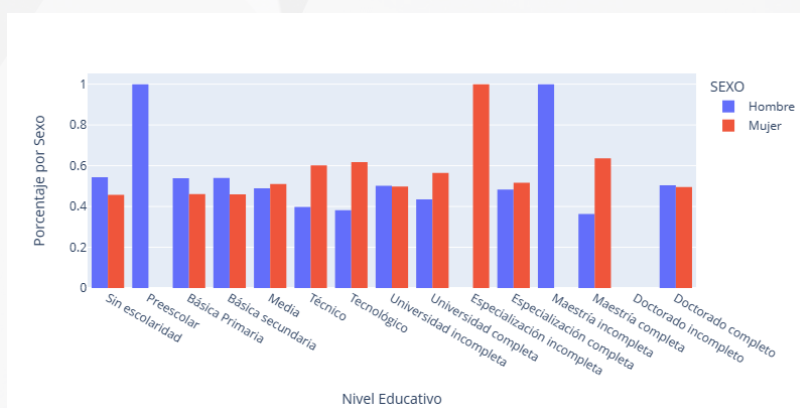
Figura 2. Frecuencia promedio del uso del internet según edad.



Fuente: Elaboración propia con base en DANE (2021).

Otra de las variables demográficas de interés que se utilizaron para realizar las caracterizaciones fue la variable de sexo. Esta variable categoriza el sexo de las personas encuestadas como hombre, mujer o intersexual, sin embargo, dentro de la muestra no se encontró población identificada como intersexual, por lo que, los análisis hechos sobre esta variable solamente tomaron en cuenta a aquellas personas identificadas como hombre o mujer. Al comparar esta variable con la de nivel educativo, se observó que, en niveles educativos superiores a la educación media, las mujeres estuvieron más presentes en la proporción de estudiantes, siendo ellas la mayoría de la población en seis de los diez niveles educativos y los hombres la mayoría en tres de ellos. La muestra no capturó información de mujeres con doctorado sin completar (ver Figura 3).

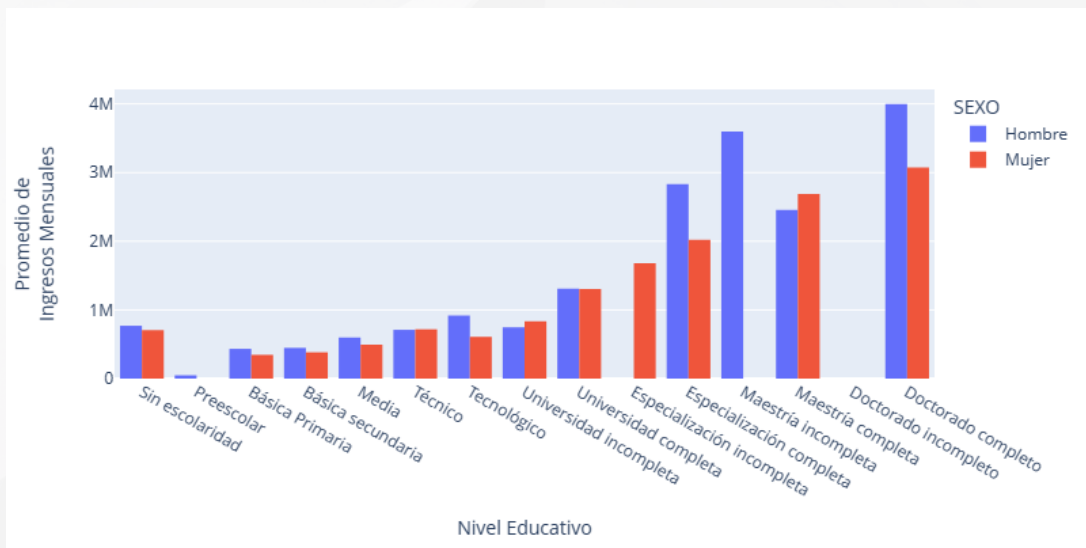
Figura 3. Nivel educativo por sexo.



Fuente: Elaboración propia con base en DANE (2021).

Al analizar la variable de sexo y el promedio salarial por nivel educativo, se encontró que el promedio salarial aumenta conforme aumenta el nivel educativo. Sin embargo, como se muestra en la Figura 4, se identificó que en niveles donde no hay escolaridad, el promedio salarial se asemeja al promedio salarial de un nivel de técnico o tecnológico, indicando que allí, el nivel educativo no está comprometido con el ingreso salarial de las personas. En otras palabras, en este grupo, el ingreso salarial no necesariamente está ligado solamente a su nivel educativo, si no que otras variables influyen (Ej.: La experiencia laboral puede tener aquí un mayor peso). Adicionalmente, se encuentra que, en diez de los quince niveles educativos, son los hombres quienes devengan más ingresos, dejando a las mujeres con ventaja solamente en cuatro de los niveles, lo que puede ser una representación de la brecha salarial por sexo en Bogotá y en los municipios de Cundinamarca.

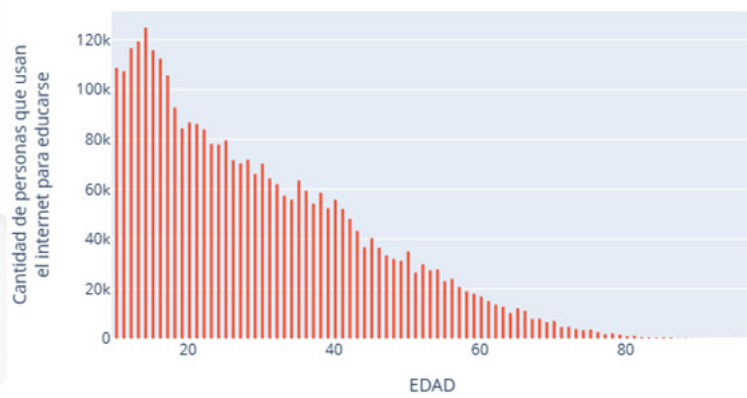
Figura 4. Promedio salarial por nivel educativo más alto alcanzado.



Fuente: Elaboración propia con base en DANE (2021).

La EM2021 del DANE también suministró datos sobre cómo las personas encuestadas utilizaron el internet, ya sea para fines educativos, laborales o para el entretenimiento. Procesando estas variables, se encontró que las personas que respondieron que utilizaban el internet para educarse están acumuladas hacia edades menores, siendo los 14 años la edad en la que las personas utilizan más el internet para la educación. De forma similar a lo observado en la Figura 2, se evidenció que, a medida que se alcanza una edad mayor, el uso de internet para la educación disminuye (ver Figura 5).

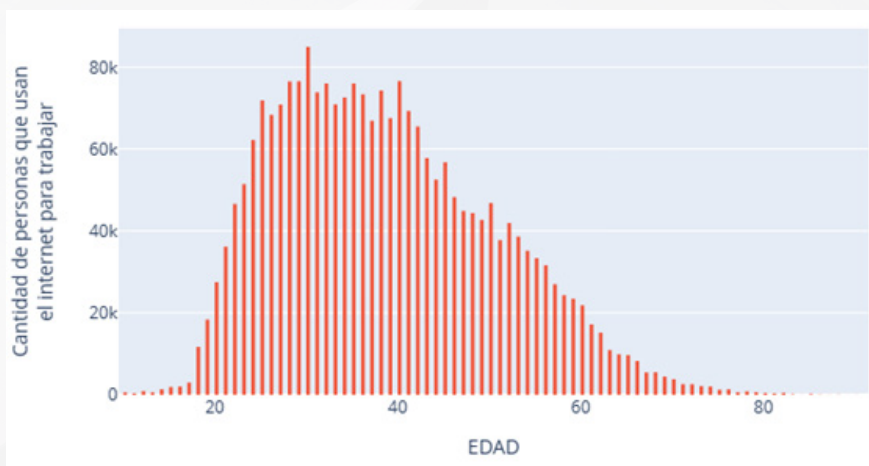
Figura 5. Uso del internet para la educación según la edad.



Fuente: Elaboración propia con base en DANE (2021).

Como se observó en la Figura 5, las personas que utilizaban el internet para educarse son en mayoría adolescentes con edades alrededor de los 14 años. En complementación, la Figura 6 muestra que el grupo de edad que utiliza el internet para trabajar ya no se encuentra en un rango de edad adolescente o juvenil, sino en un rango de edad adulta, superior a los 18 años y que se extiende hasta alrededor de los 63 años. Lo anterior puede estar relacionado con la entrada a una etapa productiva de las personas de estas edades y las oportunidades de empleo en línea que existen en la actualidad.

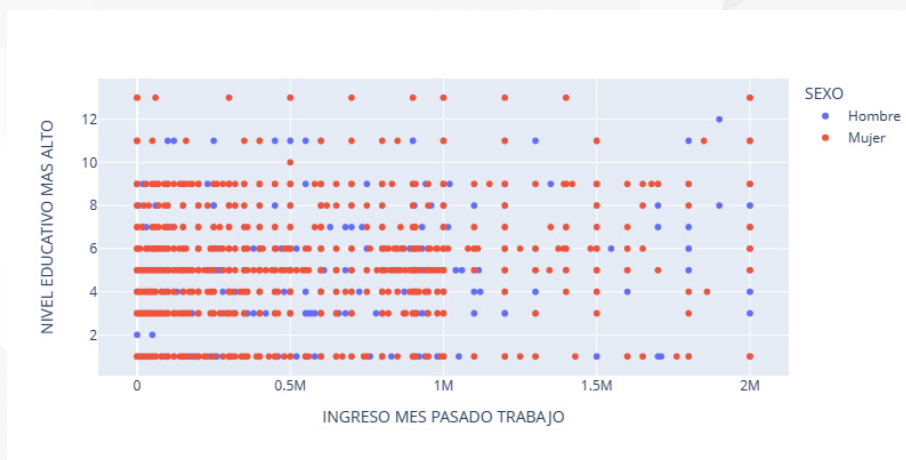
Figura 6. Uso del internet para trabajar según la edad.



Fuente: Elaboración propia con base en DANE (2021).

Considerando ahora, la variable de ingresos mensuales junto con la variable de nivel educativo más alto alcanzado, se encontraron personas con escolaridad nula que devenguen el mismo salario que personas que alcanzaron niveles de maestría completa (Figura 7), a manera de referencia este salario se encontraba alrededor de 511 USD al mes. Lo anterior puede sugerir que un nivel educativo más alto no necesariamente es un indicador de ingresos más altos.

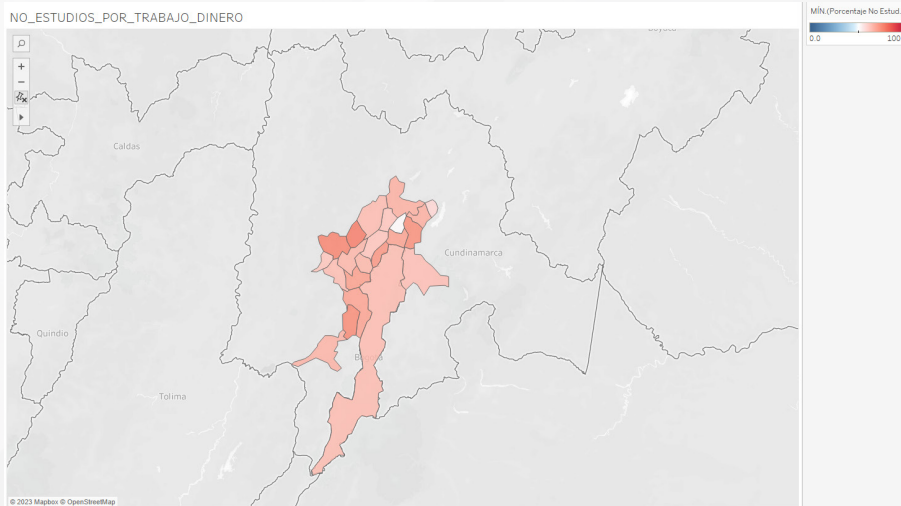
Figura 7. Relación ente Ingresos mensuales y nivel educativo.



Fuente: Elaboración propia con base en DANE (2021).

Para el análisis de las causas por las que las personas dejaron de estudiar en los municipios se hizo el mapa de calor de la Figura 8. En el mapa se evidencia que, tanto en los municipios como en la ciudad de Bogotá, las principales razones para no estudiar que manifiestan las personas es que no cuentan con los recursos necesarios para hacerlo o porque deben trabajar. En otras palabras, las principales razones para no estudiar en los municipios incluidos en la investigación corresponden a un determinante socioeconómico y están asociadas a la generación de recursos monetarios en más del 50% en todos los municipios (Mínimo: Cajicá 50,79%, Máximo: El Rosal 76,56%).

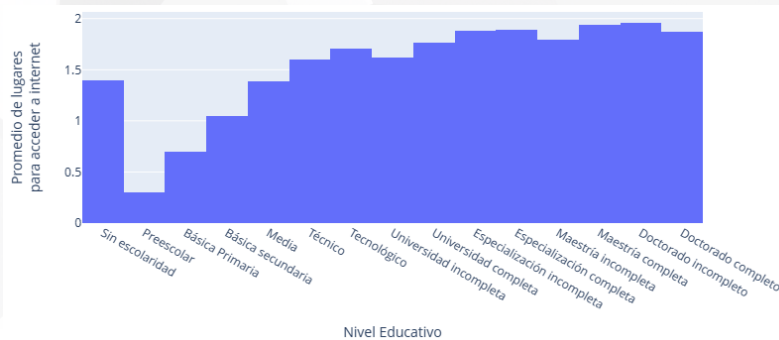
Figura 8. Mapa de calor de motivo de no estudio por trabajo.



Fuente: Elaboración propia con base en DANE (2021).

Al realizar una comparación entre el nivel educativo de las personas y el promedio de la cantidad de lugares en los que ellas pueden acceder a internet, se encontró que conforme aumenta el nivel educativo, también aumenta este promedio de lugares. Las personas con nivel de doctorado incompleto tenían el más alto con un promedio de 1,96 lugares para acceder a internet. Llama la atención el comportamiento de esta variable para el nivel educativo Sin escolaridad, ya que, como se observa en la Figura 9, en este nivel la cantidad de lugares para acceder a internet se asemeja a niveles como el de la educación Media y la Técnica. Lo anterior puede relacionarse con el uso de internet para actividades diferentes a las académicas, como las de entretenimiento o las laborales.

Figura 9. Nivel educativo respecto a la cantidad de lugares para acceder a internet.

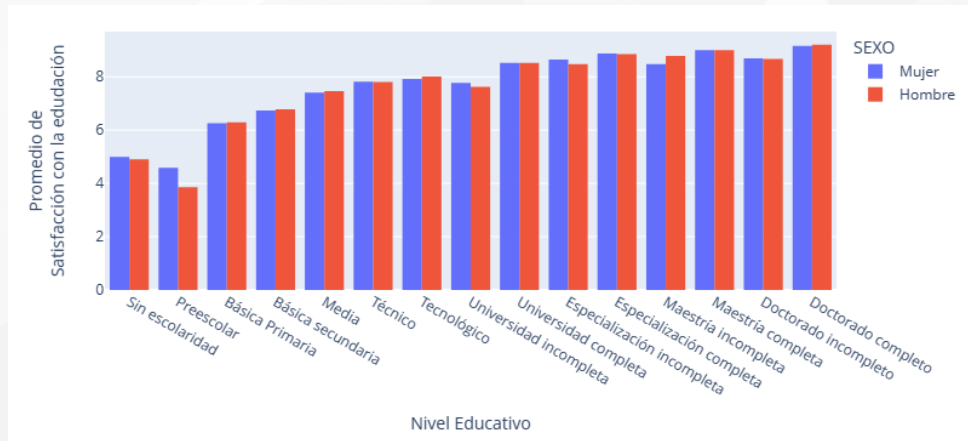


Fuente: Elaboración propia con base en DANE (2021).

Dentro de las variables disponibles en la EM2021, se observó una que describe la satisfacción de la muestra encuestada respecto de su educación. Comparando el promedio de esta variable con los niveles educativos, se encuentra que existe una relación entre la satisfacción en la educación de las personas y su nivel académico. La Figura 10 muestra que, aquellas personas con menor satisfacción se encuentran en los niveles más bajos de escolaridad, sin escolaridad y

preescolar. A medida que los niveles educativos crecen, la satisfacción en la educación también lo hace.

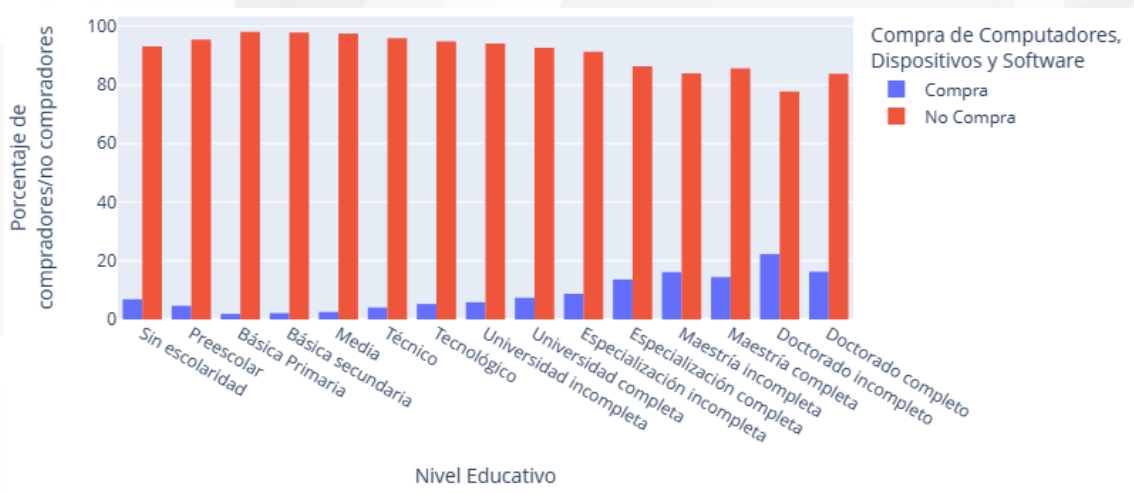
Figura 10. Promedio de satisfacción en la educación respecto a nivel educativo más alto alcanzado.



Fuente: Elaboración propia con base en DANE (2021).

Otro grupo de variables que se identificaron dentro de la EM2021 están relacionadas con la compra y adquisición de computadores y software. Al indagar si la compra de estos recursos tiene alguna incidencia sobre el nivel educativo de las personas, se encontró que la mayoría de ellas, en todos los niveles educativos, no compran computadores, dispositivos o software. Sin embargo, en los niveles educativos más altos, la compra de estos insumos está más presente (Figura 11).

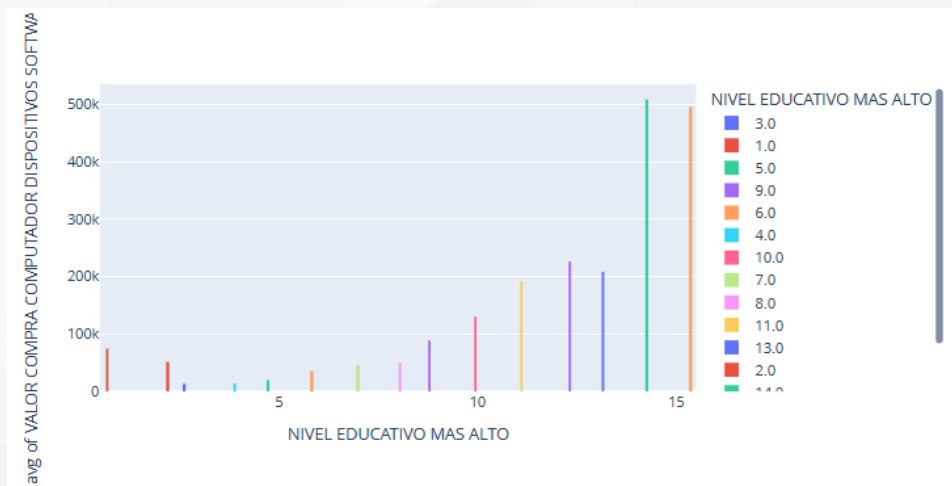
Figura 11. Porcentaje de compradores de computadores y software por nivel educativo.



Fuente: Elaboración propia con base en DANE (2021).

Adicionalmente, considerando la variable de valor anual de inversión en computadores y software, la Figura 12 revela que, las personas que se encuentran en niveles educativos más altos suelen invertir cantidades mayores de dinero en dispositivos, como computadores, tablets o software, en comparación con personas que tienen niveles educativos más bajos.

Figura 12. Valor anual promedio de inversión en computadores y software por nivel educativo.



Fuente: Elaboración propia con base en DANE (2021).

Con base en las figuras 1 a 12 se definieron las variables independientes y la variable dependiente que sirvieron para realizar la parametrización del árbol de decisión. Las variables independientes seleccionadas fueron: la edad, el sexo, la frecuencia de uso del internet, el uso del internet para educación, el municipio de origen, la existencia acceso a internet en el hogar, la razón para no estudiar, la cantidad de lugares para acceso a internet, el valor invertido en dispositivos digitales o software adquiridos por las personas en Bogotá y los municipios circundantes de Cundinamarca.

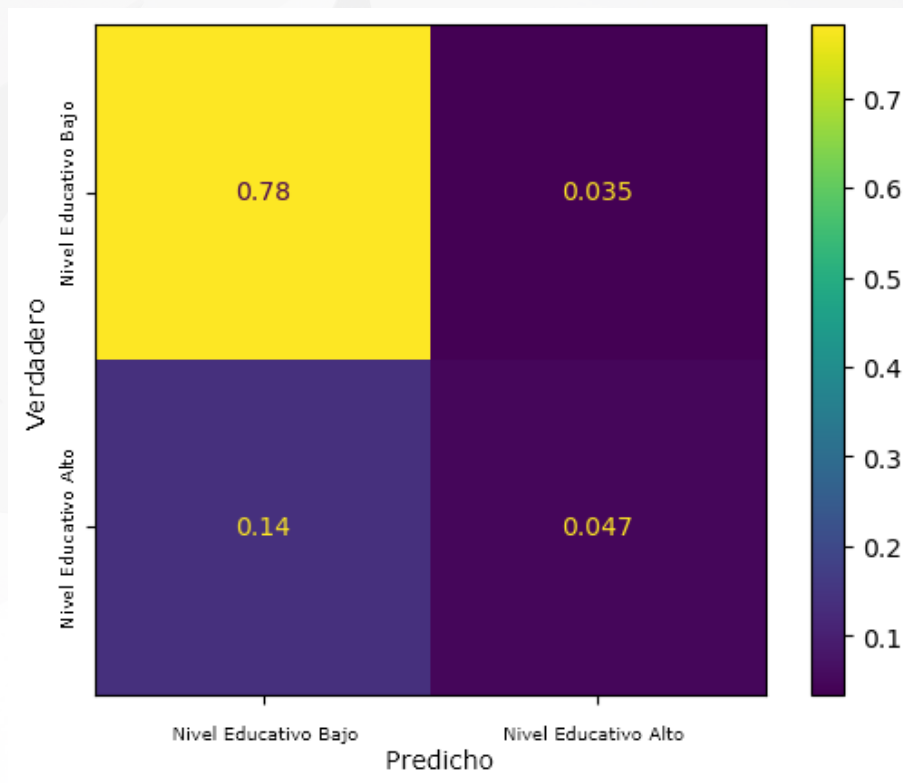
Es de anotar que, conforme al análisis realizado, se descartaron variables demográficas y del uso de las TIC que demostraron no ser indicadoras o caracterizadoras de una persona con nivel educativo alto. Estas variables son los ingresos mensuales y la compra de dispositivos digitales o software. Asimismo, variables como la pertenencia a la población de los sectores LGBTQ+ o a una etnia no resultaron relevantes pues la participación de las personas con estos rasgos dentro de la muestra encuestada fue muy pequeña. Finalmente, con el análisis de la variable de Satisfacción en la educación, se pudo observar que esta puede tener un comportamiento explicativo de la variable de Nivel Educativo Alto, ya que, al tener un nivel educativo más alto, también se tendrá un nivel de satisfacción mayor, aunque esta relación puede ser una realimentación positiva entre ambas variables, es decir, a mayor satisfacción con la educación, se busca tener un mayor nivel educativo y a mayor nivel educativo, mayor es la satisfacción con la educación por los beneficios sociales e individuales que ella retorna.

El árbol de decisión generado a partir de las variables independientes (de entrada) seleccionadas, con respecto de la variable de salida tuvo un nodo padre que inició el recorrido del árbol a partir de la variable de edad, al preguntar si una persona es menor de 23 años. A partir de este nodo, el árbol comenzó a generar otros nodos evaluando un valor específico en cada uno de estos para variables como sexo, la cantidad de lugares para acceder a internet, si una persona vive en un municipio, el valor que la persona invierte en computadores y software, si una persona utiliza el internet para educarse y los ingresos mensuales de la persona.

Finalmente, el árbol arrojó como resultado que aquellas personas que tenían un nivel educativo alto suelen invertir más en dispositivos y software, tenía por lo menos un lugar para acceder a

internet, lo utilizaba para educarse y tenía 29 años o más. Del mismo modo, este árbol también mostró que las personas que, con edades menores a 25 años, aunque utilizaban el internet para educarse y tenían por lo menos 2 lugares para utilizar el internet, si no invierten dinero en dispositivos y software, se asociarían con un nivel académico bajo. Es de anotar que la garantía en la calidad de la generación del árbol de decisión recae en métricas como su porcentaje de correcta clasificación del 82,85%, representada como la suma de los porcentajes de los registros clasificados como verdaderos positivos y verdaderos negativos en la matriz de confusión presentada en la Figura 13.

Figura 13. Matriz de confusión para el árbol de decisión generado.



Fuente: Elaboración propia.

De igual forma, implementando un método de validación cruzada, se encontró que la precisión del modelo generado supera el promedio de métricas de precisión de los modelos propuestos por este método, al ser este promedio del 82,53% de correcta clasificación. Las medidas de precisión para cada modelo propuesto por el método de validación cruzada se presentan en la Tabla 2.

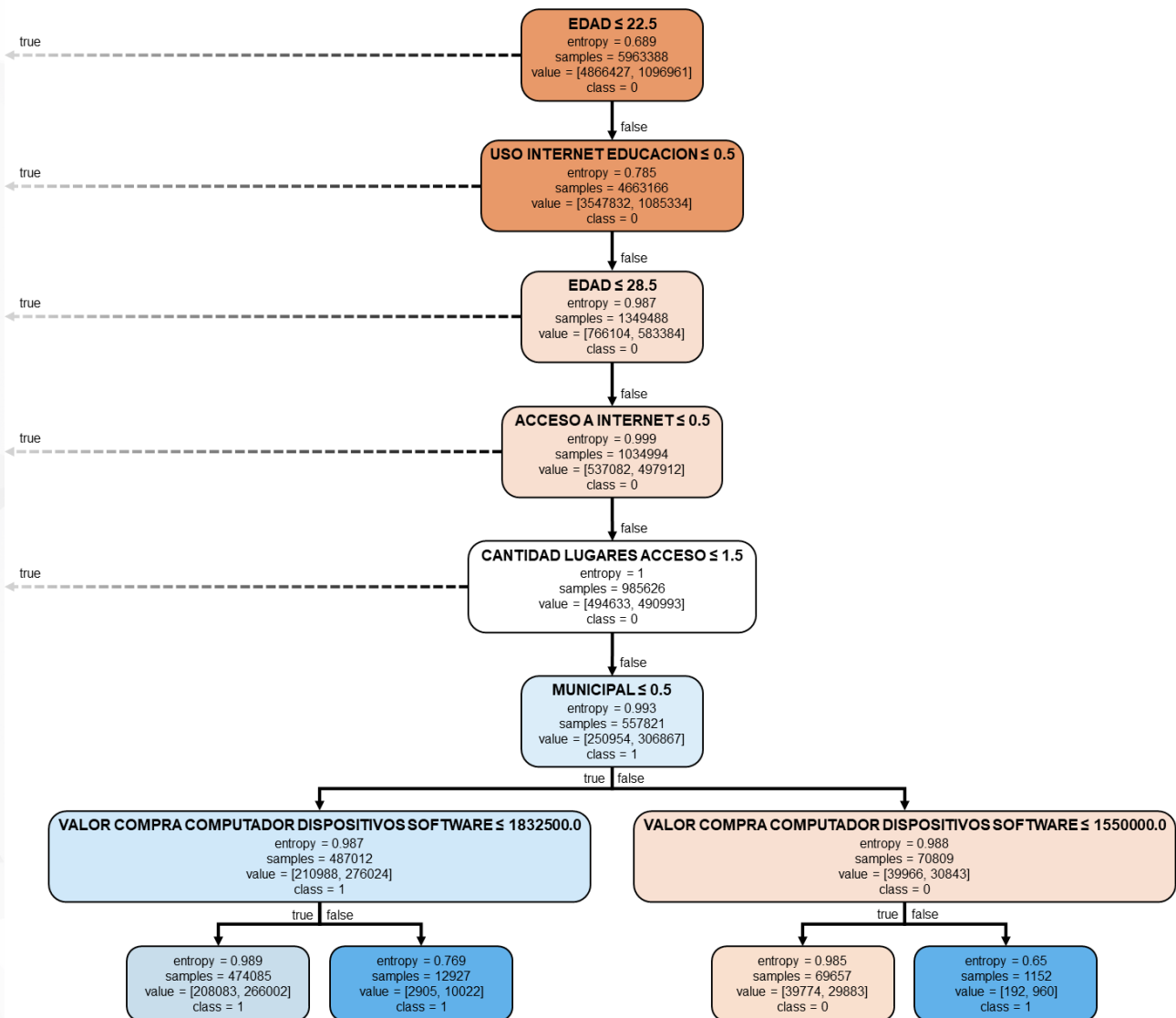
Tabla 2. Distribución poblacional por sexo de la muestra encuestada y el universo obtenido a partir del factor de expansión.

Método	Precisión	Método	Precisión
1	82,135	6	82,197
2	82,590	7	82,047
3	82,692	8	83,363
4	81,743	9	83,414
5	83,557	10	81,597
Promedio		82,534%	

Las validaciones de la precisión del modelo generado presentadas anteriormente proporcionan una idea de la confiabilidad y calidad de los resultados generados a partir de este modelo, al observar que su precisión es alta y, además, supera al promedio de precisiones posibles generadas a partir de otras muestras.

Debido a la amplitud del árbol de decisión generado, el recorrido realizado por este y los nodos de decisión que lo componen pueden ser observados dentro del Anexo 1, al final de este artículo. En la Figura 14, se muestra una versión acotada del árbol, representando los nodos principales, con los cuales se pudieron generar las conclusiones centrales de la investigación.

Figura 14. Árbol de decisión generado. Parte 1.



5. DISCUSIÓN Y CONCLUSIONES

De manera propositiva, y a raíz de los hallazgos realizados en esta investigación, se plantea que la Secretaría Distrital de Planeación, con apoyo del gobierno nacional, podría ampliar la zona de operación de la EM2021 a la extensión completa del territorio colombiano a nivel de municipios. Lo anterior debido a que se considera importante estudiar las dinámicas que se pueden observar dentro del conjunto de datos recolectados no solamente en grandes ciudades o en municipios cercanos a estas sino también en municipios más alejados. Esto podría generar insumos para el diseño y la formulación de políticas públicas educativas en torno al acceso a herramientas digitales, la calidad de la educación y, en general, la calidad de vida de las personas, tanto en áreas metropolitanas como en áreas rurales o apartadas, ya que, como lo sugieren Hernández y Esparza (2022), comprender los diferentes contextos sociales en los que viven las personas en territorios más alejados puede posibilitar una mejor atención a las necesidades de educación y en otros aspectos para las personas de estos territorios.

Similar a lo expuesto anteriormente, se encontró que poblaciones minoritarias, como la compuesta por personas de los sectores LGBTIQ+ o por personas pertenecientes a una etnia no representaban un porcentaje importante de la población encuestada (0,08% y 1,51%, respectivamente). Es de vital importancia tener información concluyente sobre estas poblaciones para realizar investigaciones que sean confiables y den mejor respuesta a las necesidades puntuales que afrontan, tal y como lo resaltaron Ghorbanian (2022) y Robbin (1999) en sus investigaciones.

Por otro lado, si bien la EM2021 suministró un conjunto de variables relacionadas con el nivel educativo, se consideró que las redes sociales incluidas en las preguntas no representaban de manera fiel el uso de estas redes en todos los rangos de edad. Redes sociales como TikTok o Instagram no fueron incluidas dentro de la encuesta, siendo estas dos de las redes sociales más utilizadas alrededor del mundo (Kemp, 2023). Por lo anterior, se considera importante incluir dentro de las preguntas relacionadas con redes sociales, a las más relevantes de estas, enmarcadas en el periodo en el que se esté realizando el levantamiento de los datos para la encuesta. Lo anterior permitiría realizar trabajos de investigación de manera más específica sobre cómo las personas, de distintos rangos de edad y diferentes trasfondos demográficos, utilizan estas redes para distintas actividades como la obtención de información, adquisición de conocimientos o para el entretenimiento, ya que algunas redes sociales, como Facebook o Twitter (ahora X) pueden ser utilizadas por comunidades poco diversas respecto a la demografía de sus usuarios/as, como sugirieron Mellon y Prosser (2017).

El alcanzar el objetivo de segmentar y caracterizar la población de Bogotá y sus municipios circundantes respecto al nivel académico y otras variables demográficas, considerando el uso de herramientas digitales y la conectividad, permitió concluir que, las personas, al alcanzar edades mayores, suelen dejar de utilizar el internet o simplemente no tienen la costumbre de hacerlo. Además, la compra de computadores, tablets y software puede sugerir un crecimiento en la calidad de vida de las personas, aunque no necesariamente en sus niveles académicos. También se concluyó que, al relacionar diversas variables demográficas y del uso de tecnologías de la información, las personas que tienen un nivel educativo alto tendían a ser mayores de 29 años, tenían 2 o más lugares donde se podían conectar a internet, incluyendo sus hogares,

y también, solían invertir una cantidad considerable de dinero en dispositivos, computadores y software. En adición, las personas menores de 25 años, que no invertían mucho dinero en computadores o software, y que solamente tenían un lugar para acceder a internet, estaban asociadas a niveles bajos de educación.

Respecto al ámbito municipal, llamó poderosamente la atención la poca incidencia de la variable Municipio, que indicaba si una persona es oriunda de un municipio o no. Lo anterior permitió concluir que, no hay diferencia de oportunidades en el ámbito educativo relacionado con el uso de las TIC entre Bogotá y los municipios circundantes, cuyos datos se recopilaban en la EM2021. Es posible que esto se relacione con la cercanía de estos municipios a la ciudad capital del país y a la facilidad de acceso a recursos provenientes de esta (algunos de los municipios participantes son conurbados o son dormitorio). Lo anterior plantea una oportunidad de investigación más amplia alrededor de las dinámicas de acceso al internet y su impacto en el nivel educativo en poblaciones de municipios alejados de ciudades principales, ya que, como lo plantean Galpin *et al.* (2022), las oportunidades de acceso a las TIC en un municipio más lejano a una ciudad principal pueden variar, debido a factores relacionados con el acceso a servicios como la electricidad o la disponibilidad de equipos de cómputo e internet en estos municipios.

Otra oportunidad que se identificó como trabajo futuro es la de analizar el comportamiento de las personas respecto al uso de herramientas digitales y del internet con el pasar de los años. Lo anterior con el fin de establecer si el declive en el uso de estas herramientas es debido a factores de edad o es más bien debido a factores culturales y de oportunidad de acceso a estas, teniendo en cuenta que para el año 2021, las personas en edades mayores a los 60 años, en su juventud, no tuvieron la misma facilidad de acceso a estas herramientas que las personas jóvenes en la actualidad. En los estudios realizados por Perrin (2015) y Zickuhr & Madden (2012) se demostró que el uso de herramientas digitales como las redes sociales y el internet, por parte de personas pertenecientes a este grupo de edad ha crecido, lo que sugiere que en el futuro esta parte de la población puede tomar un papel más significativo en el uso de estas herramientas.

Se identificó como oportunidad el análisis de datos relacionados al uso de los dispositivos digitales, específicamente en la navegación en redes sociales, ya sea para entretenimiento, obtención de información y acceso a la educación un enfocado en diferentes grupos etarios. Además, la segmentación y la caracterización de la población de Bogotá y sus municipios circundantes en cuanto a nivel académico y otras variables demográficas, considerando el uso de herramientas digitales y la conectividad, mostró que los datos recolectados en la EM2021 son una fuente valiosa para estudiar diversas dinámicas. Esto posibilita el estudio de estas dinámicas, con la oportunidad de abordar estas desde un tema o campo de interés establecido. Tal fue el caso de esta investigación, en el que se pudieron utilizar variables de diferentes tipos, ya sean de demografía, de educación y fuerza de trabajo con una orientación al uso de las herramientas digitales y el internet, con el fin de estudiar la dinámica o relación entre el nivel educativo de las personas, el uso de estas herramientas digitales y otros factores demográficos.

Como oportunidad de trabajo futuro, se identificó la implementación de otros modelos estadísticos y técnicas de aprendizaje automático como redes neuronales, regresión logística, y clustering. Esto no solo aumentaría la precisión de las conclusiones, sino que también permitiría una

comprensión más holística de las dinámicas estudiadas. Además, incorporar análisis longitudinal, considerando las diferentes aplicaciones de la Encuesta Multipropósito (cada dos años), podría proporcionar una visión temporal de las tendencias y cambios en el uso de herramientas digitales y niveles académicos, aportando información valiosa para la formulación de políticas públicas y estrategias educativas más efectivas.

6. REFERENCIAS

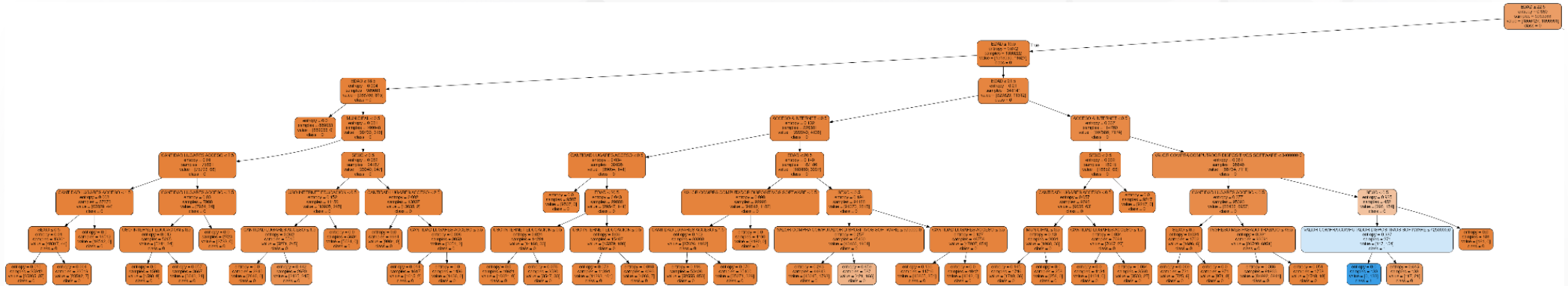
- Andrade, M., Valdivieso, D., & Zambrano, C. (2023). Redes sociales y el comportamiento de los adolescentes en la comunicación digital. *Polo del Conocimiento*, 8(10), 357-371 <http://doi.org/10.23857/pc.v8i10.6129>
- Barragán, S., & González, L. (2021). Factores que inciden en la deserción estudiantil de las personas de los sectores LGBTI en Bogotá. *Sophia-Educación*, 7(1), 1-16. <http://dx.doi.org/10.18634/sophiaj.17v.1i.994>
- Barragán, S., & González, L. (2023). El nivel educativo en su papel mediador en la discriminación observada por orientación sexual. *Revista Colombiana de Sociología*, 6(1), 221-242. <https://doi.org/10.15446/rcs.v46n1/84549>
- Cicik Wijayanti, T., Naim, S., Hendayani, N., & Hanum, F. (2024). Identify the use of economics for family financial management in digital days. *Indonesian Interdisciplinary Journal of Sharia Economics (IIJSE)*, 7(1). <https://doi.org/10.31538/iijs.v7i1.3754>
- De Zúñiga, H. G., Mateos, A., & Inguanzo, I. (2022). Rethinking social capital in the digital era and diverse societies. *Revista Internacional de Sociología*, 80(4). <https://doi.org/10.3989/ris.2022.80.4.MI22-0001>
- Departamento Administrativo Nacional de Estadística-DANE. (2018). *Marco Geoestadístico Nacional 2018*. <https://bit.ly/3Uzy4EA>
- Departamento Administrativo Nacional de Estadística-DANE. (2022). *Encuesta multipropósito Bogotá Cundinamarca EM 2021*. <https://bit.ly/3wgoZbH>
- Dirección de Metodología y Producción Estadística - DIMPE. (2022). *Ficha metodológica encuesta multipropósito Bogotá Cundinamarca 2021*. <https://bit.ly/44ujyT7>
- Esteve-Mon, F. M., Postigo-Fuentes, A. Y., & Castañeda, L. (2023). A strategic approach of the crucial elements for the implementation of digital tools and processes in higher education. *Higher Education Quarterly*, 77(3), 558-573. <https://doi.org/10.1111/hequ.12411>
- Galpin, I., Morales Dussan, S. D., Leon, M., & Garcia-Bedoya, O. (2022). Exploring the Colombian digital divide using Moodle logs through supervised learning. *Interactive Technology and Smart Education*, 19(3), 281-299. <https://doi.org/10.1108/ITSE-03-2021-0052>
- García-Martínez, J.-A., Rosa-Napal, F.-C., Romero-Tabayo, I., López-Calvo, S., & Fuentes-Abeledo, E.-J. (2020). Digital Tools and Personal Learning Environments: An Analysis in Higher Education. *Sustainability*, 12(19), 8180. <https://doi.org/10.3390/su12198180>
- Gatto, S. L., & Tak, S. H. (2008). Computer, Internet, and E-mail Use Among Older Adults: Benefits and Barriers. *Educational Gerontology*, 34(9), 800-811. <https://doi.org/10.1080/03601270802243697>
- Ghorbanian, A., Aiello, B., & Staples, J. (2022). Under-Representation of Transgender Identities in Research: The Limitations of Traditional Quantitative Survey Data. *Transgender Health*, 7(3), 261-269. <https://doi.org/10.1089/trgh.2020.0107>

- Greener, S., & Wakefield, C. (2015). Developing confidence in the use of digital tools in teaching. *The Electronic Journal of e-Learning*, 13(4), 260-267 <https://bit.ly/3UoW78U>
- Guevara, E., Lucero, E., Noroña, F., & Quishpe, M. (2024). Herramientas Web 2.0: una revisión bibliográfica sobre su aplicación en la metodología de la clase inversa. *MENTOR Revista de investigación educativa y deportiva*, 3(7), 187-201. <https://doi.org/10.56200/mried.v3i7.6734>
- Hernández Herrera, M. T., & Esparza Urzúa, G. A. (2022). La calidad de la educación en territorios rurales desde las políticas públicas. *Sophia*, 32, 171-193. <https://doi.org/10.17163/soph.n32.2022.05>
- Hernández, J., & Rayón, L. (2021). Teléfonos móviles, redes sociales y praxis en adolescentes. *Educatio Siglo XXI*, 39(3), 135-156. <https://doi.org/10.6018/educatio.427011>
- Hrynevych, L., Morze, N., Vember, V., & Boiko, M. (2021). Use of digital tools as a component of STEM education ecosystem. *Educational Technology Quarterly*, 2021(1), 118-139. <https://doi.org/10.55056/etq.24>
- Kemp, S. (2023). *We Are Social, DataReportal, & Meltwater. Biggest social media platforms 2023. Most Popular Social Networks Worldwide as of October 2023, Ranked by Number of Monthly Active Users*. <https://bit.ly/3wpXvQM>
- Kingsford, C., & Salzberg, S. L. (2008). What are decision trees? *Nature Biotechnology*, 26(9), 1011-1013. <https://doi.org/10.1038/nbt0908-1011>
- Levano, L., Sanchez, S., Guillén, P., Tello, S., Herrera, N., & Collantes, Z. (2019). Competencias digitales y educación. *Propósitos y Representaciones*, 7(2). <https://doi.org/10.20511/pyr2019.v7n2.329>
- Mellon, J., & Prosser, C. (2017). Twitter and Facebook are not representative of the general population: Political attitudes and demographics of British social media users. *Research & Politics*, 4(3), 205316801772000. <https://doi.org/10.1177/2053168017720008>
- Mero-Ponce, J. I. (2021). Herramientas digitales educativas y el aprendizaje significativo en los estudiantes. *Dominio de las ciencias*, 7(1), 712-724. <https://doi.org/10.23857/dc.v7i1.1735>
- Miguel, D. (2020). Academic achievement and technology: Evolution of the debate in the last decades. *Cadernos de Pesquisa*, 50(178), 1122-1137. <https://doi.org/10.1590/198053147144>
- Mohammed, Q. A., Rao Naidu, V., Hasan, R., Mustafa, M., & Jesrani, K. A. (2019). Digital education using free and open-source tools to enhance collaborative learning. *IJAEDU-International E-Journal of Advances in Education*, 5(13), 50-57. <https://doi.org/10.18768/ijaedu.531636>
- Mucundanyi, G., & Woodley, X. (2021). Exploring Free Digital Tools in Education. *International Journal of Education and Development using Information and Communication Technology (IJE-DICT)*, 17(2), 96-103. <https://bit.ly/3ycShZh>
- Nkansah, J. O., & Oldac, Y. I. (2024). *Unraveling the attributions of digital literacy skills and knowledge gap in Ghana's higher education: Undergraduate students voices in a phenomenological study*. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12483-8>

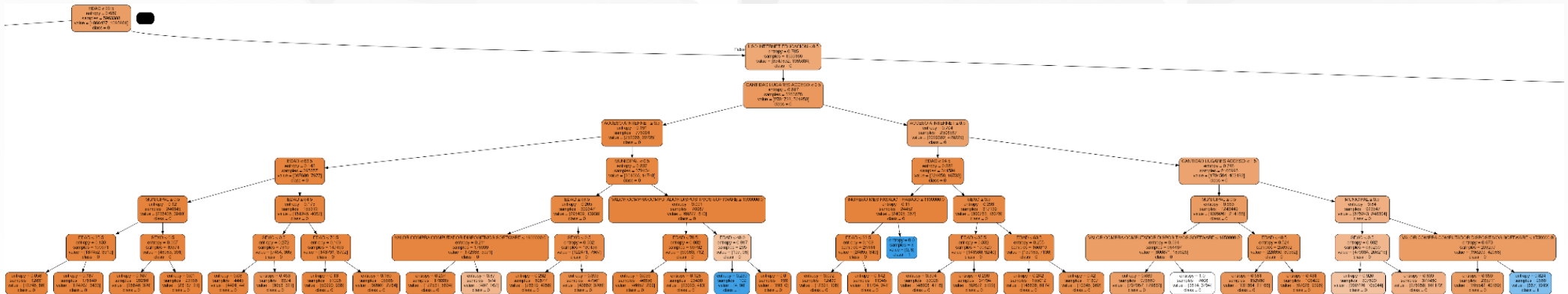
- Perrin, A. (2015). *Social Media Usage: 2005-2015*. Pew Research Center. <https://bit.ly/4bt43wV>
- Pillajo, G., Torres, L., Vera, J., & Grunauer, G. (2023). Una revisión sistemática de las estrategias de aprendizaje enfocadas en las personas adultas con escolaridad inconclusa del tercero BGU. *Polo del conocimiento*, 9(1), 24-48. <https://doi.org/10.23857/pc.v9i1.6366>.
- Robbin, A. (1999). The problematic status of U.S. statistics on race and ethnicity. *Journal of Government Information*, 26(5), 467-483. [https://doi.org/10.1016/S1352-0237\(99\)00078-7](https://doi.org/10.1016/S1352-0237(99)00078-7)
- Shwedeh, F. (2024). Harnessing digital issue in adopting metaverse technology in higher education institutions: Evidence from the United Arab Emirates. *International Journal of Data and Network Science*, 8(1), 489-504. <https://doi.org/10.5267/j.ijdns.2023.9.007>
- Simões, F., & Marta, E. (2024). Public Employment Services and Vulnerable Youth in the EU: The Case of Rural NEETs. *Politics and Governance*, 12. <https://doi.org/10.17645/pag.7432>
- Ventrella, F. M., & Cotnam-Kappel, M. (2024). Examining digital capital and digital inequalities in Canadian elementary schools: Insights from teachers. *Telematics and Informatics*, 86, 102070. <https://doi.org/10.1016/j.tele.2023.102070>
- Wirth, R., & Hipp, J. (2000). *CRISP-DM: Towards a Standard Process Model for Data Mining*. 1-11. <https://bit.ly/4acfOGU>
- Zickuhr, K., & Madden, M. (2012). *Older adults and internet use*. Pew Research Center. <https://bit.ly/3yafs6q>

7. ANEXOS

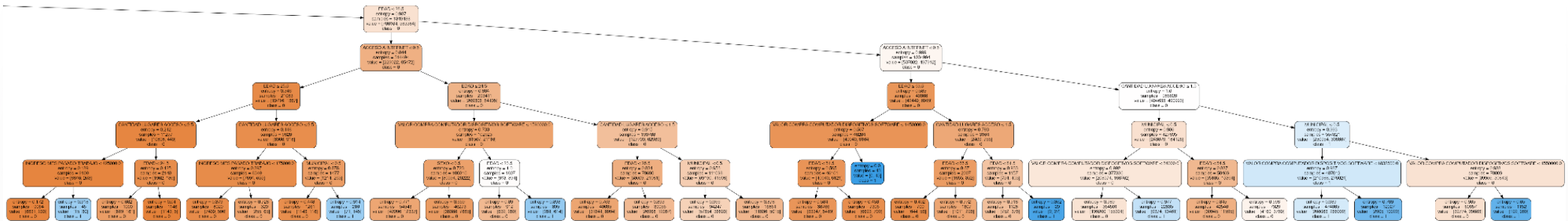
ÁRBOL DE DECISIÓN PARTE 1:



ÁRBOL DE DECISIÓN PARTE 2:



ÁRBOL DE DECISIÓN PARTE 3:



ÁRBOL DE DECISIÓN COMPLETO:

