

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

ÍNDICE PARA LA MEDICIÓN DE LA POBREZA DIGITAL DE LOS HOGARES EN
COSTA RICA

Trabajo final de investigación aplicada sometido a la consideración de la Comisión del
Programa de Posgrado en Estadística para optar al grado y título de Maestría Profesional en
Estadística

MARICRUZ LEÓN MIRANDA

Ciudad Universitaria Rodrigo Facio, Costa Rica

2025

DEDICATORIA

Dedico este logro a Dios y la Virgen María porque han sido mi soporte espiritual diario.

A mi mamá, por impulsarme a estudiar y enseñarme que perder una batalla no significa perder la guerra.

A mi papá (q.d.D.g) por inculcarme a perseverar para alcanzar.

A mi esposo Leslie y a mis hijos Diana y Adrián, por su amor, apoyo, compañía y motivación.

AGRADECIMIENTO

Al Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones por el interés en el desarrollo de esta investigación, en especial a la Carla Valverde Barahona y Elidier Moya Rodríguez, las jefaturas de la Dirección de Evolución y Mercado de Telecomunicaciones.

A mi tutor M.Sc. Rodrigo Corrales Mejías, por su orientación, apoyo y valiosas sugerencias a lo largo de todo el proceso de esta investigación.

A mi lector Dr. Gilbert Brenes Camacho, por su honestidad, empatía y recomendaciones personales y metodológicas.

A mi lectora M.Sc. Paola Omodeo Cubero, por su ayuda desinteresada y por los años de amistad incondicional.

Este trabajo final de investigación aplicada fue aceptado por la Comisión del Programa de Estudios de Posgrado en Estadística de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Profesional en Estadística.

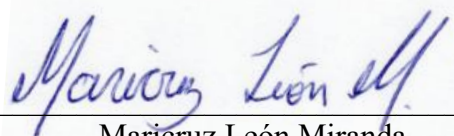
Dra. Jackeline García Fallas
**Representante de la Decanatura
Sistema de Estudios de Posgrado**

M.Sc. Rodrigo Corrales Mejías
Profesor Guía

Dr. Gilbert Brenes Camacho
Lector

M.Sc. Paola Omodeo Cubero
Lectora

M.Sc. Ericka Méndez Chacón
**Representante del Director del
Programa de Posgrado en Estadística**



Maricruz León Miranda
Sustentante

TABLA DE CONTENIDO

DEDICATORIA	ii
AGRADECIMIENTO	iii
HOJA DE APROBACIÓN	iv
TABLA DE CONTENIDO	v
RESUMEN EN ESPAÑOL	viii
ABSTRACT.....	ix
LISTA DE CUADROS.....	x
LISTA DE TABLAS	xii
LISTA DE ILUSTRACIONES	xiii
LISTA DE ABREVIATURAS	xiv
LISTA DE GRÁFICOS	xv
CAPÍTULO I. INTRODUCCIÓN	1
CAPÍTULO II. OBJETIVOS.....	2
2.1 Objetivo General.....	2
2.2 Objetivos Específicos	2
CAPÍTULO III. ORIENTACIÓN CONCEPTUAL	3
3.1 Tecnologías digitales	3
3.2 Tecnologías de la Información y Comunicación	4
CAPÍTULO IV. METODOLOGÍA	6
4.1 Fuente datos	6
4.1.1 Encuesta Nacional de Ingresos y Gastos 2018-2019	6
4.1.2 Encuesta Nacional de Hogares 2018	7
4.1.3 Calidad de los datos	8

4.2	Variables de estudio.....	9
4.2.1	Acceso a las TIC	10
4.2.2	Uso de las TIC	11
4.2.3	Gasto en TIC	12
4.3	Técnicas estadísticas: descripción de las técnicas usadas.....	18
4.3.1	Análisis Factorial Exploratorio	18
4.3.2	Modelo de Regresión Múltiple para la variable Gasto TIC.....	20
4.3.3	<i>Cluster Analysis</i>	23
4.3.4	Softwares utilizados.....	24
4.4	Alcances y Limitaciones.....	25
CAPÍTULO V. RESULTADOS.....		27
5.1	Análisis descriptivo de las variables de estudio.....	27
5.1.1	Variables sobre el Acceso a las TIC	27
5.1.2	Variables sobre el Uso a las TIC	28
5.1.3	Variable Gasto en TIC	31
5.2	Estimación de la variable del gasto TIC por medio del modelo de regresión múltiple.....	32
5.2.1	Relaciones bivariadas del modelo de regresión para el Gasto TIC	33
5.2.2	Ecuación de regresión para estimar el Gasto TIC.....	35
5.2.2.1	Ecuación del Gasto TIC en hogares con jefatura ocupada	35
5.2.2.2	Ecuación del Gasto TIC en hogares con jefatura no ocupada	37
5.2.2.3	Validación de las ecuaciones de regresión en segunda submuestra	38
5.3	Estimación de la variable del gasto TIC por medio del modelo de regresión múltiple en la ENAHO.....	38
5.4	Construcción de las dimensiones.....	41
5.4.1	Análisis de Confiabilidad.....	42

5.4.2	Construcción de Dimensiones mediante la aplicación del Análisis de Factorial Exploratorio	44
5.4.2.1	Componente TIC en la ENIGH	44
5.4.2.2	Componente TIC en la ENAHO	46
5.5	Construcción del IPD según la ENIGH 2018.	49
5.6	Construcción del IPD según la ENAHO 2018.....	53
5.7	Comparación de comportamiento del IPD según encuesta.	57
5.7.1	Componentes <i>TIC</i> y Gasto en TIC	58
5.7.1.1	Componente TIC.....	58
5.7.1.2	Componente Gasto TIC	60
5.7.2	IPD _{Enigh_z} versus IPD _{Enaho_z}	62
5.7.3	Niveles de <i>IPD_{Enigh}</i> versus Niveles del <i>IPD_{Enaho}</i>	63
5.8	Caracterización sociodemográfica de los hogares de Costa Rica según pobreza digital.	64
5.8.1	Caracterización de los hogares según <i>IPD_{Enigh}</i>	64
5.8.2	Caracterización de los hogares según <i>IPD_{Enaho}</i>	67
CAPÍTULO VI. CONCLUSIONES		71
CAPÍTULO VII. BIBLIOGRAFÍA		74
ANEXOS		78

RESUMEN EN ESPAÑOL

El apropiado acceso y uso de las Tecnologías de Información y Comunicación se ha relacionado como un determinante del progreso de un país. La medición de la evolución del sector de telecomunicaciones orienta la generación de políticas públicas enfocadas hacia al mejoramiento del acceso y uso de las herramientas tecnológicas, y, por ende, mejorar las condiciones en los hogares y la población de Costa Rica. En el presente estudio se categorizan los hogares del país según pobreza digital por medio de IPC construido en la Encuesta Nacional de Ingresos y Gastos 2018 y en la Encuesta Nacional de Hogares 2018. El índice integra dos componentes: TIC y Gasto TIC. El componente TIC es construido mediante la aplicación del AFE y constituye una aproximación de nivel de acceso y uso de los servicios y dispositivos en los hogares. El componente Gasto TIC representa el desembolso monetario mensual de los hogares, el cual fue calculado directamente en la ENIGH, mientras que en la ENAHO su estimación se obtiene mediante la aplicación de un modelo de regresión múltiple. Estos componentes fueron combinados para construir el IPD en cada encuesta, y se estandariza en escala de 0 a 100, donde mayores valores representan mayor nivel de pobreza digital. La media nacional del IPD_{Enigh} se estima en 67,3 y de 65,9 con el IPD_{Enaho} . Se realiza la segmentación de hogares mediante el análisis de conglomerados (k-medias) en cada encuesta, obteniéndose dos grupos en la ENIGH (pobre digital y no pobre digital) y tres grupos en la ENAHO (pobre digital alto, pobre digital medio y no pobre digital). Los resultados demuestran que el nivel de pobreza digital está asociada a las variables la zona y región de residencia, nivel educativo, edad y condición de actividad del jefe del hogar. Se sugiere al MICITT considerar la posibilidad de incorporar el IPD_{Enaho} como un indicador de referencia para el monitoreo de la pobreza digital, alineándolo con las herramientas de planificación institucional.

ABSTRACT

The adequate access to, and use of, Information and Communications Technologies is firmly and directly related to a country's development. Analyzing the evolution of the telecommunications industry has helped create public policies that focus on promoting ICT access and usage and, therefore, has also contributed to the improvement of the overall conditions of households nationwide and the Costa Rican population. This study will categorize households across the country, according to their level of digital poverty, by using the Digital Poverty Index [DPI] obtained from the 2018 National Household Income and Expenditures Survey ["ENIGH" as per its acronym in Spanish] and the 2018 National Household Survey ["ENAHO" as per its acronym in Spanish]. This DPI is comprised of two key components, namely: ICT and ICT Expenditure. The former of these two components, ICT, measures the overall application of the Exploratory Factor Analysis (EFA) and estimates the level of access to, and the usage of, ICT services and devices in Costa Rican households. While the latter, ICT Expenditure, represents the overall monthly spending of Costa Rican households in Information and Communications Technologies. ICT Expenditure is directly calculated in ENIGH and is estimated from the data obtained from the ENAHO through the use of a multiple regression model. These two components were then consolidated and used to construct the DPI of each survey, on a scale of 0 to 100, where the higher the value the higher the level of digital poverty. The national average for the DPI_{Enigh} was estimated to be 67.3, while the national average for the DPI_{Enaho} was estimated to be 65.9. Households were then partitioned via K-means clustering for each survey, yielding two segments in the ENIGH ("digitally poor" and "not digitally poor") and three segments in the ENAHO ("high digital poverty", "intermediate digital poverty" and "no digital poverty"). These results show that the level of digital poverty is directly related to a number of variables in the area/region of residence, such as level of education, age, and the head of household's employment status. It is therefore suggested that MICITT consider the possibility of incorporating the DPI_{Enaho} as a benchmark indicator to monitor the level of digital poverty, while also making sure that the indicator is aligned with the institution's planning tools.

LISTA DE CUADROS

Cuadro 1: Porcentaje de hogares sin acceso a dispositivos y servicios TIC	28
Cuadro 2: Porcentaje de hogares sin uso de los dispositivos y servicios TIC.....	29
Cuadro 3: Porcentaje de hogares según nivel de hacinamiento tecnológico.....	30
Cuadro 4: Estadísticas Descriptivas del Gasto mensual en TIC de los hogares de la ENIGH	31
Cuadro 5: Estadísticos descriptivos de las variables para el modelo de regresión para la estimación del gasto TIC de la submuestra 1.	33
Cuadro 6: Coeficientes de correlación de las variables para el modelo de regresión para la estimación del gasto TIC de la submuestra 1.	34
Cuadro 7: Resultados del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición ocupada (Submuestra 1).....	36
Cuadro 8: Resultados del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición no ocupada (Submuestra 1).	37
Cuadro 9: Estadísticas Descriptivas de la estimación Gasto TIC en los hogares de la ENAHO	41
Cuadro 10: Cambios en el Alfa de Cronbach al suprimir cada variable de la ENIGH.	42
Cuadro 11: Cambios en el Alfa de Cronbach al suprimir cada variable de la ENAHO.	42
Cuadro 12: Resultados del AFE aplicado a la ENIGH	44
Cuadro 13: Cargas Factoriales y análisis de fiabilidad para las variables de la ENIGH 2018.	45
Cuadro 14: Estadísticas Descriptivas del componente TIC de la ENIGH.	46
Cuadro 15: Resultados del AFE aplicado a la ENAHO.....	47
Cuadro 16: Cargas Factoriales y análisis de fiabilidad para las variables de la ENAHO 2018.	48
Cuadro 17: Estadísticas Descriptivas del componente TIC según encuesta.	48
Cuadro 18: Estadísticas Descriptivas del Índice de Pobreza Digital en los hogares de la ENIGH 2018.....	50
Cuadro 19: Estadísticas Descriptivas del <i>IPDENigh_z</i> y gasto mensual en TIC de los hogares de la ENIGH según nivel de pobreza digital.....	52

Cuadro 20: Estadísticas Descriptivas del <i>IPDEnigh_z</i> , <i>TICEnigh_z</i> y <i>Gasto TICEnigh_z</i> de los hogares de la ENIGH 2018 según nivel.	52
Cuadro 21: Estadísticas Descriptivas del Índice de Pobreza Digital en los hogares de la ENAHO 2018.	54
Cuadro 22: Estadísticas Descriptivas del <i>IPDEnaho</i> y gasto mensual estimado en TIC de los hogares de la ENAHO según nivel de pobreza digital.....	56
Cuadro 23: Estadísticas Descriptivas del <i>IPDEnaho_z</i> , <i>TICEnaho_z</i> y <i>Gasto TICEnaho_z</i> de los hogares de la ENAHO 2018 según nivel.	57
Cuadro 24: Prueba de igualdad de medias para el componente TIC en muestras independientes.	59
Cuadro 25: Prueba de igualdad de medias para el componente Gasto TIC en muestras independientes.	61
Cuadro 26: Comparación de los valores del <i>IPDEnigh</i> y <i>IPDEnaho</i> nacional, por zona y por región.	62
Cuadro 27: Comparación de los conglomerados según IPD y sus componentes según encuesta.....	63
Cuadro 28: Distribución porcentual de los hogares por conglomerados según variables sociodemográficas.	66
Cuadro 29: Estadísticas Descriptivas del gasto medio en TIC de los hogares de la ENIGH según conglomerado	67
Cuadro 30: Distribución porcentual de los hogares por conglomerados según variables sociodemográficas de la ENAHO.....	69
Cuadro 31: Estadísticas Descriptivas del gasto medio en TIC de los hogares de la ENAHO según conglomerado	70

LISTA DE TABLAS

Tabla 1: Clasificación de equipos y servicios de TIC, basada en la CCIF 2018	13
Tabla 2: Escalamiento de la variable nivel educativo del jefe del hogar de la ENIGH.....	21
Tabla 3: Escalamiento de la variable nivel educativo del jefe del hogar de la ENAHO.	39
Tabla 4: Variables resultantes del AFE que conforman el componente TIC según encuesta.	58

LISTA DE GRÁFICOS

Gráfico 1: Distribución porcentual de los hogares según nivel de pobreza digital para la ENIGH 2018	51
Gráfico 2: Distribución porcentual de los hogares según nivel de pobreza digital para la ENAHO 2018	55
Gráfico 3: Distribución porcentual de los hogares según intervalos del componente TIC por encuesta.....	60
Gráfico 4: Distribución porcentual de los hogares según intervalos del componente Gasto TIC según encuesta.....	61

LISTA DE ILUSTRACIONES

Ilustración 1: Diagrama Propuesto del Índice de Pobreza Digital	10
--	----

LISTA DE ABREVIATURAS

AFC. Análisis Factorial Confirmatorio

AFE. Análisis Factorial Exploratorio

ENAH. Encuesta Nacional de Hogares

ENIGH. Encuesta Nacional de Ingresos y Gastos de los Hogares

INEC. Instituto Nacional de Estadísticas y Censos

IPD. Índice de Pobreza Digital

MICITT. Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones

TIC. Tecnologías de la Información y Comunicación

UIT. Unión Internacional de Telecomunicaciones

CAPÍTULO I. INTRODUCCIÓN

En los últimos años, el desarrollo de la tecnología digital, así como el acceso y uso apropiado de las Tecnologías de la Información y Comunicación (TIC) se ha concebido como uno de los determinantes del progreso de un país. Conocer el nivel de digitalización de éstos, permite la generación de políticas públicas enfocadas hacia al mejoramiento del acceso y uso de las herramientas tecnológicas, y, por ende, mejorar las condiciones para la población (Otero, 2018).

En Costa Rica, el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT) publica el Índice de Brecha Digital (IBD) (Ministerio de Ciencia, Tecnología y Telecomunicaciones, 2019), basado en el Índice de Desarrollo de las TIC (IDI por sus siglas en inglés) de la Unión Internacional de Telecomunicaciones (UIT), para el monitoreo de la evolución del sector de telecomunicaciones. Sin embargo, este índice se limita a un valor global de la situación del país, omitiendo las realidades de los hogares en cuanto a la limitación en el uso, acceso y gasto en las tecnologías digitales, así como la diferenciación que se presenta entre las zonas (urbano y rural) o entre las regiones del país.

Un concepto en pleno desarrollo que intenta describir estas realidades, es el de pobreza digital, el cual incluye las carencias a nivel de accesibilidad a las TIC, así como las de utilización y el consumo en estas herramientas (Alva de la Selva, 2015) (Coria, Pérez Meza, Mendoza Cortés, & Martínez Peláez, 2011) (Barrantes, 2007) (Barja & Gigler, 2006). El desarrollo y aplicación de este concepto en Costa Rica es limitado en la actualidad y disponer de alguna herramienta que permita analizar exhaustivamente las carencias de TIC en los hogares costarricenses, permitiría orientar la formulación de las políticas públicas en el ámbito de las telecomunicaciones.

Por tanto, ante la carencia de información sobre la pobreza digital en los hogares del país, en esta investigación se propone la creación de un índice para la medición de la Pobreza Digital de los Hogares de Costa Rica, que permita reflejar las condiciones de los hogares en cuanto al acceso, uso y gasto en las TIC, e identificar los principales determinantes de la pobreza digital según zona (urbana y rural) y región de planificación.

CAPÍTULO II. OBJETIVOS

2.1 Objetivo General

Caracterizar los hogares de Costa Rica mediante el Índice Propuesto de Pobreza Digital para la orientación de la toma de decisiones y políticas públicas del sector de Telecomunicaciones.

2.2 Objetivos Específicos

- i. Construir las dimensiones de Acceso, Uso y Gasto en los Servicios de Telecomunicaciones a partir de los datos de la Encuesta Nacional de Hogares 2018 y la Encuesta Nacional de Ingresos y Gastos 2018.
- ii. Estimar el Índice de Pobreza Digital de los Hogares, según zona (urbano y rural) y región de residencia, utilizando los datos de la Encuesta Nacional de Ingresos y Gastos 2018.
- iii. Estimar el Índice de Pobreza Digital de los Hogares, según zona (urbano y rural) y región de residencia, utilizando los datos de la Encuesta Nacional de Hogares y con el valor imputado del Gasto TIC.
- iv. Comparar el comportamiento del IPD aplicado en cada encuesta.
- v. Caracterizar a los hogares de Costa Rica de acuerdo con los valores del Índice de Pobreza Digital.

CAPÍTULO III. ORIENTACIÓN CONCEPTUAL

3.1 Tecnologías digitales

Las tecnologías digitales es un término asociado al concepto de Sociedad de la Información, la cual ha cobrado auge y difusión en las últimas décadas. Según Reusser, la definición de Sociedad de la Información es compleja ya que su nivel de desarrollo es aún incipiente: además de que diversas opiniones afirman que se trata de una sociedad en formación, en la que se integran las nociones de información, comunicación y nuevas tecnologías (Reusser Monsálvez, 2003).

Este mismo autor, hace referencia a las definiciones de otros investigadores; entre ellos a Masuda la cual desde una óptica humanista expresa que se trata de una sociedad “que crece y se desarrolla alrededor de la información y aporta un florecimiento general de la creatividad intelectual humana, en lugar de un aumento del consumo material” (Masuda, 1984). También hace referencia a la Comisión de la Sociedad de la Información del Ministerio de Ciencias de Portugal, la cual en 1997 brindó una definición más amplia indicando que "se refiere a una forma de desarrollo económico y social en el que la adquisición, almacenamiento, procesamiento, evaluación, transmisión, distribución y diseminación de la información con vistas a la creación de conocimiento y a la satisfacción de las necesidades de las personas, de las organizaciones y juega un papel central en la actividad económica, en la creación de riqueza y en la definición de la calidad de vida y las prácticas culturales de los ciudadanos”.

Por su parte, la UIT publica los acuerdos generados en la Cumbre Mundial Sobre la Sociedad de la Información, celebrada en Ginebra 2003 y Túnez 2005 (Unión Internacional de Telecomunicaciones, 2005), entre los cuales logra formular una definición de la sociedad de la información convenida internacionalmente, con líneas de acción específicas, que abarcó, entre otros temas los siguientes:

- La función de los gobiernos en la promoción de las Tecnologías de la Información y Comunicación para el desarrollo.
- El acceso a la información y al conocimiento.

- El fomento de la confianza y seguridad en la utilización de las Tecnologías de la Información y Comunicación.
- Aplicaciones de las Tecnologías de la Información y Comunicación y sus beneficios en todos los aspectos de la vida.

Más recientemente, esta misma organización publicó en el Informe de Medición de la Sociedad de la Información (Unión Internacional de Telecomunicaciones, 2018) la evaluación del estado de la sociedad de la información en todo el mundo. Esta evaluación examina el acceso a las TIC, la infraestructura de comunicaciones, el uso de Internet y otros indicadores clave en diferentes países. Estos indicadores se utilizan para medir el progreso hacia una sociedad de la información más inclusiva y equitativa.

Dentro de los hallazgos reportados, se menciona el crecimiento de la conectividad, el acceso y uso de dispositivos móviles lo que permite un mayor acceso a los servicios en línea, así como el crecimiento del uso de Internet para actividades como el comercio electrónico, las redes sociales y el acceso a servicios gubernamentales en línea.

El informe también proporciona recomendaciones para abordar los desafíos y promover un mayor desarrollo de la sociedad de la información, como la necesidad de políticas y regulaciones efectivas, inversiones en infraestructura y capacidad, así como iniciativas para fomentar la alfabetización y la inclusión digitales (Unión Internacional de Telecomunicaciones, 2018).

3.2 Tecnologías de la Información y Comunicación

Las TIC se definen como un conjunto de herramientas, dispositivos, infraestructuras y sistemas que se utilizan para procesar, almacenar, transmitir y gestionar la información de manera digital (Belloch, 2012). Es decir, las TIC abarcan una amplia gama de tecnologías y aplicaciones que permiten la comunicación, el acceso a la información y la automatización de procesos.

Según la UIT, organismo especializado de las Naciones Unidas para las tecnologías de la información y la comunicación, éstas incluyen la penetración de Internet, el acceso a servicios de banda ancha y la disponibilidad de teléfonos móviles. Además, el uso de estas tecnologías en el acceso a servicios en línea, en el comercio electrónico y en la participación en redes sociales (Unión Internacional de Telecomunicaciones, 2018).

Por su parte, el Banco Mundial afirma que, si bien las TIC han traído consigo beneficios y oportunidades, también existe la posibilidad de una "brecha digital" que puede ampliar las desigualdades existentes entre los países y dentro de ellos. Esta misma entidad enfatiza la importancia de promover una "transformación digital inclusiva", que garantice que todos los sectores de la sociedad puedan beneficiarse de las TIC. Lo que implica fomentar la adopción de tecnología en empresas y promover la alfabetización digital de la población y la educación en habilidades digitales (Banco Mundial, 2016).

CAPÍTULO IV. METODOLOGÍA

En este apartado, se presenta una descripción de los datos y variables que se utilizarán para el análisis, además de las técnicas estadísticas empleadas y los paquetes estadísticos empleados para el procesamiento.

4.1 Fuente datos

Para el desarrollo de esta investigación se requiere el uso de dos encuestas nacionales realizadas por el Instituto Nacional de Estadísticas y Censos (INEC): Encuesta Nacional de Ingresos y Gastos de los Hogares 2018-2019 (ENIGH) y Encuesta Nacional de Hogares (ENAHO).

Estas fuentes de datos comparten similitudes y diferencias en el diseño metodológico y en los temas de investigación. En cuanto a las similitudes a nivel metodológico, la población de estudio, el marco muestral de viviendas, los dominios de estudio, la estratificación del marco muestral y el diseño de muestreo, son equivalentes. En cuanto a los temas de investigación, tanto de la ENAHO como de la ENIGH es posible obtener datos sobre las características sociodemográficas y socioeconómicas de los hogares y sus miembros, así como características de las viviendas y los servicios que dispone. Para una mejor comprensión, se describen brevemente las estas fuentes de datos.

4.1.1 Encuesta Nacional de Ingresos y Gastos 2018-2019

La primera fuente de datos corresponde a la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018-2019 cuyo objetivo principal es proporcionar información oportuna y detallada sobre la procedencia y estructura de los ingresos y los patrones de gasto de los hogares, así como otros aspectos relacionados a la medición del bienestar. Para la edición más reciente, la ENIGH requirió la colaboración de 9 828 hogares seleccionados en la muestra y distribuidos en todo el territorio nacional; el trabajo de campo se desarrolló de

febrero 2018 a marzo 2019, para lograr la captura de los rubros de ingreso y de gasto que son estacionales (Instituto Nacional de Estadísticas y Censos, 2019).

Los datos disponibles de esta encuesta se organizan en las siguientes bases:

- *ENIGH Base Hogares*: comprende la información de los hogares de las viviendas seleccionadas: características sociodemográficas y socioeconómicas, el ingreso según fuente y el tipo de gasto del hogar.
- *ENIGH 2018 Base Gastos*: contiene la información detallada de en qué gastan los hogares y comprende información sobre el tipo de artículo o servicio adquirido por el hogar y un detalle del gasto en colones mensuales, entre otros.
- *ENIGH Base Personas*: esta base contiene la información de las características sociodemográficas de los miembros del hogar, transferencias sociales, condición de actividad, características del trabajo e ingresos personales.

Para esta investigación se utilizará la base ENIGH 2018-2019 Base Gastos, para estimar el Gasto TIC en los hogares de Costa Rica, mediante la suma simple de los gastos registrados en dispositivos y servicios TIC dentro de cada hogar, según la metodología del indicador Gasto TIC de la UIT. Adicionalmente, se construirá una base personalizada con las variables referidas al hogar (extraídas de la ENIGH Base Hogares) y variables del jefe del hogar (extraídas de la ENIGH Base Personas), que en conjunto con la variable del Gasto TIC serán la base de datos para el ajuste del modelo de regresión para estimar el gasto TIC.

4.1.2 Encuesta Nacional de Hogares 2018

La segunda fuente de datos a utilizar es la Encuesta Nacional de Hogares (ENAH), la cual es un programa de recolección de datos cuyo enfoque está asociado al nivel de bienestar de la población, especialmente centrados en la conformación del ingreso de los hogares, su distribución y características de los hogares y la población en situación de pobreza. Además, se contempla el estudio de la tenencia de vivienda y sus características, el acceso de las

personas a la educación y al seguro social, así como la población que trabaja y las condiciones de esos trabajos, entre otros (Instituto Nacional de Estadísticas y Censos, 2022). La ENAHO se realiza en julio de cada año, y permite obtener información a nivel nacional, regional y para la zona urbana y rural. El tamaño de muestra de esta encuesta es cercano a las 13 400 viviendas, distribuidas a nivel nacional por región de planificación y zona (urbana y rural). De esta encuesta se empleará la información de los hogares correspondientes a las variables relacionadas con el acceso y uso de los servicios y dispositivos TIC. Adicionalmente, se imputará el gasto TIC, cuyo valor también será requerido para la construcción del IPD.

Sin embargo, cada fuente de datos responde a objetivos diferentes. La ENIGH se especializa en recolectar información sobre los gastos e ingresos en los hogares permitiendo el análisis del consumo y la satisfacción de necesidades básicas. De esta encuesta es posible la aproximación del gasto TIC, así como indagar sobre el acceso a los dispositivos y servicios TIC, requeridos para esta investigación. Por su parte, la ENAHO se especializa en recolectar información para determinar la situación de pobreza de los hogares y las personas e incluye temas de educación, seguridad social, cobertura de programas sociales, empleo, entre otros. De la ENAHO es posible obtener la información sobre el dispositivos y servicios TIC y el uso de éstos; sin embargo, carece de datos sobre el gasto en este rubro.

Por tanto, resulta de interés realizar una comparación entre el IPD obtenido con la ENIGH y el IPD con la ENAHO.

4.1.3 Calidad de los datos

La calidad de los datos o la calidad de la información que se analiza en toda investigación requieren de una valoración previo al inicio de los análisis y procesamientos. Se ha mencionado que asegurar la calidad de los datos en cuanto a la confiabilidad, integridad y exactitud, permitan asegurar que los datos serán de calidad para el uso que se les quiera dar (Troncoso-González, Rodríguez, & Caro, 2022).

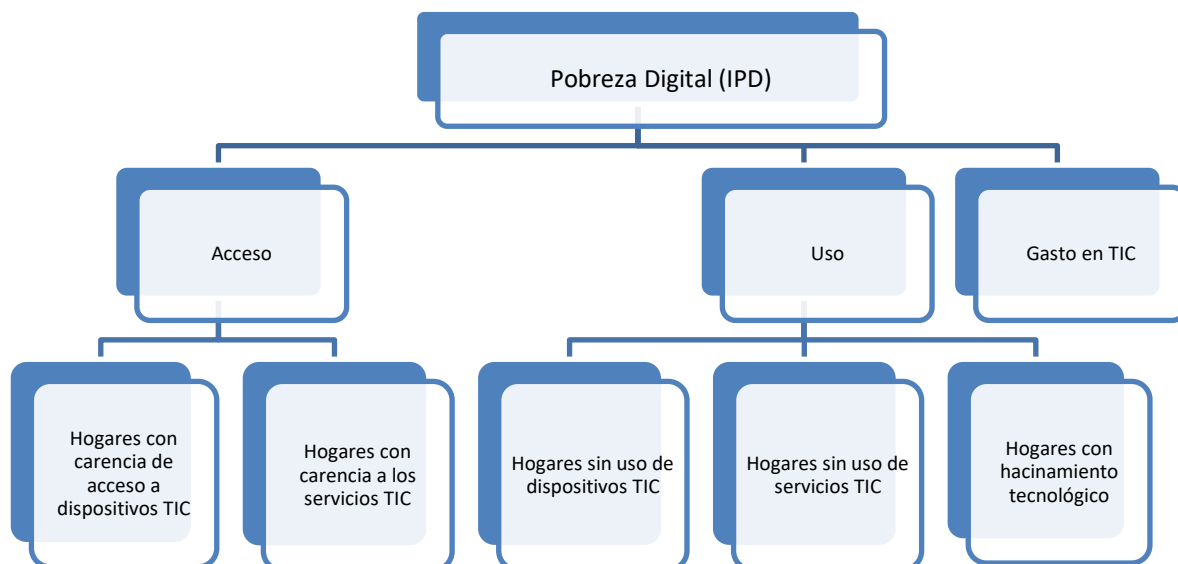
Las bases de datos de la ENAHO y la ENIGH son elaboradas por el INEC. En ambas encuestas, el INEC sigue los procedimientos estadísticos y recomendaciones internacionales

en todas las fases metodológicas: diseño de la muestra, diseño del cuestionario, recolección de datos y procesamiento; con el fin de obtener estadísticas que respondan a estándares de calidad (Instituto Nacional de Estadísticas y Censos, 2019) (Instituto Nacional de Estadísticas y Censos, 2022).

Para esta investigación, previo a la aplicación de las técnicas estadísticas mencionadas, se realizó una valoración general de los datos que se utilizarán mediante frecuencias simples y cruces de variables, entre otros.

4.2 Variables de estudio

En la medición de la Pobreza Digital en Costa Rica se emplea como unidad de análisis el hogar; por tanto, las variables que se analizarán se refieren a valores a este nivel. Basado en la revisión bibliográfica sobre el tema, se propone la siguiente estructura teórica:

Ilustración 1: Diagrama Propuesto del Índice de Pobreza Digital

A continuación, se describen las variables que constituirán cada uno de los componentes del IPD.

4.2.1 Acceso a las TIC

La carencia de acceso a los servicios y dispositivos de las TIC en los hogares es considerada una forma de pobreza digital. Los hogares con limitado o nulo acceso a las TIC tendrán menos posibilidades de aprovechar los beneficios de la sociedad de la información y conocimiento, incidiendo en que sus miembros tengan menos oportunidades de optar por mejores condiciones de vida (Alva de la Selva, 2015).

Aplicando las recomendaciones de la UIT en cuanto al acceso a los servicios y dispositivos y con los datos disponibles de la ENIGH y la ENAHO, para el componente Acceso se operacionalizaron las siguientes variables según subdimensión:

- Hogar con carencia de acceso a dispositivos TIC. Son aquellos hogares que no cuentan con:

- Computadora portátil
 - Computadora de escritorio
 - Tableta
 - Radio o equipo de sonido
 - Televisor
- Hogar con carencia de acceso a servicios TIC. Son los hogares que no cuentan con:
 - Línea telefónica celular
 - Línea de teléfono residencial
 - Internet

4.2.2 Uso de las TIC

Aunado a las carencias en el acceso a los servicios y dispositivos TIC, se ha mencionado sobre las barreras en el uso de estas herramientas y las brechas existentes entre las personas quienes las utilizan y aquellos que no (Alva de la Selva, 2015). En los hogares cuyos miembros tengan un limitado uso de las herramientas tecnológicas, les limita la capacidad de explorar los recursos de las TIC y aplicarlos a sus necesidades particulares (Alva de la Selva, 2015).

Siguiendo esta línea, para el componente de Uso se analizaron las variables:

- Hogar sin uso de dispositivos TIC. Son los hogares donde alguno de sus miembros mayores a 5 años no ha utilizado en los últimos 3 meses cada uno de los siguientes dispositivos:
 - Celular

- Computadora portátil
 - Tableta
 - Computadora de escritorio
- Hogar sin uso de servicios TIC. Son aquellos hogares donde alguno de sus miembros mayores a 5 años no ha utilizado Internet en los últimos 3 meses.
 - Hacinamiento tecnológico. Se encuentra en pobreza digital aquellos hogares donde la cantidad total de dispositivos (computadora de escritorio, portátil, tableta o celular) no es el suficiente para que la totalidad de miembros mayores a 5 años puedan utilizarlos ya sea para recibir lecciones o teletrabajar (Roman Vega, 2021). Esta variable se calcula con la siguiente fórmula:

$$\text{Hacinamiento Tecnológico} = \frac{\text{Total de dispositivos en el hogar}}{\text{Total de miembros de 5 o más años en el hogar}}$$

Para los análisis posteriores, la variable de hacinamiento tecnológico se categorizará asignando el valor 1 a los hogares cuyo cociente de la formula anterior sea superior sea menor a la unidad.

4.2.3 Gasto en TIC

La determinación del gasto en los hogares es una de las variables de mayor uso para el desarrollo de diversos indicadores para la toma de decisiones de política pública. La determinación del gasto en TIC resulta una variable determinante de la pobreza digital de los hogares entendiendo que los hogares que no satisfaga sus necesidades de comunicación e información por medios digitales son catalogados como pobres digitales (Barrantes, 2007).

Para el desarrollo esta investigación, el gasto TIC solo es posible estimarlo en la ENIGH. Dicha estimación se realizó siguiendo los lineamientos plasmados en el Manual para la Medición del Acceso y Uso de la TIC en los hogares y por las personas publicado en el 2020 por la Unión Internacional de Telecomunicaciones mediante la cuantificación del gasto de

los hogares en equipos y servicios TIC (Unión Internacional de Telecomunicaciones, 2020). En dicho Manual, propone utilizar la clasificación CCFI 2018 de equipos y servicios TIC para medir el Gasto en TIC, los cuales se detallan en el siguiente cuadro.

Tabla 1: Clasificación de equipos y servicios de TIC, basada en la CCIF 2018

Código CCFI	Categorías de equipos y servicios de TIC
08 Información y comunicación	<p>08.1) Equipos de información y comunicación</p> <ul style="list-style-type: none"> • Equipos de telefonía fija: teléfonos, radioteléfonos, máquinas telefax, contestadores automáticos y altavoces. • Equipos de telefonía móvil: teléfonos móviles, incluidos los dispositivos polivalentes, y teléfonos inteligentes. • Equipos de procesamiento de la información: computadoras personales, impresoras, escáneres, monitores, proyectores, cascos de realidad aumentada (RA) y de realidad virtual (RV), módems, encaminadores, conmutadores de red y similares, teclados, ratones, digitalizadores; tabletas; calculadoras, incluidas las de bolsillo; máquinas de escribir y procesadores de texto (dispositivo); tóner y cartuchos de tinta, tambores de impresión láser, cintas de máquinas de escribir; e instalaciones de telefax y contestador automático en computadoras. • Equipos para la recepción, la grabación y la reproducción de imagen y sonido: <ul style="list-style-type: none"> ○ televisores, reproductores y grabadoras de cintas de vídeo, grabadoras de vídeo digital, reproductores de DVD, reproductores Blu-ray, reproductores Blue-ray de ultraalta definición, dispositivos de difusión continua, antenas de televisión de todo tipo; ○ receptores de radio (radios, radios digitales, radios Internet, radios por satélite, autorradios, radiorelojes, radios

Código CCFI	Categorías de equipos y servicios de TIC
	<p>bidireccionales, walkie-talkies, receptores y transmisores de radioaficionados);</p> <ul style="list-style-type: none"> ○ reproductores de CD portátiles y no portátiles; ○ reproductores de sonido portátiles y no portátiles; ○ equipos estéreo y grabadoras de CD, radio y cintas; ○ tocadiscos, sintonizadores, amplificadores, platinas, micrófonos y altavoces, equipos de DJ, sistemas de karaoke; ○ sistemas de audio y vídeo para vehículos; ○ descodificadores, receptores de satélite, receptores de TVIP, convertidores de televisión; ○ reproductores de medios digitales; ○ cascos, auriculares y auriculares inalámbricos/Bluetooth. <ul style="list-style-type: none"> • Medios de grabación vírgenes: <ul style="list-style-type: none"> ○ CD (R y RW); ○ DVD (R y RW); ○ discos Blu-ray (R y RE); ○ cintas de vídeo; ○ cintas de audio, cintas de casete, DAT; ○ discos duros externos y discos de estado sólido, NAS (almacenamiento anexo a la red); ○ llaves/memorias USB; ○ tarjetas SD, memorias flash, etc.; ○ cintas de datos magnéticas; ○ otros medios de grabación magnéticos; ○ otros medios de grabación ópticos; ○ otros medios de grabación (medios de grabación por cambio de fase, medios de grabación holográficos, medios de grabación moleculares).

Código CCFI	Categorías de equipos y servicios de TIC
	<p>08.2) Software, excluidos los paquetes de software de juegos informáticos, como sistemas operativos, aplicaciones, lenguajes de programación, etc.</p> <ul style="list-style-type: none"> • Abonos de software y utilización de software en línea. • Aplicaciones.
	<p>08.3) Servicios de información y comunicación</p> <ul style="list-style-type: none"> • Servicios de comunicación fija: <ul style="list-style-type: none"> ○ gastos de instalación y abono de equipos de telefonía personal; ○ llamadas telefónicas desde una línea privada o de una línea pública (cabina telefónica, locutorio, etc.); ○ llamadas locales, regionales, nacionales e internacionales; ○ llamadas telefónicas desde hoteles, cafés, restaurantes y similares. • Servicios de comunicación móvil: <ul style="list-style-type: none"> ○ llamadas locales, regionales, nacionales e internacionales, incluidas las llamadas de voz y vídeo; ○ mensajes, incluidos los mensajes de voz, los mensajes escritos (SMS) y las imágenes (MMS), tasa de abono a otras mensajerías; ○ funcionalidades de llamada adicionales, como el correo vocal y la identificación de llamada, vendidas por separado o integradas en el plan de servicio móvil local; ○ planes de telefonía móvil de voz y mensajes que también comprenden datos limitados; ○ planes de telefonía móvil de voz, texto y datos; ○ otros servicios de telefonía móvil; ○ coste del teléfono, si se incluye en el precio del abono;

Código CCFI	Categorías de equipos y servicios de TIC
	<ul style="list-style-type: none"> ○ teléfono móvil incluido en un paquete, es decir, paquetes de prepago o postpago, generalmente vinculados a un operador concreto durante un periodo de tiempo determinado, si no se paga por separado. • Servicios de acceso a Internet y servicios de almacenamiento en la red: <ul style="list-style-type: none"> ○ Servicios de acceso a Internet prestados por operadores de infraestructura alámbrica, inalámbrica o de satélite. ○ Servicios de almacenamiento en la nube, alojamiento de ficheros y alojamiento web. ○ Abonos a servicios de correo electrónico. <p>Comprende también:</p> <ul style="list-style-type: none"> ○ Tasas y cuotas mensuales de activación e instalación. • Servicios de telecomunicaciones agrupados: <ul style="list-style-type: none"> ○ Paquetes de telefonía/Internet/televisión. ○ Cualquier combinación de paquetes de telecomunicaciones. • Reparación y arrendamiento de equipos de información y comunicación: <ul style="list-style-type: none"> ○ Se incluye el coste de los materiales sólo si no se factura por separado. Comprende: <ul style="list-style-type: none"> ○ reparación de todos los equipos de información y comunicación; ○ arrendamiento de teléfonos, máquinas telefax, contestadores automáticos y altavoces telefónicos; ○ arrendamiento de equipos de telefonía inalámbrica; ○ arrendamiento de equipos de acceso a Internet; ○ arrendamiento de equipos de telegrafía, télex, telefax, radiotelefonía, radiotelegrafía y radiotélex.

Código CCFI	Categorías de equipos y servicios de TIC
	<ul style="list-style-type: none"> • Otros servicios de información y comunicación: <ul style="list-style-type: none"> ○ servicios de telegrafía, télex y telefax; ○ VoIP (Voz por el protocolo Internet) (uso nómada); ○ licencias de radio y televisión; ○ abonos a televisión por cable, televisión por satélite, TVIP y televisión de pago; ○ servicios de difusión en flujo continuo; ○ servicios de grabación de vídeo en línea (servicios DVR web); ○ servicios VOD (vídeo a la carta); ○ arrendamiento o abono de CD, cintas de vídeo, DVD, discos Blu-ray, software (excluido el software de juegos); ○ servicios de radiotelefonía, radiotelegrafía y radiotélex; ○ servicios de instalación de software; ○ arrendamiento/alquiler de descodificadores, adaptadores de televisión, etc.
09.1 Otros bienes recreativos	09.2.1) Juegos, juguetes y aficiones: <ul style="list-style-type: none"> ○ Videojuegos; consolas de videojuegos que se conectan a un televisor; cartuchos de videojuegos, CD-ROM de videojuegos, descargas de videojuegos. • Aplicaciones lúdicas. • Gamepads, joysticks, volantes y otros accesorios de videojuegos. • Juegos electrónicos.

Fuente: Manual para la medición del acceso y uso de las TIC en los hogares y por las personas. Edición 2020, UIT.

Sin embargo, dado que el gasto TIC es uno de los componentes del IPD y que no es posible obtener este valor en la ENAHO, se propone ajustar un modelo de regresión múltiple para aproximar el gasto TIC a partir de la ENIGH 2018 y aplicarlo a la ENAHO para la imputación

del gasto TIC en esta encuesta. Para lograr esto, se propone que el modelo de regresión estime el gasto TIC a partir de las variables del hogar y del jefe del hogar.

4.3 Técnicas estadísticas: descripción de las técnicas usadas

En la construcción del IPD se seguirán los pasos para la construcción de un índice sugeridos en la literatura (Schuschny & Soto, 2009). Se inició con el desarrollo de un marco conceptual sobre el concepto de Pobreza Digital y sus dimensiones. Con este marco, se enlistaron aquellas variables disponibles relacionadas con el tema propuesto, y mediante la aplicación del Análisis Factorial Exploratorio se seleccionaron aquellas que ayuden a explicar el concepto.

4.3.1 Análisis Factorial Exploratorio

En la literatura se reconocen dos tipos Análisis de Factores: exploratorio (AFE) y confirmatorio (AFC). En el primero, se analiza un conjunto de datos sin tener ninguna hipótesis previa acerca de su estructura o la relación de las variables (Ferrando & Anguiano-Carrasco, 2010); mientras que el AFC se realiza cuando la estructura de los datos se deriva de una teoría definida (Nunnally & Bernstein, 1995).

En esta investigación, se empleará sólo el AFE, el cual es una técnica estadística utilizada para identificar las dimensiones subyacentes o constructos latentes en un conjunto de variables observadas (DeVellis & Thorpe, 2021). Estos constructos latentes pueden ser utilizados posteriormente para construir índices o medidas compuestas que representen de manera más precisa los conceptos que se desean medir.

Esta técnica de análisis ha sido empleada en diversas investigaciones. En Argentina se empleó esta técnica para examinar la pobreza multidimensional y así identificar las dimensiones subyacentes a la pobreza (Fagnola & Moneta, 2021). Otro ejemplo del uso del AFE es la evaluación de un Modelo de Calidad de Vida Construido realizado en Chile (Cancino, González, Gallardo, & Estrada, 2016). En Sudáfrica se investigaron los indicadores críticos para evaluar los proyectos de infraestructura de transporte (Okoro,

Musonda, & Agumba, 2019). A nivel nacional, en la investigación de Gómez, se utilizó esta técnica para explorar las dimensiones que componen los índices de intensidad de uso de Internet, frecuencia de uso de medios, participación en organizaciones y de esparcimiento, relacionados con el uso del tiempo de las personas jóvenes (Gómez Meléndez, 2014). Por su parte Montero y Jiménez, aplicaron el análisis de factores exploratorio para verificar la unidimensionalidad de la prueba de “Diagnóstico de conocimientos y destrezas en matemática del estudiante al ingresar a la universidad”, de la Escuela de Matemática de la Universidad de Costa Rica (Jiménez Alfaro & Montero Rojas, 2013).

A nivel general, la aplicación del Análisis Factorial supone que para un vector de respuesta \mathbf{x} , de p variables, de una población con media $\boldsymbol{\mu}$ y matriz de varianzas-covarianzas $\boldsymbol{\Sigma}$, de la cual se puede obtener m factores subyacentes (donde se desea que $m < p$), denotados por f_1, f_2, \dots, f_m tales que (Johnson, 1998):

$$x_j = \mu_j + \lambda_{j1}f_1 + \lambda_{j2}f_2 + \dots + \lambda_{jm}f_m + \eta_j \text{ para } j = 1, 2, \dots, p$$

Donde se supone que:

1. los f_k son independientes e idénticamente distribuidos, con media 0 y varianza 1, para $k = 1, 2, \dots, m$;
2. los η_j están independientemente distribuidos, con media 0 y varianza Ψ_j , para $j = 1, 2, \dots, p$; y
3. f_k y η_j tienen distribuciones independientes para todas las combinaciones de k y j , $k = 1, 2, \dots, m$ y $j = 1, 2, \dots, p$.

En la terminología del análisis por factores presentada por Johnson (1998), las variables f_1, f_2, \dots, f_m son denominados *factores comunes*, y $\eta_1, \eta_2, \dots, \eta_p$ son los *factores específicos*; el valor de cada η_j describe la variación residual específica de la j -ésima variable respuesta y Ψ_j es la varianza específica de la j -ésima variable. Además, los λ_{jk} son llamados *cargas de los factores*, es decir miden la contribución del k -ésimo factor común a la j -ésima variable respuesta. Para este caso particular, al tratarse de variables dicotómicas se utilizará

la matriz de correlaciones tetracórica dado que genera mejores resultados (Dominguez Lara, 2014).

4.3.2 Modelo de Regresión Múltiple para la variable Gasto TIC

Por otro lado, se mencionó la necesidad de imputar el gasto TIC en la ENAHO, dado que esta encuesta se realiza anualmente pero no recolecta información sobre los gastos de los hogares; por tanto, se propone el ajuste de un Modelo de Regresión Múltiple. Este modelo, estudia la relación entre una variable dependiente y (gasto TIC, en este caso) y diversas variables explicativas x_1, x_2, \dots, x_k y se denota de la siguiente manera (Maddala, 1996):

$$y_i = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i, \quad i = 1, 2, \dots, n$$

Donde los componentes α y β_k son los coeficientes o parámetros de regresión, estimados a partir de los datos observados y los u_i son los errores de medición en y y a la especificación de la relación entre y (dependiente) y las x_k (explicativas).

Por consiguiente, se ajustó una ecuación de regresión lineal múltiple, que estimó el gasto en el hogar en TIC con las siguientes variables independientes:

- Relacionadas con el hogar: número de miembros ocupados, ingreso bruto, zona de residencia.
- Relacionadas con el jefe del hogar: nivel educativo del jefe, rama de actividad y ocupación del jefe.

El procedimiento para el tratamiento de las variables independientes que se incluirán en el modelo de regresión es de la siguiente forma.

El número de miembros ocupados (num_ocup) del hogar es la cantidad de habitantes del hogar que cumplen con la condición de actividad ocupado. Se espera que a mayor número de ocupados el gasto TIC sea mayor o a la inversa.

La variable del ingreso total bruto del hogar (ingr_total) está definida como la suma del ingreso corriente bruto con valor locativo más el ingreso por transacciones de capital. Teóricamente

se ha mencionado una relación recíproca entre el gasto del hogar y el ingreso de este. Siendo el gasto en TIC un componente del gasto del hogar, se espera que esta relación de reciprocidad en el gasto en TIC se mantenga con el ingreso del hogar. A los valores de esta variable se les aplicará la transformación del logaritmo natural para mejorar el ajuste del modelo de regresión.

En cuanto al nivel educativo del jefe (*educ_jefe*) del hogar, la ENIGH recolecta la información sobre el último año aprobado por este. Para su utilización, se escaló de la siguiente manera:

Tabla 2: Escalamiento de la variable nivel educativo del jefe del hogar de la ENIGH.

Categorías		Escalamiento
0 Ningún grado 1 Preparatoria	2 Enseñanza especial 99 Ignorado	0 Sin estudios
11 Primer año de primaria 12 Segundo año de primaria 13 Tercero año de primaria	14 Cuarto año de primaria 15 Quinto año de primaria 19 Primaria ignorados	1 Primaria incompleta
16 Sexto año de primaria		2 Primaria completa
21 Primer año de secundaria académica 22 Segundo año de secundario académica 23 Tercer año de secundaria académica 24 Cuarto año de secundaria académica 25 Quinto año de secundaria académica	31 Primer año de secundaria técnica 32 Segundo año de secundaria técnica 33 Tercer año de secundaria técnica 34 Cuarto año de secundaria técnica 35 Quinto año de secundaria técnica	3 Secundaria incompleta

Categorías		Escalamiento
29 Secundaria académica ignorado	36 Sexto año de secundaria técnica	
26 Tiene el bachillerato de secundaria académica	37 Tiene el bachillerato de secundaria técnica	4 Secundaria completa
41 Primer año en parauniversitaria 42 Segundo año en parauniversitaria 43 Tercer año en parauniversitaria	51 Primer año de universidad 52 Segundo años de universidad 53 Tercer año de universidad	5 Universitaria incompleta
54 Cuarto año de universidad 55 Quinto año de universidad 56 Sexto año de universidad 57 Séptimo año de universidad 61 Especialidad un año 62 Especialidad dos años 64 Especialidad cuarto año	71 Maestría, doctorado un año 72 Maestría, doctorado dos años 73 Maestría, doctorado tres años 74 Maestría, doctorado cuatro años 75 Maestría, doctorado cinco años 79 Maestría, doctorado Ignorada	6 Universitaria completa

Se espera que, a mayor nivel educativo del jefe del hogar, se requiera una mayor demanda de servicios y uso de dispositivos TIC y por ende el gasto en este rubro aumente.

Finalmente, con las variables Rama de actividad y Ocupación del jefe del hogar, se procede a la construcción de un índice (Ind_Ocup) que relacione estas variables con el ingreso corriente bruto del jefe del hogar, dado que es de esperar que los gastos del hogar estén relacionados con el ingreso. Para la construcción de este índice se aplicó la técnica del *Cluster*

Analysis a la estimación del promedio del ingreso corriente bruto del jefe del hogar por rama de actividad y por ocupación (ver Anexo 1), con el objetivo de clasificar las ocupaciones y las ramas de actividad de la manera óptima, procurando obtener una menor cantidad de grupos de categorías logrando que la variancia del ingreso dentro de los grupos sea la más pequeña posible y entre los grupos sea la mayor. Se realizaron 21 clasificaciones diferentes, formando desde 30 hasta 45 grupos. Para determinar la cantidad adecuada de grupos o clúster para los análisis subsiguientes se correlacionó cada agrupación con la variable Gasto TIC, oscilando este valor entre 0,421 y 0,440, por lo que se decidió utilizar la clasificación de 40 grupos.

4.3.3 *Cluster Analysis*

El análisis de *clusters* también conocido como análisis de conglomerados, es una técnica estadística utilizada para agrupar objetos o datos en grupos o clusters basados en sus similitudes. El objetivo principal del análisis de clusters es encontrar patrones o estructuras intrínsecas en los datos sin tener conocimiento previo y agrupar a los objetos o datos de manera que los elementos dentro de cada cluster sean similares entre sí y que haya diferencias significativas entre los clusters (Vicente Villardón, 2007).

Este agrupamiento es una técnica de aprendizaje no supervisado, lo que significa que no requiere datos de etiquetados o identificados (Jain, Murty, & Flynn, 1999).

El análisis de clusters se utiliza en diversas áreas de investigación desde la segmentación de mercado hasta el análisis de redes sociales y detección de anomalías. Por ejemplo, los investigadores Terrones y Vargas (Terrones & Vargas, 2013), aplicaron el análisis de clúster para obtener una clasificación de los bancos comerciales del sistema bancario peruano en grupos homogéneos utilizando la información anual para el período 2010-2012 de 15 bancos comerciales y 16 indicadores financieros clasificados en 7 grupos: calidad de activos, liquidez, riesgo cambiario, rentabilidad, solvencia, estructura de activo y estructura de pasivo. Por otro lado, los puertos españoles se clasificaron utilizando una serie de indicadores que caracterizan la actividad portuaria y se logró demostrar que los puertos españoles se

pueden caracterizar correctamente por medio de indicadores físicos y de explotación (Camarero Orive, Cerbán Jiménez, Turias Domínguez, González Cancelas, & Camarero Orive, 2016). Investigadores en China aplicaron el análisis de agrupamiento de similitud multivariante (MSCA, por sus siglas en inglés) para estimar un esquema de división biogeográfica para la fauna terrestre, obteniéndose un total de siete reinos y 20 sub-reinos. Además, hallaron que los resultados de agrupación para los animales superiores e inferiores no diferían significativamente, lo que nos lleva a considerar que el resultado del análisis es convincente como el primer esquema de división zoogeográfica para todos los animales terrestres (Shen, y otros, 2022).

En Costa Rica, Madrigal en su investigación “Estratificación de hogares y segmentos por niveles de ingreso en el censo 2000”, empleó esta técnica para la creación del índice de ocupación y también en la estratificación de los hogares, aplicando el procedimiento de K-medias (Madrigal, 2000). Por su parte, Monge y Rodríguez aplicaron el análisis clúster para la caracterización de 809 empresas PYME según su grado de madurez, su permanencia y éxito en el mercado en que operan, siendo esta clasificación útil para la evaluación del acceso a financiamiento de las empresas o a programas de incentivos por parte de instituciones públicas (Monge González & Rodríguez Álvarez, 2011).

Por tanto, para la estratificación de los hogares según el nivel de Pobreza Digital, se aplicó el *Cluster Analysis*, con las dimensiones de la ENIGH y con las correspondientes de la ENAHO. Según indica Dillon y Goldstein, el proceso de *Cluster Analysis* toma una matriz de p variables y n sujetos, a la cual se le aplica algún algoritmo de clasificación para formar subgrupos de sujetos (*clusters*). El objetivo es que la variación dentro de cada subgrupo sea mínima y que la variación entre subgrupos sea máxima (Dillon & Goldstein, 1984)

4.3.4 Softwares utilizados

El principal software utilizado es el SPSS (*Statistical Package for the Social Sciences*). Adicionalmente se emplea el software R empleado a través de la interfase *Rstudio* versión

2024.12.1+563 (Posit Software, 2024) para la estimación de la matriz de correlaciones tetracórica que se usa en el AFE.

4.4 Alcances y Limitaciones

Esta investigación presentó aportes al tema de la pobreza digital, no obstante, es necesario reconocer los alcances y limitaciones que restringen la generalización de los resultados.

En primer lugar, el estudio constituye una propuesta metodológica para medir la pobreza digital en Costa Rica, a nivel de los hogares. A partir de las bases de datos de la ENAHO y la ENIGH, se seleccionaron las variables que según el marco teórico son las idóneas para construir un índice que permita visibilizar desigualdades en el acceso y uso de las tecnologías de la información y la comunicación.

La propuesta posee una vinculación parcial con las mediciones de pobreza (línea de pobreza y pobreza multidimensional) desde una perspectiva metodológica; sin embargo, el ámbito digital es considerado de forma limitada. De esta manera, esta investigación ofrece un referente inicial que puede ser replicado y perfeccionado en estudios posteriores. Asimismo, se sugiere que futuras investigaciones consideren la metodología Alkire-Foster, con el fin de fortalecer la medición de la pobreza digital desde una perspectiva multidimensional (Alkire, Conconi, & Seth, 2014).

Con respecto a las limitaciones, se aclara que éstas se derivan principalmente de las fuentes de información utilizadas. Al tratarse de encuestas nacionales de hogares (ENAHO y ENIGH), el estudio se restringe al nivel de los hogares y a las variables disponibles en dichos instrumentos. Algunas dimensiones relevantes para caracterizar la pobreza digital no se encuentran incluidas en estas encuestas, como la finalidad específica del uso de las TIC, la frecuencia de uso y la diversidad de los trámites realizados en línea.

Asimismo, la aplicación de las técnicas estadísticas exigió la exclusión de ciertas variables, lo que redujo la cobertura de aspectos vinculados al fenómeno. Existen módulos especializados en el uso de TIC incluidos en la ENAHO por la Superintendencia de

Telecomunicaciones (SUTEL), sin embargo, al no estar disponibles para su análisis, no pudieron ser incorporados en la propuesta. En consecuencia, si bien el índice constituye un avance en la medición de la pobreza digital, sus resultados deben interpretarse considerando estas limitaciones y la necesidad de complementarlo con nuevas fuentes y metodologías en el futuro.

CAPÍTULO V. RESULTADOS

5.1 Análisis descriptivo de las variables de estudio

El análisis descriptivo de las variables de estudio permite determinar el comportamiento de dichas variables en cuanto a su dispersión, detección de valores extremos, valores atípicos y determinar algunas relaciones a priori. En este apartado, se presenta el análisis descriptivo de las variables que se analizaron.

5.1.1 Variables sobre el Acceso a las TIC

Las variables seleccionadas sobre el Acceso a las Tecnologías de Información y Comunicación que se utilizaron se describen en el cuadro 1. Se presenta la distribución de frecuencia simple para las variables seleccionadas, tanto para la ENAHO y la ENIGH.

De este cuadro, se observa que cerca de 8 de cada 10 hogares no cuenta con Tablet, el 85% no poseen computadora de escritorio y el 58% carece de computadora portátil y que el radio o equipo de sonido no está presente en 4 de cada 10 hogares. El dispositivo más frecuente es el televisor, sólo el 3% de los hogares carecen de este.

Cuadro 1: Porcentaje de hogares sin acceso a dispositivos y servicios TIC

	2018	
	ENIGH	ENAHO
Total de Hogares	1 538 704	1 561 637
Hogares <u>sin acceso</u> a los dispositivos:		
Computadora de escritorio	85,7	84,9
Computadora portátil	58,3	59,8
Tableta	80,9	80,2
Radio o equipo de sonido	38,1	40,0
Televisor	2,9	3,2
Hogares <u>sin acceso</u> a los servicios de:		
Línea telefónica fija	66,3	66,8
Línea telefónica celular	3,2	4,1
Internet	22,5	26,8

Fuentes: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos de los hogares 2018, INEC. [Archivo de datos].

En cuanto a los servicios, 2 de cada 3 hogares no tiene acceso al servicio de telefonía fija y entre 3% y 4% indican que no cuentan con línea telefónica celular. En relación con el servicio de Internet, se tiene una ligera discrepancia en los porcentajes obtenidos de ambas encuestas, ya que la ENIGH reporta que el 22,5% de los hogares entrevistados no cuenta con este servicio, mientras que la ENAHO estima este porcentaje en 26,8%.

5.1.2 Variables sobre el Uso a las TIC

El uso de dispositivos e Internet se limita a los miembros del hogar con edades iguales o superiores a los 5 años, ya que la metodología de la ENAHO así lo contempla. El porcentaje que se presenta en el cuadro 2 hace referencia a los hogares en los cuales alguno de sus miembros no hizo uso del dispositivo o servicio. Así, en los últimos 3 meses en 9 de cada 10 hogares alguno de sus miembros no usó la tableta ni la computadora de escritorio; en el

81,7% de los hogares algún miembro no hizo uso de la computadora portátil. En el caso del celular, en el 25,3% de los hogares indicaron algún miembro no hizo uso de este dispositivo.

Cuadro 2: Porcentaje de hogares sin uso de los dispositivos y servicios TIC

	2018
Total de Hogares	1 561 637
Hogares <u>sin uso</u> de los dispositivos:	
Computadora de escritorio	90,0
Computadora portátil	81,7
Tableta	90,7
Celular	25,3
Hogares <u>sin uso</u> del servicio:	
Internet	44,0

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

En cuanto al uso del servicio TIC, se obtiene que en el 44% de los hogares alguno de sus miembros no ha usado Internet en los últimos 3 meses (Cuadro 2).

Por otro lado, el termino de hacinamiento tecnológico hace referencia a la cantidad de dispositivos disponibles en el hogar en comparación con la cantidad de los miembros del hogar con edades superiores a los 5 años. Específicamente, se estable que los dispositivos adecuados para acceder a los beneficios de la era digital son el celular, las computadoras (escritorio y portátil) y la Tablet. Así, según la formula descrita en el apartado 3.2, se define:

- Hogar con hacinamiento tecnológico: aquellos hogares cuyo cociente es menor al valor 1; es decir, cuando la cantidad de dispositivos en el hogar es menor a la cantidad de miembros de este con edades superiores a los 5 años.
- Hogares con equidad tecnológica: aquellos hogares con igualdad de cantidad de dispositivos y de miembros del hogar y por tanto su cociente es igual al valor 1.

- Hogares sin hacinamiento tecnológico: aquellos hogares con una cantidad de dispositivos superior a la cantidad de miembros del hogar.
- Hogares sin dispositivos tecnológicos: aquellos hogares que no poseen ninguno de los dispositivos mencionados.

En el cuadro 3, se desglosa la distribución porcentual de los hogares según las categorías mencionadas.

Cuadro 3: Porcentaje de hogares según nivel de hacinamiento tecnológico.

	2018	
	ENIGH	ENAHO
Total de Hogares	1 538 704	1 561 637
Hogares con hacinamiento tecnológico	25,7	26,3
Hogares con equidad tecnológica	31,6	29,5
Hogares sin hacinamiento tecnológicos	39,6	40,2
Hogares sin dispositivos tecnológicos	3,1	4,1

Fuentes: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos de los hogares 2018, INEC. [Archivo de datos].

Para efectos de los análisis posteriores, se agrupan en dos categorías

- Hogares con hacinamiento tecnológico y sin dispositivos tecnológicos se agrupan en una misma categoría. Así el 28,8% de los hogares de la ENIGH y el 30,4% de los hogares de la ENAHO poseen hacinamiento tecnológico.
- Hogares sin hacinamiento y con equidad tecnológica. El 69,7% de los hogares de la ENAHO y el 71,2% de los hogares de la ENIGH pertenecen a esta categoría.

5.1.3 Variable Gasto en TIC

Como se expuso en el apartado 3.2.iii Gasto en TIC, la variable fue construida con la suma simple de los montos registrados en la ENIGH para los equipos y servicios de TIC definidos en el Manual para la Medición del Acceso y Uso de las TIC de la UIT. En el cuadro 4, se presentan las estadísticas descriptivas del gasto medio del hogar en TIC desglosada por zona y región de planificación.

Cuadro 4: Estadísticas Descriptivas del Gasto mensual en TIC de los hogares de la ENIGH¹

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
Gasto TIC (€)	49 183,2	900,5	47 413,7	50 952,8	1,8	2,0	6 482
Zona							
Urbana	54 103,4	1 093,6	51 954,5	56 252,3	2,0	2,0	4 361
Rural	36 543,4	1 011,2	34 556,4	38 530,5	2,8	1,6	2 121
Región							
Central	54 637,0	1 337,8	52 008,1	57 265,8	2,4	2,3	1 991
Chorotega	42 462,0	1 532,1	39 451,4	45 472,7	3,6	1,1	954
Pacífico Central	43 130,0	1 544,7	40 094,6	46 165,5	3,6	1,0	815
Brunca	37 218,0	1 254,1	34 753,5	39 682,4	3,4	1,0	909
Huetar Atlántico	38 930,7	1 735,7	35 520,0	42 341,5	4,5	1,4	910
Huetar Norte	38 560,9	1 282,4	36 040,9	41 080,9	3,3	1,0	903

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC. [Archivo de datos].

¹ Se excluyen los valores atípicos menores al percentil 2 (€4330) y mayores al percentil 99,9 (€279 167).

La estimación del gasto en TIC de los hogares es de ¢49 183,2 mensuales. Por zona, en la urbana se estima un gasto de ¢54 103, mientras que en la rural este gasto es de ¢36 543.

5.2 Estimación de la variable del gasto TIC por medio del modelo de regresión múltiple.

Con la estimación del gasto en TIC por hogar, se procede al ajuste de un modelo de regresión para esta variable. Para ejecutar este paso, se construye una base con las variables referidas al hogar extraídas de la ENIGH Base Hogares y de la ENIGH Base Persona se extrae la información relacionada con el jefe del hogar. Esta base personalizada, se divide en dos submuestras para aplicar en la primera mitad, los análisis respectivos y en la segunda, se valida. Para realizar la división de la muestra, primero se ordenan geográficamente las UPM y se aplicó una selección sistemática de uno cada dos.

Un aspecto por considerar para el desarrollo del modelo de regresión para la variable gasto TIC en hogares, es el tratamiento del jefe del hogar. El modelo planteado incluye las variables de condición de actividad y ocupación del jefe del hogar; sin embargo, el 31,9% de los jefes de hogar de la ENIGH están desocupados o fuera de la fuerza de trabajo. Ante este panorama, se modifica la definición del jefe del hogar de la siguiente forma (Madrigal, 2000):

- a) Se selecciona al jefe del hogar cuya condición de actividad sea ocupado (68,1% de los hogares).
- b) En los hogares cuyo jefe está en desocupado o fuera de la fuerza de trabajo, se selecciona como jefe al residente de mayor edad cuya condición de actividad sea ocupado (14,6% de los hogares).
- c) Se selecciona como jefe al residente de mayor edad si en el hogar no hay personas ocupadas (17,3%).

Esto produce que se requieran de dos ecuaciones para estimar el gasto en TIC; una para aquellos hogares cuyo jefe se encuentra ocupado (grupos a y b) y otra para los hogares cuyo jefe no está ocupado y es sustituido por el residente de mayor edad (grupo c). Con esta estrategia se logra estimar el gasto TIC en todos los hogares de la ENIGH 2018.

5.2.1 Relaciones bivariadas del modelo de regresión para el Gasto TIC

En este apartado, se analiza las relaciones bivariadas de la variable Gasto TIC con las variables propuestas para el modelo, enfatizando en la dirección e intensidad en la correlación y así determinar si el planteamiento teórico expuesto es adecuado, así como la revisión de la multicolinealidad entre variables.

Para la aplicación del modelo de regresión, se procede a la transformación de la variable Gasto TIC empleando el logaritmo natural para mejorar el ajuste del modelo de regresión.

En el cuadro 5, se presenta los estadísticos descriptivos de las variables independientes del modelo de regresión para la estimación del gasto TIC. Se observan diferencias según zona de residencia. Las estimaciones del ingreso total bruto, del número de miembros ocupados y la escolaridad del jefe del hogar para los datos correspondientes a la zona urbana fueron superiores a las respectivas estimaciones de la zona rural. En la zona rural se estima un valor mayor de miembros del hogar y de porcentaje de dependientes.

Cuadro 5: Estadísticos descriptivos de las variables para el modelo de regresión para la estimación del gasto TIC de la submuestra 1.

Variables	Promedio			Mín.	Máx.	Error estándar
	Total	Urbano	Rural			
Miembros ocupados	1,421	1,440	1,370	0	7	0,022
LN Ingreso total bruto del hogar	13,659	13,800	13,248	10,620	16,890	0,019
Escolaridad del jefe	3,312	3,583	2,577	0	6	0,041
Índice ocupación	7,952	8,943	5,292	1	40	0,226

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

Las correlaciones de Pearson de las variables independientes y el LN del Gasto TIC se detallan en el cuadro 6. Entre las variables LN ingreso bruto del hogar y LN Gasto TIC se

presentó una correlación de 0,655. El número de miembros ocupados del hogar, la escolaridad del jefe y el índice de ocupación también correlacionaron positivamente con el LN Gasto TIC con valores entre 0,346 y 0,426. Las correlaciones resultantes entre las variables independientes y la variable de estudio se dan en la dirección esperada y sus valores indican relaciones moderadas.

Cuadro 6: Coeficientes de correlación de las variables para el modelo de regresión para la estimación del gasto TIC de la submuestra 1.

	Miembros ocupados	LN Ing. total bruto del hogar	Escol. del jefe	Índ. Ocupación
Total				
Miembros ocupados	1,000			
LN Ing. total bruto del hogar	0,409	1,000		
Escol. del jefe	0,052	0,564	1,000	
Índ. Ocupación	0,001	0,556	0,687	1,000
LN (Gasto TIC)	0,346	0,655	0,426	0,416
Urbano				
Miembros ocupados	1,000			
LN Ing. total bruto del hogar	0,395	1,000		
Escol. del jefe	0,035	0,553	1,000	
Índ. Ocupación	-0,011	0,568	0,686	1,000
LN (Gasto TIC)	0,342	0,653	0,420	0,418
Rural				
Miembros ocupados	1,000			
LN Ing. total bruto del hogar	0,474	1,000		
Escol. del jefe	0,076	0,465	1,000	
Índ. Ocupación	-0,005	0,419	0,634	1,000
LN (Gasto TIC)	0,366	0,565	0,283	0,282

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

De las correlaciones entre las variables independientes se observa que la máxima y positiva se presenta entre el índice de ocupación y la escolaridad del jefe (0,687). El LN del ingreso bruto total también muestra una correlación positiva con la escolaridad y el índice de ocupación, con valores de 0,564 y 0,556, respectivamente. Este patrón se mantiene entre en la zona urbana y en la zona rural.

5.2.2 Ecuación de regresión para estimar el Gasto TIC

La estimación del gasto TIC se realiza mediante un modelo de regresión múltiple con dos ecuaciones, una correspondiente a los hogares cuya jefatura se encontraba en condición de actividad ocupado y otra para los hogares que no tenían personas ocupadas, ambas para la primera submuestra. Adicionalmente se incluyó la variable zona de residencia con las categorías urbana y rural, con el objetivo de estimar el gasto en TIC para todos los hogares.

5.2.2.1 Ecuación del Gasto TIC en hogares con jefatura ocupada

En este apartado se presentan los resultados del ajuste de la ecuación de regresión para estimar el Gasto TIC en los hogares con las siguientes condiciones:

- Jefatura del hogar cuya condición de actividad sea ocupado (68,1% de los hogares).
- Jefatura del hogar está en desocupado o fuera de la fuerza de trabajo sustituido por el residente de mayor edad cuya condición de actividad sea ocupado (14,6% de los hogares).

En el Cuadro 7 se desglosa los resultados. Como primer hallazgo se tiene que el modelo de regresión es significativo, lo que implica que las variables incluidas contribuyen significativamente a explicar la variación del logaritmo natural del Gasto TIC del hogar. El segundo hallazgo, se observa que los coeficientes de las variables son significativos por lo que es razonable incluirlos en el modelo. En el Anexo 2, se desglosan el resumen del modelo, las estadísticas de colinealidad y el gráfico de residuos.

Cuadro 7: Resultados del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición ocupada (Submuestra 1)

Variable	Coef. Beta	Error Estándar	Coef. Beta estandarizado	t	Sig.	IC para Beta	
						Inferior	Superior
Constante	5,003	0,013	--	372,988	0,000	4,977	5,029
Miembros ocupados (<i>Nocup</i>)	0,092	0,001	0,111	99,462	0,000	0,091	0,094
LN Ing. total bruto del hogar (<i>LNIng</i>)	0,390	0,001	0,503	358,149	0,000	0,388	0,392
Escol. del jefe (<i>Escol</i>)	0,007	0,000	0,085	59,676	0,000	0,006	0,007
Índ. Ocupación (<i>IOcup</i>)	0,019	0,001	0,050	34,246	0,000	0,018	0,020
Zona ¹ (<i>Zon</i>)	0,112	0,002	0,078	74,895	0,000	0,109	0,115

¹ Urbana =1, Rural =0.

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

La ecuación resultante para estimar el logaritmo natural del Gasto TIC del hogar con jefatura ocupada es la siguiente:

$$LN(\text{GastoTIC}) = 5,003 + 0,092 * \text{Nocup} + 0,390 * \text{LNIng} + 0,007 * \text{Escol} + 0,019 * \text{IOcup} + 0,112 * \text{Zon}$$

$$(R^2 = 0,427)$$

El coeficiente de determinación indica que, para los hogares con jefatura ocupada, las variables incluidas explican el 42,7% de la variabilidad del logaritmo natural del Gasto TIC.

Con respecto a los supuestos para este modelo, se tiene que:

- los residuos presentan autocorrelación; sin embargo, por tratarse de datos de corte transversal, su influencia no es preocupante para el análisis,

- la multicolinealidad se determinó mediante el estadístico VIF cuyos valores fueron menores a 5, por lo que no se requiere intervención.
- Las gráficas para los residuos del modelo y los gráficos de residuos contra predichos no parecen tener un patrón, así que sí puede suponer homocedasticidad.

5.2.2.2 Ecuación del Gasto TIC en hogares con jefatura no ocupada

Los resultados del ajuste del modelo de regresión para los hogares con jefatura no ocupada se muestran en el cuadro 8 y el resumen del modelo, las estadísticas de colinealidad y el gráfico de residuos se encuentran en el Anexo 3.

Cuadro 8: Resultados del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición no ocupada (Submuestra 1).

Variable	Coef. Beta	Error Estándar	Coef. Beta estandarizado	t	Sig.	IC para Beta	
						Inferior	Superior
Constante	5,577	0,030	--	188,401	0,000	5,519	5,635
LN Ing. total bruto del hogar (<i>LNIng</i>)	0,345	0,002	0,473	142,555	0,000	0,341	0,350
Escol. del jefe (<i>Escol</i>)	0,055	0,001	0,169	52,399	0,000	0,053	0,057
Zona ¹ (<i>Zon</i>)	0,062	0,004	0,047	16,849	0,000	0,055	0,069

¹ Urbana =1, Rural =0.

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

El ajuste de la ecuación de regresión para estimar el Gasto TIC en los hogares con jefatura no ocupada se desglosa es la siguiente:

$$LN(\text{GastoTIC}) = 5,577 + 0,345 * LN\text{Ing} + 0,055 * \text{Escol} + 0,062 * \text{Zon} \quad (R^2 = 0,366)$$

En el Anexo 3 se desglosa los resultados. Similar a los resultados de la ecuación anterior, se tiene que el modelo de regresión resulta significativo y los coeficientes de las variables son significativos por lo que es razonable incluirlos en el modelo. El coeficiente de determinación indica que, para los hogares con jefatura desocupada, las variables incluidas explican el 36,6% de la variabilidad del logaritmo natural del Gasto TIC. Con respecto a los supuestos para este modelo, se tiene que:

- los residuos presentan autocorrelación; sin embargo, por tratarse de datos de corte transversal, su influencia no es preocupante para el análisis,
- la multicolinealidad se determinó mediante el estadístico VIF cuyos valores fueron menores a 5, por lo que no se requiere intervención
- las gráficas para los residuos del modelo y los gráficos de residuos contra predichos no parecen tener un patrón, así que sí puede suponer homocedasticidad

5.2.2.3 Validación de las ecuaciones de regresión en segunda submuestra

Los modelos ajustados fueron validados en la segunda submuestra, para lo cual se utilizaron las ecuaciones construidas se estimó el logaritmo natural del ingreso del Gasto TIC para posteriormente correlacionarlo con el valor transformado estimado en la Encuesta. La correlación de Pearson a nivel nacional fue de 0,674; para la zona urbana fue de 0,666 y para la rural de 0,601.

Adicionalmente, en la segunda submuestra se estimó los modelos propuestos para determinar la validez de estos. Los resultados corroboran que la magnitud de los coeficientes, el signo y la significancia, al igual que los coeficientes de determinación, fueron muy similares a los obtenidos en la primera submuestra. En el Anexo 4 y 5, se detalla los resultados.

5.3 Estimación de la variable del gasto TIC por medio del modelo de regresión múltiple en la ENAHO.

Para la aplicación de las ecuaciones del modelo de estimación del gasto TIC obtenido de la ENIGH en la base de la ENAHO, se necesitó la modificación de la variable Nivel de Instrucción según la siguiente categorización:

Tabla 3: Escalamiento de la variable nivel educativo del jefe del hogar de la ENAHO.

Categorías		Escalamiento
Años de Escolaridad	Nivel de Instrucción	
0 Sin escolaridad, preescolar o enseñanza especial	0 Sin nivel de instrucción	0 Sin estudios
1 Un año 2 Dos años 3 Tres años 4 Cuatro años 5 Cinco años 99 Ignorado	1 Primaria incompleta	1 Primaria incompleta
6 Seis años	2 Primaria completa	2 Primaria completa
7 Siete años 8 Ocho años 9 Nueve años 10 Diez años 99 Ignorado	3 Secundaria académica incompleta	3 Secundaria incompleta
7 Siete años 8 Ocho años 9 Nueve años 10 Diez años 11 Once años	5 Secundaria técnica incompleta	
11 Once años	4 Secundaria académica completa	4 Secundaria completa

Categorías		Escalamiento
Años de Escolaridad	Nivel de Instrucción	
12 Doce años	6 Secundaria técnica completa	
12 Doce años 13 Trece años 14 Catorce años 99 Ignorado	7 Educación superior de pregrado y grado	5 Universitaria incompleta
15 Quince años 16 Dieciseis años 17 Diecisiete años	7 Educación superior de pregrado y grado	6 Universitaria completa
16 Dieciseis años 17 Diecisiete años 18 Dieciocho años 20 Veinte años 21 Veintiuno años 22 Veintidós años	8 Educación superior de posgrado	

En el caso de la variable Índice de ocupación, se aplicó la misma categorización utilizada en la ENIGH que se detalla en el Anexo 1.

En el cuadro 9 se desglosa la estimación del gasto de los hogares en TIC. A nivel nacional, se obtiene que el gasto es de ¢49 576. Por zona, se tiene que en la urbana este gasto es de ¢54 927 y en la rural de ¢35 611.

Cuadro 9: Estadísticas Descriptivas de la estimación Gasto TIC en los hogares de la ENAHO¹

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
Gasto TIC (€)	49 576,3	565,7	48 466,2	50 686,3	1,1	1,7	10 762
Zona							
Urbana	54 927,2	677,1	53 598,8	56 255,7	1,2	1,6	7 394
Rural	35 611,7	517,0	34 597,2	36 626,1	1,5	1,3	3 368
Región							
Central	55 274,9	762,2	53 779,4	56 770,4	1,4	1,7	6 050
Chorotega	43 912,0	1 613,9	40 745,4	47 078,7	3,7	1,5	996
Pacífico Central	42 726,1	1 679,7	39 430,4	46 021,8	3,9	1,5	769
Brunca	38 262,9	1 353,7	35 606,8	40 919,0	3,5	1,5	994
Huetar Atlántico	39 991,6	1 195,7	37 645,6	42 337,7	3,0	1,5	1 076
Huetar Norte	37 889,8	1 203,1	35 529,3	40 250,3	3,2	1,4	877

Fuente: elaboración propia con datos de la Encuesta Nacional de los Hogares 2018, INEC. [Archivo de datos].

5.4 Construcción de las dimensiones

En este apartado se compone de dos secciones. La primera sección contiene los resultados del análisis de confiabilidad para determinar la consistencia interna y en la segunda se presenta el AFE para la definición de las dimensiones. Es necesario indicar que en la evidencia teórica existen modelos confirmatorios sobre las dimensiones de acceso y uso de los dispositivos y servicios de telecomunicaciones; sin embargo, por las características del sector en Costa Rica, esta investigación se desarrolla desde la perspectiva exploratoria.

¹ Se excluyen los valores atípicos mayores al percentil 99 (€206 190).

5.4.1 Análisis de Confiabilidad

En el apartado 5.1 Análisis Descriptivo de las variables de estudio, se desglosan los porcentajes de hogares con acceso y uso de los dispositivos TIC, los cuales se categorizaron asignando el valor 1 si el hogar tiene el dispositivo y 0 si carece del mismo.

El análisis de confiabilidad aplicado a las variables de la ENIGH produjo un alfa de Cronbach de 0,575 lo que indica un nivel bajo de confiabilidad. En el en el cuadro 10 se desglosa los valores que toma el alfa al suprimir cada variable, obteniéndose mejoras al excluir la variable de tenencia de radio o equipo de sonido.

Cuadro 10: Cambios en el Alfa de Cronbach al suprimir cada variable de la ENIGH.

	Alfa de Cronbach si el elemento se ha suprimido
Total	0,575
Línea telefónica celular	0,566
Tableta	0,530
Computadora de escritorio	0,542
Computadora portátil	0,487
Radio o equipo de sonido	0,592
Televisor	0,570
Línea telefónica fija	0,572
Internet	0,523
Hacinamiento tecnológico	0,515

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC. [Archivo de datos].

Por su parte, para las variables analizadas de la ENAHO, el alfa de Cronbach resultante es de 0,743. Del cuadro 11, se obtiene que la exclusión de las variables sobre radio o equipo de sonido, televisor, línea telefónica celular y línea telefónica fija, mejorarían la consistencia.

Cuadro 11: Cambios en el Alfa de Cronbach al suprimir cada variable de la ENAHO.

Alfa de Cronbach si el elemento se ha suprimido

Total	0,743
Tableta	0,723
Radio o equipo de sonido	0,756
Línea telefónica celular	0,738
Televisor	0,745
Línea telefónica fija	0,757
Internet	0,718
Computadora portátil	0,730
Computadora de escritorio	0,729
Hacinamiento tecnológico	0,707
Uso de celular	0,721
Uso de computadora de escritorio	0,725
Uso de Internet	0,703
Uso de tableta	0,725
Uso de computadora portátil	0,715

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Estos resultados, sugieren que:

- La exclusión de la variable tenencia de radio o equipo de sonido podría mejorar la consistencia del índice aplicado a la ENIGH; y
- La exclusión de las variables tenencia de radio o equipo de sonido, tenencia de televisor y acceso a línea telefónica fija, mejoraría la consistencia del índice aplicado a la ENAHO.

Por tanto, se procede a la exclusión de dichas variables según corresponda.

5.4.2 Construcción de Dimensiones mediante la aplicación del Análisis de Factorial Exploratorio

Tomando como base la teórica se ejecuta el AFE para determinar las dimensiones considerando las variables de los cuadros 11 y 11 del apartado anterior. Es necesario indicar que para la aplicación del AFE se utilizó la matriz de correlaciones tetracórica, dado que las variables son dicotómicas. Para mejor interpretación, se desglosan los resultados según encuesta.

5.4.2.1 Componente TIC en la ENIGH

La aplicación del AFE a la ENIGH presentó limitantes en la estimación, los cuales se detallan en el cuadro 12.

Cuadro 12: Resultados del AFE aplicado a la ENIGH

Variable	Factor		h2	Peso Factorial		Var
	1	2		1	2	
Celular	0,992	-0,230	1,038	678 237 385	227 008 709	0,033
Hacinamiento	0,732	0,146	0,557	-367 625 121	-123 045 567	0,217
Internet	0,708	0,110	0,513	-323 384 104	-108 237 926	0,189
Computadora portátil	0,684	0,293	0,553	38 487 156	12 881 801	0,234
Tableta	0,612	0,317	0,475	-79 996 954	-26 775 293	0,132
Computadora de escritorio	0,482	0,390	0,385	1 106 387	370 311	0,093
Televisor	0,315	0,384	0,247	-146 485 058	-49 029 122	0,041
Línea telefónica fija	-0,080	1,010	1,026	339 206 066	113 533 599	0,198

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC. [Archivo de datos].

Por lo anterior se decide que:

- a) La variable de acceso de línea celular se excluye porque el valor de la comunalidad es superior a 1 y su varianza es baja.
- b) La variable línea telefónica fija se excluye por el valor de la comunalidad.
- c) Las variables acceso a computadora de escritorio y acceso a televisor, se excluyen porque la variabilidad es menor al 10%.

Con las variables restantes, se obtiene un alfa de Cronbach de 0,594 y del AFE se logra extraer un único factor que explica el 52% de la varianza total con cargas factoriales entre 0,626 y 0,866, siendo estos resultados adecuados para una investigación exploratoria (Cuadro 13).

Cuadro 13: Cargas Factoriales y análisis de fiabilidad para las variables de la ENIGH 2018.

Variable	Carga Factorial	Comunalidad	Peso Factorial
Computadora portátil	0,866	0,751	0,563
Hacinamiento	0,707	0,500	0,194
Tableta	0,646	0,417	0,217
Internet	0,626	0,392	0,126

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

Por tanto, para la ENIGH el componente TIC se construye con la siguiente ecuación:

$$TIC_{Enigh} = 0,563 * Comp_{portátil} + 0,194 * Hacin + 0,217 * Tablet + 0,126 * Internet$$

Para facilitar la interpretación de este componente, se procede a estandarizar el resultado de la ecuación anterior de manera que oscile entre 0 (menor nivel TIC) y 100 (mayor nivel TIC).

El valor medio de componente TIC_{Enigh_z} es de 46,5 según se observa en el cuadro 14.

Cuadro 14: Estadísticas Descriptivas del componente TIC de la ENIGH.

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
TIC_{Enigh_z}	46,5	0,8	44,8	48,2	1,9	2,1	7046

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

5.4.2.2 Componente TIC en la ENAHO

De forma similar, los resultados de AFE aplicado a la ENAHO presentaron problemas de estimación en esta encuesta, como se puede observar en el Cuadro 15. Por lo que se decide que:

- La variable uso del celular dado se excluye dado que el valor de la comunalidad asociada sobrepasa la unidad.
- Las variables tenencia de celular, uso de tableta y uso de computadora de escritorio se excluyen dado que la variabilidad es baja lo que afecta la estimación de la matriz tetracórica.

Cuadro 15: Resultados del AFE aplicado a la ENAHO

Variable	Factor		h2	Peso Factorial		Var
	1	2		1	2	
Uso de celular	1,003	0,080	1,013	37 2975 329	-263 653 733	0,190
Uso de Internet	0,818	0,428	0,853	-178 521 336	126 195 523	0,247
Uso de portátil	0,603	0,586	0,707	100 103 569	-70 762 534	0,142
Hacinamiento	0,653	0,533	0,711	38 462 161	-27 188 639	0,215
Internet	0,448	0,568	0,523	263 157 831	-186 024 486	0,200
Celular	0,621	0,422	0,564	-403 908 770	285 520 370	0,040
Computadora portátil	0,200	0,490	0,281	-31 304 286	22 128 787	0,238
Computadora de escritorio	0,255	0,530	0,345	54 610 184	-38 603 571	0,121
Tableta	0,110	0,837	0,712	187 863 504	-132 799 436	0,152
Uso de tableta	0,377	0,768	0,732	-14 4439 143	102 103 049	0,078
Uso computadora escritorio	0,555	0,569	0,633	-187 598 156	132 611 867	0,084

Fuente: elaboración propia con datos de la Encuesta Nacional de los Hogares 2018, INEC. [Archivo de datos].

Con las variables seleccionadas, se estima nuevamente el alfa de Cronbach, alcanzado el valor de 0,714, considerándolo un valor aceptable.

Se tiene que el análisis sólo extrajo un único factor, cuyas cargas fueron superiores a 0,5; además, 6 de las 7 comunalidades superaron el valor de 0,3. Adicionalmente, este AFE produjo que 51% de la varianza explicada siendo este valor adecuado para esta investigación exploratoria (Cuadro 16).

Cuadro 16: Cargas Factoriales y análisis de fiabilidad para las variables de la ENAHO 2018.

Variable	Carga Factorial	Comunalidad	Peso Factorial
Uso de portátil	0,832	0,693	0,305
Uso de Internet	0,819	0,670	0,120
Hacinamiento	0,799	0,639	0,279
Internet	0,753	0,566	0,247
Tableta	0,649	0,421	0,107
Computadora de escritorio	0,558	0,311	0,070
Computadora portátil	0,546	0,298	0,056

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Con estos hallazgos, se plantea la creación del componente TIC, mediante la siguiente ecuación:

$$TIC_{Enaho} = 0.305 * U_{port} + 0,120 * U_{inter} + 0,279 * Hacin + 0.247 * Inter + 0,107 * Tablet + 0,07 * Comp_{escrit} + 0,056 * Comp_{portátil}$$

Al igual que en la sección anterior, el resultado de la ecuación anterior se estandariza a valores entre 0 (menor nivel TIC) y 100 (mayor nivel TIC) para una mejor interpretación. En este caso, la estimación de la media TIC_{Enaho_z} es 46,7 (Cuadro 17).

Cuadro 17: Estadísticas Descriptivas del componente TIC según encuesta.

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
TIC_{Enaho_z}	46,7	0,4	45,8	47,5	1,0	1,6	10 942

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Tal y como se expuso al inicio de esta sección, el planteamiento teórico de esta investigación intentaba explorar la construcción de las dimensiones de acceso y uso de dispositivos y servicios TIC; sin embargo, los resultados indican que, para el caso de Costa Rica se tiene un único componente que fusiona el acceso y uso a las TIC.

5.5 Construcción del IPD según la ENIGH 2018.

Para la elaboración del IPD_{Enigh} se procede con la combinación de los componentes TIC_{Enaho_z} y Gasto TIC. Dado que estos componentes poseen escalas diferentes, se decide estandarizar la variable Gasto TIC ($Gasto\ TIC_{Enigh_z}$) para que oscile en el rango de 0 a 100, asignado 0 a los hogares con menor gasto y 100 a los hogares con mayor gasto.

Por tanto, el cálculo del índice se obtiene de la suman ambos componentes, asignándole igual peso y su resultado se estandariza en una escala de 0 a 100, donde 0 representa pobreza digital menor y 100 pobreza digital mayor (IPD_{Enigh_z}).

En el cuadro 18 se desglosa las estadísticas descriptivas del índice las cuales consideran el diseño muestral complejo de esta encuesta en su estimación. Obsérvese que la media del IPD_{Enigh_z} para Costa Rica es de 67,3.

La media del IPD_{Enigh_z} para la zona urbana es 64,0 siendo menor al límite inferior del respectivo valor nacional; mientras que, en la zona rural, el valor obtenido de 75,5 es mayor al límite superior del índice nacional. Los resultados indica que existe una diferencia significativa en cuanto al índice de pobreza digital entre ambas regiones.

En cuanto a las regiones de planificación, en la Región Central, se tiene un resultado similar, dado el valor de 63,5 es menor al límite inferior nacional. Para el resto de las regiones, todos los valores superan el límite superior nacional. Se evidencia que existe una diferencia significativa de mayor incidencia de hogares con pobreza digital en las regiones distintas a la Central (Anexo 6).

Cuadro 18: Estadísticas Descriptivas del Índice de Pobreza Digital en los hogares de la ENIGH 2018.

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
<i>IPD_{Enigh_z}</i>	67,3	0,6	66,1	68,4	0,9	2,2	6 482
Zona¹							
Urbana	64,0	0,7	62,7	65,4	1,1	2,2	4 361
Rural	75,5	0,7	74,2	76,9	0,9	1,5	2 121
Región²							
Central	63,5	0,8	61,9	65,2	1,3	2,5	1 991
Chorotega	72,0	1,0	70,0	74,0	1,4	1,2	954
Pacífico Central	71,5	1,1	69,4	73,6	1,5	1,1	815
Brunca	73,9	0,8	72,2	75,5	1,1	0,9	909
Huetar Atlántico	75,5	1,2	73,0	77,9	1,6	1,6	910
Huetar Norte	74,4	1,0	72,4	76,4	1,4	1,2	903

¹ Diferencia significativa, $t=12,9$; $gl=455$; $Sig.=0,00$

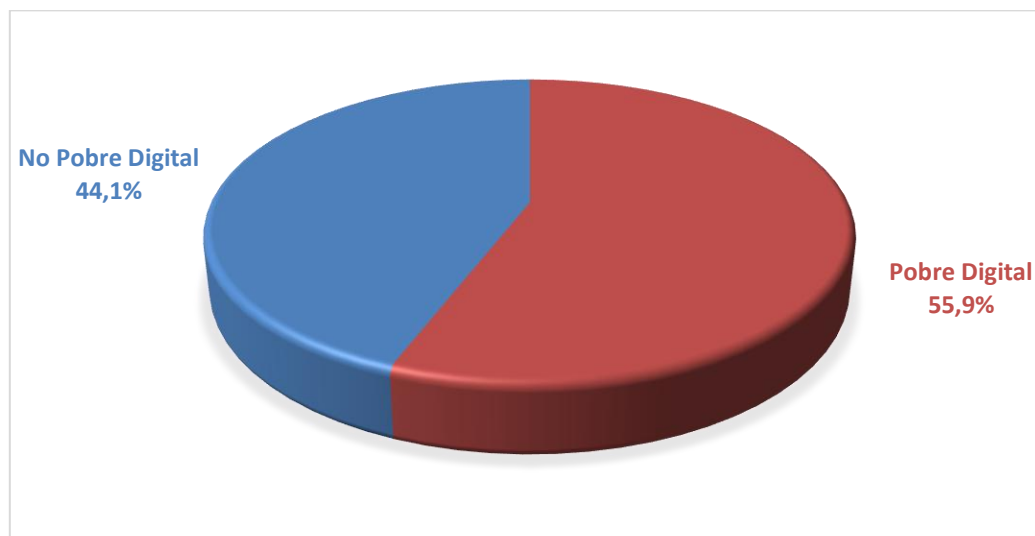
² Test de Wald, $F=27,7$; $gl1=5$, $gl2=451$; $Sig=0,00$

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

Con el propósito de caracterizar de manera más detallada de los hogares de la ENIGH se aplicó el procedimiento de clasificación k-medias para la conformación de conglomerados con los componentes del índice: TIC_{Enigh_z} y $Gasto\ TIC_{Enigh_z}$.

El procedimiento produjo dos conglomerados alcanzando un 99,8% de casos agrupados correctamente según el análisis discriminante aplicado (ver Anexo 7). Bajo el contexto de la investigación, es posible denominar los conglomerados como pobre digital y no pobre digital. En el gráfico 3, se detalla la distribución porcentual de los hogares según nivel de pobreza digital.

Gráfico 1: Distribución porcentual de los hogares según nivel de pobreza digital para la ENIGH 2018



Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

Según el gráfico anterior y los resultados del cuadro 19, los hogares de la ENIGH se distribuyen de la siguiente forma:

- El 55,9% de los hogares se ubican en Pobreza Digital, con un valor promedio en el IPD_{Enigh_z} de 84,2.
- En el conglomerado de No Pobreza Digital se catalogan al 44,1% de los hogares y el valor del IPD_{Enigh_z} se estima en 45,8.

Cuadro 19: Estadísticas Descriptivas del IPD_{Enigh_z} y gasto mensual en TIC de los hogares de la ENIGH según nivel de pobreza digital

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
IPD_{Enigh_z}	67,3	0,6	66,1	68,4	0,9	2,2	6 482
Pobre digital	84,2	0,2	83,8	84,6	0,3	1,5	3 874
No pobre digital	45,8	0,3	45,2	46,5	0,7	1,7	2 608

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

En el cuadro 20 se presenta el desglose de las estadísticas descriptivas de los componentes del IPD_{Enigh_z} según los niveles de pobreza digital.

Cuadro 20: Estadísticas Descriptivas del IPD_{Enigh_z} , TIC_{Enigh_z} y $Gasto TIC_{Enigh_z}$ de los hogares de la ENIGH 2018 según nivel.

		Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
				Inferior	Superior			
Pobre Digital	IPD_{Enigh_z}	84,2	0,2	83,8	84,6	0,3	1,5	3 874
	TIC_{Enigh_z}	20,2	0,3	19,6	20,8	1,5	1,4	
	$Gasto TIC_{Enigh_z}$	11,4	0,2	11,0	11,9	2,0	1,5	
No pobre Digital	IPD_{Enigh_z}	45,8	0,3	45,2	46,5	0,7	1,7	2 608
	TIC_{Enigh_z}	84,8	0,3	84,2	85,5	0,4	1,6	
	$Gasto TIC_{Enigh_z}$	23,5	0,4	22,7	24,4	1,8	1,6	

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

5.6 Construcción del IPD según la ENAHO 2018.

De forma similar a la descrita en el apartado anterior, para el cálculo del índice aplicado a la ENAHO, se procede a combinar los componentes TIC_{Enaho_z} y $GastoTIC_{Enaho}$. Para unificar las escalas de estos componentes, se resuelve estandarizar la variable gasto para que oscile de 0 en los hogares con mínimo gasto y 100 en los hogares con máximo gasto ($GastoTIC_{Enaho_z}$). Finalmente, para el cálculo del IPD_{Enaho} se suman ambos componentes TIC_{Enaho_z} y $Gasto_{Enaho_z}$ y su resultado se convierte a escala de 0 a 100, donde 0 representa mínima pobreza digital y 100 representa máxima pobreza digital (IPD_{Enaho_z}).

Del cuadro 21, se obtiene que el IPD_{Enaho_z} nacional es de 65,9; en la zona urbana es de 62,5 y para la zona rural es de 74,8 mayor al límite superior del valor nacional (66,5), esto indica que entre zonas se evidencia diferencia significativa en el valor del índice.

Por su parte, la media del índice para la región Central es menor a la media nacional; mientras que en el resto de las regiones los valores del índice superan el límite superior de la media nacional y se comprueba la existencia de diferencia significativa de la Región Central con respecto al resto regiones del país, según se muestra en el Anexo 8.

Cuadro 21: Estadísticas Descriptivas del Índice de Pobreza Digital en los hogares de la ENAHO 2018.

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
<i>IPD</i>_{Enaho_z}	65,9	0,3	65,2	66,5	0,5	1,7	10 762
Zona¹							
Urbana	62,5	0,4	61,7	63,2	0,6	1,6	7 394
Rural	74,8	0,4	74,0	75,6	0,5	1,4	3 368
Región²							
Central	62,1	0,4	61,3	62,9	0,7	1,7	6 050
Chorotega	71,3	1,0	69,3	73,2	1,4	1,6	996
Pacífico Central	70,1	0,9	68,4	71,9	1,3	1,3	769
Brunca	71,7	1,0	69,9	73,6	1,3	1,6	994
Huetar Atlántico	71,8	0,9	70,2	73,5	1,2	1,6	1 076
Huetar Norte	74,5	0,8	72,9	76,2	1,1	1,5	877

¹Diferencia significativa, $t=23,2$; $gl=1106$, $p\leq 0,001$

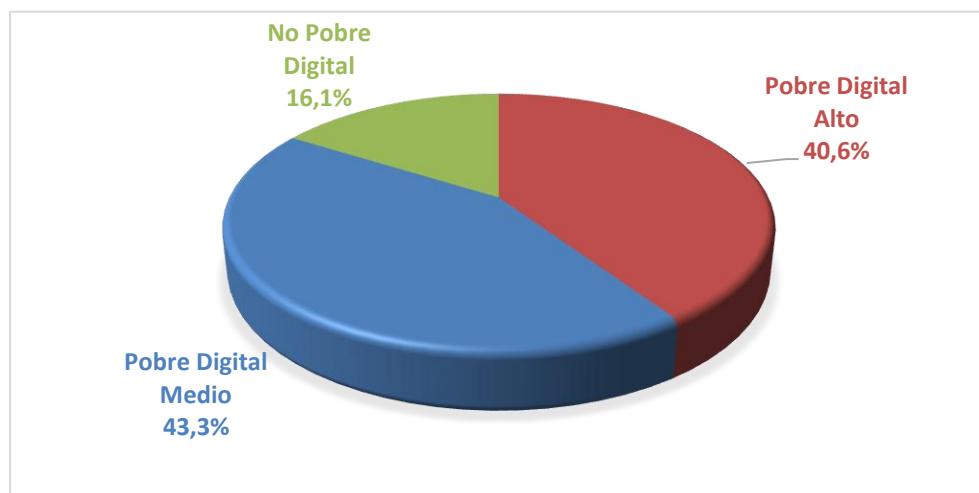
²Test de Wald, $F=71,9$; $gl1=5$; $gl2=1102$, $p<0,001$

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Replicando el procedimiento del apartado anterior, se procede a clasificar los hogares de la ENAHO aplicando a los componentes TIC_{Enaho_z} y $Gasto\ TIC_{Enaho_z}$ el análisis de conglomerados con el procedimiento k-medias.

En este caso, el resultado del procedimiento de conglomerados obtuvo tres conglomerados, con un 98,8% de los casos correctamente agrupados, según el análisis discriminante que se presenta en el Anexo 9.

Gráfico 2: Distribución porcentual de los hogares según nivel de pobreza digital para la ENAHO 2018



Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

De los resultados que se presentan en el cuadro 22 y el gráfico anterior, los hogares de la ENAHO se distribuyen de la siguiente manera:

- En pobreza digital alta se catalogan al 40,6% de los hogares, con un valor medio en el IPD_{Enaho_z} de 84,6.
- El 43,3% de los hogares se consideran en pobreza digital media y su valor medio en el IPD_{Enaho_z} de 61,0.
- El 16,1% de los hogares se ubican en el conglomerado de no pobre digital, poseen un valor medio en el IPD_{Enaho_z} de 31,7.

Cuadro 22: Estadísticas Descriptivas del IPD_{Enaho} y gasto mensual estimado en TIC de los hogares de la ENAHO según nivel de pobreza digital

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
IPD_{Enaho}	65,9	0,3	65,2	66,5	0,5	1,7	10 762
Pobre Digital Alto	84,6	0,1	84,3	84,9	0,2	1,1	4 515
Pobre Digital Medio	61,0	0,1	60,8	61,3	0,2	1,2	4 635
No Pobre Digital	31,7	0,3	31,2	32,3	1,0	1,2	1 612

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

El desglose de las estadísticas descriptivas de los componentes del IPD_{Enaho_z} según nivel de pobreza se encuentran en siguiente cuadro.

Cuadro 23: Estadísticas Descriptivas del IPD_{Enaho_z} , TIC_{Enaho_z} y $Gasto TIC_{Enaho_z}$ de los hogares de la ENAHO 2018 según nivel.

		Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
				Inferior	Superior			
Pobre	IPD_{Enaho_z}	84,6	0,1	84,3	84,9	0,2	1,1	
Digital	TIC_{Enaho_z}	17,6	0,2	17,2	18,0	1,2	1,1	4 515
Alto	$Gasto TIC_{Enaho_z}$	13,1	0,1	12,8	13,4	1,0	1,2	
Pobre	IPD_{Enaho_z}	61,0	0,1	60,8	61,3	0,2	1,2	
Digital	TIC_{Enaho_z}	59,1	0,2	58,8	59,5	0,3	1,1	4 635
Medio	$Gasto TIC_{Enaho_z}$	18,7	0,2	18,4	19,0	0,8	1,2	
No Pobre	IPD_{Enaho_z}	31,7	0,3	31,2	32,3	1,0	1,2	
Digital	TIC_{Enaho_z}	84,4	0,4	83,7	85,0	0,4	1,1	1 612
	$Gasto TIC_{Enaho_z}$	52,0	0,5	50,9	53,0	1,1	1,1	

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC.

[Archivo de datos].

5.7 Comparación de comportamiento del IPD según encuesta.

Considerando los resultados de los apartados anteriores, se realiza una comparación de IPD_{Enigh_z} y IPD_{Enaho_z} , desde varios enfoques. Primero, se analiza el comportamiento de sus componentes. Seguidamente, se comparan, los valores a nivel nacional, así como los desglosados por zona y región planificación Y se finaliza, contrastando los niveles de pobreza digital según encuesta.

5.7.1 Componentes TIC y Gasto en TIC

Para una mejor comprensión de las diferencias entre el IPD_{Enigh_z} y el IPD_{Enaho_z} , se analizan las partes que conforma el componente TIC y el componente Gasto TIC . Este análisis se basa en una exploración descriptiva del comportamiento de estas en cada una de las encuestas.

5.7.1.1 Componente TIC

El componente TIC es el resultado de la aplicación del AFE; sin embargo, por los objetivos de cada encuesta, no se dispone de las mismas variables. Como se muestra en la siguiente ilustración, en el componente TIC de la ENIGH contiene las variables de tenencia de computadora, tableta e internet, así como la variable hacinamiento. Por su parte, el componente TIC en la ENAHO incluye, además de las anteriores, tenencia de computadora de escritorio, uso de portátil y uso de Internet. Es oportuno recordar, que la ENIGH no incluye variables sobre el uso de los dispositivos y servicios TIC .

Tabla 4: Variables resultantes del AFE que conforman el componente TIC según encuesta.

TIC_{Enigh}	TIC_{Enaho}
Computadora portátil	Computadora portátil
Hacinamiento	Hacinamiento
Tableta	Tableta
Internet	Internet
	Uso de portátil
	Uso de Internet
	Computadora de escritorio

El cuadro 24, presenta la prueba de igualdad de medias para los componentes TIC_{Enigh_z} y TIC_{Enaho_z} , obteniéndose que no existe diferencia significativa.

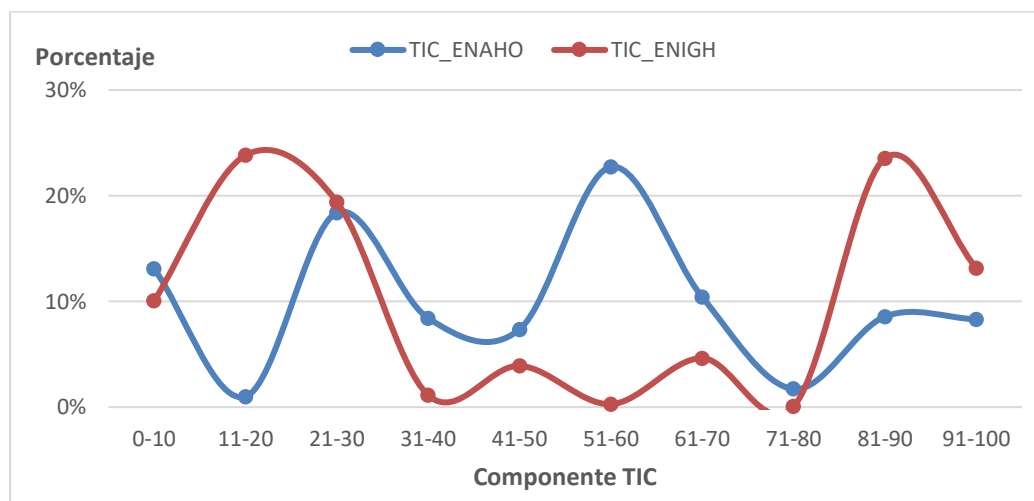
Cuadro 24: Prueba de igualdad de medias para el componente TIC en muestras independientes.

	Media	Error Típico	t	Grados de libertad	Sig.	Intervalo de confianza para la diferencia	
						Inferior	Superior
TIC_{Enigh_z}	46,5	0,8	0,15	1 430	0,88	-1,68	1,96
TIC_{Enaho_z}	46,7	0,4					

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

Sin embargo, para una mejor visualización del comportamiento de este componente, se representa en el gráfico 5, la distribución porcentual de los hogares en diez intervalos equitativos. Se observa que, para el TIC_{Enigh_z} se tiene mayor concentración de hogares en los extremos del rango; específicamente en el intervalo de menores de 30 y mayores a 80. Por su parte, la distribución del TIC_{Enaho_z} es más oscilante.

Gráfico 3: Distribución porcentual de los hogares según intervalos del componente TIC por encuesta.



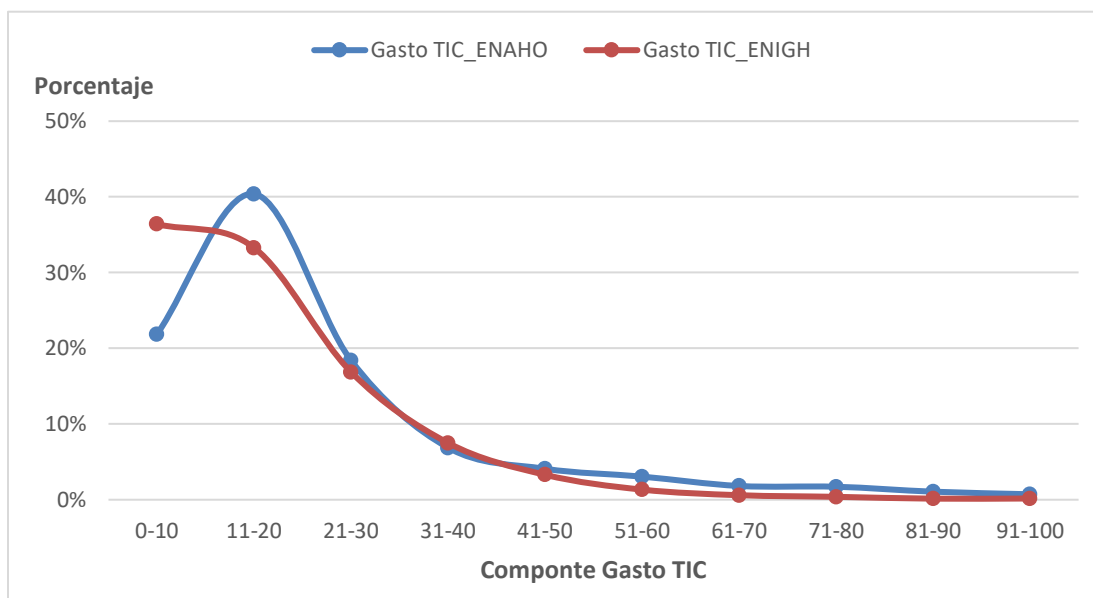
Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

5.7.1.2 Componente Gasto TIC

Para comparar este componente, es necesario recordar que en cada encuesta se estimó de diferente manera. Así, el gasto en TIC de la ENIGH fue construida a partir de la suma de los montos de gastos en dispositivos y servicios TIC según las recomendaciones de la UIT; mientras que el gasto TIC de la ENAHO es el resultado de la aplicación del modelo de regresión.

El gráfico 6 muestra la distribución porcentual de los hogares según valor en el componente Gasto TIC por encuesta. El primer hallazgo que se observa que la forma de ambas curvas es similar, especialmente a partir del tercer rango del componente. También, se tiene que en el primer rango del componente (hogares con los menores gastos en TIC), se obtiene mayor porcentaje hogares de la ENIGH que en la ENAHO, mientras que el rango de 11 a 20 sucede lo contrario.

Gráfico 4: Distribución porcentual de los hogares según intervalos del componente Gasto TIC según encuesta.



Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

Para comprobar si la diferencia en la forma de las curvas del gráfico anterior existe, se realiza la prueba de igualdad de medias. Los resultados indican que existen una diferencia significativa para este componente según encuesta, como se detalla en el Cuadro 25.

Cuadro 25: Prueba de igualdad de medias para el componente Gasto TIC en muestras independientes.

	Media	Error Típico	t	Grados de libertad	Sig.	Intervalo de confianza para la diferencia	
						Inferior	Superior
<i>Gasto TIC_{Enigh_z}</i>	16,8	0,3	18,0	1 430	0,0	5,7	7,1
<i>Gasto TIC_{Enaho_z}</i>	21,8	0,3					

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

5.7.2 IPD_{Enigh_z} versus IPD_{Enaho_z}

Como se detalla en el cuadro 26, la diferencia en el valor del índice de pobreza digital según encuesta y se desglosa por zona y región. Se observa que el IPD a nivel nacional tiene una diferencia significativa siendo el IPD_{Enigh_z} es mayor al IPD_{Enaho_z} .

Por zona de residencia se tiene que existe evidencia de diferencia en entre IPD_{Enigh_z} y IPD_{Enaho_z} para la zona urbana; sin embargo, en la zona rural los resultados demuestran que ambos índices no difieren.

Por región de residencia, se comprueba que sólo en la Huetar Atlántica existe diferencia estadística entre los valores IPD_{Enigh_z} y IPD_{Enaho_z} , para el resto de las regiones las discrepancias observadas resultaron no significativas.

Cuadro 26: Comparación de los valores del IPD_{Enigh} y IPD_{Enaho} nacional, por zona y por región.

	IPD_{Enigh}	IPD_{Enaho}	Dif	t	Grados de libertad	Sig.	Intervalo de confianza	
							Inf.	Sup.
Nacional	67,3	65,9	1,4	-2,2	1 430	0,03	-2,6	-0,1
Zona								
Urbana	64,0	62,5	1,6	-2,0	964	0,04	-3,1	-0,1
Rural	75,5	74,8	0,7	-0,9	465	0,35	-2,3	0,8
Región								
Central	63,5	62,1	1,4	-1,6	679	0,12	-3,2	0,4
Chorotega	72,0	71,3	0,7	-0,5	156	0,61	-3,6	2,1
Pacífico Central	71,5	70,1	1,4	-1,0	126	0,31	-4,2	1,4
Brunca	73,9	71,7	2,1	-1,7	156	0,09	-4,6	0,4
Huetar Atlántico	75,5	71,8	3,6	-2,4	164	0,02	-6,6	-0,7
Huetar Norte	74,4	74,5	-0,1	0,1	144	0,93	-2,5	2,7

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

5.7.3 Niveles de IPD_{Enigh} versus Niveles del IPD_{Enaho}

Se analiza a continuación, una comparación de las estimaciones del IPD, sus componentes y el gasto mensual en TIC para cada uno de los grupos resultantes del análisis de conglomerados aplicado a cada encuesta. Como se indicó anteriormente, los resultados de este análisis en la ENIGH produjeron dos conglomerados: Pobre Digital y No Pobre Digital; mientras que en la ENAHO se hallaron tres conglomerados denominados: Pobre Digital Alto, Pobre Digital Medio y No Pobre Digital. Lo anterior produce una limitación en cuanto a la comparación entre conglomerados alcanzados.

Del desglose que se observa en el cuadro 27, se logra clasificar que tanto en la ENIGH como en la ENAHO se obtuvo un conglomerado que posee características muy similares. Se aprecia que los conglomerados Pobre Digital-ENIGH y Pobre Digital Alto-ENAHO, poseen una media estimada en el índice cercana de 84,2 y 84,6, respectivamente, y comprueba la no existencia de diferencia significativa (ver Anexo 10). Siguiendo con estos conglomerados, se observa que tanto en el componente TIC y como en el componente Gasto TIC las medias obtenidas son las más bajas en comparación con el resto de los conglomerados.

Cuadro 27: Comparación de los conglomerados según IPD y sus componentes según encuesta.

	IPD	Componente	
		TIC	Gasto TIC
ENIGH.			
Pobre Digital	84,2	20,2	11,4
No pobre Digital	45,8	84,8	23,5
ENAHO			
Pobre Digital Alto	84,6	17,6	13,1
Pobre Digital Medio	61,0	59,1	18,7
No pobre Digital	31,7	84,4	52,0

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

5.8 Caracterización sociodemográfica de los hogares de Costa Rica según pobreza digital.

Para la caracterización de los hogares, se procede a separar los análisis por encuesta, dado que se obtuvo diferencias en cuanto a la clasificación en conglomerados. Se analizarán las variables sociodemográficas geográficas como la zona y región de residencia y sociales como condición de actividad, sexo, nivel educativo y grupos de edad del jefe del hogar.

5.8.1 Caracterización de los hogares según IPD_{Enigh}

En el cuadro 28 se describen los principales resultados de los conglomerados obtenidos para el IPD_{Enigh_z} según variables sociodemográficas. Se realizó un análisis de asociación entre las variables, resultando que sólo en la variable sexo del jefe del hogar no se hallaron diferencia significativa, es decir se tiene una composición similar entre los grupos de hogares pobres y no pobres digitales. Para el resto de las variables analizadas, las diferencias sí fueron significativas (ver Anexo 11).

Al analizar la distribución por zona, se tiene que en la zona urbana la distribución de los hogares es relativamente equitativa entre los conglomerados; por el contrario, en la zona rural se observa que una mayor proporción de hogares clasificados como pobre digital (70,0%).

Por región de planificación, en la región Central se tiene que la distribución de hogares por conglomerados es similar; en el resto de las regiones se tiene mayor porcentaje de hogares pobres digitales.

Al analizar la distribución de los hogares según la condición de actividad del jefe, se observa que en aquellos cuya jefatura está ocupada la distribución es relativamente equitativa, 51,7% son pobres digitales y 48,3% no pobre digitales. En los hogares con jefatura desocupada o fuera de la fuerza laboral, la mayoría se clasifica como pobre digital.

Los resultados muestran una notable asociación entre el nivel educativo del jefe y la pobreza digital. A medida que aumenta el nivel educativo, disminuye notoriamente el porcentaje de

hogares clasificados como pobres digitales. Específicamente, entre los hogares con jefatura sin estudios 9 de cada 10 son clasificados como pobre digital. En contraste, entre los hogares cuyo jefe posee educación universitaria, esta proporción se reduce a 2 de cada 10.

Finalmente, se evidencia una relación entre la edad del jefe y la condición de pobreza digital. Se observa que conforme aumenta la edad del jefe del hogar, disminuye la proporción de hogares pobres digitales. En general, los hogares con jefaturas menores a 25 años o mayores a 55 años presentan mayor incidencia de pobreza digital. En los hogares con jefaturas con edades intermedias se tiene porcentajes más similares entre los conglomerados pobre digital y no pobre digital.

En resumen, estos resultados indican que la condición de actividad, el nivel educativo y la edad del jefe, podrían influir en el nivel de pobreza digital del hogar. Además, los hogares de la zona rural y los que se sitúan en las regiones distintas a la Región Central, poseen mayor incidencia de hogares en pobreza digital.

Cuadro 28: Distribución porcentual de los hogares por conglomerados según variables sociodemográficas.

		Conglomerado	
		Pobre Digital	No Pobre Digital
Total		55,9	44,1
Zona	Urbana	50,4	49,6
	Rural	70,0	30,0
Región	Central	49,6	50,4
	Chorotega	63,2	36,8
	Pacífico Central	63,1	36,9
	Brunca	65,0	35,0
	Huetar Atlántico	72,6	27,4
	Huetar Norte	66,8	33,2
Condición de actividad del jefe	Ocupado	51,7	48,3
	Desocupado	74,7	25,3
	Fuera de la fuerza laboral	78,4	21,6
Sexo del jefe	Hombre	55,6	44,4
	Mujer	56,4	43,6
Nivel educativo del jefe	Sin estudios	90,4	9,6
	Primaria	74,5	25,5
	Secundaria	56,9	43,1
	Universitaria	20,4	79,6
Grupos de edad del jefe	Menos de 18	100,0	0,0
	18 a 24	66,2	33,8
	25 a 34	53,0	47,0
	35 a 44	49,6	50,4
	45 a 54	48,2	51,8
	55 o más	66,6	33,4

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

Para finalizar, se observa en el cuadro 29, el gasto medio estimado en TIC que realizan los hogares según conglomerado. Los hogares pobres digital se estima que el gasto es de ¢34 966 mensuales, mientras que en los hogares no pobre digital, este gasto se duplica.

Cuadro 29: Estadísticas Descriptivas del gasto medio en TIC de los hogares de la ENIGH según conglomerado

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
Gasto TIC Total	49 183,2	900,5	47 413,7	50 952,8	1,8	2,0	6 482
Pobre Digital	34 966,4	613,1	33 761,7	36 171,1	1,8	1,5	3 874
No pobre digital	67 194,6	1 155,6	64 923,9	69 465,3	1,7	1,6	2 608

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

5.8.2 Caracterización de los hogares según IPD_{Enaho}

La descripción de las variables sociodemográfica de los conglomerados del $IPD_{Enaho-z}$ se presentan en el cuadro 30. En el Anexo 12 se detallan las pruebas de asociación entre estas variables y los tres conglomerados, hallándose que existe diferencias estadísticas en todas las variables analizadas.

Al analizar por zona de residencia, se observa que el 45,4% de los hogares de la zona urbana son clasificados como pobre digital medio, el 34,4% como pobre digital alto y el 20,3% se catalogan como no pobre digital. En la zona rural, la mayoría de los hogares (56,9%) se ubican en el conglomerado pobre digital alto y se destaca que sólo el 5,2% de los hogares de esta zona es no pobre digital.

Al analizar por región de planificación, los resultados reflejan que en la región Central la mayoría de los hogares se ubica en el conglomerado de pobre digital medio (45,2%) y posee la proporción más alta de hogares no pobres digitales (21,1%). En el contraste, en el resto de

las regiones, predominan los hogares catalogados como pobres digitales con porcentajes entre el 46,1% en la Pacífico Central y el 58,9% en la Huetar Norte.

Por condición de actividad del jefe, se observa que los hogares con jefatura ocupada el mayor porcentaje de estos se ubica en pobre digital medio (45,2%). Entre los hogares con jefatura desocupada, la distribución es relativamente equitativa entre los conglomerados hogar pobre digital alto y medio, pero sólo un 2,3% de estos hogares son no pobres digitales. Mientras que en los hogares con jefes fuera de la fuerza laboral, la mayor proporción de estos hogares se clasifican como pobres digital alto y el 4,1% son no pobres digital.

Los resultados por sexo del jefe del hogar muestran diferencias moderadas pero consistentes. Los hogares con jefatura mujer, el 41,3% son pobre digital alto, el 44,6% pobre digital medio y el 14,1% no pobre. Por su parte, los hogares con jefes hombre, el 40,0% son pobre digital alto, 42,5% pobre digital medio y 17,5% no pobre digital.

El nivel educativo del jefe muestra una asociación notable con el nivel de pobreza digital. Los resultados indican que conforme aumenta el nivel educativo del jefe disminuye el porcentaje de hogares con pobreza digital alta y, de forma similar, aumenta la proporción de hogares no pobres digital.

Finalmente, la distribución de los hogares según edad del jefe del hogar y el nivel de pobreza digital es heterogénea. En los grupos de edad hasta los 54 años, la mayoría de los hogares son pobre digital medio. En los hogares con jefes menores de 18 años no se ubicaron hogares en la categoría no pobre digital. Los hogares con jefaturas menores de 18 años o mayores de 55 años concentran mayores niveles de pobreza digital alta, con un 48,6% y 49,6%, respectivamente. Los hogares liderados por personas de 35 a 44 años, se ubica el mayor porcentaje de hogares no pobres digitales.

Cuadro 30: Distribución porcentual de los hogares por conglomerados según variables sociodemográficas de la ENAHO.

		Conglomerado		
		Pobre Digital	Pobre Digital	No Pobre
		Alto	Medio	Digital
Total		40,6	43,3	16,1
Zona	Urbana	34,3	45,4	20,3
	Rural	56,9	37,9	5,2
Región	Central	33,7	45,2	21,1
	Chorotega	51,9	38,1	10,0
	Pacífico Central	46,1	44,6	9,3
	Brunca	49,8	41,6	8,7
	Huetar Atlántico	50,4	42,6	7,0
	Huetar Norte	58,9	35,2	5,9
Condición de actividad del jefe	Ocupado	33,8	43,9	22,3
	Desocupado	47,9	49,7	2,3
	Fuera de la fuerza laboral	54,1	41,7	4,1
Sexo del jefe	Hombre	40,0	42,5	17,5
	Mujer	41,3	44,6	14,1
Nivel educativo del jefe	Sin estudios	78,6	21,4	0,0
	Primaria	56,4	42,3	1,3
	Secundaria	32,6	55,5	11,9
	Universitaria	9,2	31,4	59,4
Grupos de edad del jefe	Menos de 18	48,6	51,4	0,0
	18 a 24	33,3	54,9	11,8
	25 a 34	36,4	43,3	20,3
	35 a 44	33,8	42,7	23,5
	45 a 54	32,2	49,0	18,8
	55 o más	49,6	40,2	10,2

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Finalmente, las estadísticas descriptivas del gasto medio de los hogares en TIC se detallan en el cuadro 31; teniéndose que el menor gasto se estima para el conglomerado de pobre digital alto y el gasto mayor en el conglomerado no pobre digital.

Cuadro 31: Estadísticas Descriptivas del gasto medio en TIC de los hogares de la ENAHO según conglomerado

	Media	Error típico	Intervalo de confianza		Coef. de variación	Raíz del efecto del diseño	Muestra
			Inferior	Superior			
Gasto TIC Total	49 576,3	565,7	48 466,2	50 686,3	1,1	1,7	10 762
Pobre Digital Alto	32 196,5	272,7	31 661,5	32 731,6	0,8	1,2	4 515
Pobre Digital Medio	43 433,2	305,2	42 834,3	44 032,2	0,7	1,2	4 635
No pobre digital	109 954,3	1 092,1	107 811,6	112 097,1	1,0	1,1	1 612

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

CAPÍTULO VI. CONCLUSIONES

Al analizar los resultados de la propuesta para medición de la pobreza digital se alcanza hallazgos relevantes. Se logra ajustar un modelo de regresión múltiple utilizando la ENIGH 2018 para la estimación del gasto promedio mensual de los hogares en dispositivos y servicios TIC. Este modelo utilizar variables del hogar y del jefe de hogar que también están incluidas en la ENAHO, lo que permite realizar estimación del gasto TIC en las ediciones después del 2018.

El modelo desarrollado incluyó dos ecuaciones según condición de actividad del jefe del hogar, las cuales mostraron una adecuada capacidad predictiva, lo que permite la estimación del gasto TIC en las ediciones posteriores de la ENAHO y el monitoreo de su evolución.

Los análisis de confiabilidad aplicados a las variables analizadas de ambas encuestas, produjo niveles distintos de confiabilidad. En la ENIGH se logra un alfa de Cronbach aceptable para estudios exploratorios, mientras que en la ENAHO el valor es adecuado. Los resultados de este análisis sugieren la exclusión de las variables de tenencia de telefonía fija, televisor y radio o equipo de sonido para la mejora de la consistencia interna.

El AFE aplicado a las variables de la ENIGH extrajo un factor con las variables de acceso a computadora portátil, tableta e Internet, además de la variable hacinamiento. Tanto el porcentaje de varianza explicada como el alfa de Cronbach del factor extraído, se considera apropiado para la investigación exploratoria. A este factor estandarizado se denomina el componente TIC_{Enigh_z} .

Y de forma similar, la aplicación del AFE en la ENAHO plantea la creación de un factor con las variables hacinamiento, uso de computadora portátil e Internet, acceso a computadora portátil, computadora de escritorio, tableta e Internet. El valor del alfa de Cronbach para este factor resultó aceptable y el porcentaje de varianza explicada del factor es apropiado para la investigación exploratoria. Este factor estandarizado constituye el componente TIC_{Enaho_z} .

Aunque el planteamiento teórico de esta investigación pretendía la construcción de dos dimensiones: una relacionada al acceso de dispositivos y otra al uso de los dispositivos y servicios TIC, los resultados indican que, en el caso de Costa Rica, se tiene un único

componente que concentra estas dimensiones. Por tanto, se puede interpretar que los componentes TIC_{Enigh_z} y TIC_{Enaho_z} , representan una aproximación del nivel TIC de los hogares costarricenses.

Al examinar la estimación de la pobreza digital mediante el IPD_{Enigh_z} se extrae un valor promedio relativamente alto y se evidencian diferencias a nivel geográfico tanto a nivel de zona como de región de planificación. El análisis de conglomerados propuso la segmentación de los hogares en dos grupos: pobre digital y no pobre digital.

Con el IPD_{Enaho_z} se alcanza también un valor alto de la pobreza digital en el país, con diferencias marcadas entre los hogares de la zona urbana y rural y entre los hogares de la región Central y del resto de regiones. Para este índice, el análisis de conglomerados clasificó adecuadamente a los hogares en tres grupos denominados: pobre digital alto, pobre digital medio y no pobre digital.

Al comparar los componentes que conforman la pobreza digital, se observan descubrimientos importantes. El primero indica que a pesar de las discrepancias en cuanto a las variables que constituyen el TIC_{Enigh_z} y el TIC_{Enaho_z} , las medias estimadas a nivel nacional no muestran diferencias estadísticas. El segundo hallazgo se relaciona con el valor medio $Gasto\ TIC_{Enigh_z}$ y $Gasto\ TIC_{Enaho_z}$ que resultan estadísticamente diferentes.

A nivel nacional, la estimación de la pobreza digital con el IPD_{Enigh_z} difiere significativamente con la estimación obtenida con el IPD_{Enaho_z} , obteniéndose un valor medio mayor al calculado a partir de la ENIGH. Al examinar los promedios del IPD_{Enigh_z} y IPD_{Enaho_z} por zona, no se encuentran diferencias significativas en la zona rural; y sólo en la región Huetar Atlántica se detecta discrepancia estadística entre estos índices.

La comparación de los análisis de conglomerados entre encuestas es limitada, por la diferencia en la cantidad de clústeres extraídos. Sin embargo, se destaca la alta similitud observada entre los grupos Pobre Digital-ENIGH y el Pobre Digital Alto-ENAHO cuyos valores medios del IPD no difieren estadísticamente.

Se concluye que los resultados de la caracterización de los hogares según IPD_{Enigh_z} y IPD_{Enaho_z} produjo resultados consistentes. Se verifica la pobreza digital de los hogares guarda estrecha vinculación con las variables geográficas de zona y la región, así como la condición de ocupación, nivel educativo y edad del jefe del hogar.

Adicionalmente, los resultados de esta investigación aplicada arrojan aspectos valiosos a nivel metodológico, a saber:

- En la medición de la pobreza digital propuesta con estas encuestas, se logra mantener una coherencia metodológica y conceptual entre los IPD contruidos ya que los componentes TIC y Gasto TIC muestran comportamientos similares en la distribución de los hogares, especialmente en los grupos con mayor pobreza digital.
- El AFE aplicado para la construcción del componente TIC produjo resultados estables, a pesar de la discrepancia en las variables utilizadas.
- La diferencia en el componente Gasto TIC se explica por la diferencia metodológica en su estimación: en la ENIGH se construyó con base en montos reportados de gasto, mientras que en la ENAHO se estimó mediante un modelo de regresión.
- El modelo de regresión para la estimación del gasto mensual en TIC de los hogares es adecuado y permite su aplicación a las encuestas más reciente.

Finalmente, se sugiere al MICITT considerar la posibilidad de incorporar el IPD_{Enaho_z} como un indicador de referencia para el monitoreo de la pobreza digital, alineándolo con las herramientas de planificación y con los instrumentos del Plan Nacional de Desarrollo de las Telecomunicaciones. Esto permitiría al país contar con un mecanismo consistente, replicable y sensible a las desigualdades territoriales y sociodemográficas que afectan la transformación digital.

CAPÍTULO VII. BIBLIOGRAFÍA

- Alkire, S., Conconi, A., & Seth, S. (2014). *Multidimensional Poverty Index 2014: Brief methodological note and results*. Oxford Poverty and Human Development Initiative.
- Alva de la Selva, A. R. (2015). Los nuevos rostros de la desigualdad en el siglo xxi: la brecha digital. *Revista Mexicana de Ciencias Políticas y Sociales*, LX(223), 265-286.
Obtenido de <https://www.revistas.unam.mx/index.php/rmcpys/article/view/45387/40864>
- Banco Mundial. (2016). *Informe sobre el Desarrollo Mundial 2016: Dividendos Digitales*. Washington DC.
- Barja, G., & Gigler, B.-S. (2006). Qué es y cómo medir la pobreza de información y comunicación en el contexto Latinoamericano. (D. IDRC, Ed.) *POBREZA DIGITAL: Las perspectivas de América Latina y El Caribe*, 1-19.
- Barrantes, R. (2007). Analysis of ICT Demand: What is Digital Poverty and How to measure it? En *Digital Poverty: Latin American and Caribbean Perspective* (págs. 29-54). International Development Research Centre.
- Belloch, C. (2012). *Las Tecnologías de la Información y Comunicación en el aprendizaje*. Obtenido de Universidad de Valencia: <https://www.uv.es/bellohc/pedagogia/EVA1.pdf>
- Camarero Orive, A., Cerbán Jiménez, M., Turias Domínguez, I., González Cancelas, N., & Camarero Orive, A. (2016). Metodología para la clasificación de los puertos mediante indicadores de explotación utilizando análisis de conglomerados. *INGE CUC*, 41-49. doi:10.17981/ingecuc.12.2.2016.04
- Cancino, N., González, C., Gallardo, I., & Estrada, C. (2016). Evaluación de un Modelo de Calidad de Vida Construido desde los Datos. *Acta Colombiana de Psicología*, 297-309. doi:10.14718/ACP.2016.19.1.13
- Coria, S. R., Pérez Meza, M., Mendoza Cortés, E., & Martínez Peláez, R. (2011). Brecha Digital y Pobreza Digital en el Estado de Oaxaca. *Conciencia Tecnológica*, 19-25. Obtenido de <https://www.redalyc.org/articulo.oa?id=94421442004>
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale Development: Theory and Applications*. North Caroline.

- Dillon, W., & Goldstein, M. (1984). *Multivariate Analysis: Methods and Applications*. United States: Jhon Wiley & Sons.
- Dominguez Lara, J. A. (2014). ¿Matrices Policóricas/Tetracóricas o Matrices Pearson? Un estudio metodológico. *Revista Argentina de Ciencias del Comportamiento*, 6(1), 39-48. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=4800004>
- Fagnola, B., & Moneta, A. (2021). Identificación de la Pobreza Multidimensional en Argentina con Métodos Robustos de Análisis Factorial. *Cuadernos del CIMBAGE*(23), 1-20.
- Ferrando, P., & Anguiano-Carrasco, C. (2010). El Análisis Factorial como técnica de Investigación en Psicología. *Papeles del Psicólogo*, 31(1), 18-33.
- Gómez Meléndez, A. (2014). El uso de las nuevas tecnologías y del tiempo como factores determinantes de la formación y educación de los y las jóvenes en Costa Rica. *Revista del Consejo de la Persona Joven*, 58-77. Obtenido de <https://cpj.go.cr/>
- Instituto Nacional de Estadísticas y Censos. (2019). *Encuesta Nacional de Ingresos y Gastos de los Hogares ENIGH, 2018: Metodología*. San José.
- Instituto Nacional de Estadísticas y Censos. (2022). *Encuesta Nacional de Hogares*. Obtenido de <https://inec.cr/>
- Instituto Nacional de Estadísticas y Censos. (2022). *Encuesta Nacional de Ingresos y Gastos de los Hogares*. Obtenido de <https://inec.cr/>
- Jain, A., Murty, M., & Flynn, P. (1999). Data clustering: a review. *Association for Computing Machinery Journals*, 31, 264-323. doi:10.1145/331499.331504
- Jiménez Alfaro, K., & Montero Rojas, E. (2013). Aplicación del modelo de Rasch, en el análisis psicométrico de una prueba de diagnóstico en matemática. *Revista digital Matemática, Educación e Internet*, 1-24. Obtenido de www.tec-digital.itcr.ac.cr
- Johnson, D. (1998). *Métodos Multivariados Aplicados al Análisis de Datos*. International Thomson.
- Maddala, G. (1996). *Introducción a la Econometría*. Prentice-Hall, .
- Madrigal, J. (2000). Estratificación de hogares y segmentos por niveles de ingreso en el censo 2000. *Simposio Costa Rica a la Luz del Censo 2000*, (pág. 34). San José.
- Masuda, Y. (1984). *La sociedad informatizada como sociedad post-industrial* (Vol. 7). Madrid: Tecnos.

- Ministerio de Ciencia, Tecnología y Telecomunicaciones. (2019). *IBD Índice de Brecha Digital 2016-2018*. Obtenido de https://www.micitt.go.cr/wp-content/uploads/2022/04/indice_de_brecha_digital_2016-2018_0.pdf
- Monge González, R., & Rodríguez Álvarez, J. A. (Abril de 2011). Caracterización de las PYMES en Costa Rica mediante el empleo del Análisis de Conglomerados o Cluster. *Tec Empresarial*, 5(1), 8-19. Obtenido de <https://dialnet.unirioja.es/servlet/oaiart?codigo=5761526>
- Nunnally, J., & Bernstein, I. (1995). *Teoría Psicométrica*. Estados Unidos: McGraw-Hill.
- Okoro, C., Musonda, I., & Agumba, J. (2019). An exploratory factor analysis of transportation project sustainability indicators: A case of projects in South Africa. *CIB World Building Congress 2019*. Hong Kong.
- Otero, J. (22 de Agosto de 2018). *El Economista*. Recuperado el 2022, de <https://www.economista.com.mx/opinion/Pobreza-digital-20180823-0006.html>
- Posit Software. (2024). *RStudio: Integrated development environment for R (Version 2024.12.1+563)*. Obtenido de <https://posit.co/products/open-source/rstudio/>
- Reusser Monsálvez, C. (2003). ¿Qué es la sociedad de la información? *Revista Chilena De Derecho Informático*(2). doi:<https://doi.org/10.5354/rchdi.v0i2.10650>
- Roman Vega, I. (9 de marzo de 2021). *Programa Estado de la Nación*. Obtenido de <https://estadonacion.or.cr/educacion-remota-y-hacinamiento-tecnologico-el-celular-esta-salvando-la-clase/>
- Schuschny, A., & Soto, H. (2009). *Guía metodológica - Diseño de indicadores compuestos de desarrollo sostenible*. Santiago: Publicación de la Naciones Unidas.
- Shen, Q., Lu, J., Zhang, S., You, Z., Ren, Y., & Shen, X. (2022). The Bio-Geographical Regions Division of Global Terrestrial Animal by Multivariate Similarity Clustering Analysis Method. *Open Journal of Ecology*, 12(3). doi:10.4236/oje.2022.123014
- Superintendencia de Telecomunicaciones. (2023). *Estadísticas del Sector de Telecomunicaciones* (1er ed.). San José, Costa Rica.
- Terrones, C., & Vargas, P. (2013). Clasificación de la banca comercial peruana: un análisis de clúster jerárquico. *SBS Documentos de Trabajo*.
- Troncoso-González, A., Rodríguez, A., & Caro, A. (2022). valuación de la calidad de datos abiertos en el contexto municipal: Una revisión sistemática de la literatura. *Ingeniare*.

Revista chilena de ingeniería, 255-265. doi:<https://dx.doi.org/10.4067/S0718-33052022000200255>

Unión Internacional de Telecomunicaciones. (2005). Documentos Finales. *Cumbre Mundial sobre la Sociedad de la Información* (págs. 14-23). Tuñez: UIT. Recuperado el 21 de Junio de 2023, de <https://www.itu.int/net/wsis/outcome/booklet-es.pdf>

Unión Internacional de Telecomunicaciones. (2018). *Informe sobre Medición de la Sociedad de la Información*. Ginebra: UIT. Obtenido de <https://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2018/MISR2018-ES-PDF-S.pdf>

Unión Internacional de Telecomunicaciones. (2020). Manual para la medición del acceso y el uso de las TIC en los hogares y por las personas. Recuperado el 2022, de <https://www.itu.int/pub/D-IND-ITCMEAS-2020/es>

Vicente Villardón, J. (2007). *Introducción al Análisis de Cluster*. Obtenido de Universidad de Salamanca: <https://acortar.link/VBMe7F>

ANEXOS

Anexo 1: Promedio del ingreso corriente bruto del jefe del hogar por rama de actividad y por ocupación

Rama de actividad	Ocupación Principal									Total
	Dir. y Gerentes	Prof. científicos e intelectuales	Técnicos y prof. de nivel medio	Personal de Apoyo Adm.	Trab. de los serv. y vendedores de comercios y mercado	Agric. y trabajadores calificados agropecuarios, forestales y pesqueros	Ofi., operarios y artesanos de artes mecánicas y de otros oficios	Operad. Instalac. y máq. y ensamb.	Ocupac. elementales	
Agricultura, ganadería, caza	1 329 720	1 572 515	764 094	439 679	323 658	383 374	429 971	377 777	304 325	658 346
Minas y Canteras			1 421 496	1 394 792			567 591	318 627	208 297	782161
Industrias Manufactureras	3 569 585	1 858 346	856 463	430 308	464 524	476 313	395 637	429 523	366 176	982986
Suministro Electricidad, gas y vapor	2 740 668	1 789 429	948 438	699 027	518 684		746 686	995 636	444 472	1 110 380
Suministro de agua, gestión de desechos		649 050	849 030	452 259	339 690		415 610	365 459	298 913	481 430
Construcción	1 854 576	1 628 717	445 199	269 502	680 717		394 803	477 203	304 902	756 952
Comercio	1 810 411	1 179 934	840 797	546 354	532 497	228 816	447 180	456 511	279 353	702 428
Transporte y Almacenamiento	2 012 354	2 320 400	1 044 742	498 392	583 371		435 583	501 661	290 045	960 819
Alojamiento y servicio de Comidas	2 094 825	755 668	730 059	574 848	439 069	241 916	267 171	453 286	339 765	655 178
Información y comunicación	3 950 975	1 369 331	792 104	746 762	356 959		387 897	295 834		1 128 552
Actividades financieras y seguros	2 640 437	1 297 034	1 043 995	903 930	808 392				369 256	1 177 174
Actividades Inmobiliarias	4 090 639	1 003 598	1 134 669	896 195	357 084		734 801		457 713	1 239 243
Act Profesionales, científicas y técnicas	2 483 564	1 387 351	634 162	754 219	516 347		525 481	598 821		985 706

Rama de actividad	Ocupación Principal									Total
	Dir. y Gerentes	Prof. científicos e intelectuales	Técnicos y prof. de nivel medio	Personal de Apoyo Adm.	Trab. de los serv. y vendedores de comercios y mercado	Agric. y trabajadores calificados agropecuarios, forestales y pesqueros	Ofi., operarios y artesanos de artes mecánicas y de otros oficios	Operad. Instalac. y máq. y ensamb.	Ocupac. elementales	
Servicios administrativos y de apoyo	1 539 181	1 594 847	854 560	605 925	441 272	225 464	452 797	584 577	251 634	727 806
Administración pública y planes de seguridad social	2 202 061	1 865 367	923 129	690 548	776 066		690 756	542 665	452 467	1 017 882
Enseñanza	1 726 475	1 280 495	855 780	629 848	429 969		767 585	835 432	430 393	869 497
Salud y Asistencia Social	3 804 185	2 403 949	998 688	657 031	479 230		895 058	700 766	739 097	1 334 750
Artísticas, entretenimiento	1 404 168	327 606	670 714	1 350 464	435 912		281 450	390 441	367 824	653 572
Otras de Servicios	1 537 521	1 139 813	713 649	472 441	365 411		260 270	259 293	182 147	616 318
Hogares como empleadores					216 891			178 575	215 910	203 792
Organizaciones extraterritoriales		1 568 369					343 499			955 934
Total	2 399 491	1 420 622	869 567	684 870	477 144	311 176	496 833	498 015	347 291	865 645

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

Anexo 2: Resumen del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición ocupada (Submuestra 1).

Resumen del modelo ^{b,c}										
Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Estadísticos de cambio					Durbin-Watson
					Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. Cambio en F	
1	0,653 ^a	0,427	0,427	0,1322037	0,427	87051,69 2	5	584032	0,000	0,010

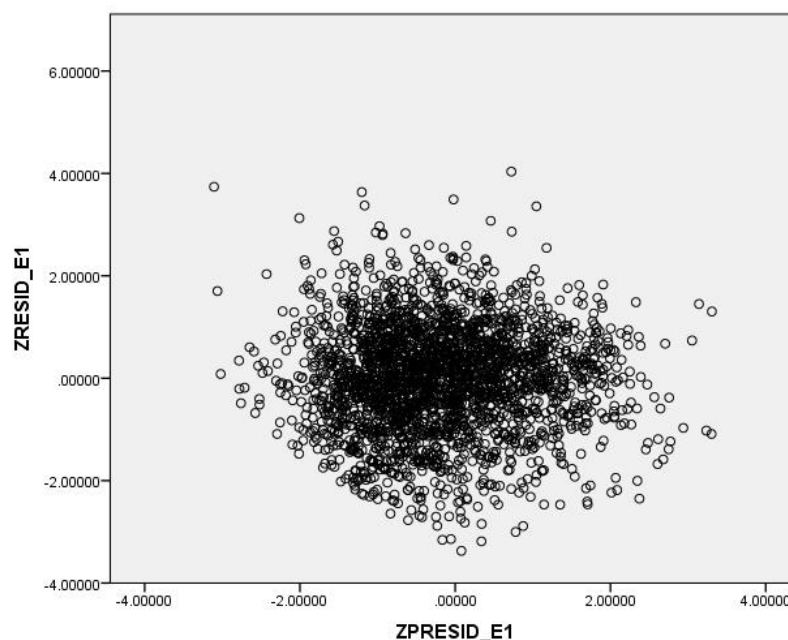
a. Predictores: (Constante), Zona (*Zon*), Miembros Ocupados (*Nocup*), Índice de Ocupación (*IOcup*), LN Ing. total bruto del hogar (*LNIng*), Escolaridad (*Escol*)

b. Variable dependiente: LN_Gasto_y

c. Regresión de mínimos cuadrados ponderada - Ponderada por ponderacMCP

Modelo		Estadísticas de colinealidad	
		Tolerancia	VIF
1	LN Ing. total bruto del hogar (<i>LNIng</i>)	0,791	1,265
	Escol. del jefe (<i>Escol</i>)	0,498	2,009
	Índ. Ocupación (<i>IOcup</i>)	0,483	2,069
	Zona ¹ (<i>Zon</i>)	0,466	2,144
	Miembros ocupados (<i>Nocup</i>)	0,906	1,103

Gráfico de Dispersión de los valores predichos y los residuos del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición ocupada.



Anexo 3: Resumen del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición no ocupada (Submuestra 1).

Resumen del modelo^{b,c}

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Estadísticos de cambio					Durbin-Watson
					Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. Cambio en F	
1	0,605 ^a	0,366	0,366	0,1285291	0,366	17856,494	3	92893	0,000	,010

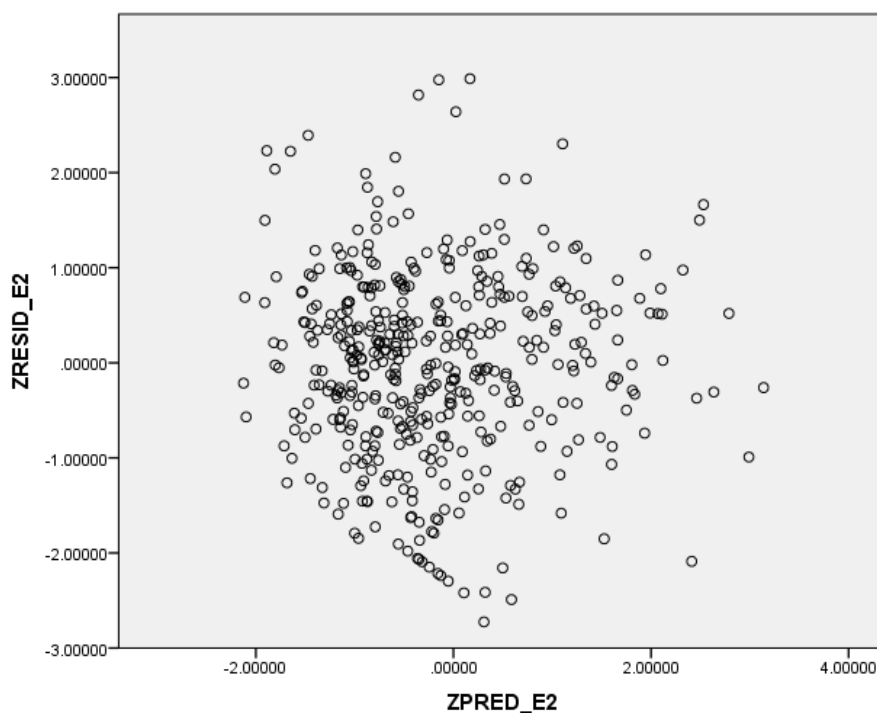
a. Predictores: (Constante), Zona (*Zon*), LN Ing. total bruto del hogar (*LNIng*), Escolaridad (*Escol*)

b. Variable dependiente: LN_Gasto_y

c. Regresión de mínimos cuadrados ponderada - Ponderada por ponderacMCP

Modelo		Estadísticas de colinealidad	
		Tolerancia	VIF
1	LN Ing. total bruto del hogar (<i>LNIng</i>)	0,621	1,609
	Escol. del jefe (<i>Escol</i>)	0,659	1,517
	Zona ¹ (<i>Zon</i>)	0,878	1,139

Gráfico de Dispersión de los valores predichos y los residuos del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición no ocupada.



Anexo 4: Resultados del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición ocupada (Submuestra 2).

Variable	Coef. Beta	Error Estándar	Coef. Beta estandarizado	t	Sig.	IC para Beta (95%)	
						Inferior	Superior
Constante	4,807	0,014	--	332,402	0,000	4,779	4,835
Miembros ocupados (<i>Nocup</i>)	0,102	0,001	0,131	115,312	0,000	0,100	0,103
LN Ing. total bruto del hogar (<i>LNIng</i>)	0,401	0,001	0,490	341,409	0,000	0,399	0,403
Escol. del jefe (<i>Escol</i>)	0,007	0,000	0,078	55,513	0,000	0,007	0,007
Índ. Ocupación (<i>IOcup</i>)	0,023	0,001	0,057	39,814	0,000	0,022	0,024
Zona ¹ (<i>Zon</i>)	0,122	0,002	0,084	81,209	0,000	0,120	0,125

¹ Urbana =1, Rural =0.

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

Resumen del modelo^{b,c}

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Estadísticos de cambio					Durbin-Watson
					Cambio en R cuadrado	Cambio en F	g1	g2	Sig. Cambio en F	
1	0,655 ^a	0,429	0,429	,13359885	0,429	86451,938	5	576041	0,000	,010

a. Predictores: (Constante), Zona (*Zon*), Miembros Ocupados (*Nocup*), Índice de Ocupación (*IOcup*), LN Ing. total bruto del hogar (*LNIng*), Escolaridad (*Escol*)

b. Variable dependiente: LN_Gasto_y

c. Regresión de mínimos cuadrados ponderada - Ponderada por ponderacMCP

Anexo 5: Resultados del modelo de regresión múltiple para el LN (Gasto TIC) del hogar con jefatura en condición no ocupada (Submuestra 2).

Variable	Coef. Beta	Error Estándar	Coef. Beta estandarizado	t	Sig.	IC para Beta (95%)	
						Inferior	Superior
Constante	5,171	0,026		195,990	0,000	5,119	5,223
LN Ing. total bruto del hogar (<i>LNIng</i>)	0,377	0,002	0,511	176,345	0,000	0,372	0,381
Escol. del jefe (<i>Escol</i>)	0,040	0,001	0,117	40,428	0,000	0,038	0,042
Zona ¹ (<i>Zon</i>)	0,116	0,003	0,085	33,692	0,000	0,110	0,123

¹ Urbana =1, Rural =0.

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares 2018, INEC (50% de la muestra). [Archivo de datos].

Resumen del modelo^{b,c}

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación	Estadísticos de cambio					Durbin-Watson
					Cambio en R cuadrado	Cambio en F	gl1	gl2	Sig. Cambio en F	
1	0,605 ^a	0,366	0,366	0,1285291	0,366	17856,494	3	92893	0,000	,010

a. Predictores: (Constante), Zona (*Zon*), LN Ing. total bruto del hogar (*LNIng*), Escolaridad (*Escol*)

b. Variable dependiente: LN_Gasto_y

c. Regresión de mínimos cuadrados ponderada - Ponderada por ponderacMCP

Anexo 6: Prueba de contraste de medias del IPD_{Enigh_z} según región.

Contraste	Dif	Error Estándar	Grados de libertad	t	Sig.
Central - Chorotega	-8,46	1,23	451	-6,89	0,00
Central – Pacífico Central	-8,01	1,21	451	-6,61	0,00
Central – Brunca	10,33	1,03	451	10,15	0,00
Central – Huetar Atlántico	-11,92	1,35	451	-8,83	0,00
Central – Huetar Norte	-10,89	1,22	451	-8,94	0,00
Chorotega – Pacífico Central	0,45	1,32	451	0,34	1,00
Chorotega – Brunca	1,86	1,15	451	1,61	0,59
Chorotega – Huetar Atlántico	-3,46	1,45	451	-2,39	0,16
Chorotega – Huetar Norte	-2,42	1,33	451	-1,83	0,45
Pacífico Central – Brunca	2,31	1,14	451	2,07	0,32
Pacífico Central – Huetar Atlántico	3,91	1,43	451	2,72	0,07
Pacífico Central – Huetar Norte	2,87	1,31	451	2,19	0,24
Brunca – Huetar Atlántico	-1,59	1,28	451	-1,24	0,82
Brunca – Huetar Norte	-0,56	1,14	451	-0,49	1,00
Huetar Atlántico – Huetar Norte	1,03	1,44	451	0,72	0,98

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC.
[Archivo de datos].

Anexo 7: Resultados del análisis de conglomerados aplicado a los componentes del

$$IPD_{Enigh_z}$$

Centro de los conglomerados finales para ENIGH 2018.

Componente	Clúster	
	1	2
TIC_{Enigh}	84,8	20,2
$Gasto\ TIC_{Enigh}$	23,5	11,4

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

Análisis Discriminante de la clasificación de la ENIGH 2018.

Clúster original		Pertenencia a grupo pronosticado		Total
		1	2	
Recuento	1	629 712	2 748	632 460
	2	0	801 267	801 267
Porcentaje	1	99,6	0,4	100,0
	2	0,0	100,0	100,0

99,8% de los casos agrupados originales clasificados correctamente.

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

Anexo 8: Prueba de contraste de medias del IPD_{Enaho_z} según región.

Contraste	Dif	Error Estándar	Grados de libertad	t	Sig.
Central - Chorotega	-9,18	0,96	1102	-9,53	0,00
Central – Pacífico Central	-8,04	0,96	1102	-8,37	0,00
Central – Brunca	9,63	0,90	1102	10,70	0,00
Central – Huetar Atlántico	-9,75	0,89	1102	-10,90	0,00
Central – Huetar Norte	-12,44	0,84	1102	-14,81	0,00
Chorotega – Pacífico Central	1,14	1,24	1102	0,92	0,94
Chorotega – Brunca	0,46	1,19	1102	0,38	1,00
Chorotega – Huetar Atlántico	-0,57	1,19	1102	-0,48	1,00
Chorotega – Huetar Norte	-3,26	1,15	1102	-2,84	0,05
Pacífico Central – Brunca	1,59	1,19	1102	1,34	0,76
Pacífico Central – Huetar Atlántico	1,70	1,19	1102	1,44	0,70
Pacífico Central – Huetar Norte	4,40	1,14	1102	3,84	0,00
Brunca – Huetar Atlántico	-0,11	1,14	1102	-0,01	1,00
Brunca – Huetar Norte	-2,80	1,09	1102	-2,56	0,11
Huetar Atlántico – Huetar Norte	-2,69	1,09	1102	-2,47	0,13

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Anexo 9: Resultados del análisis de conglomerados aplicado a los componentes delIPD_{Enaho}

Centro de los conglomerados finales para ENAHO 2018.

Componente	Clúster		
	1	2	3
TIC_{Enaho}	17,6	59,1	84,4
$Gasto\ TIC_{Enaho}$	13,1	18,7	52,0

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018, INEC. [Archivo de datos].

Análisis Discriminante de la clasificación en la ENAHO 2018.

Clúster original	Pertenenencia a grupo			Total	
	pronosticado				
	1	2	3		
Recuento	1	621 233	926	425	622 584
	2	272	659 555	5 519	665 346
	3	0	11 344	235 560	246 904
Porcentaje	1	99,8	0,1	0,1	100,0
	2	0,0	99,1	0,8	100,0
	3	0,0	4,6	95,4	100,0

98,8% de casos agrupados originales clasificados correctamente.

Fuente: elaboración propia con datos de la Encuesta Nacional de Ingresos y Gastos de Hogares 2018, INEC. [Archivo de datos].

Anexo 10: Prueba de igualdad de medias para el valor del IPD entre conglomerados de la ENIGH y ENAHO.

	Media	Error Típico	t	Grados de libertad	Sig.	Intervalo de confianza para la diferencia	
						Inferior	Superior
Pobre Digital-ENIGH	84,2	0,3					
Pobre Digital Alto - ENAHO	84,6	0,1	1,6	1 368	0,1	-0,1	0,9

Fuente: elaboración propia con datos de la Encuesta Nacional de Hogares 2018 y Encuesta Nacional de Ingresos y Gastos 2018, INEC. [Archivo de datos].

Anexo 11: Resultados de la prueba de independencia de las variables sociodemográficas y los conglomerados de dos grupos para la ENIGH.

		Estadístico	gl	Sig.
Región*Nivel_IPD	Rao & Scott ajustado (χ^2)	204,1	5	,00
	Rao & Scott ajustado (F)	33,2	4,0; 1823,1	,00
Zona * Nivel_IPD	Rao & Scott ajustado (χ^2)	222,97	1,0	,00
	Rao & Scott ajustado (F)	105,8	1,0; 456	,00
Cond_Activ. *	Rao & Scott ajustado (χ^2)	269,2	2,0	,00
	Rao & Scott ajustado (F)	88,0	1,9; 900	,00
Sexo_jefe *	Rao & Scott ajustado (χ^2)	0,4	1,0	,63
	Rao & Scott ajustado (F)	0,2	1,0; 456	,63
Escolaridad *	Rao & Scott ajustado (χ^2)	1406,3	3,0	,00
	Rao & Scott ajustado (F)	264,7	2,9; 1338,8	,00
Grupo_edad *	Rao & Scott ajustado (χ^2)	188,8	5,0	,00
	Rao & Scott ajustado (F)	22,9	4,5; 2062,9	,00

Anexo 12: Resultados de la prueba de independencia de las variables sociodemográficas y los conglomerados de dos grupos para la ENAHO.

		Estadístico	gl	Sig.
Región*Nivel_IPD	Rao & Scott ajustado (χ^2)	521,6	10	0,00
	Rao & Scott ajustado (F)	35,9	9,5; 10474	0,00
Zona * Nivel_IPD	Rao & Scott ajustado (χ^2)	615,5	2	0,00
	Rao & Scott ajustado (F)	224,7	1,9; 2126,0	0,00
Cond_Activ. *	Rao & Scott ajustado (χ^2)	756,3	4	0,00
	Rao & Scott ajustado (F)	167,7	3,97; 4400,7	0,00
Sexo_jefe *	Rao & Scott ajustado (χ^2)	22,6	2	0,00
	Rao & Scott ajustado (F)	10,4	2; 2201,1	0,00
Escolaridad *	Rao & Scott ajustado (χ^2)	4606,7	6	0,00
	Rao & Scott ajustado (F)	659,0	6; 6593,3	0,00
Grupo_edad *	Rao & Scott ajustado (χ^2)	413,4	10	0,00
	Rao & Scott ajustado (F)	38,2	9,8; 10843,6	0,00