



□ UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN □ FACULTAD DE ECONOMÍA □ CENTRO DE INVESTIGACIONES ECONÓMICAS □

Identificación de patrones de inclusión financiera en México con análisis multivariado: creación de dimensiones aplicando PCA para variables mixtas

Identifying financial inclusion patterns in Mexico with multivariate analysis: creating dimensions by applying PCA to mixed variables

Elizabeth Ximena Elizarrarás Barbosa* Ericka García Blanquel** Ambrosio Ortiz Ramírez***

Información del artículo	Resumen
Recibido: 31/12/2025 Aceptado: 13/03/2026 Clasificación JEL: C38, G21, C45 Palabras clave: Inclusión financiera, PCA, Dimensiones, Estructuras latentes	Este estudio aborda la complejidad del análisis de la inclusión financiera en México por la multidimensionalidad de los datos de la ENIF 2024 y la cantidad de variables cualitativas contenidas, las cuales componen las categorías que limitan la participación de la población en el sistema financiero formal. El objetivo principal es la identificación de patrones aplicando análisis multivariado para seleccionar adecuadamente el conjunto de dimensiones que conforman las categorías de estudio y extraer con mayor exactitud sus características. Aplicando el Análisis de Componentes Principales para variables mixtas, se trabajan eficientemente las variables categóricas de la encuesta considerando que la mayoría de las técnicas multivariadas trabaja con variables numéricas. Los hallazgos demuestran que esta metodología es eficiente para la detección de patrones y reducción de dimensionalidad lo que resulta útil para emplearse en investigaciones con mayor profundidad.
Article information	Abstract
Received: 31/12/2025 Accepted: 13/03/2026 JEL Classification: C38, G21, C45 Keyword: Financial inclusion, PCA, Dimensions, Latent structures	This study addresses the complexity of financial inclusion analysis in Mexico due to the multidimensionality of the ENIF 2024 data and the number of qualitative variables contained, which compose the categories that limit the population's participation in the formal financial system. The main objective is the identification of patterns by applying multivariate analysis to properly select the set of dimensions that make up the categories under study and to extract their characteristics with greater accuracy. By applying Principal Component Analysis for mixed variables, the categorical variables of the survey are efficiently handled, considering that most multivariate techniques work with numerical variables. The findings demonstrate that this methodology is efficient for pattern detection and dimensionality reduction, which is useful for use in more in-depth research.

*Autora de correspondencia
Instituto Politécnico Nacional
elizarrarasb1200@alumno.ipn.mx
<https://orcid.org/0009-0009-0784-834X>

**Instituto Politécnico Nacional
egarciab@ipn.mx
<https://orcid.org/0009-0006-0581-4427>

***Ambrosio Ortiz Ramírez
amortiz@ipn.mx
<https://orcid.org/0000-0002-3698-2873>

1. Introducción

La inclusión financiera en México representa un pilar estratégico para el desarrollo económico y la reducción de la desigualdad social. La Encuesta Nacional de Inclusión Financiera (ENIF, 2024) es una herramienta clave para diagnosticar este fenómeno. El último levantamiento, en 2024, proporciona información más detallada sobre categorías como el acceso a productos financieros, factores socioeconómicos, geográficos y demográficos de la población, así como de la educación financiera. Sin embargo, la naturaleza mixta de sus variables y su longitud dificultan la interpretación de patrones subyacentes.

La literatura previa se ha centrado principalmente en el uso de modelos econométricos y la construcción de índices sintéticos para identificar determinantes socioeconómicos básicos (Carballo, 2020; Pasto, 2022). Estos trabajos han permitido diagnosticar barreras generales, pero aún presentan limitaciones para gestionar la alta dimensionalidad y la multicolinealidad de las bases de datos por la cantidad de variables cualitativas de la ENIF. Así, otras investigaciones han implementado el Análisis de Componentes Principales (PCA) casi en su totalidad para la creación de indicadores de intensidad, omitiendo el potencial de la reducción de ruido en procesos que segmentan la población (Pérez-García et al., 2021).

Esto genera una brecha en la búsqueda que integre el análisis multivariado para identificar patrones latentes. La originalidad de este estudio radica en superar la descripción univariada tradicional mediante el uso de PCA mixto para clasificar y agrupar adecuadamente las dimensiones de estudio, sin generar información redundante y con la mayor aportación de características para la realización de estudios más especializados.

El objetivo principal es la identificación de patrones aplicando un análisis multivariado para reducir adecuadamente el conjunto de datos que conforman las categorías de estudio y extraer con mayor exactitud sus características. El PCA es una herramienta eficaz, la cual, mediante transformaciones matriciales, genera dimensiones para reducir la cantidad de los datos. Con esta metodología se facilita el análisis y visualización de los datos, así como la interpretación de los patrones para explicar diversos fenómenos con notoria complejidad, como el análisis de la desigualdad del acceso y uso al sistema financiero, sin perder información relevante.

Finalmente, el principal hallazgo fue identificar las dimensiones latentes de inclusión financiera mediante el PCA mix; esto facilita perfilar grupos de población con patrones comunes en materia, haciendo su análisis más factible con otras herramientas para identificar brechas por grupos de población con mayor precisión aplicando ciencia de datos.

2. Revisión de la literatura

La inclusión financiera es aquel acceso y uso de productos y servicios financieros adecuados y accesibles para toda la población, así como la educación financiera y protección al consumidor. Dentro de la literatura académica se presentan variaciones significativas en la operacionalización del concepto. En el estudio de Honoban (2008), la inclusión financiera se basa únicamente en la dimensión de uso, mientras que en el estudio de Allen et al. (2012) se desarrolla en tres dimensiones: si la persona tiene un producto financiero formal, si lo ha utilizado para ahorrar o si ha realizado tres o más disposiciones bancarias por mes. A pesar de estos avances, la selección de dimensiones y sus variables suele estar condicionada por la naturaleza de los datos o las herramientas analíticas empleadas.

En México, la investigación reciente subraya la necesidad de analizar las barreras estructurales que condicionan el acceso y el uso, lo cual se alinea con las dimensiones exploradas en el presente estudio (Carballo, 2020; Pasto, 2022). Por otro lado, la ENIF proporciona información clave sobre la inclusión financiera, lo que la convierte en una de las principales bases de datos para estos estudios (Comisión Nacional Bancaria y de Valores e Instituto Nacional de Estadística y Geografía [CNBV & INEGI], 2024). Aplicar PCA a los datos de la ENIF 2024 permite manejar adecuadamente las variables categóricas, mejorar la comprensión de estos patrones y utilizarlos como base para realizar estudios con mayor profundidad.

En el ámbito internacional, un estudio realizado en la India permitió identificar la proximidad geográfica y la falta de infraestructura y de educación financiera como factores críticos que limitan el acceso (Lenka, 2015), mientras que en África subsahariana se demostró que esta técnica sirve para cuantificar el impacto del sector financiero en el crecimiento económico y en la mitigación de la pobreza (Akpan et al., 2024).

3. Datos y metodología

El presente estudio se desarrolló a partir del análisis de los microdatos de la ENIF 2024, encuesta diseñada por la CNBV y levantada en conjunto con el INEGI. La ENIF proporciona información representativa a nivel nacional y regional sobre las características sociodemográficas, el acceso y uso de productos financieros, las barreras percibidas por los agentes económicos, así como la educación y la capacidad financiera de la población adulta en México (de 18 años y más). De las 389 variables¹ utilizadas, 35 son cuantitativas y 354 cualitativas.

3.1. Datos y Variables

Las variables utilizadas se seleccionaron conforme a la estructura temática definida en la metodología oficial del levantamiento, con el fin de capturar las categorías de la inclusión financiera (INEGI, 2024). En la Tabla 1 se presenta un desglose y una breve descripción del conjunto de variables que conforman cada una de ellas.

Tabla 1: Descripción de categorías utilizadas

Categorías	Descripción	Variables ENIF	Variables Categóricas	Variables Numéricas	Variables en PCA
Características sociodemográficas	Explica el perfil sociodemográfico de la persona encuestada.	22	20	2	91
Vulnerabilidad	Conocer en qué forma las personas administran sus gastos, la capacidad de pago.	29	27	2	102
Ahorro Formal	Se realiza cuando una parte del ingreso se ahorra u opta por gastar menos y este ahorro se realiza mediante alguna institución formal.	67	64	3	139
Crédito Formal	Confianza que se le otorga a un agente económico en la capacidad, voluntad y solvencia para el cumplimiento de una obligación contraída.	65	53	12	132
Pagos	Acción de solventar algo que se debe.	39	37	2	73
Seguros	Busca disminuir las consecuencias económicas adversas de distintos eventos.	51	49	2	107
Afore	Contempla la acción de ahorro a largo plazo para utilizarse posteriormente en su retiro.	31	29	2	100
Uso de canales financieros	Contemplan tres tipos de canales, sucursales bancarias, cajeros automáticos y corresponsales bancarios.	8	6	2	28
Confianza	Refleja la confianza de los agentes económicos al solicitar servicios de alguna institución financiera.	30	28	2	74
Protección	Expresa el conocimiento de los agentes económicos de las instituciones que los protegen en cuanto al ámbito financiero.	16	14	2	42

¹En este estudio, las variables corresponden a los indicadores medidos en la ENIF, las cuales pueden ser numéricas, binarias o categóricas. Por su parte, las dimensiones son los ejes factoriales generados por el análisis multivariado empleando la técnica PCAmix la cual resume la estructura común de variación del conjunto de variables. Por tanto, las variables son datos observados y las dimensiones son datos construidos por medio de transformaciones lineales y estadística.

Tabla 1: Descripción de categorías utilizadas (continuación)

Capacidades y educación financiera	Forma en que los agentes económicos aplican sus conocimientos y habilidades en la toma de decisiones financieras.	12	10	2	47
Propiedad de activos	Tenencia de algún tipo de activo que respalde su ingreso al sistema financiero formal.	19	17	2	63
TOTAL		389	354	35	998

Fuente: Elaboración propia con datos de ENIF (2024).

Es importante señalar que las variables utilizadas pueden ser categóricas o numéricas; esta distinción es crucial para comprender cómo las diferentes categorías y sus variables interactúan y afectan la inclusión financiera. Por lo tanto, en este análisis se realiza una clara distinción entre ambos tipos de variables para aplicar el PCA correspondiente, ya que una variable categórica se presenta como varias columnas, una por clase.

3.2. Algoritmo para la creación de dimensiones

Considerando que hay variables categóricas y numéricas, se aplicó el PCA para Datos Mixtos (PCA mix) como paquete principal en el lenguaje de programación R. Este permite manejar ambos tipos de variables de manera eficiente y es un tipo de extensión del PCA tradicional que adapta la técnica para tratar datos mixtos. Su diseño integra adecuadamente las diferencias de escala y tipo de datos.

El PCA mix busca encontrar un vector de componentes $f \in \mathbb{R}^n$ que maximice la relación entre variables de tipos diferentes; es decir, maximiza la suma de los cuadrados de las correlaciones con las variables continuas y la suma de las razones de correlación con las variables categóricas (Chaven et al., 2017).

$$PCAmix(f) = \sum_{j \in J_{cont}} cor^2(f, x_j) + \sum_{j \in J_{cat}} \eta^2(f, x_j)$$

Donde, J_{cont} refleja el conjunto de variables cuantitativas, J_{cat} es el conjunto de variables cualitativas, $cor^2(f, x_j)$ muestra el cuadrado del coeficiente de correlación de Pearson y $\eta^2(f, x_j)$ es la razón de correlación, que mide la varianza de f explicada por la variable categórica x_j . Para la resolución de la optimización, se preprocesan los datos mediante una matriz global $Z = [X_1|X_2]$, donde X_1 contiene las variables cuantitativas estandarizadas (con media 0 y varianza 1) y X_2 contiene las variables cualitativas transformadas en una matriz de indicadores (dummies) reescalada. La matriz, al descomponerse en valores singulares, permite obtener eigenvalores λ , que reflejan la varianza explicada por cada componente principal y, así, reducir los datos sin una pérdida significativa de información (Pagés, 2014).

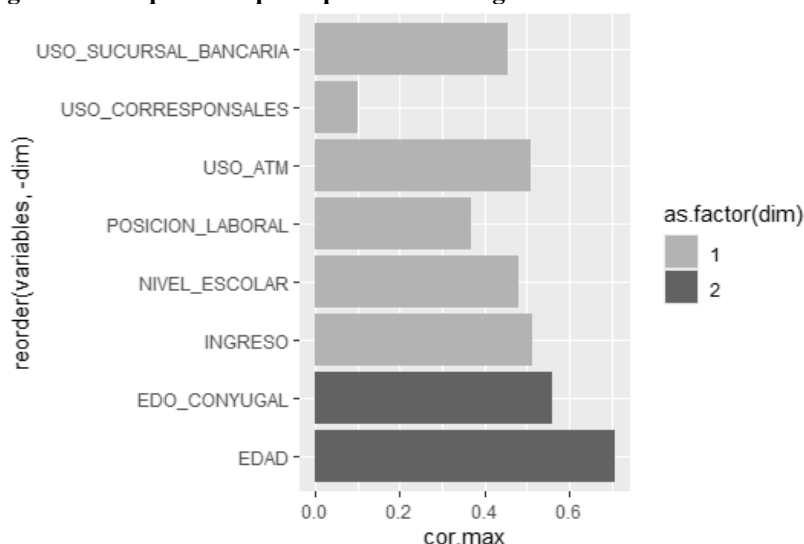
La metodología consta de 3 etapas para este análisis. Los datos se preprocesaron para su correcta preparación y tratamiento. En la primera etapa, se realizó el proceso de Extracción, Transformación y Carga (ETL): se extrajeron los datos del repositorio (INEGI), se aplicó el proceso de transformación (limpieza, enriquecimiento y eliminación de datos) y, finalmente, se cargaron los datos para proceder con el análisis, la visualización y la detección de patrones. En la segunda etapa, se realizó la selección de variables, se incluyeron aquellas contempladas en la metodología de la ENIF 2024, siguiendo criterios de relevancia teórica y consistencia conceptual. La etapa tres se enfocó en la técnica PCA, aplicando transformaciones lineales de las variables originales a los componentes principales. De acuerdo con la metodología del INEGI, el procesamiento de las

variables categóricas mediante la creación de variables ficticias expandió la matriz original de 389 a 998², como resultado del proceso que sigue la metodología utilizada (PCA mix).

Para esta investigación, se optó por realizar 13 pruebas en cada una de las 12 categorías de la Tabla 1 y se utilizó el criterio de varianza explicada acumulada, que indica la retención de los componentes que explican una proporción considerable de la varianza total. En este apartado se identificaron los componentes que explican el 70% de la varianza total.

La Figura 1 muestra los componentes principales de la categoría de uso de canales financieros, utilizando las características de dichas variables en este trabajo. El primer componente (gris claro) refleja el perfil de inclusión y el nivel socioeconómico, mientras que el segundo (gris oscuro) refleja el momento de vida de los agentes económicos. Estos componentes principales servirán como nuevas variables al aplicar otros modelos de evaluación.

Figura 1: Componentes principales de la categoría de uso de canales financieros.



Fuente: Elaboración propia con datos de ENIF (2024).

Una vez que se obtuvieron los componentes principales, se analizó la carga de cada una de las variables que contemplaban cada categoría para interpretar las dimensiones que subyacen a la inclusión financiera en México. Dentro de estos componentes se identificaron distintos patrones que reflejan aspectos sociodemográficos y de acceso a productos financieros, así como barreras y percepciones de los usuarios, incluida la educación financiera. El modelo se validó a través del corte de varianza explicada y la permutación de los componentes principales. Se realizó una validación cruzada (*k-fold cross-validation*, $k=5$) para asegurar que los componentes identificados fueran robustos y no producto del azar.

4. Resultados y discusión

El análisis de los microdatos de la ENIF 2024 con PCA mix permitió realizar la identificación de los componentes principales que subyacen a la inclusión financiera en México, extrayendo aquellos componentes que explican el 70% de la varianza total, asegurando que la reducción de los datos no haya comprometido la naturaleza o esencia de estos. Luego de este análisis, se identificaron los factores con mayor carga factorial o relevancia para cada categoría evaluada (véase Tabla 2, en donde se indican los principales).

²Las 998 variables en el PCA se generaron al momento en que el PCAmix transformó las variables categóricas creando una columna o variable independiente por cada categoría de la variable en la ENIF. Los valores que toman estas nuevas variables es 1 si el individuo pertenece a esa categoría y 0 en caso contrario.

Tabla 2. Categorías y sus componentes principales.

Categorías	Componentes principales	Principales factores subyacentes
Características sociodemográficas	8	Situación laboral y nivel escolar
Vulnerabilidad	15	Condición laboral y condición de frases financieras
Ahorro Formal	15	Destino de ahorro y condición de cuenta de nómina
Crédito Formal	15	Situación de tarjeta y razón de no tenencia
Pagos	15	Transferencias y tipo de compras
Seguros	13	Tenencia y comparación de seguros
Afore	13	Tipo de ingreso con el que cubre gastos y situación de afore
Uso de canales financieros	2	Ciclo de vida y perfil técnico de uso
Confianza	5	Condición de reclamo y tipo de problema
Protección	5	Conocimiento de institución que protege el ahorro y ciclo de vida
Capacidades	3	Conocimiento financiero y condición geográfica
Propiedad de activos	7	Condición de vivienda y auto
TOTAL	116	

Fuente: Elaboración propia con base en datos de INEGI (2024).

Al aplicar el método, se reduce la dimensionalidad sin perder información relevante, lo que da como resultado la columna *Componentes principales*, donde se expresa el conjunto subyacente de variables para cada categoría. La aplicación del método es eficiente para reducir la dimensionalidad de los datos en cada categoría, con 116 dimensiones al final de la implementación de cada PCA, aportando óptimamente a las aplicaciones futuras de un análisis más profundo sin perder la esencia de los datos o valores relevantes. Cabe resaltar que la reducción de las dimensiones dio como resultado factores subyacentes que corresponden a la parte descrita en la tabla (la descripción detallada y los datos utilizados se encuentran en los Apéndices A y B).

4.1. Evaluación

El presente trabajo presenta un análisis multivariado mediante la aplicación de PCA para datos mixtos, que es una herramienta de suma utilidad y eficiencia para abordar la complejidad de medir la inclusión financiera en México. Al aplicar este método a los microdatos de la ENIF 2024, se identificaron patrones y dimensiones subyacentes clave que afectan el acceso y el uso de productos y servicios financieros en el país. La reducción de las dimensiones de estas categorías permite realizar análisis posteriores con mayor profundidad y menor complejidad al interpretar los datos, gracias a la aplicación del modelo PCAmix.

La técnica de análisis multivariado aplicada a la ENIF 2024 ha demostrado ser eficaz para reducir la dimensionalidad de los datos sin perder información relevante, permitiendo que los datos puedan ser analizados con mayor facilidad mediante modelos econométricos o que se puedan identificar patrones y brechas por grupos de población con mayor precisión, aplicando herramientas de ciencia de datos. Esto permite incluir una cantidad considerablemente menor en el estudio y que los resultados no se vean afectados, pues una de las principales ventajas del modelo es que se genera la reducción mediante dimensiones que contemplan toda la información de cada una de las variables registradas, sin importar si es cuantitativa o cualitativa.

Se sugiere que en futuras indagaciones se incorporen variables adicionales al análisis, que aborden el impacto de factores externos, como las políticas públicas de inclusión financiera, las iniciativas de inclusión digital o los servicios financieros digitales, como las Fintech.

5. Conclusiones y recomendaciones

El presente trabajo presenta un análisis multivariado mediante la aplicación de PCA a datos mixtos, una herramienta de suma utilidad y eficiencia para abordar la complejidad de medir la inclusión financiera en México. Al aplicar este método a los microdatos de la ENIF 2024, se identificaron patrones y dimensiones subyacentes clave que afectan el acceso y el uso de productos y servicios financieros en el país.

La reducción de las dimensiones influye en análisis posteriores, con mayor profundidad y menor complejidad al interpretar los datos, por lo que es necesario considerar algunos aspectos clave para mejorar tanto la interpretación como los resultados. Esta indagación, al concentrarse en la ENIF, puede generar sesgos en los datos, a pesar de que sea una muestra representativa a nivel nacional, pues existe población encuestada que no responde a las preguntas, lo que puede provocar subrepresentación de ciertos grupos poblacionales, afectar la precisión de las estimaciones y excluir variables que no se miden en el análisis. Estos métodos pueden ser esenciales para el diseño de políticas públicas focalizadas y efectivas, reduciendo las brechas en el acceso a servicios y productos financieros, lo cual puede lograrse mediante la aplicación de algoritmos de clustering (k-modas) o de redes neuronales, sin enfrentar el problema de la cantidad y el tipo de variables.

Los hallazgos sugieren que la inclusión financiera en México puede mejorar al abordar tanto las barreras económicas y de infraestructura como las educativas y culturales. Asimismo, el uso del PCA mix ha demostrado ser una estrategia eficiente y eficaz para reducir la dimensionalidad de los datos, al tiempo que ayuda a identificar patrones que, de otro modo, podrían no haberse considerado. Es necesario continuar explorando y validando los patrones mediante estudios adicionales e incorporar nuevas variables que puedan influir en la inclusión financiera. Se recomienda realizar una desagregación por región geográfica y género para obtener una comprensión más detallada de cómo estas dimensiones se incluyen en distintos contextos o de cómo se comportan ciertos grupos de sectores con una mayor vulnerabilidad (un ejemplo de su aplicación puede observarse en el Apéndice C).

Agradecimientos

Este trabajo fue financiado por el Instituto Politécnico Nacional, proyecto IPN-SIP20253432 y la Secretaría de Ciencia, Humanidades e Innovación.

Referencias

- Akpan, E. O., Akpan, S. A., & Ifeanyichukwu, C. N. (2024). The impact of financial inclusion on economic growth and poverty reduction: Empirical evidence from sub-Saharan Africa. *Economic Research-Ekonomska Istraživanja*.
- Banco Mundial. (2017). Financial inclusion for development. World Bank Group. Recuperado de [13 de noviembre de 2025]: <https://www.worldbank.org/en/topic/financialinclusion>
- Carballo, I. E. (2020). Inclusión financiera y tecnologías digitales: una mirada a la literatura actual. *Revista de Ciencias Sociales (Ve)*, 26(4), 183-199. <https://doi.org/10.31876/rcs.v26i4.34658>
- Chavent, M., Kuentz-Simonet, V., Labenne, A., & Saracco, J. (2017). Multivariate Analysis of Mixed Data: The PCAmixdata R Package. Cornell University. <https://arxiv.org/abs/1411.4911>
- CNBV & INEGI. (2024). *Encuesta Nacional de Inclusión Financiera (ENIF) 2024: Reporte de Resultados*. https://www.cnbv.gob.mx/Inclusi%C3%B3n/Anexos%20Inclusin%20Financiera/Reporte_ENIF2024.pdf
- CNBV & INEGIa. (2024). *Encuesta Nacional de Inclusión Financiera (ENIF) 2024: Documento conceptual*. Comisión Nacional Bancaria y de Valores; Instituto Nacional de Estadística y Geografía. <https://www.inegi.org.mx/contenidos/programas/enif/2024/doc/889463923121.pdf>

- INEGI. (2024). Encuesta Nacional de Inclusión Financiera (ENIF) 2024. Diseño conceptual 2025. <https://www.inegi.org.mx/contenidos/programas/enif/2024/doc/889463923121.pdf>
- Jolliffe, I. T. (2002). Principal component analysis (2nd ed.). Springer.
- Lenka, S. K. (2015). Measuring financial development in India: A PCA approach. *Theoretical and Applied Economics*, 22(1), 187–198.
- Pagès, J. (2014). *Multiple Factor Analysis by Example Using R*. (1st Ed.) CRC Press
- Pastor, F. (2022). Índice de Inclusión Financiera: una Medición Multidimensional y Global (Documento de Trabajo N° 195). Red Nacional de Investigadores en Economía (RedNIE). <https://rednie.eco.unc.edu.ar/files/DT/195.pdf>
- Pérez-García, J., Medina, A., & Gallegos, R. (2021). Análisis multivariado de la inclusión financiera en economías emergentes: Un enfoque de componentes principales. *Revista Mexicana de Economía y Finanzas*, 16(2), 1-22. <https://doi.org/10.21919/remef.v16i2.603>
- Allen, F., Carletti, E., Cull, R., Qian, J., Senbet, L., & Valenzuela, P. (2014). The African financial development and financial inclusion gaps. *Journal of African Economies*, 23(5), 614-642
- Honohan, P. (2008). Cross-country variation in household access to financial services. *Journal of Banking and Finance*, 32, 2493-2500.

Apéndices

Apéndice A

El PCA mix está diseñado para procesar datos numéricos y categóricos sin perder la esencia de ninguno. Para los datos numéricos utiliza la lógica de un PCA convencional y para los datos categóricos realiza un análisis de correspondencias múltiples (MCA). En el caso de “uso de canales financieros”, el PCA no puede sumar “casado” con “34” años, por esta razón, el algoritmo realiza una codificación binaria. Es decir, cada categoría de las variables cualitativas se convierte en una columna propia (dimensión), como se muestra en la Tabla A1, en donde la suma de las dimensiones para ambos tipos de variables es 28. Es importante mencionar que el algoritmo toma como una dimensión aquellas variables cualitativas que tienen solo dos categorías y en variables con más de dos categorías, se omite una de ellas (n-1) para evitar multicolinealidad perfecta. Para esta variable, el total de 28 dimensiones se deriva de la siguiente estructura:

- Tres variables binarias (3 dimensiones).
- Tres variables con 7, 13 y 6 clases, a las que se les resta una dimensión por cada variable (6, 12 y 5 dimensiones respectivamente).
- Dos variables numéricas (2 dimensiones).

Tabla A1: Descripción de variables y dimensiones del uso de canales financieros

Nomenclatura	Variable	Descripción	Dimensiones
USO_SUCURSAL_BANCARIA	Uso de sucursales bancarias	Binaria	1
USO_CORRESPONSALES	Uso de correspondientes bancarios	Binaria	1
USO_ATM	Uso de cajeros automáticos	Binaria	1
POSICION_LABORAL	Puesto laboral	Empleada, jornalera, ayudante con pago, patrona, trabajadora por su cuenta, trabajadora sin pago, blanco por secuencia.	6

Tabla A1: Descripción de variables y dimensiones del uso de canales financieros (continuación)

NIVEL_ESCOLAR	Nivel escolar	Ninguno, preescolar, primaria, secundaria, normal básica, estudios técnicos (sec.), preparatoria, estudios técnicos (prepa.), licenciatura, especialidad, maestría, doctorado, no sabe.	12
EDO_CONYUGAL	Situación conyugal	Unión libre, separada, divorciada, viuda, casada, soltera.	5
INGRESO	Ingreso	Continua	1
EDAD	Edad	Continua	1
TOTAL			28

Fuente. Elaboración propia.

El PCA mix tiene como objetivo encontrar patrones de correlación entre esas 28 dimensiones, en el caso de esta categoría el algoritmo detecta que las personas con alto “Nivel Escolar” suelen tener un mayor “Ingreso” y usan más las “Sucursales Bancarias”, es decir, en lugar de ver estos 3 componentes separados los agrupa en “componente principal” o una “super variable”. Por otro lado, esta agrupación nos ayuda a identificar dos ejes, el primero de ellos engloba qué tanto el individuo se integra al sistema según su estatus socioeconómico y la segunda muestra cómo influye el momento de su vida sobre su comportamiento financiero, dando así el origen de los nombres de los dos componentes obtenidos.

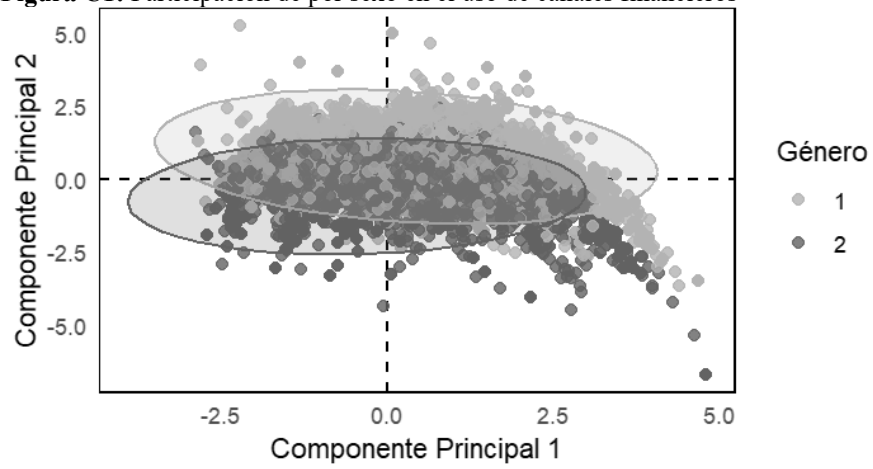
Apéndice B

Para consultar bases de datos completas, código, categorías y dimensiones desglosadas véase el GitHub encriptado en el siguiente código QR.



Apéndice C

La figura siguiente muestra, mediante un análisis de clusters de acuerdo con los scores de los componentes principales, la participación de los hombres y mujeres para el uso de canales financieros, lo que visualmente sugiere que las mujeres (1) tienen menos participación que los hombres (2) puesto que se encuentra una cantidad superior de puntos grises claro por fuera de la unión de los óvalos, más a la periferia que los puntos gris oscuro, representando como este método puede aportar a la medición de la brecha de género.

Figura C1: Participación de por sexo en el uso de canales financieros

Fuente. Elaboración propia con datos de ENIF (2024).