

## ANEXO 6

### FORMATO DE INFORME DE AVANCE FÍSICO-FINANCIERO

#### N°\_3\_\_ INFORME FINAL DE AVANCE FÍSICO Y FINANCIERO

#### Proyecto: “IMPLEMENTACIÓN DE UN SISTEMA DIGITAL AUTOMATIZADO PARA MEDIR PARÁMETROS SOCIODEMOGRÁFICOS EN TIEMPO REAL EN LA REGIÓN PIURA”

##### I. Descripción del Proyecto

En el marco del proyecto de implementación de un sistema digital automatizado para la medición en tiempo real de parámetros sociodemográficos en la Región Piura, se desarrolló la fase final, orientada al procesamiento analítico de la información recolectada y a la producción de artículos científicos. Esta etapa comprendió la depuración, sistematización y análisis estadístico de los datos obtenidos mediante el sistema digital automatizado, a partir de una muestra de 365 hogares, seleccionados y distribuidos en diez zonas rurales de la región: Pacaipampa, Frías, Montero, Canchaque, San Miguel del Faique, Sondorillo, El Tallán, Cura Mori, Santo Domingo y Buenos Aires.

El procesamiento de la información permitió la estimación de indicadores demográficos clave, incluyendo variables asociadas a vivienda, estructura poblacional y fecundidad, garantizando consistencia interna, trazabilidad de los datos y reducción de errores no muestrales derivados del levantamiento de información. Sobre esta base empírica, se desarrollaron dos productos científicos. El primero, titulado “*Vivienda, estructura poblacional y fecundidad en diez zonas rurales de Piura, 2025: un análisis sociodemográfico integral*”, se fundamenta en el análisis inferencial de los indicadores demográficos generados. El segundo, titulado “*Digital versus paper data collection in large-scale household surveys: time efficiency, heterogeneity by complexity, and modelling of the within-pair difference*”, aborda una evaluación comparativa del desempeño del sistema digital frente a métodos tradicionales de recolección de datos, incorporando análisis de eficiencia temporal y modelamiento de diferencias intra-par.

El objetivo del presente informe es evaluar el desempeño operativo y analítico del sistema digital automatizado en el contexto de la recolección de datos de la muestra del proyecto, con énfasis en la calidad, consistencia y validez de los indicadores sociodemográficos generados en las diez zonas rurales de la Región Piura.

## **II. Objetivos del Proyecto**

### **2.1 Objetivo General:**

- Implementar un sistema digital automatizado para la toma de parámetros sociodemográficos.

### **2.2 Objetivos Específicos:**

- Validar instrumentos para la toma de parámetros sociodemográficos.
- Diseñar un sistema digital automatizado para la toma de parámetros sociodemográficos.
- Probar en campo un sistema digital automatizado para la toma de parámetros sociodemográficos.
- Comparar tiempos en la recolección de información e inicio de análisis.
- Analizar parámetros sociodemográficos recopilados.

## **III. Estado Actual del Proyecto**

El presente informe expone el estado actual del proyecto en su fase final de ejecución, precisando las actividades desarrolladas durante esta etapa. En este contexto, se ha concluido el procesamiento, depuración y análisis de la información recolectada, lo que ha permitido la construcción de indicadores demográficos correspondientes a la muestra de estudio.

De manera paralela, se ha avanzado en la elaboración de artículos científicos sustentados en los resultados obtenidos; sin embargo, dichos manuscritos aún no han sido sometidos a evaluación en revistas indexadas, encontrándose actualmente en fase de revisión técnica, metodológica y editorial.

Adicionalmente, se viene ejecutando una segunda capacitación dirigida al equipo de investigación, en el marco de las actividades contempladas en el proyecto.

Esta capacitación, denominada "Análisis de datos con Python", tiene como finalidad fortalecer las competencias técnicas en el procesamiento, análisis y gestión de bases de datos generadas mediante el sistema digital automatizado, contribuyendo así a garantizar la calidad, consistencia y rigurosidad de los resultados obtenidos, así como su adecuada difusión en espacios académicos especializados.

#### **IV. Avance Físico del Proyecto**

**Actividades Realizadas:** Generación de indicadores demográficos y redacción de artículos científicos.

**Porcentaje de Avance:** 90%

Tabla 1. Descripción de Actividades Realizadas

<b>N°</b>	<b>Actividades Realizadas</b>	<b>Porcentaje de Avance</b>
1	Generación de indicadores demográficos	100%
2	Redacción de artículos científicos (en proceso de envío)	80%
3	Capacitación: "Análisis de datos con Python" (en ejecución)	90%
	Total	90%

Fuente: Elaboración propia.

#### **V. Avance Financiero del Proyecto**

**Gastos Realizados:** Registro detallado de los gastos efectuados hasta el momento.

**Presupuesto Ejecutado:** Comparación entre el presupuesto planificado y el real.

Tabla 2. Detalle de Gastos por Categoría

**Adjunto el cuadro detallado de los gastos hasta la fecha:**



Gestión de Proyectos de Investigación en Ciencia Aplicada  
 Instituto de Investigación  
 Vicerrectorado de Investigación  
 Universidad Nacional de Piura

INVESTIGADOR PRINCIPAL: RICARDO ANTONIO ARMAS JUAREZ

TÍTULO DEL PROYECTO: "IMPLEMENTACIÓN DE UN SISTEMA DIGITAL AUTOMATIZADO PARA MEDIR PARÁMETROS SOCIODEMOGRÁFICOS EN TIEMPO REAL EN LA REGIÓN PIURA"

FUENTE DE FINANCIAMIENTO: RECURSOS DETERMINADOS

RESOLUCIÓN DE AUTORIZACIÓN GASTOS: N° 629-R-2024

CONTRATO: CONTRATO N° 017-2024-LNP

BASES DEL CONCURSO 2024: N° 161-CU-2024

REGLAMENTO: N° 644-CU-2023

SEGUIMIENTO FINANCIERO																
N°	GERÉRICA DE GASTO	PRESUPUESTO APROBADO	PRESUPUESTO EJECUTADO	SALDO	AÑO	SIAF	ORDEN DE COMPRA	CLASIFICADOR	PROVEEDOR	DESCRIPCIÓN	UNIDAD DE MEDIDA	CANTIDAD	PRECIO UNITARIO	PRECIO TOTAL	MONTO TOTAL	OBSERVACIONES
2024 - 2025																
1					2024	12312	370	2.3.1.5.1.2	PEGASUS PRODUCTOS Y SERVICIOS E.I.R.L	PAPEL A4	EMP X 500	20	\$/17.50	\$/350.00		Iniciales
2					2024	14137	493	2.3.1.5.1.2	P&S OFFICE	PACK DE TINTAS PARA IMPRESORA	JUEGO	1	\$/192.00	\$/192.00	\$/542.00	Iniciales
3					2024	17781	6510	2.3.2.7.14.98	JORGE ANDRE ALVARADO APONTE	SERVICIOS DE ASESORIA DE UN PROFESIONAL DESARROLLADOR/PROGRAMADOR	INFORME	1	\$/5,000.00	\$/5,000.00	\$/5,000.00	Iniciales
4					2024	17768	6477	2.3.2.7.14.98	ALEJOS CABRERA RUTH HILLAGROS	SERVICIOS DE CAPACITACION PROFESIONAL EN REDACCION DE ARTICULOS/CIENFICOS	INFORME	1	\$/5,000.00	\$/5,000.00	\$/5,000.00	Iniciales
5	2.3	5/23,300.00	5/23,242.00	5/5,058.00	2024	15012	520	2.3.1.7.1.1	CAUDERON CORREA (GRGE BRAYHAM	Fundas Protectoras para PC	JUEGO	1	\$/100.00	\$/100.00	\$/900.00	Iniciales
6					2025	5871	309	2.3.1.5.1.1	B2B PIURA E.I.R.L	Admisión de 01 KIT DE 04 CODER PARA GABINETE	GUIP	1	\$/850.00	\$/850.00	\$/850.00	Durante
					2025	5863	2508	2.3.2.7.13.98	B2B PIURA E.I.R.L	SERVICIOS DE CONSULTORIA PARA LA SEGUNDA VERSION DEL SISTEMA DIGITAL	INFORME	1	\$/5,000.00	\$/5,000.00	\$/5,000.00	Durante
					2025	10802	-	2.3.2.1.2.1	RICARDO ANTONIO ARMAS JUAREZ	PASAJES	PASAJE	5	\$/100.00	\$/500.00	\$/2,380.00	Durante
					2025	10803	-	2.3.2.1.2.2	RICARDO ANTONIO ARMAS JUAREZ	PASAJES	PASAJE	5	\$/100.00	\$/500.00	\$/2,420.00	Durante
					2025	11442	453	2.3.1.9.1.2	YULLY GONZA HERREPA	VIAJES	UNID	12	\$/75.00	\$/900.00	\$/900.00	Durante
					2024	13633	-	2.5.3.1.1.1	DIANA EVELYN HIDALGO ECHE	ESTIPENDIO ALUMNOS	INFORME	2	\$/500.00	\$/1,000.00	\$/900.00	Durante
					2024	14342	-	2.5.3.1.1.1	DIANA EVELYN HIDALGO ECHE	ESTIPENDIO ALUMNOS	INFORME	1	\$/500.00	\$/500.00	\$/2,500.00	Iniciales
					2025	15951	-	2.5.3.1.1.1	GABRIELA IPANAQUE CARDEMAS	SUBVENCION ALUMNOS	INFORME	1	\$/500.00	\$/500.00	\$/2,500.00	Durante
					2025	17667	-	2.5.3.1.1.1	GABRIELA IPANAQUE CARDEMAS	SUBVENCION ALUMNOS	INFORME	1	\$/500.00	\$/500.00	\$/2,500.00	Durante
					2024	966	363	2.6.3.2.9.4	B2B PIURA E.I.R.L	LUPS B00VA1360V3	UNID	1	\$/800.00	\$/800.00	\$/800.00	Iniciales
					2024	2340	394	2.6.3.2.3.1	B2B PIURA E.I.R.L	Servidor-Inel Xeon Silver 430972 800GHZ	UNID	1	\$/33,984.00	\$/33,984.00	\$/33,984.00	Iniciales
					2024	4112	379	2.6.3.2.3.1	B2B PIURA E.I.R.L	LAPTOP_CORE I9 CPU_CORE I9	UNID	2	\$/8,969.40	\$/17,938.80	\$/17,938.80	Iniciales
					2024	5617	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	7008	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	12173	363	2.6.3.2.9.4	B2B PIURA E.I.R.L	LUPS B00VA1360V3	UNID	1	\$/800.00	\$/800.00	\$/800.00	Iniciales
					2024	12406	394	2.6.3.2.3.1	B2B PIURA E.I.R.L	Servidor-Inel Xeon Silver 430972 800GHZ	UNID	1	\$/33,984.00	\$/33,984.00	\$/33,984.00	Iniciales
					2024	12286	379	2.6.3.2.3.1	B2B PIURA E.I.R.L	LAPTOP_CORE I9 CPU_CORE I9	UNID	2	\$/8,969.40	\$/17,938.80	\$/17,938.80	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales
					2024	14389	502	2.6.3.2.3.1	EMCONORT S.R.L	TABLET	UNID	6	\$/9,899.80	\$/5,939.40	\$/5,939.40	Iniciales</

## **VI. Problemas y Desviaciones del plan original**

Durante la ejecución del proyecto, se identificaron algunas desviaciones respecto al plan original, principalmente asociadas a factores administrativos externos y a oportunidades de mejora en el proceso analítico.

En relación con el componente de capacitación, se presentó un desfase temporal debido a la demora en la emisión de la orden de servicio por parte de la entidad ejecutora. De acuerdo con la documentación administrativa correspondiente, la orden de pago para la capacitación en análisis de bases de datos fue emitida con fecha 22 de abril de 2026, lo que generó un retraso en la implementación oportuna de esta actividad. En consecuencia, la capacitación denominada "Análisis de datos con Python" no pudo ejecutarse en el momento inicialmente previsto en el cronograma del proyecto.

Respecto al procesamiento de la información, es importante señalar que esta actividad fue culminada en su totalidad, alcanzando un nivel de ejecución del 100%. No obstante, se reconoce que la incorporación oportuna de la capacitación mencionada habría permitido optimizar aún más los procedimientos de análisis, particularmente en términos de eficiencia, automatización y profundidad analítica.

En cuanto a la elaboración de artículos científicos, se ha evidenciado una extensión en los tiempos inicialmente programados, debido a los procesos de revisión técnica, metodológica y adecuación a estándares de publicación en revistas indexadas. Esta situación ha implicado que los manuscritos aún no hayan sido sometidos a evaluación externa.

En términos generales, las desviaciones identificadas no han comprometido el cumplimiento de los objetivos fundamentales del proyecto, aunque sí han generado ajustes en los tiempos de ejecución y en la secuencia de algunas actividades. Asimismo, han permitido identificar oportunidades de mejora orientadas al fortalecimiento de capacidades técnicas y a la optimización de los procesos analíticos.

## **VII. Conclusiones**

El proyecto de implementación del sistema digital automatizado para la medición en tiempo real de parámetros sociodemográficos en la Región Piura ha alcanzado satisfactoriamente sus objetivos en términos de recolección, procesamiento y análisis de la información. La aplicación del sistema permitió obtener datos consistentes y de alta calidad, lo que se refleja en la generación de indicadores demográficos confiables correspondientes a la muestra de 365 hogares en las diez zonas rurales de estudio.

El procesamiento de la información se ejecutó en su totalidad, alcanzando un nivel de cumplimiento del 100%, lo que evidencia la operatividad y eficiencia del sistema digital automatizado en contextos reales. Asimismo, se logró consolidar una base de datos estructurada y validada, apta para el análisis estadístico y la producción de evidencia científica.

En relación con la producción académica, se ha avanzado significativamente en la elaboración de artículos científicos basados en los resultados obtenidos; sin embargo, estos aún se encuentran en proceso de revisión y adecuación a estándares editoriales, previo a su envío a revistas indexadas.

Por otro lado, se identificaron limitaciones de carácter administrativo que generaron retrasos en la implementación de la capacitación especializada en análisis de datos. No obstante, esta actividad se encuentra actualmente en ejecución y contribuirá al fortalecimiento de las capacidades técnicas del equipo de investigación, permitiendo optimizar futuros procesos analíticos.

En términos generales, el proyecto demuestra la viabilidad, pertinencia y eficiencia del uso de sistemas digitales automatizados para la recolección y análisis de información sociodemográfica en zonas rurales, constituyéndose en una herramienta innovadora que puede ser replicada en otros contextos de investigación.

## IX. Anexos

[Los anexos deben incluir órdenes de compra, órdenes de servicio, boletas de compra, factura, recibos u Otros documentos relacionados con los gastos realizados en el proyecto, actas de conformidad de recepción de bienes y servicios, hasta la fecha. De ser posible incluir gráficos, tablas de datos, fotografías, imágenes que ilustren el estado actual del proyecto].

Fecha: 30 de abril de 2026



---

**Dr. Ricardo Antonio Armas Juárez Ing.**  
DNI 42238686  
**Investigador principal**



---

**Diana Evelyn Hidalgo Eche.**  
DNI 70575452  
**Estudiante**



---

**Gabriela Ipanaque Cárdenas.**  
DNI 75167467  
**Estudiante**

**UNIVERSIDAD NACIONAL DE PIURA  
FACULTAD DE CIENCIAS**



**PROYECTO DE INVESTIGACIÓN**

**IMPLEMENTACIÓN DE UN SISTEMA DIGITAL AUTOMATIZADO PARA  
MEDIR PARÁMETROS SOCIODEMOGRÁFICOS EN TIEMPO REAL EN LA  
REGIÓN PIURA**

**INFORME FINAL**

**PRESENTADO POR**

**Dr. Ricardo Antonio Armas Juárez (Investigador Principal)  
Diana Evelyn Hidalgo Eche. (Estudiante)  
Gabriela Ipanaque Cárdenas. (Estudiante)**

**Piura, Perú, 2026**

## Piura, Perú, 2026

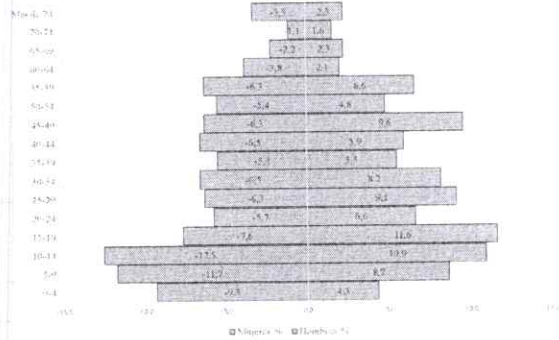
### ANÁLISIS DE ÍNDICES DEMOGRÁFICOS

#### INDICADORES DE LA POBLACION

##### PIRAMIDE TOBLACIONAL

Intervalo de edad	Hombres		Mujeres	
	Cantidad	%	cantidad	%
0-4	34	-9,3	19	4,3
5-9	43	-11,7	28	6,7
10-14	46	-12,5	48	10,9
15-19	28	-7,6	51	11,6
20-24	21	-5,7	29	6,6
25-29	23	-6,1	40	9,1
30-34	24	-6,5	36	8,2
35-39	20	-5,4	24	5,5
40-44	24	-6,5	26	5,9
45-49	23	-6,1	42	9,6
50-54	20	-5,4	21	4,8
55-59	23	-6,1	29	6,6
60-64	14	-3,8	9	2,1
65-69	8	-2,2	10	2,3
70-74	4	-1,1	7	1,6
May de 74	12	-3,3	10	2,3
<b>TOTAL</b>	<b>367</b>	<b>-100</b>	<b>439</b>	<b>100</b>

Pirámide poblacional del distrito de Pacaipampa,



La pirámide poblacional del distrito de Pacaipampa presenta una estructura expansiva, caracterizada por una base ancha y una cúspide estrecha. Esto indica una alta proporción de población joven (niños y adolescentes) y una baja proporción de adultos mayores, lo que sugiere una población en crecimiento con altas tasas de natalidad.

**INDICE DE MASCULINIDAD**

$$\text{Relacion hombres - mujeres} = \frac{\text{numero de hombres}}{\text{numero de mujeres}} * 100$$

IM= 83,60

En el distrito de Pacaipampa, el índice de masculinidad es de 83.6, lo que significa que por cada 100 mujeres hay aproximadamente 84 hombres. Esto indica que las mujeres son mayoría en la población encuestada de este distrito

**INDICE DE ENVEJECIMIENTO**

$$IV = \frac{\text{Suma de la Poblacion de 65 a mas años}}{\text{Suma de la Poblacion de 0 a 15 años}} * 100$$

IV= 22,37

El índice de envejecimiento es 22.37, lo que significa que hay aproximadamente 22 personas adultas mayores por cada 100 niños y niñas. Este resultado indica que la población aún es joven, pero ya se observa una tendencia de envejecimiento demográfico.

**TASA DE DEPENDENCIA**

$$TDI = \frac{\text{suma de personas entre 0 y 14 años} + \text{suma de personas de 65 a mas}}{\text{suma de personas entre 15 a 64 años}}$$

TDI= 52,94

Por cada 100 personas en edad de trabajar (15 a 64 años), hay aproximadamente 53 personas dependientes (niños o adultos mayores) que necesitan ser sostenidas.

**INDICADORES RELACIONADOS A LA FECUNDIDAD**

**TASA BRUTA DE NATALIDAD**

$$\text{Tasa bruta de natalidad} = \frac{\text{Numero total de nacimientos en un año}}{\text{Total de poblacion}} * 100$$

TBN= 1,86

La tasa de natalidad en Pacaipampa es moderada, con aproximadamente 2 (1.86) nacimientos por cada 100 habitantes

**TASA ESPECIFICA DE FECUNDIDAD POR EDAD**

EDAD	Nacimientos	5fx
15-19	1	1,96
20-24	1	3,45
25-29	4	10,00
30-34	3	8,33
35-39	2	8,33
40-44	3	11,54
45-49	1	2,38
TOTAL	15	

La distribución de los nacimientos según grupo de edad muestra que la mayor concentración de nacimientos en los últimos 12 meses se dio en los grupos de edad de 25-29 años, con 4 nacimientos, seguidos por los grupos de 30-34 y 40-44 años con 3 nacimientos

TASA DE FECUNDIDAD GENERAL.

$$TFG = \frac{\text{Nacimientos total de nacimientos en un año}}{\text{Poblacion femenina(15 a 49 años)}} \cdot 100$$

TFG= 6,05

La Tasa General de Fecundidad en Pacaipampa es de 6 nacimientos por cada 100 mujeres en edad fértil

EDAD MEDIA DE LA FECUNDIDAD

EDAD	X	5fx	
15-19	17,5	0,0196	0,3
20-24	22,5	0,0345	0,8
25-29	27,5	0,1000	2,8
30-34	32,5	0,0833	2,7
35-39	37,5	0,0833	3,1
40-44	42,5	0,1154	4,9
45-49	47,5	0,0238	1,1
TOTAL		0,46	15,7

$$EMF = \frac{\sum(sX_x + sf_x)}{\sum sf_x} \quad EMF= 34,21$$

Este valor quiere decir que si las mujeres en edad fértil de Pacaipampa, tuvieran todos sus hijos a una sola edad, lo harian, en promedio, a los 34.2 años exactos.

TASA GLOBAL DE FECUNDIDAD

$$TGF = 5 \cdot \sum sf_x$$

TGF= 2,2998

La Tasa Global de Fecundidad (TGF) en Pacaipampa es de 2 (2.2998) hijos por mujer. Esto significa que, en promedio, cada mujer tendria aproximadamente dos hijos durante su vida reproductiva, si las condiciones actuales de fecundidad se mantuvieran constantes.

TASA BRUTA DE REPRODUCCION

$$TBR = 0,4878 \cdot TGF$$

TBR= 1,1218

En Pacaipampa se presentó un promedio de 1 (1.1218) hija por mujer durante su período fértil, suponiendo que no están expuestas a la mortalidad y que durante todo este período su patrón reproductivo se mantendrá constante.

INDICADORES RELACIONADOS A LA MORTALIDAD

TASA BRUTA DE MORTALIDAD

$$d^t = \frac{D^t}{N^t} \cdot 1000$$

TBM= 19,85

La Tasa Bruta de Mortalidad fue de 19.85 por cada 1,000 habitantes durante el último año. Esto implica que, por cada 1000 personas, fallecieron alrededor de 19 personas.

Los resultados presentados corresponden a los indicadores demográficos estimados para la zona rural de Pacaipampa, los cuales han sido obtenidos a partir del procesamiento y análisis de la información recolectada mediante el sistema digital automatizado. Estos indicadores comprenden variables fundamentales como características de la vivienda, estructura poblacional, composición por edad y sexo, así como niveles de fecundidad, entre otros aspectos relevantes para el análisis sociodemográfico.

Es importante señalar que la misma metodología de recolección, depuración y análisis de datos fue aplicada de manera estandarizada en las demás zonas rurales consideradas en el estudio. En este sentido, los procedimientos utilizados en Pacaipampa fueron replicados en las otras nueve zonas rurales de la Región Piura —Frias, Montero, Canchaque, San Miguel del Faique, Sondorillo, El Tallán, Cura Mori, Santo Domingo y Buenos Aires— garantizando la comparabilidad, consistencia y validez de los resultados obtenidos.

De esta manera, los indicadores generados no solo permiten caracterizar la realidad sociodemográfica de cada zona en particular, sino también realizar análisis comparativos entre territorios, identificar patrones y tendencias, y aportar evidencia empírica sólida para la toma de decisiones en políticas públicas y planificación territorial en contextos rurales.

## REDACCION DEL PRIMER ARTICULO CIENTIFICO

### **Vivienda, estructura poblacional y fecundidad en diez zonas rurales de Piura, 2025: Un análisis sociodemográfico integral.**

Housing, population structure and fertility in ten rural areas of Piura, 2025: A comprehensive sociodemographic analysis.

Gabriela Ipanaqué Cárdenas,

Ricardo Antonio Armas Juárez,

Diana Evelyn Hidalgo Eche

*Universidad Nacional de Piura, Perú*

#### *Resumen*

El presente estudio analiza las características sociodemográficas de diez zonas rurales de la ciudad de Piura, con el propósito de identificar similitudes y diferencias entre estas localidades, así como de contrastarlas con los resultados obtenidos en el Censo Nacional de Población y Vivienda 2017. La información utilizada en esta investigación fue recolectada mediante una encuesta aplicada a través de un sistema digital automatizado, lo que permitió garantizar la precisión y eficiencia en la captura de datos. Este análisis busca proporcionar una visión detallada de la estructura poblacional, los patrones de vivienda, la educación, el acceso a servicios básicos y otros indicadores sociodemográficos, contribuyendo así a una comprensión más profunda de la realidad de las zonas rurales de Piura y sus posibles implicancias para la planificación y formulación de políticas públicas.

*Palabras claves:* Estructura poblacional, indicadores de fecundidad, zonas rurales.

#### *Abstract*

This study analyzes the sociodemographic characteristics of ten rural areas in the city of Piura, with the aim of identifying similarities and differences among these localities and comparing them with the results obtained in the 2017 National Population and Housing Census. The information used in this research was collected through a survey administered via an automated digital system, ensuring accuracy and efficiency in data collection. This analysis seeks to provide a detailed view of the population structure, housing patterns, education, access to basic services, and other sociodemographic indicators, thus contributing to a deeper understanding of the realities of rural areas in Piura and their potential implications for planning and formulating public policies.

*Keywords:* Population structure, fertility indicators, rural areas.

## **INTRODUCCION**

Los indicadores sociodemográficos que abarcan dimensiones como la salud, la educación, la vivienda, la economía y la estructura familiar, sintetizan los principales aspectos del bienestar colectivo y la organización social, siendo pilares fundamentales para el desarrollo sostenible de una comunidad (Cedeño et al, 2019). Estos indicadores permiten conocer la realidad de la población y sus condiciones de vida, ofreciendo una visión integral de los factores que influyen en el desarrollo humano y territorial.

Para poder conocer y analizar los indicadores sociodemográficos, es necesario contar con información confiable obtenida a través de encuestas o registros sistemáticos. Tradicionalmente, estas encuestas se realizaban de manera presencial o manual; sin embargo, en los últimos años los avances tecnológicos han transformado significativamente este proceso, integrando herramientas digitales y sistemas automatizados que optimizan el levantamiento y la validación de los datos. Este progreso, impulsado por el mayor acceso a Internet y a dispositivos tecnológicos, ha permitido aplicar cuestionarios de forma remota y en tiempo real, haciendo más ágil, precisa y segura la obtención de información mediante funciones como validaciones automáticas y georeferenciación de viviendas (Lobe et al, 2020; Lugo & Pérez, 2025; Morán, 2025).

En ese contexto, se desarrolló el proyecto "Implementación de un sistema digital automatizado para medir parámetros sociodemográficos en la región Piura", cuyo objetivo principal fue diseñar e implementar una plataforma tecnológica capaz de recopilar, procesar y analizar información demográfica y social de manera automatizada. Esta innovación permitió obtener una base de datos que facilitó la comparación de indicadores sociodemográficos entre diez zonas de la ciudad de Piura, generando un panorama integral sobre las características poblacionales y las condiciones de vida.

Por lo tanto, el objetivo del presente estudio es describir los principales indicadores sociodemográficos obtenidos mediante el sistema digital automatizado en diez zonas rurales de la ciudad de Piura, identificando las diferencias, similitudes y patrones estructurales que caracterizan a cada una, y demostrando la utilidad del sistema como herramienta para el análisis de información sociodemográfica.

## **ANTECEDENTES TEORICOS Y EMPIRICOS**

### **Demografía y población: una mirada conceptual**

Según Hauser y Duncan (1962), la demografía se ocupa de analizar el tamaño, la distribución y la estructura de la población, así como las variaciones que experimentan estos elementos y los factores que las originan, entre ellos la movilidad social, las migraciones, la mortalidad y la fecundidad. (Arquimigo, 2022). Aunque su base es esencialmente estadística, su propósito va más allá de la simple descripción numérica: busca comprender cómo se desarrolla la vida humana dentro de una sociedad específica, influida por factores familiares, económicos, sociales, culturales y ambientales. (Comisión Económica para América Latina y el Caribe [CEPAL], 2014). Comprender estos procesos es fundamental para interpretar la dinámica poblacional y prever sus efectos en distintos sectores de la sociedad. La información demográfica permite identificar tendencias, anticipar necesidades y analizar cómo los cambios en la población se relacionan con fenómenos económicos y sociales, facilitando así la formulación de políticas y acciones orientadas al desarrollo. (Instituto nacional de estadística y censo [INEC], 2000)

En este sentido, la población constituye el eje fundamental de la demografía, pues es a partir de sus características que se pueden comprender los procesos sociales que la afectan y transforman. Por ello, como menciona Rupert B. Vance, "el desarrollo de un concepto científico de la población es la primera tarea de la demografía". En sus orígenes etimológicos, el término *populatio* hacía referencia a acciones como "saqueo" o "devastación", un sentido muy distinto al actual. Con el tiempo, adquirió nuevos significados: por un lado, aludió a los habitantes de un territorio; por otro, a la acción de poblar, es decir, a un proceso dinámico y social (Vieira, 1973). Esta doble connotación ha sido la base de su desarrollo conceptual contemporáneo, llegando a definirse por la Real Academia Española (RAE) como "el conjunto de personas que habitan en un lugar determinado"

Para Miro (2007), la población dentro de la demografía se divide en tres grupos. El primer grupo lo conforman los adolescentes y adultos jóvenes, los cuales necesitan políticas orientadas a sus necesidades ya que se trata de una población expuesta a múltiples vulnerabilidades. El segundo grupo lo conforma la población en edad de trabajar, que en teoría podría considerarse, como señalan algunos autores, un "bono demográfico". Sin embargo, en contextos donde prevalecen altos niveles de desempleo y subempleo, esta potencial ventaja no se materializa. Por el contrario, el crecimiento de este grupo termina generando más problemas, como la subutilización de la fuerza laboral, la reducción de aportes a los sistemas de seguridad social y el agravamiento de la

precariedad económica. Y por último tenemos el grupo de los Adultos mayores, cuya creciente presencia en aumento representa un desafío significativo, pues se trata de una población con mayor dependencia y necesidad específicas de atención.

Teniendo en cuenta los siguientes contextos, para poder comprender a la población y sus distintos grupos requiere no solo describir sus dinámicas, sino también contar con información que permita retratar su situación en un momento específico. En este sentido, las estadísticas de población ofrecen una visión detallada de sus características y composición, a partir de datos obtenidos mediante censos y encuestas (Arquimigo, 2022). Bajo este enfoque, en el presente estudio se aplicó una encuesta que permitió reunir la información necesaria para caracterizar a la población analizada.

#### **Hallazgos previos sobre la estructura y comportamiento de la poblacional en el ámbito rural**

El análisis de los indicadores demográficos permite comprender no solo cómo está compuesta una población, su objetivo es proporcionar las principales características, la dinámica y la evolución de los fenómenos demográficos (Instituto Nacional de Estadística [INE], 2024).

En este sentido, tenemos a los indicadores de estructura poblacional, los cuales tienen el objetivo de facilitar el análisis y comprender cómo se organizan las poblaciones según el sexo y los distintos grupos de edad. (Departamento Administrativo Nacional de Estadística [DANE], s.f.). Entre los más utilizados se encuentran la relación de masculinidad, la cual es necesaria analizar dado el comportamiento diferencial que presenta hombres y mujeres tanto desde el punto de vista demográfico como desde el económico y social. (Bel & Gomez, 1998). También tenemos el índice de envejecimiento, la relación de dependencia demográfica y la edad media de la población.

La vivienda constituye otro eje clave para interpretar las condiciones de vida, así como la accesibilidad y calidad de los servicios básicos, para medirlos, se tienen en cuenta varios factores como la disponibilidad de una variedad de servicios de saneamiento, la fuente de agua destinada al consumo humano y el material de construcción de la vivienda. (Arquimigo, 2022)

Según el Censo de Población y Vivienda (Instituto Nacional de Estadística e Informática [INEI], 2018), en las zonas rurales del Perú, el material predominante en las paredes es el adobe o tapia (69,6%), seguido por la madera (13,9%). En cuanto a los pisos,

la mayoría de las viviendas cuenta con piso de tierra (78,8%), mientras que el piso de cemento alcanza el 15,6% de los hogares. Por otro lado, los techos de las viviendas rurales son mayormente de planchas de calamina, seguidos por las tejas (20,7%). En cuanto a la accesibilidad y calidad de los servicios básicos, el abastecimiento de agua proviene principalmente de la red pública (36,1%) y de fuentes naturales como ríos o manantiales (33,4%), mientras que otros hogares usan pozo, pilón o camión cisterna. Respecto a los servicios higiénicos, las viviendas con desagüe por red pública experimentaron un incremento de 194,1% entre 2007 y 2017, del mismo modo el alumbrado eléctrico, se registró un crecimiento de 179,2%, en el mismo periodo.

Otras de las variables fundamentales en el análisis demográfico es el estado conyugal, pues conocer la situación marital aporta información clave sobre las redes de apoyo disponibles en caso de aumento dependencia (Del Barrio & Abellan, 2006). De manera complementaria, el tipo de seguro de salud permite identificar el nivel de cobertura y protección sanitaria con el que cuenta la población, mientras que los indicadores de alfabetismo y grado de instrucción ofrecen una aproximación a las capacidades educativas y a las brechas existentes entre zonas rurales. Asimismo, la religión y la autoidentificación étnica permiten reconocer la diversidad cultural presente en el territorio y los valores que influyen en las prácticas sociales, comunitarias y familiares de la población.

Para Dolores (2017), los flujos de nacimientos es otro factor que interviene en la estructura de la población, y trae consigo dos fenómenos fundamentales, la natalidad, que mide la frecuencia con que se produce el suceso en el conjunto de una población y la fecundidad que mide la frecuencia del suceso en relación con las mujeres en edad fértil. A partir de estos dos conceptos se construyen indicadores tales como la tasa bruta de natalidad, la tasa de fecundidad general, la tasa global de fecundidad, la tasa bruta de natalidad y la edad media a la maternidad. Respecto a estos se encontró que en el año 2017, el promedio de hijos por mujer en las zonas rurales estudiadas es 1,7, lo que confirma una disminución de la fecundidad respecto al Censo 2007. Esto indica que las mujeres están teniendo familias más pequeñas, un patrón asociado al avance de la transición demográfica, mayor acceso a métodos anticonceptivos y cambios en las condiciones sociales y educativas. Hay que mencionar, además, que existen diferencias internas importantes. Por ejemplo, Ayabaca presenta el promedio más alto (2,5 hijos por mujer), lo que refleja una fecundidad rural todavía elevada. En contraste, Piura y Sullana, que no cuentan con una proporción tan amplia de zonas rurales, registran los valores más

bajos (1,6 hijos por mujer), evidenciando patrones reproductivos más cercanos a los observados en áreas urbanas. (INEL, 2018)

#### MATERIAL Y MÉTODO

El estudio se desarrolló bajo un enfoque descriptivo, cuyo propósito es caracterizar con precisión los atributos y perfiles de la población de interés. De acuerdo con Hernández et al (2014), este tipo de estudios, "únicamente pretenden medir o recoger información de manera independiente o conjunta sobre los conceptos o las variables a las que se refiere, su objetivo no es indicar cómo se relacionan éstas". Bajo este enfoque, el presente estudio describe y compara los principales indicadores sociodemográficos de los hogares pertenecientes a las zonas rurales evaluadas.

El tamaño de la muestra inicialmente se determinó utilizando la fórmula para poblaciones finitas.

$$n = \frac{N \cdot Z^2 \cdot p(1-p)}{e^2(N-1) + Z^2 \cdot p(1-p)}$$

Para el cálculo se consideró una población total de 7158 viviendas rurales, que corresponde a la suma de las viviendas de las 10 zonas rurales, un nivel de confianza del 95% ( $Z = 1.96$ ), una proporción esperada de  $p = 0.50$  y un margen de error permitido de 0.05, dando como resultado un tamaño muestral aproximado de 365 viviendas rurales.

$$n_{total} = \frac{7158 \cdot (1.96)^2 \cdot 0.50(1-0.50)}{0.05^2(7158-1) + 1.96^2 \cdot 0.50(1-0.50)} = 365$$

Además se empleó un muestreo estratificado, decisión que se sustenta tanto en criterios teóricos como operativos. Según Portela y Villeta (2007), uno de los principales motivos para utilizar este diseño muestral es "la necesidad de obtener estimaciones separadas para cada subpoblación o estrato". En coherencia con ello, cada una de las diez zonas rurales del estudio fue considerada como un estrato independiente, permitiendo obtener estimaciones más precisas y comparables entre áreas. La asignación del tamaño muestral se realizó mediante afijación proporcional, método que, de acuerdo con Pérez López (2005), "consiste en asignar a cada estrato un número de unidades muestrales proporcional a su tamaño; las  $n$  unidades de la muestra se distribuyen en función del tamaño de cada estrato expresado en número de unidades". La fórmula usada para hallar el tamaño de la muestra ( $n_i$ ) para cada una de las zonas rurales  $i$  fue:

$$n_i = \left(\frac{N_i}{N}\right) * n_{total}$$

Donde  $N_i$  es el número de viviendas en la zona rural  $i$ ,  $N = 7158$  y  $n_{total} = 365$ , el tamaño de muestra total. Obteniendo así la muestra que se observa en la tabla 1.

Tabla 1: Muestra inicial para cada una de las zonas rurales

Nº Zona	Zonas rurales	Número de viviendas ocupadas	Proporción	Muestra
1	Pacainpampa	353	4.93%	18
2	Frias	614	8.58%	31
3	Montero	334	4.67%	17
4	Canchaque	415	5.80%	21
5	San Miguel del Frique	333	4.65%	17
6	Sondorillo	165	2.31%	8
7	El Tallan	286	4.00%	15
8	Las Lomas	3161	44.16%	161
9	Santo Domingo	412	5.76%	21
10	Buenos Aires	1085	15.16%	55
	TOTAL	7158	100%	365.00

Fuente: Elaboración propia en base a los resultados del Censo 2017

No obstante, durante la fase de ejecución del trabajo de campo se presentaron limitaciones en la zona rural de Las Lomas, inicialmente considerada en el diseño muestral. Estas limitaciones estuvieron asociadas a dificultades en el acceso a las viviendas y a la disponibilidad efectiva de unidades de observación que cumplieran con los criterios definidos para el estudio. En este contexto, se decidió sustituir dicha zona por la zona rural de Cura Mori, la cual presenta características sociodemográficas comparables y pertenece al mismo distrito de Piura.

Este ajuste se produjo cuando las encuestas correspondientes a las cinco primeras zonas rurales ya habían sido aplicadas, por lo que no resultó metodológicamente viable modificar el tamaño muestral previamente ejecutado.

Ante esta situación, se optó por realizar un remuestreo parcial, manteniendo fijo el tamaño muestral ya levantado y conservando el tamaño total de la muestra ( $n = 365$ ). En consecuencia, las encuestas correspondientes a las cinco primeras zonas rurales se

mantuvieron sin alteración, mientras que la fracción muestral restante fue redistribuida entre las zonas aún no encuestadas, incorporando a la nueva zona rural de Cura Mori.

Considerando que el tamaño muestral ya ejecutado en las cinco primeras zonas rurales fue de 104 encuestas, la muestra restante por asignar se determinó como:  $n_{i total}^* = 365 - 104 = 261$ . La redistribución de la muestra restante se realizó aplicando nuevamente el criterio de asignación proporcional al tamaño del estrato, considerando únicamente las zonas rurales pendientes de encuesta. Para cada zona rural  $i$  el tamaño muestral ajustado ( $n_i^*$ ) se obtuvo mediante la siguiente expresión

$$n_i^* = \left(\frac{N_i}{N^*}\right) * n_{i total}^*$$

Donde  $N_i$  es el número de viviendas la zona rural  $i$ ,  $N^*$  corresponde al total de viviendas ocupadas en las zonas rurales restantes y  $n_{i total}^* = 261$  el tamaño de muestral pendiente de asignación. Como resultado de este procedimiento, se obtuvo la distribución muestral que se presenta en la Tabla 2.

Tabla 2: Muestra de las cinco zonas rurales restantes.

	Zonas rurales	Número de viviendas ocupadas	Proporción	Muestra
6	Sondorillo	165	5.81%	15
7	El Tallan	286	10.07%	26
8	Cura Mori	893	31.43%	82
9	Santo Domingo	412	14.50%	38
10	Buenos Aires	1085	38.19%	100
	TOTAL	2841	100%	261.00

Fuente: Elaboración propia en base a los resultados del Censo 2017

Finalmente, dentro de cada estrato, los hogares fueron seleccionados mediante muestreo por conveniencia, priorizando aquellos de mayor cercanía y accesibilidad para el equipo de recolección. Este procedimiento facilitó la conformación de una base de datos viable y metodológicamente comparable entre las áreas analizadas.

La recolección de información se realizó mediante un sistema digital automatizado que integró un cuestionario orientado a registrar las principales dimensiones sociodemográficas. Para este fin, se diseñó un instrumento de recolección de datos basado en el cuestionario censal utilizado por el INEI, incorporando preguntas precisas y adaptadas a la realidad de las zonas rurales de Piura. Dicho cuestionario fue

estructurado en 5 secciones las cuales fueron: Características y servicios de la vivienda, características del hogar, características de la población, características de fecundidad y características de mortalidad, permitiendo obtener una visión integral de la dinámica social y demográfica de cada zona estudiada.

El procedimiento consistió en la programación del cuestionario en la plataforma digital, la capacitación del equipo encuestador y la aplicación del instrumento mediante dispositivos móviles en cada uno de los estratos definidos. Los datos recolectados fueron posteriormente depurados y verificados con el fin de asegurar su consistencia y evitar registros incompletos o inconsistentes.

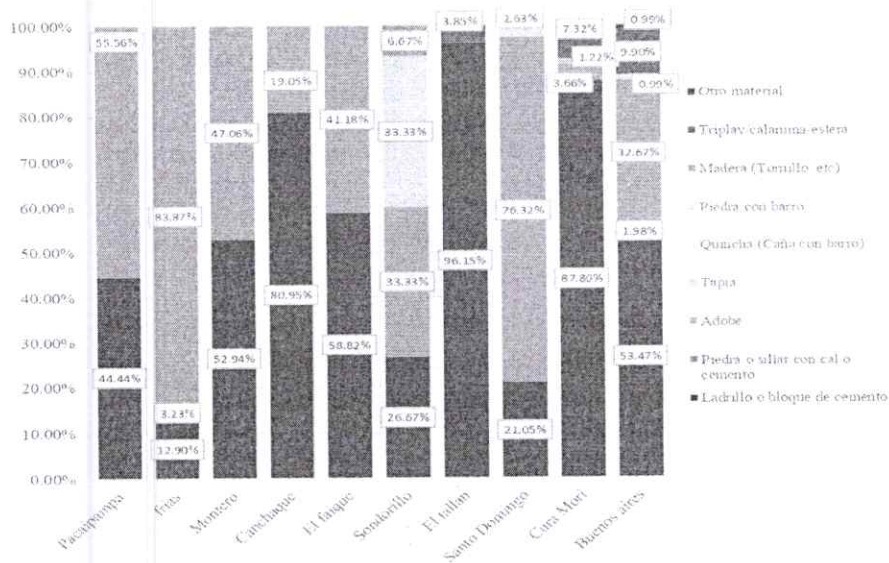
El análisis de la información se llevó a cabo utilizando Excel y el software estadístico SPSS, aplicando medidas descriptivas como frecuencias, porcentajes, promedios y gráficos comparativos. Asimismo, se elaboraron tabulados para contrastar las diferencias entre estratos y analizar patrones sociodemográficos relevantes.

## ANÁLISIS DE LOS RESULTADOS

### Indicadores de Calidad de Vivienda Basados en Materiales Constructivos

En la figura 1 se observa los porcentajes de los materiales predominantes en las paredes exteriores de las viviendas de cada una de las zonas rurales estudiadas, se observa que en zonas como El Tallán (96,15%), Cura Mori (87,8%) y Canchaque (80,95%), presentan un claro predominio de construcciones con ladrillo o bloque de cemento, lo que refleja mejor acceso a materiales industrializados y mayores niveles de inversión familiar en infraestructura. En contraste, otras zonas como Frías (83,87%), Pacaipampa (55,56%), Montero (47,06%) y Santo Domingo (76,32%) muestran un uso considerable de adobe, tapia o quincha, característicos de contextos rurales con recursos más limitados.

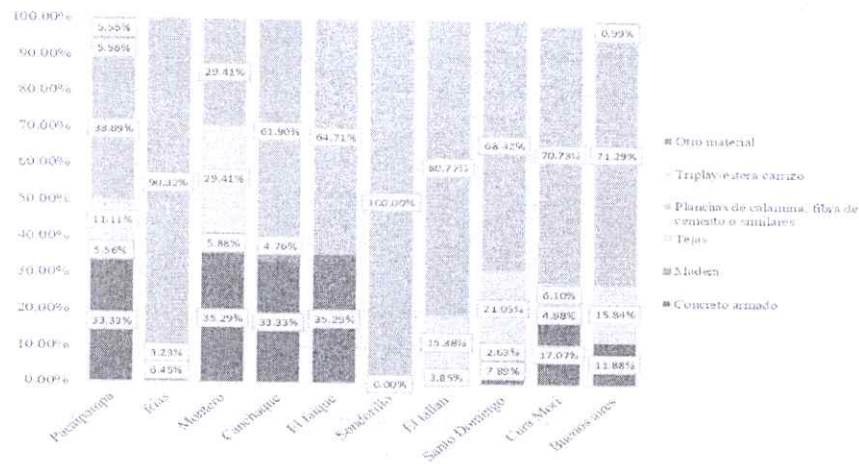
Figura 1. Material predominante en las paredes exteriores por zona rural



Fuente: Elaboración propia con la data obtenida de la encuesta realizada

En la figura 2 se observa el material predominante en los techos, donde el material que predomina es la calamina o fibra de cemento, superando el 60% en casi todas las localidades y alcanzando incluso el 100% en Sondorillo. Si bien este material es económico y de fácil instalación, presenta limitaciones en aislamiento térmico y acústico. Pese a ello, algunas zonas como Canchaque (33,33%), Pacaipampa (33,33%), Montero (35,29%) y El Faique (35,29%) registran un porcentaje importante de techos de concreto armado, lo que constituye un indicador de mayor inversión estructural y mejores condiciones de seguridad. Asimismo, la presencia de tejas en localidades como Buenos Aires, Montero y Santo Domingo refleja estilos tradicionales propios de zonas altoandinas y rurales.

Figura 2. Material predominante en los techos por zona rural



Fuente: Elaboración propia con la data obtenido de la encuesta realizada

En relación con los pisos, se observa en la Figura 3 que existe una tendencia general al uso de cemento, especialmente en Canchaque (76,19%), Pacaipampa (72,22%), El Tallán (73,08%) y El Faique (70,59%), lo que demuestra un nivel básico, pero aceptable, de consolidación estructural. No obstante, el uso de pisos de tierra, es aún significativo en varias zonas, como Sondorillo (46,67%), Santo Domingo (34,21%), Cura Mori (30,49%), Buenos Aires (26,73%), y El Faique (23,53%), la presencia de este tipo de piso considerado un indicador de pobreza y vulnerabilidad sanitaria, implica mayor exposición a humedad, polvo, parásitos e infecciones respiratorias, lo que evidencia limitaciones económicas en estas zonas.

Figura 3. Material predominante en los pisos por zona rural



Fuente: Elaboración propia en el software SPSS

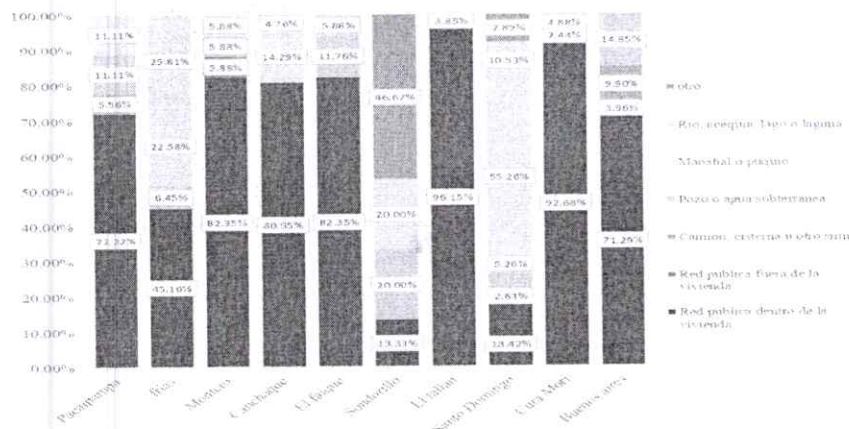
En conjunto, estos patrones reflejan diferencias marcadas en la calidad habitacional entre las zonas analizadas, donde algunas muestran avances hacia viviendas más sólidas y seguras, mientras que otras que son elaboradas con materiales rústicos suelen carecer de un adecuado aislamiento, lo que permite el ingreso de aire frío y la aparición de puentes térmicos en diversas partes de la estructura. Esto genera ambientes interiores fríos que, junto con hábitos de ventilación poco controlados, incrementan el riesgo de infecciones respiratorias, especialmente en niños y adultos mayores. (Tito, 2018)

#### Entre la Cobertura y la Carencia: Condiciones de Acceso a Agua, Saneamiento, Electricidad y combustible

El acceso a los servicios básicos como el agua revela profundas diferencias territoriales. En la Figura 4 se puede observar que en las zonas mejor abastecidas, como El Tallán (96,15%), Cura Mori (92,68%), El Faique (82,35%), Canchaque (80,95%) y Buenos Aires (71,29%), predomina el suministro de agua por red pública dentro de las viviendas.

indicando una adecuada cobertura formal del servicio. En contraste, las zonas altoandinas dependen en mayor medida de fuentes naturales como manantiales, ríos o acequias, como ocurre en Santo Domingo (55,26% manantial). Frías (25,81% río/acequia; 22,58% manantial), Sondorillo (46,67% "otro") y Pacaipampa (11,11% manantial). Este patrón revela una limitada cobertura de red pública, lo que obliga a muchas familias a desplazarse largas distancias para obtener agua, frecuentemente en condiciones poco seguras. Para Saavedra (2018) esta situación se asocia a mayores niveles de vulnerabilidad, ya que el acceso informal al recurso expone a la población a riesgos sanitarios, tensiones comunitarias y dificultades que inciden directamente en la calidad de vida de los habitantes

Figura 4. Servicio de agua por zona rural

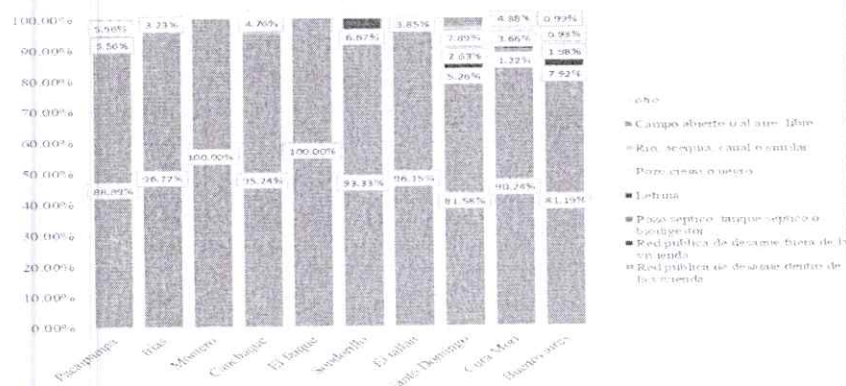


Fuente: Elaboración propia en el programa SPSS

En cuanto al saneamiento, la mayoría de territorios poseen conexión a la red pública de desagüe dentro de la vivienda, destacando El Faique y Montero con un 100% de cobertura, así como Frías (96,77%), El Tallan (96,15%) y Canchaque (95,24%).

reflejando mejores condiciones estructurales y un mayor nivel de urbanización. Sin embargo, persisten prácticas de disposición inadecuada de excretas, como en Sondorillo (7,89%) y Canchaque (4,76%) donde utiliza ríos o acequias, representando un riesgo ambiental severo, mientras que en Buenos Aires (6,93%) y Cura Mori (4,88%) aún se emplean pozos ciegos o letrinas como alternativas transitorias. (Figura 5)

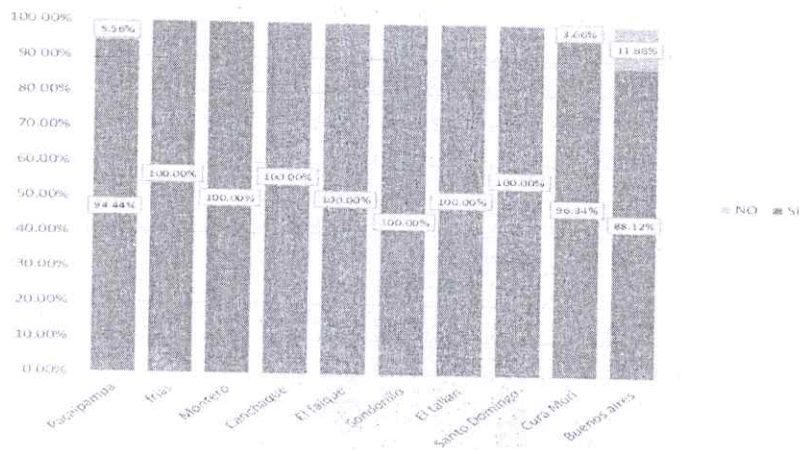
Figura 5. Condición de saneamiento por zona rural



Fuente: Elaboración propia en el programa SPSS

Respecto al servicio de alumbrado eléctrico, la cobertura es casi universal, con valores entre 94% y 100% en la mayoría de zonas; no obstante, Buenos Aires registra un 11,88% sin acceso, lo que evidencia brechas persistentes en áreas urbano-marginales que aún requieren atención. (Figura 6)

Figura 6. Servicio de alumbrado público por zona rural



Fuente: Elaboración propia en el programa SPSS

En la Figura 7 se observa que el uso de combustibles para cocinar muestra un predominio general del gas en balón (GLP), especialmente en Canchaque (90,48%), Pacaipampa (88,9%), Buenos Aires (78,2%) y El Tallán (73,08%), lo que indica un avance hacia fuentes más limpias y seguras. Sin embargo, persiste un uso elevado de leña en varias zonas rurales como Santo Domingo (36,8%), Sondorillo (40%), Montero (35,29%), y Cura Mori (34,15%) reflejando limitaciones económicas, dificultades de acceso a GLP o mantenimiento de prácticas tradicionales. El carbón es poco empleado, con valores que no superan el 12%, y el gas natural prácticamente no está presente, evidenciando ausencia de infraestructura para su distribución.

Figura 7. Energía o combustible usado por zona rural



Fuente: Elaboración propia en el programa SPSS

El análisis comparativo de las diez zonas estudiadas revela una marcada heterogeneidad en las condiciones de habitabilidad, reflejando desigualdades estructurales asociadas a nivel de urbanización, disponibilidad de servicios básicos y capacidad económica de los hogares.

#### Características Esenciales de la Población Rural por Zonas

La población se estudia a través de diversas variables e indicadores, como la edad, la identidad de género, la ocupación, el lugar de nacimiento o el estado civil (Sposob, 2025), lo que permite reconocer las particularidades sociales, culturales y educativas que distinguen a cada zona. Las siguientes tablas comparan estas características entre zonas rurales, destacando similitudes, diferencias y brechas que ayudan a comprender mejor su realidad

Tabla 3: Distribución del tipo de seguro de salud según zona rural

	Buenos Aires	Canchaque	Cura Mori	El Faique	El Tallán	Frias	Montero	Pacapaupa	Santo Domingo	Sondorillo
Seguro integral de salud (SIS)	n 196	30	149	38	46	60	22	37	66	37
	% 76,9%	61,2%	74,9%	74,5%	76,7%	68,2%	64,7%	68,5%	67,3%	78,7%
Es salud	n 45	17	47	11	12	21	7	12	25	6
	% 17,6%	34,7%	23,6%	21,6%	20,0%	23,9%	20,6%	22,2%	25,5%	12,8%
Seguro de fuerzas armadas o policiales	n 0	0	1	0	0	0	0	0	1	1
	% 0,0%	0,0%	0,5%	0,0%	0,0%	0,0%	0,0%	0,0%	1,0%	2,1%
Otro seguro	n 1	1	0	0	0	1	0	1	0	0
	% 0,4%	2,0%	0,0%	0,0%	0,0%	1,1%	0,0%	1,9%	0,0%	0,0%
Ninguno	n 13	1	2	2	2	6	5	4	6	3
	% 5,1%	2,0%	1,0%	3,9%	3,3%	6,8%	14,7%	7,4%	6,1%	6,4%
Total	n 255	49	199	51	60	88	34	54	98	47
	% 100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fuente: Elaboración propia

La tabla 3 presenta la distribución según tipo de seguro de salud y zona rural, se observa que el Seguro Integral de Salud (SIS) es el tipo de aseguramiento predominante en todas ellas, aunque con diferencias relevantes en su distribución. Las zonas rurales con mayor proporción de afiliados al SIS son Sondorillo (78,7%), Buenos Aires (76,9%), y El Tallán (76,7%), evidenciando una fuerte dependencia del sistema público. En cuanto a EsSalud, que constituye la segunda forma de aseguramiento más frecuente, sobresale Canchaque con 34,7%, seguido de Sondor con 25,5%. Por su parte, zonas como Sondorillo (12,8%) y Buenos Aires (17,6%) presentan los niveles más bajos, lo que podría estar asociado a una menor presencia de empleo formal o vínculos laborales con instituciones aseguradoras. Los seguros de Fuerzas Armadas/Policiales y otros seguros mantienen niveles casi nulos en todas las zonas rurales. Otro punto importante a resaltar son las personas sin ningún tipo de seguro, en este grupo Montero alcanza el nivel más alto (14,7%), mientras que la mayoría de zonas mantiene porcentajes bajos, entre 1% y 7%.

Tabla 4: Distribución de la población que sabe leer y escribir según zona rural

	Buenos Aires	Canchaque	Cura Mori	El Faique	El Tallán	Frias	Montero	Pacapaupa	Santo Domingo	Sondorillo
Si	n 218	45	161	49	57	78	31	47	90	39
	% 85,5%	91,8%	80,9%	96,1%	95,0%	88,6%	91,2%	87,0%	91,8%	83,0%
No	n 37	4	38	2	3	10	3	7	8	8
	% 14,5%	8,2%	19,1%	3,9%	5,0%	11,4%	8,8%	13,0%	8,2%	17,0%
Total	n 255	49	199	51	60	88	34	54	98	47
	% 100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fuente: Elaboración propia

La tabla 4 muestra la distribución del alfabetismo en las distintas zonas rurales, diferenciando entre quienes si saben leer y escribir y quienes no. En todos los distritos se observa una clara predominancia de población alfabetizada, aunque la magnitud varía según el territorio.

Los porcentajes más altos de personas que si saben leer y escribir se presentan en El Faique (96,1%), El Tallán (95,0%), Canchaque (91,8%) y Santo Domingo (91,8%). Estos valores reflejan que en estos distritos el alfabetismo está prácticamente consolidado, mostrando una presencia muy reducida de personas con limitaciones educativas básicas.

Por otro lado, distritos como Cura Mori y Sondorillo muestran los porcentajes más elevados de población que no sabe leer y escribir, con 19,1% y 17,0% respectivamente. Aunque la mayoría de la población de estos lugares es alfabetizada, estas proporciones indican que aún existen brechas educativas importantes que podrían requerir intervenciones específicas.

Tabla 5: Distribución del grado de instrucción según zona rural

	Buenos Aires	Canchaque	Cura Mori	El Faique	El Tallán	Frias	Montero	Pacapaupa	Santo Domingo	Sondorillo
Sin nivel	n 25	3	29	1	2	7	4	3	4	6
	% 9,8%	6,1%	14,6%	2,0%	3,3%	8,0%	11,8%	5,6%	4,1%	12,8%
Inicial	n 9	1	10	2	0	1	0	4	3	5
	% 3,5%	2,0%	5,0%	3,9%	0,0%	1,1%	0,0%	7,4%	3,1%	10,6%

Continúa

Tabla 5: Continuación

		Prinoria									
		n	80	6	53	16	20	23	6	17	27
	%	31,4%	12,2%	26,6%	31,4%	33,3%	26,1%	17,6%	31,5%	27,6%	17,0%
Secundaria	n	99	18	84	8	23	22	13	15	33	14
	%	38,8%	46,7%	42,2%	15,7%	38,3%	25,0%	38,2%	27,8%	33,7%	29,8%
Sup. Universitaria incompleta	n	4	4	3	2	0	3	3	2	5	2
	%	1,6%	8,2%	1,5%	3,9%	0,0%	3,4%	8,8%	3,7%	5,1%	4,3%
Sup. No universitaria completa	n	16	7	9	6	7	14	3	5	15	5
	%	6,3%	14,3%	4,5%	11,8%	11,7%	15,9%	8,8%	9,3%	15,3%	10,6%
Sup. Universitaria incompleta	n	6	1	3	5	7	3	1	2	1	3
	%	2,4%	2,0%	1,5%	9,8%	11,7%	3,4%	2,9%	3,7%	1,0%	6,4%
Sup. Universitaria completa	n	15	8	8	11	1	15	4	6	10	4
	%	5,9%	16,3%	4,0%	21,6%	1,7%	17,0%	11,8%	11,1%	10,2%	8,5%
Maestría o posgrado	n	1	1	0	0	0	0	0	0	0	0
	%	0,4%	2,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Total	n	255	49	199	51	60	88	34	54	98	47
	%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fuente: Elaboración propia

La distribución del nivel educativo según zona rural revela un panorama heterogeneo en las diez zonas analizadas. En términos generales, la educación secundaria constituye el nivel formativo más frecuente, con porcentajes especialmente elevados en Cura Mori (42,2 %), Canchaque (36,7 %), Montero (38,2 %) y Santo Domingo (33,7 %).

No obstante, se observan zonas donde persiste un rezago formativo importante. Distritos como Cura Mori, Sondorillo y Frías concentran proporciones destacables de personas sin nivel educativo o con solo primaria, lo cual evidencia desigualdades internas

que podrían estar vinculadas a limitaciones económicas, distancia a los centros educativos o disponibilidad reducida de servicios públicos. En contraste, algunas zonas muestran mayores niveles de escolaridad superior. El Faique presenta el porcentaje más alto de educación universitaria completa (21,6 %), seguido por Canchaque y Frías. Aun así, la educación de posgrado es prácticamente inexistente en la muestra, reflejando la dificultad de acceder a niveles formativos avanzados en ámbitos rurales. (Tabla 5)

Tabla 6: Distribución del estado civil según zona rural

		Buenos Aires	Canchaque	Cura Mori	El Faique	El Tallán	Frías	Montero	Pacaipampa	Santo Domingo	Sondorillo
Conviviente	n	71	13	45	8	22	25	2	9	24	8
	%	27,8%	26,5%	22,6%	15,7%	36,7%	28,4%	5,9%	16,7%	24,5%	17,0%
Separado(a)	n	10	3	10	2	3	5	5	2	0	1
	%	3,9%	6,1%	5,0%	3,9%	5,0%	5,7%	14,7%	3,7%	0,0%	2,1%
Casado(a)	n	45	8	64	13	14	11	9	15	21	6
	%	17,6%	16,3%	32,2%	25,5%	23,3%	12,5%	26,5%	27,8%	21,4%	12,8%
Viudo(a)	n	20	3	12	3	5	8	5	2	6	6
	%	7,8%	6,1%	6,0%	5,9%	8,3%	9,1%	14,7%	3,7%	6,1%	12,8%
Divorciado(a)	n	0	0	0	0	0	0	0	0	2	0
	%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	2,0%	0,0%
Soltero(a)	n	109	22	68	25	16	39	13	26	45	26
	%	42,7%	44,9%	34,2%	49,0%	26,7%	44,3%	38,2%	48,1%	45,9%	55,3%
Total	n	255	49	199	51	60	88	34	54	98	47
	%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fuente: Elaboración propia

Respecto al estado civil, la tabla 6 nos muestra que el estado soltero(a) es el más frecuente en casi todas las zonas, con porcentajes que superan el 40 % en Buenos Aires (42,7 %), Canchaque (44,9 %), El Faique (49,0 %), Pacaipampa (48,1 %) y Sondorillo (55,3 %). Este predominio sugiere que una parte importante de la población adulta mantiene estructuras familiares no formales o no ha consolidado una unión estable. El estado conviviente ocupa el segundo lugar en varias zonas, destacando El Tallán (36,7

%), Frías (28,4 %) y Buenos Aires (27,8 %). En cuanto al estado casado(a), se observa una mayor concentración en Cura Mori (32,2 %), El Faique (25,5 %) y Pacaipampa (27,8 %). Estas cifras indican que, si bien el matrimonio formal mantiene presencia en algunas zonas, su proporción es menor comparada con la convivencia y la soltería. Los estados separados(a) y viudo(a) muestran porcentajes bajos en la mayoría de distritos. El estado divorciado es prácticamente inexistente, registrándose únicamente un 2 % en Santo Domingo.

Tabla 7: Distribución de la identidad étnica en las zonas rurales.

	Buenos Aires	Canchaque	Cura Mori	El Faique	El Tallán	Frías	Montero	Pacaipampa	Santo Domingo	Sondorillo
Blanco	n	1	0	0	0	0	3	1	7	0
	%	0,4%	0,0%	0,0%	0,0%	0,0%	8,8%	1,9%	7,1%	0,0%
Mestizo	n	254	49	199	51	60	88	53	91	47
	%	99,6%	100,0%	100,0%	100,0%	100,0%	91,2%	98,1%	92,9%	100,0%
Total	n	255	49	199	51	60	88	54	98	47
	%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%	100,0%

Fuente: Elaboración propia

La distribución de la autoidentificación étnica muestra una marcada homogeneidad en todas las zonas analizadas, como se observa en la tabla 7, en prácticamente todos los distritos, la categoría mestiza es predominante, alcanzando valores cercanos o iguales al 100 %. Zonas como Canchaque, Cura Mori, El Faique, El Tallán, Frías y Sondorillo registran un 100 % de población que se identifica como mestiza, mientras que en otras zonas, como Buenos Aires, Pacaipampa y Santo Domingo, esta categoría también supera el 92 %.

La categoría blanca presenta una presencia mínima, con porcentajes muy bajos en Montero (8,8 %), Santo Domingo (7,1 %), Pacaipampa (1,9 %) y Buenos Aires (0,4 %). En los demás distritos, no se reporta población identificada como blanca. Esta baja frecuencia evidencia que la autoidentificación blanca no constituye un grupo representativo dentro las zonas rurales estudiadas (Tabla 7)

Tabla 8: Autoidentificación religiosa de la población en zonas rurales.

	Buenos Aires	Canchaque	Cura Mori	El Faique	El Tallán	Frías	Montero	Pacaipampa	Santo Domingo	Sondorillo	
Católica	n	218	37	179	39	60	85	34	26	44	
	%	85,5%	75,5%	89,9%	76,5%	100,0	96,6%	100,0	48,1%	90,8%	93,6%
Evangélica	n	20	12	16	11	0	2	0	25	7	2
	%	7,8%	24,5%	8,0%	21,6%	0,0%	2,3%	0,0%	46,3%	7,1%	4,3%
Otra	n	6	0	1	0	0	0	0	0	0	0
	%	2,4%	0,0%	0,5%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%
Ninguna	n	11	0	3	1	0	1	0	3	2	1
	%	4,3%	0,0%	1,5%	2,0%	0,0%	1,1%	0,0%	5,6%	2,0%	2,1%
Total	n	255	49	199	51	60	88	34	54	98	47
	%	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0	100,0

Fuente: Elaboración propia

La afiliación religiosa en las zonas rurales se caracteriza por un marcado predominio de la religión católica, que supera el 80 % en la mayoría de distritos y alcanza el 100 % en El Tallán y Montero. La religión evangélica aparece como la segunda más frecuente, especialmente en Canchaque (24,5 %), El Faique (21,6 %) y Pacaipampa, donde llega al 46,3 %. Las categorías "otra religión" y "ninguna" presentan proporciones muy bajas y no representan grupos relevantes dentro de la población rural. (Tabla 8)

#### Diferencias Estructurales en la Población Rural

La tabla 9 presenta los indicadores de estructura de población más importantes por cada una de las zonas rurales estudiadas. En primer lugar, el Índice de Masculinidad (IM) muestra un rango considerable entre zonas, desde valores extremadamente bajos como en Canchaque (40) hasta valores muy elevados como en Frías (125,64). Un Índice de masculinidad por debajo de 100 indica predominio femenino, mientras que valores superiores implican sobre representación masculina. Zonas como Canchaque, Montero y El Tallán presentan déficit de población masculina, lo que podría estar asociado a migración laboral de varones hacia centros urbanos o a mayores índices de mortalidad masculina. En contraste, Frías y El Faique exhiben superávit masculino, lo cual puede

corresponder a economías locales basadas en actividades agrícolas o extractivas donde predomina la participación de hombres.

Tabla 9 Indicadores de población por zona rural

Zonas rurales	Índice de	Índice de	Índice de	Edad Media
	Masculinidad	Envejecimiento	dependencia	
	IM	IV	TDI	EM
Buenos Aires	91.73	111.11	59.38	41.6
Canchaque	40	275	44.12	43.92
Cura Mori	85.92	89.47	56.69	38.31
El Faique	112.5	142.86	50	42.59
El tallan	76.47	116.67	27.66	40.08
Frías	125.64	123.08	49.15	42.27
Montero	78.95	300	54.55	48.38
Pacaiyampa	100	26.67	54.29	34.81
Santo Domingo	108.51	371.43	50.77	46.47
Sondorillo	80.77	26.67	67.86	29.57

Fuente: Elaboración propia

Respecto al índice de Envejecimiento (IV), los contrastes son aún más notables. Mientras Santo Domingo (371.43), Montero (300) y Canchaque (275) muestran estructuras poblacionales marcadamente envejecidas, donde la población de 60 años o más supera ampliamente a la infantil, otras zonas como Pacaiyampa (26.67) y Sondorillo (26.67) mantienen una estructura joven. La elevada proporción de adultos mayores en localidades altoandinas como Santo Domingo, Montero y Canchaque sugiere salidas constantes de jóvenes a otras zonas, limitando la base productiva y aumentando la carga social sobre las familias. Por el contrario, zonas jóvenes como Sondorillo y Pacaiyampa podrían encontrarse en una etapa previa de cambios demográficos, con alta fecundidad o menor migración de jóvenes.

En cuanto al índice de Dependencia Total (TDI), zonas como Sondorillo (67.86) y Buenos Aires (59.38) presentan una carga demográfica elevada, lo que implica que la población económicamente activa debe sostener un número considerable de dependiente es decir a jóvenes menores de 14 años y personas mayores de 50. En cambio, El Tallan registra el valor más bajo (27.66), lo que sugiere una estructura más equilibrada. Es importante destacar que, en zonas como Santo Domingo, la relación entre envejecimiento

y dependencia no sigue un patrón lineal. A pesar de presentar un nivel muy alto de envejecimiento (371.43), la dependencia general es moderada (50.77), esto sugiere que la carga demográfica recae principalmente en la población adulta mayor y no en la población infantil.

Finalmente, la Edad Media (EM) refuerza los patrones previamente descritos. Las zonas con mayor envejecimiento muestran edades medias elevadas, como Montero (48.38) y Santo Domingo (46.47). Por el contrario, las zonas jóvenes como Sondorillo (29.57) y Pacaiyampa (34.81) exhiben edades medias considerablemente menores.

### Brechas en el Comportamiento Reproductivo de la Población Rural

Tabla 10: Indicadores de fecundidad por zona rural

Zonas rurales	Tasa bruta de natalidad	Tasa de fecundidad general	Edad Media de la fecundidad	Tasa global de fecundidad	Tasa bruta de reproducción
	TBN	TFG	EMF	TGF	TBR
Buenos Aires	1.96	8.47	23.16	2.40	1.17
Canchaque	4.08	11.11	31.07	2.91	1.42
Cura Mori	2.51	9.62	24.82	2.92	1.42
El Faique	3.92	22.22	30.00	5.00	2.43
El tallan	1.67	5.56	22.50	2.50	1.21
Frías	3.41	16.67	31.05	5.16	2.50
Montero	2.34	25.00	47.50	1.66	0.81
Pacaiyampa	1.85	6.67	22.50	2.5	1.21
Santo Domingo	2.04	9.52	26.00	2.38	1.16
Sondorillo	6.38	27.27	28.21	5.83	2.84

Fuente: Elaboración propia

Los indicadores de fecundidad permiten describir el comportamiento reproductivo de las zonas rurales analizadas. En la Tabla 10 se presentan las principales medidas asociadas a este fenómeno. La Tasa Bruta de Natalidad (TBN), por ejemplo, refleja el número de nacidos vivos registrados por cada 100 habitantes, constituyéndose en un indicador básico para aproximar la dinámica reproductiva de la población; por lo tanto, valores elevados reflejan mayor intensidad reproductiva general. En este indicador, Sondorillo (6.38),

Canchaque (4.08) y El Faique (3.92) presentan las mayores tasas, evidenciando una dinámica demográfica expansiva. En contraste, El Tallán (1.67) y Pacaipampa (1.85) registran niveles bajos, asociados a menor ocurrencia de nacimientos en relación con su tamaño poblacional.

La Tasa de Fecundidad General (TFG), que mide los nacimientos por cada 100 mujeres en edad fértil, muestra una marcada variabilidad territorial. Las tasas más altas se observan en Sondorillo (27.27), Montero (25.00) y El Faique (22.22), lo que indica un mayor ritmo reproductivo femenino. Por el contrario, El Tallán (5.56), Pacaipampa (6.67) y Buenos Aires (8.47) presentan niveles considerablemente inferiores.

Respecto a la Edad Media de la Fecundidad (EMF), esta permite identificar la estructura del patrón reproductivo según tres formas clásicas: cúspide temprana (mayor aporte entre 20–24 años), cúspide dilatada (aportación distribuida entre 20–24 y 25–29 años) y cúspide tardía (máximo entre 25–29 años) (Instituto Nacional de Estadística e Informática, 2017). Bajo esta clasificación, Buenos Aires, El Tallán y Pacaipampa, con EMF entre 22 y 24 años, se ajustan a un patrón de cúspide temprana, asociado a maternidad más precoz. Por su parte, Cura Mori, Santo Domingo y Sondorillo, con EMF alrededor de 25–26 años, muestran un patrón de cúspide dilatada, donde la fecundidad se concentra entre los dos grupos centrales de edad. Finalmente, Canchaque, El Faique y Frías, con EMF cercanas a 30 años, presentan un patrón de cúspide tardía, indicando que los nacimientos se desplazan hacia edades reproductivas más avanzadas.

Por otro lado la Tasa Global de Fecundidad (TGF) estima el número promedio de hijos que tendría una mujer a lo largo de su vida. Los valores más altos se registran en Sondorillo (5.83), Frías (5.16) y El Faique (5.00), consistentes con patrones reproductivos propios de contextos rurales con familias numerosas. En contraste, Montero (1.66) y Buenos Aires (2.40) muestran niveles bajos de 1 a 3 hijos durante la vida reproductiva de la mujer.

Por último, la Tasa Bruta de Reproducción (TB), que mide el número promedio de hijas por mujer, confirma estas diferencias territoriales. Zonas como Sondorillo, Frías y El Faique presentan tasas superiores a 2, lo que implica capacidad de crecimiento poblacional sostenido. Por el contrario, Montero (0.81) muestra un nivel insuficiente para asegurar el reemplazo, y áreas como Buenos Aires, Santo Domingo y El Tallán presentan valores moderados.

## DISCUSIÓN Y CONCLUSIONES

El presente estudio tuvo como objetivo comparar los indicadores demográficos en cada una de las zonas rurales analizadas y contrastar los resultados obtenidos con la información del Censo Nacional de Población y Vivienda 2017.

En primer lugar, se evaluaron las características relacionadas con las condiciones de la vivienda y el acceso a los servicios básicos. En este aspecto, el material predominante en las paredes presenta una distribución heterogeneidad territorial importante, mientras que localidades como El Tallán (96,15%), Cura Mori (87,8%) y Canchaque (80,95%) registran un claro predominio de ladrillo o bloque de cemento, en lugares como Frías (83,87%), Montero (47,06%) y Pacaipampa (55,56%), el uso de adobe y quincha continúa siendo predominante, lo que refleja condiciones económicas más limitadas, dicho patrón coincide con lo descrito por el Censo 2017 cuyos registros evidencian un uso mayoritario de materiales como el adobe o la tapia en un 69,9 % de las zonas rurales, tendencia que se ve reforzada por lo señalado por Tito (2018), quien explica que en diversas localidades el adobe sigue siendo utilizado debido a que la tierra necesaria para su elaboración se obtiene directamente de las parcelas agrícolas, lo que simplifica su preparación y reduce los costos de construcción.

En cuanto a los materiales predominantes en los techos, se observa que las planchas de calamina o fibra de cemento constituyen el material claramente dominante, con porcentajes elevados en Buenos Aires (71,3%), Canchaque (61,9%), Cura Mori (70,7%), El Faique (64,7%), El Tallán (80,8%), Frías (90,3%), Montero (90,3%), Santo Domingo (68,4%) y alcanzando el 100% en Sondorillo. En contraste, el concreto armado presenta una participación mucho menor y solo destaca en Canchaque, El Faique, Montero y Pacaipampa, con porcentajes que fluctúan entre el 33% y el 35%. Estos resultados coinciden con el Censo 2017, que señala que en las zonas rurales el 61,3% de las viviendas utiliza planchas de calamina o fibras de cemento como principal cobertura.

En relación con los pisos de las viviendas, los resultados muestran una tendencia marcada hacia el uso del cemento, especialmente en Canchaque (76,19%), Pacaipampa (72,22%), El Tallán (73,08%) y El Faique (70,59%), lo que refleja un nivel básico, aunque aceptable, de consolidación estructural. No obstante, el uso de pisos de tierra continúa siendo considerable en varias zonas, destacando Sondorillo (46,67%), Santo Domingo (34,21%), Cura Mori (30,49%), Buenos Aires (26,73%) y El Faique (23,53%), lo que revela condiciones de mayor vulnerabilidad. Estos hallazgos defieren parcialmente de lo

reportado por el Censo 2017 donde en las áreas rurales predomina ampliamente la tierra en los pisos de las viviendas (78,8%) y el cemento solo alcanza un 15,6%. En este sentido, aunque persisten niveles importantes de precariedad, especialmente en las zonas donde la tierra aún es frecuente, los resultados sugieren ciertos avances en la mejora de los pisos en comparación con el año 2017, evidenciando un proceso gradual de mejoramiento de la calidad de las viviendas.

Respecto al acceso al agua, se observa que la red pública es claramente predominante en distritos como El Tallán, Cura Mori, El Faique, Canchaque y Montero, mientras que en Sondorillo y Buenos Aires aún destaca el uso de pozo o agua subterránea, y en zonas como Santo Domingo y Frías prevalece el abastecimiento mediante manantiales o puquiales. Estos resultados muestran una similitud con Censo 2017, donde en el área rural nacional el 52,2% de los hogares accede al agua por red pública, el 29,7% utiliza pilón, camión o pozo y el 17,1% depende de fuentes naturales. Esto indica que los datos siguen la misma tendencia general, aunque con variaciones propias del territorio.

En cuanto a la conexión de los servicios higiénicos, se observa que predomina ampliamente la conexión a la red pública de desagüe, incluso con coberturas superiores al 95% en varios distritos, aunque aún persisten casos aislados de uso de ríos, acequias o pozos ciegos. Dichos resultados defieren con el Censo 2017, donde solo el 17,3% de los hogares rurales contaban con alcantarillado y el 63,2% usaban otras formas de disposición.

En cuanto al alumbrado público, las zonas estudiadas presentan una cobertura casi universal, con valores que oscilan entre el 94% y el 100%. Al contrastar estos resultados con el Censo 2017, donde se indica que en el área rural el 74,1% de los hogares dispone de electricidad, cifra muy superior al 27,5% registrado en el Censo 2007, se observa que la situación en nuestros distritos es aún más favorable, esto permite concluir que el acceso al alumbrado público continúa en expansión y que las viviendas con servicio eléctrico siguen aumentando de manera sostenida.

Otras de las características analizadas en el estudio son las relacionadas con la población, donde se evidencia una clara coincidencia en el orden de predominio de varias variables al compararlas con lo reportado en el Censo 2017. En primer lugar, en ambas fuentes el alfabetismo constituye la condición mayoritaria, ya que en los distritos analizados la mayoría de personas sabe leer y escribir. En cuanto al nivel educativo, los resultados muestran que la educación secundaria es el nivel más frecuente, seguida por primaria. Este patrón difiere del orden reportado por el Censo 2017, donde en las zonas

rurales la educación primaria ocupa el primer lugar y la secundaria el segundo. Esta diferencia sugiere que, en los últimos ocho años, ha ocurrido un avance educativo en las zonas rurales. Respecto a la identidad étnica, en ambas fuentes la categoría mestiza es claramente la predominante, muy por encima de otras identidades, lo que indica coherencia entre los datos del estudio y la caracterización nacional. Una tendencia comparable también se aprecia en el estado civil, donde el grupo soltero(a) aparece como la primera categoría, seguido por casados y convivientes.

Respecto a los indicadores de estructura poblacional, los distritos rurales muestran estructuras poblacionales muy heterogéneas, lo que evidencia dinámicas locales distintas. Se observan distritos con alto envejecimiento y fuerte dependencia, como Canchaque, Montero y Santo Domingo, donde la elevada proporción de adultos mayores podría estar asociada a factores como la urbanización, la expansión de la educación superior y el acceso a la planificación familiar, los cuales han propiciado el descenso de la natalidad y, en consecuencia, el aumento de la población de 60 años y más. (Organización Mundial de la Salud [OMS], 2025). En contraste, Pacaipampa y Sondorillo muestran un envejecimiento muy bajo y una estructura poblacional más joven, lo cual coincide con lo descrito por el Censo 2017, que señalan que en el área rural del país todavía predomina una población relativamente joven, pese a ciertos cambios en su composición en los últimos años.

Asimismo, el índice de masculinidad presenta comportamientos mixtos entre los distritos evaluados. Algunos muestran un ligero predominio masculino, como El Faique, Frías, Pacaipampa y Santo Domingo, lo que coincide con el Censo 2017, donde en las zonas rurales, la mayoría de grupos de edad predominan los hombres. En contraste, otros distritos evidencian una mayor presencia femenina, como Canchaque y Sondorillo, comportamiento que se asemeja más al de las zonas urbanas del país, donde este índice alcanza 96,8, reflejando que el número de hombres es ligeramente menor al de mujeres. Esta combinación de escenarios muestra que no existe un patrón demográfico único en la zona estudiada, sino dinámicas diferenciadas posiblemente asociadas a procesos migratorios y a las oportunidades económicas locales.

Los indicadores de fecundidad muestran también patrones contrastantes entre los distritos rurales. A nivel nacional, según los datos del Censo, la Tasa Global de Fecundidad (TGF) mostró una disminución de 2,97 hijos por mujer en 2007 a 2,46 en 2017, reflejando un patrón general de reducción de la fecundidad en el país. (INEI, 2017).

Al comparar esta tendencia con los datos de las zonas rurales, se observa que algunos distritos mantienen niveles de fecundidad superiores a la media nacional. Por ejemplo, Sondorillo presenta una TGF de 5,83 hijos por mujer y una tasa bruta de reproducción (TBR) de 2,84, mientras que El Faique y Frías también registran valores elevados de TGF (5,00 y 5,16 respectivamente) y TBR (2,43 y 2,50). En contraste, distritos como Buenos Aires, El Tallan, Pacaipampa y Santo Domingo muestran TGF más cercanas al promedio nacional de 2,46, con valores entre 2,38 y 2,50, reflejando una fecundidad moderada. Canchaque y Cura Mori se sitúan en un nivel intermedio, con TGF de 2,91 y 2,92, y TBR de 1,42. Montero presenta un patrón atípico, con una elevada TFG de 25,00 y una edad media de fecundidad de 47,5 años.

#### Agradecimientos

Queremos agradecer a todas las personas que aceptaron participar en el estudio, proporcionando la información necesaria para el desarrollo de la investigación. Asimismo, expresamos nuestro agradecimiento a los estudiantes de la Escuela Profesional de Estadística por su apoyo en el proceso de recolección de la información y en las actividades realizadas durante el trabajo de campo. Financiación: El presente estudio fue financiado por la Universidad Nacional de Piura.

#### REFERENCIAS BIBLIOGRAFICAS

- Arquínigo, P. (2022). *Demografía, principales indicadores demográficos de las condiciones de vida*. (Tesis de Licenciatura). Universidad nacional de educación Enrique Guzmán y Valle, Lima. <https://repositorio.une.edu.pe/server/api/core/bitstreams/61111414-ceed-4865-bf29-a088e7625377/content>
- Bel, C., & Gomez, J. (1998). "Indicadores demograficos y su distribucion territorial en el municipio de Murcia". En *Papeles de geografia*, (28): 5-25. <https://revistas.um.es/geografia/articulo/view/45381>
- Cedeño, N. V., Loor, L. E., & Romero, S. A. (2019). "Contexto sociodemografico y situacion organizacional de la poblacion Maconta". En *Revista Caribeña de Ciencias Sociales*,(9): 21. <https://www.eumed.net/rev/caribe/2019/09/situacion-organizacional-maconta.html>
- Comisión Económica para América Latina y el Caribe [CEPAL]. (2014). *Los datos demográficos. Alcances, limitaciones y metodos de evaluacion*. Santiago de Chile: Comisión Económica para América Latina y el Caribe (CEPAL).
- Del Barrio, E., & Abellan, A. (2006). "Indicadores demograficos". Inverso, España. <http://hdl.handle.net/11181/3253>
- DANE (s.d). *Indicadores de estructura*. Tomo Estadístico. Colombia. Departamento Administrativo Nacional de Estadística (DANE).
- Dolores, J. V. (2017). *Técnicas y ejercicios de Demografía*. Madrid, INE. [https://www.ine.es/ine/opusosiciones/temario\\_2021/Tecnicas\\_ejercicios\\_demografia.pdf](https://www.ine.es/ine/opusosiciones/temario_2021/Tecnicas_ejercicios_demografia.pdf)
- Hernandez, R., Fernandez, C., & Baptista, P. (2014). *Metodología de la investigación*. (6ª. Edc). Mexico, McGraw-Hill / Interamericana Editores, S.A.
- INEI (2017). *Características de la población*. Republica del Perú: Instituto Nacional de Estadística e Informática (INEI). [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib1539/cap01.pdf](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1539/cap01.pdf)
- INEI (2018). *Resultados definitivos Piura*, Republica del Perú: Instituto Nacional de Estadística e Informática (INEI). [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib1553/](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1553/)
- INEI (2018). *Características de la vivienda*. Piura, Republica del Perú: Instituto Nacional de Estadística e Informática (INEI). [https://www.inei.gob.pe/media/MenuRecursivo/publicaciones\\_digitales/Est/Lib1711/cap04.pdf](https://www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1711/cap04.pdf)
- Instituto Nacional de Estadística e informática. (2017). *Estimación y análisis de la Fecundidad segun diversas fuentes*, Republica del Perú: Instituto Nacional de Estadística e Informática (INEI). <https://www.inei.gob.pe/media/MenuRecursivo/boletines/fecundidad.pdf>
- INE. (2024). *Indicadores Demograficos Basicos*, Madrid: instituto Nacional de Estadística (INE). [https://www.ine.es/metodologia/t20/metodologia\\_idb.pdf](https://www.ine.es/metodologia/t20/metodologia_idb.pdf)
- INEC. (2000). *El sistema de registro civil y estadísticas vitales de Costa Rica*. Costa Rica: Instituto nacional de estadística y censo (INEC).
- Lobe, B., Morgan, D., & Hoffman, K. (2020). "Qualitative Data Collection in an Era of Social Distancing". En *International Journal of Qualitative Methods*, 19. <https://doi.org/10.1177/1609406920937875>
- Lugo, I. V., & Perez, Y. Y. (2025). Herramientas digitales para encuestas y evaluaciones. En *Psicología e investigación digital: Aplicaciones, retos y perspectivas*. 76-99. México. UNAM. [https://www.researchgate.net/publication/395695157\\_Herramientas\\_digitales\\_para\\_encuestas\\_y\\_evaluaciones](https://www.researchgate.net/publication/395695157_Herramientas_digitales_para_encuestas_y_evaluaciones)
- Miró, c. (2007). "La demografía en el siglo XXI en América Latina". En *Revista Latinoamericana de Población*, 1(1): 1-9. <http://www.redalyc.org/articulo.oa?id=323827539002>
- Moran, G. H. (23 de mayo del 2025). Censos Nacionales 2025: incorporarán herramientas digitales para modernizar la recolección de datos. *Instituto Nacional de Estadística e informática*. <https://www.gob.pe/institucion/inei/noticias/1173864-censos-nacionales->

2025-incorporaran-herramientas-digitales-para-modernizar-la-recoleccion-de-datos?utm\_source=chatgpt.com

Organización Mundial de la Salud [OMS]. (21 de febrero de 2025). Envejecimiento: Población mundial. *Organización mundial de la salud*. <https://www.who.int/es/news-room/questions-and-answers/item/population-ageing>

Perez Lopez, C. (2005). *Muestreo Estadístico, concepto y problemas resueltos*. Madrid, Pearson Educación, S.A

Portela Garcia, J., & Villeta Lopez, M. (2007). *Técnicas básicas de muestreo con SAS*.

Saavedra, G. N. (2018). *Propuesta técnica para el mejoramiento y ampliación del servicio de agua potable en los centros poblados rurales de Culqui y Culqui Alto en el distrito de Palmás, provincia de Ayabaca – Piura* (Tesis de Licenciatura). Universidad Nacional de Piura, Piura. <https://repositorio.unp.edu.pe/handle/UNP/1249>

Sposob, G. (17 de julio del 2025). Poblacion. *Enciclopedia Concepto*. Recuperado el 8 de noviembre de 2025 de <https://concepto.de/poblacion/>

Tito, E. (2018). *Viviendas rurales y su impacto social en Marccaracay - distrito de Santillana - Huanta - Ayacucho - 2016* (Tesis de Licenciatura). Universidad Nacional San Cristobal de Huamanga, Ayacucho. <http://repositorio.unsch.edu.pe/handle/UNSCH/3536>

Vieira, A. (1973). *El pensamiento crítico en Demografía*. Santiago de Chile: Centro Latinoamericano de Demografía. [https://cep.uer.ac.cr/cursos/icamacho/public\\_html/sociopob\\_II\\_03/contenido/demografi\\_a.pdf](https://cep.uer.ac.cr/cursos/icamacho/public_html/sociopob_II_03/contenido/demografi_a.pdf)

#### RESUMEN CURRICULAR DE LOS AUTORES

*Ricardo Antonio Armas Juárez*

Es Doctor en Ciencias de la Educación por la Universidad Nacional de Piura (UNP), Maestro en Ciencias con mención en Estadística Aplicada e Ingeniero Estadístico por la Universidad Nacional de Trujillo (UNT), Perú. Es Investigador RENACYT - CONCYTEC y Docente Ordinario Auxiliar adjunta a la Escuela Profesional de Estadística de la Facultad de Ciencias de la UNP. Con 17 años de experiencia en la docencia universitaria, participado en diversos eventos nacionales e internacionales, fortaleciendo mi actualización y desempeño profesional. Sus líneas de investigación están relacionadas con la Educación, la Estadística Aplicada y la Demografía en las que desarrolla sus actividades científicas.

Dirección electrónica: [rarnasj@gmail.com](mailto:rarnasj@gmail.com); [rarnasj@unp.edu.pe](mailto:rarnasj@unp.edu.pe)

Registro ORCID: <https://orcid.org/0000-0002-0048-2711>

*Gabriela Ipanaqué Cárdenas*

Es egresada de Estadística por la Universidad Nacional de Piura, ganadora de un concurso universitario orientado a la implementación de un sistema digital automatizado para la medición de parámetros sociodemográficos en la región Piura. Participó en un intercambio académico en la Pontificia Universidad Católica del Perú (PUCP), fortaleciendo sus competencias analíticas. Actualmente se desempeña en el ámbito de la salud, donde se encarga de la elaboración y análisis de indicadores estadísticos orientados al apoyo de la gestión, la planificación y la toma de decisiones. Sus intereses académicos se centran en la estadística aplicada, el análisis de datos, los estudios poblacionales y la producción de información estadística para la investigación.

Dirección electrónica: [gabrielaipanaquecardenas@gmail.com](mailto:gabrielaipanaquecardenas@gmail.com), [gabyipan5@gmail.com](mailto:gabyipan5@gmail.com)

Registro ORCID: <https://orcid.org/0009-0002-3750-4707>

*Diana Evelyn Hidalgo Eche*

Es Bachiller en Estadística por la Universidad Nacional de Piura, con formación sólida en análisis de datos, estadística aplicada y toma de decisiones. Ingresante al Programa de Extensión Universitaria (PEU) del INEI, con enfoque en la producción, análisis y uso de estadísticas oficiales. Experiencia en prácticas preprofesionales en el sector público, realizando análisis y procesamiento de información para la gestión institucional. Ganadora de un proyecto universitario orientado a la implementación de un sistema digital automatizado para la medición de parámetros sociodemográficos en tiempo real en la región Piura. Participación en intercambio académico en la Pontificia Universidad Católica del Perú (PUCP). Manejo de Python, Excel avanzado y técnicas de análisis estadístico. Experiencia en web scraping y análisis de precios.

Dirección electrónica: [dianahe2003@gmail.com](mailto:dianahe2003@gmail.com); [a20233251@pucp.edu.pe](mailto:a20233251@pucp.edu.pe)

Registro ORCID: <https://orcid.org/0009-0009-0480-0957>

Survey Research Methods (2024)  
Vol. 17, No. 1, pp. 10000-10000  
doi:10.1515/sur-2024-0001  
© 2024 Survey Research Methods  
www.surveys.researchmethods.org

Open Access  
This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Digital versus paper data collection in large-scale household surveys: time efficiency, heterogeneity by complexity, and modelling of the within-pair difference

Ricardo Armas Juárez<sup>1</sup>, Diana Hidalgo Eche<sup>1</sup>, and Gabriela Ipanaque Cardenas<sup>1</sup>  
<sup>1</sup>Faculty of Sciences, Universidad Nacional de Piura, Peru

This study compares pencil-and-paper data collection (PAPI) with tablet-assisted data collection (CAPI) in a large-scale household survey, based on 363 matched observations collected in the Piura Region, Peru.

The digital method reduces mean administration time by 43.41% (from 1:39:67 to 79:47 minutes).

A paired *t*-test ( $t = 47.15$ ,  $p < .001$ ) and the Wilcoxon signed-rank test ( $V = 66,966$ ,  $p < .001$ ) both confirm the difference.

Cohen's *d* for paired samples is  $d = 2.475$  (95% CI: [2.265, 2.680]), indicating a large effect size; the 95% confidence interval for the mean difference ( $\mu_{diff}$ ) is [57.70, 62.72] minutes.

Additionally, a simple linear regression model shows that the effective number of questions in the questionnaire explains 85.48% of the variance of the within-pair difference ( $\beta = 0.824$  min/question,  $t = 46.10$ ,  $p < .001$ ), while the percentage reduction remains statistically constant across the three complexity levels ( $F = 401.20$ ,  $p < .001$  for absolute difference;  $t = 197.7$ ,  $p < .001$  for percentage reduction), demonstrating that the efficiency gain of the digital method is proportional and robust to household complexity.

These findings provide solid empirical evidence in favour of the digitalisation of data collection in official statistics.

**Keywords:** digital survey, CAPI, PAPI, household survey, paired *t*-test, linear regression, questionnaire complexity, data collection efficiency, official statistics.

### 1 Introduction

The transformation of data collection processes in official statistics has placed the choice of survey method at the centre of methodological and operational debate. The pencil-and-paper interview (PAPI) has been the standard in household survey operations for decades; however, the expansion of portable digital devices has enabled the widespread adoption of computer-assisted personal interviewing (CAPI) and tablet-based data capture systems (Couper, 2017; Groves et al., 2009).

The decision to migrate from manual to digital methods is not purely technological. It involves trade-offs affecting data quality, field costs, interviewer training, and the statis-

tical properties of the collected information (R. Baker et al., 2010). For national statistical institutes and survey organisations in emerging economies, these trade-offs are particularly relevant given budget constraints and varying levels of digital literacy among interviewees (United Nations Statistics Division, 2014).

Comparisons between CAPI and PAPI have been carried out in diverse geographical and demographic settings. European studies (Heerwegh & Loosveldt, 2008; Lynn et al., 2014) have documented time savings and reductions in item nonresponse associated with digital modes; research in sub-Saharan Africa (Caeyers et al., 2012) and Latin America (Ballester et al., 2015) has highlighted challenges such as device failures, connectivity limitations, and interviewer resistance. In particular, Caeyers et al. (2012) demonstrate through a randomised experiment that CAPI reduces routing errors by 94% and interview times by 10–14% compared to PAPI, and that errors in paper data are not randomly distributed across the sample, which may introduce bias in estimates if problematic observations are dropped.

In the Peruvian context, this modernisation process has acquired first-order institutional significance. The National Institute of Statistics and Informatics (INEI) incorporated digi-

#### Contact information:

Ricardo Armas Juárez, Faculty of Sciences, Universidad Nacional de Piura, Peru. (E-mail: [ramas@unp.edu.pe](mailto:ramas@unp.edu.pe))  
Diana Hidalgo Eche (E-mail: [dhidalgo@unp.edu.pe](mailto:dhidalgo@unp.edu.pe))  
Gabriela Ipanaque Cardenas (E-mail: [gipanaque@unp.edu.pe](mailto:gipanaque@unp.edu.pe))  
This research was funded by Universidad Nacional de Piura under Research Project Resolution No. 385-CU-2024.

10000

ARMAS JUÁREZ, HIDALGO ECHE, AND IPANAQUE CARDENAS

tal tablets as the sole data collection tool in the 2025 National Censuses—XIII Population, VIII Housing and IV Indigenous Communities—marking the first time in the country's history that all census information was collected without paper forms (Instituto Nacional de Estadística e Informática—INEI, 2025). The operation, carried out between August and October 2025, deployed more than 30,000 enumerators equipped with tablets connected to institutional servers and cloud storage, with the projected reduction in the time to publication of first results from two years to five months.

Despite the growing body of evidence, rigorous matched comparisons—which hold the respondent population constant by applying both methods to the same household during the same visit—remain relatively scarce (Tourangeau et al., 2013). Additionally, none of the previous studies has explicitly modelled the relationship between questionnaire complexity, operationalised as the number of questions effectively administered to each household, and the magnitude of the temporal difference between modes, which limits the ability to extrapolate efficiency gains to different types of statistical operations.

The present study contributes to this literature in three dimensions. First, through a paired and controlled comparison of administration time under PAPI and CAPI with 363 pairs, confirming the temporal reduction with convergent parametric and non-parametric tests. Second, through a simple linear regression model that quantifies the effect of the number of effective questions on the within-pair difference, explaining 85.48% of its variance. Third, through a heterogeneity analysis by complexity tertiles that demonstrates that the percentage reduction is statistically constant across all strata, reinforcing the operational generalisability of the results.

The article is structured as follows: Section 2 reviews the relevant methodological literature; Section 3 describes the data and methods; Section 4 presents the results; Section 5 discusses the implications; and Section 6 concludes.

### 2 Literature Review

#### 2.1 PAPI versus CAPI: General Overview

The systematic comparison between paper-based and computer-assisted interview modes has a history of more than three decades in survey methodology. Early evaluations of CAPI in government surveys (R. P. Baker et al., 1995; Catlin & Ingram, 1988) produced mixed evidence on data quality; however, the majority of subsequent studies reported advantages of digital modes in following skip patterns and reducing out-of-range responses (de Leeuw et al., 2008).

A meta-analysis by Heerwegh (2009) found that electronic modes were associated with significantly lower item nonresponse rates compared to self-administered paper questionnaires, a result consistent with the routing logic and automatic validation available in CAPI systems. Studies

with random assignment of interviewers to mode (Jäckle et al., 2010) additionally demonstrated that differences in data quality between modes are partly attributable to the differential burden imposed on interviewers when managing complex skip patterns.

The randomised experiment by Caeyers et al. (2012) in Tanzania, the methodologically most rigorous study available in the African context, compared full CAPI, restricted CAPI (without consistency checks), and PAPI in 1,840 households, finding that routing errors are reduced by 94% with CAPI and that the mean interview time decreases by 10–14%. A particularly relevant finding of that study is that errors in paper data are systematically concentrated in households with certain characteristics—large households, female-headed, and non-agricultural—which introduces sample bias if observations with problematic data are dropped. This evidence underlines that comparisons of data quality between modes cannot ignore the structure of errors in relation to household characteristics.

#### 2.2 Time Efficiency and Questionnaire Complexity

Several studies have quantified the time savings associated with digital data collection. Fuchs et al. (2000) examined whether CAPI or PAPI interviews take longer, finding that the digital mode consistently reduced administration time, with gains attributable primarily to automated skip navigation and the elimination of manual data handling steps. Ballester et al. (2015) reported similar reductions in household surveys in Latin America, attributing the gains mainly to the elimination of manual transcription and automated navigation of complex questionnaire structures. More recently, Adali et al. (2021) evaluated the transition from PAPI to CAPI in a national survey context, documenting a substantial reduction in mean administration time in favour of digital collection after controlling for questionnaire length and respondent characteristics.

An aspect that the prior literature has not systematically addressed is the relationship between within-sample questionnaire complexity—measured as the effective number of questions administered to each household, which varies according to its size and composition—and the absolute magnitude of the temporal difference between modes. This relationship is theoretically predictable: if the digital mode automates the navigation of skip patterns and response validation, the time saving per additional question should be positive and approximately constant, generating a linear relationship between number of questions and within-pair difference. The present study tests this hypothesis formally through simple linear regression.

### 2.3 Statistical Considerations in Matched Mode Comparisons

Matched experimental designs for mode comparison have been advocated by Tourangeau et al. (2013) as the gold standard for isolating method effects. When matched data are available, the paired  $t$ -test and its non-parametric counterpart, the Wilcoxon signed-rank test, are the most widely used inferential tools (Conover, 1999; Hollander et al., 2014). Effect size estimation via Cohen's  $d$  is increasingly recommended alongside hypothesis tests to provide interpretable magnitude estimates beyond statistical significance (Cohen, 1988; Ellis, 2010).

The presence of positive skewness and excess kurtosis in interview duration data, documented in studies of mode effects on interview timing (Adali et al., 2021) and confirmed in the present data set, motivates the joint use of parametric and non-parametric tests. When both yield concordant results, confidence in the inferential conclusions is substantially strengthened (Wilcox, 2012).

## 3 Data and Methods

### 3.1 Study Design and Sample

The data were collected within the research project "Implementation of an automated digital system to measure socio-demographic parameters in real time in the Piura Region" (Resolution No. 385-CU-2024), funded by the Universidad Nacional de Piura (Hidalgo et al., 2024). Fieldwork was conducted in urban areas of the Piura Region, Peru. A total of 363 surveys were administered under a matched design, in which each dwelling was visited by two interviewers: one using the paper questionnaire (manual method) and another with a tablet running a CAPI application (digital method). The order of administration was randomised at the dwelling level to control for learning and fatigue effects. All interviewers received standardised training before fieldwork.

The sample comprises 363 matched observations ( $n = 363$  pairs). Administration time was recorded in decimal minutes for both methods. No expansion factors were applied, as the focus of the analysis is on within-pair differences rather than population-level estimates.

The instrument was a demographic household survey composed of six thematic modules: (i) dwelling characteristics and services (10 items); (ii) household characteristics and assets (4 items, including a durable goods inventory of 14 items); (iii) population characteristics administered to each household member (11 items per person, up to 10 members); (iv) employment and marital status for persons aged 12 and over (4 items per person); (v) fertility characteristics for women aged 12 and over (8 items); and (vi) mortality characteristics (5 items). Given the list nature of modules (iii) and (iv), the effective number of questions varied according to household composition, ranging from 32 to 210 questions

per survey (mean: 70.20,  $s = 27.29$ ). This variability serves as the indicator of questionnaire complexity in the regression analysis.

### 3.2 Variables

The main outcome variable is survey duration, measured in decimal minutes from the start of the first item to the recording of the last response. For the manual method, duration was recorded by a field supervisor with a stopwatch, for the tablet method, timestamps were extracted automatically from the CAPI system log. Although these two procedures differ in their recording mechanisms, both instruments capture elapsed clock time and the resulting measurement error is expected to be non-differential with respect to the mode comparison of interest (Groves et al., 2009). The within-pair difference  $D_i = X_{i, \text{manual}} - X_{i, \text{tablet}}$  constitutes the estimator of the method effect for pair  $i = 1, \dots, 363$ .

The main explanatory variable in the regression analysis is  $Q_i$ , the number of questions effectively administered in survey  $i$ , which operates as an indicator of questionnaire complexity. This variable is directly observable in the database and varies exogenously according to household composition, allowing it to be interpreted as a structural moderator of the mode effect.

### 3.3 Statistical Analysis

Descriptive statistics were computed for both methods, including the mean ( $\bar{X}$ ), median, standard deviation ( $s$ ), variance ( $s^2$ ), coefficient of variation (CV), quartiles (P25, P50, P75), skewness ( $g_1$ ), and kurtosis ( $g_2$ ).

#### Main Hypothesis Test

Given the matched design, the null hypothesis

$$H_0: \mu_D = 0 \quad \text{versus} \quad H_1: \mu_D \neq 0 \quad (1)$$

was tested with a two-tailed paired  $t$ -test with test statistic

$$t = \frac{\bar{D}}{s_D / \sqrt{n}} \quad (2)$$

where  $\bar{D}$  is the mean of the differences,  $s_D$  their standard deviation, and  $n = 363$ . In view of the distributional characteristics (positive skewness; kurtosis  $> 3$ ), the Wilcoxon signed-rank test was additionally applied as a non-parametric verification. The test statistic is

$$W^* = \sum_{i: D_i > 0} R_i^* \quad (3)$$

where  $R_i^*$  is the rank of  $|D_i|$  among all non-zero differences. Under  $H_0$ , for large  $n$ , the standardised statistic  $z = (W^* - \mu_{W^*}) / \sigma_{W^*}$  is approximately standard normal (Conover,

1999; Hollander et al., 2014). Effect size was quantified via Cohen's  $d$  for paired samples (Cohen, 1988):

$$d = \frac{\bar{D}}{s_D} \quad (4)$$

#### Regression Model: Within-Pair Difference and Complexity

To quantify the effect of the effective number of questions  $Q_i$  on the within-pair temporal difference, the simple linear regression model is estimated:

$$D_i = \alpha + \beta Q_i + \varepsilon_i, \quad \varepsilon_i \stackrel{iid}{\sim} (0, \sigma^2) \quad (5)$$

where  $\alpha$  is the intercept,  $\beta$  is the marginal effect of an additional question on the difference (in minutes), and  $\varepsilon_i$  is the error term. The coefficient  $\beta > 0$  would indicate that longer questionnaires generate larger absolute differences between modes, consistent with the hypothesis that the digital mode automates a constant fraction of time per question. Parameters are estimated by ordinary least squares (OLS), significance is evaluated through two-tailed  $t$ -tests.

#### Heterogeneity Analysis by Complexity

To examine whether the percentage reduction is homogeneous across different levels of questionnaire complexity, the sample is stratified into tertiles according to  $Q_i$ : Low (32–57 questions,  $n = 122$ ), Medium (58–76 questions,  $n = 124$ ), and High (77–210 questions,  $n = 117$ ). Descriptive statistics are computed by stratum and the hypothesis of equality of means of the absolute difference is tested via one-way analysis of variance (ANOVA), supplemented by the Pearson correlation between  $Q_i$  and the percentage reduction  $R_i$ .

## 4 Results

### 4.1 Descriptive Statistics

Table 1 presents the main descriptive statistics for both methods. The mean administration time was 139.67 minutes ( $s = 52.15$ ) for the manual method and 79.47 minutes ( $s = 29.37$ ) for the digital method, representing an absolute reduction of 60.21 minutes and a relative reduction of 43.11%. The coefficient of variation is practically identical in both groups (37.34% for manual and 36.95% for digital), indicating that relative dispersion remains stable despite the marked difference in mean levels.

Table 1

#### Descriptive statistics of administration time by method

Statistic	Manual	Tablet	Diff.	Red. (%)
Mean	139.67	79.47	60.21	43.11
Median	129.52	73.83	56.30	43.13
SD	52.15	29.37	24.33	
Variance	2719.37	862.33	591.92	
CV	37.34	36.95	40.41	
Minimum	89.47	33.32	17.81	
Maximum	417.46	228.95	188.50	
P25	104.05	59.21	42.51	
P75	165.37	95.33	73.05	
Skewness ( $g_1$ )	1.24	1.31	1.09	
Kurtosis ( $g_2$ )	2.69	2.91	2.06	
Total	50701.90	28846.80	21855.10	

CV = coefficient of variation; Diff. = difference (Manual – Tablet); Red. = percentage reduction; Time measured in minutes.

Both distributions exhibit positive skewness ( $g_1 = 1.24$  and  $g_1 = 1.31$ , respectively) and kurtosis above 3 ( $g_2 = 2.66$  and  $g_2 = 2.91$ ), indicating heavier right tails than expected under normality. This pattern is consistent with the literature on interview durations (Adali et al., 2021; Fuchs et al., 2000) and motivates the complementary use of non-parametric methods.

Figure 1 complements Table 1 with a graphical representation of the time distributions for each method. The box plot allows visual appreciation of the difference in medians, the greater dispersion of the manual method, and the presence of upper outliers in both groups, more pronounced in the paper case.

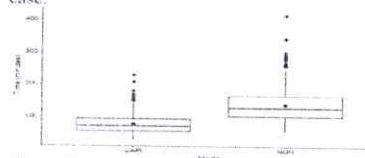


Figure 1

Comparison of survey administration time by method. Comparative box plot of administration time (in minutes) for the manual method and the digital method (tablet). Boxes represent the interquartile range (P25–P75); the central line indicates the median; points correspond to outliers. Source: own elaboration with R (R Core Team, 2024).

Figure 2 presents frequency histograms for both methods in separate panels. The distribution of the manual method is more dispersed and exhibits a longer right tail, while the

distribution of the digital method is concentrated at lower values, confirming the positive skewness pattern reported in Table 1.

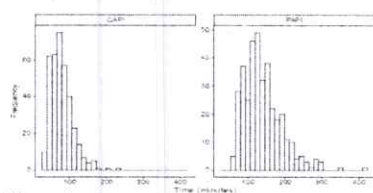


Figure 2

**Distribution of administration time by data collection method.** Frequency histograms of administration time (in minutes) by method. Left and right panels correspond to the manual and digital methods, respectively. Source: own elaboration with R (R Core Team, 2024).

#### 4.2 Distribution of Within-Pair Differences

Table 1 summarises the distribution of the difference variable ( $D_i = X_{i(\text{manual})} - X_{i(\text{tablet})}$ ) for the 363 pairs. The minimum observed difference was 17.81 minutes and the maximum was 188.50 minutes, indicating that in no case was the manual method faster than the digital method. The median difference was 56.30 minutes and the mean 60.21 minutes, with positive skewness ( $g_1 = 1.08$ ) suggesting that some complex households generated particularly large differences.

Figure 3 shows the estimated density curves for both methods superimposed on the same plane. The leftward shift of the digital method distribution relative to the manual method is unequivocal: most of the probability mass of the tablet is concentrated between 40 and 120 minutes, while that of the manual method extends between 80 and 250 minutes, with a notably longer right tail.

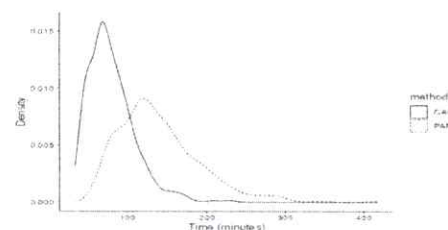


Figure 3

**Estimated density curves of administration time for manual and digital methods.** Estimated density curves (Gaussian kernel) of administration time for the manual and digital methods superimposed. The leftward shift of the digital distribution reflects the systematic reduction in collection time. Source: own elaboration with R (R Core Team, 2024).

#### 4.3 Hypothesis Tests and Effect Size

Table 2 summarises the inferential results. The paired  $t$ -test yielded  $t(362) = 47.15$  ( $p < .001$ ), broadly rejecting the null hypothesis (1). The Wilcoxon signed-rank test ( $V = 66,066$ ,  $p < .001$ ) confirms this result without assuming normality of the differences. Cohen's  $d$  is 2.475 (95% CI: [2.265, 2.680]), and the 95% confidence interval for  $\mu_D$  is [57.70, 62.72] minutes.

Table 2

Results of hypothesis tests and effect size estimation

Test	Statistic	df	p-value	Lower	Upper	Units
Paired $t$ -test	47.15***	362	< 0.001			
Wilcoxon test	66066***		< 0.001			
Cohen's $d$	2.475			2.265	2.680	
Mean difference ( $\mu_D$ )				57.70	62.72	min
Mean reduction	42.78					%

$H_0: \mu_D = 0$ . CI = confidence interval. \*\*\*  $p < 0.001$ .

The convergence of both tests—parametric and non-parametric—substantially strengthens inferential confidence, as recommended by Wilcox (2012) when data depart from normality. The value  $V = 66,066$  of the Wilcoxon test constitutes particularly solid evidence of the systematic superiority of the digital method.

#### 4.4 Regression of the Difference on Number of Questions

Table 3 presents the results of the simple linear regression model (5). The estimated coefficient for number of ques-

tions is  $\hat{\beta} = 0.824$  (SE = 0.018,  $t = 46.10$ ,  $p < .001$ ), indicating that each additional question in the questionnaire is associated with approximately 49.4 seconds of greater time difference between the manual and digital methods. The intercept is not significantly different from zero ( $\hat{\alpha} = 2.356$ ,  $p = .081$ ), consistent with the hypothesis that the time difference between modes is essentially zero when the questionnaire is empty. The model explains 88.48% of the variance of the within-pair difference ( $R^2 = .885$ ), revealing that questionnaire complexity is the overwhelming determinant of the magnitude of the time saving.

Table 3

Simple linear regression: within-pair difference on effective number of questions

Variable	(1)
Intercept	2.356 (1.346)
Number of questions ( $Q_i$ )	0.824*** (0.018)
$R^2$	0.885

Dependent variable:  $D_i$  (difference in minutes; Manual – Tablet). Standard errors in parentheses. \*\*\*  $p < 0.001$ .

Figure 4 presents the scatter plot of the within-pair difference as a function of the number of questions, together with the fitted regression line. The relationship is markedly linear, with no visual evidence of severe heteroscedasticity or systematic non-linearity.

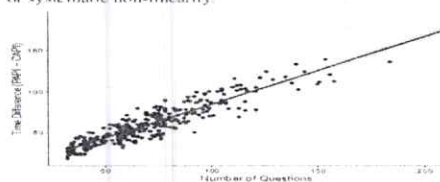


Figure 4

**Linear regression of within-pair time difference on number of effective questions.** Scatter plot of the within-pair difference (Manual – Tablet, in minutes) as a function of the effective number of questions ( $Q_i$ ), with OLS regression line ( $R^2 = .885$ ). Source: own elaboration with R (R Core Team, 2024).

#### 4.5 Heterogeneity Analysis by Complexity Tertiles

Table 4 disaggregates administration times and the within-pair difference according to three questionnaire complexity strata defined by tertiles of the effective number of questions. The results reveal a systematic pattern: the absolute difference between methods scales monotonically with complexity—from 38.82 minutes in the Low tertile to 57.03 in the Medium and 85.87 in the High—a difference confirmed by ANCOVA ( $F(2, 360) = 301.20$ ,  $p < .001$ ). In contrast, the percentage reduction remains remarkably constant across the three strata: 42.24%, 42.95%, and 43.18%, respectively. The Pearson correlation between number of questions and percentage reduction is  $r = .077$  ( $p = .141$ ), non-significant at the 5% level, formally confirming that the proportion of time saved by the digital mode does not depend on questionnaire complexity.

Table 4

Administration time statistics by questionnaire complexity tertile

Tertile ( $Q_i$ )	Manual (min)		Tablet (min)	
	Mean	SD	Mean	SD
Low (32–87), $n = 122$	91.23	16.49	52.41	9.00
Medium (58–76), $n = 124$	132.09	14.33	75.06	7.96
High (77–210), $n = 117$	198.23	44.69	112.36	25.64
ANCOVA $F(2, 360)$	301.20***			
Pearson $r(Q_i, R)$	0.077 ( $p = .141$ )			

$Q_i$  = effective number of questions. \*\*\*  $p < .001$ .

## 5 Discussion

The results of this study confirm, in the context of a large-scale household survey in the Piura Region, Peru, the operational superiority of the digital method over the manual method in terms of time efficiency. The 43.13% reduction in mean administration time is consistent with the findings of Adoh et al. (2021) and exceeds those reported by Fuchs et al. (2000) and Ballivian et al. (2015), possibly due to the higher average number of questions per household in the present instrument and the interviewers' familiarity with the device.

**Mechanisms of the temporal reduction.** Several mechanisms may explain this reduction. First, CAPT systems automate the navigation of skip patterns, eliminating the time interviewers would otherwise spend manually identifying and skipping inapplicable sections (de Leeuw et al., 2008). Second, automatic range checks prevent re-readings caused by out-of-range entries (R. P. Baker et al., 1995). Third, the elimination of paper handling generates cumulative time savings that are proportionally greater in more extensive instruments, as Fuchs et al. (2000) demonstrated by examining whether CAPT or PAPI interviews take longer

across instruments of varying length. This third mechanism is precisely what explains the regression coefficient  $\beta = 0.824$  min/question: for each additional question the questionnaire requires, the digital mode saves nearly one additional minute compared to the manual mode.

**Key finding: proportionality of the gain.** The most operationally relevant result of this study is the constancy of the percentage reduction across complexity tertiles ( $r = 0.77$ ,  $p = .141$ ). This finding implies that the relative gain of the digital mode is robust and independent of whether the surveyed household is simple or complex. From the perspective of a statistical operation planner, this means that time and cost savings can be forecast with precision: a reduction close to 43% can be expected regardless of the distribution of household sizes in the sample. This property contrasts with what might be expected if the digital mode advantage were concentrated in complex households: in that case, the projected reduction would depend critically on the sample composition.

The fact that  $R^2 = .855$  in regression (5) is so high has an additional implication: the effective number of questions is a sufficiently precise predictor of administration time under the digital mode—and, by extension, of the difference between modes—to allow logistical planning based on the expected questionnaire length. Survey organisations can use the relationship  $\hat{D} = 2.36 + 0.824Q$  to project time savings before field deployment, given a demographic module of known composition.

**Operational savings.** The mean difference of 60.21 minutes per pair has direct implications for field costs. In a hypothetical operation of 500 interviewers each completing ten surveys per day, the digital method would save approximately 5,017 interviewer-hours daily; over a thirty-day fieldwork period, this would amount to 150,510 total interviewer-hours, with direct implications for the labour cost per completed survey and the overall operation timeline. Additionally, the elimination of manual transcription and the automatic transfer of digital records to the central server reduce the time between fieldwork completion and data availability, compressing the statistical production cycle (United Nations Statistics Division, 2014).

**Comparison with the international literature.** In the framework of the Caeyers et al. (2012) experiment, the time reduction observed for full CAPI compared to PAPI was of the order of 10%. The difference with the 43% found in the present study is notable and is probably attributable to three factors: (i) the questionnaire used in Piura includes list modules per household member that generate a strong dependence between household complexity and total time, amplifying the advantage of automatic routing; (ii) the mean number of questions per survey in the Peruvian sample (70.20) is significantly higher than the 84 fixed questions of the Tanzanian questionnaire, with lower within-sample variance in

the latter; and (iii) field conditions in urban areas of Piura may favour typing speed compared to rural settings in East Africa, as suggested by evidence on interviewer workload and device familiarity (Lynn et al., 2014).

**Limitations.** Among the study's limitations, the following are noted. First, the nature of the data does not allow identification of whether the time reduction translates directly into higher quality of substantive responses; future research should incorporate quality indicators, such as item nonresponse rates and internal consistency measures, as done by Caeyers et al. (2012). Second, the study lacks information on interviewers' prior experience with digital devices, a factor that could moderate the effect found (Heerwegh, 2009); interviewers more familiar with touch interfaces may achieve greater reductions, while those with limited digital literacy may initially show temporal penalties during the learning curve. Third, generalisation to rural or hard-to-reach settings requires caution, as connectivity limitations, device battery constraints, and the availability of technical support may alter the behaviour of the digital method (Caeyers et al., 2012).

Finally, the matched design of this study, which randomises the order of method administration at the dwelling level, constitutes a relevant methodological strength compared to uncontrolled comparisons. This decision eliminates selection bias between groups and allows causal attribution of the observed differences to the data collection method rather than to the differential characteristics of the interviewed households (Tourangeau et al., 2013).

## 6 Conclusions

This study has demonstrated, based on 363 matched observations collected in the Piura Region within the project funded by the Universidad Nacional de Piura (Resolution No. 385-CU-2024) (Hidalgo et al., 2024), that the digital tablet data collection method reduces the administration time of a demographic household survey by 43.11% compared to the paper manual method. The mean difference of 60.21 minutes is statistically significant ( $p < .001$ ) and of large magnitude according to Cohen's  $d$  ( $d = 2.475$ ). Both tests—parametric ( $t = 47.15$ ) and non-parametric ( $V = 66.066$ )—converge in rejecting the null hypothesis with  $p < .001$ , and the Wilcoxon statistic  $V = 66.066$  constitutes particularly compelling evidence that the advantage of the digital mode was universal across all observed pairs.

Beyond the global comparison, the present study makes two new contributions to the literature. First, it demonstrates that the effective number of questionnaire questions predicts 85.48% of the variance of the within-pair difference ( $\beta = 0.824$  min/question,  $p < .001$ ), providing a directly applicable tool for planning digital field operations: expected savings can be projected with precision from the anticipated questionnaire length. Second, it demonstrates that the percentage reduction is statistically constant across all levels of

household complexity ( $r = .077$ ,  $p = .141$ ), meaning that the efficiency gains of the digital mode are predictable and robust regardless of sample composition.

These results support the adoption of CAPI technology in official household survey operations, especially in contexts where field cost, processing speed, and respondent burden are critical constraints (Couper, 2017; United Nations Statistics Division, 2014). The consistency of the effect across all pairs and the proportionality of the gain at different complexity levels suggest that the time savings are robust, operationally predictable, and generalisable to instruments of different lengths.

Future research should expand this analysis by incorporating substantive data quality indicators (item nonresponse rates, internal consistency, and transcription errors) and moderating variables such as interviewers' digital experience, following the framework of Caeyers et al. (2012). Extending the matched design to rural and peri-urban settings would allow assessment of the robustness of the findings under greater logistical complexity (Ballivian et al., 2015).

### Acknowledgements

The authors acknowledge the support of the Universidad Nacional de Piura under Research Project Resolution No. 385-CU-2024. The fieldwork team is thanked for the careful collection of matched observations.

### References

- Adali, T., Türkülmüz, A. S., & Lepkowski, J. M. (2021). Evaluating the demographic and health surveys mode switch from PAPI to CAPI: An experiment from Turkey. *Social Science Computer Review*, 40(6). <https://doi.org/10.1177/0894432321100956>
- Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., Dillman, D., Frankel, M. R., Garland, P., Groves, R. M., Kennedy, C., Kosnick, J., Lavrakas, P. J., Lee, S., Link, M., Piekarski, L., Rao, K., Thomas, R. K., & Zahs, D. (2010). Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly*, 74(4), 711–781. <https://doi.org/10.1093/poq/nfq048>
- Baker, R. P., Bradburn, N. M., & Johnson, R. A. (1995). Computer-assisted personal interviewing: An experimental evaluation of data quality and cost. *Journal of Official Statistics*, 11(4), 413–431.
- Ballivian, A., Azevedo, J. P., & Darbin, W. (2015). Using mobile phones for high-frequency data collection. In P. de Pádua, D. Tonielli, & R. Prater (Eds.), *Mobile research methods: Opportunities and challenges of mobile research methodologies* (pp. 21–39). Ubiquity Press. <https://doi.org/10.5334/bar.e>

- Caeyers, B., Chalmers, N., & De Weerd, J. (2012). Improving consumption measurement and other survey data through CAPI: Evidence from a randomised experiment. *Journal of Development Economics*, 95(1), 19–33. <https://doi.org/10.1016/j.jdevco.2011.12.001>
- Catlin, G., & Ingram, S. (1988). The effects of CATI on costs and data quality: A comparison of CATI and paper methods in centralized interviewing. In *Telephone survey methodology*. Wiley.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9780203771587>
- Couper, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). Wiley.
- Couper, M. P. (2017). New developments in survey data collection. *Annual Review of Sociology*, 43, 121–145. <https://doi.org/10.1146/annurev-soc-060116-053643>
- de Leeuw, E. D., Hox, J. J., & Dillman, D. A. (Eds.). (2008). *International handbook of survey methodology*. Taylor & Francis. <https://doi.org/10.4324/9780203843124>
- Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511761676>
- Fuchs, M., Couper, M., & Hansen, S. E. (2000). Technology effects: Do CAPI or PAPI interviews take longer? *Journal of Official Statistics*, 16(3), 273–286.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Wiley.
- Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research*, 21(1), 111–121. <https://doi.org/10.1093/ijpor/edn054>
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, 72(5), 836–846. <https://doi.org/10.1093/poq/nfn045>
- Hidalgo, D., Ipanaque, G., & Armas Juárez, R. (2024). *Implementation of an automated digital system to measure sociodemographic parameters in real time in the Piura Region* [Basic and Applied Research Project 2024, approved by University Council Resolution No. 385-CU-2024, Universidad Nacional de Piura, Piura, Peru].
- Hollander, M., Wolfe, D. A., & Chicken, E. (2014). *Non-parametric statistical methods* (3rd ed.). Wiley.