

UNIVERSIDAD PERUANA UNIÓN

ESCUELA DE POSGRADO

Unidad de Posgrado de Ingeniería y Arquitectura



**Abordaje estadístico para la evaluación de la generación de
residuos sólidos domiciliarios en hogares peruanos:**

2014 – 2021

Tesis para obtener el Título de Segunda Especialidad Profesional en
Estadística Aplicada para Investigación

Autor:

Jorge Luis Padilla Vento

Asesor:

Dr. Juan Jesús Soria Quijaite

Lima, octubre de 2024

DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

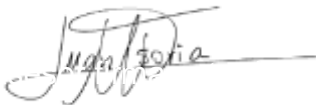
Yo Juan Jesús Soria Quijaite, docente de la Unidad de Posgrado de la Unidad de Posgrado de Ingeniería y Arquitectura, Escuela de Posgrado de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: **“Abordaje estadístico para la evaluación de la generación de residuos sólidos domiciliarios en hogares peruanos: 2014 – 2021”** del autor del autor Jorge Luis Padilla Vento tiene un índice de similitud de 2 % verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad

u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima, a los 11 días del mes de noviembre del año 2024



Dr. Juan Jesús Soria Quijaite

ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Ñaña, Villa unión a 3 días del mes de octubre del año 2024, siendo las 8:40 horas, se reunieron de forma online sincrónica, bajo la dirección del presidente del jurado Dr. Josué Edison Turpo Chaparro, el secretario Mtro. Carlos Daniel Abanto Ramirez; los demás miembros: y el Mg. Lizeth Huanca López y el PhD. Javier Linkolk López Gonzales y el asesor Dr. Juan Jesús Soria Quijaite con el propósito de administrar el acto académico de sustentación de Tesis de la Segunda Especialidad titulada "Abordaje estadístico para la evaluación de la generación de residuos sólidos domiciliarios en hogares peruanos: 2014 – 2021", conducente a la obtención del Título de Segunda Especialidad Profesional de ingeniería: Estadística Aplicada para Investigación.

El presidente inició el acto académico de sustentación invitando al candidato hacer uso del tiempo determinado para su exposición. Concluido la exposición, el Presidente invitó a los demás miembros del Jurado a efectuar las preguntas, cuestionamientos y aclaraciones pertinentes, los cuales fueron absueltos por el candidato. Luego se produjo un receso para las deliberaciones y la emisión del dictaminador del Jurado.

Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Candidato: Jorge Luis Padilla Vento

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	19	A	Nominación de excelente	Excelencia

Finalmente, el Presidente del Jurado invitó al candidato a ponerse de pie, para recibir la evaluación final. Además, el Presidente del Jurado concluyó el acto académico de sustentación, procediéndose a registrar a registrar las firmas respectivas.



Presidente



Secretario



Asesor



Miembro



Miembro



Candidato

Resumen

Este estudio aborda la evaluación de la generación de residuos sólidos domiciliarios en los hogares peruanos utilizando técnicas estadísticas y la metodología de minería de datos SEMMA y PCA. El objetivo es explorar cómo la gestión de residuos, la población y el índice GPC per cápita influyen en la producción de estos residuos en los departamentos peruanos. La muestra se obtuvo de la base de datos de informes anuales presentados por los municipios distritales y provinciales al MINAM a través del Sistema de Información para la Gestión de Residuos Sólidos (SIGERSOL), incluyendo datos de los 24 departamentos del Perú, con un total de 14,852 registros organizados en 196 fichas de registro. Se aplicaron técnicas estadísticas y la adaptación de la metodología SEMMA junto con el Análisis de Componentes Principales (PCA), El ANOVA de Welch mostró diferencias significativas en la generación de residuos sólidos domiciliarios entre los departamentos peruanos [$F(6, 94.310) = 790.444$; $p = 0.0$, $p < 0.05$]. Además, un Eta cuadrado de 99.09% Se encontró una relación positiva entre la producción de residuos sólidos domiciliarios RSD y el número de habitantes. Lima destacó con el mayor promedio de RSD 13,220.47 toneladas y un índice per cápita GPC del 50%. En el Grupo a Lambayeque 5,616.48 toneladas, Loreto 2,946.44 toneladas y San Martín 1,596.07 toneladas registraron los promedios de RSD más altos, mientras que Amazonas 441.1 toneladas obtuvo el más bajo. Ucayali 60%, Loreto 58% y San Martín 57% mostraron los índices GPC más altos. En el Grupo b Ayacucho 701.81 toneladas tuvo el promedio de RSD más alto y Apurímac 497 toneladas el más bajo. Tacna y Apurímac con 44% y Moquegua 43% registraron los índices GPC más altos, mientras que Huancavelica 42% y Pasco 41% tuvieron los más bajos. En el Grupo C Piura 4,476.53 toneladas y La Libertad 3,478.46 toneladas mostraron los promedios de RSD más altos, mientras que Huánuco 859.41 toneladas y Cajamarca 812.74 toneladas registraron los más bajos. Ica y Piura lideraron con un GPC promedio del 48%, mientras que Puno y Junín con 43% tuvieron los valores más bajos.

1 Introduction

La generación y gestión de residuos sólidos domiciliarios es un problema mundial que se agrava por factores como el crecimiento demográfico, la expansión económica (Abushammala y Ghulam, 2023). C. H. Li et al. (2023) anticipan un aumento en la generación de desechos en los futuros años, considerando que la población global ya supera los 8 mil millones de habitantes. Li y Wang (2022) resaltan que este fenómeno potencia la producción de residuos sólidos domésticos. Cheng et al. (2022) y Pillay et al. (2023) han sentado una base teórica al tratar temas asociados a la influencia de factores psicológicos, sociales y económicos en la conducta de segregación de residuos. Adicionalmente, se han reconocido teorías significativas, tales como la evaluación crítica de los estudios relacionados con la educación moral y ambiental (Begum et al., 2022), y la relevancia de la humildad como una virtud esencial para la sensibilidad ambiental (Bonnett, 2023).

El Ministerio del Ambiente MINAM (2023) señala que en Perú El 70% de los residuos municipales son producidos por hogares; el diario El Peruano (2024) informa que durante el año 2021 se generaron aproximadamente 8 millones 214,355.90 toneladas de desperdicios sólidos municipales, lo que equivale a un promedio de 22,505.08 toneladas por día. Estos residuos estaban compuestos en un 56.70% por materia orgánica, 20.94% inorgánica, 12.66% no aprovechable y un 9.71% correspondiente a desechos peligrosos. El medio Gestión (2023) reporta que el MINAM ha instruido a las autoridades locales de San Vicente de Cañete, en Lima, para que elaboren un plan de acción que permita enfrentar el problema de la recolección de los residuos sólidos (RS) en dicho distrito.

A pesar de los esfuerzos realizados, persisten lagunas en la comprensión de la relación entre El índice GPC per cápita y la producción de RS domiciliarios. Esta brecha de conocimiento conlleva una deficiente gestión de RS y una infraestructura inadecuada (W. Li & Wang, 2022). A esto se suma que la pandemia ha alterado la composición de los desechos y ha aumentado la demanda de plásticos (Nayak & Uppal, 2023). Requena-Sanchez et al. (2023) señalan que la falta de información esencial para entender la dinámica de la administración de RS dificulta la puesta en marcha de estrategias; Zulkipli et al. (2022) resaltan la importancia de la gestión de RS y el cuidado del medio ambiente.

En este contexto, el presente estudio se propone llenar estos vacíos de conocimiento y proporcionar una visión más completa y detallada de la problemática de la generación de RS domiciliarios en los hogares peruanos. Para lograr esto, hemos recurrido al abordaje estadístico para la evaluación de datos con el apoyo de la metodología de minería de datos SEMMA, con el fin de mejorar la interpretación de datos y así optimizar la toma de decisiones en la gestión de RS. En este sentido, el principal aporte de este trabajo es ofrecer un esquema de evaluación de datos que permita visualizar la producción de RS y se espera explicar la influencia del índice GPC per cápita y el nivel de población en la generación de RS domiciliarios en los hogares peruanos, constituyendo una herramienta valiosa para la visualización y focalización de políticas de apoyo a la gestión de RS, ofreciendo un enfoque integral para observar la distribución de los RS domiciliarios. Esta técnica consta de las etapas: Muestra, Explora, Modifica, Modela y Evalúa, establecida por el Instituto SAS, para realizar las tareas esenciales que faciliten la comprensión, la estructuración, la implementación y el mantenimiento de proyectos de minería de datos

(Tapia, 2024). También utilizamos el análisis de componentes principales (PCA) es una de las estrategias de aprendizaje sin supervisión se utilizan en el análisis exploratorio de datos (Martínez, 2018). El Minan se compromete con la Agenda 2030 para el Desarrollo Sostenible (SIAR, 2024) y nuestra investigación se alinea específicamente con el decimotercer objetivo de Desarrollo Sostenible que consiste en emprender acciones urgentes para lidiar con el cambio climático y sus repercusiones. El resto de este estudio se organiza de la siguiente forma: la Sección 2 expone los diversos trabajos relacionados que preceden a este estudio. La Sección 3 explica la adaptación de la metodología SEMMA de la minería de datos empleada para una interpretación más efectiva de los datos de la producción de RS domésticos. para la etapa de evaluación solo se consideró la interpretación de los resultados, se generó clústeres por el método de Ward y se realizó PCA para una mejor observación de la totalidad de los residuos sólidos. la Sección 4 se contrasta la discusión entre los resultados obtenidos y otros estudios semejantes. la Sección 5 expone las conclusiones.

2. Trabajo relacionado

Zhong et al. (2022) propone un conjunto de modelos generales de análisis de eventos de ciberseguridad combinados con la visualización de ciberseguridad y el modelo analítico clásico SEMMA Para estandarizar el proceso de análisis. se implementan la segmentación de activos de red, la extracción de eventos anómalos de red y la evolución de eventos de ataque mediante la sinergia de múltiples vistas. Los resultados muestran la validez del modelo analítico. En 2017, China decidió introducir nuevas normativas para la categorización de los RS urbanos. Mediante un análisis de texto y un modelo de Coherencia de Políticas Medidas (PMC), se concluyó que ninguna de las 11 políticas evaluadas era de alta calidad o de poca calidad, obteniendo un promedio de 7.71 en el índice PMC (Xu et al., 2023). Popli et al. (2021) construyen un modelo de predicción de RS urbanos en Laos desde 1995 hasta 2050 usando regresión multilínea (MLR) con seis parámetros socioeconómicos. Se crean modelos en cuatro escenarios y se evalúa su precisión con R^2 para el rendimiento. El modelo incluye población y PIB per cápita, es el más efectivo para estimar la tasa de generación de residuos. Se proyecta que los residuos sólidos urbanos llegarán a 0,98 millones de toneladas en 2030, 1,26 en 2040 y 1,52 en 2050. Asimismo, Karadimas et al. (2007) utilizó el Algoritmo Genético (GA) para identificar rutas óptimas en la recolección de Residuos Sólidos Municipales (RSU). Considerando que entre el 60 y el 80% del gasto en gestión de residuos se destina a la recolección, una pequeña mejora en esta etapa puede generar un ahorro considerable.

Fontaine et al. (2024) desarrollaron modelos predictivos para la gestión de residuos al final del ciclo de vida de los productos, subrayando la relevancia de un análisis previo detallado; se presenta un estudio de caso en Gatineau. Este modelo permite evaluar el impacto ambiental de las estrategias de gestión de residuos los datos revelaron una desviación estándar real un 84 % mayor que la prevista calculando medidas como asimetría 0,072 que indica una distribución simétrica. Del mismo modo, Rifaldi et al. (2021) llevó a cabo un estudio en Bekasi Regency, utilizando un enfoque descriptivo cualitativo y técnicas de diagrama de bucle causal (CLD) para entender el sistema de gestión de residuos. Los resultados señalan que la pandemia afecta variables como los patrones de consumo, las condiciones económicas y el volumen de residuos, se identificaron factores como el número de personas, bienestar y salud públicos, consumo público, volumen de residuos, área de vertedero, costos de gestión de residuos, tecnología, condiciones económicas y presupuestos gubernamentales

En su investigación, Requena-Sanchez et al.(2023) emplearon un formulario online para reunir y generar información sobre la producción y características de los residuos sólidos domésticos en América Latina. Se descubrió que el 43% de los encuestados reportó un incremento en el volumen total de sus desechos. La pandemia de COVID-19 ha exacerbado el uso y desecho de equipos de protección personal (EPP), un análisis en Lima, Perú, reveló la presencia de 138 elementos de EPP en 11 playas durante un período de 12 semanas. Las playas de recreo fueron las más perjudicadas, seguidas por las áreas de surf y pesca. Este fenómeno podría estar asociado a la alta densidad de visitantes en estos lugares. La mayoría del EPP fue descartado por los visitantes de la playa (De-la-Torre et al., 2021). Asimismo, se utilizó la técnica de Liderazgo Comunitario en Saneamiento Ambiental Urbano en una pequeña comunidad en Malawi para evaluar su factibilidad técnica, política y económica. Los hallazgos mostraron que, aunque la calidad y el acceso al agua eran satisfactorios, la cantidad de agua disponible era insuficiente y la administración de residuos necesitaba mejoras (Mtika & Tilley, 2020).

Maalouf y Mavropoulos (2023) definen los residuos sólidos como la suma de materiales no líquidos ni gaseosos que se desechan, tales como embalajes, papel, plástico, vidrio, metal y sobras de comida. En la misma línea, el MINAM (2023) alude a los desperdicios generados en toneladas por acciones de limpieza en ámbitos comerciales, institucionales y espacios públicos. Según el Centro Panamericano de Ingeniería Sanitaria y Ciencias del Ambiente (CEPIS) el 70% de los residuos sólidos municipales proceden de las viviendas, mientras que el 30% restante pertenece a los residuos no domésticos del total de los desechos municipales.

La gestión de RS implica todas las acciones desde la generación hasta la eliminación final de los RS. MINAM (2023) examina la generación de residuos por habitante en el Perú. Patrao y Karnik (2023) destacan la importancia de la conciencia sobre el reciclaje y los efectos negativos de una mala gestión de los residuos electrónicos. Mumtaz et al. (2023) en Polonia demuestran que una adecuada gestión de los Residuos Sólidos Urbanos puede tener un impacto económico significativo, y sugieren la oxidación húmeda como una técnica eficaz para manejar grandes cantidades de RS. También se plantea la reutilización de los residuos en ingeniería geotécnica como una solución a la creciente producción de residuos sólidos debido a la urbanización y el crecimiento demográfico (Liu & Hung, 2023). Fernández-Domínguez et al. (2021) se analizaron 28 pares de sustrato/digestato en términos de calidad de RS orgánica. Utilizando análisis de componentes principales y agrupamiento jerárquico, los resultados indicaron una variabilidad del 58,02% en la accesibilidad y una disminución de la accesibilidad entre el 16-66% y un aumento de la complejidad entre el 34-98%. Además, se observó un incremento de compuestos no biodegradables entre el 17-66%. Esta información se utilizó para un modelo que predice la mineralización de carbono orgánico en el suelo. Kim y de Vasconcelos Barros (2023) analizó los efectos de la Política Nacional de Residuos Sólidos (PNRS) en los municipios populosos de Minas Gerais, Brasil, mediante PCA. se encontró que dos componentes representaban más del 70% de la variabilidad de los datos. Los componentes estaban relacionados con los objetivos y principios del PNRS. Zhang et al. (2022) aplicó una novedosa nariz electrónica (EN) para extraer información del gas del espacio de cabeza del lixiviado de incineración de residuos sólidos municipales. Combinando múltiples técnicas de aprendizaje automático y PCA, el modelo UMAP-XGBT basado en EN tuvo el mejor rendimiento de clasificación, con una tasa de precisión del 99,95% en el conjunto de entrenamiento y del 95,83% en el de prueba.

3. Hipótesis

La investigación tiene como hipótesis Existen diferencias significativas en la generación de residuos sólidos domiciliarios entre los departamentos de Perú durante el periodo 2014-2021; La cantidad de población influye significativamente en la generación de residuos sólidos domiciliarios en los departamentos del Perú en el periodo 2014 al 2021; El índice GPC per cápita influye significativamente en la generación de residuos sólidos domiciliarios de los departamentos en Perú en el periodo 2014 al 2021 y Existe correlación entre los diferentes cantidades de residuos sólidos y la cantidad de habitantes de los departamentos en Perú durante el período de 2014 a 2021.

4. Objetivo

Los objetivos de la investigación es formular un esquema de evaluación de datos que adapte la metodología SEMMA explorando la influencia de la gestión de la producción de residuos sólidos domiciliarios en los departamentos de Perú durante el periodo 2014-2021; Determinar la influencia de la cantidad de población en la generación de residuos sólidos domiciliarios en los departamentos del Perú durante el periodo de 2014 a 2021; Evaluar cómo el índice GPC (Gasto Per Cápita) influye significativamente en la generación de residuos sólidos domiciliarios en los departamentos del Perú durante el periodo de 2014 a 2021 y Examinar la correlación entre los diferentes cantidades de residuos sólidos y la cantidad de habitantes en los departamentos del Perú durante el periodo de 2014 a 2021.

3. Metodología de Minería de datos

3.1 Muestra

Esta etapa implica la selección de un subconjunto representativo de la población de estudio para realizar un análisis exploratorio (Shin et al., 2024). La muestra se obtuvo de la base de datos provenientes de los informes anuales presentados por los municipios distritales y provinciales al MINAM a través del Sistema de Información para la Gestión de Residuos Sólidos (SIGERSOL); disponible en la plataforma nacional de datos abiertos, En cumplimiento del D.S. 157-2021-PCM, se encuentra disponible para las organizaciones de la sociedad civil, la ciudadanía, la academia y el sector privado. Cada registro, que representa una localidad específica, se consideró como unidad de análisis y contiene información sobre la cantidad de habitantes y la generación de residuos sólidos (RS) en dicha localidad. La población de interés para este estudio se definió como todos los hogares peruanos que produjeron residuos sólidos entre 2014 y 2021. Para asegurar la representatividad de la muestra, se incluyeron datos de todos los 24 departamentos del Perú. Se incluyeron los siguientes departamentos: Amazonas, Ancash, Apurímac, Arequipa, Ayacucho, Cajamarca, Cusco, Huancavelica, Huánuco, Ica, Junín, La Libertad, Lambayeque, Lima, Loreto, Madre de Dios, Moquegua, Pasco, Piura, Puno, San Martín, Tacna, Tumbes y Ucayali.

Se excluyeron los datos de la Provincia Constitucional del Callao (56 registros), datos sin información sobre la cantidad de residuos (41 registros) y datos con falta de consistencia lógica (29 registros) Tras la aplicación de estos criterios, de un total de 14,978 registros iniciales. La muestra se conformó por un total de 14,852 registros. Estos registros se organizaron en 196 fichas de registro, con 8 fichas por cada departamento. Por lo tanto, la muestra para este estudio quedó definida en 196 fichas de registro.

3.2 Análisis Exploratorio de Datos (EDA)

En la Figura 1 parte (a) revela que la mayor producción de RS domésticos se halla en Lima, superando las 200 mil toneladas al año, mientras que las menores cantidades se encuentran en Huancavelica, Amazonas, Apurímac y Ayacucho; (b) los departamentos que tuvieron el mayor número de registros son Lima, Ancash y Cajamarca. Los que poseen una menor frecuencia son Pasco, Tacna, Moquegua, Ucayali, Tumbes y Madre de Dios (c) Existen fuertes correlaciones entre los RS domésticos, los RS no domésticos y los RS municipales; sin embargo, la correlación de estos con el índice GPC doméstico es muy débil.

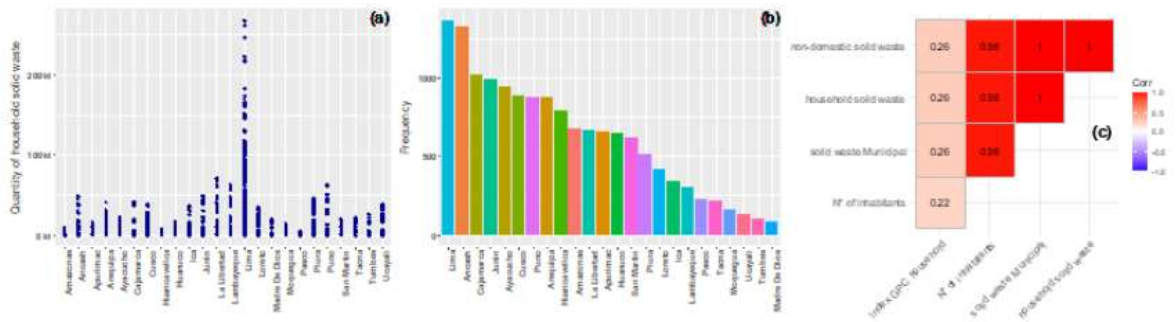


Figura 1. Características de los RS en el período comprendido entre 2014 y 2021: (a) Gráfico de dispersión de la cantidad de RS domiciliarios por departamento; (b) Gráfico de barras de la frecuencia de registro de RS por departamento; y (c) Gráfico de correlación entre variables de RS.

La Figura 2 muestra una relación positiva entre la producción promedio anual de RS domiciliarios y el tamaño de la población de cada departamento. Los cinco departamentos con mayor producción anual de RS domiciliarios son Lima (13,220.47 toneladas y 59,813 habitantes en promedio), Ucayali (5,667.83 toneladas y 31,756 habitantes), Lambayeque (5,616.48 toneladas y 33,710 habitantes), Piura (4,476.53 toneladas y 29,579 habitantes) y La Libertad (3,478.46 toneladas y 23,261 habitantes). Por otro lado, los departamentos con menor producción anual de RS domiciliarios son Huancavelica (265.95 toneladas y 4,319 habitantes en promedio), Amazonas (441.10 toneladas y 5,046 habitantes), Apurímac (497.00 toneladas y 5,349 habitantes) y Ayacucho (701.81 toneladas y 5,688 habitantes).

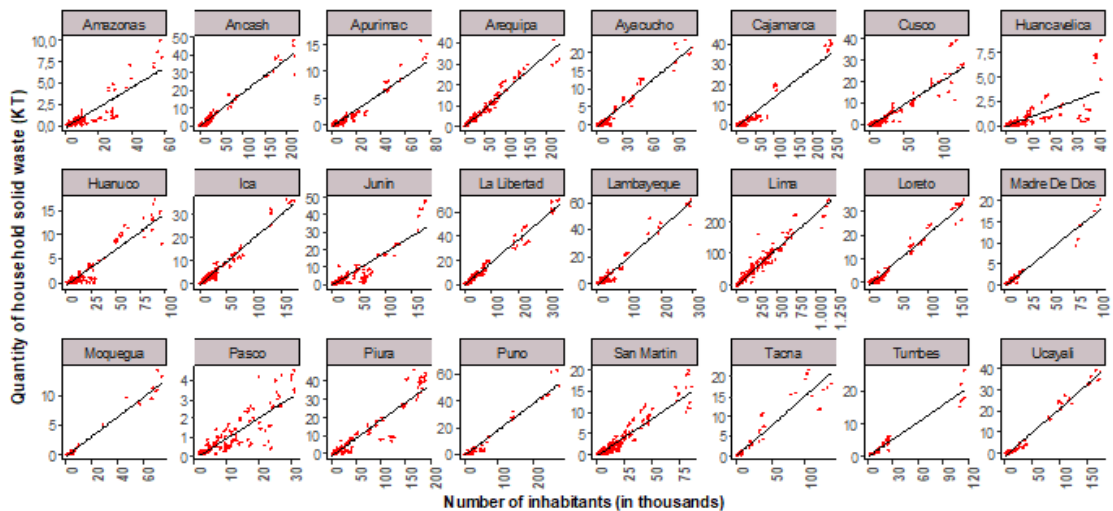


Figura 2: Grafica de dispersión de la cantidad de residuos por población

En la tabla 1 proporciona estadísticas sobre RS. los departamentos varían en población entre 131 y 1,203,125 habitantes, con una media de 16,540 y una mediana de 4,223 habitantes. El índice per cápita GPC de hogares varían entre 0,01 y 2,00 kg/hab./día, con una distribución razonablemente simétrica en torno a las respectivas mediana y media de estas tasas de 0,44 y 0,47 kg/hab./día. La generación de RS domésticos con una media de 2.671,02 toneladas y una mediana de 207,98 toneladas, la producción de RS domésticos varía entre 2,62 y 267.875,78 toneladas, lo que indica un sesgo a favor de las zonas con mayor producción de RS. De forma similar, los RS no doméstica tenía una media de 1.144,72 toneladas y un rango de 1,12 a 114.803,91 toneladas, con 89,13 toneladas como mediana. En cuanto a los RS municipales en total, el rango va desde 3.80 hasta 382,679.70 toneladas, con una mediana de 297.10 toneladas y un promedio de 3,815.70 toneladas.

Tabla 1 Estadística de resumen de la Gestión de Residuos sólidos

	Nº of inhabitants	GPC	household SW (tons)	non-domestic SW (tons)	Municipal SW (tons)
Min.	131	0.01	2.62	1.12	3.8
1st Qu.	1,675	0.43	85.46	36.62	122.1
Median	4,223	0.44	207.98	89.13	297.1
Mean	16,540	0.47	2,671.02	1,144.72	3,815.70
3rd Qu.	10,919	0.51	767.23	328.81	1,096.00
Max.	1,203,125	2	267,875.78	114,803.91	382,679.70

El grupo con la mayor mediana en el índice GPC per cápita de RS domiciliarios es "46 % - 50.99% kg/hab./día ", lo que sugiere que este grupo de población produce más RS domiciliarios. En contraste, la mediana más baja se detecta en el grupo "0 % - 43.99% kg/hab./día ". Además, se percibe una menor variabilidad en el cuartil inferior de este último grupo. El grupo "46 % - 50.99% kg/hab./día " también exhibe una distribución equilibrada y su rango intercuartílico es superior en comparación a los otros grupos La Figura 3.

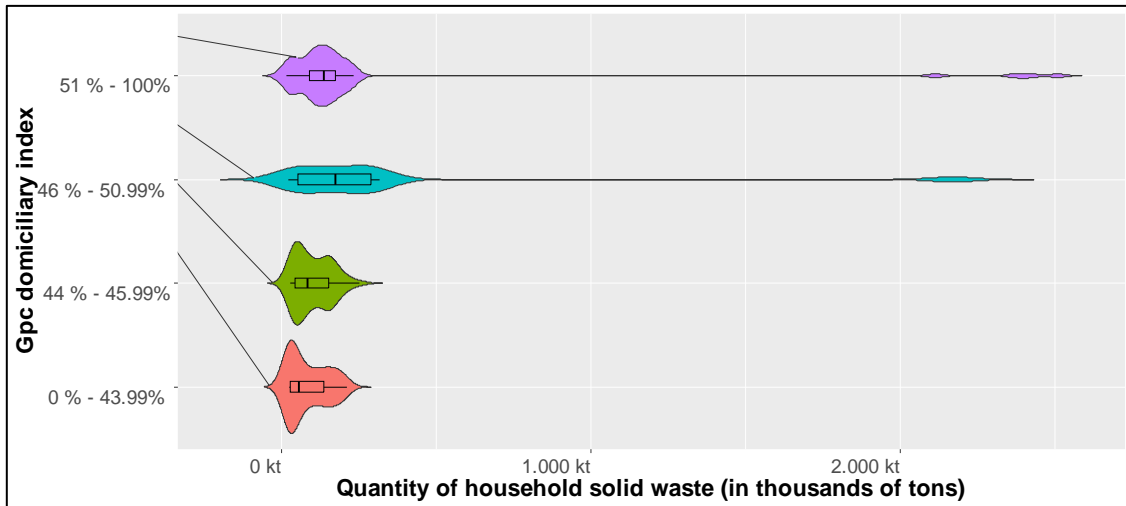


Figura 3: Grafica de violín del índice GPC per cápita de RS domiciliario.

Se observa una relación positiva entre el tamaño poblacional y la cantidad de RS domiciliarios en la Figura 4. El grupo con poblaciones entre 9,685,000 - 10,850,000 habitantes muestran la mediana más alta, mientras que el grupo de 134,000 - 256,000 habitantes exhibe la menor mediana. Además, el grupo 9,685,000 - 10,850,000 presenta una mayor variabilidad en los cuartiles y un rango intercuartílico más amplio en comparación con otros grupos. Se observa un sesgo positivo en este rango poblacional, indicando que la mayoría de los pobladores generan residuos sólidos por debajo del promedio de grupo.

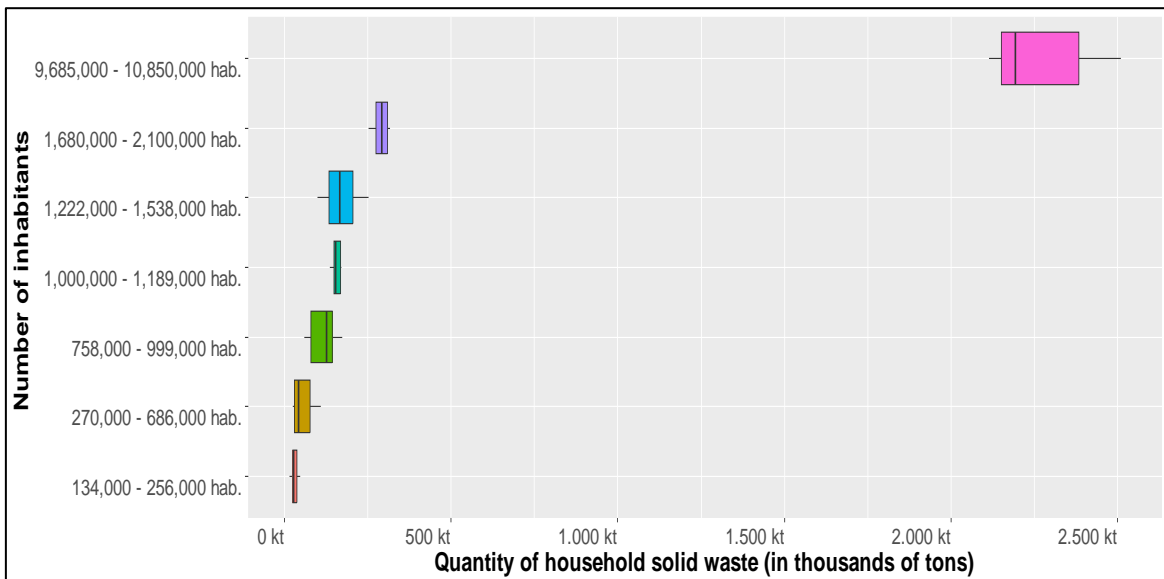


Figura 4: Grafica de caja y bigotes número de habitantes y la cantidad RS domiciliario.

4. Modificación

En esta etapa, los datos fueron transformados de tal manera que facilitaron la extracción de información. De un total de 14,852 registros, se organizaron 196 fichas de registro. Se formaron siete grupos para las poblaciones y cuatro para el índice GPC tal como se observa en la Tabla 2.

Tabla 2 Niveles y estadística de resumen de la Población y GPC

Grupos Población	Grupo GPC	n	Media	DE
134,000 - 256,000 hab.	0 % - 43.99%	6	25,135.01	2,022.73
134,000 - 256,000 hab.	44 % - 45.99%	6	34,037.52	3,817.47
134,000 - 256,000 hab.	46 % - 50.99%	7	36,443.78	10,121.25
134,000 - 256,000 hab.	51 % - 100.00%	5	21,831.59	6,752.18
270,000 - 686,000 hab.	0 % - 43.99%	25	44,479.63	24,170.17
270,000 - 686,000 hab.	44 % - 45.99%	12	48,863.24	16,127.07
270,000 - 686,000 hab.	46 % - 50.99%	6	41,477.51	9,278.57
270,000 - 686,000 hab.	51 % - 100.00%	13	73,287.75	31,087.42
758,000 - 999,000 hab.	44 % - 45.99%	10	89,838.42	33,187.30
758,000 - 999,000 hab.	46 % - 50.99%	4	131,059.28	47,258.41
758,000 - 999,000 hab.	51 % - 100.00%	11	133,864.54	23,105.03
1,000,000 - 1,189,000 hab.	0 % - 43.99%	2	150,650.96	6,447.90
1,000,000 - 1,189,000 hab.	44 % - 45.99%	6	157,749.80	14,254.05
1,000,000 - 1,189,000 hab.	51 % - 100.00%	7	156,576.32	10,120.70
1,222,000 - 1,538,000 hab.	0 % - 43.99%	15	160,846.99	32,073.34
1,222,000 - 1,538,000 hab.	44 % - 45.99%	14	154,336.29	47,829.82
1,222,000 - 1,538,000 hab.	46 % - 50.99%	11	166,091.33	53,913.85
1,222,000 - 1,538,000 hab.	51 % - 100.00%	8	213,426.19	15,427.38
1,680,000 - 2,100,000 hab.	46 % - 50.99%	16	289,563.54	20,695.12
9,685,000 - 10,850,000 hab.	46 % - 50.99%	4	2,168,068.49	36,029.00
9,685,000 - 10,850,000 hab.	51 % - 100.00%	4	2,353,332.45	169,895.15

Transformamos los datos para mejorar la visualización de comparaciones significativas. Usamos el método PCA para reducir la dimensionalidad de los datos cambiando el conjunto original de 5 variables a un conjunto(p) de 2 variables incorrelacionadas llamadas componentes principales (q). los datos modificados muestran en la Figura 5 muestra una alta correlación positiva entre los RS domiciliarios, RS no domiciliarios y RS municipales y la cantidad de habitantes; esta contribuye de gran manera a Dim1. Lima es el departamento que tiene valores muy altos en su producción de RS domiciliarios, teniendo valores atípicos. Por otra parte, Piura, La Libertad, Arequipa y Lambayeque poseen valores altos en su producción de RS domiciliarios y valores más bajos Madre de Dios, Huancavelica, Pasco y Moquegua. Por otra parte, los departamentos con mayor porcentaje GPC son Ucayali, Loreto, San Martín, Lambayeque, Amazonas y Madre de Dios. Cabe la posibilidad de que se tenga que revisar la manera en que se están registrando los datos ya que contribuye a la Dim2.

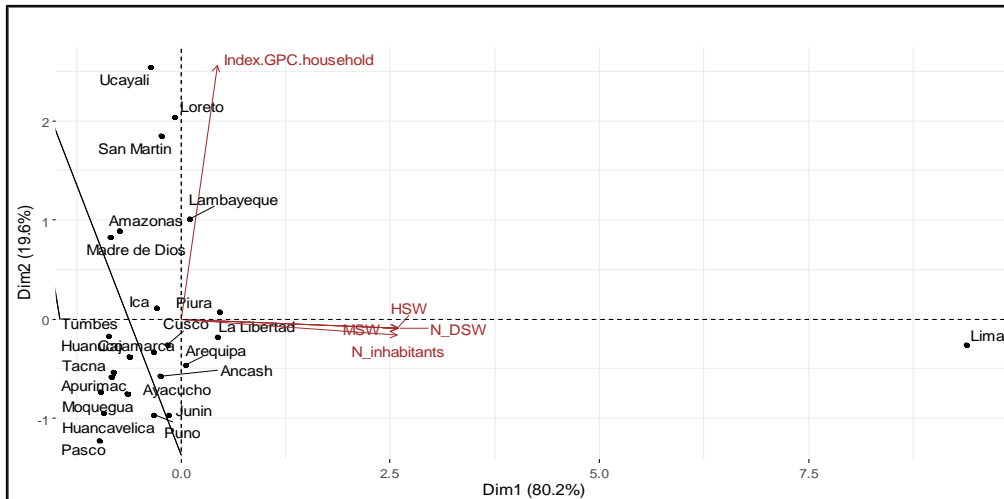


Figura 5: Gráfica PCA muestra los residuos sólidos de los departamentos del Perú. household solid waste (HSW), non-domestic solid waste (N_DSW), solid waste municipal (MSW). la primera componente explica el 80.2% de la inercia, un porcentaje significativo que muestra que esta componente principal recoge gran parte de la información esencial en los datos. Las dos primeras componentes en total representan el 99.8% de la varianza total, lo que indica que se puede preservar la mayoría de la información relevante con sólo dos dimensiones.

5. Modelación

Esta etapa implica la aplicación de diversos modelos matemáticos y estadísticos para describir y predecir patrones en los datos que se han recogido y preparado en las etapas anteriores para lo cual se seleccionó el estadístico adecuado para los datos y se verificaron los supuestos para el contraste de hipótesis. la Tabla 3 muestran que todos los datos tienen una distribución normal.

Tabla 3 Pruebas de normalidad - Shapiro-Wilk para la Cantidad RS domiciliario

Departamento	w	p-value	Departamento	w	p-value	Departamento	w	p-value
Amazonas	0.89	0.26	Huánuco	0.85	0.09	Moquegua	0.91	0.33
Ancash	0.92	0.41	Ica	0.86	0.13	Pasco	0.9	0.26
Apurímac	0.95	0.67	Junín	0.92	0.42	Piura	0.92	0.45
Arequipa	0.89	0.22	La Libertad	0.89	0.25	Puno	0.95	0.74
Ayacucho	0.86	0.12	Lambayeque	0.93	0.51	San Martín	0.89	0.25
Cajamarca	0.84	0.08	Lima	0.87	0.14	Tacna	0.93	0.49
Cusco	0.98	0.98	Loreto	0.87	0.14	Tumbes	0.88	0.19
Huancavelica	0.84	0.08	Madre de Dios	0.89	0.22	Ucayali	0.92	0.39

Verificamos los supuestos para el análisis de varianza (ANOVA):

- a) Se cumplió el supuesto de que la media de errores es cero.
- b) La varianza de los errores no es igual entre grupos, ya que la Prueba de Levene ($p=2e-16$) y la Prueba de Bartlett ($p < 2.2e-16$) demuestran que este supuesto no se cumple, evidenciando heterogeneidad de varianzas entre los grupos.
- c) Los errores están correlacionados, dado que la prueba de Durbin-Watson ($DW = 0.848$, $p\text{-value} = 2.2e-16$) muestra una autocorrelación positiva significativa en los residuos del

modelo, violando el supuesto de independencia. d) Los errores se distribuyen de acuerdo a una distribución normal, según los resultados de las pruebas de normalidad de Shapiro-Wilk aplicadas a los residuos por departamento, donde todos los valores p son mayores que el nivel de significancia de 0.05, cumpliendo este supuesto. Dada la evidencia de heterogeneidad de varianzas entre los grupos y la presencia de autocorrelación en los residuos, lo cual viola dos supuestos clave para el ANOVA tradicional, se decidió realizar el ANOVA de Welch el cual es más robusto ante violaciones de estos supuestos en lugar del ANOVA clásico (Zaiontz, 2024). Es un enfoque para tratar con varianzas heterogéneas en un análisis de varianza sobre las medias de los grupos (Howell, 2010, p. 336). se define:

- (1) K representa el número de grupos o muestras.
- (2) W representa los pesos asociados a cada grupo o muestra.

$$(3) F = \frac{\frac{1}{k-1} \sum_{j=1}^k w_j (\bar{x}_j - \bar{x}')^2}{1 + \frac{2(k-2)}{k^2-1} \sum_{j=1}^k \left(\frac{1}{n_j-1} \right) \left(1 - \frac{w_j}{w} \right)^2}$$

$$(4) w_j = \frac{n_j}{s_j^2}$$

$$(5) w = \sum_{j=1}^k w_j \bar{x}' = \frac{\sum_{j=1}^k w_j \bar{x}_j}{w}$$

$$(6) F \sim F(k-1, df)$$

$$(7) df = \frac{k^2 - 1}{3 \sum_{j=1}^k \left(\frac{1}{n_j - 1} \right) \left(1 - \frac{w_j}{w} \right)^2}$$

Donde F representa la estadística de prueba ANOVA de Welch, w_j son los pesos para cada grupo, \bar{x}_j es la media muestral del grupo j, \bar{x}' es la media ponderada general, n_j es el tamaño de la muestra del grupo j, s_j^2 es la varianza muestral del grupo j, y df son los grados de libertad para la distribución F aproximada.

Para buscar patrones y relaciones en los datos, se utilizaron las técnicas de modelización Cluster Average y Cluster Ward.D2 para identificar grupos de hogares que presentan patrones similares de generación de RS domésticos según el departamento donde se produjeron Figura 6. Cluster Average Method este método calcula la distancia promedio entre cada par de objetos en diferentes grupos. Cluster Ward.D2 Method busca minimizar la suma de las diferencias al cuadrado dentro de cada grupo.

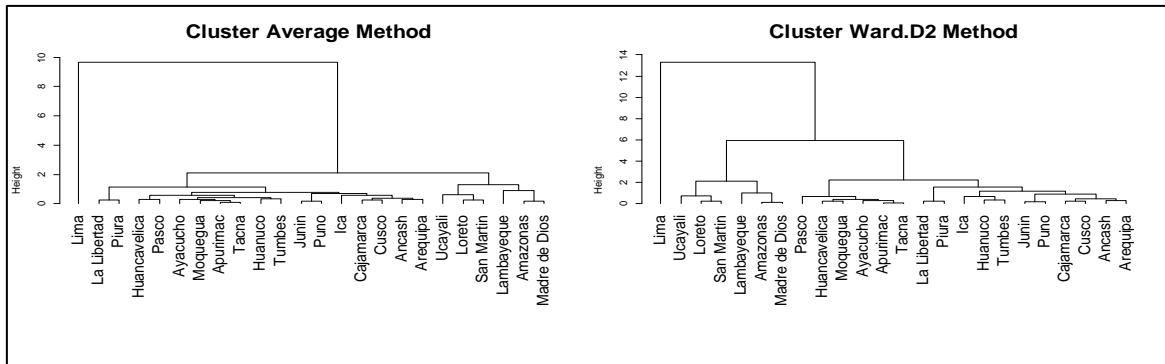


Figura 6: Dendrogram de los Cluster

6. Evaluación

Los resultados del análisis estadístico respaldan la hipótesis de que la cantidad de población influye significativamente en la generación de residuos sólidos domiciliarios en los departamentos del Perú durante el periodo 2014-2021. El Anova de Welch [$F(6, 94,310) = 790,444$; $p = 0,0$, $p < 0,05$] y un Eta cuadrado de 0,9909 revelan un efecto muy grande de la Cantidad de población sobre la generación de RS domiciliarios entre departamentos. el análisis post hoc mediante la prueba Games-Howell mostró diferencias estadísticamente relevantes ($p < 0,05$) en los volúmenes de RS domésticos producidos entre la mayoría de los pares de grupos de población comparados. Sin embargo, se encontró una notable excepción entre los grupos "1.000.000 - 1.189.000 habitantes" y "1.222.000 - 1.538.000 habitantes", donde la comparación dio como resultado un valor p de 0,564 ($p > 0,05$), lo que sugiere que no hay diferencias estadísticamente relevantes en la cantidad de RS doméstica producida entre estos dos estratos. La prueba t de Student para muestras independientes reveló una diferencia estadísticamente significativa ($t(38) = -4,0833$; $p = 0,0001102$, $p < 0,05$) en la generación de residuos sólidos entre los grupos "758.000 – 999.000 hab." y "1.000.000 – 1.189.000 habitantes". El intervalo de confianza $[-Inf, -23748,71]$ indica que el grupo "758.000 – 999.000 habitantes" produce menos residuos sólidos domiciliarios que el grupo "1.000.000 - 1.189.000 habitantes", con un tamaño del efecto d de Cohen = $-1,33359$, lo que sugiere un gran efecto. La prueba t Student para muestras independientes mostró una diferencia estadísticamente significativa ($t(21) = -54,919$; $p = 2,2e-16$, $p < 0,05$) en la generación de residuos sólidos entre los grupos "1.000.000 - 1.189.000 hab." y "9.685.000 - 10.850.000 hab.". El intervalo de confianza $[-Inf, -2038507]$ indica que el grupo "1.000.000 - 1.189.000 hab." produce menos residuos sólidos domiciliarios que el grupo "9.685.000 - 10.850.000 hab.", con un tamaño del efecto d de Cohen = $-24,04328$, Esto indica un gran efecto. Por otra parte, Los resultados apoyan la hipótesis de que el índice GPC per cápita influye en la generación de residuos sólidos domiciliarios de los departamentos del Perú durante el periodo 2014-2021. El análisis de varianza de Welch [$F(3, 94,310) = 5,359$; $p = 0,02$, $p < 0,05$] y un Eta cuadrado de 0,0705 revelan la existencia de una influencia significativa del GPC sobre la producción de residuos sólidos domiciliarios, explicando el 7,05% de la varianza total en los niveles de producción entre departamentos, lo que sugiere un efecto cercano a medio o moderado sobre los volúmenes de generación de residuos domiciliarios. El análisis post hoc de Games-Howell identificó diferencias estadísticamente significativas en los volúmenes de residuos sólidos, concretamente al comparar los grupos "0% - 43,99%" y "45,98% - 50,53%" ($p = 0,017$), y entre los grupos "44% - 45,99%" y "46,0% - 50,99%" ($p = 0,03$). La prueba t de Student para muestras independientes reveló una diferencia estadísticamente significativa ($t(94) = -3,0795$; $p = 0,001359$, $p < 0,05$) en la generación de residuos sólidos entre los grupos "0% - 43,99%" y "46,00% - 50,99%". El intervalo de confianza $[-Inf, -116898,7]$ indica que el grupo "0% - 43,99%" produce menos residuos sólidos domiciliarios que el grupo "46,00% - 50,99%", con un tamaño del efecto d de Cohen = $-0,6285902$, lo que sugiere un efecto medio. la prueba t de

Student para muestras independientes entre los grupos "44% - 45,99%" y "46% - 50,99%" ($t(94) = -2,8744$; $p = 0,002503$, $p < 0,05$) reveló una diferencia estadísticamente significativa. El intervalo de confianza $[-\text{Inf}, -99927,99]$ indica que el grupo "44% - 45,99%" produce menos residuos sólidos domésticos que el grupo "46% - 50,99%", con un tamaño del efecto d de Cohen = $-0,586743$, lo que sugiere un efecto medio.

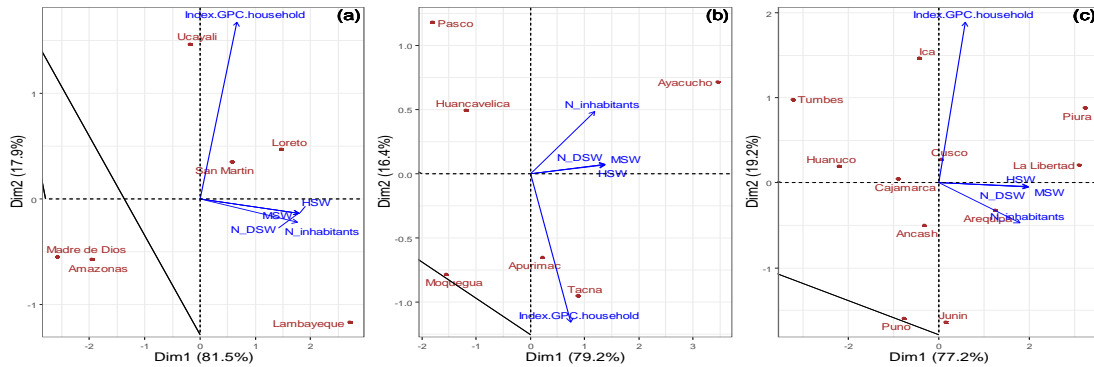


Figura 7: PCA del Cluster Ward.D2 Method de los departamentos (a) Ucayali, Loreto, San Martín, Lambayeque, Amazonas y Madre de Dios ;(b) Pasco, Huancavelica, Moquegua, Ayacucho, Apurímac y Tacna; (c) Ancash, La Libertad, Piura, Ica, Huánuco, Puno, Junín, Cajamarca, Cusco, Tumbes y Arequipa. Para una mejor visualización se llevó a cabo el PCA, nos apoyamos en el método de mínima varianza de Ward, el cual suele generar conglomerados equilibrados y compactos formamos tres grupos: En el Grupo (a) la primera componente explica el 81.5% de la inercia y las dos primeras componentes representan el 99.4% de la varianza total. Lambayeque (5,616.48 toneladas), Loreto (2,946.44 toneladas) y San Martín (1,596.07 toneladas) presentaron los promedios más altos de RSD, mientras que Amazonas (441.1 toneladas) tuvo el más bajo. Ucayali (60%), Loreto (58%) y San Martín (57%) registraron los índices GPC más elevados. Grupo (b) las cifras correspondientes son 79.2% y 95.6%. Ayacucho (701.81 toneladas) mostró el promedio más alto de RSD, y Apurímac (497 toneladas) el más bajo. Tacna y Apurímac (44%) y Moquegua (43%) exhibieron los índices GPC más altos, mientras que Huancavelica (42%) y Pasco (41%) los más bajos. Grupo (c) la primera componente recoge el 77.2% de la inercia y las dos primeras el 96.4% de la varianza. Piura (4,476.53 toneladas) y La Libertad (3,478.46 toneladas) presentaron los promedios más altos de RSD, y Huánuco (859.41 toneladas) y Cajamarca (812.74 toneladas) los más bajos. Ica y Piura lideraron con un GPC promedio del 48%, mientras que Puno y Junín (43%) registraron los valores más bajos. En todos los casos, con solo dos dimensiones se preserva la mayor parte de la información esencial y se observa una correlación positiva entre los tipos de RSD y la cantidad de habitantes.

7. Discusión

El análisis de los datos recolectados y el uso de técnicas estadísticas permitieron corroborar la que la cantidad de población influye significativamente en la generación de RS domiciliarios en los departamentos del Perú. El Anova de Welch y el Eta cuadrado proporcionaron evidencia sólida de un efecto muy grande de la cantidad de población en la generación de RS domiciliarios entre departamentos. Estos hallazgos son consistentes con la literatura existente que sugiere una relación directa entre el tamaño de la población y la generación de residuos sólidos. A medida que la población aumenta, también lo hace la demanda de bienes y servicios, lo que a su vez lleva a un aumento en la producción de residuos (Hoornweg y Bhada-Tata, 2012). El

análisis post hoc mediante la prueba Games-Howell reveló diferencias estadísticamente significativas en los volúmenes de RS domésticos producidos entre la mayoría de los pares de grupos de población comparados. Sin embargo, hubo una excepción notable entre los grupos "1.000.000 - 1.189.000 habitantes" y "1.222.000 - 1.538.000 habitantes", donde no se encontraron diferencias estadísticamente significativas en la cantidad de RS domiciliarios producidos. Esto puede deberse a factores no considerados en este estudio, como la variabilidad en las prácticas de gestión de residuos o las diferencias en los niveles de consumo. En armonía Quispe et al. (2023) en el distrito de Santa Rosa, Ayacucho realizó un análisis demostró diferencias significativas entre los RS. Obtuvo una alta correlación entre las zonas A y B indica que los desechos orgánicos superan a los inorgánicos.

Por otra parte, el análisis de la varianza de Welch, el análisis post hoc de Games-Howell y las pruebas t de Student han proporcionado evidencia estadística sólida que el índice GPC per cápita tiene una influencia significativa en la generación de RS domiciliarios. El hecho de que el GPC explique el 7,05% de la varianza total en los niveles de producción de residuos sólidos entre departamentos sugiere que, aunque existen otros factores que también influyen en la generación de residuos, el GPC per cápita tiene un efecto moderado sobre los volúmenes de generación de residuos domiciliarios. Esta afirmación coincide lo Storto (2024) descubrió que el indicador era relativo al impacto sobre la ecoeficiencia de las variables que miden el tamaño o la densidad de los servicios de RSU, los indicadores siguen siendo no concluyentes.

Se observaron diferencias estadísticamente significativas en los volúmenes de residuos sólidos entre los diferentes grupos de GPC. Los grupos con un GPC más alto ("45,98% - 50,53%" y "46,0% - 50,99%") generaron más residuos sólidos que los grupos con un GPC más bajo ("0% - 43,99%" y "44% - 45,99%"). En concordancia C. H. Li et al. (2023) Se prevé que la cantidad de desechos aumentará en los próximos años. Los resultados del estudio corroboran esta teoría, ya que muestran una correlación positiva entre los diferentes tipos de residuos sólidos RS domiciliarios, RS no domiciliarios y RS municipales y la cantidad de habitantes en los departamentos de Perú.

El análisis PCA y el método de mínima varianza de Ward revelaron que Lima, siendo el departamento con la mayor cantidad de habitantes, es también el que presenta los valores más altos en la producción de RS domiciliarios, constituyendo un grupo por sí misma. Esta observación sugiere que la generación de residuos sólidos está fuertemente influenciada por la densidad poblacional. en acuerdo Cerna-Cueva et al. (2022) debido al aumento de la población y los hábitos de consumo la gestión de RS son desafíos ambientales en el mundo. En el Grupo a) la primera componente explicó el 81.5% de la inercia y las dos primeras el 99.4% de la varianza. Lambayeque 5,616.48 toneladas, Loreto 2,946.44 y San Martín 1,596.07 tuvieron los promedios de RSD más altos, mientras que Amazonas 441.1 el más bajo. Ucayali 60%, Loreto 58% y San Martín 57% registraron los índices GPC más elevados. En el Grupo b) Ayacucho 701.81 toneladas mostró el promedio RSD más alto y Apurímac 497 el más bajo. Tacna, Apurímac 44% y Moquegua 43% exhibieron los índices GPC más altos, mientras que Huancavelica 42% y Pasco 41% los más bajos. En el Grupo c) Piura 4,476.53 toneladas y La Libertad 3,478.46 con promedios RSD más altos, Huánuco 859.41 y Cajamarca 812.74 los más bajos. Ica y Piura lideraron con GPC de 48%, mientras que Puno y Junín 43% los más bajos. Se observó correlación positiva entre RSD y cantidad de habitantes. Estos hallazgos son consistentes con los reportados por Fernández-Domínguez et al. (2021), quienes utilizaron el ACP y el análisis de agrupamiento jerárquico para caracterizar la calidad de los residuos orgánicos y encontraron una variabilidad del 58.02% en la

accesibilidad, así como cambios significativos en la complejidad y los compuestos no biodegradables después del tratamiento. Asimismo, Kim y de Vasconcelos Barros (2023) aplicaron el ACP para evaluar los efectos de la Política Nacional de Residuos Sólidos en municipios de Minas Gerais, Brasil, y lograron representar más del 70% de la variabilidad en los datos. Además, el estudio de Zhang et al. (2022) demuestra la utilidad del ACP en combinación con técnicas de aprendizaje automático para la clasificación de lixiviados de incineración de residuos sólidos municipales, logrando tasas de precisión superiores al 95% en los conjuntos de entrenamiento y prueba.

8. Conclusiones

A través de la aplicación de técnicas estadísticas y la adaptación de la metodología de minería de datos SEMMA y el Análisis de Componentes Principales (PCA), se ha logrado obtener una visión más detallada y precisa sobre la generación de residuos sólidos domiciliarios en Perú. Este estudio ha revelado las influencias significativas tanto de la cantidad de población y índice GPC per cápita con la generación de RS domiciliarios. La existencia de una alta correlación positiva entre los RS domiciliarios, RS no domiciliarios y RS municipales.

Las pruebas t dan evidencia del efecto del tamaño de la población en la generación de residuos sólidos. Los grupos con poblaciones más grandes produjeron significativamente más residuos que los grupos con poblaciones más pequeñas, lo que se refleja en los grandes tamaños de efecto encontrados. El GPC per cápita tiene un efecto moderado, lo que da pie a otros estudios para explorar más a fondo las causas subyacentes de estas diferencias regionales y evaluar la eficacia de las medidas de gestión de residuos implementadas hasta la fecha.

El estudio indica una correlación positiva entre la generación de residuos sólidos (RS) residenciales y el número de habitantes en los departamentos de Perú, lo cual subraya la repercusión del volumen demográfico en la elaboración de RS domiciliarios. En particular Lima, con un promedio de 13,220.47 toneladas, 59,813 residentes y un coeficiente de GPC del 50%, se distingue por su alta producción de RS residenciales, representando un caso singular.

Se realizó PCA para tres grupos como resultado de agruparlos por método de mínima varianza de Ward se detalla los resultados: En el grupo (a) Lambayeque (5,616.48 toneladas), Loreto (2,946.44 toneladas) y Madre de Dios (2,091.20 toneladas), en estos se muestra la mayor media en la producción de Residuos Sólidos domiciliarios. Por otro lado, Amazonas (441.1 toneladas) y San Martín (1,596.07 toneladas) exhiben los promedios más bajos. En lo que respecta al índice de producción per cápita (GPC) en promedio, Ucayali 60%, Loreto 58% y San Martín 57% tienen los valores más elevados, en comparación con Madre de Dios 51%, Amazonas 52% y Lambayeque 53% presentan los valores más bajos. En el grupo (b) Ayacucho (701.81 toneladas) presenta el promedio más alto de Residuos Sólidos (RS) domiciliarios, mientras que Apurímac (497 toneladas) registra el menor número. En relación con el indicador GPC los departamentos de Tacna con un 44%, Apurímac con un 44% y Moquegua con un 43% exhiben los porcentajes promedios más elevados, en contraste con Huancavelica con un 42% y Pasco con un 41%, que despliegan los más reducidos. En el grupo (c), los departamentos con los valores promedio más altos de residuos sólidos domiciliarios son Piura (4,476.53 toneladas) y La Libertad (3,478.46 toneladas), mientras que los de menor volumen son Huánuco (859.41 toneladas) y Cajamarca (812.74 toneladas). En lo que respecta al GPC, Ica y Piura lideran con un promedio del 48%, en contraposición a Puno y Junín, que ostentan los valores más bajos con un 43%.

Referencias Bibliográficas

- Abushammala, H., & Ghulam, S. T. (2023). Impact of Residents' Demographics on Their Knowledge, Attitudes, and Practices towards Waste Management at the Household Level in the United Arab Emirates. *Sustainability (Switzerland)*, 15(1). Scopus. <https://doi.org/10.3390/su15010685>
- Al Mamun, A., Hayat, N., Masud, M. M., Makhbul, Z. K. M., Jannat, T., & Salleh, M. F. M. (2022). Modelling the Significance of Value-Belief-Norm Theory in Predicting Solid Waste Management Intention and Behavior. *FRONTIERS IN ENVIRONMENTAL SCIENCE*, 10, 906002. <https://doi.org/10.3389/fenvs.2022.906002>
- Begum, A., Liu, J., Qayum, H., & Mamdouh, A. (2022). Environmental and Moral Education for Effective Environmentalism: An Ideological and Philosophical Approach. *International Journal of Environmental Research and Public Health*, 19(23). Scopus. <https://doi.org/10.3390/ijerph192315549>
- Bonnett, M. (2023). Environmental consciousness, nature and the philosophy of education: Matters arising. *Environmental Education Research*. Scopus. <https://doi.org/10.1080/13504622.2023.2225807>
- Cerna-Cueva, A. F., Rosas-Echevarría, C., Perales-Flores, R., Ataucusi-Flores, P. L., Cerna-Cueva, A. F., Rosas-Echevarría, C., Perales-Flores, R., & Ataucusi-Flores, P. L. (2022). Predicción de la generación de residuos sólidos domiciliarios con machine learning en una zona rural de Puno. *Tecnia*, 32(1), 44-52. <https://doi.org/10.21754/tecnica.v32i1.1378>
- Cheng, K. M., Tan, J. Y., Wong, S. Y., Koo, A. C., & Sharji, E. A. (2022). A Review of Future Household Waste Management for Sustainable Environment in Malaysian Cities. *SUSTAINABILITY*, 14(11), 6517. <https://doi.org/10.3390/su14116517>
- Chowdhury, T., Chowdhury, H., Rahman, M. S., Hossain, N., Ahmed, A., & Sait, S. M. (2022). Estimation of the healthcare waste generation during COVID-19 pandemic in Bangladesh. *SCIENCE OF THE TOTAL ENVIRONMENT*, 811, 152295. <https://doi.org/10.1016/j.scitotenv.2021.152295>
- De-la-Torre, G. E., Rakib, M. R. J., Pizarro-Ortega, C. I., & Dioses-Salinas, D. C. (2021). Occurrence of personal protective equipment (PPE) associated with the COVID-19 pandemic along the coast of Lima, Peru. *SCIENCE OF THE TOTAL ENVIRONMENT*, 774, 145774. <https://doi.org/10.1016/j.scitotenv.2021.145774>
- El Peruano. (2024). Situación actual de los residuos sólidos. <https://elperuano.pe/noticia/216136-situacion-actual-de-los-residuos-solidos>
- Fontaine, L., Legros, R., & Frayret, J.-M. (2024). Solid waste generation prediction model framework using socioeconomic and demographic factors with real-time MSW collection data. *Waste Management and Research*. Scopus. <https://doi.org/10.1177/0734242X241231414>
- GESTIÓN, N. (2023, abril 6). Residuos en Cañete | Minam: San Vicente de Cañete en emergencia por gestión y manejo de residuos sólidos | Gobierno | San Vicente de Cañete | Minam | PERU. Gestión; NOTICIAS GESTIÓN. <https://gestion.pe/peru/residuos-en-canete->

minam-san-vicente-de-canete-en-emergencia-por-gestion-y-manejo-de-residuos-solidos-gobierno-san-vicente-de-canete-minam-noticia/

He, R., Sandoval-Reyes, M., Scott, I., Semeano, R., Ferrao, P., Matthews, S., & Small, M. J. (2022). Global knowledge base for municipal solid waste management: Framework development and application in waste generation prediction. *JOURNAL OF CLEANER PRODUCTION*, 377, 134501. <https://doi.org/10.1016/j.jclepro.2022.134501>

Karadimas, N. V., Papatzelou, K., & Loumos, V. G. (2007). Genetic algorithms for municipal solid waste collection and routing optimization. *IFIP International Federation for Information Processing*, 247, 223-231. Scopus. https://doi.org/10.1007/978-0-387-74161-1_24

Li, C. H., Lee, T. T., & Lau, S. S. Y. (2023). Enhancement of Municipal Solid Waste Management in Hong Kong through Innovative Solutions: A Review. *Sustainability (Switzerland)*, 15(4). Scopus. <https://doi.org/10.3390/su15043310>

Li, W., & Wang, J. (2022). Feasibility study of the unit pricing system for household wastes sorting in the context of China. *WASTE MANAGEMENT & RESEARCH*, 40(5), 586-595. <https://doi.org/10.1177/0734242X20988783>

Liu, C.-H., & Hung, C. (2023). Reutilization of solid wastes to improve the hydromechanical and mechanical behaviors of soils—A state-of-the-art review. *Sustainable Environment Research*, 33(1). Scopus. <https://doi.org/10.1186/s42834-023-00179-6>

lo Storto, C. (2024). Measuring ecoefficiency of municipal solid waste management in Apulia to account for governance heterogeneities. *Cleaner Waste Systems*, 7. Scopus. <https://doi.org/10.1016/j.clwas.2024.100131>

Maalouf, A., & Mavropoulos, A. (2023). Re-assessing global municipal solid waste generation. *WASTE MANAGEMENT & RESEARCH*, 41(4), 936-947. <https://doi.org/10.1177/0734242X221074116>

Martínez. (2018). RPubS—Análisis de componentes principales (PCA). https://rpubs.com/Cristina_Gil/PCA

MINAM. (2023). Ministerio del Ambiente. <https://www.datosabiertos.gob.pe/group/ministerio-del-ambiente>

Mtika, W. M., & Tilley, E. (2020). Environmental Sanitation Planning: Feasibility of the CLUES Framework in a Malawian Small Town. *FRONTIERS IN ENVIRONMENTAL SCIENCE*, 7, 204. <https://doi.org/10.3389/fenvs.2019.00204>

Mumtaz, H., Werle, S., & Sobek, S. (2023). A waste wet oxidation technique as a solution for chemical production and resource recovery in Poland. *Clean Technologies and Environmental Policy*. Scopus. <https://doi.org/10.1007/s10098-023-02520-4>

Nayak, S. P., & Uppal, S. B. (2023). Preliminary assessment of COVID-19 Waste management scenario during lockdown in Chandigarh & nearby areas. 1110(1). Scopus. <https://doi.org/10.1088/1755-1315/1110/1/012067>

Patrao, G., & Karnik, A. (2023). Identifying Drivers and Hindrances to the Disposal of Used Mobile Phones: A Study of User Behavior in the UAE. *SAGE Open*, 13(3). Scopus. <https://doi.org/10.1177/21582440231196757>

Pillay, R., Hansraj, R., & Rampersad, N. (2023). Disposal of spectacles and contact lenses: Optometrist and lens wearer perspectives. *African Vision and Eye Health*, 82(1). Scopus. <https://doi.org/10.4102/aveh.v82i1.784>

Popli, K., Park, C., Han, S.-M., & Kim, S. (2021). Prediction of solid waste generation rates in urban region of Laos using socio-demographic and economic parameters with a multi linear regression approach. *Sustainability (Switzerland)*, 13(6). Scopus. <https://doi.org/10.3390/su13063038>

Quispe, K., Martínez, M., da Costa, K., Romero Giron, H., Via y Rada Vittes, J. F., Mantari Mincami, L. D., Hadi Mohamed, M. M., Huamán De La Cruz, A. R., & López-Gonzales, J. L. (2023). Solid Waste Management in Peru's Cities: A Clustering Approach for an Andean District. *Applied Sciences*, 13(3), Article 3. <https://doi.org/10.3390/app13031646>

Requena-Sanchez, N., Carbonel, D., Moonsammy, S., Demel, L., Vallester, E., Velasquez, D., Toledo Cervantes, J. A., Diaz Nunez, V. L., Vasquez Garcia, R., Santa Cruz, M., Visbal, E., & Ng, K. T. W. (2023a). COVID-19 impacts on household solid waste generation in six Latin American countries: A participatory approach. *ENVIRONMENTAL MONITORING AND ASSESSMENT*, 195(1), 155. <https://doi.org/10.1007/s10661-022-10771-9>

Requena-Sanchez, N., Carbonel, D., Moonsammy, S., Demel, L., Vallester, E., Velasquez, D., Toledo Cervantes, J. A., Diaz Nunez, V. L., Vasquez Garcia, R., Santa Cruz, M., Visbal, E., & Ng, K. T. W. (2023b). COVID-19 impacts on household solid waste generation in six Latin American countries: A participatory approach. *ENVIRONMENTAL MONITORING AND ASSESSMENT*, 195(1), 155. <https://doi.org/10.1007/s10661-022-10771-9>

Rifaldi, M., Zid, M., & Sumargo, B. (2021). CAUSAL LOOP DIAGRAM AS APPROACHING MODEL ANALYSIS IN INCREASING THE WASTE VOLUME AT THE COVID19 PANDEMIC PERIOD. *Jurnal Kesehatan Lingkungan*, 13(3), 180-185. Scopus. <https://doi.org/10.20473/jkl.v13i3.2021.180-185>

SIAR. (2024). MINAM contribuye al conocimiento y puesta en marcha de la Agenda 2030 a través de nueva publicación sobre los Objetivos de Desarrollo Sostenible | SIAR Cajamarca | Sistema de Información Ambiental Regional de Cajamarca. <https://siar.regioncajamarca.gob.pe/novedades/minam-contribuye-conocimiento-puesta-marcha-agenda-2030-traves-nueva>

Shin, S. S., Park, S. D., & Kim, G. (2024). Applicability Comparison of GIS-Based RUSLE and SEMMA for Risk Assessment of Soil Erosion in Wildfire Watersheds. *Remote Sensing*, 16(5). Scopus. <https://doi.org/10.3390/rs16050932>

Taherdoost, H. (2022). What are Different Research Approaches? Comprehensive Review of Qualitative, Quantitative, and Mixed Method Research, Their Applications, Types, and Limitations (SSRN Scholarly Paper 4178694). <https://papers.ssrn.com/abstract=4178694>

Tapia, J. A. C. (2024). Modelo de Aprendizaje Profundo para identificar plagas en la producción de quinua. *Revista Ingeniería*, 8(20), Article 20. <https://doi.org/10.33996/revistaingenieria.v8i20.116>

Xu, J., Zhang, Z., Xu, Y., Liu, L., & Pei, T. (2023). Quantitative evaluation of waste sorting management policies in China's major cities based on the PMC index model. *Frontiers in Environmental Science*, 11. <https://doi.org/10.3389/fenvs.2023.1065900>

Zaiontz, C. (2024). Prueba de Welch | Estadísticas reales usando Excel. <https://real-statistics.com/one-way-analysis-of-variance-anova/welchs-procedure/>

Zhong, Y., Wang, S., Wu, H., Cheng, Z., & Li, X. (2022). SEMMA-Based visual exploration of cyber security event. *Journal of Zhejiang University, Science Edition*, 49(2), 131-140. Scopus. <https://doi.org/10.3785/j.issn.1008-9497.2022.02.001>

Zulkipli, F., Nopiah, Z. M., Jamian, N. H., Basri, N. E. A., & Kie, C. J. (2022). Sustainable Public Awareness on Solid Waste Management and Environmental Care using Logistics Regression. *JURNAL KEJURUTERAAN*, 5(2), 139-148. [https://doi.org/10.17576/jkukm-2022-si5\(2\)-15](https://doi.org/10.17576/jkukm-2022-si5(2)-15)

1. Anexo A

Instrumentos de Recolección de Datos

The screenshot displays the 'Datos Abiertos' (Open Data) portal for the Peruvian government. The main content area features a dataset titled 'Residuos municipales generados anualmente' (Annual municipal waste generated). The table lists various municipalities and their corresponding waste generation statistics. The data is organized into columns for different categories of waste and their respective quantities.

Municipio	Residuos sólidos	Residuos orgánicos	Residuos inorgánicos	Residuos peligrosos	Residuos electrónicos	Residuos textiles	Residuos de construcción	Residuos de demolición	Residuos de otros tipos
2000001	1000000	500000	300000	100000	50000	20000	10000	5000	2000
2000002	1200000	600000	350000	120000	60000	25000	12000	6000	2500
2000003	1500000	750000	450000	150000	75000	30000	15000	7500	3000
2000004	1800000	900000	550000	180000	90000	35000	18000	9000	3500
2000005	2000000	1000000	600000	200000	100000	40000	20000	10000	4000

The page also includes a sidebar with navigation options and a main text area providing context and details about the data. The text describes the scope of the data, its source, and how it is used for public access and analysis.

<https://www.datosabiertos.gob.pe/dataset/residuos-municipales-generados-anualmente>



Esta página también está disponible en inglés

Instituto Nacional de Estadística e Informática

Encuesta Nacional de Hogares - ENAHO 2023

Informe

31 de enero de 2023

Descripción

La Encuesta Nacional de Hogares se inicia en el año 1995, es a partir del año 2003 que se realiza de forma continua, en el área urbana y rural de todo el país y se obtiene información anual con desagregación departamental. Mensualmente se visitan aproximadamente 3,070 viviendas seleccionadas por un método aleatorio, tiene un componente panel de 30% de la muestra.

Es una actividad estadística continua, que permite obtener información actualizada sobre las condiciones de vida, gastos e ingresos de los hogares, y fundamentalmente conocer la evolución de los niveles de pobreza monetaria en el país.

La carga de trabajo asignada al/la encuestador/a en el área urbana son 6 viviendas y en el área rural 8 viviendas; en cada vivienda se recopila información de todos los miembros del hogar, mediante entrevistas presenciales que se complementan con entrevistas telefónicas.

Objetivos

1. Generar indicadores que permitan conocer la evolución de la pobreza monetaria, del bienestar y de las condiciones de vida de los hogares.
2. Efectuar diagnósticos sobre las condiciones de vida y pobreza monetaria de la población.
3. Medir el alcance de los programas sociales alimentarios y no alimentarios en la mejora de las condiciones de vida de la población.

Nº	Variable	Unidad	Valor	Fecha	Categoría	Detalle
1	AGE	Categoría	0	0	Alfabeto Encuesta	Región: Región 0
2	ME5	Categoría	0	0	Nivel de Educación de la Encuesta	Región: Región 0
3	CONG03M	Categoría	0	0	Número del Congresista	Región: Región 00
4	VUE03A	Categoría	0	0	Número de Vecinos de la Vivienda	Región: Región 00
5	MOG04	Categoría	0	0	Número de miembros del hogar	Región: Región 0
6	MOG05	Categoría	0	0	Ubicación geográfica	Región: Región 0
7	CONG05	Región	0	0	Categoría Geográfica	Región: Región 0
8	ED05A070	Región	0	0	Edad	Región: Región 0
9	MOG0502	Región	0	0	Presencia de personas de la encuesta fuera del hogar	Región: Región 0
10	ED05C	Región	0	0	Tipo de Educación del Congresista	Región: Región 0
11	MOG0504	Región	0	0	Punto de contacto del hogar de la encuesta (Muestreo)	Región: Región 00
12	MOG0507	Región	0	0	Residencia física de la Encuesta	Región: Región 0
13	MOG0508	Región	0	0	¿El hogar fue entrevistado en otro país?	Región: Región 0
14	MOG0509	Región	0	0	Además de esta vivienda ¿tienen otra vivienda en el que viven?	Región: Región 0
15	MOG0510	Región	0	0	¿En qué idioma, género y desplazamiento está el jefe de hogar?	Región: Región 0
16	MOG0511	Región	0	0	¿Las personas entrevistadas de la encuesta fueron invitadas? ¿Por quién?	Región: Región 0
17	MOG0512	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
18	MOG0513	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
19	MOG0514	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
20	MOG0515	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
21	MOG0516	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
22	MOG0517	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
23	MOG0518	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
24	MOG0519	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
25	MOG0520	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
26	MOG0521	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
27	MOG0522	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
28	MOG0523	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
29	MOG0524	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
30	MOG0525	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
31	MOG0526	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
32	MOG0527	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
33	MOG0528	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
34	MOG0529	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
35	MOG0530	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
36	MOG0531	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
37	MOG0532	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
38	MOG0533	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
39	MOG0534	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
40	MOG0535	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
41	MOG0536	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
42	MOG0537	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
43	MOG0538	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
44	MOG0539	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
45	MOG0540	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
46	MOG0541	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
47	MOG0542	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
48	MOG0543	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
49	MOG0544	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
50	MOG0545	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
51	MOG0546	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
52	MOG0547	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
53	MOG0548	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
54	MOG0549	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0
55	MOG0550	Región	0	0	¿La cabeza de familia, por cuestiones de salud no está en la vivienda?	Región: Región 0

<https://www.gob.pe/institucion/inei/informes-publicaciones/3364523-encuesta-nacional-de-hogares-ena-2023>

2. Anexo B

Conciencia Ambiental por departamento	Pruebas de normalidad - Shapiro-Wilk			
	Cantidad de residuos de origen domiciliario		Total de residuos solidos	
	w	p-value	w	p-value
Amazonas	0.89	0.26	0.89	0.26
Ancash	0.92	0.41	0.92	0.41
Apurímac	0.95	0.67	0.95	0.67
Arequipa	0.89	0.22	0.89	0.22
Ayacucho	0.86	0.12	0.83	0.06
Cajamarca	0.84	0.08	0.84	0.08
Cusco	0.98	0.98	0.98	0.98
Huancavelica	0.84	0.08	0.84	0.08
Huánuco	0.85	0.09	0.85	0.09
Ica	0.86	0.13	0.84	0.08
Junín	0.92	0.42	0.92	0.42
La Libertad	0.89	0.25	0.89	0.25
Lambayeque	0.93	0.51	0.93	0.51
Lima	0.87	0.14	0.87	0.14
Loreto	0.87	0.14	0.87	0.14
Madre de Dios	0.89	0.22	0.89	0.22
Moquegua	0.91	0.33	0.91	0.33
Pasco	0.90	0.26	0.90	0.26
Piura	0.92	0.45	0.92	0.45
Puno	0.95	0.74	0.95	0.74
San Martín	0.89	0.25	0.89	0.25
Tacna	0.93	0.49	0.93	0.49
Tumbes	0.88	0.19	0.88	0.19
Ucayali	0.92	0.39	0.92	0.39

Grupo	Año	Residuos domiciliarios	Población	Residuos solidos	is.outlier	is.extreme
Cajamarca	2021	110416.31	1455245	315475.04	VERDADERO	FALSO
Tacna	2021	57277.03	377,842	163648.67	VERDADERO	FALSO

Pruebas robustas de igualdad de medias

Total de residuos sólidos (Ton/Año)

	Estadístico ^a	gl1	gl2	Sig.
Welch	557,878	23	60,544	,000

a. F distribuida de forma asintótica