



**INSTITUTO DE GOBIERNO Y DE GESTIÓN PÚBLICA
UNIDAD DE POSGRADO**

**EQUIDAD DE INGRESOS POR PENSIONES EN EL
SISTEMA PREVISIONAL PERUANO: PERIODO 2012-**

2022



**PRESENTADO POR
MARCO AURELIO ESPINOZA RIVERA**

**ASESOR
Dr. ISAAC ROBERTO ANGELES LAZO**

**TESIS
PARA OPTAR EL GRADO ACADÉMICO
DE DOCTOR EN GOBIERNO Y POLÍTICA PÚBLICA**

LIMA – PERÚ

2023



CC BY-NC-ND

Reconocimiento – No comercial – Sin obra derivada

El autor sólo permite que se pueda descargar esta obra y compartirla con otras personas, siempre que se reconozca su autoría, pero no se puede cambiar de ninguna manera ni se puede utilizar comercialmente.

<http://creativecommons.org/licenses/by-nc-nd/4.0/>



**INSTITUTO DE GOBIERNO Y DE GESTIÓN PÚBLICA
SECCIÓN DE POSGRADO**

**EQUIDAD DE INGRESOS POR PENSIONES EN EL
SISTEMA PREVISIONAL PERUANO: PERIODO 2012-2022**

**TESIS PARA OPTAR
EL GRADO ACADÉMICO DE
DOCTOR EN GOBIERNO Y POLÍTICA PÚBLICA**

**PRESENTADO POR:
MARCO AURELIO ESPINOZA RIVERA**

**ASESOR:
Dr. ISAAC ROBERTO ANGELES LAZO**

POLÍTICA PÚBLICA

**LIMA – PERÚ
2023**

Dedicatoria

A mis padres, que han sabido formarme con buenos sentimientos, hábitos y valores lo cual me ha ayudado a seguir adelante en los momentos difíciles.

A mi esposa e hijos, quienes son mi mayor motivación.

Agradecimientos

A mi dilecto asesor de tesis por su comprensión y estímulo constante, brindándome la guía necesaria para la culminación de este trabajo que concreta mi contribución a la discusión de mejorar la política y gestión pública del país en materia previsional.

Índice de contenido

Dedicatoria	ii
Agradecimientos	iii
Índice de tablas	vi
Índice de figuras	vii
Resumen	viii
Abstract	ix
INTRODUCCIÓN	1
Descripción de la situación problemática	1
Formulación del problema	3
Objetivo General	4
Objetivos Específicos	4
Hipótesis principal	5
Hipótesis Específicas	5
Importancia del estudio	6
Enfoque y tipo de diseño metodológico	7
Estructura de la tesis	8
CAPÍTULO I - MARCO TEÓRICO	10
1.1. Antecedentes de la Investigación	10
1.2. Bases Teóricas	13
1.3. Definición de términos básicos	23
CAPÍTULO II - HIPÓTESIS Y VARIABLES	28
2.1. Formulación de hipótesis	28
2.2. Variables y definición operacional	29
CAPITULO III – METODOLOGÍA DE LA INVESTIGACIÓN	31

3.1. Diseño metodológico.....	31
3.2. Diseño muestral	31
3.3. Técnicas de recolección de datos	32
3.4. Técnicas estadísticas para el procesamiento de información...	35
3.5. Aspectos éticos.....	37
CAPITULO IV – RESULTADOS.....	38
4.1. Ingreso por pensión, como factor de equidad	38
4.2. Coeficiente GINI de equidad de Ingresos por pensión.....	44
4.3. Exploración de factores de equidad mediante modelos de regresión de árboles de decisión.....	48
4.4. Determinación de factores de mayor impacto sobre los ingresos mediante modelo XGBoost.....	51
4.5. Factores de mayor impacto sobre los ingresos	53
CAPITULO V – DISCUSIÓN	59
5.1. Con respecto al objetivo general de investigación	59
5.2. Con respecto a los objetivos específicos de investigación.....	62
5.3. Con respecto a la hipótesis principal de la investigación	66
CAPITULO VI – CONCLUSIONES.....	69
CAPITULO VII – RECOMENDACIONES	71
FUENTES DE INFORMACIÓN.....	73
ANEXOS	1

Índice de tablas

Tabla 1. Ingreso mensual promedio, por fondo previsional, Perú.....	38
Tabla 2. Ingreso mensual promedio, por fondo previsional, Costa	40
Tabla 3. Ingreso mensual promedio, por fondo previsional, Sierra	41
Tabla 4. Ingreso mensual promedio, por fondo previsional, Selva.....	42
Tabla 5. Ingreso mensual promedio, por fondo previsional, Lima	43
Tabla 6. Coeficiente de equidad GINI según segmentos	44
Tabla 7. Ingresos y coeficiente de equidad GINI 2022	46
Tabla 8. Bondad de ajuste de modelos de regresión	50

Índice de figuras

Figura 1. Ingreso mensual promedio, por fondo previsional, Perú	39
Figura 2. Ingreso mensual promedio, por fondo previsional, Costa	40
Figura 3. Ingreso mensual promedio, por fondo previsional, Sierra.....	41
Figura 4. Ingreso mensual promedio, por fondo previsional, Selva	42
Figura 5. Ingreso mensual promedio, por fondo previsional, Lima	43
Figura 6. Coeficiente de equidad GINI según segmentos	45
Figura 7. Ingresos y coeficiente de equidad GINI 2022	46
Figura 8. Factores de equidad de ingresos por pensión.....	52
Figura 9. Valor SHAP para el factor: Nivel de estudio	54
Figura 10. Valor SHAP para el factor: Vive en Lima metropolitana.....	55
Figura 11. Valor SHAP para el factor: Edad.....	56
Figura 12. Valor SHAP para el factor: Sexo	57
Figura 13. Valor SHAP para el factor: Tipo de entidad.....	58

Resumen

Este estudio profundiza en el sistema previsional peruano entre 2012 y 2022, con un énfasis particular en la equidad de los ingresos por pensiones de jubilación y los factores determinantes de estas desigualdades. A través de un enfoque cuantitativo y la implementación de técnicas avanzadas de machine learning, incluyendo modelos de regresión XGBoost y análisis SHAP, se analizó una muestra representativa de 858,474 pensionistas. Los resultados revelan una discrepancia significativa entre los pensionistas del Sistema Nacional de Pensiones y aquellos del Sistema Privado de Pensiones, siendo las pensiones de este último en promedio más elevadas.

El estudio identifica como factores cruciales en la equidad de ingresos por pensiones el nivel educativo, la ubicación geográfica (Lima versus otras regiones), la edad, el género y el sector de empleo (público o privado). Sorprendentemente, el nivel educativo de los pensionistas influye en casi un 20% del monto de la pensión por jubilación, lo que subraya su papel significativo en la determinación de los ingresos de jubilación. Este hallazgo es de particular importancia para la formulación de políticas públicas, sugiriendo que mejorar el acceso a la educación podría ser una estrategia clave para aumentar la equidad en el sistema previsional peruano.

Palabras clave: Sistema previsional, Inequidad de ingresos por pensión, Igualdad y justicia en pensiones.

Abstract

This study delves into the Peruvian pension system from 2012 to 2022, focusing particularly on the equity of retirement pension incomes and the determining factors of these disparities. Employing a quantitative approach and advanced machine learning techniques, including XGBoost regression models and SHAP analysis, a representative sample of 858,474 pensioners was analyzed. The results reveal a significant discrepancy between pensioners in the National Pension System and those in the Private Pension System, with pensions in the latter being on average higher.

The study identifies key factors affecting pension income equity as educational level, geographical location (Lima versus other regions), age, gender, and employment sector (public or private). Remarkably, the educational level of pensioners influences nearly 20% of the retirement pension amount, highlighting its significant role in determining retirement incomes. This finding is particularly important for public policy formulation, suggesting that improving access to education could be a key strategy for increasing equity in the Peruvian pension system.

Keywords: Pension system, Pension income inequality, Equality and justice in pensions.

EQUIDAD DE INGRESOS POR PENSIONES EN EL
SISTEMA PREVISIONAL PERUANO: PERIODO 2012-2022
por MARCO AURELIO ESPINOZA RIVERA

Fecha de entrega: 21-mar-2024 05:48p.m. (UTC-0500)

Identificador de la entrega: 2327241315

Nombre del archivo: Equidad de Ingresos por Pensiones en el Sistema Previsional Peruano-Periodo 2012-2022.pdf (556.84K)

Total de palabras: 16427

Total de caracteres: 85687

Revisión Antiplagio

INFORME DE ORIGINALIDAD

13% INDICE DE SIMILITUD **12%** FUENTES DE INTERNET **4%** PUBLICACIONES **%** TRABAJOS DEL ESTUDIANTE

FUENTES PRIMARIAS

1	repositorio.usmp.edu.pe Fuente de Internet	1 %
2	www.slideshare.net Fuente de Internet	1 %
3	www.coursehero.com Fuente de Internet	1 %
4	issuu.com Fuente de Internet	<1 %
5	repositorio.unsaac.edu.pe Fuente de Internet	<1 %
6	dokumen.pub Fuente de Internet	<1 %
7	www.repositorioacademico.usmp.edu.pe Fuente de Internet	<1 %
8	estadisticamente.com Fuente de Internet	<1 %
9	hdl.handle.net Fuente de Internet	<1 %

INTRODUCCIÓN

Descripción de la situación problemática.

Se puede decir que el número de personas mayores está creciendo rápidamente en este siglo. A fines del año 2022, una de cada nueve personas en el mundo tenía más de 60 años y este número seguirá aumentando si la gente sigue teniendo menos hijos y vive más años; esto significa que en un futuro cercano habrá cada vez más personas mayores en la población global (Fondo de Población de las Naciones Unidas, 2022). Este aumento en la población ha ocasionado implicaciones significativas en sistemas previsionales, al no haber considerado esta nueva realidad demográfica.

Vivir más tiempo no es malo; de hecho, es algo que la humanidad ha querido lograr desde hace tiempo, la ciencia está haciendo un gran trabajo para alcanzar esa meta. Sin embargo, este cambio trajo desafíos para la economía y la sociedad, especialmente en países como el nuestro donde hay mucha desigualdad. Este problema podría empeorar para los grupos más frágiles, como las personas mayores. Hay planes y acuerdos internacionales, como el Plan de acción de Madrid y la Convención Interamericana de 2015, que buscan proteger los derechos y mejorar la calidad de vida de las personas mayores. Estos acuerdos nos guían sobre cómo crear un entorno más amigable para el envejecimiento (Organismo de Estados Americanos, 2022). La prolongación de la esperanza de vida, aunque sea un logro humano, plantea retos en la gestión pública y en sistemas como el previsional; por lo que requieren una acción coordinada y fundamentada en el respeto de los derechos humanos.

En esta línea de crear entornos favorables para el envejecimiento, el investigador consideró pertinente posicionar el tema del pensionamiento con equidad, para su discusión en la política pública del país. El Perú empezó su sistema de pensiones a mediados del siglo pasado con un sistema basado en contribuciones colectivas (SNP), que luego de algunos años empezó a tener problemas financieros. Para solucionarlos, en 1992 se creó un nuevo sistema basado en cuentas individuales de ahorro (SPP), que empezó a funcionar al lado del antiguo sistema desde junio de 1993. Al momento de la investigación, los trabajadores tienen que elegir entre estos dos sistemas para hacer sus aportes obligatorios, pese a que existen otros 16 más pequeños sistemas que en total cubren alrededor del 2.4% de la fuerza laboral (Ministerio de Economía y Finanzas, 2004). Esto crea un panorama complejo que la administración pública debe afrontar con estrategias y políticas innovadoras.

Según la Encuesta Nacional de Hogares (Instituto Nacional de Estadística e Informática, 2017) actualizada al 2022, la población mayor de 14 años y en capacidad de trabajar fue de 18.600 millones personas, de los cuales 13.184 millones declararon estar afiliados a algún sistema de pensiones (70.9%).

A junio de 2022 (Asociación de AFP, 2022), existían 4.706 millones de afiliados al SNP, de los cuales 1.459 millones (31.0%) son cotizantes activos. La cotización obligatoria es del 13% de la remuneración que perciben. El SNP tiene 660,998 pensionistas cuya pensión promedio es de S/ 767.40 soles y que representan el 77.0% de los pensionistas del Sistema Previsional Peruano (Oficina de Normalización Previsional, 2022).

A la misma fecha, existían 8.478 millones de personas afiliadas al Sistema de cuentas individuales de ahorro (SPP), de los cuales, 3.561 millones (42.0%) son cotizantes activos o regulares. El aporte a su fondo individual de pensión es del 10% de la remuneración que perciben, pero también deben pagar otros costos como una tarifa de administración y un seguro que cubre invalidez y otros gastos, costos adicionales que hacen que el aporte sea casi el 13% de la remuneración mensual. A junio del 2022, el SPP tiene 197,476 pensionistas cuya pensión promedio es de S/ 987.00 soles y que representan el 23.0% de los pensionistas del Sistema Previsional Peruano (Superintendencia de Banca, Seguros y AFP, 2022).

Al momento de la presente investigación, el sistema previsional presenta rasgos de fragmentación e inequidad, dado que, dos personas con situaciones parecidas reciben pensiones muy distintas dependiendo si están en el sistema privado (SPP) o en el público (SNP), pudiendo resultar que las pensiones del sistema privado son bajas y en el sistema público podría quedarse sin pensión si no se ha contribuido por al menos 10 años. Este panorama muestra la importancia de entender bien las complejidades y diferencias entre los sistemas de pensiones disponibles.

Formulación del problema.

En base a lo anteriormente expresado, el investigador consideró necesario el formular el problema general de investigación bajo los siguientes términos: ¿De qué manera la distribución de ingresos por pensiones en el sistema previsional peruano ha sido equitativa o inequitativa durante el periodo de 2012 a 2022, y qué factores clave, incluyendo condiciones socioeconómicas y demográficas, han influido en estas disparidades?

En la misma línea, se consideran como problemas específicos de investigación, los siguientes: ¿Cómo la zona geográfica de residencia de los pensionistas influencia las diferencias en sus ingresos en el Sistema Previsional Peruano?; ¿De qué manera el tipo de entidad laboral (pública o privada) donde trabajó el pensionista afecta las diferencias en sus ingresos de pensiones?; ¿Cuál es el impacto del sistema previsional (público o privado) en el que el pensionista aportó, sobre las diferencias en sus ingresos?; ¿Cómo el género del pensionista afecta la distribución de los ingresos en el Sistema Previsional Peruano?; y ¿De qué manera el nivel de estudios alcanzado por los pensionistas influye en las diferencias de sus ingresos de pensiones?"

Objetivo General.

Los objetivos de investigación deben estar dirigidos hacia la obtención de conocimientos (Hernández Sampieri y otros, 2014), y es en ese sentido que el investigador consideró como objetivo general el enunciado a continuación:

Analizar la equidad en la distribución de ingresos por pensiones en el sistema previsional peruano, identificando los factores clave que influyen las disparidades y evaluar cómo estos factores se correlacionan con las condiciones socioeconómicas y demográficas de los pensionistas, durante el periodo comprendido entre los años 2012 al 2022.

Objetivos Específicos.

Analizar el impacto de la zona geográfica en la que vive el pensionista del Sistema Previsional Peruano para determinar cómo este factor contribuye a las diferencias en los ingresos del pensionista.

Analizar el impacto del tipo de entidad laboral (público o privado) en la que prioritariamente trabajó el pensionista del Sistema Previsional Peruano para determinar cómo este factor contribuye a las diferencias en los ingresos del pensionista.

Analizar el impacto del Sistema Previsional (público o privado) receptor primario de los aportes del pensionista del Sistema Previsional Peruano para determinar cómo este factor contribuye a las diferencias en los ingresos del pensionista.

Analizar el impacto del género (hombre o mujer) del pensionista del Sistema Previsional Peruano para determinar cómo este factor contribuye a las diferencias en los ingresos del pensionista.

Analizar el impacto del nivel de estudios alcanzado por el pensionista del Sistema Previsional Peruano para determinar cómo este factor contribuye a las diferencias en los ingresos del pensionista.

Hipótesis principal.

La equidad en la distribución de ingresos por pensiones en el sistema previsional peruano entre 2012 y 2022 ha sido significativamente influenciada por factores socioeconómicos y demográficos específicos, resultando en disparidades notables.

Hipótesis Específicas.

En concordancia y consistencia con los objetivos de investigación, y la hipótesis principal, se consideraron las siguientes hipótesis específicas:

- a) Los pensionistas del Sistema Previsional Peruano que residen en ciertas zonas geográficas reciben ingresos significativamente diferentes, contribuyendo a la disparidad en la distribución de ingresos.

- b) El tipo de entidad laboral (pública o privada) en la que trabajó el pensionista tiene un impacto considerable en la cantidad de sus ingresos de pensiones.
- c) Existe una correlación significativa entre el sistema previsional (público o privado) receptor de los aportes y la cantidad de ingresos que recibe el pensionista.
- d) El género del pensionista es un factor determinante en la disparidad de los ingresos por pensiones en el sistema previsional peruano.
- e) El nivel de educación alcanzado por los pensionistas afecta de manera significativa sus ingresos por pensiones.

Importancia del estudio.

El estudio aborda un tema crítico para la sociedad peruana, dada la importancia de los sistemas de pensiones en la seguridad económica y el bienestar de los ciudadanos mayores. La equidad en los ingresos por pensiones es fundamental para garantizar una vejez digna y segura para todos los peruanos.

La investigación pone de relieve las disparidades existentes en el sistema de pensiones, incluyendo diferencias basadas en la ubicación geográfica, género, nivel educativo, y entre los sistemas de pensiones público y privado. Esta visibilidad es esencial para informar y motivar reformas necesarias.

Los hallazgos proporcionan una base empírica para el diseño y la implementación de políticas públicas más equitativas y efectivas en el ámbito de las pensiones, contribuyendo así a un sistema previsional más justo y sostenible.

Este estudio aporta significativamente al conocimiento académico en el campo de la economía social y de las pensiones, llenando un vacío en la comprensión de cómo los diferentes factores afectan la equidad de los ingresos por pensiones en Perú.

Los resultados son de gran valor para una variedad de stakeholders, incluyendo formuladores de políticas, administradores de sistemas de pensiones, organizaciones de la sociedad civil, y los propios pensionistas, al ofrecer datos e insights detallados sobre el estado actual del sistema.

Al iluminar las áreas problemáticas y las oportunidades de mejora en el sistema de pensiones, la investigación puede servir como un catalizador para el cambio, impulsando debates y acciones hacia un sistema más equitativo y sostenible.

Al final, el estudio tiene el potencial de contribuir a mejoras en la calidad de vida de los pensionistas en Perú, especialmente para aquellos en situaciones vulnerables, a través de una distribución más justa de los recursos de pensiones.

Enfoque y tipo de diseño metodológico.

El estudio de investigación adopta un enfoque cuantitativo, lo que es apropiado para el análisis estadístico de datos y para examinar las relaciones entre variables.

El diseño es principalmente descriptivo y analítico, se describen las características demográficas, socioeconómicas y del sistema previsional de los pensionistas peruanos, y se analizan las relaciones entre estas características y la equidad en los ingresos por pensiones.

Se utilizaron datos secundarios de la Encuesta Nacional de Hogares

(ENAHO), que proporciona información detallada y representativa a nivel nacional sobre las condiciones de vida de los hogares peruanos, incluyendo datos sobre pensiones e ingresos.

Se seleccionaron datos relevantes para la población pensionista, incluyendo variables como zona geográfica, tipo de entidad laboral, sistema previsional, género, nivel de estudios, edad, duración del aporte al sistema de pensiones, estado de salud, acceso a servicios de salud y dependencia económica; datos que se sometieron a una fase de limpieza y preparación para asegurar su calidad y relevancia para el análisis.

Estructura de la tesis.

La tesis se estructura en base al Manual para la elaboración de las tesis y los trabajos de investigación para obtener los grados académicos de maestro y de doctor del Instituto de Gobierno y de Gestión Pública, aprobada según resolución directoral N° 137-2019-IGGP-USMP del 18 de octubre de 2019; cuyo esquema es el siguiente:

- Páginas preliminares (portada, dedicatoria, agradecimientos, índice de contenidos, índice de tablas y figuras, resumen-abstract)
- Introducción
- Capítulo I - Marco teórico
- Capítulo II - Hipótesis y variables
- Capítulo III - Metodología de la investigación
- Capítulo IV - Resultados
- Capítulo V - Discusión
- Conclusiones
- Recomendaciones

- Fuentes de información
- Anexos

CAPÍTULO I - MARCO TEÓRICO

1.1. Antecedentes de la Investigación

(Otero-Bahamón y otros, 2022) en un estudio académico titulado ¿Cómo se reduce la desigualdad en las ciudades colombianas?, analizaron cómo se distribuyen los ingresos en Colombia. Usaron datos de dos grandes encuestas hechas entre los años 2002 y 2018 y se centraron en el impacto de las pensiones utilizando el Coeficiente de Gini. Encontraron que este coeficiente cambia cuando se consideran las pensiones. Específicamente, al incluir las pensiones en el cálculo, la desigualdad, medida por el Gini, disminuye ligeramente. Según los investigadores, esto sugiere que las pensiones son un factor clave para entender la desigualdad en Colombia. Este estudio resalta el papel crucial que desempeñan las pensiones en el contexto de la desigualdad social y económica, lo cual es una observación relevante para cualquier país en el proceso de reformar o ajustar sus sistemas previsionales.

Asimismo, y la misma línea de búsqueda de causas de desigualdad en los ingresos, (Cenas Rodriguez, 2019) examina cómo el crecimiento económico y la educación afectan la desigualdad de ingresos en Perú entre 1997 y 2017. Para hacer esto, usa un modelo matemático sencillo y datos de varias organizaciones importantes como el Banco Mundial y el Instituto Nacional de Estadística. El estudio se centra en cómo han cambiado con el tiempo la desigualdad de ingresos, medida por el coeficiente de Gini, la tasa de analfabetismo y el ingreso promedio por persona. Los hallazgos revelan que cuando el ingreso promedio por persona aumenta, la desigualdad tiende a disminuir; pero cuando la tasa de analfabetismo sube, la desigualdad de

ingresos también aumenta. Este análisis subraya cómo elementos como el crecimiento económico y la educación pueden influir directamente en la desigualdad, lo que podría tener implicaciones importantes para políticas públicas y reformas educativas.

Otra de las investigaciones que contribuyen a la búsqueda de causas de presencia de desigualdad, es la de (Chavez Jorge, 2018), quien investigó cómo el gasto público afecta la desigualdad de ingresos en Perú desde 1997 hasta 2017. Usó información del Banco Mundial y del Ministerio de Economía y Finanzas de Perú para su estudio. Los resultados muestran que el gasto público en Perú ha ayudado a disminuir la desigualdad económica. En base a esto, el investigador sugiere que se deberían crear y poner en marcha políticas más efectivas para reducir la desigualdad, especialmente en áreas donde la brecha es más grande. Esta investigación resalta la importancia del gasto público como herramienta para combatir la desigualdad y sugiere un camino para políticas más efectivas, especialmente en áreas con desigualdades más marcadas.

Años antes, (Rodriguez Weber, 2014) también llevó a cabo un estudio profundo sobre cómo ha evolucionado la desigualdad en Chile desde 1850 hasta 2009. Su meta era entender las raíces de la desigualdad y sus efectos a lo largo del tiempo. Utilizó una técnica que consiste en agrupar a la gente según sus trabajos y estimar cuánto gana cada grupo para entender cómo se distribuye el ingreso año tras año. Llegó a la conclusión de que la desigualdad no es solo el resultado de las fuerzas del mercado, sino también de aspectos sociales y políticos, particularmente del balance de poder entre diferentes grupos. En otras palabras, la desigualdad se forma por una mezcla de factores

económicos, sociales y políticos. También destaca que, para entender completamente la desigualdad, es esencial considerar su desarrollo histórico y cómo influye y es influenciada por otros aspectos sociales. Este estudio ofrece una perspectiva multidimensional sobre la desigualdad, al considerar no solo los aspectos económicos sino también los sociales y políticos, lo que podría ser útil para entender cómo abordar este complejo problema.

Por otro lado, (Mesa-Lago, 2022) en su artículo “Desempeño de pensiones privatizadas en América Latina, 1980-2020: el caso de México”, analiza cómo han funcionado los sistemas de pensiones privados en 10 países de América Latina durante 40 años, desde 1980 hasta 2020. El autor destaca que estos países adoptaron uno de los tres modelos. El primer modelo reemplazó completamente el sistema público con uno privado, como ocurrió en Chile y México. El segundo modelo mantuvo el sistema público y añadió un componente privado, que es el caso de países como Argentina y Uruguay. El tercer modelo dejó que el sistema público y el privado coexistieran y compitieran entre sí, como sucede en Colombia y Perú. El investigador evaluó cada sistema con base en criterios de la Organización Internacional del Trabajo, como la cantidad de personas cubiertas y si las pensiones son suficientes. Uno de los hallazgos más notables es que en siete de los diez países, las pensiones se ajustan anualmente de acuerdo con ciertos factores como la inflación, mientras que, en tres países, como Perú, el gobierno decide si se ajustan o no (OCDE, 2021).

Una de las investigaciones que brinda luces sobre mecanismos de medición de inequidad, es la realizada por (Rodríguez López, 2018) quien utilizó información de salud del Banco Mundial para medir las diferencias en

el bienestar entre dos grupos de personas. Los resultados apuntaron a dos áreas clave. La primera tiene que ver con enfermedades y muertes que podrían evitarse mediante programas de salud preventiva. La segunda área está relacionada con aspectos como instalaciones médicas adecuadas y personal capacitado en los sistemas de salud.

Finalmente, un trabajo de interés por el enfoque metodológico empleado, y que serviría de base para el que la presente investigación, es el que llevó a cabo (Rosas Febres, 2018), que se propone determinar la frecuencia y evolución de los pagos no oficiales realizados por los pacientes que utilizan el Sistema Integral de Salud entre 2008 y 2010. Utiliza como fuente los resultados de la Encuesta Nacional de Hogares de esos años. En el estudio de dos variables, descubre que estos pagos informales están fuertemente vinculados con factores como el año en cuestión, estar en la edad de trabajar, nivel educativo más alto, haber sufrido accidentes o enfermedades crónicas, no vivir en la pobreza, recibir atención en hospitales y la percepción de que el hospital está lejos. La investigación concluye que hubo un aumento en la tendencia de los pagos informales a nivel nacional entre 2008 y 2010, con diferencias notables entre las diferentes regiones administrativas.

1.2. Bases Teóricas.

Teoría de la Justicia (Equidad) de John Rawls

La teoría de la justicia formulada por John Rawls es una de las obras más influyentes en el campo de la filosofía política y la teoría social. Su trabajo busca definir qué es una sociedad justa y cómo debería ser diseñada para garantizar la igualdad y la libertad básica de sus ciudadanos.

En su libro "Teoría de la Justicia", (Rawls, Teoría de la Justicia, 2012) Rawls establece que, una sociedad justa debe estar basada en dos principios de justicia: el principio de libertad igual para todos y el principio de diferencia. El primero de ellos establece que todas las personas deben tener el mismo conjunto básico de libertades, siempre y cuando no interfieran con las libertades de los demás. Esto significa que todos deberíamos tener la libertad de expresión, la libertad de religión, la libertad de asociación, entre otras. El segundo principio, por su parte, establece que las desigualdades sociales y económicas deben ser diseñadas para beneficiar a los menos afortunados en la sociedad. Esto significa que la sociedad debería estar diseñada de tal forma que aquellos que están en una posición más desventajosa tengan más oportunidades y recursos disponibles para ellos.

Rawls argumenta que estos principios de justicia deben ser elegidos por personas imparciales y racionales que se encuentren en una situación inicial de igualdad. En esta situación, los individuos no saben qué posición ocuparán en la sociedad o qué recursos tendrán, lo que les permite elegir los principios de justicia de manera imparcial y justa.

Para Rawls, la justicia como equidad es la forma más adecuada de garantizar una sociedad justa. La equidad implica que los ciudadanos tienen las mismas oportunidades y recursos, lo que permite una mayor movilidad social y promover cambios positivos en las condiciones de vida de todas las personas en la sociedad.

La teoría de la justicia de John Rawls propone un marco teórico basado en los principios de libertad igual para todos y diferencia para los menos afortunados. Su teoría busca garantizar que las instituciones sociales y

políticas de una sociedad estén diseñadas para maximizar la libertad y la igualdad básica de sus ciudadanos. Con esto, busca crear una sociedad donde todos puedan tener las mismas oportunidades y recursos, independientemente de su posición social o económica. (Vidal, 2009)

Teoría de la Justicia (Igualdad) de Amartya Sen.

Amartya Sen es un filósofo y economista indio que ha invertido una buena parte de su trayectoria profesional en la investigación del concepto de justicia y cómo se puede lograr en la sociedad. En su libro "La Idea de Justicia", Sen propone una teoría de la justicia que se centra en el desarrollo de las capacidades y libertades de las personas (Rodríguez, 2015).

Sen critica las teorías de la justicia que se enfocan en la distribución igualitaria de recursos, ya que estas teorías no tienen en cuenta las diferencias en las necesidades y capacidades de las personas. Por ejemplo, una persona con una discapacidad puede necesitar más recursos que una persona sin discapacidad para tener las mismas oportunidades de desarrollar sus capacidades y libertades.

En cambio, Sen argumenta que la justicia debe basarse en el concepto de "capacidad", que se refiere a las oportunidades y libertades que las personas tienen para desarrollarse plenamente como seres humanos. Estas capacidades incluyen cosas como la educación, la salud, la seguridad, la participación política y la libertad de expresión.

Por lo tanto, la justicia no solo implica distribuir recursos de manera igualitaria, sino también garantizar que todas las personas cuenten con las mismas posibilidades de potenciar sus capacidades y libertades. Sen llama a esto "justicia de las capacidades".

Además, Sen sostiene que la justicia debe ser evaluada desde una perspectiva comparativa, es decir, debemos analizar cómo las diferentes políticas y decisiones afectan a las personas con mayor vulnerabilidad de la sociedad. No se trata solo de lograr una igualdad abstracta, sino de reducir las desigualdades reales que existen en la sociedad. En otras palabras, la justicia no se trata solo de distribuir recursos, sino de garantizar que todos los integrantes de una comunidad tengan oportunidad de potenciar sus capacidades y libertades, independientemente de su origen socioeconómico o de cualquier otra condición que pueda afectar sus oportunidades.

Sen resalta también la relevancia de la participación democrática en la búsqueda de la equidad. En su perspectiva, la equidad solo puede ser alcanzada cuando todas las personas tienen la ocasión de involucrarse en el proceso de tomar elecciones que impactan sus vidas. Esto engloba la contribución en asuntos políticos y en la sociedad, junto con la habilidad de compartir sus puntos de vista y ser tomados en cuenta.

La teoría de la justicia de Amartya Sen se enfoca en el desarrollo de las capacidades y libertades de las personas, y en la necesidad de reducir las desigualdades reales que existen en la sociedad. Para Sen, la justicia no es algo estático, sino que debe ser evaluada constantemente a la luz de las necesidades y realidades cambiantes de la sociedad. La justicia solo se puede lograr a través de la participación democrática y la garantía de igualdad de oportunidades para todos.

Igualdad y Justicia según Martha Nussbaum.

Martha Nussbaum es una filósofa contemporánea que ha desarrollado una teoría de la justicia que se centra en la importancia de las capacidades

humanas y la interdependencia entre los seres humanos. Su enfoque es interesante porque desafía la idea convencional de que, la justicia se reduce a la distribución equitativa de recursos materiales.

Para Nussbaum, la justicia no se trata solo de la distribución equitativa de bienes y recursos, sino que también implica garantizar que todas las personas tengan la oportunidad de potenciar sus capacidades humanas básicas. Estas capacidades incluyen cosas como la capacidad de tener relaciones significativas con otros, de controlar nuestro destino y de experimentar el placer y el dolor. Según Nussbaum, estas capacidades son esenciales para llevar una vida plena y satisfactoria.

Además, Nussbaum sostiene que la justicia también implica reconocer nuestra interdependencia como seres humanos. Todos dependemos de otros en algún momento de nuestras vidas, ya sea para satisfacer nuestras necesidades físicas o emocionales. Como resultado, la justicia requiere que nos aseguremos de que todas las personas tengan acceso a los recursos necesarios para llevar una vida plena. Esto incluye no solo los recursos materiales, sino también los sociales, emocionales y educativos.

Por lo tanto, la teoría de la justicia de Nussbaum implica una responsabilidad colectiva para garantizar que todas las personas tengan las oportunidades necesarias para desarrollar sus capacidades humanas básicas. Esta responsabilidad no recae solo en el gobierno, sino que también incluye a la sociedad en su conjunto.

La teoría de la justicia de Martha Nussbaum es una alternativa interesante a la idea convencional de que, la justicia se reduce a la distribución equitativa de bienes y recursos. Al enfatizar la importancia de las capacidades

humanas básicas y la interdependencia entre los seres humanos, Nussbaum nos desafía a considerar nuevas formas de abordar la justicia y la igualdad en nuestra sociedad. (Nussbaum, *Human Functioning and Social Justice: In Defense of Aristotelian Essentialism*, 1992)

Teorías de la Justicia.

A pesar de que en los últimos tiempos han emergido enfoques meta sobre la justicia, que buscan armonizar o conciliar diversas teorías de la justicia, como lo ilustra el caso de (Stilman, 2022), quien argumenta que una sociedad justa será aquella que logre amalgamar y amalgamar, bajo el principio democrático, las concepciones particulares de la justicia sostenidas por sus miembros. En este estudio, el investigador sostiene que las teorías de justicia más influyentes en la actualidad son: la noción de "justicia como equidad" de John Rawls y la teoría de la "igualdad de capacidades" de Amartya Sen. Ambas teorías se insertan en la corriente del igualitarismo liberal.

Rawls formula una teoría distributiva de la justicia que sigue líneas liberales, buscando una aceptación universal y otorgando prioridad a la libertad sobre la igualdad. No obstante, contempla la necesidad de corregir las desigualdades en la libertad mediante una apuesta por la igualdad de oportunidades.

Amartya Sen, por su parte, dirige el debate sobre la igualdad no hacia la pregunta de "¿por qué igualdad?", sino más bien hacia la interrogante de "¿igualdad de qué?". Para él, todas las teorías sobre la estructura social tienden a exigir un tratamiento equitativo de los individuos en alguna

dimensión relevante, en relación a alguna variable crucial para la teoría en cuestión.

Equidad

La equidad es un concepto complejo y multifacético que ha sido definido de diversas maneras por diferentes autores. Algunos de ellos, como Amartya Sen, Martha Nussbaum, John Rawls y Paulo Freire, han argumentado que la equidad implica tratar a todas las personas de manera justa e imparcial, pero han enfatizado diferentes aspectos de este concepto.

Para (Sen, 1999) la equidad no se limita a la igualdad formal, sino que también incluye la igualdad de oportunidades y el acceso a recursos básicos. Según la teoría de Sen, la equidad se alcanza cuando todas las personas tienen la capacidad de tomar decisiones y llevar a cabo sus propias vidas de acuerdo con sus propios valores y objetivos. En contraste, (Nussbaum, Crear capacidades: Propuesta para el desarrollo humano, 2012) sostiene que la equidad significa tratar a todas las personas como seres humanos dignos de respeto y consideración, y se enfoca en promover y proteger las capacidades humanas básicas como la salud, la educación y la libertad de expresión.

(Rawls, Teoría de la Justicia, 2012) por su parte, propone una teoría de la justicia que se basa en el concepto de equidad y establece que, las desigualdades sociales y económicas solo son justas si benefician a las personas menos favorecidas de la sociedad. Según Rawls, la equidad implica tratar a todas las personas de manera justa e imparcial, independientemente de su posición social o económica. Por otro lado, (Freire, 2005) se enfoca en la liberación de todas las personas a través de la promoción de la conciencia crítica y la participación activa en la toma de decisiones que afectan sus vidas.

Estos autores coinciden en que la equidad implica tratar a todas las personas de manera justa e imparcial, sus enfoques difieren en cuanto a los aspectos clave de este concepto. Mientras que Sen enfatiza la igualdad de oportunidades y el acceso a recursos básicos, Nussbaum se enfoca en promover y proteger las capacidades humanas básicas. Rawls propone un enfoque de la equidad basado en el principio de diferencia, y Freire se enfoca en la liberación a través de la conciencia crítica y la participación activa. En última instancia, todas estas perspectivas pueden contribuir a una comprensión más completa y enriquecedora de la equidad como un valor fundamental en nuestra sociedad.

La creciente disparidad en la disponibilidad de recursos entre países y personas pone de manifiesto la necesidad de centrarse en la equidad como un componente esencial para el desarrollo sostenible. La inequidad, que básicamente es una forma de desigualdad, se ve como una manifestación de injusticia. Esta no es una cuestión nueva; ha existido desde los albores de la civilización y es prevalente en diversas partes del mundo, algunas más que otras. A pesar de los esfuerzos considerables de organizaciones dedicadas a erradicarla, el desafío de eliminar la inequidad sigue siendo grande.

En el ámbito social es donde más se observa esta disparidad. Las personas con más recursos económicos pueden acceder fácilmente a bienes y servicios en su región, mientras que otros grupos socialmente desfavorecidos luchan por obtener lo mismo. Esto ilustra de forma clara el problema de inequidad que prevalece. En este contexto, los que carecen de medios financieros suficientes suelen ser tratados de manera diferente en

comparación con aquellos que sí tienen la capacidad económica para obtener lo que necesitan.

Envejecimiento poblacional

El creciente número de personas mayores dentro de una población es un fenómeno demográfico que surge debido a la caída de las tasas de natalidad y al incremento de la esperanza de vida. Este cambio demográfico lleva consigo una serie de consecuencias que impactan diversos ámbitos, desde lo económico hasta lo político, y que requieren una atención cuidadosa por parte de los gobiernos y de la sociedad en su conjunto.

Desde una perspectiva económica, el envejecimiento de la población conlleva retos significativos. La combinación de una menor tasa de natalidad y una vida más larga genera una reducción en la mano de obra disponible y aumenta la presión sobre los sistemas de seguridad social, como los de salud y pensiones. Además, el consumo de bienes y servicios por parte de la población mayor puede disminuir, lo que podría afectar negativamente la economía en general.

En el ámbito social, el envejecimiento de la población presenta también desafíos importantes. Las personas mayores son más susceptibles a enfrentar condiciones como la pobreza, la soledad y el aislamiento social, lo cual puede mermar su bienestar y salud. La demanda de atención médica y de cuidados a largo plazo también tiende a incrementar, lo cual podría ser un desafío para ciertas comunidades que carecen de los recursos necesarios para cubrir estas necesidades.

Desde un punto de vista político, los gobiernos tienen la tarea de satisfacer las demandas específicas de una población que envejece, la cual

podría tener expectativas más altas en relación a servicios sociales y atención médica. Las personas mayores también suelen ser una fuerza política considerable, con la capacidad de influir en decisiones y en la distribución de recursos.

Para hacer frente al envejecimiento poblacional, es importante afrontar estos retos desde una perspectiva multidisciplinaria. En primer lugar, se necesitan políticas que fomenten una tasa de natalidad sostenible, para garantizar un equilibrio adecuado entre la población joven y la población mayor. En segundo lugar, se deben establecer sistemas de seguridad social y atención médica adecuados, para garantizar una atención de calidad a los ancianos y reducir la vulnerabilidad de esta población. En tercer lugar, se deben fomentar políticas de integración social, para hacer frente a la soledad y el aislamiento social de los adultos mayores (Cardona & Peláez, 2012)

Esperanza de vida

La gente en todo el mundo está viviendo más tiempo que antes. Ahora, la mayoría de las personas pueden esperar vivir al menos hasta los 60 años. Y no solo hay más personas mayores, sino que su proporción dentro de la población total también está creciendo. Según la Organización Mundial de la Salud, se espera que para 2050 haya tres veces más personas mayores de 80 años que en 2020, llegando a unos 426 millones (Organización Mundial de la Salud, 2022).

Cuando se habla de 'esperanza de vida', usualmente los autores se refieren al número promedio de años que se espera que una persona viva. Este número se calcula basándose en las tasas actuales de mortalidad para diferentes edades. Es una manera de resumir cuán bien o mal está un país

en términos de salud y mortalidad. Este dato ayuda a comparar países y ver cómo están avanzando con el tiempo.

La esperanza de vida es tan importante que forma parte de un indicador más grande llamado 'Índice de Desarrollo Humano', que también incluye la educación y el Producto Bruto Interno (PBI). Vivir una vida larga y saludable es una de las mejores maneras de medir el progreso social de un país. Este aumento en la esperanza de vida se debe a varias razones, como mejores condiciones de vida, avances en la atención médica y una disminución general en las tasas de mortalidad.

Hay algo más que se debería tener en cuenta: cuando la esperanza de vida aumenta, no es solo porque las personas mayores están viviendo más tiempo, también significa que menos jóvenes están muriendo. En resumen, un aumento en la esperanza de vida es una buena noticia para todos, sin importar la edad¹.

1.3. Definición de términos básicos

Política Social

La política social se refiere a cómo el Estado organiza y financia servicios públicos esenciales para mejorar la vida de los ciudadanos. Algunos ejemplos de áreas en las que se enfoca la política social son la educación, la salud, la vivienda, el trabajo y la seguridad social.

En este sentido, (Esping-Andersen, 1990), uno de los autores más reconocidos en el estudio de la política social, argumenta que, el gobierno tiene entre otras, la responsabilidad de garantizar la protección social de los

¹ <https://cenie.eu/es/blogs/envejecer-en-sociedad/la-esperanza-de-vida-que-es-y-por-que-importa>

ciudadanos, independientemente de su estatus socioeconómico. Esping-Andersen sostiene que los estados de bienestar pueden clasificarse en tres categorías: el liberal, el conservador y el socialdemócrata, según la forma en que se organiza la política social.

En el caso del estado de bienestar liberal, el enfoque principal es proporcionar seguridad social a aquellos que no pueden trabajar o que no tienen trabajo. Por otro lado, el estado de bienestar conservador se centra en apoyar a las familias y fomentar la participación activa en el mercado laboral. Finalmente, el estado de bienestar socialdemócrata tiene como objetivo proporcionar servicios públicos universales y garantizar una protección social completa para todos los ciudadanos.

Otro autor relevante en el estudio de la política social es T.H. Marshall (Marshall & Bottomore, 2023), quien argumenta que la ciudadanía no solo implica la pertenencia a un Estado-nación, sino también la participación activa en la vida social, política y económica del país. Para Marshall, la política social es fundamental para garantizar que todos los ciudadanos tengan igualdad de oportunidades y acceso a los recursos necesarios para una vida digna. Marshall destaca la importancia de los derechos civiles, además de los derechos políticos y sociales en la formación de una sociedad justa.

Coeficiente de GINI

El coeficiente de Gini es como un termómetro que mide la desigualdad en una comunidad, un país o incluso en todo el mundo. Este coeficiente es un número que va del 0 al 1. Si el número es 0, significa que todo el mundo tiene la misma cantidad de dinero o recursos, es como si todos compartieran el

pastel por igual. Pero si el número llega a 1, es como si una persona se llevara todo el pastel y no dejara nada para los demás.

¿Cómo se llega a este número?, se usa la 'curva de Lorenz', que es una gráfica que muestra cómo está repartido el dinero o los recursos entre la gente. Se compara esta curva con una línea recta que representaría la situación perfecta donde todo el mundo tiene la misma cantidad. Cuanto más se aleje la curva de Lorenz de esta línea recta, más alto será el coeficiente de Gini, y eso significa que hay más desigualdad en esa comunidad o país.

Sistema contributivo de reparto

Un sistema contributivo de reparto es un sistema de pensiones en el que los trabajadores en activo aportan una parte de sus ingresos a un fondo común, que se utiliza para financiar las pensiones de los jubilados. El monto de la pensión que recibe cada jubilado depende de la cantidad de años cotizados y del sueldo promedio que percibió durante su vida laboral.

Desde la perspectiva de un sistema previsional, un sistema contributivo de reparto es un sistema solidario, en tanto que las pensiones de los jubilados se financian con las cotizaciones de los trabajadores en activo, independientemente de sus ingresos o de su situación laboral. También es un sistema redistributivo, en tanto los trabajadores que tienen ingresos más altos cotizan más que los trabajadores que tienen ingresos más bajos, lo que permite que las pensiones de los jubilados sean más equitativas.

Sistema de capitalización individual

Un sistema de capitalización individual es un sistema de pensiones en el que los trabajadores en activo aportan una parte de sus ingresos a una cuenta personal, que se utiliza para financiar su pensión de jubilación. El

monto de la pensión que recibe cada trabajador depende de la cantidad de dinero que haya acumulado en su cuenta personal.

Desde la perspectiva de un sistema previsional, en un sistema de capitalización individual los trabajadores deben ahorrar durante su vida laboral para poder financiar su pensión. Los trabajadores pueden elegir la administradora de fondos de pensiones (AFP) que les ofrezca las mejores condiciones.

Ahorro previsional

El ahorro previsional es el dinero que una persona destina a un fondo de pensiones o sistema de seguridad social para asegurarse un ingreso o pensión en el futuro, cuando deje de trabajar.

En muchos países, existen sistemas de seguridad social o fondos de pensiones que permiten a las personas ahorrar una parte de sus ingresos para financiar su

retiro. Estos ahorros previsionales suelen ser obligatorios y pueden ser administrados por el Estado o por entidades privadas autorizadas. El objetivo del ahorro previsional es asegurarse una fuente de ingresos durante la etapa de jubilación o retiro, ya que en general se asume que la persona tendrá menos ingresos que cuando estaba trabajando y necesitará una fuente de financiamiento para mantener su calidad de vida. Además, el ahorro previsional también puede servir como una forma de protección ante imprevistos como enfermedades o accidentes que impidan a la persona seguir trabajando.

Previsión Social

La previsión social es un término que se utiliza para referirse a un conjunto de medidas y servicios que se organizan con el fin de proteger a las

personas contra situaciones de riesgo social. Estas situaciones pueden incluir la enfermedad, el desempleo, la vejez, la discapacidad, entre otros.

La previsión social busca garantizar que todas las personas tengan acceso a recursos y apoyos que les permitan hacer frente a situaciones adversas y mantener un nivel de vida digno. Estos recursos y apoyos pueden ser proporcionados por el Estado, las empresas y otros actores sociales.

Algunas de las formas en que se puede expresar la previsión social incluyen la seguridad social, las prestaciones laborales, las ayudas sociales y las políticas de vivienda. La seguridad social es un sistema de protección social que garantiza el acceso a servicios de salud, pensiones, seguro de desempleo, entre otros. Las prestaciones laborales son beneficios que se otorgan a los trabajadores, como días de vacaciones, aguinaldo, seguro médico, entre otros. Las ayudas sociales son recursos económicos o apoyos en especie que se proporcionan a personas en situación de vulnerabilidad, como personas mayores, personas con discapacidad, madres solteras, entre otros. Las políticas de vivienda son medidas que buscan garantizar el acceso a viviendas dignas y adecuadas para todos los sectores de la población.

CAPÍTULO II - HIPÓTESIS Y VARIABLES

2.1. Formulación de hipótesis

De acuerdo con (Hernández Sampieri y otros, 2014) las hipótesis son como los mapas de un tesoro en una investigación. Nos dicen hacia dónde vamos y qué es lo que estamos intentando demostrar. Son ideas iniciales que tenemos sobre lo que podría estar pasando en la situación que estamos estudiando. Ahora, según (Espinoza, 2018), algunas hipótesis van más allá y hablan sobre la 'causalidad'. Esto significa que no solo señalan que hay una conexión entre ciertas cosas (las variables), sino que una de ellas podría ser la razón por la que la otra sucede; y siendo del parecer del investigador que este tipo de hipótesis es el que mejor orientará la investigación, la formula en los siguientes términos:

Hipótesis principal:

La equidad en la distribución de ingresos por pensiones en el sistema previsional peruano entre los años 2012 y 2022 ha sido significativamente influenciada por factores socioeconómicos y demográficos específicos, resultando en disparidades notables.

Hipótesis Específicas:

Hipótesis Específica 1 - Los pensionistas del Sistema Previsional Peruano que residen en ciertas zonas geográficas reciben ingresos significativamente diferentes, contribuyendo a la disparidad en la distribución de ingresos.

Hipótesis Específica 2 - El tipo de entidad laboral (pública o privada) en la que trabajó el pensionista tiene un impacto considerable en la cantidad de sus ingresos de pensiones.

Hipótesis Específica 3 - Existe una correlación significativa entre el sistema previsional (público o privado) receptor de los aportes y la cantidad de ingresos que recibe el pensionista.

Hipótesis Específica 4 - El género del pensionista es un factor determinante en la disparidad de los ingresos por pensiones en el sistema previsional peruano.

Hipótesis Específica 5 - El nivel de educación alcanzado por los pensionistas afecta de manera significativa sus ingresos por pensiones.

2.2. Variables y definición operacional.

(Hernández Sampieri y otros, 2014) dicen que una variable es como un rasgo o característica que puede cambiar, y que podemos medir o ver cómo cambia. Estas variables se vuelven realmente importantes para la ciencia cuando empiezan a interactuar entre sí. Es decir, cuando forman parte de una idea más grande o de una teoría sobre cómo funcionan las cosas.

Por otro lado, (MARKA Investigación, 2022) nos dice que hay diferentes tipos de variables. La 'variable independiente' es como el jefe; no depende de nadie más. La 'variable dependiente' es como el empleado, su comportamiento cambia según lo que haga la variable independiente, o el jefe. Y luego está la 'variable interviniente', que es como un tercer personaje que se mete en la relación entre el jefe y el empleado, cambiando la dinámica entre ellos.

Variable dependiente:

Equidad de ingresos por pensiones: Se refiere al grado de justicia o imparcialidad en la distribución de ingresos por pensiones entre los pensionistas.

Variables independientes e intervinientes:

- Ingresos por pensión: Monto de la pensión recibida por los pensionistas.
- Zona geográfica: Ubicación geográfica del pensionista (costa, sierra, selva, Lima metropolitana).
- Tipo de entidad en la que trabajó: Clasificación del lugar de trabajo anterior del pensionista (entidad pública, empresa privada).
- Sistema previsional al que pertenece el pensionista: Clasificación del sistema de pensiones (Sistema público ONP, Sistema privado AFP).
- Sexo: Género del pensionista.
- Nivel de estudios alcanzado: Grado académico o nivel de educación completado por el pensionista.
- Monto de pensión expresado en soles.

CAPITULO III – METODOLOGÍA DE LA INVESTIGACIÓN

3.1. Diseño metodológico

El enfoque de investigación es de tipo cuantitativo, de diseño no experimental, transversal; descriptivo y analítico, en tanto recopila información cuantificable para ser utilizada en el análisis estadístico (Hernández Sampieri y otros, 2014), es asimismo explicativo, en tanto busca identificar aquellos factores que influyen en la calidad de vida del pensionista peruano que inciden con mayor determinación en la equidad de ingresos por pensión de jubilación.

Se utilizaron datos secundarios de la Encuesta Nacional de Hogares (ENAH), que proporciona información detallada y representativa a nivel nacional sobre las condiciones de vida de los hogares peruanos, incluyendo datos sobre pensiones e ingresos.

Se seleccionaron datos relevantes para la población pensionista, incluyendo variables como zona geográfica, tipo de entidad laboral, sistema previsional, género, nivel de estudios, edad, duración del aporte al sistema de pensiones, estado de salud, acceso a servicios de salud y dependencia económica; datos que se sometieron a una fase de limpieza y preparación para asegurar su calidad y relevancia para el análisis.

3.2. Diseño muestral

(Hernandez Sampieri y otros, 2021) definen el muestreo censal como "el procedimiento mediante el cual se selecciona a la totalidad de los elementos de una población" (p. 256). Los autores argumentan que el muestreo censal es el tipo de muestreo más preciso, ya que permite obtener resultados que representan fielmente a la población. Sin embargo, también señalan que es el tipo de muestreo más costoso y laborioso.

(Gómez Pérez, 2015) define el muestreo censal como "la técnica que consiste en estudiar a la totalidad de los elementos de una población" (p. 105). El autor argumenta que el muestreo censal es el tipo de muestreo más adecuado cuando es necesario obtener resultados precisos y representativos de la población.

Estos autores coinciden en que el muestreo censal es el tipo de muestreo más preciso, pero que también es el más costoso y laborioso. Su uso está recomendado cuando es necesario obtener resultados precisos y representativos de la población, o cuando la población es pequeña o manejable.

En esta investigación, la población de estudio estuvo conformada por todos aquellos que a junio de 2022 son pensionistas del sistema previsional peruano. Según datos de la Asociación de AFP, a dicha fecha, existían 858,474 pensionistas. De ese total, 660,998 pertenecen al sistema nacional de pensiones (SNP), y 197,476 al sistema privado de pensiones (SPP) (Asociación de AFP, 2022).

El censo nacional 2017 actualizado al 2022, recogió información de las personas que vivieron al menos un mes en una casa y que dijeron que reciben una pensión. Según (Ramírez, 2010), cuando un estudio toma en cuenta a todo el grupo objetivo, se le llama 'muestra censal'.

3.3. Técnicas de recolección de datos

En el contexto de la metodología de investigación científica, el uso de bases de datos estructurados como fuente de datos como alternativa al trabajo de campo, es una tendencia que ha ido ganando aceptación y reconocimiento en el ámbito académico y de la investigación. Uno de los

autores más destacados en esta área es Yin, quien argumenta que las bases de datos estructurados pueden ser una fuente de datos válida y confiable para la investigación científica, y pueden ser utilizadas para realizar una amplia gama de investigaciones, incluyendo estudios de caso, estudios transversales y estudios longitudinales (Yin, 2014). Quinn Patton, señala que las bases de datos estructurados pueden ser una herramienta valiosa para la recopilación y organización eficiente y efectiva de datos cualitativos (Quinn Patton, 2015). Collier, señala que las bases de datos estructurados son una herramienta esencial para la gestión de datos de investigación, puesto que pueden ayudar a los investigadores a organizar, almacenar, preservar y compartir sus datos de manera segura y eficiente (Collier, 2016).

En general, los autores que defienden el uso de bases de datos estructurados en la investigación científica argumentan que esta tendencia se basa en varios puntos clave:

- a) Las bases de datos estructuradas ofrecen acceso a una gran cantidad de información que puede ser difícil o imposible de obtener a través del trabajo de campo tradicional. Por ejemplo, bases de datos que recopilan datos a nivel nacional o de largos períodos históricos permiten a los investigadores explorar tendencias y patrones que de otro modo serían inaccesibles.
- b) Los datos almacenados en bases de datos suelen seguir un formato estandarizado, lo que facilita su análisis y comparación. Además, la disponibilidad de estos datos permite a otros investigadores validar o reproducir estudios, un aspecto fundamental del método científico.

- c) Realizar trabajo de campo, especialmente en ubicaciones remotas o para grandes muestras, puede ser muy costoso y llevar mucho tiempo. Las bases de datos proporcionan una forma de realizar análisis significativos con una inversión de recursos considerablemente menor.
- d) El uso de bases de datos permite combinar datos de diferentes disciplinas, abriendo caminos para enfoques de investigación innovadores y multidisciplinarios.
- e) Con el avance de las tecnologías de la información, ahora es posible manejar y analizar grandes volúmenes de datos (Big Data). Esto ha abierto nuevas posibilidades para análisis complejos y profundos que antes no eran factibles.

El investigador está investigando que la recolección de datos en entornos donde los conjuntos de datos son masivos, como una base de datos estructurada con millones de registros, es necesario contar con herramientas y técnicas eficientes que permitan extraer los datos de manera fiable y efectiva. Python, un lenguaje de programación versátil y ampliamente utilizado, ofrece numerosas ventajas para la recolección de datos en este tipo de escenarios. (Smith y otros, 2019). Según (McKinney, 2017), Python se ha convertido en una opción destacada para la recolección de datos desde bases de datos estructuradas con millones de registros, debido a su capacidad de conectarse a estas bases de datos y a la disponibilidad de gran variedad de bibliotecas especializadas (Deep learning, Machine learning, Cálculo numérico, Visualización, Inteligencia artificial explicable, Procesamiento de lenguaje natural).

Habiéndose planteado que para la investigación se utilizaría la Encuesta Nacional de Hogares (ENAH) cuyos datos y resultados se encuentran en grandes bases de datos que son administradas por el Instituto Nacional de Estadística e Informática (INEI), dependencia oficial de gobierno que genera informes técnicos de propósito específico para los diferentes aspectos de política y gestión pública del país; el investigador consideró pertinente utilizar la técnica “extracción de datos mediante programación” para la obtención de los datos necesarios para el logro de los objetivos de investigación.

Los datos del INEI son oficiales y son publicados en la dirección: <https://proyectos.inei.gob.pe/microdatos/>, que cuenta con más de 54 millones de registros. La data utilizada, es aquella que se despliega en la sección denominada “Metodología ACTUALIZADA” y considerando los datos relacionados a las condiciones de vida y pobreza en el periodo comprendido entre los años 2012 a 2022, y específicamente el módulo correspondiente a empleo e ingresos.

3.4. Técnicas estadísticas para el procesamiento de información

Una vez que los datos han sido recolectados, existen diversas técnicas de procesamiento que se pueden utilizar para analizar y obtener información de los datos. Para efectos de esta investigación, se utilizaron las siguientes:

- a) *Limpieza de datos*: Se identificaron y corrigieron problemas o errores en los datos recolectados, como valores atípicos, datos faltantes o inconsistentes. Se reemplazaron valores faltantes, y se detectaron y eliminaron valores atípicos.

- b) *Transformación de datos*: Se codificaron las variables categóricas y se crearon nuevas variables derivadas a partir de las existentes. Se generan variables dummy a las características: zona geográfica, entidad de trabajo, sistema pensionario, sexo.
- c) *Minería de datos*: Esta técnica implica descubrir patrones, relaciones y conocimientos útiles a partir de data ENAHO. Se utilizaron algoritmos de aprendizaje automático y técnicas estadísticas para explorar los datos y encontrar información oculta. Se emplearon como filtros básicos, los datos correspondientes a monto de pensión, y sistema previsional (AFP/ONP); y como características asociadas a los ingresos del pensionista aquellas que plantea el investigador en su hipótesis principal: zona geográfica (ubigeo), entidad donde trabajó la persona antes de su jubilación, género, edad, y nivel de estudio alcanzado.
- d) *Ingeniería de datos*: Se utilizó el software R 4.2.2 para la determinación del porcentaje de participación, ingresos para cada sistema de pensiones, en la ventana de tiempo conformada entre los años 2012 al 2022 con intervalo de 2 años, además del coeficiente Gini para determinar la equidad entre los sistemas pensionarios, la información fue segregada por zonas: costa, sierra, selva, Lima metropolitana y total Perú.
- e) *Modelado y predicción*: Se utilizó específicamente la técnica de árboles de regresión potenciados por gradientes XGBoost para modelar la relación entre la variable dependiente y las variables independientes. Finalmente, se trabajó con modelamientos de Machine Learning por ser técnicas modernas con mayor capacidad para generar predicciones con cercanía a

la realidad; además, de ser el estado del arte en analítica de datos, se propuso modelos de regresión de árboles de decisión

- f) *Analítica de datos*: Para la evaluación de los factores de impacto se utilizó el diagrama de cascada (waterfall), y todo el proceso de analítica de datos fue llevado a cabo mediante el software Python 3.11.2 con la librería SHAP.

3.5. Aspectos éticos

El investigador declara que toda la investigación es original y creada por su propio esfuerzo intelectual. Al utilizar información de otros sitios, como libros, estudios o revistas, se aseguró de dar el crédito correspondiente a los autores citándolas en las notas al pie y en la bibliografía.

Si alguna persona tuviese en el futuro, problemas con la ausencia o mala acreditación de texto de su autoría utilizada en la investigación, el investigador se hace cargo del asunto, y de esta manera, la Universidad de San Martín de Porres y su Instituto de Gobierno y Gestión Pública no tienen que preocuparse por posibles problemas legales relacionados con derechos de autor.

Además, el investigador declara que este estudio no ha sido presentado a otra entidad académica, para la obtención de grado académico alguno.

CAPITULO IV – RESULTADOS

Teniendo en consideración que la hipótesis de investigación es “La equidad en la distribución de ingresos por pensiones en el sistema previsional peruano entre los años 2012 y 2022 ha sido significativamente influenciada por factores socioeconómicos y demográficos específicos, resultando en disparidades notables”, se describen en consistencia; los resultados obtenidos.

4.1. Ingreso por pensión, como factor de equidad

Para realizar este primer análisis de datos, se utilizó técnicas de minería de datos con la finalidad de ubicar indicadores que permitan conocer la evolución de la pobreza monetaria, del bienestar y de las condiciones de vida de los hogares, además de efectuar diagnósticos sobre las condiciones de vida y pobreza monetaria de la población. Se consideraron los datos relacionados a ingresos por pensiones de jubilación, segmentándolos asimismo por fondo pensionario (AFP, ONP). Finalmente, se identificaron las características que, según juicio experto del investigador, impactan en los ingresos de los pensionistas: zona geográfica (ubigeo), tipo de entidad donde trabajó la persona antes de su jubilación, sistema pensionario, sexo, edad, nivel de estudio e ingreso por pensión de jubilación en soles (S/).

Tabla 1

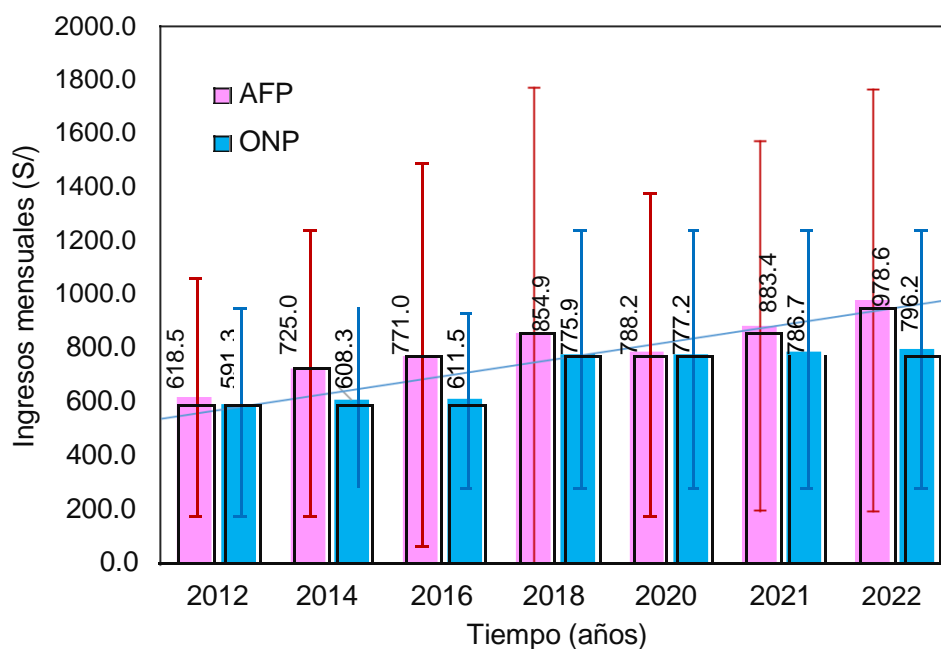
Ingreso mensual promedio, por fondo previsional, Perú, 2012 al 2022

Periodo	Ingresos (S/)		Participación (%)	
	AFP	ONP	AFP	ONP
2012	618.5	591.3	13.9	86.1
2014	725.0	608.3	14.8	85.2
2016	771.0	611.5	17.2	82.8
2018	854.9	775.9	13.5	86.5

2020	788.2	777.2	13.8	86.2
2021	883.4	786.7	13.2	86.8
2022	978.6	796.2	12.6	87.4

Figura 1

Ingreso mensual promedio, por fondo previsional, Perú, 2012 al 2022



La tabla 1 y la figura 1 muestran la media de ingresos por concepto de pensión tanto en el fondo de pensiones que administra el estado peruano, como el que es administrado por el sector privado en el periodo comprendido entre los 2012 al 2022, donde se aprecia una tendencia creciente en ambos fondos, siendo más notorio en la AFP (valores estuvieron entre S/ 618.5 a S/ 978.6) a comparación de la ONP (valores estuvieron entre S/ 519.3 a S/ 796.2).

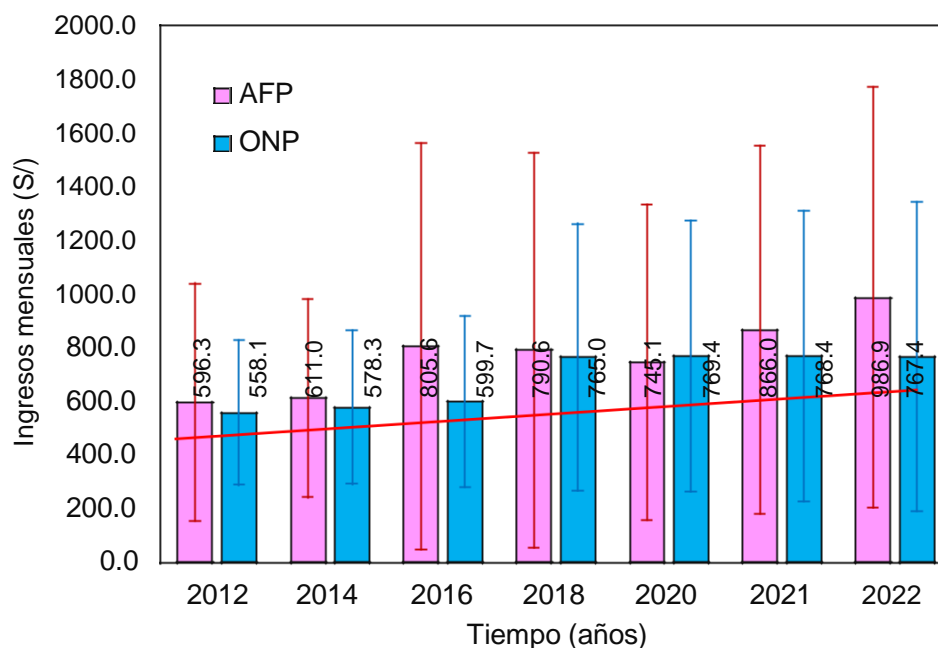
Tabla 2

Ingreso mensual promedio, por fondo previsional, región costa, 2012 al 2022

Periodo	Ingresos (S/)		Participación (%)	
	AFP	ONP	AFP	ONP
2012	596.3	558.1	11.9	88.1
2014	611.0	578.3	10.1	89.9
2016	805.6	599.7	14.2	85.8
2018	790.6	765.0	10.8	89.2
2020	745.1	769.4	11.6	88.4
2021	866.0	768.4	11.5	88.5
2022	986.9	767.4	11.4	88.6

Figura 2

Ingreso mensual promedio por fondo previsional, región costa, 2012 al 2022



En la tabla 2 y figura 2 se puede apreciar una tendencia creciente en ambos sistemas pensionarios, siendo más notorio en la AFP (valores estuvieron entre S/ 596.3 a S/ 986.9) a comparación de la ONP ((valores estuvieron entre S/ 558.1 a S/ 767.4).

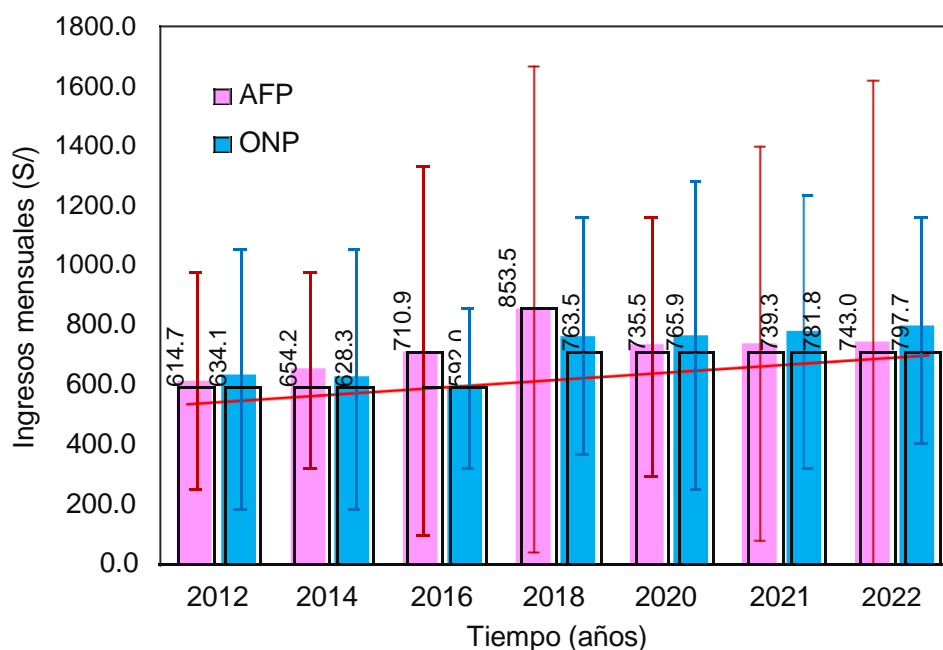
Tabla 3

Ingreso mensual promedio, por fondo previsional, región sierra, 2012 al 2022

Periodo	Ingresos (S/)		Participación (%)	
	AFP	ONP	AFP	ONP
2012	614.7	634.1	10.9	89.1
2014	654.2	628.3	15.3	84.7
2016	710.9	592.0	16.2	83.8
2018	853.5	763.5	12.4	87.6
2020	735.5	765.9	12.5	87.5
2021	739.3	781.8	9.1	90.2
2022	743.0	797.7	5.8	94.2

Figura 3

Ingreso mensual promedio, por fondo previsional, región sierra, 2012 al 2022



En la tabla 3 y figura 3, se observa también una tendencia creciente en ambos sistemas pensionarios, tanto para la ONP (valores estuvieron entre S/ 592.0 a S/ 797.7) y la AFP (valores estuvieron entre S/ 614.7 a S/ 853.5).

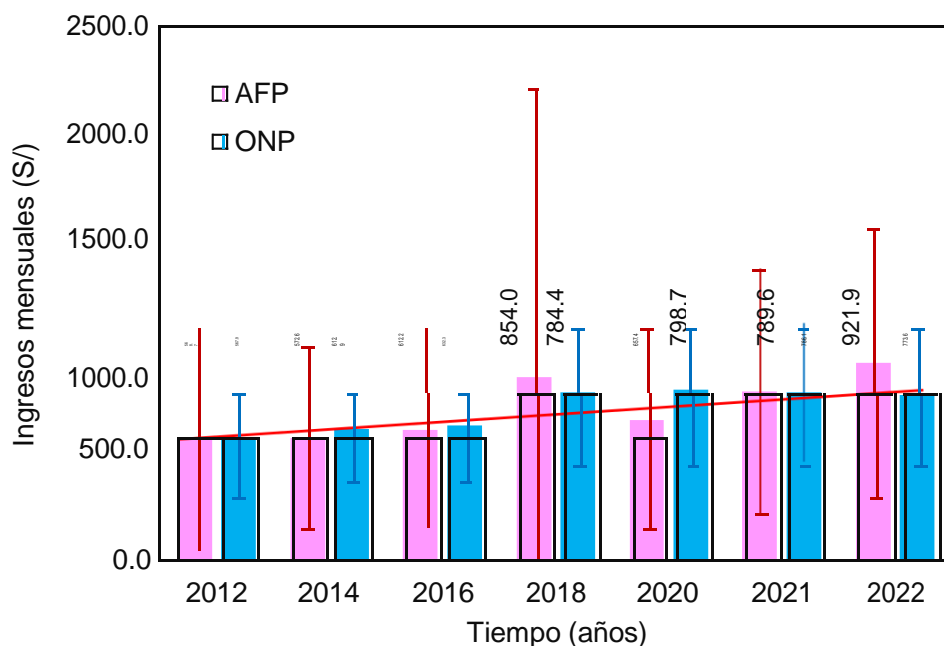
Tabla 4

Ingreso mensual promedio, por fondo previsional, región selva, 2012 al 2022

Periodo	Ingresos (S/)		Participación (%)	
	AFP	ONP	AFP	ONP
2012	568.7	567.9	20.8	79.2
2014	572.6	612.9	41.3	58.7
2016	612.2	632.3	39.8	60.2
2018	854.0	784.4	21.1	78.9
2020	657.4	798.7	21.3	78.7
2021	789.7	786.1	22.9	77.1
2022	921.9	773.6	24.4	75.6

Figura 4

Ingreso mensual promedio, por fondo previsional, región selva, 2012 al 2022



En la tabla 4 y figura 4, se puede observar que para la región selva, existe, asimismo, una tendencia creciente con respecto a los ingresos por concepto de pensión, siendo más notorio en la AFP (valores estuvieron entre S/ 568.7 a S/ 921.9) a comparación de la ONP (valores estuvieron entre S/ 567.9 a S/ 798.7).

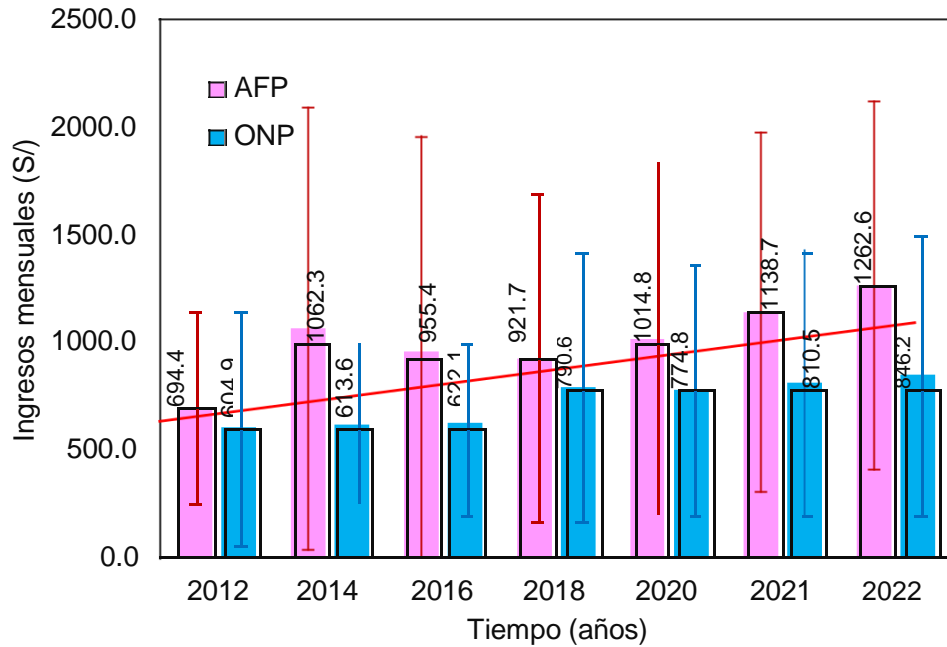
Tabla 5

Ingreso mensual promedio, por fondo previsional, región Lima metropolitana, 2012 al 2022

Periodo	Ingresos (S/)		Participación (%)	
	AFP	ONP	AFP	ONP
2012	694.4	604.9	18.6	81.4
2014	1062.3	613.6	17.1	82.9
2016	955.4	622.1	19.1	80.9
2018	921.7	790.6	16.5	83.5
2020	1014.8	774.8	16.1	83.9
2021	1138.7	810.5	17.5	82.5
2022	1262.6	846.2	18.9	81.1

Figura 5

Ingreso mensual promedio, por fondo previsional, región Lima metropolitana, 2012 al 2022



En la tabla 5 y figura 5, se observa que en la región Lima metropolitana existe una tendencia creciente en ambos sistemas pensionarios, siendo más

notorio en la AFP (valores estuvieron entre S/ 694.4 a S/ 1262.6) a comparación de la ONP (valores estuvieron entre S/ 604.9 a S/ 846.2).

4.2. Coeficiente GINI de equidad de Ingresos por pensión

Tabla 6

Coeficiente de equidad Gini según segmentos, 2012 al 2022

Periodo	Coeficiente de equidad Gini				
	Costa	Sierra	Selva	Lima metropolitana	Total Perú
2012	0.033	0.016	0.001	0.069	0.022
2014	0.027	0.020	0.034	0.268	0.088
2016	0.147	0.091	0.016	0.211	0.115
2018	0.016	0.056	0.042	0.077	0.048
2020	0.016	0.020	0.097	0.134	0.007
2021	0.071	0.028	0.092	0.166	0.055
2022	0.125	0.036	0.087	0.197	0.103

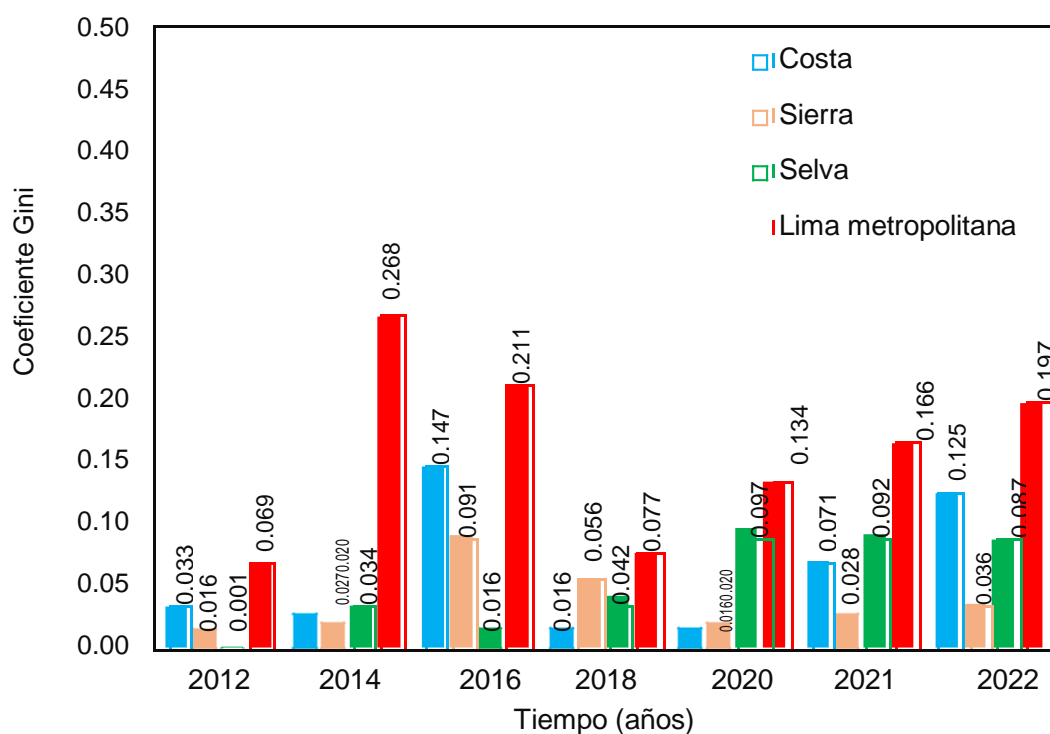
El coeficiente de Gini es una medida de desigualdad de ingresos. El valor de este coeficiente varía entre 0 y 1, donde 0 refleja perfecta igualdad (todos tienen el mismo ingreso) y 1 refleja perfecta desigualdad (una persona tiene todo el ingreso y todos los demás no tienen ingreso).

En la tabla 6 se observa que entre los años 2012 a 2016 se produce un aumento consistente y significativo en la desigualdad entre los pensionistas. En el periodo comprendido entre el año 2016 a 2020 la tendencia se invierte dramáticamente, con una fuerte caída en el coeficiente de Gini. El 2020, la desigualdad entre los pensionistas es casi inexistente, lo que indica una distribución de ingresos entre pensionistas muy equitativa. En el periodo 2020-2022, se observa nuevamente un aumento en la desigualdad, alcanzando niveles cercanos a los de 2016 en el año 2022. Estas fluctuaciones pueden

ser resultado de diversos factores que merecerían un análisis más detallado del contexto socioeconómico, político y regulatorio en esos años.

Figura 6

Coefficiente de equidad Gini según segmentos, 2012 al 2022



En la figura 6 se observa que la desigualdad en Lima Metropolitana es consistentemente más alta que la desigualdad promedio en el país. Esto podría indicar que dentro de la capital hay grandes disparidades entre los pensionistas, posiblemente debido a la heterogeneidad del mercado laboral, las diferencias en oportunidades educativas y laborales, entre otros factores.

La selva, por otro lado, tuvo la desigualdad más baja en 2012, pero ha visto un aumento en los años posteriores, aunque sigue siendo relativamente baja. La costa y sierra han mostrado fluctuaciones, pero en general la sierra ha mantenido un coeficiente Gini más bajo en comparación con la costa, lo que sugiere una distribución más equitativa entre pensionistas en la sierra.

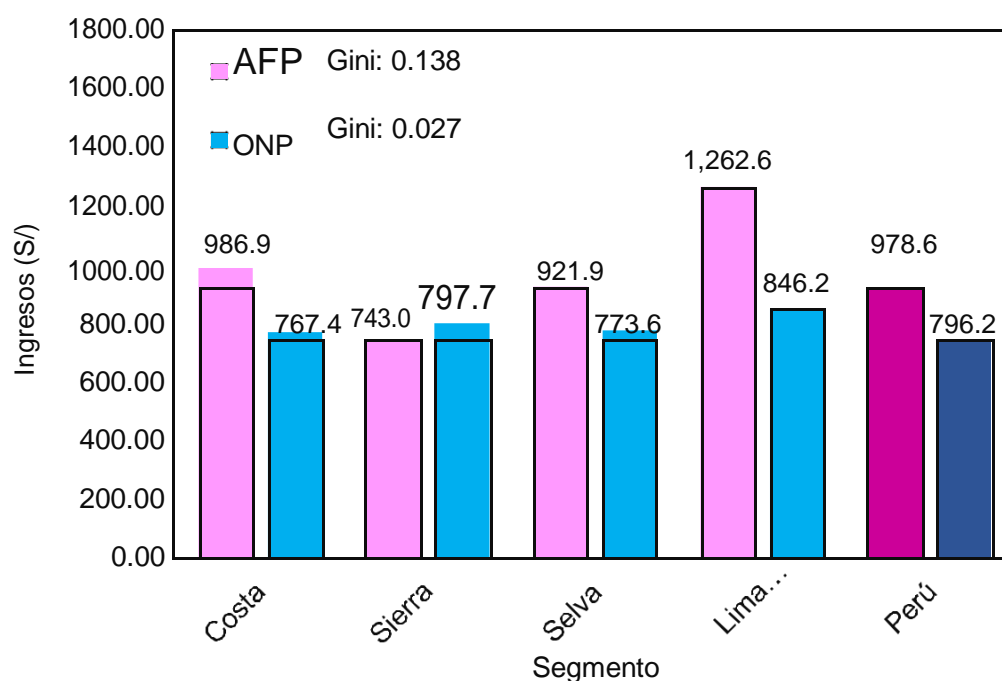
Tabla 7

Ingresos y Coeficiente de equidad Gini según segmentos, 2022

Segmento	Sistema de pensión	%	S/	Coeficiente de equidad Gini
Costa	AFP	11.4	986.9	0.125
	ONP	88.6	767.4	
Sierra	AFP	5.8	743.0	0.035
	ONP	94.2	797.7	
Selva	AFP	24.4	921.9	0.087
	ONP	75.6	773.6	
Lima metropolitana	AFP	18.9	1262.6	0.197
	ONP	81.1	846.2	
Total, Perú	AFP	12.6	978.6	0.103
	ONP	87.4	796.2	

Figura 7

Ingresos promedio por tipo de fondo previsional, según segmentos, 2022



En la tabla 7 y figura 7 se puede observar que, en todas las regiones, la mayoría de los pensionistas están en el sistema público (ONP). Esto refleja la naturaleza del sistema pensional en Perú, donde el sistema público es más

popular y accesible para la mayoría de la población. En cuanto a las pensiones promedio, el sistema privado (AFP) tiende a pagar pensiones más altas en todas las regiones. Sin embargo, hay que destacar que el sistema privado representa a un menor porcentaje de pensionistas en todas las regiones. Esto podría indicar que aquellos en el sistema privado tienen condiciones económicas previas más ventajosas o diferentes condiciones de aporte.

Lima Metropolitana presenta la brecha más grande entre ONP y AFP, con una diferencia de 416.4 soles. Además, es la región con la pensión promedio más alta en ambos sistemas. Esto puede reflejar el costo de vida más alto en la capital y/o la concentración de trabajadores con salarios más altos que se jubilan con pensiones más sustanciales. Sorprendentemente, en la Sierra, el sistema público (ONP) paga una pensión promedio más alta que el sistema privado (AFP). Esta es la única región donde esto ocurre. La razón podría ser la representación demográfica o laboral particular de aquellos que están en el sistema privado en esta región, o puede estar relacionado con las particularidades del mercado laboral en la Sierra.

A nivel nacional, las cifras reflejan lo que se ve regionalmente: la mayoría están en el sistema público, pero las pensiones en el sistema privado son, en promedio, más altas. Esto evidencia desigualdad entre los pensionistas de ONP y AFP en cuanto a las cantidades de pensiones. Esto puede ser una señal de que hay inequidades en el acceso a oportunidades laborales remuneradas, en la capacidad de contribuir a un sistema privado y en las rentabilidades o beneficios del sistema privado.

Finalmente, si se busca ajustar los valores reales de ingresos promedio por pensión en función de su proporcionalidad, se debería ponderar los

ingresos promedio de cada sistema (ONP y AFP) por la proporción de pensionistas en cada sistema, lo que resulta en S/ 820.35 soles.

4.3. Exploración de factores de equidad mediante modelos de regresión de árboles de decisión

Para este análisis se utilizó técnicas de ingeniería de datos sobre la data de la ENAHO, por lo cual se generaron variables dummy a las características: zona geográfica obteniendo nuevas variables de costa, sierra, selva y Lima metropolitana (donde 0 se consideró como ausencia de esa variable y 1 como presencia de esa variable).

Para la característica entidad de trabajo al ser una variable de escala nominal se consideró como 0 = entidad pública y 1 = entidad privada. Para la característica sistema pensionario se consideró 0 = ONP y 1 = AFP. Para la característica sexo se codificó 0 = mujer y 1 = hombre. Para la característica edad se encontró en escala numérica discreta. Y finalmente, para la característica nivel de estudio, una escala ordinal de 10 puntos (1 = educación inicial, 2 = primaria incompleta, 3 = primaria completa, 4 = secundaria incompleta, 5 = secundaria completa, 6 = superior no universitaria incompleta, 7 = superior no universitaria completa, 8 = superior universitaria incompleta, 9 = superior universitaria completa, hasta 10 = maestría/doctorado).

Los modelos de regresión de árboles de decisión son herramientas poderosas en el campo de la estadística y el análisis de datos, son una extensión de los árboles de decisión y generalmente se emplean para predecir valores numéricos en función de variables independientes. Capturan patrones no lineales y complejos en los datos, lo que los hace aptos para problemas en los que las relaciones entre variables son más intrincadas.

Los árboles de regresión fueron propuestos por Leo Breiman en el libro *Classification and Regression Trees* (Breiman y otros, 1984) y son una técnica de aprendizaje supervisado que predice valores de respuestas mediante el aprendizaje de reglas de decisión derivadas de características.

Por otro lado, los algoritmos de gradient boosting son técnicas avanzadas de aprendizaje automático que se utilizan para construir modelos predictivos precisos y complejos, donde el objetivo es predecir valores numéricos o asignar etiquetas a diferentes categorías. La idea fundamental detrás de estos algoritmos es combinar múltiples modelos más simples, generalmente árboles de decisión, en un modelo más fuerte y preciso. El proceso se realiza en etapas iterativas, donde cada nuevo modelo se construye para corregir los errores del modelo anterior. Los modelos XGBoost, LightGBM y CatBoost son tres implementaciones avanzadas de algoritmos de gradient boosting que se han vuelto muy populares en la comunidad de ciencia de datos y análisis de datos.

XGBoost (Extreme Gradient Boosting) es una librería optimizada y eficiente para gradient boosting, combina múltiples modelos más simples, generalmente árboles de decisión, y una de sus características distintivas es su capacidad para manejar tanto problemas de regresión como de clasificación. Utiliza una función de pérdida personalizable y utiliza técnicas de regularización para controlar el sobreajuste, así como manejar valores faltantes en los datos de manera efectiva.

LightGBM (Light Gradient Boosting Machine), es otra librería de gradient boosting que se centra en la velocidad y la eficiencia. Utiliza técnicas como "binning" (agrupación) y "leaf-wise" (crecimiento de hojas) para reducir

el tiempo de entrenamiento y mejorar el uso de la memoria. Además de regresión y clasificación, también puede manejar tareas como ranking y predicción de tiempo de supervivencia.

CatBoost (Categorical Boosting), es un algoritmo de gradient boosting que se destaca por su capacidad para manejar características categóricas de manera eficiente sin la necesidad de convertirlas en números. Esto es especialmente útil en conjuntos de datos que contienen muchas variables categóricas, implementando regularización y ajustando automáticamente los hiperparámetros.

Tabla 8

Bondad de ajuste de modelos de regresión de árboles de decisión

<i>Árboles de decisión potenciados por gradientes</i>	<i>Modelos</i>	<i>R²</i>
Extreme Gradient Boosting	XGBoost	0.897
Light gradient-boosting machine	LightGBM	0.470
Categorical Boosting	CatBoost	0.759

R cuadrado (R²), es una métrica estadística utilizada en el análisis de regresión para evaluar cuánta variabilidad en la variable dependiente (la que estamos tratando de predecir) puede ser explicada por las variables independientes (las que se utilizan para hacer la predicción), vale decir; mide qué tan bien el modelo de regresión se ajusta a los datos observados. Tiene valores que van desde 0 hasta 1, donde R² = 0 significa que el modelo no puede explicar nada de la variabilidad de los datos y todas las predicciones son equivalentes al valor promedio de la variable dependiente, y R² = 1 significa que el modelo explica toda la variabilidad de los datos y las predicciones coinciden perfectamente con los valores reales.

Consecuentemente, un R^2 más alto indica que el modelo es capaz de capturar y explicar una mayor cantidad de la variación en los datos observados.

En la tabla 8, se presenta la bondad de ajuste de modelos de regresión de árboles de decisión potenciados por gradientes para la evaluación de factores en la determinación del nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022, donde se observa que el modelo con mayor coeficiente de determinación fue el Extreme Gradient Boosting (XGBoost) con R^2 de 0.897 (debiendo ser superior a 0.85 para ser considerado un modelo con buena bondad de ajuste), a comparación de LightGBM (0.470) y CatBoost (0.759).

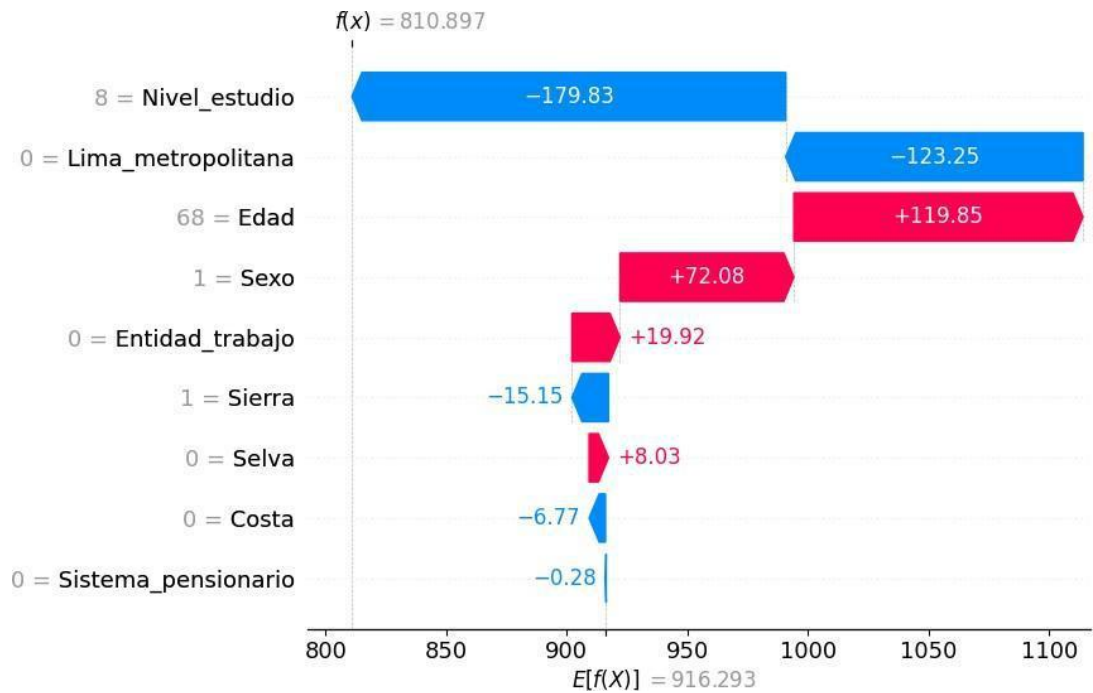
4.4. Determinación de factores de mayor impacto sobre los ingresos mediante modelo XGBoost

Como se ha mencionado, XGBoost, es una aplicación de tipo machine learning. Se sustenta en generar múltiples modelos secuenciales de predicción “sencilla”, y cada nuevo modelo toma el resultado del modelo anterior, haciendo que los resultados obtenidos en cada secuencia sean cada vez más “robustos”, y, por tanto, obteniendo en cada iteración resultados más exactos.

Al aplicar este modelo mediante el software R v4.2.2 (herramienta muy popular en la ciencia de datos, utilizada para el análisis de aprendizaje automático, clustering, segmentación de datos y para extraer conocimientos y patrones de grandes conjuntos de datos), sobre la data publicada en la Encuesta Nacional de Hogares (ENAH) del Instituto Nacional de Estadística e Informática del Perú (INEI), se obtuvieron los siguientes resultados:

Figura 8

Significancia de factores de equidad de ingresos por pensión, en el sistema previsional peruano, 2022



En la Figura 8 se tiene al diagrama de cascada para la evaluación de los factores de mayor significancia en la determinación del nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022; donde se observa que el ingreso promedio por pensiones de jubilación es de S/ 916.293 ($E[f(X)]$); evaluando la importancia, se tiene que el factor de mayor impacto es el nivel de estudio con coeficiente estimado de -179.83 (8 = superior no universitaria completa), esto quiere decir que a menor nivel de estudio los ingresos por pensiones tienden a ser menores (por ejemplo si pasamos de secundaria completa a secundaria incompleta las pensiones disminuyen en promedio S/ 179.83); como segundo factor de importancia está la jubilación en Lima metropolitana (categorizado nominalmente como no jubilado en Lima metropolitana = 0 y jubilado en Lima

metropolitana = 1) con coeficiente estimado de -123.25 (0 = No Lima metropolitana), esto quiere decir que las personas que no se jubilaron en Lima metropolitana tienen en promedio menos S/ 123.25 en sus ingresos por pensiones a comparación de las personas que si se jubilaron en este segmento geográfico; como tercer factor de importancia se encuentra la edad con coeficiente estimado de 119.85 (edad = 68 años), esto quiere decir que por cada año de vida los ingresos por pensiones aumentan en promedio S/ 119.85; como cuarto factor de importancia está el sexo del jubilado (categorizado nominalmente como mujer = 0 y hombre = 1) donde el coeficiente estimado fue de 72.08 (1 = hombre), esto quiere decir que los hombres tienen en promedio S/ 72.08 más de ingresos por jubilación que las mujeres; y como último factor de importancia la entidad de trabajo (categorizado nominalmente como entidad pública = 0 y entidad privada = 1) con coeficiente estimado de 19.92 (0 = entidad pública), esto indica que jubilarse trabajando en una entidad pública mejora los ingresos por pensiones en S/ 19.92, no se tomó los demás factores como importantes debido a los bajos coeficientes estimados (en valor absoluto inferiores a 16).

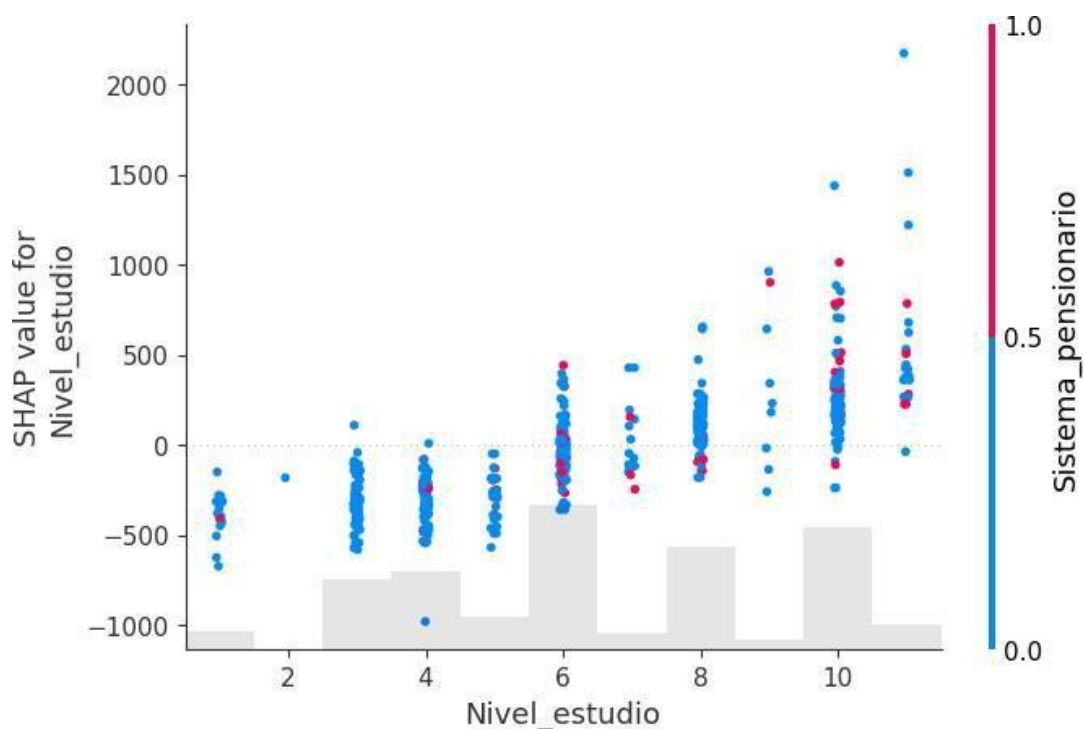
4.5. Factores de mayor impacto sobre los ingresos

En la vida cotidiana, a menudo se hacen planes o 'modelos' para enfrentar desafíos. A veces, estos planes son tan complicados que cuesta entender cómo funcionan. Pero hay formas de entender estos modelos complicados. Uno de los métodos más populares para hacerlo se llama SHAP, que son las iniciales de Shapley Values.

Al usar el modelo XGBoost sobre la data ENAHO del INEI, y asignar Spapley Values a las características y variables de la hipótesis, se obtuvieron los siguientes resultados:

Figura 9

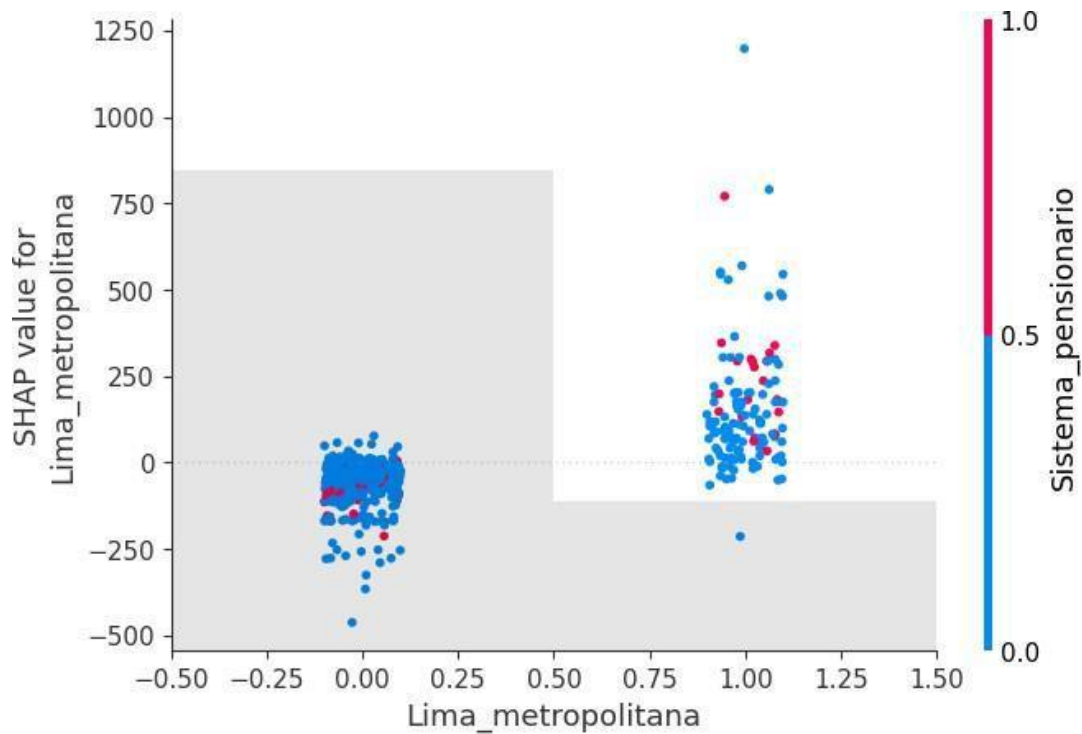
Valor Shap para el factor nivel de estudio respecto a ingresos por pensión, 2022



En la Figura 9 se tiene a la evaluación del factor nivel de estudio con respecto al nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022, donde se observa que al aumentar el nivel de estudios los ingresos tienden a ser mayores (mayores valores SHAP). Además, se observan mayores ingresos en jubilados afiliados a la ONP con mayor nivel de estudio (puntos de color azul) a comparación de los pensionarios en la AFP (puntos rojos).

Figura 10

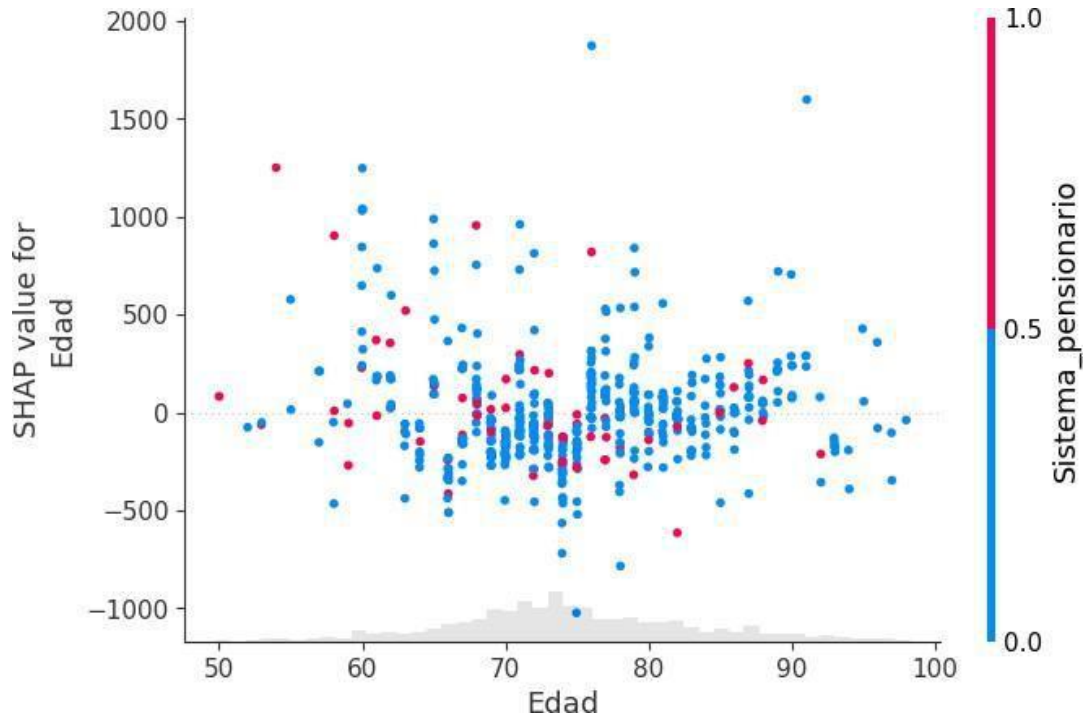
Valor Shap para el factor vive en lima metropolitana respecto a ingresos por pensión, 2022



En la Figura 10 se tiene a la evaluación del factor zona de residencia Lima metropolitana con respecto al nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022, donde se observa que al estar jubilado en Lima metropolitana los ingresos por pensiones son mayores (mayores valores SHAP) (categorizado nominalmente como no jubilado en Lima metropolitana = 0 y jubilado en Lima metropolitana = 1). Además, se observan mayores ingresos en jubilados afiliados a la ONP (puntos de color azul) a comparación de los pensionarios en la AFP (puntos rojos).

Figura 11

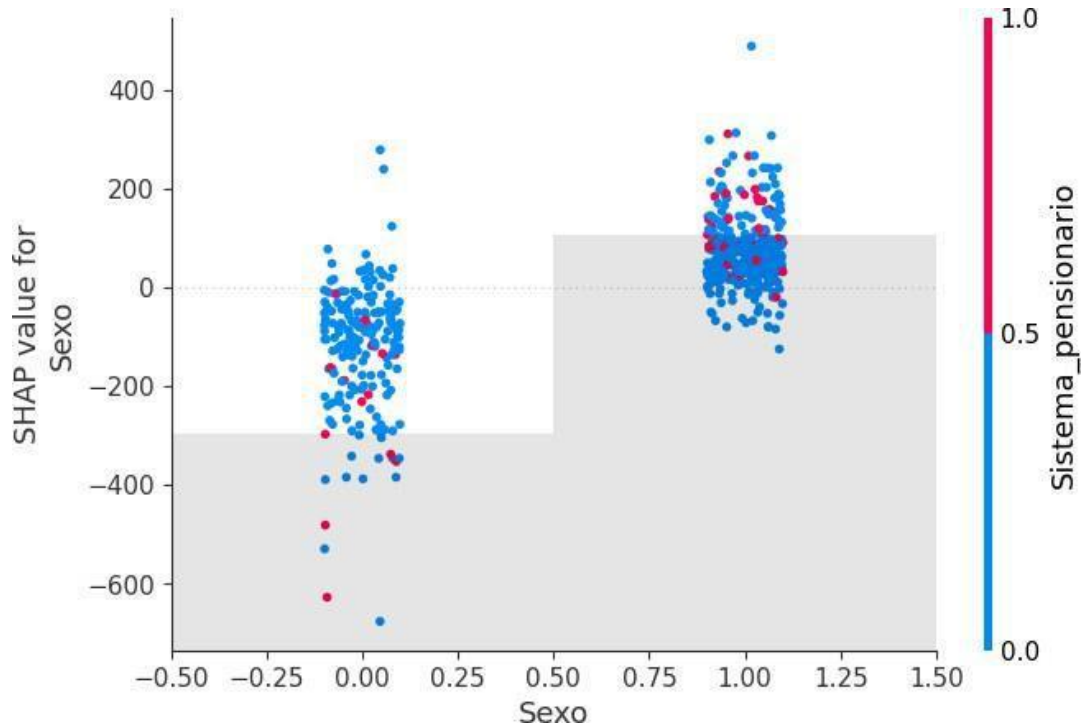
Valor Shap para el factor edad respecto a ingresos por pensión, 2022



En la Figura 11 se tiene a la evaluación del factor edad con respecto al nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022, donde se observa que al aumentar la edad los ingresos tienden a ser ligeramente mayores (mayores valores SHAP). Además, se observan mayores ingresos en jubilados afiliados a la ONP (puntos de color azul) a comparación de los pensionarios en la AFP (puntos rojos).

Figura 12

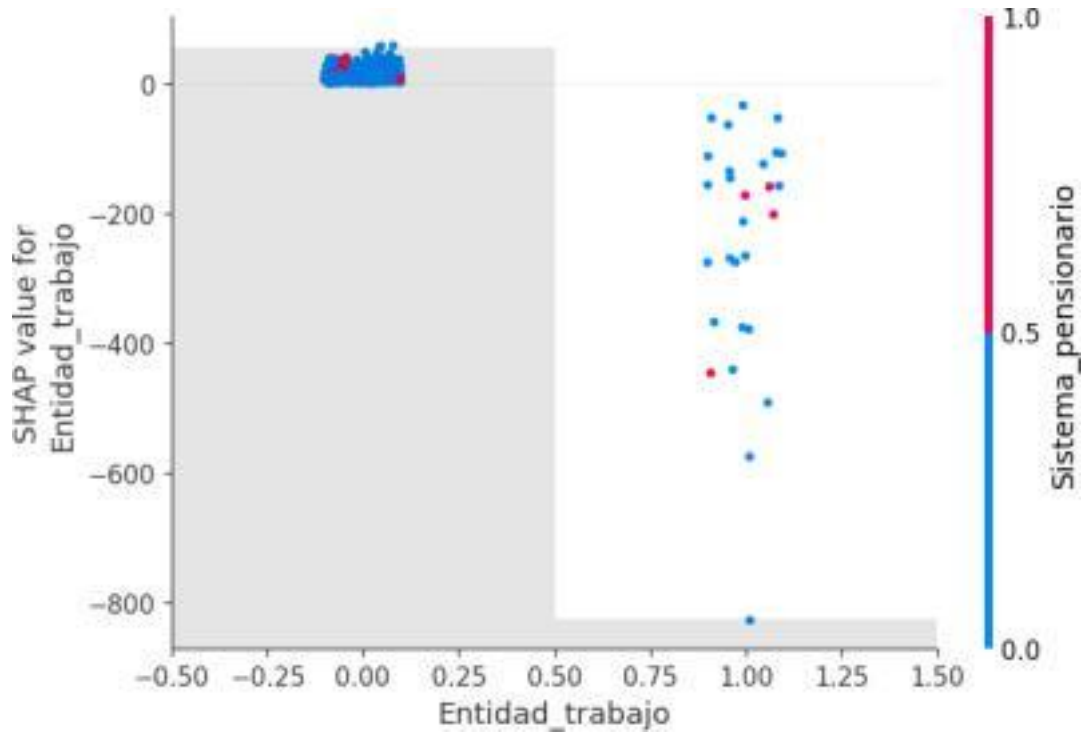
Valor Shap para el factor sexo respecto a ingresos por pensión, 2022



En la Figura 12 se tiene a la evaluación del factor sexo con respecto al nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022, donde se observa que al ser hombre los ingresos por pensiones son mayores (mayores valores SHAP) (categorizado nominalmente como mujer = 0 y hombre = 1). Además, se observan mayores ingresos en jubilados afiliados a la ONP (puntos de color azul) a comparación de los pensionarios en la AFP (puntos rojos).

Figura 13

Valor Shap para el factor entidad en la que trabajó respecto a ingresos por pensión, 2022



En la Figura 13 se tiene a la evaluación del factor entidad de trabajo con respecto al nivel de equidad de ingresos por pensiones otorgados por el sistema previsional peruano para el periodo 2022, donde se observa que al ser jubilado en una entidad pública las pensiones son mayores (mayores valores SHAP) (categorizado nominalmente como entidad pública = 0 y entidad privada = 1). Además, se observan mayores ingresos en jubilados afiliados a la ONP (puntos de color azul) a comparación de los pensionarios en la AFP (puntos rojos).

CAPITULO V – DISCUSIÓN

5.1. Con respecto al objetivo general de investigación

El objetivo general de la investigación, es analizar la equidad en la distribución de ingresos por pensiones en el sistema previsional peruano, identificando los factores clave que influyen las disparidades y evaluar cómo estos factores se correlacionan con las condiciones socioeconómicas y demográficas de los pensionistas, durante el periodo comprendido entre los años 2012 al 2022; lo que implica examinar el sistema de pensiones en Perú con un enfoque en evaluar si los ingresos que reciben las personas que son parte de ese sistema previsional están distribuidos de manera justa y si existe alguna disparidad significativa en estos ingresos.

John Rawls, propone que la justicia social implica que las desigualdades deben ser estructuradas de manera que beneficien a los menos privilegiados y bajo el principio de diferencia de Rawls, cualquier desigualdad económica debe ser justificable solo si beneficia a los menos privilegiados. En este caso, la diferencia en las pensiones entre los sistemas AFP y ONP no parece estar diseñada con ese propósito, puesto que el sistema privado AFP se basa en la capitalización de aportes individuales, mientras que el sistema público ONP se basa en un sistema de reparto, y en teoría, la capitalización debería permitir una mayor acumulación de fondos y, por lo tanto, pensiones más altas, pero esta desigualdad puede ser problemática desde la perspectiva de Rawls, especialmente si los que acceden a la AFP tienden a ser individuos con mayores ingresos.

Por otro lado, se debe tener en cuenta que el Sistema privado, AFP, fue creado en 1993 y aún no ha completado un ciclo de vida laboral completo,

esto puede hacer que la comparación sea algo desigual, ya que no se han percibido completamente los posibles beneficios o desventajas a largo plazo de la AFP; aun así, la diferencia actual en las pensiones debe ser considerada seriamente, y podría argumentarse que refleja una falta de equidad en el sistema. En consecuencia; desde la perspectiva de Rawls, parece haber una inequidad en los ingresos por pensión de jubilación en el sistema previsional peruano, y la desigualdad existente entre los dos sistemas no parece beneficiar principalmente a los menos privilegiados, lo cual estaría en conflicto con los principios de equidad de Rawls; sin embargo, la complejidad del sistema y la falta de un ciclo completo de vida laboral en la AFP pueden requerir un análisis más profundo y continuo para obtener una imagen completa y justa.

Por otro lado, Amartya Sen en su obra teoría de la justicia, pone énfasis en las capacidades y oportunidades reales que tienen las personas, en lugar de centrarse únicamente en los recursos o bienes distribuidos, argumenta que lo que es verdaderamente importante es qué son capaces de hacer y ser las personas, y si tienen la libertad y oportunidad de elegir esas vidas. Si los pensionistas en el sistema ONP (el 87.4%) tienen menores pensiones y por ende menos recursos, esto podría limitar su capacidad para llevar una vida digna y plena en la jubilación, esto puede verse como una clara inequidad en términos de capacidades. Por otro lado, Sen también destaca la importancia de la comparación social en la evaluación de la justicia; la existencia de una brecha significativa en las pensiones entre los dos sistemas puede llevar a una sensación de injusticia y desigualdad entre los diferentes grupos de pensionistas, la comparación directa entre los que están en AFP y ONP podría

resaltar una disparidad en las oportunidades y capacidades, lo que se traduce en una inequidad. Asimismo, Sen se enfoca en el bienestar humano y cómo las personas realmente viven y funcionan, la diferencia en las pensiones no es solo una cuestión de números; afecta la calidad de vida, la salud, la educación, y otros aspectos del bienestar humano, si el sistema ONP no proporciona suficientes recursos para asegurar un nivel aceptable de bienestar, esto es inequitativo. Consecuentemente, desde la perspectiva de Amartya Sen, el sistema previsional peruano parece mostrar inequidad en términos de capacidades, libertad de elección, comparación social, y bienestar humano; la brecha en las pensiones y las limitaciones en la elección y oportunidad reflejan una falta de justicia en cómo las personas son capaces de vivir y funcionar en su jubilación. La teoría de Sen nos invita a mirar más allá de los números y considerar cómo estas diferencias afectan las vidas reales y las libertades de las personas.

La teoría de Martha Nussbaum, similar a la de Sen, se centra en las capacidades, pero va más allá al establecer una lista de capacidades básicas que deben ser garantizadas para todos los individuos para promover la justicia y la igualdad, estas capacidades básicas incluyen cosas como la vida, salud, integridad corporal, control sobre el entorno, y participación social y política. La diferencia en las pensiones entre la AFP y la ONP puede influir directamente en las capacidades básicas de los pensionistas, si la pensión más baja en la ONP limita la capacidad de los pensionistas para acceder a una atención médica adecuada, vivienda, o participación social, esto sería una inequidad en términos de la teoría de Nussbaum. Por otro lado, la participación en la toma de decisiones es una de las capacidades básicas en la teoría de

Nussbaum; si los pensionistas, especialmente aquellos en el sistema ONP, no tienen voz en las políticas y decisiones que afectan sus pensiones, esto podría ser una violación de su capacidad de participación y una forma de inequidad; y si los individuos no tienen control o elección real en la selección de su sistema de pensiones, o si están atrapados en un sistema que no se ajusta a sus necesidades y deseos, esto podría ser visto como una restricción de su capacidad de controlar su entorno, lo cual es inequitativo según la teoría de Nussbaum. Consecuentemente, desde la perspectiva de Martha Nussbaum, el sistema previsional peruano muestra signos de inequidad, ya que las diferencias en las pensiones entre la AFP y la ONP pueden estar en conflicto con varias capacidades básicas que Nussbaum identifica como esenciales para una vida digna; la inequidad en este caso no es solo una cuestión de recursos financieros, sino de cómo esos recursos afectan la capacidad de los individuos para llevar una vida plena y digna, participar en la sociedad, y tener control sobre su entorno.

Aunque las teorías de Rawls, Sen y Nussbaum tienen diferencias importantes en su enfoque y énfasis, comparten preocupaciones comunes sobre la justicia social, el bienestar humano, la atención a los menos privilegiados, y cómo los sistemas y políticas afectan las vidas reales y las oportunidades de las personas.

5.2. Con respecto a los objetivos específicos de investigación

El objetivo subsidiario del presente estudio, es, identificar los factores que son determinantes del nivel de equidad de ingresos por pensiones que tienen los pensionistas del sistema previsional peruano durante el periodo 2012 al 2022.

El análisis de datos realizado en diversas bases de datos (INEI) y publicaciones estadísticas (SBS, ONP, Asociación de AFP) y específicamente las bases de datos resultantes de la Encuesta Nacional de Hogares (ENAHOG), específicamente sobre los montos de pensión y otros aspectos relacionados, se realizó aplicando mecanismos de machine learning supervisado (algoritmo de gradient boosting, XGBoost), mediante el software R, obteniéndose que el factor de mayor impacto en la equidad de los ingresos por pensión de los pensionistas del sistema previsional peruano es el nivel de estudio con coeficiente estimado de -179.83 y correspondiente a la escala ordinal 8 (superior universitaria incompleta).

Estos resultados indican que a menor nivel de estudio los ingresos por pensiones tienden a ser menores. Por ejemplo, los pensionistas con estudios superiores incompletos, tienen ingresos que superan en S/ 179.83 de aquellos que tienen estudios no superiores completos -escala inmediata anterior-. Esto se explica fundamentalmente porque las personas con un nivel de estudios más elevado suelen acceder a trabajos que requieren habilidades especializadas y conocimientos avanzados, por lo que suelen estar mejor remunerados. Dado que las personas con niveles de educación más altos tienen salarios más altos en promedio, sus contribuciones también serán más altas; como resultado, el cálculo de sus pensiones estará basado en un monto de contribuciones más sustancial, lo que podría resultar en una pensión más elevada. Por otro lado, las personas con niveles de educación más altos también pueden tener más oportunidades de ascenso y promoción en sus trabajos; esto puede llevar a un aumento gradual en sus ingresos a lo largo de su carrera, lo que a su vez influye en las contribuciones al sistema de

pensiones y en la base de cálculo para su pensión futura. Finalmente, también se podría decir que los individuos con educación superior a menudo tienen acceso a trabajos más estables y de mayor calidad, lo que puede resultar en carreras laborales más prolongadas y contribuciones al sistema de pensiones durante más tiempo; contribuyendo a acumular un historial de contribuciones más amplio, lo que puede impactar positivamente en el monto de la pensión al momento de jubilarse.

En esta misma corrida de machine learning, aparece como segundo factor de importancia el lugar en el que se desempeñó laboralmente el pensionista (Lima metropolitana, resto del país) siendo que aquellos pensionistas que trabajaron en el resto del país tienen ingresos inferiores en S/ 123.25 que aquellos que trabajaron en Lima metropolitana. Respecto a este factor, se puede señalar que Lima Metropolitana suele tener un costo de vida más elevado en comparación con otras regiones del país; los gastos asociados con vivienda, alimentos, transporte y otros servicios suelen ser más altos en una gran ciudad como Lima, esto significa que los trabajadores que han laborado en Lima Metropolitana podrían haber enfrentado mayores gastos a lo largo de su vida laboral en comparación con aquellos que trabajaron en áreas con un menor costo de vida; sin embargo, las remuneraciones en Lima Metropolitana tienden a ser más altas para compensar el mayor costo de vida. Esto significa que, los pensionistas que trabajaron en Lima podrían experimentar una menor diferencia entre sus ingresos por pensión y los gastos necesarios para mantener su calidad de vida, en comparación con aquellos que trabajaron en áreas con un costo de vida más bajo.

Como tercer factor de importancia se encuentra la edad con coeficiente estimado de 119.85 (68 años), esto quiere decir que por cada año de vida los ingresos por pensiones disminuyen en promedio S/ 119.85. En relación a estos resultados, se debe mencionar que la edad del pensionista puede influir en los ingresos por pensión debido a los gastos médicos adicionales que tienden a aumentar con la edad, los pensionistas mayores podrían enfrentar mayores costos de atención médica y una mayor vulnerabilidad financiera en comparación con los más jóvenes. Estos factores pueden afectar su capacidad para mantener su calidad de vida después de cubrir los gastos médicos y otros costos de vida, lo que podría requerir un ingreso por pensión más alto para garantizar su bienestar en la jubilación.

Como cuarto factor de importancia se encuentra el género del pensionista donde el coeficiente estimado fue de 72.08 (hombre), esto quiere decir que los hombres tienen en promedio S/ 72.08 más de ingresos por pensión que las mujeres. Respecto a esto, podemos mencionar que, en muchos contextos, existe una brecha salarial de género, lo que significa que las mujeres tienden a ganar menos que los hombres por el mismo trabajo o por trabajos de valor similar, esta brecha salarial se refleja en las contribuciones que las mujeres realizan al sistema de pensiones a lo largo de su vida laboral, lo que obviamente puede dar como resultado una base de cálculo más baja para sus pensiones. Por otro lado, las mujeres a menudo enfrentan interrupciones en sus carreras laborales debido a responsabilidades familiares, como el cuidado de los hijos o familiares mayores; estas interrupciones pueden resultar en períodos de ingresos más bajos o en la falta de contribuciones al sistema de pensiones durante ciertos periodos, afectando

directamente el monto de la pensión. Así mismo; las mujeres tienden a tener una mayor esperanza de vida en comparación con los hombres, a medida que las mujeres viven más tiempo en la jubilación, es posible que necesiten una pensión más alta para cubrir los gastos durante un período más prolongado.

El quinto factor en importancia es el tipo de sector al que pertenece la organización en la que trabajó el pensionista resultando el coeficiente estimado de 19.92 (entidad pública), esto indica que jubilarse trabajando en una entidad pública mejora los ingresos por pensiones en S/ 19.92. Se Podría decir que, en algunos casos, el sector público puede ofrecer salarios y beneficios más atractivos en comparación con el sector privado, los empleados del sector público pueden tener acceso a salarios más estables, y mejores prestaciones. El sector público a menudo se caracteriza por una mayor estabilidad laboral en comparación con el sector privado, los empleados del sector público podrían tener carreras laborales más prolongadas y contribuir durante más tiempo al sistema de pensiones; esto podría resultar en un historial de contribuciones más amplio y, por lo tanto, en una base de cálculo más sólida para la pensión.

Los demás factores resultantes en la misma corrida de datos aplicando machine learning no fueron considerados importantes debido a los bajos coeficientes estimados (en valor absoluto inferiores a 16).

5.3. Con respecto a la hipótesis principal de la investigación

En esta investigación se declaró como hipótesis principal, que la equidad en la distribución de ingresos por pensiones en el sistema previsional peruano entre 2012 y 2022 ha sido significativamente influenciada por factores

socioeconómicos y demográficos específicos, resultando en disparidades notables.

Los factores socioeconómicos y demográficos se refieren a características de una población que pueden influir en su comportamiento, necesidades y dinámica. Los factores socioeconómicos considerados por el investigador son: el monto de pensión, tipo de entidad laboral (sector público o privado), zona geográfica en la que reside, y nivel de educación. Los factores demográficos considerados son: edad del pensionista, y género.

La diferencia en los costos de vivienda entre Lima y otras regiones del país puede tener un impacto significativo en la cantidad real de dinero que un pensionista puede gastar; en Lima, los costos pueden ser más altos, y si un pensionista de la AFP (con una pensión promedio más alta) vive en Lima, sus gastos de vivienda podrían consumir una mayor proporción de su pensión que alguien en la ONP viviendo fuera de Lima, esto crea una disparidad en la cantidad de dinero disponible para otros gastos esenciales.

Los gastos del hogar, como servicios públicos y mantenimiento, también varían según la región y el tipo de propiedad en la que vive el pensionista; la capacidad de mantener una calidad de vida aceptable dependerá de cuánto de su pensión se destine a estos gastos, la inequidad puede surgir de diferencias en la eficiencia y costo de los servicios en diferentes regiones, así como de diferencias entre los dos subsistemas de pensiones, con los pensionistas de la AFP posiblemente teniendo más capacidad para cubrir estos costos.

Los costos de atención médica y la calidad de los servicios de salud también pueden variar significativamente en Perú, y esto se refleja en la

equidad entre los pensionistas; los pensionistas de la AFP podrían tener acceso a mejores servicios de salud privados o mayores ahorros para cubrir estos costos, mientras que los de la ONP podrían depender más de los servicios de salud públicos, que pueden variar en calidad y disponibilidad según la región.

Estos diferentes factores contribuyen a la inequidad entre los pensionistas en Perú al crear variaciones en los gastos esenciales en función de la ubicación y del subsistema de pensiones al cual pertenecen; la vivienda, el hogar y la salud son áreas vitales que consumen una proporción significativa de los ingresos de una persona jubilada, y las diferencias en estos gastos se magnifican por las disparidades regionales y las diferencias entre los sistemas de AFP y ONP, la combinación de estos factores puede dar lugar a una experiencia muy diferente de la jubilación según dónde y cómo vive una persona, y qué tipo de pensión recibe. Esto podría ser una llamada a revisar cómo el sistema de pensiones podría ser reformado o ajustado para reducir estas disparidades y asegurar una jubilación más equitativa para todos.

CAPITULO VI – CONCLUSIONES

Podemos concluir en base a lo investigado y discutido en este estudio, lo siguiente:

PRIMERA. -

Se ha constatado que la equidad en la distribución de ingresos por pensiones en el sistema previsional peruano (2012-2022) ha sido significativamente influenciada por factores socioeconómicos y demográficos específicos, lo que ha resultado en disparidades notables. Este hallazgo subraya la complejidad del sistema previsional y la necesidad de considerar múltiples dimensiones en la formulación de políticas.

SEGUNDA. -

Los pensionistas en diferentes zonas geográficas experimentan variaciones significativas en sus ingresos por pensiones. Esta disparidad regional indica la necesidad de políticas que aborden específicamente las desigualdades geográficas para garantizar una distribución más equitativa de los recursos previsionales.

TERCERA. -

La naturaleza de la entidad laboral (pública o privada) donde trabajó el pensionista influye considerablemente en la cuantía de sus pensiones. Este resultado sugiere que las diferencias estructurales entre los sectores público y privado tienen un papel clave en la determinación de los beneficios previsionales.

CUARTA. -

Se observa una correlación significativa entre el tipo de sistema previsional (público o privado) receptor de los aportes y la cantidad de

ingresos que recibe el pensionista. Esto resalta la importancia de evaluar y potencialmente reformar las estructuras y políticas de ambos sistemas para lograr una distribución más equitativa.

QUINTA. -

El género del pensionista es un determinante clave en la disparidad de los ingresos por pensiones. Este hallazgo evidencia la existencia de brechas de género en el sistema previsional y la necesidad de implementar medidas para abordar estas desigualdades.

SEXTA. -

El nivel educativo alcanzado por los pensionistas afecta significativamente sus ingresos por pensiones. Este aspecto subraya la interconexión entre la educación y la seguridad económica en la vejez, sugiriendo la importancia de políticas educativas y laborales inclusivas.

CAPITULO VII – RECOMENDACIONES

En base a las conclusiones de la investigación, estamos en capacidad de realizar las siguientes recomendaciones:

PRIMERA. -

En respuesta a la variabilidad geográfica en los ingresos por pensiones, se recomienda al Poder Ejecutivo la creación de políticas que atiendan las necesidades específicas de cada región, considerando factores como el costo de vida y el acceso a servicios básicos.

SEGUNDA. -

Se recomienda al Poder Ejecutivo, implementar programas que aborden las disparidades de género en el sistema previsional, incluyendo incentivos para la participación equitativa de mujeres en el mercado laboral y ajustes en los cálculos de pensiones que compensen las interrupciones de carrera típicamente experimentadas por las mujeres.

TERCERA. -

Se recomienda al Poder Legislativo, proponer y aprobar reformas legislativas que garanticen una mayor equidad entre los trabajadores de los sectores público y privado en términos de beneficios previsionales.

CUARTA. -

Se recomienda al Poder Legislativo, establecer normativas que aseguren la transparencia en la gestión de los fondos previsionales y promuevan una distribución más equitativa de los mismos.

QUINTA. -

Se recomienda al Poder Ejecutivo y al Poder Legislativo, implementar programas educativos que mejoren la comprensión de los ciudadanos sobre

los sistemas previsionales, con especial énfasis en la planificación financiera para la jubilación.

SEXTA. -

Se recomienda al Poder Ejecutivo y al Poder Legislativo, establecer mecanismos para el seguimiento y evaluación periódica del sistema previsional, permitiendo ajustes basados en evidencia y en respuesta a cambios demográficos y socioeconómicos.

SETIMA. -

Se recomienda al Poder Judicial, asegurar que las leyes y regulaciones relacionadas con el sistema previsional se apliquen de manera justa y equitativa, proporcionando un mecanismo de revisión para casos de discriminación o injusticia en la asignación de pensiones.

FUENTES DE INFORMACIÓN

- Alonso, J., Sanchez, R., & Tuesta, D. (2014). Un modelo para el sistema de pensiones en el Perú: Diagnóstico y recomendaciones. *Estudios Económicos*(27), 81-98.
- Arrieta, A., Jaramillo, M., Prieto, L., Seinfeld, J., De la Torre, A., & Tuesta, D. (2017). *Propuestas de reformas en el sistema de pensiones, financiamiento en la salud y seguro de desempleo*. Lima: Comisión de Protección Social.
- Asociación de AFP. (13 de Setiembre de 2022). *Asociación de AFP: Estadísticas*.
<http://estadisticas.asociacionafp.pe/ActiveAffiliatesAccordingAfp>
- Banco Mundial. (13 de Setiembre de 2022). *Banco Mundial Datos*.
<https://datos.bancomundial.org/indicador/SP.DYN.LE00.FE.IN?view=chart>
- Bernal, N. (2016). Los gastos públicos en pensiones en América Latina y sus proyecciones al año 2075: evidencia de Chile, Perú, Colombia y México. *Apuntes, XLIII*(79), 79-128.
<https://doi.org/http://dx.doi.org/10.21678/apuntes.79.867>
- Breiman, L., Friedman, R., Olshen, A., & Stone, J. (1984). *Classification and Regression Trees*. Hall/CRC.
- Brower Beltramin, J. (2020). La Justicia como equidad. Una reformulación. *Lus et Praxis*, 597-600.
- Cardona, D., & Peláez, E. (2012). Envejecimiento poblacional en el siglo XXI: oportunidades, retos y preocupaciones. *Salud Uninorte*, 335-348.

- Cenas Rodriguez, S. (2019). *Desigualdad del ingreso en el Perú: 1997-2017*.
Universidad Nacional de Trujillo.
- Chavez Jorge, D. (2018). *Gasto público y desigualdad de ingreso: Perú, 1997-2017*. Lima: Universidad César Vallejo.
- Collier, A. (2016). *Data Management for Researchers*.
- Congreso de la República del Perú. (5 de Mayo de 2016). *LEY Nº 30478*.
<https://busquedas.elperuano.pe/normaslegales/ley-que-modifica-el-articulo-40-y-la-vigesimo-cuarta-disposi-ley-n-30478-1398360-6/>
- Couture de Troismonts, R. (1975). *Manual de técnicas de documentación*.
Marymar.
- Esping-Andersen, G. (1990). *Los tres mundos del estado del bienestar*.
Edicions Alfons el Magnanim. <https://doi.org/84-7822-097-6>
- Espinoza, E. (2018). La hipótesis en la investigación. *Mendive*, 122-139.
<https://doi.org/1815-7696>
- Fabrigar, L. (1999). Evaluating the use of Exploratory Factor Analysis in
Psychologic Research. *Psychological Methods*, 272–299.
- Fondo de Población de las Naciones Unidas. (13 de Setiembre de 2022).
Envejecimiento de la población.
<https://www.unfpa.org/es/envejecimiento-de-la-poblaci%C3%B3n>
- Freire, P. (2005). *Pedagogía del oprimido*. Siglo XXI Editores S.A.
<https://doi.org/968-23-2589-7>
- Gestión, R. (8 de Junio de 2020). Perú será el país con mayor caída de PBI
en Sudamérica este año, según BM. *Gestión*, pág. 4.
- Gómez Pérez. (2015). *Introducción a la metodología de la investigación científica*.

- Hernandez Sampieri, Fernandez Collado, & Baptista Lucio. (2021). *Metodología de la Investigación*.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la Investigación*. McGraw-Hill. <https://doi.org/978-1-4562-2396-0>
- HoyosGómez, D. (2008). Elementos para una teoría de justicia: una comparación entre John Rawls y Amartya Sen. *Desafíos*, 156-181.
- Instituto Nacional de Estadística e Informática. (2017). *Estadísticas*. Encuesta Nacional de Hogares: <https://www.inei.gob.pe/>
- Instituto Nacional de Estadística e Informática. (3 de noviembre de 2022). *ENAHO, Condiciones de vida en el Perú*. <https://m.inei.gob.pe/biblioteca-virtual/boletines/condiciones-de-vida/1/#lista>
- Instituto Nacional de Estadística e Informática. (3 de noviembre de 2022). *Ficha Técnica de la Encuesta Nacional de Hogares 2017*. <http://inei.inei.gob.pe/inei/srienaho/Descarga/DocumentosMetodologicos/2017-55/FichaTecnica.pdf>
- MARKA Investigación. (3 de noviembre de 2022). *Blog de MARKA Investigación*. <https://markainvestigacion.wordpress.com/>
- Marshall, T., & Bottomore, T. (2023). *Ciudadanía y clase social*. Alianza Editorial. <https://doi.org/9788411481809>
- McKinney, W. (2017). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media.

- Mesa-Lago, C. (2022). Desempeño de pensiones privatizadas en América Latina, 1980-2020: el caso de México. *El Trimestre Económico*, LXXXIX(355), 755-794. <https://doi.org/10.20430/ete.v89i355.1477>
- Ministerio de Economía y Finanzas. (2004). *Los Sistemas de Pensiones en Perú*. Lima: Dirección General de Asuntos Económicos y Sociales.
- Murphy, T. (2016). Dos teorías sobre la justicia social. *Economía*, 78-89. Nunnally, J., & Bernstein, I. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Nussbaum, M. (1992). Human Functioning and Social Justice: In Defense of Aristotelian Essentialism. *Political Theory*, 20(2), 202-246.
- Nussbaum, M. (2012). *Crear capacidades: Propuesta para el desarrollo humano*. Espasa Libros. <https://doi.org/978-84-493-0988-5>
- OCDE. (2021). *Pensions at a Glance 2021: OECD and G20 Indicators*. París: OCDE.
- Oficina de Normalización Previsional. (13 de Setiembre de 2022). *Sistema previsional y sus actores*. https://www.onp.gob.pe/pensiones_peru_onp/sistema_previsional
- Organismo de Estados Americanos. (13 de Setiembre de 2022). *Tratados multilaterales*. https://www.oas.org/es/sla/ddi/tratados_multilaterales_interamericanos_A-70_derechos_humanos_personas_mayores.asp
- Organización Mundial de la Salud. (9 de Noviembre de 2022). *Envejecimiento y salud*. <https://www.who.int/es>
- Otero-Bahamón, S., Álvarez-Altamiranda, L., Sampayo, A., & Alvis, J. (2022). *Trayectorias divergentes de la desigualdad en las ciudades*

intermedias: el rol de las pensiones y las ayudas en la reducción de la desigualdad de ingresos. Universidad del Rosario, Bogotá, Colombia.
<https://doi.org/10.18046/recs.iEspecial.4932>

Palomar, J. (9 de Noviembre de 2022). *Etimologías Filosóficas*. Etimologías Filosóficas:

<http://etimologiaspalomar.blogspot.com/2011/10/equidad.html>

Peterson, R. (2000). Meta-Analysis of Variance Accounted for and Factor Loadings in Exploratory Factor Analysis. *Marketing Letters*, 261-275.

Quinn Patton, M. (2015). *Qualitative Research & Evaluation Methods*.

Ramirez, T. (2010). *Como hacer un proyecto de investigacion*. Caracas: Editorial Panapo. <https://doi.org/980-733903-2>

Rawls, J. (1999). Justicia como equidad. *Revista española de control*, 129-158.

Rawls, J. (2012). *Teoría de la Justicia*. México D.F.: FCE - Fondo de Cultura Económica. <https://doi.org/607-16-0903-8>

Rodríguez López, M. (2018). *Diseño y validación de un índice focalizado y compuesto basado en el índice de concentración para la medición de inequidad en salud*. Bogotá: Universidad Nacional de Colombia.

Rodriguez Weber, J. (2014). *La Economía Política de la Desigualdad de ingreso en Chile, 1850-2009*. Montevideo, Uruguay: Universidad de la República.

Rodríguez, M. (2015). Reseña de libro la Idea de Justicia de Amartya Sen. *Recerca : revista de pensament i anàlisi*, 17(7), 129-147.
<https://doi.org/10.6035>

- Rosas Febres, M. (2018). *Inequidad dentro de la equidad, pagos informales durante la atención de los afiliados al seguro integral de salud 2008 a 2010*. Universidad Peruana Cayetano Heredia.
- Sen, A. (1999). *Desarrollo y libertad*. 2000, Editorial Planeta S.A.
<https://doi.org/950-49-0473-4>
- Silva, Y. (2017). *Igualdad, equidad, derecho y justicia en el régimen pensional colombiano*. Bogotá: Universidad Católica de Colombia.
- Smith, J., Johnson, A., & Rodriguez, M. (2019). Comparative analysis of programming languages for efficient data collection from structured databases. *Data Science*, 10(3), 345-362.
- Social, C. d. (2017). *Propuestas de reforma en el sistema de pensiones, financiamiento en la salud y seguro de desempleo*. Lima: CPS.
- Stilman, G. (5 de Octubre de 2022). *Justicia de justicias: Etica universal y progreso moral en un mundo plural*.
https://www.academia.edu/11274307/Justicia_de_justicias_Etica_universal_y_mejoramiento_moral_en_un_mundo_plural
- Superintendencia de Banca, Seguros y AFP. (13 de Setiembre de 2022). *Comisiones y Primas de Seguro del SPP*.
https://www.sbs.gob.pe/app/spp/empleadores/comisiones_spp/paginas/comision_prima.aspx
- Superintendencia de Banca, Seguros y AFP. (14 de Setiembre de 2022). *SBS para ciudadanos*. <https://www.sbs.gob.pe/usuarios/informacion-de-pensiones/otros-beneficios-del-spp/entrega-de-hasta-el-955-del-fondo-de-pensiones>

UCLA, I. f. (16 de julio de 2020). *A practical introduction to factor analysis: Exploratory factor analysis.*

<https://stats.idre.ucla.edu/spss/seminars/introduction-to-factor-analysis/a-practical-introduction-to-factor-analysis/>

Unidas, N. (2020). *El impacto del COVID-19 en América Latina y el Caribe.* New York: Naciones Unidas.

Urbano-Guzmán, M. (2014). El concepto de igualdad en algunas teorías contemporáneas de la justicia. *Criterio libre jurídico*, 123-139.

Varea, J. (15 de Junio de 2020). ¿Cómo afectará la crisis de la Covid-19 a nuestro sistema de pensiones? *La Vanguardia*, pág. 7.

Vickery, B. (1970). *Techniques of information retrieval.* Londres: Butterworths.

Vidal, P. F. (2009). *La teoría de la justicia social en Rawls.* Universidad Federal de Rio de Janeiro.

Yamada, G., Castro, J., & Oviedo, N. (2016). *Revisitando el coeficiente de Gini en el Perú: El rol de las políticas públicas en la evolución de la desigualdad.* Lima: Universidad del Pacífico.

Yin, R. (2014). *Case Study Research, Design and Methods.*

ANEXOS

ANEXO 1- MATRIZ DE OPERACIONALIZACIÓN DE VARIABLES

Variable	Tipo de Variable	Definición Conceptual	Definición Operacional	Indicadores
Equidad de ingresos por pensiones	Dependiente	Grado de justicia o imparcialidad en la distribución de ingresos por pensiones entre los pensionistas.	Evaluación cualitativa y cuantitativa de la distribución.	<ul style="list-style-type: none"> ▪ Índice de Gini ▪ Coeficiente de variación
Ingresos por pensión	Independiente	Monto de la pensión recibida por los pensionistas.	Monto mensual de la pensión en soles.	<ul style="list-style-type: none"> ▪ Monto de pensión mensual
Zona geográfica	Independiente	Ubicación geográfica del pensionista.	Clasificación por región: costa, sierra, selva, Lima metropolitana.	<ul style="list-style-type: none"> ▪ Región de residencia
Tipo de entidad en la que trabajó	Independiente	Tipo de organización donde trabajó el pensionista.	Categorización: entidad pública, empresa privada.	<ul style="list-style-type: none"> ▪ Tipo de entidad laboral
Sistema previsional al que pertenece el pensionista	Independiente	Sistema de pensiones al que pertenece el pensionista.	Categorización: Sistema público (ONP), Sistema privado (AFP).	<ul style="list-style-type: none"> ▪ Tipo de sistema previsional
Sexo	Independiente	Género del pensionista.	Clasificación: masculino, femenino.	<ul style="list-style-type: none"> ▪ Género

Nivel de estudios alcanzado	Independiente	Grado académico o nivel de educación del pensionista.	Clasificación: educación inicial, primaria incompleta, primaria completa, secundaria incompleta, secundaria completa, superior no universitaria incompleta, superior no universitaria completa, superior universitaria incompleta, superior universitaria completa, maestría/doctorado.	▪ Nivel de educación
-----------------------------	---------------	---	---	----------------------

ANEXO 2- MATRIZ DE CONSISTENCIA

Problema de Investigación	Objetivo General	Objetivos Específicos	Hipótesis	Variables	Metodología
¿De qué manera la distribución de ingresos por pensiones en el sistema previsional peruano ha sido equitativa o inequitativa durante el periodo de 2012 a 2022, y qué factores clave, incluyendo condiciones socioeconómicas y demográficas, han influido en estas disparidades?	Analizar la equidad de la distribución de ingresos por pensiones en el sistema previsional peruano, identificando los factores clave que influyen en las disparidades y evaluar cómo estos factores se correlacionan con las condiciones socioeconómicas y demográficas de los pensionistas, durante el periodo comprendido entre los años 2012 al 2022.	<ol style="list-style-type: none"> Analizar el impacto de la zona geográfica en la que vive el pensionista para determinar cómo este factor contribuye a las diferencias en sus ingresos. Analizar el impacto del tipo de entidad laboral (público o privado) en la que trabajó el pensionista, para determinar cómo este factor contribuye a las diferencias en sus ingresos. Analizar el impacto del Sistema Previsional (público 	<ol style="list-style-type: none"> La equidad en la distribución de ingresos por pensiones en el sistema previsional peruano entre 2012 y 2022 ha sido significativamente influenciada por factores socioeconómicos y demográficos específicos, resultando en disparidades notables. Los pensionistas que residen en ciertas zonas geográficas reciben ingresos significativamente diferentes, contribuyendo a la disparidad en la distribución de ingresos. El tipo de entidad laboral en la que trabajó el pensionista tiene un 	<p>Dependiente:</p> <ul style="list-style-type: none"> ▪ Equidad de ingresos por pensiones <p>Independientes:</p> <ul style="list-style-type: none"> ▪ Ingresos por pensión ▪ Zona geográfica ▪ Tipo de entidad en la que trabajó ▪ Sistema previsional al que pertenece el pensionista ▪ Sexo ▪ Nivel de estudios alcanzado 	<ul style="list-style-type: none"> ▪ Enfoque Cuantitativo, ▪ Diseño no experimental, transversal, ▪ Descriptivo y analítico ▪ Uso de datos secundarios, ENAHO ▪ Uso la técnica de extracción de datos mediante programación ▪ Uso de minería de datos para el procesamiento y análisis de datos.

<p>o privado) receptor primario de los aportes del pensionista, para determinar cómo este factor contribuye a las diferencias en sus ingresos.</p>	<p>impacto considerable en la cantidad de sus ingresos.</p>	<p>Monto de pensión</p>
<p>4. Analizar el impacto del género del pensionista, para determinar cómo este factor contribuye a las diferencias en sus ingresos.</p>	<p>4. Existe una correlación significativa entre el sistema previsional receptor de los aportes y la cantidad de ingresos que recibe el pensionista.</p>	
<p>5. Analizar el impacto del nivel de estudios alcanzado por el pensionista, para determinar cómo este factor contribuye a las diferencias en sus ingresos.</p>	<p>5. El género del pensionista es un factor determinante en la disparidad de los ingresos en el sistema previsional peruano.</p>	
	<p>6. El nivel de educación alcanzado por los pensionistas afecta de manera significativa sus ingresos.</p>	

ANEXO 3- FICHA TÉCNICA, ENCUESTA NACIONAL DE HOGARES 2022

La Encuesta Nacional de Hogares (ENAH) es la investigación que permite al Instituto Nacional de Estadística e Informática (INEI) desde el año 1995, efectuar el seguimiento de los indicadores sobre las condiciones de vida.

A inicios del año 2007, con el auspicio del Banco Mundial (BM), el INEI convocó a un Comité Asesor especializado conformado por representantes de organismos internacionales, de organismos gubernamentales nacionales, representantes de la comunidad académica y de centros de investigación. El Comité Asesor de Pobreza ha venido participando cada año en verificar la calidad de la encuesta y la medición de la pobreza.

A partir del año 2010, mediante Resolución Suprema Nº 097-2010-PCM, publicado en el diario oficial El Peruano el 13 de abril del 2010, el Comité Asesor de Pobreza se constituye en Comisión Consultiva para Estimación de la Pobreza y otros indicadores relacionados en el país; cuyo objetivo es garantizar la calidad, transparencia y confianza de la información en el campo de la medición de la pobreza y otros indicadores relacionados.

Debido al estado de Emergencia Nacional por las graves circunstancias que afectan la vida de la Nación como consecuencia del brote del COVID-19, iniciado el 16 de marzo del año 2020, ampliado por el gobierno el año 2021 y en el presente año 2022 debido a que continuamos en emergencia sanitaria, el Instituto Nacional de Estadística e Informática (INEI), continúa ejecutando la Encuesta Nacional de Hogares en la modalidad de entrevista mixta, es decir, se inicia la entrevista de manera presencial (cara a cara), donde se recaba información de las variables principales de la investigación (características de la vivienda, hogar y demográficas, así como los gastos del hogar), asimismo solicitará los números telefónicos de las personas de 14 y más años de edad y continuará el recojo de información mediante entrevista telefónica, con la finalidad de disminuir el tiempo de exposición del personal de campo ante la emergencia sanitaria.

I. OBJETIVOS

- Generar indicadores que permitan conocer la evolución de la pobreza, del bienestar y de las condiciones de vida de los hogares.
- Efectuar diagnósticos sobre las condiciones de vida y pobreza de la población.
- Medir el alcance de los programas sociales alimentarios y no alimentarios en la mejora de las condiciones de vida de la población.
- Servir de fuente de información a instituciones públicas y privadas, así como a investigadores.
- Permitir la comparabilidad con investigaciones similares, en relación con las variables investigadas.

II. COBERTURA

La encuesta se realiza en el ámbito nacional, en el área urbana y rural, en los 24 departamentos del país y en la Provincia Constitucional del Callao.

III. PERIODO DE EJECUCIÓN

En forma continua, iniciándose a partir de mayo del 2003.

IV. PERIODO DE REFERENCIA

Los períodos de referencia para las variables a investigarse en la Encuesta Nacional de Hogares 2021, son los siguientes:

- Características de la vivienda y del hogar:
 - Día de la entrevista
 - Últimos 12 meses
 - Último gasto mensual
- Características de los miembros del hogar:
 - Día de la entrevista
 - Residencia habitual: Últimos 30 días
 - Trabajo infantil: Semana pasada
- Educación
 - Día de la entrevista
 - Presente año
 - Últimos 12 meses
 - Mes anterior
- Salud
 - Día de la entrevista
 - Últimas 4 semanas
 - Últimos 3 meses
 - Últimos 12 meses
- Empleo
 - Condición de actividad: Semana pasada
 - Sistema de pensiones: Día de la entrevista
- Inclusión Financiera
 - Día de la entrevista
 - Últimos 12 meses
- Ingreso por trabajo:
 - Ordinarios: Día, semana, quincena, mes anterior
 - En especie: Últimos 12 meses
 - Extraordinarios por trabajo dependiente: Últimos 12 meses
- Trabajo anterior para los desocupados: Último trabajo realizado
- Ingresos por transferencias corrientes: Últimos 6 meses
- Ingresos por rentas de la propiedad: Últimos 12 meses
- Otros ingresos extraordinarios: Últimos 12 meses
- Ingreso por actividad agropecuaria: Últimos 12 meses
- Ingresos del trabajador independiente o empleador o patrono: Mes anterior, últimos 12 meses
- Gastos:
 - Últimos 15 días
 - Mes anterior
 - Últimos 3 meses
 - Últimos 12 meses
 - Día de la entrevista
 - Semana anterior
- Programas sociales de ayuda alimentaria:
 - Últimos 3 meses

- Programas sociales no alimentarios:
 - Últimos 3 años
- Participación ciudadana:
 - Día de la entrevista
- Gobernabilidad, democracia y transparencia
 - Día de la entrevista
 - Últimos 12 meses
- Percepción del hogar
 - Día de la entrevista
 - Últimos 12 meses
- Discriminación
 - Últimos 12 meses
- Acceso a la justicia
 - Últimos 12 meses

V. POBLACIÓN OBJETIVO

La población de estudio está definida como el conjunto de todas las viviendas particulares y sus ocupantes residentes en el área urbana y rural del país.

Por no ser parte de la población de estudio, se excluye a los miembros de las fuerzas armadas que viven en cuarteles, campamentos, barcos, y otros. También se excluye a las personas que residen en viviendas colectivas (hoteles, hospitales, asilos y claustros religiosos, cárceles, etc.).

VI. DISEÑO Y MARCO MUESTRAL

Marco muestral:

El marco muestral para la selección de la muestra lo constituye la información estadística proveniente de los Censos de Población y Vivienda y material cartográfico actualizado para tal fin.

Unidades de Muestreo:

En el Área Urbana

- La Unidad Primaria de Muestreo (UPM) es el centro poblado urbano con 2 mil y más habitantes.
- La Unidad Secundaria de Muestreo (USM) es el conglomerado que tiene en promedio 120 viviendas particulares.
- La Unidad Terciaria de Muestreo (UTM) es la vivienda particular.

En el Área Rural

- La Unidad Primaria de Muestreo (UPM) es de dos tipos:
 - El centro poblado urbano con 500 a menos de 2 mil habitantes.
 - El Área de Empadronamiento Rural (AER) el cual tiene en promedio 100 viviendas particulares.
- La Unidad Secundaria de Muestreo (USM) es de dos tipos:
 - El conglomerado que tiene en promedio 120 viviendas particulares.
 - La vivienda particular
- La Unidad Terciaria de Muestreo (UTM) es la vivienda particular.

Tipo de muestra:

La muestra es del tipo probabilística, de áreas, estratificada, multietápica e independiente en cada departamento de estudio.

A fin de medir los cambios en el comportamiento de algunas características de la población, desde el año 2008 se viene implementando muestras panel de viviendas, en la cual viviendas encuestadas son nuevamente investigadas cada año.

En la muestra no panel se visitan cada año los mismos conglomerados en el mismo mes de encuesta, pero se seleccionan distintas viviendas.

El nivel de confianza de los resultados muestrales, es del 95%.

Tamaño de la muestra:

El tamaño anual de la muestra 2022 es de 36 822 viviendas particulares, correspondiendo 24 206 viviendas al área urbana y 12 616 viviendas al área rural.

Asimismo, el tamaño de la muestra panel es de 12 194 viviendas particulares mientras que el tamaño de la muestra no panel es de 24 628 viviendas particulares.

La muestra de conglomerados en el ámbito nacional es de 5 359, corresponde 3 782 conglomerados al área urbana y 1 577 conglomerados al área rural. Con respecto al tamaño de la muestra panel es de 1 769 conglomerados mientras que el tamaño de la muestra no panel es de 3 590 conglomerados.

Distribución de la muestra:

A continuación, se presenta la distribución de la muestra 2022 según departamentos de estudio:

DEPARTAMENTO	MUESTRA	
	CONGLOMERADOS	VIVIENDAS
TOTAL	5359	36 822
Amazonas	184	1 344
Ancash	206	1 464
Apurímac	134	984
Arequipa	254	1 694
Ayacucho	173	1 242
Cajamarca	213	1 560
Cusco	189	1 354
Huancavelica	144	1 092
Huánuco	183	1 324
Ica	247	1 604
Junín	230	1 596
La libertad	248	1 666
Lambayeque	215	1 436
Lima	944	6 270
Loreto	217	1 506
Madre de dios	101	694
Moquegua	166	1 094
Pasco	132	926

Piura	257	1 692
Puno	185	1 318
San Martín	202	1 400
Tacna	218	1 466
Tumbes	139	912
Ucayali	178	1 184

VII. NIVELES DE INFERENCIA

La muestra de la encuesta permite obtener estimaciones de las características socio- demográficas de la población para los niveles de inferencia siguientes:

a. De la muestra Integrada (panel y no panel)

Anual

- Nacional
 - Urbano Nacional
 - Rural Nacional
 - 24 departamentos, cada uno como dominio de estudio
 - Costa Urbana
 - Costa Rural
 - Sierra Urbana
 - Sierra Rural
 - Selva Urbana
 - Selva Rural
 - Área Metropolitana de Lima y Callao

Trimestral

- Nacional
 - Urbano Nacional
- Rural Nacional

b. De la muestra panel

Anual

- Nacional
 - Urbano Nacional
 - Rural Nacional
 - Costa
 - Sierra
 - Selva

De acuerdo con el diseño muestral, se podrá producir resultados para diferentes “arreglos” de unidades y su nivel de desagregación dependerá fundamentalmente de la precisión (error de muestreo) con que se estime el dato, y este del tamaño de la muestra para cada caso.

VIII. UNIDAD DE INVESTIGACIÓN

La unidad de investigación está constituida por:

- a. Los integrantes del hogar familiar,
- b. Los trabajadores del hogar con cama adentro, reciban o no pago por sus servicios,
- c. Los integrantes de una pensión familiar que tienen como máximo 9 pensionistas, y
- d. Las personas que no son miembros del hogar familiar, pero que estuvieron presentes en el hogar los últimos 30 días.

No serán investigados:

- e. Los integrantes de una pensión familiar que tiene de 10 a más pensionistas, y
- f. Los trabajadores del hogar con cama afuera.

IX. INFORMANTES

- a. Jefe del Hogar
- b. Cónyuge
- c. Perceptores (se refiere a la población de 14 años y más de edad que recibe ingresos monetarios)
- d. Personas de 12 años y más (a partir de esta edad se realiza la entrevista directa de la población, con respecto a los menores de 12 años la información la proporciona una persona responsable del hogar)
- e. Personas de 18 años y más (corresponde a la población que es seleccionada para proporcionar información del módulo de opinión).

X. NÚMERO DE PREGUNTAS E ÍTEMS: 404 preguntas y 329 ítems

Preguntas:

ENAH0. 01	= 143
ENAH0. 01-A	= 171
ENAH0. 01-B	= 43
ENAH0. 02	= 25
ENAH0. 04	= 22

Ítems:

Gastos 601	= 203 ítems
Gastos 602-612	= 126 ítems

XI. CARACTERÍSTICAS DE LA ENCUESTA

Método de Entrevista: Directa utilizando equipos móviles para captura de datos (TABLET).

Tipo de Encuesta: Encuesta de derecho, es decir, la población de estudio estará constituida por todos los residentes habituales del hogar.

Personal de Campo: Coordinadores Departamentales, Supervisores y Encuestadores.

XII. TEMAS INVESTIGADOS

1. Carátula (5 preguntas)

- El hogar fue entrevistado el año pasado

- Si existe otra vivienda en la cual algún miembro del hogar vive regularmente
- Ubicación vivienda secundaria
- Características de la fachada de la vivienda: Con tarrajeo y/o pintadas (por observación)
- Características de la calle donde se encuentra la vivienda: pista asfaltada, afirmada, veredas, alumbrado público (por observación)

2. Características de la vivienda y del hogar (32 preguntas)

Vivienda (7 preguntas)

- Tipo de vivienda
- Material de construcción predominante: paredes, pisos y techos
- Número total de habitaciones
- Número de habitaciones para dormir
- Licencia de construcción / Asistencia técnica

Hogar (25 preguntas)

- Régimen de tenencia
- Alquiler mensual y alquiler imputado
- Título de propiedad
- Registro del título de propiedad en la SUNARP
- Obtención de créditos para compra de vivienda, mejoramiento, o construcción de vivienda nueva
- Institución que le otorga el crédito y monto total del crédito
- Abastecimiento de agua en el hogar
- Potabilidad del agua
- Acceso al Servicio de Agua por red pública
- Horas al día que cuenta con el servicio (cuando accede todos los días de la semana)
- Días de la semana que cuenta con el servicio
- Horas al día (cuando no tiene acceso al servicio todos los días de la semana)
- Pagan por el servicio de agua
- A quien pagan por el servicio de agua
- Disponibilidad de servicio higiénico
- Tipo de alumbrado del hogar
- Tipo de servicio eléctrico del hogar según tipo de medidor
- Tipo de combustible para cocinar los alimentos
- Tenencia de teléfono fijo, celular, TV. cable y/o internet
- Último gasto mensual por abastecimiento de agua, alumbrado, combustible para cocinar, teléfono, celular, Internet, etc.
- Evaluación del nivel de cloro residual² en el agua que consume el hogar
- Quién extrajo la muestra de agua
- De dónde se extrajo la muestra de agua

3. Características de los miembros del hogar (21 preguntas)

² La evaluación del nivel de cloro en el agua de los hogares hasta febrero 2016 se realizaba con tres categorías: 1 = Seguro cloro, 2 = Inadecuada dosificación de cloro y 3 = Sin cloro. A partir de marzo de 2016, se realiza la transcripción de los valores que registra el kit comparador de cloro, con el fin de garantizar una mejor calidad en el recojo de la información. En la base de datos la información se presenta en dos variables, la P110A que presenta la información recodificada en tres categorías y la P110A_Modificada que presenta los valores que registró el kit comparador de cloro.

- Identificación de los miembros de hogar
- Relación de parentesco
- Número de núcleo familiar
- Relación de parentesco con el jefe del núcleo familiar
- Miembros del Hogar
- Tiempo de ausencia en el hogar (últimos 30 días)
- Tiempo de permanencia en el hogar (últimos 30 días)
- Sexo
- Edad
- Estado civil o conyugal
- Condición de ocupación menores de 5 a 17 años de edad
- Condición de actividad menores de 5 a 17 años de edad
- Descripción de tareas realizadas en el hogar o fuera del hogar
- Número de horas que realizó las tareas
- Número de orden de personas de 3 años y más de edad que son residentes habituales, que les corresponde aplicar el capítulo educación
- Número de orden de todas las personas residentes habituales que les corresponde aplicar el capítulo salud
- Número de orden de personas de 14 años y más de edad que son residentes habituales que les corresponde aplicar el capítulo empleo e ingreso
- Número de orden en el cap. 200 de la ENAHO 2021
- Persona nueva en la ENAHO 2022
- Motivo por el que la persona ya no vive en el hogar (Vivienda PANEL)
- Número de celular

4. Educación - Para personas de 3 años y más de edad (43 preguntas)

- Lengua materna que aprendió en su niñez
- Año o grado de estudios y nivel alcanzado
- Carrera superior universitaria o no universitaria que estudia o ha estudiado
- Nombre del centro de estudios donde estudia o ha estudiado la carrera superior universitaria o no universitaria
- Nivel de alfabetismo
- Programa de alfabetización
- Matriculados en algún programa de educación básica o superior el año pasado
- Año o grado de estudios al que asistió el año pasado
- Resultado que obtuvo el año pasado
- Matriculado en algún centro o programa de educación básica o superior en este año
- Año o grado de estudios al que está matriculado actualmente
- Asistencia actual a algún centro o programa de educación básica o superior a distancia
- Opinión sobre el servicio educativo
- Desplazamiento para asistir a su centro de estudios
- Asistencia en los últimos 12 meses a algún centro de enseñanza cuya duración sea menor a 3 años como academia, CETPROs u otros.
- En el presente año o años anteriores recibió enseñanza de nivel auxiliar técnico o técnico, en algún centro o programa de estudios cuya duración sea menor a 3 años
- Nombre de la carrera de nivel auxiliar técnico o técnico que estudia o estudió

- Tiempo de duración de la carrera de nivel auxiliar técnico o técnico
- Nombre del centro de estudios donde estudia o estudió la carrera de nivel auxiliar técnico o técnico
- Gasto en estudios realizado en los últimos 12 meses
- Gasto en estudios realizado en los últimos 12 meses como lo obtuvo
- Gasto en estudios realizado en los últimos 12 meses cuanto fue el monto S/.
- Gasto en estudios realizado en los últimos 12 meses donde lo compró
- Gasto en estudios realizado en los últimos 12 meses valorización si tuviera que comprarlo
- Gasto en estudios realizado en los últimos 12 meses valorización si tuviera número de veces
- Gasto en estudios en el mes anterior
- Gasto en estudios en el mes anterior como lo obtuvo
- Gasto en estudios en el mes anterior cuánto fue el monto S/,
- Gasto en estudios en el mes anterior valorización si tuviera que comprarlo
- Gasto en estudios en el mes anterior número de veces
- Razón por la que no asiste a un Centro de Enseñanza Regular (para menores de 25 años)
- Uso del servicio de Internet
- Dónde usó el servicio de Internet
- Equipo que utilizó para ingresar a internet (computadora, celular, Tablet, etc.)
- Frecuencia de uso del servicio de Internet
- Cómo obtuvo el servicio de Internet en la cabina pública
- Monto mensual por servicio de Internet en cabina pública
- Monto estimado por servicio de Internet en cabina pública
- Para qué uso el Internet
- Uso de celular de la población de 6 años y más edad
- Número de celular para las personas de 14 y más edad
- Uso de dispositivos de computadora (Tablet, computador, laptop, o similares)
- Actividades informáticas realizadas (copiar archivo, enviar correo, crear presentaciones, transferir archivos entre dispositivos, etc.)

5. Salud - Para todas las personas (29 preguntas)

- Fecha de nacimiento
- Tenencia de DNI
- Razones por las que no tiene DNI
- Razones por las que no tiene partida de nacimiento
- Hace 5 años vivía en este distrito
- Distrito, provincia y departamento donde vivía hace 5 años
- Donde vivía la madre cuando nació (migración)
- Lugar donde vivía la madre
- Limitación o dificultad de forma permanente que le impida o dificulte desarrollarse normalmente en sus actividades diarias.
- Personas con alguna enfermedad o malestar crónico
- Sufrió algún síntoma, enfermedad o accidente en las últimas 4 semanas
- Lugar de la consulta
- Personal que le dio atención
- Tiempo para obtener la cita de atención
- Tiempo de desplazamiento hasta el establecimiento de salud
- Tiempo de espera para ser atendido

- Razones por las que no acudió a un centro o establecimiento de salud
- Consulta por control del niño sano, programa de vacunas en los últimos 3 meses
- Consulta por control del niño sano, programa de vacunas en los últimos 3 meses como lo obtuvo
- Consulta por planificación familiar, suplemento de hierro en los últimos 12 meses
- Consulta por planificación familiar, suplemento de hierro en los últimos 12 meses como lo obtuvo
- Servicios de consulta, medicinas, análisis, etc., que recibió en las últimas 4 semanas. Servicio dental y conexos, oftalmológico, vacunas, etc., que recibió en los últimos 3 meses. Servicio de hospitalización, controles de embarazo o parto en los últimos 12 meses.
- Quién efectuó el gasto de los servicios de: consulta, dental, hospitalización, etc.
- Monto total de la compra o servicio
- Lugar donde adquirió los servicios de consulta, dental, hospitalización, etc.
- El monto estimado de los servicios de consulta, dental, hospitalización, etc.
- Afiliación al sistema de prestaciones de salud
- Aportación de las cuotas de afiliación
- Desplazamiento para acudir a un centro de salud

6. Empleo e Ingreso - Para personas de 14 años y más de edad (62 preguntas)

Condición de actividad - Semana Pasada (4 preguntas)

- Tenencia de trabajo
- Tenencia de algún empleo fijo al que próximamente volverá
- Tenencia de negocio propio al que próximamente volverá
- Actividad realizada al menos una hora para obtener ingresos (Incluye trabajador familiar no remunerado)

Ocupados

- Ocupación principal (14 preguntas)

- a) Ocupación (tareas realizadas en su ocupación)
- b) Rama de actividad
- c) Categoría de ocupación
- d) Tenencia de trabajador familiar no remunerado
- e) Tenencia de trabajadores remunerados
- f) Empresa o entidad donde trabaja en su actividad
- g) Registrado como persona jurídica
- h) Negocio o empresa que lleva los libros de ingresos y gastos de la SUNAT
- i) Tipo de pago o ingreso
- j) Tipo de contrato
- k) Tamaño del establecimiento
- l) Total horas trabajadas
- m) Tiempo que trabaja en esta ocupación principal

- Ocupación secundaria (9 preguntas)

- a) Tenencia de ocupación secundaria
- b) Actividad realizada al menos una hora para obtener ingresos

- c) Rama de actividad
 - d) Categoría de ocupación
 - e) Empresa o entidad donde trabaja en su actividad
 - f) Registrado como persona jurídica
 - g) Negocio o empresa que lleva los libros de ingresos y gastos de la SUNAT
 - h) Tamaño del establecimiento
 - i) Total de horas trabajadas
- **Total Horas (5 preguntas)**
 - a) Total horas trabajadas en la actividad principal y secundaria
 - b) Horas trabajadas normalmente
 - c) Razón por la que no trabajó el número total de horas
 - d) Deseo de trabajar más horas de lo normal
 - e) Disponibilidad para trabajar más horas
 - **Búsqueda de otro empleo (2 preguntas)**
 - a) Deseo y realización de acciones para cambiar de empleo
 - b) Motivo de búsqueda de otro trabajo
 - **Desocupados (7 preguntas)**
 - a) Búsqueda de trabajo
 - b) Actividades que realizó la semana pasada
 - c) Deseo de trabajar
 - d) Disponibilidad para trabajar
 - e) Razón por la que no buscó trabajo
 - f) Actividades que realizó para conseguir trabajo
 - g) Tiempo de búsqueda de trabajo
 - **Trabajo anterior (3 preguntas)**
 - a) Trabajó o no anteriormente
 - b) Rama de actividad
 - c) Categoría de ocupación

INGRESOS POR TRABAJO DEL HOGAR

a)

preguntas)

Ocupación Principal (7

Por trabajo dependiente

- Ingreso monetario
- Frecuencia de obtención de ingreso
- Ingreso total, según la frecuencia con que recibe (diario, semana, quincena o mes anterior)
- Descuento de ley Impuestos
- Otros descuentos Ingreso líquido
- Pago en especie Valor estimado
- Frecuencia con que recibe

Por trabajo independiente

- Ingreso monetario
- Ganancia Neta en el mes anterior
- Autoconsumo o autosuministro
- Utilización de Bienes Producidos y/o Adquiridos con fines comerciales Valor de los productos utilizados por Autoconsumo o autosuministro

b) Ocupación Secundaria (7 preguntas)

Ingresos por trabajo dependiente

- Ingreso ordinario en el mes anterior Ingreso total
- Descuento de ley Impuestos
- Otros descuentos Ingreso líquido
- En especie en los últimos 12 meses Frecuencia con que recibe
- Valor estimado

Ingresos por trabajo independiente

- Ingreso monetario y/o en especie Ganancia neta en el mes anterior
- Utilización de bienes producidos y/o adquiridos con fines comerciales Valor de los productos utilizados por autoconsumo o autosuministro

c) Ingresos Extraordinarios por Trabajo Dependiente (ocupación principal y/o secundaria) – (1 pregunta)

INGRESO POR TRABAJO DEL PRODUCTOR AGROPECUARIO (25 preguntas)

- Tipo de actividad agropecuaria
- Actividad agropecuaria permanente o eventual
- Superficie total de la explotación agropecuaria
- Total de parcelas que trabaja Número de parcela Extensión de la parcela
- Uso de la parcela, en los últimos 12 meses Régimen de tenencia
- Parcelas que cuentan con título de propiedad Tipo de riego de las tierras agropecuarias
- Total de la producción agrícola
- Total de los subproductos agrícolas
- Total de la producción forestal
- Total de gastos en actividad agrícola y/o forestales
- Total de la producción pecuaria
- Total de los subproductos pecuarios
- Total de gastos en actividades pecuarias

INGRESO POR TRABAJO DEL TRABAJADOR INDEPENDIENTE INFORMAL (22 preguntas)

- Características básicas del negocio o establecimiento
 - a) Negocio o establecimiento registrado en el registro público
 - b) Cómo lleva las cuentas de su negocio o actividad
 - c) Lugar donde desempeña su negocio o su actividad
 - d) Régimen de tenencia
 - e) Servicios básicos con que cuenta el local o establecimiento
 - f) Motivos por el cual inició dicho negocio o actividad
 - g) Tiempo de trabajo en su negocio o establecimiento
 - h) Total de personas remuneradas y/o no remuneradas que trabajan en el negocio o actividad
 - i) Funcionamiento del negocio en los últimos 12 meses
 - j) Tipo de actividad que realizó en su negocio o establecimiento la semana pasada
- Producción de Bienes
 - a) Ventas

- b) Autoconsumo
 - c) Gasto en materia prima e insumos
- Comercio
 - a) Ventas
 - b) Autosuministro
 - c) Compra de mercaderías
- Servicios
 - a) Ingresos totales
 - b) Autosuministro
 - c) Gastos
- Otros gastos
 - a) Envases, embalajes, combustible, electricidad, agua, etc.
- Características de la mano de obra y el empleo
- Hoja de control

INGRESOS POR TRANSFERENCIAS CORRIENTES (ÚLTIMOS 6 MESES) – 1 pregunta

- Frecuencia con que lo recibe
- Monto recibido por vez (del país o extranjero)
- Lugar donde cobra (Del extranjero)
- Destino del envío (Del extranjero)

INGRESOS POR RENTAS DE LA PROPIEDAD (ÚLTIMOS 12 MESES) - 1 pregunta

- Frecuencia con que lo recibe
- Monto recibido por vez

INGRESOS EXTRAORDINARIOS (ÚLTIMOS 12 MESES) - 1 pregunta

- Monto recibido

7. Sistema de Pensiones (2 preguntas)

- Sistema de pensiones al que está afiliado
- Último mes y año que aportó

8. Etnicidad (1 pregunta)

- Por sus antepasados y de acuerdo a sus costumbres se considera

9. Desplazamiento de la población a otros distritos por trabajo (1 pregunta)

10. Inclusión financiera (5 preguntas)

- Tenencia de cuenta de ahorros/corriente/plazo fijo/compensación por tiempo de servicios (CTS)
- Solicitud de préstamos y/o tarjetas de crédito
- Razón por la que no tiene ninguna cuenta, préstamo(s) y/o tarjeta de crédito
- Otras formas de ahorro no formal
- Medio de pago que utiliza para realizar sus compras

11. Gastos del hogar (82 preguntas y 319 ítems)

ALIMENTOS (26 preguntas)

- a) **Gastos en alimentos y bebidas consumidas dentro del hogar (últimos 15 días), (5 preguntas - 203 ítems)**
 - a) Relación de productos obtenidos, consumidos, comprados o regalados
 - b) Formas de obtención de los productos
 - c) Frecuencia de compra, cantidad y unidad de medida
 - d) Lugar de compra
 - e) Monto total de la compra
 - f) Frecuencia obtenida, cantidad y unidad de medida

- b) **Alimentos para consumir dentro del hogar obtenidos de instituciones benéficas (últimos 15 días) (7 preguntas - 3 ítems)**
 - g) Alimentos obtenidos, comprados o regalados de instituciones benéficas
 - h) Número de veces por semana
 - i) Número de raciones recibidas cada vez
 - j) Número de miembros del hogar que se beneficiaron
 - k) Pago o no por el alimento recibido
 - l) Monto pagado cada vez por el alimento
 - m) Pago mensual por otros conceptos

- c) **Alimentos consumidos fuera del hogar obtenidos de instituciones benéficas (Menores de 14 años) (7 preguntas - 3 ítems)**
 - n) Alimentos obtenidos, comprados o regalados de instituciones benéficas
 - o) Número de veces por semana
 - p) Número de raciones recibidas cada vez
 - q) Número de miembros del hogar que se beneficiaron
 - r) Pagó o no por el alimento recibido
 - s) Monto pagado cada vez por el alimento
 - t) Pago mensual por otros conceptos

- d) **Alimentos consumidos fuera hogar obtenido de restaurantes, ambulante, etc. (7 preguntas - 4 ítems)**
 - u) Alimento consumido la semana pasada
 - v) Número de veces por semana
 - w) Dónde lo consumió
 - x) Pagó por el alimento
 - y) Cuánto pagó cada vez
 - z) El consumo fue individual
 - aa) Número de personas que consumieron

OTROS GASTOS

- a) **Mantenimiento de la vivienda (mes anterior) (5 preguntas - 15 ítems)**
 - bb) Productos y/o servicios de mantenimiento de la vivienda, obtenidos, comprados y/o regalados
 - cc) Formas de obtención
 - dd) Lugar de compra
 - ee) Monto total de la compra
 - ff) Valor estimado de la adquisición

- b) Gastos en transportes y comunicaciones (mes anterior) (5 preguntas - 13 ítems)**
- gg)** Productos y/o servicios obtenidos, comprados o regalados
 - hh)** Formas de obtención de los productos y/o servicios
 - ii)** Lugar de compra
 - jj)** Monto total de la compra del producto o servicio
 - kk)** Valor estimado de la adquisición del producto y/o servicio.
- c) Gastos en transportes y comunicaciones (semana anterior) (7 preguntas - 4 ítems)**
- d) Gastos en servicios a la vivienda (mes anterior) (4 preguntas - 8 ítems)**
- ll)** Uso de los servicios a la vivienda
 - mm)** Formas de obtención de los servicios
 - nn)** Monto pagado por el servicio
 - oo)** Valor estimado de la adquisición del servicio.
- e) Esparcimiento, diversión y servicios de cultura (mes anterior) (5 preguntas - 8 ítems)**
- pp)** Productos y/o servicios obtenidos, comprados y/o regalados
 - qq)** Formas de obtención
 - rr)** Lugar de compra
 - ss)** Monto total de la compra del producto y/o servicio
 - tt)** Valor estimado de la adquisición.
- f) Bienes y servicios de cuidados personales (mes anterior) (5 preguntas - 11 ítems)**
- uu)** Productos y/o servicios obtenidos, comprados y/o regalados
 - vv)** Formas de obtención
 - ww)** Lugar de compra
 - xx)** Monto total de la compra del producto y/o servicio
 - yy)** Valor estimado de la adquisición
- g) Vestido y calzado (últimos 3 meses) (5 preguntas - 7 ítems)**
- zz)** Productos y/o servicios obtenidos, comprados y/o regalados.
 - aaa)** Formas de obtención
 - bbb)** Lugar de compra
 - ccc)** Monto total de la compra del producto y/o servicio
 - ddd)** Valor estimado de la adquisición
- h) Gastos de transferencia (últimos 3 meses) (2 preguntas - 9 ítems)**
- eee)** Gastos realizados por transferencias: alimentos, donaciones a instituciones, impuestos directos, remesas periódicas a otros miembros del hogar, etc.
 - fff)** Monto total del gasto realizado
- i) Muebles y enseres (últimos 12 meses) (5 preguntas - 6 ítems)**
- ggg)** Productos y/o servicios obtenidos, comprados o regalados
 - hhh)** Formas de obtención
 - iii)** Lugar de compra

jjj) Monto total de la compra del producto y/o servicio
kkk) Valor estimado de la adquisición.

j) Otros bienes y servicios (últimos 12 meses) (5 preguntas - 11 ítems)

III) Productos y/o servicios obtenidos, comprados o regalados
mmm) Formas de obtención
nnn) Lugar de compra
ooo) Monto total de la compra del producto y/o servicio
ppp) Valor estimado de la adquisición.

k) Equipamiento del hogar (7 preguntas - 22 ítems)

qqq) Tenencia de artefactos y/o vehículos del hogar
rrr) Número de artefactos y/o vehículos que tiene el hogar
sss) Tipo de uso
ttt) Año de adquisición
uuu) Mes de adquisición
vvv) Costo del equipo
www) Valor actual del mismo bien en las condiciones que lo adquirió

l) Venta de inmuebles, equipos (1 pregunta)

xxx) Últimos 12 meses ¿el hogar ha vendido:

12. Programas sociales de ayuda alimentaria (7 preguntas)

- Programas sociales de ayuda alimentaria que recibió
- Nombre de miembros del hogar que recibieron la ayuda alimentaria
- Tipo de programa de ayuda alimentaria recibida
- Lugar dónde recibió
- Frecuencia con que recibió
- Paga o retribución por la ayuda alimentaria
- Motivo por el que pagó

13. Programas sociales no alimentarios (4 preguntas)

- Programas sociales no alimentarios que recibió
- Nombre de miembros del hogar que recibieron los programas no alimentarios
- Tipo de programa no alimentario recibido
- Período que fue beneficiario del programa

14. Participación ciudadana (6 preguntas)

- Participación o inscripción en algún grupo de organización, asociación y/o programa social
- Número de miembros del hogar que participan o están inscritas en algún grupo, organización, asociación y/o programa social
- Número de personas que participan o están inscritas por tipo de organización, asociación y/o programas sociales
- Cargo que ocupa en la organización, asociación, y/o programa social
- Forma como accedió a la organización, asociación, y/o programa social
- Motivo por el cual no ha participado en las elecciones de las organizaciones, asociaciones, y/o programas sociales

15. MÓDULO DE OPINIÓN (43 preguntas)

- a) Gobernabilidad (Personas de 18 años y más de edad) (3 preguntas)**
- a) Grado de confianza en las Instituciones o empresas públicas
 - b) Principales problemas del país
 - c) Opinión sobre gestión del gobierno central, regional y local
- b) Corrupción (Personas de 18 años y más de edad) (5 preguntas)**
- d) Trámites realizados en instituciones públicas en los últimos 12 meses
 - e) Le solicitaron, se sintió obligado, o dio voluntariamente retribuciones (propinas, sobornos, coimas, etc.)
 - f) Denunció el hecho ante las autoridades
 - g) Razón principal por la que no denunció el hecho
 - h) Calificación de la atención recibida de las autoridades
- c) Democracia (Personas de 18 años y más de edad) (9 preguntas)**
- i) Conocimiento de democracia
 - j) Características más importantes de la democracia
 - k) Utilidad de la democracia
 - l) Importancia de la democracia en el Perú
 - m) Funcionamiento de la democracia
 - n) Motivos por los cuales funciona mal la democracia
 - o) Opiniones sobre tipos de gobierno
 - p) Existencia de respeto o no en el país a la libertad de expresión, la igualdad frente a la ley, elecciones transparentes
 - q) A qué grupo o comunidad se siente orgulloso de pertenecer
- d) Discriminación (Personas de 18 años y más de edad) (2 preguntas)**
- r) Se ha sentido discriminado en los últimos 12 meses, motivos de discriminación
 - s) Lugares donde se ha sentido discriminado
- e) Corrupción (Solo para el jefe/a del hogar y cónyuge) (1 pregunta)**
- t) En los últimos 12 meses le solicitaron, dio propinas sobornos a funcionarios del Estado.
- f) Acceso a la justicia (Solo para el jefe/a del hogar y cónyuge) (6 preguntas)**
- u) Desacuerdos o conflictos ocurridos los últimos 12 meses
 - v) Identificación del desacuerdo de mayor gravedad
 - w) Acudió a la justicia para solucionar el desacuerdo (conciliación, arbitraje, Ministerio Público, etc.)
 - x) Razones por las que no acudió a la justicia
 - y) Situación final del proceso (concluyó favorablemente, desfavorablemente, abandonado, etc.)
 - z) Grado de satisfacción de la atención recibida, orientación en el trámite, etc.
- g) Percepción del hogar (Solo para el jefe/a del hogar y cónyuge) (3 preguntas)**
- aa) Evaluación sobre la actual situación económica de su hogar
 - bb) Evaluación del nivel de vida de su hogar y la comunidad

cc) Motivo por el cual mejoró el nivel de vida de su hogar

h) Percepción de los programas no alimentarios (Jefe/a del hogar y cónyuge) (2 preguntas)

dd) Evaluación sobre el grado de satisfacción por los programas

ee) Evaluación de la contribución en la mejora del nivel de vida en su hogar

i) Percepción de los programas alimentarios (Jefe/a del hogar y cónyuge) (2 preguntas)

ff) Evaluación sobre el grado de satisfacción por los programas

gg) Evaluación de la contribución en la mejora del nivel de vida en su hogar

j) Nivel de vida / Situaciones adversas (Solo para el jefe/a del hogar y cónyuge) (9 preguntas).

hh) Evaluación del bienestar del hogar en relación al ingreso

ii) Monto mínimo mensual que requiere el hogar para vivir (alimentos, vestido, etc.)

jj) Estabilidad o no de los ingresos del hogar

kk) Calificación del hogar

ll) Afección de problemas en los últimos 12 meses

mm) Implicancia de los problemas en el hogar

nn) Qué solución se dio a estos problemas

oo) Solución de los problemas

pp) Tiempo en que se solucionarán

k) Educación de los padres del jefe/a del hogar (1

preguntas) qq) Nivel de estudios alcanzado por sus padres

PREGUNTAS INCORPORADAS A PARTIR DEL COVID-19 (14 preguntas)

A consecuencia del COVID-19, desde el último período de la muestra del mes de marzo del año 2020, se incorporaron 12 preguntas relevantes que permitan realizar el seguimiento o impacto en los hogares en el contexto de emergencia sanitaria, adicionalmente el año 2021 se incorporaron dos preguntas en el tema de salud, relacionadas al COVID-19. Estas variables continuarán investigándose en el año 2022, Las preguntas incorporadas están distribuidas en los capítulos de Educación, Salud y Empleo:

Educación - Para personas de 3 años y más de edad (2 preguntas)

- Medios con los cuales accede al contenido de las clases a distancia
- Desarrollo de las clases a distancia

Salud - Para todas las personas (4 preguntas)

- Realización de prueba para descartar el COVID-19
- Resultado de la prueba
- Toma de medicamentos como prevención o tratamiento para el COVID-19
- Medicamentos que tomó

Empleo - Para personas de 14 años y más de edad (8 preguntas) Ocupados

- Trabajó la semana anterior al estado de emergencia nacional
- Razón por la que no trabajó la semana anterior

Para trabajadores dependientes

- Modalidad en la que realizó su trabajo la semana anterior
- Razón por la que no pudo realizar ninguna labor
- El pago que recibió por la ocupación fue completo

Para trabajadores independientes

- Antes del 16 de marzo trabajaba como empleador o independiente
- Comparación de los Ingresos o ganancia actual con el mes anterior
- Motivo por el cual los ingresos/ganancias han sido menor de lo habitual o no ha tenido ganancias.

XIII. DEFINICIONES BÁSICAS

Vivienda Particular. - Es todo local o recinto estructuralmente “separado e independiente” que ocupa un edificio o una parte de él, y está conformado por una habitación o conjunto de habitaciones, usada o destinada a ser habitada por una o más personas con o sin vínculos familiares, siempre que en el período de la entrevista no se utilice para otros fines. Una vivienda particular puede servir de alojamiento a un máximo de cinco hogares. Por excepción se consideran como tal a las pensiones familiares que albergan menos de 10 pensionistas.

Hogar. - Es el conjunto de personas, sean o no parientes (padres, hijos solteros, hijos casados, hermanos, tíos etc.), que ocupan en su totalidad o en parte una vivienda, comparten las comidas principales y atienden en común otras necesidades vitales.

Se incluye también en este grupo a las personas a quienes el Jefe considera que son miembros del hogar, por razones de afecto (ahijados, compadres, padrinos, etc.). Por excepción, se considera hogar al constituido por una sola persona.

Actividad Económica. - Es la producción de bienes y servicios tal como han sido establecidos en el Sistema de Cuentas Nacionales (SCN) de las Naciones Unidas. Abarca toda la producción de mercado y algunos tipos de producción de no-mercado, incluyendo la producción y elaboración de productos primarios para autoconsumo, la construcción por cuenta propia y otras producciones de activos fijos para uso propio. Excluye las actividades no remuneradas, como las tareas domésticas no retribuidas y los servicios voluntarios a la comunidad.

Población en Edad de Trabajar (PET). - Es aquella población definida por las normas internacionales (OIT), como apta en cuanto a edad para ejercer funciones productivas (de 14 años y más de edad). Esta se subdivide en población económicamente activa (PEA) y población económicamente inactiva (PEI).

Población Económicamente Activa (PEA). - Comprende a todas las personas de catorce (14) años y más de edad que en la semana de referencia se encontraban: i. Trabajando, ii. No trabajaron, pero tenían trabajo, iii. Se encontraban buscando activamente un trabajo.

Ocupado. - Para determinar que una persona se encuentra ocupada se utilizan cuatro criterios:

- i. Ocupados son las personas de 14 años y más de edad que estuvieron participando en alguna actividad económica, en el período de referencia.
- ii. Los trabajadores dependientes que, teniendo empleo fijo, no trabajaron la semana anterior por hallarse de vacaciones, huelga, licencia por enfermedad, licencia pre y post-natal, etc., todas ellas pagadas.
- iii. Los trabajadores independientes, que estuvieron temporalmente ausentes del trabajo durante el período de referencia; pero, la empresa o negocio siguió funcionando.
- iv. A las personas que no estuvieron en ninguna de las condiciones anteriores se les indaga si realizaron alguna actividad económica en el período de referencia, al menos una hora, por lo cual recibirá pago en dinero y/o especie. El objetivo es recuperar las actividades realizadas; pero, que no son consideradas como trabajo por las personas.

Asimismo, en la definición de ocupados se incluye a las personas que trabajaron 15 horas o más como trabajador familiar no remunerado, a los practicantes con o sin remuneración y los oficiales y suboficiales de las Fuerzas Armadas y las Fuerzas Policiales.

Desocupado. - Según la OIT (1983), los desocupados son todas aquellas personas, de uno u otro sexo, que durante el período de referencia cumplen en forma simultánea con los tres requisitos siguientes:

- i. Sin empleo, es decir; que no tienen ningún empleo, como asalariado o como independiente.
- ii. Corrientemente disponible para trabajar, es decir; con disponibilidad para trabajar en un empleo asalariado o independiente, durante el período de referencia.
- iii. En busca de empleo, es decir; que habían tomado acciones concretas para buscar un empleo asalariado o independiente, en un período de referencia especificado.

Este concepto abarca tanto a las personas que buscaron trabajo pero que trabajaron antes (cesantes), como a los que buscaron trabajo por primera vez (aspirantes). Para fines de la encuesta se considera como desocupado a las personas de 14 años y más de edad, que en la semana de referencia buscaron trabajo, o hicieron gestiones para establecer un negocio por cuenta propia. Para fines comparativos se considera como desocupados a aquellos trabajadores familiares no remunerados que trabajaron menos de 15 horas semanales.

Tasa de Actividad (o de Participación). - Es la proporción de la PEA (población económicamente activa) respecto de la PET (población en edad de trabajar). La tasa de actividad es el indicador que mide el grado de participación de la población en edad de trabajar en la actividad económica.

Tasa de Empleo. - Es la proporción de personas que se encuentran activamente empleadas sobre el total de la población económicamente activa.

Tasa de Desempleo Abierto. - Es la proporción de personas desempleadas o desocupadas, que están buscando activamente un empleo, entre la PEA. Para determinar a los desempleados abiertos se consideran las condiciones siguientes: "Sin empleo", "Corrientemente disponible para trabajar" y en "Busca de empleo".

XIV. CONSTRUCCIÓN DE LOS FACTORES DE EXPANSIÓN

La metodología de estimación para procesar los datos de la Encuesta Nacional de Hogares, involucra el uso de un peso o factor de expansión para cada registro que

será multiplicado por todos los datos que conforman el registro correspondiente.

El factor final para cada registro tiene dos componentes:

- El factor básico de expansión y
- Los factores de ajuste por la no entrevista

El factor básico de expansión para cada hogar muestral es determinado por el diseño de la muestra. Equivale al inverso de su probabilidad final de selección, el mismo que es el producto de las probabilidades de selección en cada etapa.

El diseño de la muestra de la Encuesta Nacional de Hogares, involucra hasta tres etapas de muestreo donde las unidades son seleccionadas con probabilidades proporcionales al tamaño excepto la última etapa. En la última etapa se selecciona un número de viviendas para cada conglomerado teniendo en cuenta un intervalo de selección.

Por consiguiente, los **factores de expansión** básicos para la Encuesta Nacional de Hogares 2022 son ajustados teniendo en cuenta las proyecciones de población por grupos de edad y sexo para cada mes de encuesta y niveles de inferencia propuestos en el diseño de la muestra. Mayor detalle se encuentra en el documento anexo “Metodología del Cálculo del Factor de Expansión Anual”.

Cabe mencionar que se podrán obtener estimaciones para otros niveles de desagregación y su precisión o confiabilidad estadística dependerá fundamentalmente del número de casos u observaciones contenidas en la base de datos.

XV. IMPUTACIÓN

El proceso de imputación se realiza en dos etapas:

- Primera etapa corresponde a la asignación de las características cualitativas a los miembros del hogar que no informaron los módulos de educación (cap. 300), salud (cap. 400) y empleo (cap. 500) dejando como valor missing las variables cuantitativas. Para la imputación de las variables cualitativas se utiliza la técnica del Hot Deck.
- Segunda etapa se asignan valores a las variables cuantitativas que se registraron como valores missing (información no declarada por el informante). Para la imputación de las variables cuantitativas se utiliza la técnica de Matrices Medianas de esta forma se garantiza que el valor asignado no esté influenciado por los valores extremos.

IMPUTACIÓN HOT DECK (variables cualitativas)

En esta primera etapa las imputaciones se realizan a los miembros del hogar que no informaron simultáneamente los tres módulos (300, 400 y 500), no se imputan a las personas que informaron solo alguno de estos módulos.

La imputación sobre las variables cualitativas faltantes se efectúa mediante el reemplazo de las características de información cualitativa de las personas que respondieron los módulos de educación (300), salud (400) y empleo (500) con características demográficas, educativas, espaciales, socioeconómicas y temporales similares, a estas personas se les denomina “donantes”.

Procedimiento

La imputación a los miembros del hogar se realiza para tres grupos según el nivel de parentesco: jefes de hogar, cónyuge y resto de miembros.

- Los donantes para la imputación de los jefes del hogar se obtienen de la población de jefes de hogar mayores de 18 años.
- De igual manera para la imputación de los cónyuges (esposo o esposa), se obtienen donantes de la población de cónyuges (esposo o esposa), mayores de 12 años.
- La imputación del resto de miembros del hogar (hijo, yerno nuera, nietos, padres o suegros, otros parientes y otros no parientes), se efectúa a través de donantes de una población de hijos, yernos, nueras, nietos, padres o suegros, otros parientes y otros no parientes miembros del hogar.

Características de selección de los donantes

Para la selección de donantes para: jefes de hogar, esposas y los otros miembros del hogar y efectuar la imputación, se tiene en consideración las características y la jerarquía siguiente:

1. Dominio (departamento)
2. Estrato socioeconómico
3. Edad
4. Sexo
5. Número de miembros por hogar
6. Nivel educativo del jefe del hogar
7. Período (trimestre)

Orden de Selección

Para seleccionar un donante se ordena la base de datos de la siguiente manera:

- En el área urbana las variables de ordenamiento son: ubigeo, número de zona, número de manzana, vivienda, hogar y número de persona.
- En el área rural las variables de ordenamiento son: ubigeo, número de AER, número del conglomerado, número de vivienda, hogar, número de persona.

Nota:

Ubigeo : Departamento, provincia, distrito.

AER : Área de empadronamiento rural.

De no encontrar donantes de acuerdo con las características señaladas se amplía la selección al periodo anual, en segunda opción se apertura el estrato socioeconómico.

IMPUTACIÓN MATRICES MEDIANAS (variables cuantitativas)

En la segunda fase de imputación se emplea la "Imputación por Medianas", es decir, se construyen matrices de valores medianos en función a las variables con más alta correlación para la asignación de datos faltantes y posteriormente se emiten los listados que contengan las identificaciones del registro de hogares

imputados, así como indicadores sobre el porcentaje imputado a nivel de cada pregunta.

Esta metodología considera como primer nivel de corrección al conglomerado, el segundo el distrito, el tercero la provincia, el cuarto el departamento y, por último, el nacional, estos niveles están diferenciados si se trata del área urbana o rural. Este procedimiento garantiza que en la mayoría de las veces la corrección se haga con información de hogares que pertenecen al mismo estrato socioeconómico.

En la corrección de datos de educación y salud se ha incluido como variable determinante el tipo de establecimiento; es decir, si pertenece al sector público o privado; y en la corrección de datos de consumo de alimentos fuera del hogar y servicios de transporte se ha considerado, además, como variable determinante del gasto, el lugar en que se efectúa el consumo y el tipo de servicio que se recibe. Así, por ejemplo, se diferencia si se trata de alimentos consumidos en restaurantes, ambulantes, mercados u otros; en el caso del transporte público si éste proviene de microbús, taxi, colectivo u otro.

XVI. DEFLACTACIÓN

Es el proceso de transformar valores monetarios nominales en valores monetarios reales (a precios constantes de un determinado período), mediante la aplicación de un índice de precios que elimine el efecto de los precios en el período de análisis.

XVII. CÁLCULO DE ERRORES MUESTRALES

Consiste en la elaboración y aplicación de un conjunto de indicadores que permitirá cuantificar los posibles errores que puedan obtenerse como consecuencia de la selección aleatoria de los hogares informantes.

XVIII. RESULTADOS

- Indicadores de pobreza (Información estadística anual)
- Perú: Perfil de la pobreza por dominios geográficos (información estadística anual)
- Informe Técnico de estadísticas sobre condiciones de vida por trimestre
- Informe Técnico de estadísticas sobre género por trimestre
- Informe Técnico de estadísticas sobre tecnología de la información por trimestre
- Informe Técnico de estadísticas sobre la situación de la niñez y del adulto mayor por trimestre
- Informe Técnico sobre el comportamiento de los indicadores de mercado laboral a nivel nacional por trimestre
- Informe Técnico de percepción ciudadana sobre gobernabilidad, democracia y confianza en las instituciones, información semestral
- Indicadores de presupuesto por resultado por semestre y anual

Principales Indicadores que se procesan con la encuesta

- Niveles de pobreza, según gasto de consumo
- Niveles de pobreza, según el ingreso
- Indicadores de necesidades básicas insatisfechas
- Mediciones de pobreza con el método integrado

Vivienda y Hogar

- Indicadores de infraestructura de la vivienda
- Indicadores de servicios básicos de la vivienda
- Indicadores de tipos de hogar
- Indicador de tamaño del hogar
- Indicador de hacinamiento

Educación

- Tasa de analfabetismo
- Tasa de asistencia neta a educación inicial
- Tasa de asistencia neta a educación primaria
- Tasa de asistencia neta a educación secundaria
- Promedio de años de estudios
- Porcentaje de población de 12 a 24 años que no estudia ni trabaja

Salud

- Acceso a los servicios de salud

Empleo

- Población en edad de trabajar
- Tasa de participación
- Tasa de actividad
- Tasa de desempleo abierto
- Tasa de ocupación
- Ocupados con previsión social

Programas Sociales

- Cobertura de los programas sociales
- Acceso a programas sociales

XIX. COMISIÓN CONSULTIVA PARA ESTIMACIÓN DE LA POBREZA

A inicios del año 2007, con el auspicio del Banco Mundial (BM), se convocó a un Comité Asesor especializado para participar en la medición de la pobreza, conformado por representantes de organismos internacionales de organismos gubernamentales nacionales, representantes de la comunidad académica y de centros de investigación. Este Comité Asesor ha estado participando regularmente cada año en la medición de la pobreza.

El 13 de abril del año 2010, mediante Resolución Suprema N° 097-2010-PCM, publicado en el diario oficial El Peruano, el Comité Asesor de Pobreza se constituye en "**Comisión Consultiva para Estimación de la Pobreza y otros indicadores relacionados en el país**"; cuyo objetivo es garantizar la calidad, transparencia y

confianza de la información en el campo de la medición de la pobreza y otros indicadores relacionados.

Son funciones de la Comisión Consultiva:

- Asesorar en todos los procesos de medición de la pobreza y otros indicadores relacionados;
- Supervisar los procesos que se adopten;
- Evaluar y validar los resultados de las estimaciones que se efectúen periódicamente;

- Mantener reserva de los procesos de la información hasta antes de su publicación;
- Proponer modificaciones a las metodologías, cuando sea pertinente, a partir de la evaluación de su validez y la experiencia internacional, siempre que se asegure la comparabilidad en el tiempo de los indicadores.

INTEGRANTES DE LA COMISIÓN CONSULTIVA

La Comisión Consultiva está presidida por el Jefe del Instituto Nacional de Estadística e Informática (INEI) y tiene como Secretaría Técnica a la Dirección Técnica de Demografía e Indicadores Sociales.

Integran la Comisión representantes de organismos internacionales como son el Banco Mundial (BM), Banco Interamericano de Desarrollo (BID); Institut de Recherche Pour le Développement (IRD) del Gobierno Francés; Fondo de Población de las Naciones Unidas (UNFPA).

Los organismos nacionales que la integran son: Ministerio de Economía y Finanzas (MEF); Centro Nacional de Alimentación y Nutrición (CENAN); Comisión Interministerial de Asuntos Sociales (CIAS), el Banco Central de Reserva (BCR), el Consorcio de Investigación Económica y Social (CIES), GRADE, el Instituto de Estudios Peruanos (IEP), la Mesa de Concertación para la Lucha Contra la Pobreza, la Pontificia Universidad Católica del Perú (PUCP), la Universidad del Pacífico, y expertos independientes.

ANEXO 2- ANALÍTICA DE DATOS

La analítica de datos estuvo conformada por la recopilación de los datos, minería de datos, ingeniería de datos, comparación de porcentaje de participación, ingresos y coeficiente de equidad Gini y evaluación de factores de importancia, se detallan a continuación:

I. RECOPIACIÓN DE DATOS

Se recopiló de la base de datos del Instituto Nacional de Estadística e Informática (INEI) siguiendo los pasos:

1. Se busco el acceso en la siguiente dirección: <https://proyectos.inei.gob.pe/microdatos/>
2. Se busco en consulta de encuestas la ENAHO Metodología ACTUALIZADA.
3. Se seleccionó por condiciones de Vida y Pobreza – ENAHO
4. Se eligió del año 2012 al 2022 con intervalos de 2 años
5. Se eligió el trimestre 4
6. Se eligió el módulo correspondiente a Empleo e Ingresos
7. Se hizo la descarga de la base de datos en formato CSV

La Encuesta Nacional de Hogares (ENAHO), estuvo conformada por el área urbana y rural de todo el país y se obtuvo información anual con desagregación departamental.

II. MINERÍA DE DATOS

1. En primera instancia se realizó una minería de datos con la finalidad de ubicar indicadores que permitan conocer la evolución de la pobreza monetaria, del bienestar y de las condiciones de vida de los hogares, además de efectuar diagnósticos sobre las condiciones de vida y pobreza monetaria de la población.
2. En segunda instancia se filtró solo los datos correspondientes a ingresos por pensiones de jubilación.
3. En una tercera etapa de minería de datos se filtró los ingresos por AFP y ONP
4. En una cuarta etapa de minería de datos se identificó que posibles características se asociaban a ingresos por pensiones de jubilación (según la experiencia de personas en temas pensionarios); siendo estas: zona geográfica (ubigeo), entidad donde trabajó la persona antes de su jubilación, sistema pensionario, sexo, edad, nivel de estudio e ingreso por pensión de jubilación en soles (S/).

III. INGENIERÍA DE DATOS

1. En la parte de ingeniería de datos se generó variables dummy a las características: zona geográfica obteniendo nuevas variables de costa, sierra, selva y Lima metropolitana (donde 0 se consideró como ausencia de esa variable y 1 como presencia de esa variable).
2. Para las características entidad de trabajo al ser una variable de escala nominal se consideró como 0 = entidad pública y 1 = entidad privada.

3. Con referencia a la característica sistema pensionario se consideró 0 = ONP y 1 = AFP.
4. Para la característica sexo se codificó 0 = mujer y 1 = hombre.
5. La característica edad se encontró en escala numérica discreta.
6. La característica nivel de estudio estuvo en escala ordinal de 10 puntos (desde 1 = educación inicial, 2 = primaria incompleta, 3 = primaria completa, 4 = secundaria incompleta, 5 = secundaria completa, 6 = superior no universitaria incompleta, 7 = superior no universitaria completa, 8 = superior universitaria incompleta, 9 = superior universitaria completa, hasta 10 = maestría/doctorado).
7. La variable respuesta monto de ingreso por pensiones de jubilación estuvo en escala numérica discreta.

IV. COMPARACIÓN DE PORCENTAJE DE PARTICIPACIÓN, INGRESOS Y COEFICIENTE GINI

Se realizó la descripción del porcentaje de participación, ingresos para cada sistema de pensiones ONP y AFP, en primera instancia en la ventana de tiempo conformada entre los años 2012 al 2022 con intervalo de 2 años, además del coeficiente Gini para determinar la equidad entre los sistemas pensionarios, la información fue segregada por zonas: costa, sierra, selva, Lima metropolitana y total Perú. Adicionalmente, teniendo en cuenta el ultimo estado de los datos e información actual (2022) se describió el porcentaje de participación, ingresos y coeficiente Gini. Se utilizó el software R 4.2.2.

V. EVALUACIÓN DE FACTORES DE IMPORTANCIA

Se trabajó con modelamientos de Machine Learning por ser técnicas modernas con mayor capacidad para generar predicciones con cercanía a la realidad; además, de ser el "Estado del Arte" en analítica de datos, se propuso modelos de regresión de árboles de decisión potenciados por gradientes: XGBoost, LightGBM y CatBoost, teniendo en cuenta al coeficiente de determinación (R^2) (se determinó que el valor de R^2 sea mayor al 0.85 para ser considerado un buen modelo). Para la evaluación de los factores de mayor impacto se utilizó el diagrama de cascada (waterfall); una vez identificados se evaluó mediante valores SHAP el impacto de los factores sobre la variable respuesta ingresos por pensión de jubilación; para realizar el proceso se utilizó el software Python 3.11.2 con la librería SHAP.