

**Análisis de la Tenencia de Vivienda en Bogotá D.C.
mediante Modelos de Aprendizaje Automático
(ECV 2024)**

Olga Lucia Parra Sanchez

Alberto Cortes Salazar

Universidad de Santo Tomas

Curso de Econometría con Aprendizaje de Máquina

Prof. Augusto Andres Alean Romero

2025

Índice

Introducción	1
Planteamiento del problema	2
Pregunta de investigación	3
Datos y preprocesamiento	4
Fuente de datos	4
Variables utilizadas	4
Análisis descriptivo: Bogotá D.C.	5
Modelos de aprendizaje automático	6
Modelos implementados	6
Comparación de modelos	6
Resultados obtenidos en la ejecución del modelo	8
Tenencia de vivienda en Bogotá D.C.	8
Evaluación y selección del mejor modelo	9
Limitación temporal	9
Importancia de variables	10
Validación del mejor modelo	11
Conclusiones	12
Limitaciones	13
Referencias	14

Índice de figuras

1.	Tenencia de vivienda en Bogotá D.C. por sexo de jefatura (ECV 2024). . . .	5
2.	Comparación de modelos según RMSE (menor es mejor).	6
3.	Importancia de variables según Random Forest.	10
4.	Valores predichos vs reales para el modelo de regresión lineal.	11

Índice de tablas

1.	Resultados de desempeño de los modelos predictivos	9
----	--	---

Introducción

La tenencia de vivienda es un elemento estructural del bienestar y la calidad de vida de los hogares, y su análisis constituye una herramienta fundamental para evaluar las condiciones socioeconómicas en contextos urbanos y nacionales. Desde una perspectiva conceptual, la literatura define la tenencia de vivienda como la forma en que una persona o un hogar tiene derecho a habitar un inmueble, ya sea mediante propiedad propia, arrendamiento u otras modalidades de ocupación, como el usufructo u otras formas de cesión (1). Este indicador es ampliamente utilizado en estudios sociales y económicos, ya que refleja no solo la capacidad económica de los hogares, sino también las dinámicas del mercado inmobiliario, el acceso al crédito, la estabilidad residencial y la efectividad de las políticas públicas de vivienda.

En Colombia, la Encuesta de Calidad de Vida (ECV), elaborada por el Departamento Administrativo Nacional de Estadística (DANE), constituye una de las principales fuentes oficiales para la caracterización de las condiciones de vida de los hogares, incluyendo información sobre la tenencia de vivienda (8). De acuerdo con los resultados nacionales publicados por el DANE, el arriendo se ha consolidado como una modalidad predominante en los últimos años. En 2023, aproximadamente el 39,5 % de los hogares habitaba una vivienda propia, mientras que el 40,3 % residía en arriendo (7). Para 2024, estas proporciones se mantuvieron relativamente estables, con cerca del 39,6 % de los hogares en vivienda propia y alrededor del 40,4 % en arriendo (6). Esta comparación, basada en boletines técnicos oficiales, sugiere un contexto nacional caracterizado por la persistencia del arriendo como forma de tenencia dominante y una estabilidad relativa en los niveles de propiedad.

Bogotá D.C., como principal área metropolitana y capital del país, concentra dinámicas urbanas, económicas y demográficas que tienden a intensificar estos patrones observados a nivel nacional. Diversos estudios y reportes locales han señalado que en la ciudad la proporción de hogares que habitan en arriendo supera a la de aquellos que

residen en vivienda propia, fenómeno asociado con factores como los altos costos del suelo, las restricciones de acceso al crédito hipotecario, la movilidad laboral y las desigualdades de ingreso (4). En este contexto, resulta pertinente analizar de manera específica la tenencia de vivienda en Bogotá D.C., con el fin de comprender cómo se manifiestan estas dinámicas en un entorno urbano de alta complejidad.

El presente trabajo se enfoca en el análisis de la tenencia de vivienda en Bogotá D.C. para el año 2024, utilizando datos agregados por departamento y sexo de jefatura del hogar provenientes del Anexo ECV–Sexo 2024. Adicionalmente, se implementan y comparan distintos modelos de aprendizaje automático con el objetivo de identificar qué variables explican, dentro de las limitaciones del conjunto de datos disponible, el porcentaje de hogares con vivienda propia. Si bien la referencia a 2023 se emplea únicamente como contexto nacional descriptivo, el análisis empírico y la estimación de modelos se restringen al año 2024, reconociendo explícitamente las limitaciones derivadas del uso de información agregada y de la ausencia de una serie temporal.

Planteamiento del problema

Bogotá D.C. concentra una proporción importante de hogares del país y presenta dinámicas urbanas que pueden intensificar las dificultades de acceso a vivienda propia, como los altos costos del suelo, la presión del mercado de alquiler y la heterogeneidad socioeconómica. En este contexto, comprender la distribución de la tenencia de vivienda y los factores asociados resulta relevante para apoyar la interpretación de indicadores oficiales y la discusión académica sobre condiciones de vida.

Sin embargo, aunque la ECV ofrece información oficial sobre vivienda y tenencia, los anexos estadísticos suelen presentarse en forma agregada y tabular, lo cual limita el análisis directo de patrones y relaciones entre variables. En particular, no siempre es evidente qué variables (por ejemplo, la estructura de tenencia en arriendo, usufructo u otros tipos, el total de hogares o el sexo de la jefatura) se asocian con el porcentaje de hogares con vivienda propia, especialmente cuando el análisis se restringe a datos

agregados por territorio.

Adicionalmente, el uso de modelos de aprendizaje automático en este tipo de información requiere precaución metodológica: (i) la disponibilidad de un único año limita la evaluación de tendencias temporales, (ii) la agregación por departamento reduce la variabilidad observada y (iii) el conjunto de predictores está condicionado por las variables incluidas en el cuadro analizado. Por ello, se plantea la necesidad de aplicar un enfoque cuantitativo que, sin exceder las posibilidades del dato, permita caracterizar Bogotá D.C. en 2024 y comparar modelos predictivos para explicar el porcentaje de vivienda propia.

Pregunta de investigación

¿En qué medida las variables disponibles en el Cuadro 4 del Anexo ECV–Sexo 2024 permiten explicar el porcentaje de hogares con vivienda propia en Bogotá D.C., y qué modelo (lineal o no lineal) presenta el mejor desempeño bajo el esquema de evaluación aplicado?

Datos y preprocesamiento

Fuente de datos

Los datos utilizados provienen del anexo *ECV – Sexo – 2024*, (5) específicamente del **Cuadro 4: Tenencia de vivienda**, el cual presenta información agregada por departamento y sexo de jefatura del hogar.

Es importante resaltar que los datos:

- Corresponden únicamente al año 2024.
- Están agregados por departamento, no por individuo.
- Incluyen porcentajes de distintas modalidades de tenencia.

Variables utilizadas

Las variables consideradas en el análisis fueron:

- **total_hogares**: número total de hogares por departamento y sexo.
- **pct_propia**: porcentaje de hogares con vivienda propia (variable objetivo).
- **pct_arriendo**: porcentaje de hogares en arriendo.
- **pct_usufructo**: porcentaje de hogares en usufructo.
- **pct_otros**: otras formas de tenencia.
- **sexo_jefatura**: variable categórica (masculina / femenina).

Las variables categóricas fueron transformadas mediante *One-Hot Encoding*, y los valores faltantes se trataron mediante imputación por la mediana (variables numéricas).

Análisis descriptivo: Bogotá D.C.

La Figura 1 presenta la distribución porcentual de la tenencia de vivienda en Bogotá D.C. para 2024, diferenciada por sexo de jefatura.

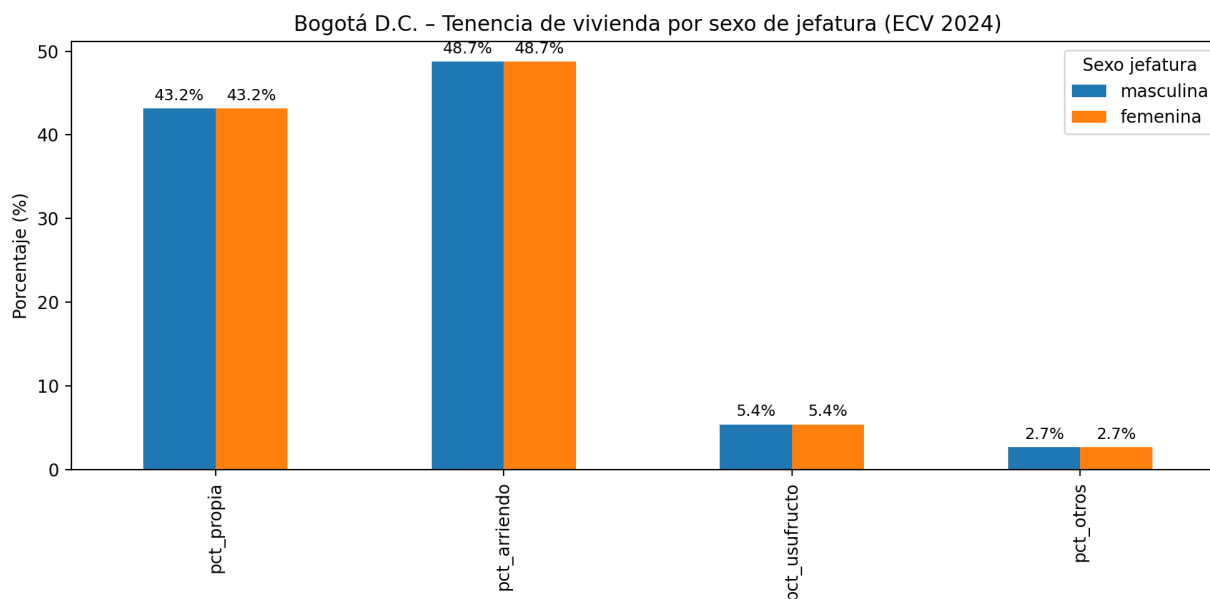


Figura 1

Tenencia de vivienda en Bogotá D.C. por sexo de jefatura (ECV 2024).

Se observa que:

- El arriendo es la modalidad predominante (aprox. 48.7%).
- La vivienda propia representa cerca del 43.2%.
- No existen diferencias porcentuales entre jefatura masculina y femenina.

Modelos de aprendizaje automático

Modelos implementados

Se entrenaron y evaluaron los siguientes modelos:

- Regresión lineal
- Ridge
- Lasso
- Elastic Net
- Árbol de decisión
- Random Forest
- K-Vecinos más cercanos (KNN)

La métrica utilizada para la comparación fue el **Error Cuadrático Medio Raíz (RMSE)**.

Comparación de modelos

La Figura 2 muestra la comparación de los modelos evaluados.

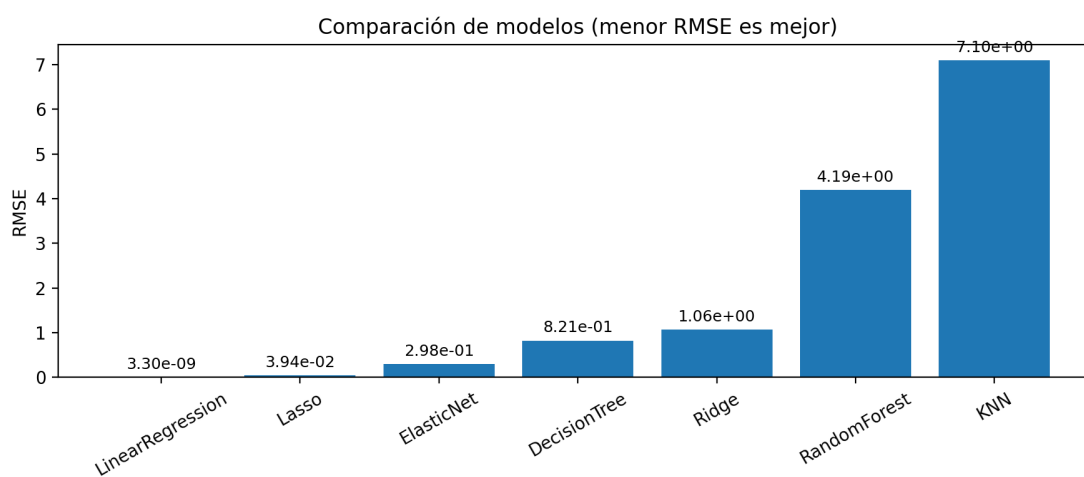


Figura 2

Comparación de modelos según RMSE (menor es mejor).

El modelo de **regresión lineal** presenta un RMSE prácticamente nulo y un coeficiente de determinación $R^2 = 1$, superando a los modelos no lineales.

Este resultado se explica por:

- El bajo número de variables explicativas.
- La estructura casi determinística de los datos agregados.

Resultados obtenidos en la ejecución del modelo

Durante la ejecución del script principal (`main.py`), se obtuvieron los siguientes resultados en consola, correspondientes al análisis de la tenencia de vivienda en Bogotá D.C. para el año 2024.

Tenencia de vivienda en Bogotá D.C.

El programa reportó la siguiente información agregada por sexo de jefatura del hogar:

- Jefatura masculina:
 - Total de hogares: 1611
 - Vivienda propia: 43.15 %
 - Arriendo: 48.75 %
 - Usufructo: 5.40 %
 - Otros tipos: 2.70 %
- Jefatura femenina:
 - Total de hogares: 1488
 - Vivienda propia: 43.15 %
 - Arriendo: 48.75 %
 - Usufructo: 5.40 %
 - Otros tipos: 2.70 %

Estos resultados confirman que, para el conjunto de datos analizado, no existen diferencias porcentuales en la tenencia de vivienda entre hogares con jefatura masculina y femenina en Bogotá D.C.

Evaluación y selección del mejor modelo

La comparación de modelos realizada mediante una partición de entrenamiento y prueba (*test split*) produjo los siguientes resultados:

Cuadro 1

Resultados de desempeño de los modelos predictivos

Modelo	RMSE	R^2
Regresión Lineal	$3,30 \times 10^{-9}$	1.0000
Lasso	$3,94 \times 10^{-2}$	0.99999
Elastic Net	$2,98 \times 10^{-1}$	0.99946
Árbol de decisión	$8,21 \times 10^{-1}$	0.99593
Ridge	1,06	0.99316
Random Forest	4,19	0.89378
KNN	7,10	0.69525

De acuerdo con estos resultados, el modelo de **regresión lineal** fue seleccionado como el mejor modelo, al presentar el menor error (RMSE) y un coeficiente de determinación igual a uno.

Limitación temporal

Es importante destacar que, tal como se indicó durante la ejecución del programa, el análisis se basa únicamente en información correspondiente al año 2024. Por esta razón, no es posible establecer si la adquisición de vivienda en Bogotá D.C. presenta una tendencia de aumento o disminución en el tiempo.

Importancia de variables

La Figura 3 presenta la importancia de las variables estimada mediante Random Forest.

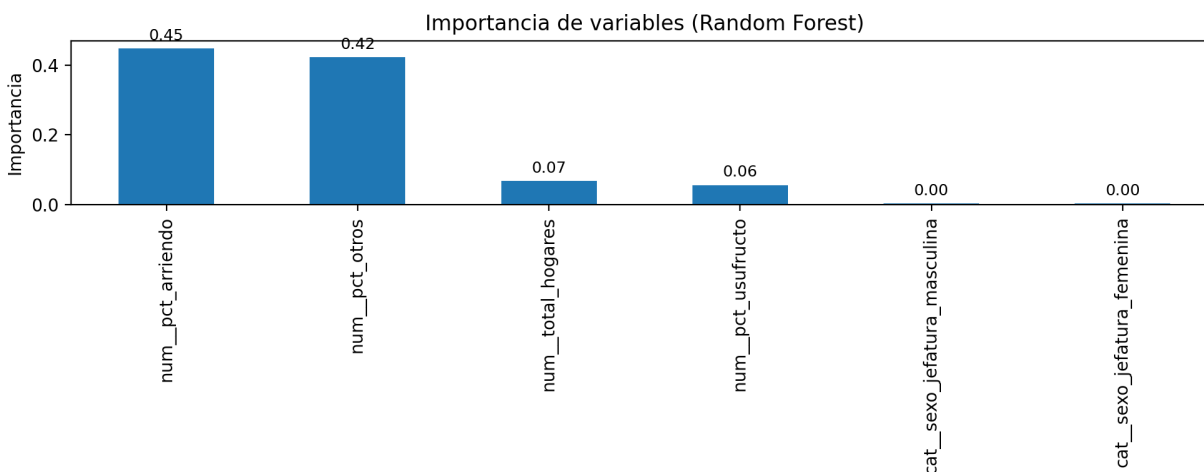


Figura 3

Importancia de variables según Random Forest.

Las variables más influyentes son:

- **pct_arriendo**
- **pct_otros**

En contraste, el sexo de la jefatura presenta una importancia prácticamente nula, indicando que no contribuye a explicar la variación del porcentaje de vivienda propia en este conjunto de datos.

Validación del mejor modelo

La Figura 4 muestra la comparación entre valores reales y predichos para el mejor modelo (regresión lineal).

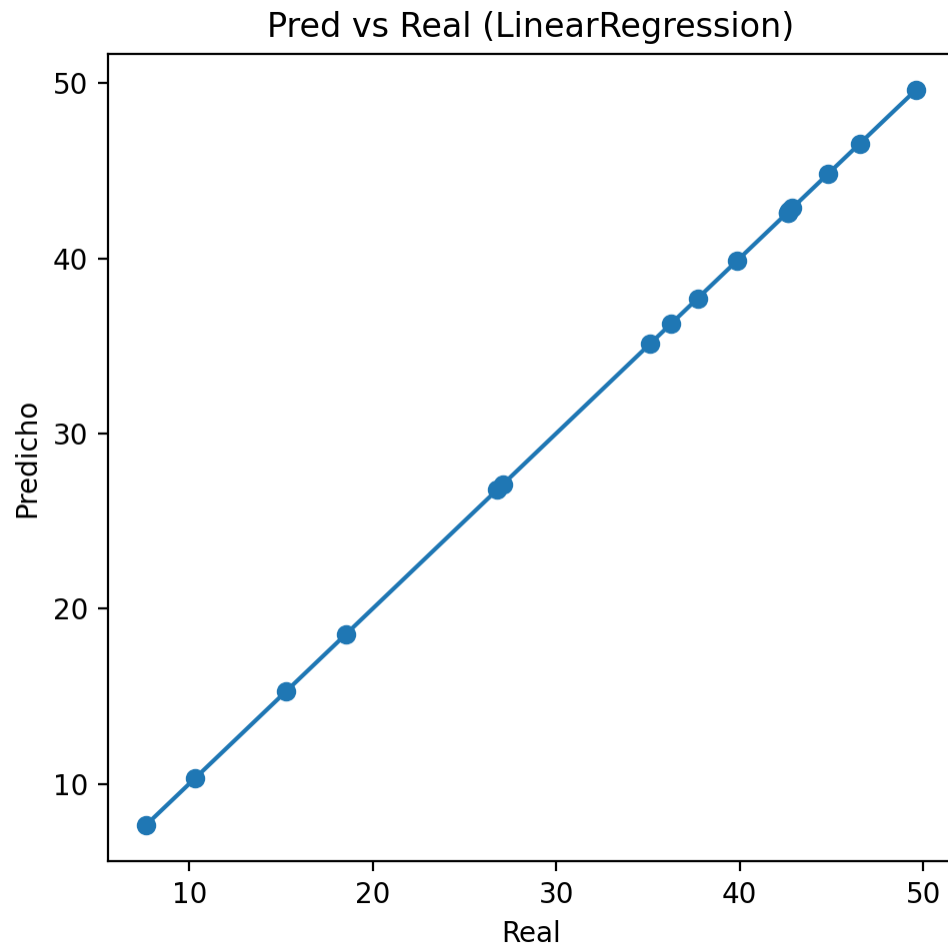


Figura 4

Valores predichos vs reales para el modelo de regresión lineal.

Los puntos se alinean prácticamente sobre la diagonal $y = x$, lo que confirma visualmente el ajuste perfecto evidenciado por las métricas cuantitativas.

Conclusiones

A partir del análisis realizado sobre los datos de la Encuesta de Calidad de Vida (ECV) correspondientes al año 2024, se concluye que:

- En Bogotá D.C., durante 2024, la modalidad de arriendo constituye la forma predominante de tenencia de vivienda, con un porcentaje cercano al 48,75 %, mientras que la vivienda propia representa aproximadamente el 43,15 % de los hogares. Este resultado es consistente con el contexto nacional descrito en los boletines oficiales del DANE para 2023 y 2024, en los cuales el arriendo se mantiene como la modalidad más frecuente a nivel país.
- No se evidencian diferencias porcentuales en la tenencia de vivienda entre hogares con jefatura masculina y femenina en Bogotá D.C. Este resultado se mantiene tanto en el análisis descriptivo como en la salida numérica reportada durante la ejecución del modelo, lo que sugiere que, a nivel agregado, el sexo de la jefatura del hogar no introduce variaciones observables en las modalidades de tenencia.
- En la comparación de modelos de aprendizaje automático, la regresión lineal presentó el mejor desempeño, con un error cuadrático medio raíz (RMSE) prácticamente nulo y un coeficiente de determinación igual a uno, superando a modelos no lineales como Random Forest y K-Vecinos más cercanos. Este comportamiento se explica por la baja dimensionalidad del conjunto de datos y por la estructura casi determinística de las variables agregadas utilizadas.
- El análisis de importancia de variables indicó que las modalidades de arriendo y otros tipos de tenencia son los principales factores asociados al porcentaje de hogares con vivienda propia, mientras que el sexo de la jefatura del hogar no aporta información explicativa relevante dentro del conjunto de predictores considerados.
- Finalmente, dado que el estudio se basa en información correspondiente a un único

año (2024) y en datos agregados por departamento, no es posible inferir tendencias temporales ni identificar condiciones individuales de adquisición de vivienda. Por tanto, los resultados deben interpretarse como una caracterización descriptiva y explicativa del estado de la tenencia de vivienda en Bogotá D.C. para 2024, contextualizada a partir de la información nacional disponible para 2023 y 2024.

Limitaciones

El estudio presenta las siguientes limitaciones:

- Ausencia de serie temporal para analizar evolución en el tiempo.
- Falta de variables socioeconómicas individuales (ingresos, educación, estrato).
- Datos agregados por departamento.

Futuras investigaciones podrían integrar otros cuadros de la ECV para ampliar el análisis multivariable.

Referencias

Housing tenure. Wikipedia. Recuperado de

https://en.wikipedia.org/wiki/Housing_tenure

Departamento Administrativo Nacional de Estadística (DANE). (2025). Encuesta Nacional de Calidad de Vida (ECV) 2024. Recuperado de

<https://www.dane.gov.co/files/operaciones/ECV/pres-ECV-2023.pdf>

La República (2025). Cuatro de cada 10 hogares en Colombia son dueños de su vivienda. Recuperado de <https://www.larepublica.co/economia/>

[cuatro-de-cada-10-hogares-son-duenos-de-su-vivienda-4118757](https://www.larepublica.co/economia/cuatro-de-cada-10-hogares-son-duenos-de-su-vivienda-4118757)

Cámara de Comercio de Bogotá (2023). En Bogotá es mayor el número de personas que viven en arriendo. Recuperado de

<https://www.ccb.org.co/specialized-information/observatory/>

[business-environment/urban-and-regional-development/](https://www.ccb.org.co/specialized-information/observatory/business-environment/urban-and-regional-development/)

[in-sabana-centro-is-greater-the-number-of-people-who-live-in-their-own-home](https://www.ccb.org.co/specialized-information/observatory/business-environment/urban-and-regional-development/in-sabana-centro-is-greater-the-number-of-people-who-live-in-their-own-home)

Departamento Administrativo Nacional de Estadística (DANE). (2024). *Anexo Encuesta de Calidad de Vida por sexo, 2024*. Recuperado de

<https://www.dane.gov.co/files/operaciones/ECV/anex-ECV-Sexo-2024.xlsx>

Departamento Administrativo Nacional de Estadística (DANE). (2025). *Boletín técnico: Encuesta Nacional de Calidad de Vida (ECV) 2024*. Recuperado de

<https://www.dane.gov.co/files/operaciones/ECV/bol-ECV-2024.pdf>

Departamento Administrativo Nacional de Estadística (DANE). (2024). *Presentación de resultados: Encuesta Nacional de Calidad de Vida (ECV) 2023*. Recuperado de

<https://www.dane.gov.co/files/operaciones/ECV/pres-ECV-2023.pdf>

Departamento Administrativo Nacional de Estadística (DANE). (s. f.). *Encuesta Nacional de Calidad de Vida (ECV)*. Recuperado de [https://www.dane.gov.co/index.php/](https://www.dane.gov.co/index.php/estadisticas-por-tema/pobreza-y-condiciones-de-vida/calidad-de-vida-ecv)

[estadisticas-por-tema/pobreza-y-condiciones-de-vida/calidad-de-vida-ecv](https://www.dane.gov.co/index.php/estadisticas-por-tema/pobreza-y-condiciones-de-vida/calidad-de-vida-ecv)