



# Perfiles de viviendas y condiciones asociadas al riesgo entomológico de Chagas

## Loja, 2005

Emilio Andrade

Carrera de Ciencia de Datos

19 de diciembre de 2025

## 1. Introducción

La enfermedad de Chagas es un problema de salud pública que se relaciona con condiciones sociales y del entorno, y de forma especial con características de la vivienda. El vector puede encontrar refugio en estructuras con materiales vulnerables, y además su presencia tiende a aumentar cuando existen factores ambientales como maleza cercana, acumulación de residuos o presencia de animales alrededor del hogar Pan American Health Organization, 2025; World Health Organization, 2025. En Ecuador el tema sigue vigente, por lo que la prevención necesita enfoques que permitan priorizar acciones, comunidades y tipos de vivienda con mayor riesgo Ministerio de Salud Pública del Ecuador, 2025

El objetivo de este informe es presentar perfiles de viviendas con condiciones similares, usando análisis de datos para agrupar hogares según materiales, prácticas de agua y saneamiento, entorno y presencia de animales. La idea es que estos perfiles ayuden a orientar intervenciones de prevención y control del vector, con una lectura clara para instituciones públicas y actores comunitarios

## 2. Metodología

Se trabajó con la base *Viviendas Loja 2005* con  $n = 1000$  registros. Para el análisis se seleccionó un subconjunto de 19 variables que representan condiciones de vivienda y entorno, por ejemplo materiales, saneamiento, agua, animales, presencia de palmas y maleza, y distancias asociadas



Primero se transformaron variables tipo sí no a una codificación 1 y 0. La variable de tratamiento del agua se manejó como ordinal simple, donde un valor mayor representa una mejor práctica. Luego se imputaron faltantes con la mediana y se estandarizaron las variables para que queden en una misma escala y así evitar que una sola magnitud domine las distancias del agrupamiento

Para identificar perfiles se compararon dos métodos de clustering. K Means, que agrupa por similitud global, y DBSCAN, que agrupa por densidad y puede separar puntos aislados. La calidad del agrupamiento se revisó con métricas de separación y cohesión como *silhouette*, Davies Bouldin y Calinski Harabasz. Para visualizar patrones se aplicó PCA en 2D, como una proyección que resume parte de la variabilidad y permite ver tendencias generales aunque exista solapamiento

### 3. Resultados principales

En la base se observó una mayoría de viviendas con estado entomológico negativo, y una fracción menor positiva. Esta diferencia es esperable y hace que el análisis se enfoque en identificar perfiles que concentren condiciones asociadas al riesgo, más que en predecir un resultado individual

#### 3.1. Resumen de los algoritmos

Con K Means se seleccionó  $k = 7$  por el mejor desempeño relativo dentro de los valores probados. La separación global fue moderada, con  $silhouette = 0,1438$ , Davies Bouldin = 1,5701 y Calinski Harabasz = 89,01. Esto sugiere que existen perfiles, pero con mezcla parcial entre grupos, lo cual es común en datos de vivienda

Con DBSCAN se trabajó con  $\epsilon = 4,1482$  y  $min\_samples = 5$ , se obtuvieron 8 clusters y un nivel de ruido bajo de 0,023. Al evaluar sin ruido la *silhouette* sube a 0,2576 y Davies Bouldin mejora a 1,274, lo que indica que DBSCAN identifica mejor núcleos densos, aunque también genera clusters pequeños que se interpretan con cautela

#### 3.2. Perfiles destacados

Al revisar los clusters de K Means, el grupo con mayor proporción de viviendas positivas fue el **cluster 2** con aproximadamente 30,4 % de positivas. Este perfil combina presencia de palmas cercanas, mayor presencia de maleza, y señales de convivencia con animales y

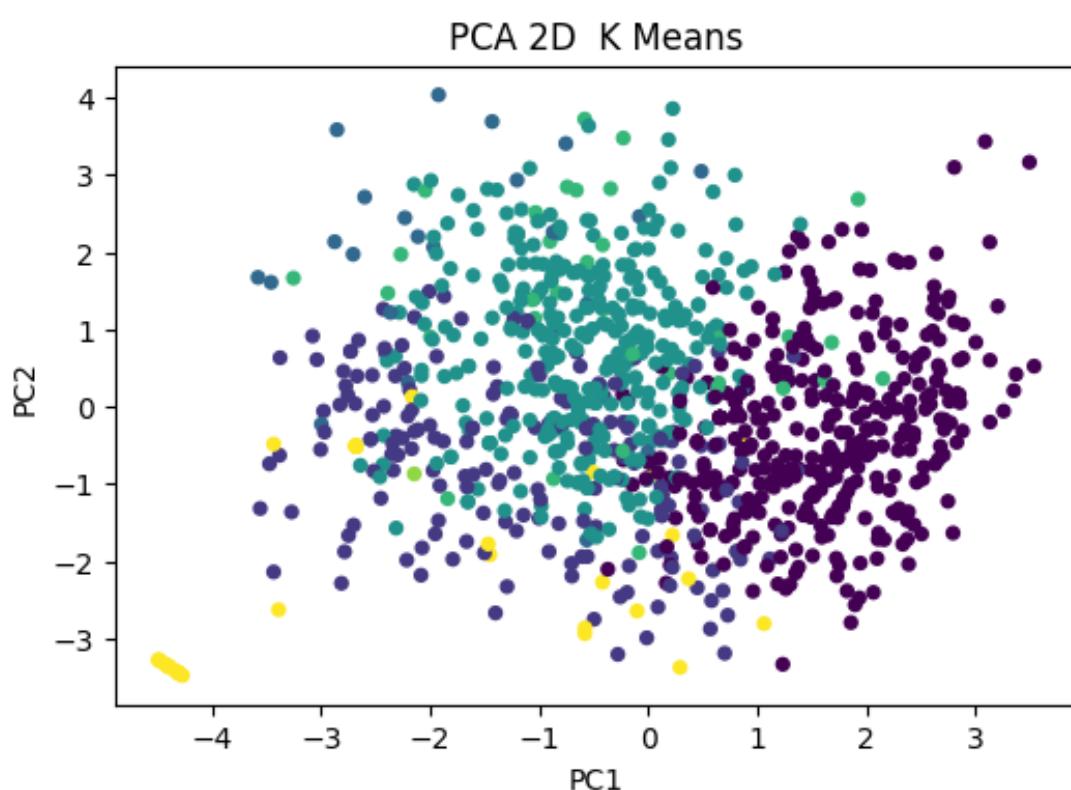


corrales, lo que en conjunto puede favorecer condiciones para el vector en ciertos entornos

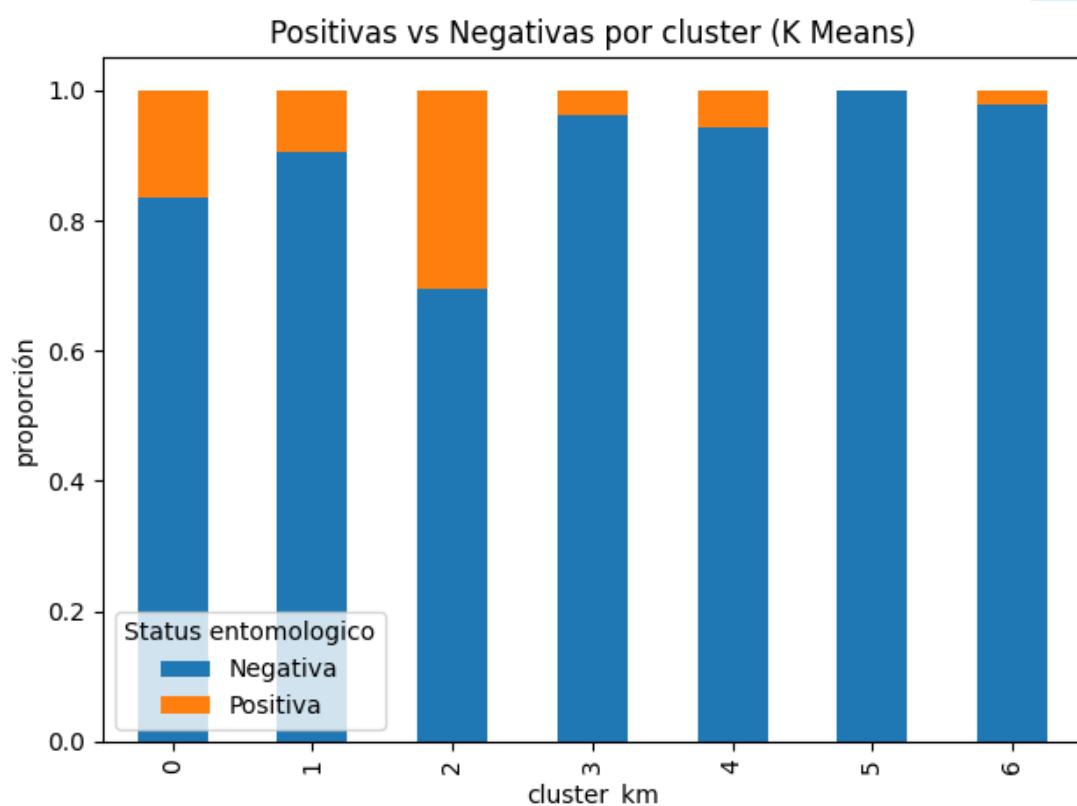
Un segundo perfil relevante fue el **cluster 0** con alrededor de 16,6 % de positivas. En este caso se observan indicadores de vulnerabilidad habitacional y sanitaria, como mayor ausencia de infraestructura, y una combinación de materiales y entorno que merece atención por riesgo potencial

Para apoyar lectura rápida, se sugiere usar estas visualizaciones del notebook

- Distribución del estado entomológico, para dimensionar el problema
- Porcentaje de positivos por cluster de K Means, para priorización
- PCA 2D coloreado por cluster, como mapa visual de perfiles



**Figura 1:** Proyección PCA 2D coloreada por cluster de K Means, útil para observar solapamiento y tendencias generales



**Figura 2:** Proporción de viviendas con estado entomológico positivo por cluster de K Means, para priorización de perfiles

## 4. Implicaciones prácticas

Los perfiles identificados permiten pasar de una lista larga de viviendas a grupos con necesidades parecidas. En la práctica esto ayuda a definir intervenciones focalizadas, por ejemplo visitas prioritarias a viviendas que combinan factores del entorno como maleza o palmas cercanas, junto con presencia de animales o corrales, y condiciones sanitarias más débiles

Además, los resultados sugieren que no existe un solo factor aislado, sino combinaciones. Por eso una campaña efectiva suele mezclar acciones de vivienda, ordenamiento del entorno y educación comunitaria, y también coordinación con autoridades locales para sostener los cambios en el tiempo

## 5. Recomendaciones

- Priorizar inspecciones y acompañamiento en perfiles con mayor proporción de viviendas positivas, en especial los grupos similares al cluster 2 y cluster 0
- Promover control del entorno cercano a la vivienda, reducción de maleza, manejo de palmas cercanas cuando corresponda, y limpieza de puntos de acumulación de residuos
- Fortalecer prácticas de saneamiento y agua segura, con mensajes claros, y acompañamiento cuando falten servicios básicos
- Ordenar la presencia de animales cerca del hogar, con corrales ubicados de forma más segura y limpieza frecuente, sin afectar la economía familiar
- Usar los gráficos por perfil como material de socialización comunitaria, porque ayudan a entender el riesgo sin entrar a términos técnicos

## 6. Conclusión

Este análisis agrupó viviendas con condiciones similares y permitió identificar perfiles que concentran combinaciones de factores vinculados al riesgo entomológico. K Means entregó grupos interpretables y DBSCAN ayudó a reconocer núcleos densos y casos aislados, aunque ambos métodos muestran que la separación no es perfecta, lo cual es coherente con la complejidad del problema



En conjunto, los resultados apoyan la toma de decisiones para prevención, porque permiten priorizar intervenciones según perfiles de vivienda y entorno, y refuerzan la idea de trabajo colaborativo entre ciencia de datos, salud pública y comunidad para reducir el riesgo de Chagas

## Referencias

Ministerio de Salud Pública del Ecuador. (2025). MSP intensifica estrategias para prevenir la Enfermedad de Chagas [Publicación institucional]. Consultado el 19 de diciembre de 2025, desde <https://www.salud.gob.ec/msp-intensifica-estrategias-para-prevenir-la-enfermedad-de-chagas/>

Pan American Health Organization. (2025). Chagas disease. Consultado el 19 de diciembre de 2025, desde <https://www.paho.org/en/topics/chagas-disease>

World Health Organization. (2025). Chagas disease (also known as American trypanosomiasis) [Fact sheet, 2 April 2025]. Consultado el 19 de diciembre de 2025, desde <https://www.who.int/news-room/fact-sheets/detail/chagas-disease-%28american-trypanosomiasis%29>