



Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Escuela de Estudios de Postgrado

Introducción a la Minería de Datos

Proyecto - parte 1

José Lester Rodríguez Culajay

Cuarto Trimestre 2024

Catedrático: M.A. Ing. Kevin Lajpop

Generalidades

Conjunto de datos

Se analizará el conjunto de datos obtenido de la encuesta nacional de condiciones de Vida ENCOVI, la cual brinda información sobre la pobreza a nivel nacional considerando características de la población y hogares. El conjunto de datos cuenta con información obtenido de diferentes años siendo estos: 2023, 2014, 2011, 2006, 2000, en este caso el análisis toma únicamente los datos 2023 para tener un enfoque de la realidad actual más próximo.

Enlace: <https://datos.minfin.gob.gt/dataset?groups=contrataciones-adquisiciones>

Conjunto: hogares

Año: 2023 – anual

Repositorio: https://github.com/lesterodgt/MIICC408_proyecto/

Reglas de asociación A priori

Regla 1

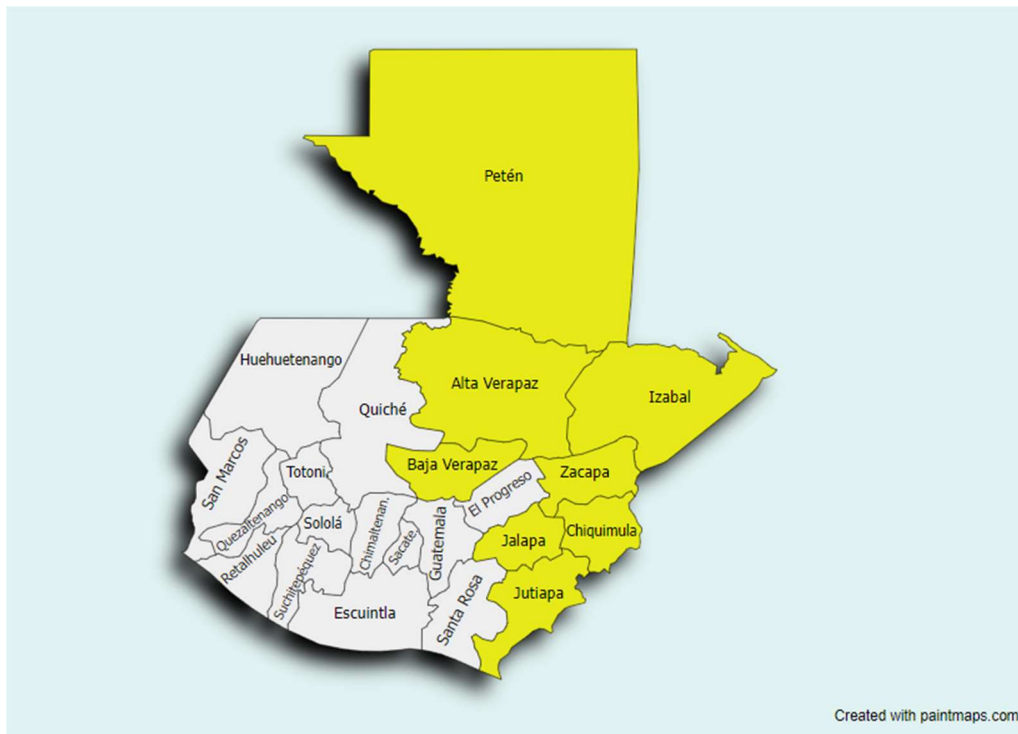
{DEPTO=[15,22]} => {PPB02=[1,5]}

DEPTO, departamento

PPB02: 2. ¿Cuántos hogares hay en esta vivienda?

Explicación: en los departamentos indicados ubicados al norte y oriente del país, en una vivienda suelen encontrarse de 1 a 5 hogares.

1	Guatemala
2	El Progreso
3	Sacatepéquez
4	Chimaltenango
5	Escuintla
6	Santa Rosa
7	Sololá
8	Totonicapán
9	Quetzaltenango
10	Suchitepéquez
11	Retalhuleu
12	San Marcos
13	Huehuetenango
14	Quiché
15	Baja Verapaz
16	Alta Verapaz
17	Petén
18	Izabal
19	Zacapa
20	Chiquimula
21	Jalapa
22	Jutiapa



Nota. Elaboración propia utilizando paintmaps.com

Regla 2

{PPB02=[1,5]} => {P01A02=[2,4]}

PPB02, ¿Cuántos hogares hay en esta vivienda?

P01A02, ¿Cuál es el material predominante en las paredes exteriores?

P01A02	1	Ladrillo
	2	Block
	3	Concreto
	4	Adobe
	5	Madera
	6	Lámina metálica
	7	Bajareque
	8	Lepa, palo o caña
	98	Otro ¿cuál?

Explicación: Cuando en una vivienda hay de 1 a 5 hogares, el material que predomina en las paredes exteriores es el block y concreto.

Regla 3

{P01D16=[2,4]} => {P01A03=[2,98]}

P01D16, ¿Qué tratamiento le dan principalmente al agua para beber?

P01A03, ¿Cuál es el material predominante en el techo?

P01D16	1	La hierven
	2	La filtran
	3	Le ponen cloro
	4	Compran agua purificada
	5	Ninguna
	98	Otro, ¿Cuál?

P01A03	1	Concreto
	2	Lámina metálica
	3	Asbesto cemento
	4	Teja
	5	Paja, palma o similar
	98	Otro, cual?

Explicación: En los hogares que filtran y ponen cloro al agua para beber, el material de los techos de las casas no es de concreto

Regla 4

P01D16=[1,2]} => {P01A04=[4,98]}

P01D16, ¿Qué tratamiento le dan principalmente al agua para beber?

P01A04, ¿Cuál es el material predominante en el piso?

P01D16	1	La hierven
	2	La filtran
	3	Le ponen cloro
	4	Compran agua purificada
	5	Ninguna
	98	Otro, ¿Cuál?

P01A04	1	Ladrillo cerámico
	2	Ladrillo de cemento
	3	Ladrillo de barro
	4	Torta de cemento
	5	Parqué
	6	Madera
	7	Tierra
	98	Otro ¿cuál?

Explicación: el material predominante en el piso de los hogares donde el agua para beber es hervida es de cemento, parque, madera, tierra entre otros.

Regla 5

{P01G01=[3,4]} => {DOMINIO=[2,3]}

P01G01, Si tuviera que comparar la situación económica de su hogar, ¿considera usted que en su hogar, respecto al año 2021

DOMINIO, Dominio de Estudio

Explicación: En el 2023 los hogares del área rural a nivel nacional y el área urbana no metropolitana consideran la situación económica del hogar permanece igual al 2021.

P01G01	1	Ha empeorado?
	2	Ha mejorado?
	3	Permanence igual?
	4	El hogar se formó después de 2021?
DOMINIO	1	Urbano Metropolitano
	2	Resto Urbano
	3	Rural Nacional

Reglas de asociación fpgrowth

Regla 1

{PPB02=[1,4]} => {P01H08=[2,99]}

PPB02, ¿Cuántos hogares hay en esta vivienda?

P01H08, ¿Usted o algún adulto en su hogar comió solo una vez al día o dejó de comer todo un día?

Explicación: En las viviendas donde hay de uno a cuatro hogares por lo general ningún miembro de la familia ha dejado de comer.

P01H08	1	Si
	2	No
	9	No sabe
	99	No responde

Regla 2

{PPB02=[1,4],P01A05A=[1,2]} => {P01D02=[1,3]}

PPB02, ¿Cuántos hogares hay en esta vivienda?

P01A05A, ¿Esta vivienda está conectada a una red de distribución de agua?

P01A05A	1	Si
	2	No

P01D02, Del total de cuartos, ¿cuántos utiliza como dormitorios?

Explicación: En las viviendas donde hay de uno a cuatro hogares sin importar si esta conectada a red de distribución de agua, las viviendas tienen de 1 a 2 habitaciones como dormitorios

Regla 3

{PPB02=[1,4]} => {P01D02=[1,3]}

PPB02, ¿Cuántos hogares hay en esta vivienda?

P01D02, Del total de cuartos, ¿cuántos utiliza como dormitorios?

Explicación: Las viviendas que tienen de 1 a 4 hogares, tienen 2 dormitorios

Regla 4

{P01A01=[1,5]} => {P01D02=[1,3]}

P01A01, ¿El tipo de vivienda que ocupa el hogar es?

P01D02, Del total de cuartos, ¿cuántos utiliza como dormitorios?

P01A01	1	Casa formal
	2	Apartamento
	3	Cuarto en casa de vecindad
	4	Rancho
	5	Casa improvisada
	98	Otro ¿cuál?

Explicación: Las viviendas, sin importar el tipo, tienen de 1 a 2 dormitorios

Regla 5

{DEPTO=[6,15]} => {POBREZA=[2,3]}

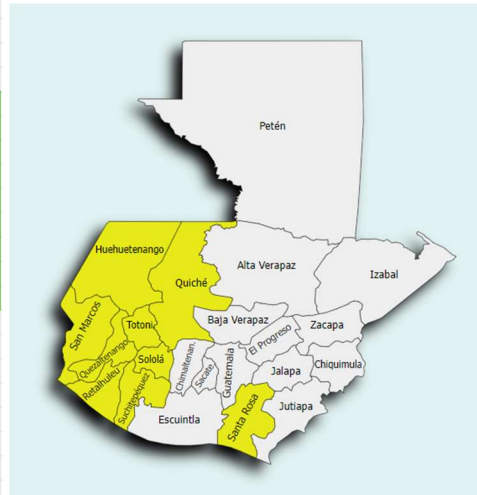
DEPTO, departamento

POBREZA, Clasificación de hogar (Pobreza)

POBREZA	1,00	Pobre extremo
	2,00	Pobre no extremo
	3,00	No pobre

Explicación: La regla indica que en los departamentos indicados (altiplano, costa sur y santa rosa) la pobreza predominante no es la extrema

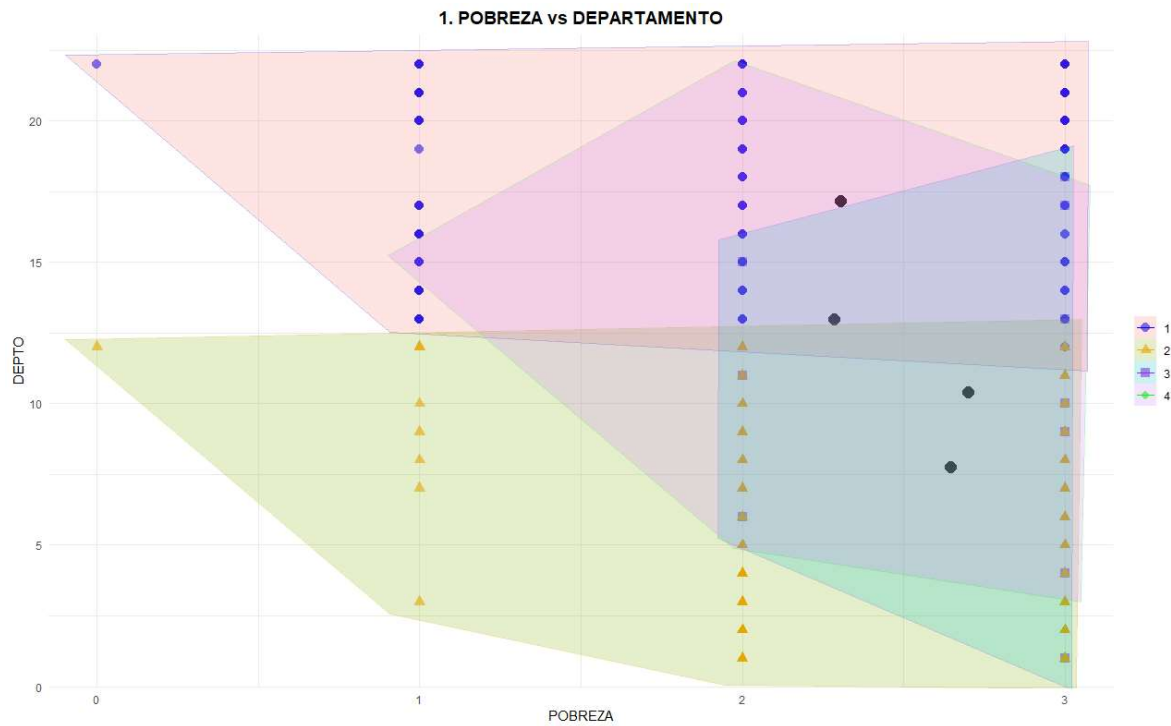
DEPTO	1	Guatemala
	2	El Progreso
	3	Sacatepéquez
	4	Chimaltenango
	5	Escuintla
	6	Santa Rosa
	7	Sololá
	8	Totonicapán
	9	Quetzaltenango
	10	Suchitepéquez
	11	Retalhuleu
	12	San Marcos
	13	Huehuetenango
	14	Quiché
	15	Baja Verapaz
	16	Alta Verapaz
	17	Petén
	18	Izabal
	19	Zacapa
	20	Chiquimula
	21	Jalapa
	22	.Iutiana



Análisis de clúster

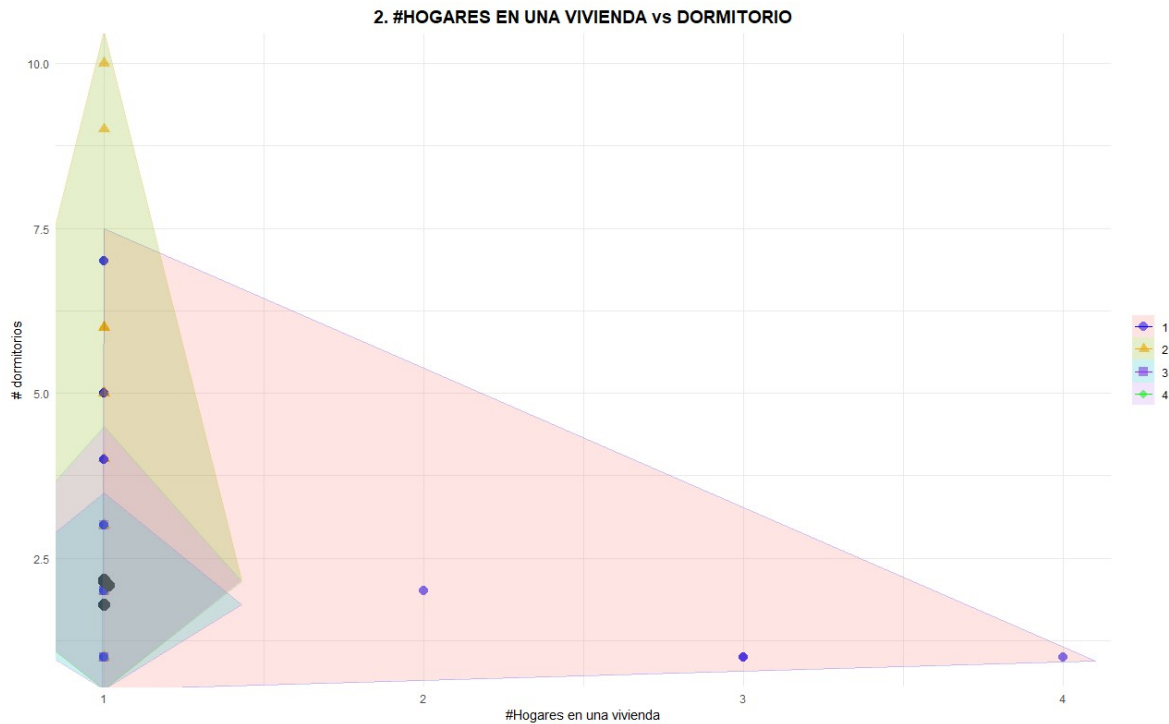
1. Pobreza vs departamento

Se puede observar una tendencia entre 2 posibles áreas del país donde la pobreza es extrema, existen áreas intermedias lo cual podría indicar las cabeceras departamentales que podrían tener mejores condiciones.



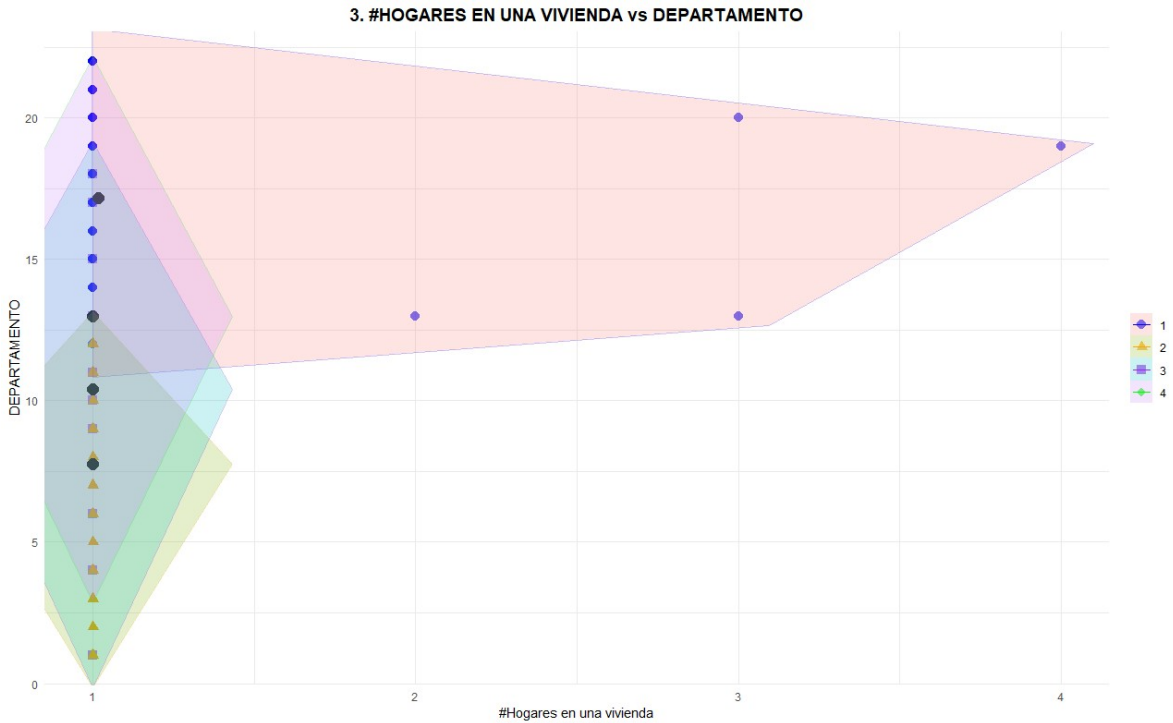
2. Numero de hogares en una vivienda y la cantidad de dormitorios

Se puede observar que cuando la vivienda cuenta con un menor numero de hogares la cantidad de contar con varios dormitorios aumenta, esto debido a que en los hogares donde las familias se ven obligadas a compartir la vivienda, generalmente por necesidades económicas, la cantidad de dormitorios tiende a ser menor, esto porque las viviendas no están en la capacidad de albergar a más de una familia.



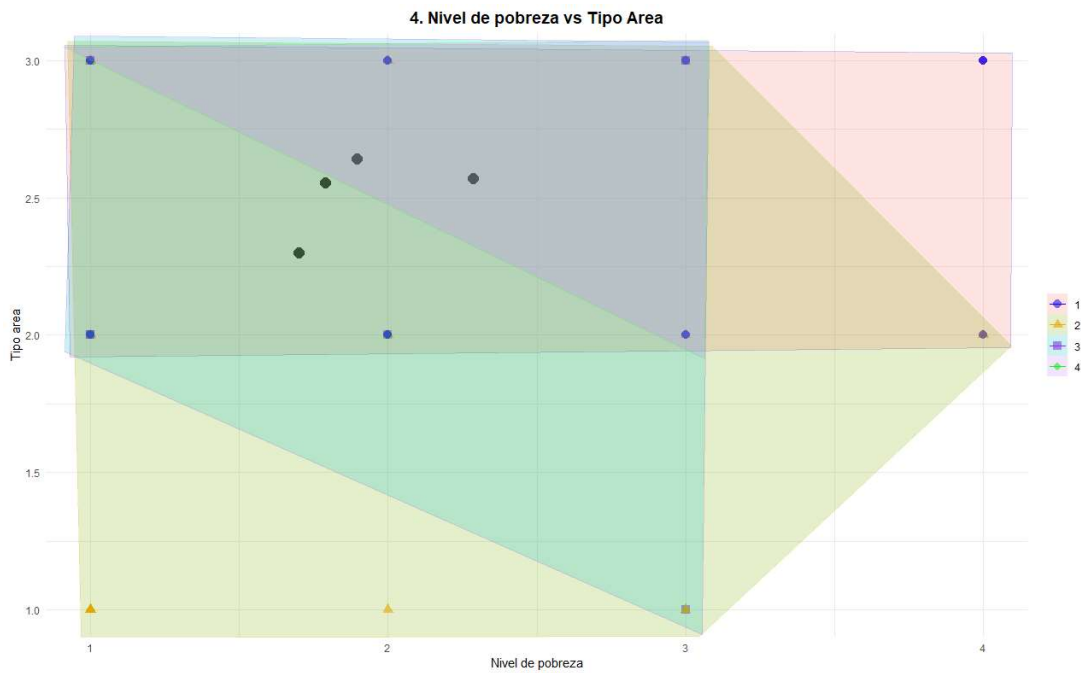
3. Numero de hogares de una vivienda versus departamento

Se observa que, en ciertos departamentos por lo general del área del norte, peten y oriente es más común encontrar viviendas con más de un hogar, esto podría ser un indicador de pobreza que podría relacionarse con el nivel de pobreza.



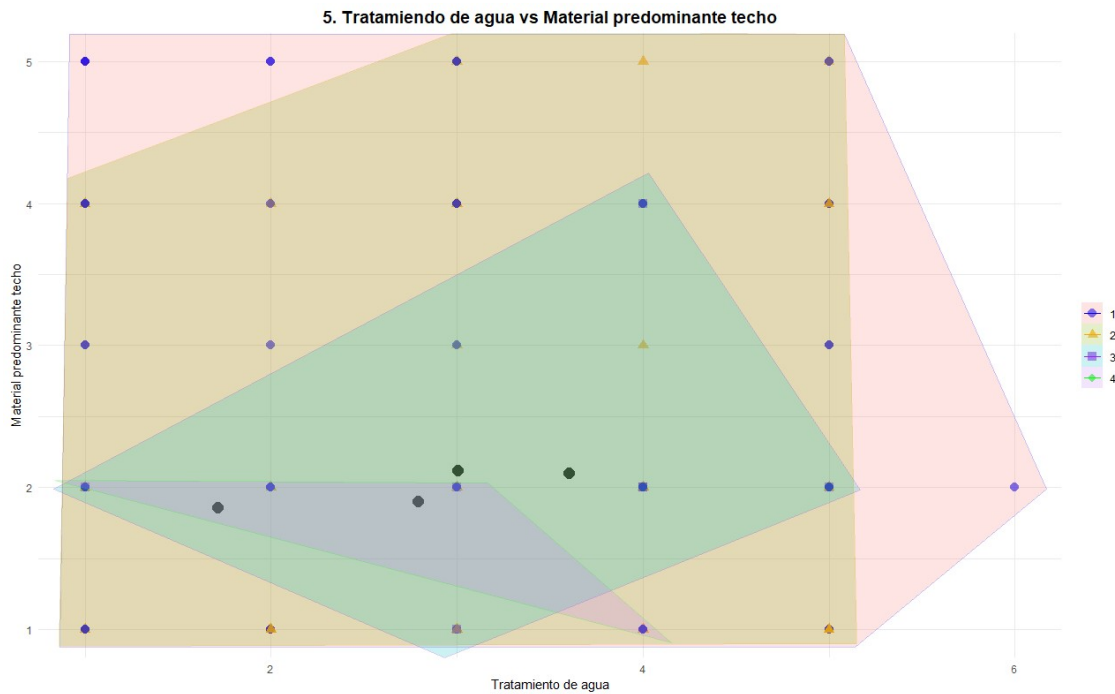
4. Nivel de pobreza por tipo de área

Nos muestra que la pobreza extrema y pobreza no extrema predomina en todas las áreas aun en el área metropolitana (3)



5. Tratamiento del agua vs material predominante del techo

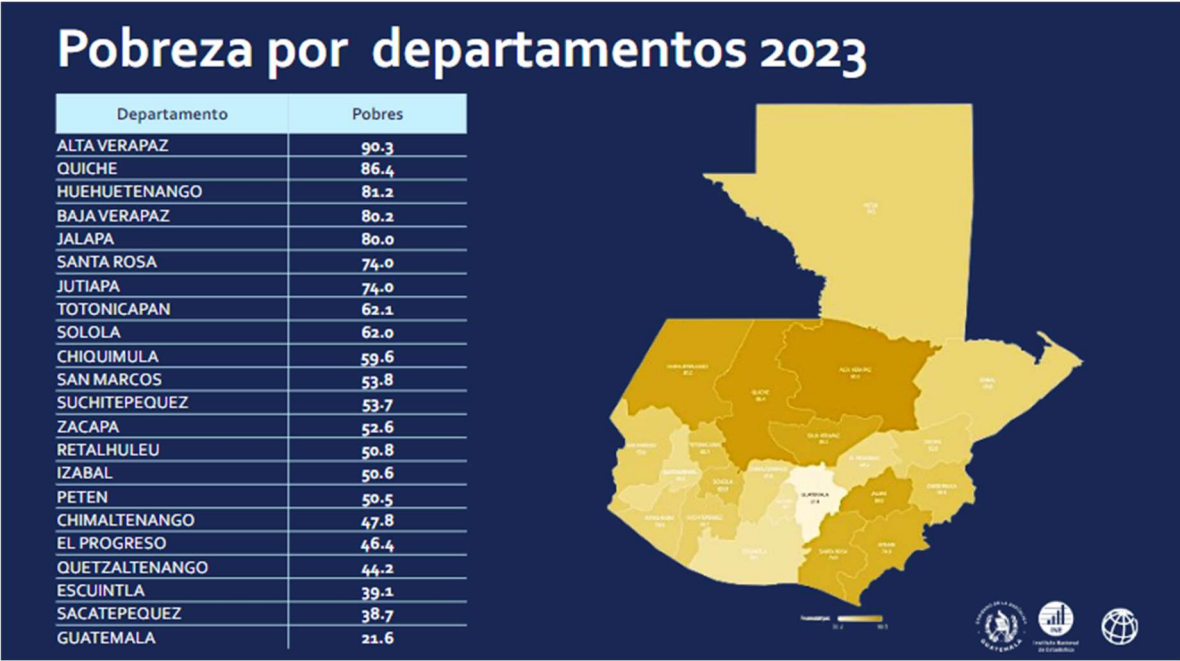
Esta grafica nos muestra que predomina el material tipo 2(lamina metálica) pero que el tratamiento del agua es diverso, sin embargo se puede visualizar una preferencia por la cloración (3).



Análisis de resultados

Considerando la regla obtenida respecto que en los departamentos Baja Verapaz, Alta Verapaz, peten, Izabal, Zacapa, Chiquimula, Jalapa Jutiapa, ubicados al norte y oriente del país, en una vivienda suelen encontrarse de 1 a 5 hogares.

Según resultados del INE respecto a la pobreza por departamentos suelen predominar departamentos del área norte (Alta Verapaz, Baja Verapaz) y región Nor-occidente, lo cual es coincidente en que muchas familias se ven en la necesidad de compartir la vivienda cuando se forman nuevos hogares.



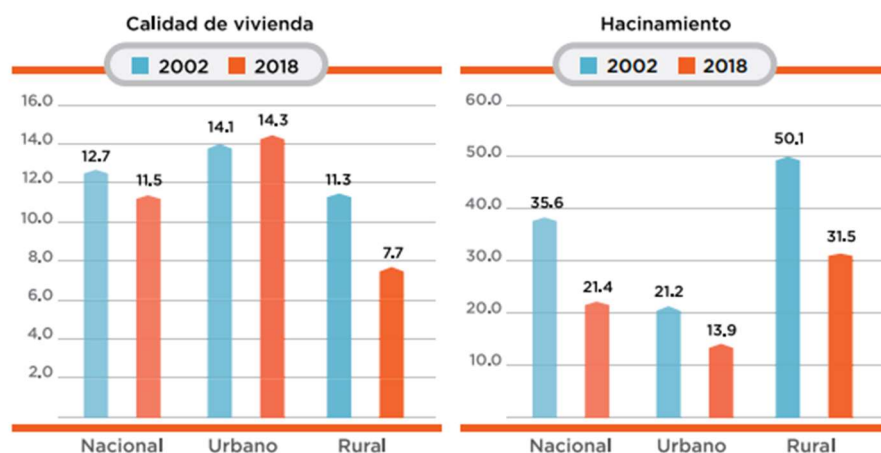
Nota. (ENCOVI, 2023) Pobreza por departamentos 2023

(<https://www.ine.gob.gt/2024/08/21/el-ine-presenta-cifras-de-pobreza-en-guatemala>)

Por otro lado, en el estudio titulado “Pueblos indígenas en Guatemala: desafíos demográficos, lingüísticos y socioeconómicos: análisis comparativo de los censos 2002 vs 2018” (Us, Mendoza, & Guzmán, 2022) se muestra el resultado de hacinamiento por vivienda, mostrando hacinamiento alto en el área rural, en especial en las áreas indígenas. Existe muchos motivos por el cual se realiza este tipo de acción, desde la necesidad económicas y falta de vivienda

como ámbitos culturales donde las nuevas familias se establecer alrededor del líder del hogar raíz por tradición.

Gráfica 44: Porcentajes de hogares con necesidad básica insatisfecha en calidad de la vivienda y hacinamiento por áreas urbana y rural, Censo 2002 vs Censo 2018.



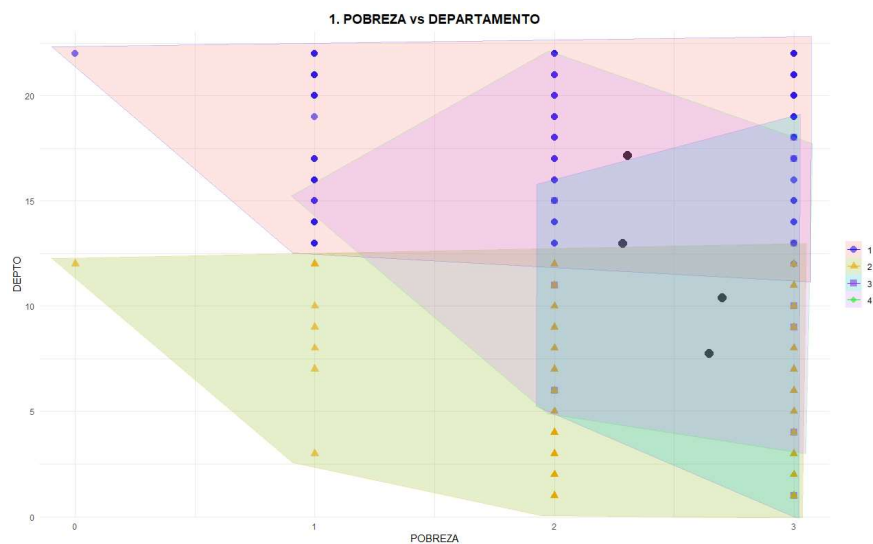
Nota. (Us, Mendoza, & Guzmán, 2022)Gráfica 44, Porcentajes de hogares con necesidad básica insatisfecha en calidad de la vivienda y hacinamiento por áreas urbana y rural, Censo 2002 vs Censo 2018 (<https://publications.iadb.org/publications/spanish/document/Pueblos-indigenas-en-Guatemala--desafios-demograficos-linguisticos-y-socioeconomicos-analisis-comparativo-de-los-censos-2002-vs-2018.pdf>)

En el caso de la regla que nos indica que ciertos departamentos no es muy común encontrar pobreza extrema en los departamentos de Huehuetenango, Quiché, San Marcos, Totonicapán, Quetzaltenango, Sololá, Retalhuleu, Sacatepéquez y Santa Rosa



Nota. Fuente elaboración propia utilizando paintmaps.com

La grafica de Clúster nos muestra un comportamiento donde

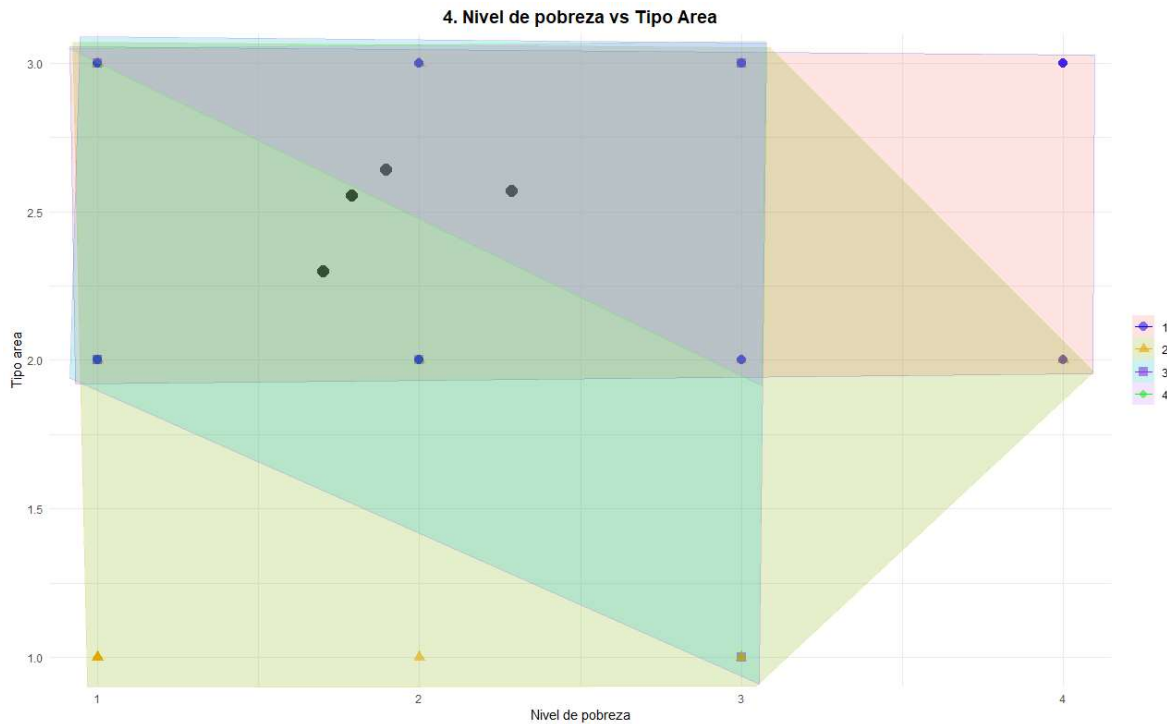


Nota: fuente elaboración propia utilizando R estudio

Los clústeres 1 y 2 pertenecen a departamentos del Altiplano y área sur oriental, que tienden a concentrar mayores niveles de pobreza extrema debido a factores sociales, históricos y económicos como el tipo predominante de agricultura y ganadería.

El clúster 3 muestra departamentos con mejor acceso a infraestructura y recursos, como los ubicados cerca de la capital o zonas productivas región central y metropolitana como Escuintla y Guatemala.

Por último, el clúster 4 incluye varios departamentos y esto podría ser un indicador que son departamentos donde existe mucha diversidad en el ámbito de pobreza

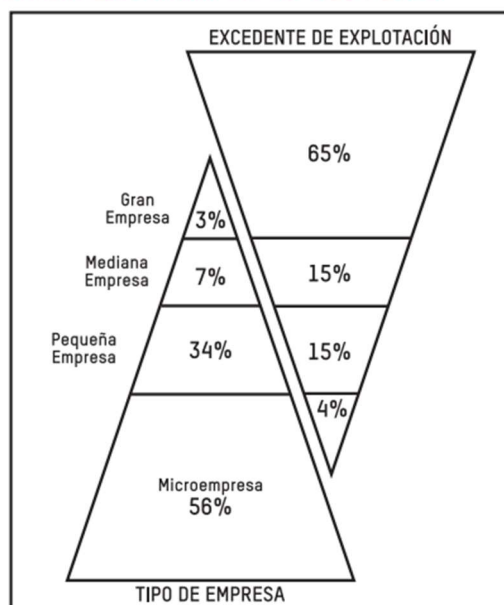


Nota: fuente elaboración propia utilizando R estudio

En el grafico de nivel de pobreza versus el tipo de área, siendo 1 pobreza extrema, 2 pobreza no extrema y 3 no pobre, para el área 1 representa área metropolitana, 2 urbano no metropolitana y 3 rural nacional. Puede observarse una predominancia en entre área rural y pobreza extrema. La mayoría de clúster están en el área rural y urbano no metropolitana con un nivel de pobreza intermedia entre extrema y no extrema. Cabe mencionar que es interesante que para el área 1 metropolitana es más común encontrar nivel de pobreza extrema y no extrema.

Esto concuerda con (OXAM INTERNACIONAL, 2019), quien en su publicación “Entre el suelo y el cielo” se hace un análisis de las condiciones de desigualdad que vive la población de Guatemala, indicando que las mujeres son quienes tienen menores oportunidades y que los mayores ingresos no están distribuidos de forma uniforme en la población. Esto da un soporte importante de la razón por la cual en el área metropolitana predomina la pobreza extrema y no extrema, atreviéndome a indicar que se centra en los alrededores de la ciudad como los barrios y asentamientos. En esta publicación se remarca que solo el 3% de las empresas se queda con el 65% del excedente generado en la actividad económica.

GRÁFICA 17: ESQUEMA PIRAMIDAL DE DISTRIBUCIÓN DEL EXCEDENTE DE EXPLOTACIÓN DE LA ECONOMÍA



Fuente: Waxenecker (2018 b), op. cit.

Nota. (OXAM INTERNACIONAL, 2019) Grafica 17 Esquema piramidal de distribución del excedente de explotación de la economía (https://www-cdn.oxfam.org/s3fs-public/file_attachments/entre_el_suelo_y_el_cielo_0.pdf)

Por tal razón solo ciertos sectores privilegiados tienen la posibilidad de mejorar sus condiciones de vida. Existen muchos factores que podrían analizarse para contrarrestar estos comportamientos, a consideración el más importante es la corrupción, acceso a la justicia justa y las políticas claras, justas y bien aplicadas que velen por el bienestar de la población en general.

Referencias

ENCOVI. (2023). *Condiciones de pobreza en Guatemala*.

OXAM INTERNACIONAL. (Marzo de 2019). *OXAM*. Obtenido de https://www-cdn.oxfam.org/s3fs-public/file_attachments/entre_el_suelo_y_el_cielo_0.pdf

Us, H., Mendoza, C., & Guzmán, V. (Diciembre de 2022). *Banco interamericano de desarrollo*. Obtenido de <https://publications.iadb.org/publications/spanish/document/Pueblos-indigenas-en-Guatemala--desafios-demograficos-linguisticos--y-socioeconomicos-analisis-comparativo-de-los-censos-2002-vs-2018.pdf>