

Propuesta de índice de marginación para los municipios en México

Armando Salinas Lorenzana

21 de febrero de 2024

1. PROBLEMA

Los datos que se encuentran en el archivo `IMM_2020.xls` corresponden al índice de marginación (IM) por cada municipio del país, calculado por el Consejo Nacional de Población (CONAPO) en 2020, y es hasta la fecha, el más reciente. El archivo contiene también los valores de varios indicadores que representan "nueve formas de exclusión de la marginación en las dimensiones: educación, vivienda, distribución de la población e ingresos monetarios;", y fueron construidos en base a la información del censo de población y vivienda 2020 realizado por INEGI. Estos indicadores se usan para calcular el IM según se describe en la Nota Técnico-Metodológica. El índice de marginación por municipio se muestra de forma categorizada en la Figura 1.

Supón que trabajas como asesor en la Secretaría del Desarrollo Social (o su equivalente) en algún estado del país, y te piden analizar formas alternativas de construir ese índice. Para eso, considera las siguientes actividades.

- a) Realiza un análisis de PCA basado en los 9 indicadores de CONAPO. ¿Qué puedes decir respecto al fenómeno de marginación en el país basado en este análisis? ¿Encuentras algún patrón interesante respecto a los municipios?
- b) Construye un IM alternativo tomando el primer componente principal que obtuviste con PCA (el indicador debe estar en $[0, 1]$ preferentemente). Compáralo con el de CONAPO y describe tus hallazgos al respecto. ¿Qué tanto cambia el resultado si tomas el segundo o tercer componente principal? ¿Tendría algún sentido hacerlo?

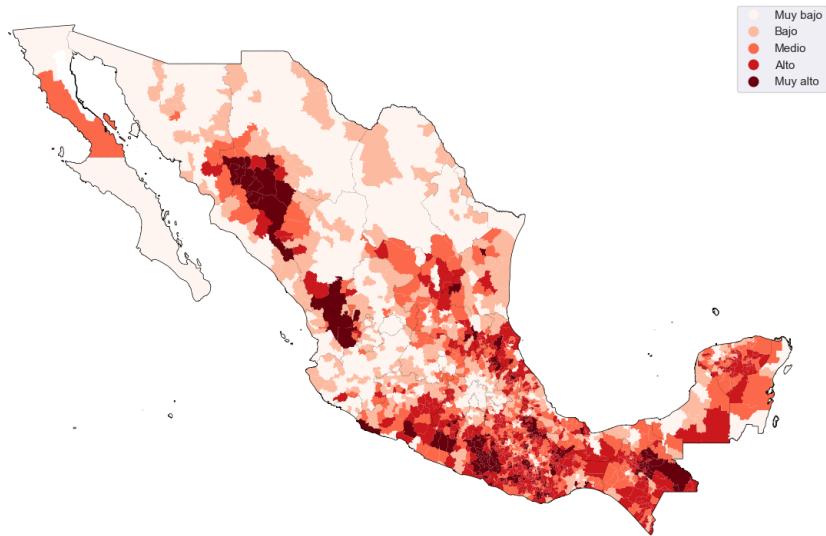


Figura 1.1: Índice de marginación 2020 a nivel municipal en 5 categorías.

- c) ¿Qué otra información propondrías que se incluyera dentro de la elaboración de tu índice (ya sea de estadísticas oficiales o de otra fuente)? ¿Estás de acuerdo con la metodología usada? ¿Tienes alguna otra propuesta para la elaboración del índice?

El resultado de este ejercicio debe ser un reporte ejecutivo con tus hallazgos y conclusiones, como si estuviera dirigido a tu hipotético jefe que no sabe nada (o casi nada) de estos métodos. Cada inciso puede ser un anexo técnico, con las tablas o ilustraciones que creas apropiadas para respaldar tu reporte.

1.1. REPORTE EJECUTIVO

1.1.1. INTRODUCCIÓN

El presente informe técnico presenta los resultados de un análisis exhaustivo sobre el Índice de Marginación en México, desglosado por municipio y proporcionado por el Consejo Nacional de Población (CONAPO). El objetivo fue aplicar técnicas avanzadas de análisis para profundizar en la estructura subyacente de estos datos complejos y ofrecer resultados valiosos.

El análisis se centró en la implementación del Análisis de Componentes Principales (PCA), una técnica fundamental en estadística multivariada, con el fin de explorar y comprender la variabilidad en el conjunto de datos. PCA no solo facilita la reducción de dimensionalidad, sino que también revela patrones y relaciones subyacentes entre las variables que contribuyen al índice de marginación.

Los resultados obtenidos no solo ofrecen una visión más clara de las dinámicas subyacentes en la marginación a nivel municipal, sino que también sirven como base para la toma de decisiones informadas y estratégicas en el ámbito de políticas públicas y desarrollo social.

1.1.2. RESULTADOS

Se realizó la técnica llamada PCA, para buscar información en los datos de las 9 índices de marginación, los resultados de dicha técnica se pueden ver a continuación

Modelando
el siglo XXI

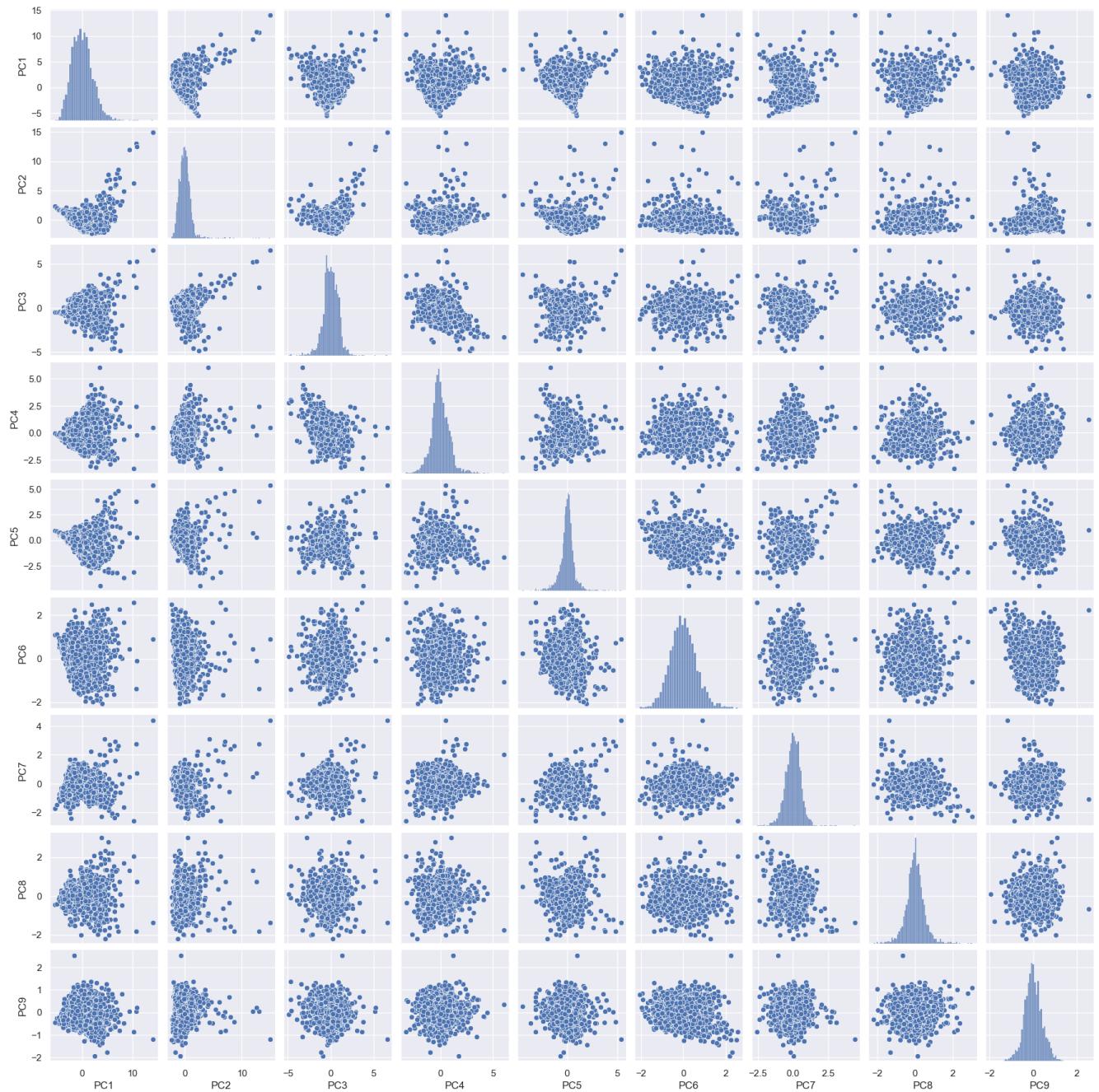


Figura 1.2: Las 9 componentes principales.

En la Figura 1.3 podemos observar la varianza de cada una de las componentes principales.

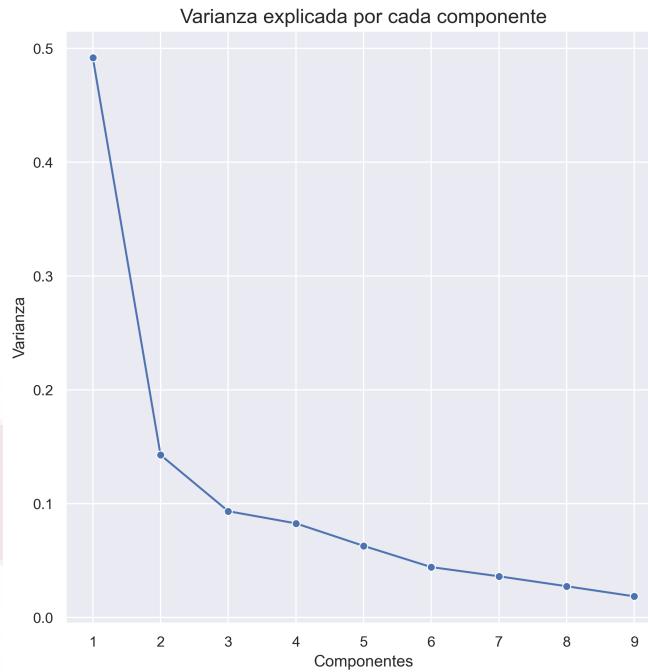


Figura 1.3: Diagrama de codo. En el diagrama se puede observar claramente como la primera componente es la componente con mayor varianza, siguiendo las segunda y así sucesivamente hasta la última componente.

De forma muy similar podemos ver que la varianza se reparte en las 9 componentes

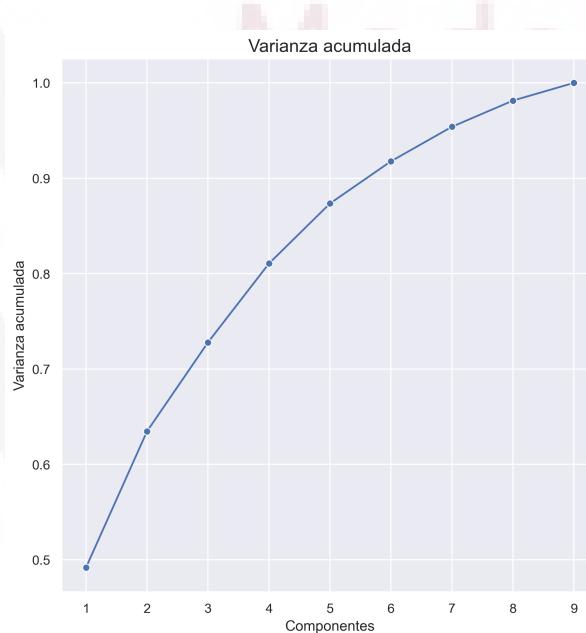


Figura 1.4: Varianza acumulada de las componentes. Podemos distinguir que a medida que tomamos más componentes, la varianza que obtenemos de las componentes es mayor y por ende las componentes contienen mayor información del conjunto de datos.

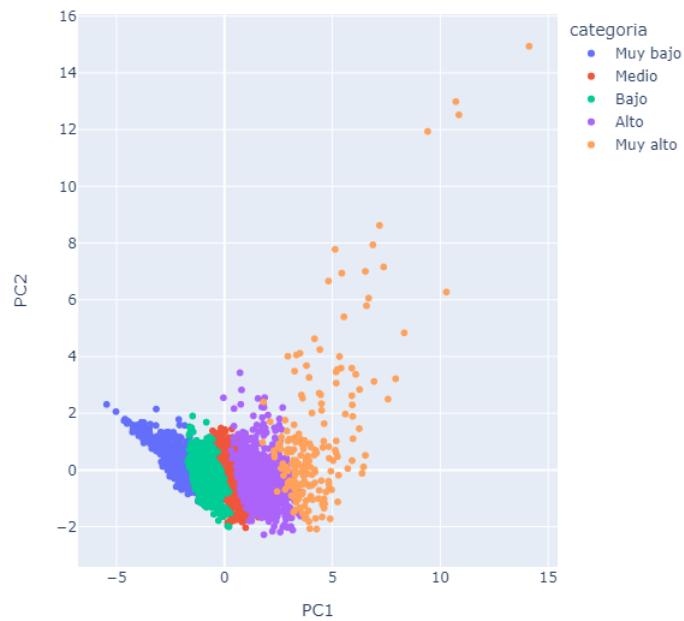


Figura 1.5: Componente principal uno contra componente principal dos.

Podemos observar un patrón en esta gráfica, podemos notar cómo se separan los grupos de municipios de acuerdo a su categoría de índice de marginación, por lo que podemos notar que la mayoría de municipios en México tienen un índice alto y muy alto de marginación, y que al menos 50 municipios tienen un índice de marginación muy alto.

Se puede distinguir que existen más municipios con índice de marginación bajos que con índice medio. De forma similar que la Figura 1.6.

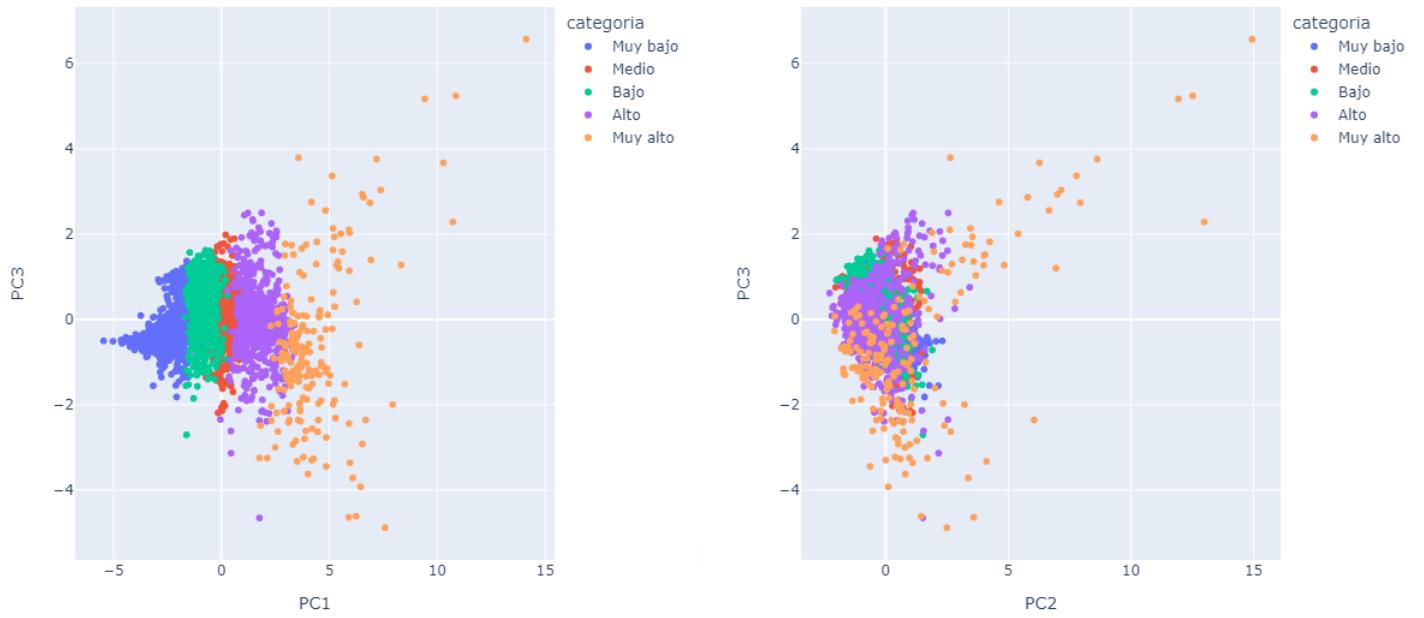


Figura 1.6: Primeras tres componentes principales.

Podemos notar en el gráfico de la componente uno contra la componente dos que se asemeja al gráfico de la componente uno contra la componente dos, sin embargo la estructura es distinta y puede observarse que se pierde, debido a que tiene menos información. La figura de PC2 y PC3 nos da mucho menos información, por lo que no podemos distinguir claramente los municipios de acuerdo a su categoría, recordemos que en estas componentes, la información es menor.

También podemos observar la relación entre las variables con el gráfico de Biplot, por lo que podemos analizar la Figura 1.7.

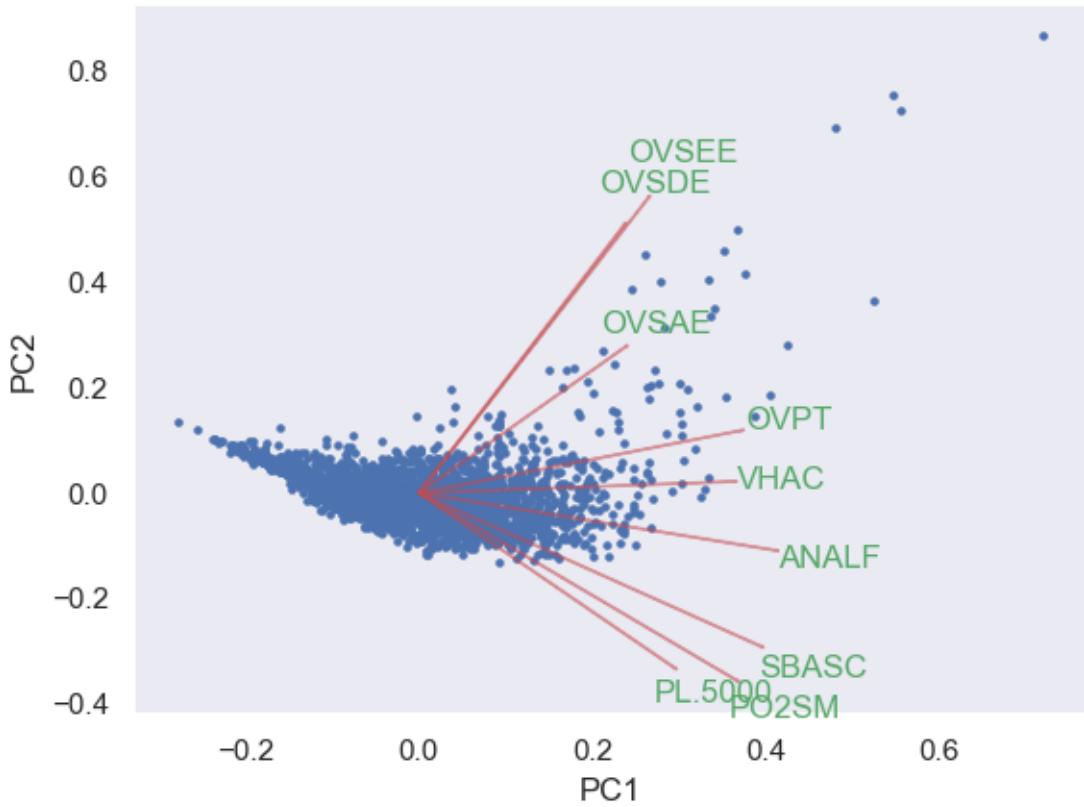
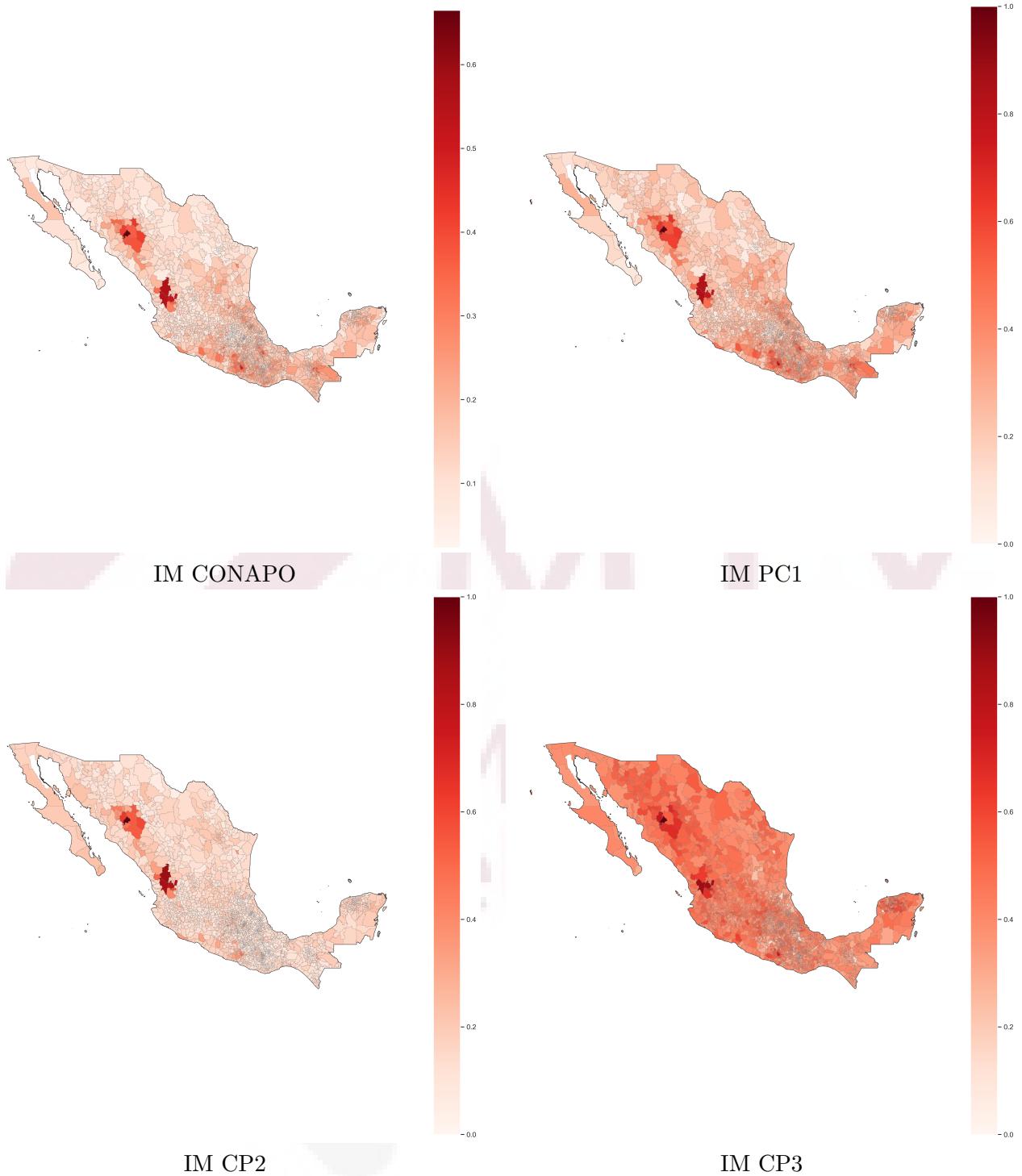


Figura 1.7: Biplot de las variables.

Podemos observar las relaciones entre las variables, primeramente podemos ver que el porcentaje de ocupantes en viviendas particulares sin drenaje ni sanitario (OVSDE) y el porcentaje de ocupantes en viviendas particulares sin energía eléctrica (OVSEE) se encuentran estrechamente relacionados por lo que la gran cantidad de viviendas sin drenaje, ni sanitario, tampoco tienen energía eléctrica, lo cual hace mucho sentido, y esto está orientado con los municipios con mayor índice de marginación. También podemos notar que el porcentaje de municipios con población menor a 5000 (PL5000) y el porcentaje de población ocupada con ingresos de hasta dos salarios mínimos (PO2SM) también están muy relacionadas, por lo que la mayoría de municipios con menor cantidad de habitantes, tienen un ingreso menor, lo cuál hace sentido para comunidades pequeñas, donde no se encuentran completamente desarrollados, por último, podemos notar que en cuanto a las viviendas sin drenaje y sanitario y sin energía eléctrica no están relacionadas con las poblaciones con habitantes menores a 5000, dado que el ángulo entre ellos es de aproximadamente 90 grados.

La comparación entre los índices de marginación calculado con los distintos componentes principales con respecto al propuesto por CONAPO se puede apreciar a continuación



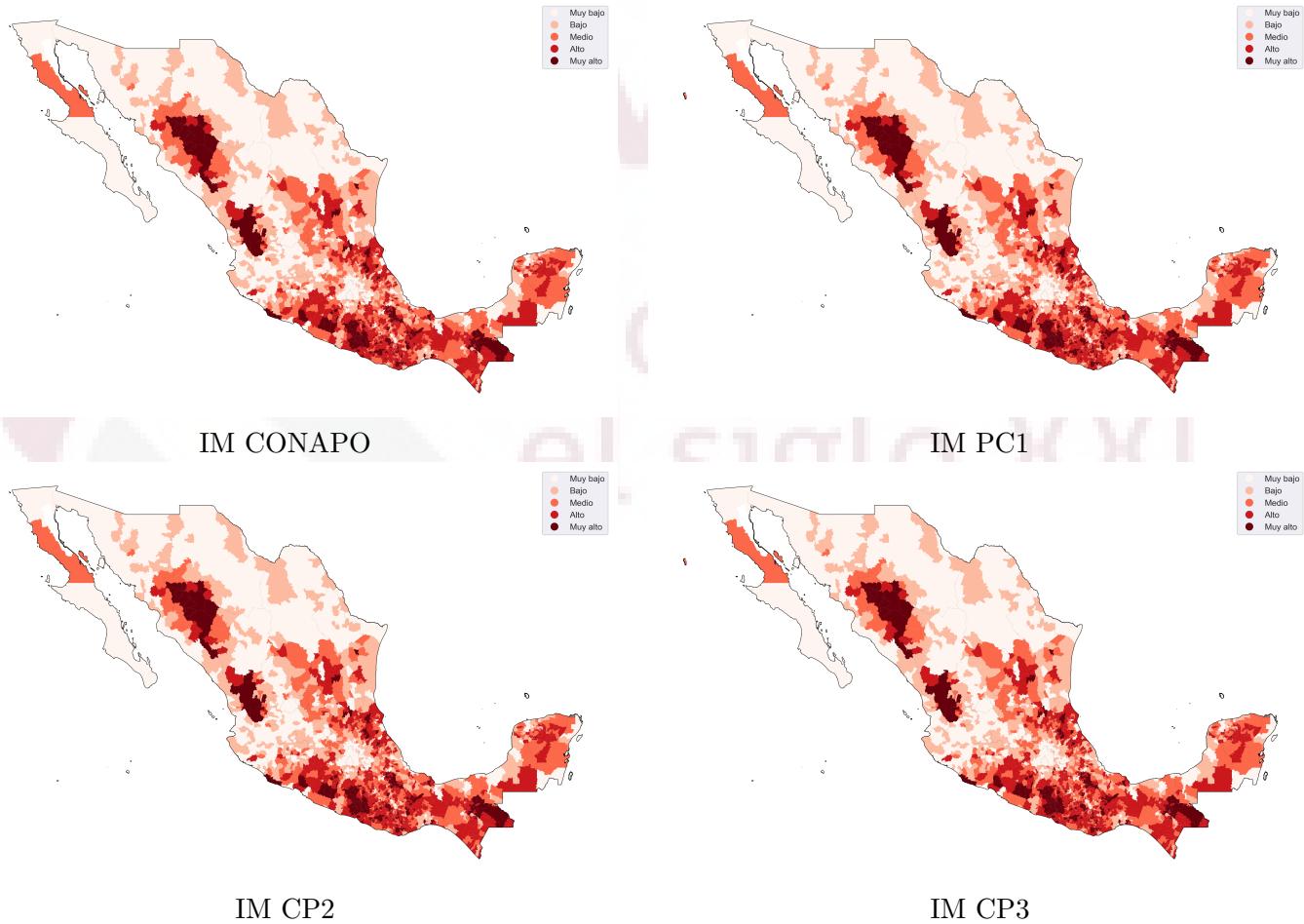
En estos mapas se puede apreciar la relación entre los índices de marginación propuesto por la CONAPO como los obtenidos de las componentes principales, podemos notar que el IM para el PC1, se asemeja más a la representación por la CONAPO, dado que preserva la mayor cantidad de características de la información de los municipios (49 % de la varianza), respetando su propia escala de color, para la componente 2 y 3 la representación del índice de marginación no resulta ser muy buena, dado que contienen menos información, por lo que no sería de gran ayuda para interpretar lo estudiado, pudiendo dar información falsa o errónea, por lo que no se recomienda usar estas dos representaciones del IM para dar estudiar el

índice de marginación.

Considerando esto, nos quedamos con la componente principal 1. Y analizando este mapa podemos encontrar que existe una zona con mucho IM en los estados de Chihuahua, Sinaloa, Durango y Nayarit, donde se ve muy remarcado este índice, así también podemos notar que la parte sur de México es la más afectada y que la parte norte de México resulta tener mucho menor índice de marginación, de esta forma podemos apreciar también que falta mucho desarrollo en el país, y qué falta desarrollarse poco más de la mitad del país.

La metodología usada por la CONAPO y los 9 índices de marginación que consideraron hacen una buena representación de la marginación del país, quizás hagan falta agregar más índices de marginación, más allá de los nueve, pero las nueve variables parecen describir aceptablemente la situación de México. Por lo que concuerdo con la representación de la CONAPO y cómo calculan el índice de marginación por municipio. De igual forma la representación encontrada con la primera componente principal, nos da una un índice de marginación parecido, el cual también puede utilizarse para darnos una referencia.

A pesar que los IM de las componentes 2 y 3 no resultan ser buena en la escala del 0 al 1 porque a simple vista las tonalidades se puede apreciar que cambian, cuando establecemos los límites de la escala del 0 al 1 para establecer las categorías de *Muy bajo*, *Bajo*, *Medio*, *Alto* y *Muy alto*, todas las gráficas coinciden como se muestra a continuación.



Estos límites para establecer las categorías en el intervalo del 0 al 1 se obtuvieron observando visualmente la base de datos de CONAPO, y lo establecieron de la siguiente manera aunque CONAPO no da una explicación de por qué realizaron una partición de esta forma:

0.87	\leq	Muy bajo
0.84	\leq	Bajo
0.82	\leq	Medio
0.765	\leq	Alto
		Muy alto
		$< 0.765.$

2. CONCLUSIONES

En conclusión, el análisis realizado mediante la técnica de Análisis de Componentes Principales (PCA) proporcionó una valiosa comprensión de la estructura y distribución de la información contenida en los datos de los 9 índices de marginación. Se observó claramente cómo la primera componente principal captura la mayor cantidad de varianza.

Las visualizaciones, como el gráfico de dispersión de las primeras dos componentes principales, permitieron identificar patrones y separación entre los municipios según sus categorías de índice de marginación. La distribución de los municipios en el plano de las dos primeras componentes reveló una concentración significativa en índices altos y muy altos de marginación, con una curiosidad adicional de la presencia de más municipios con índices bajos en comparación con los de índice medio.

La exploración de relaciones entre las variables mediante el biplot proporcionó correlaciones sobre la interconexión de ciertos indicadores de marginación, como la relación entre viviendas sin drenaje y sin energía eléctrica. La comparación de índices de marginación calculados con PCA respecto al índice propuesto por CONAPO sugiere que la primera componente principal podría ser una representación adecuada, mientras que las componentes secundarias podrían no contener información suficiente para una interpretación clara. En el análisis geográfico, se observó una distribución desigual del índice de marginación a lo largo del país, destacando áreas específicas con altos índices, como Chihuahua, Sinaloa, Durango y Nayarit, así como una disparidad entre el norte y sur de México.

En última instancia, tanto el método propuesto por CONAPO como la primera componente principal del PCA ofrecen perspectivas útiles sobre la marginación en México, brindando una representación efectiva de la situación de los municipios.